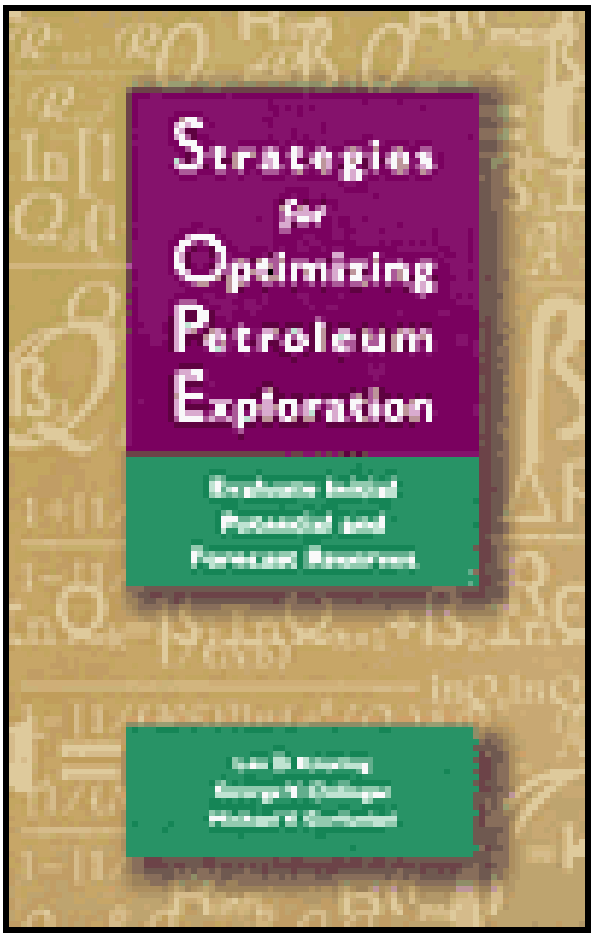
Strategies for Optimizing Petroleum Exploration: Evaluate Initial Potential and Forecast Reserves

by George V. Chilingar, Michael V. Gorfunkel, Lev Knoring



- ISBN: 0884159493
- Publisher: Elsevier Science & Technology Books
- Pub. Date: May 1999

Foreword

Strategies for Optimizing Petroleum Exploration contributes an economic/technologic evaluation methodology to the oil and gas exploration community that is unique and timely. In today's competitive market a company's exploration policy is governed by supply and demand economics. Exploration success mandates the adoption of new technologies. The book's methodology will assist exploration managers in formulating better exploration and development decisions.

The authors' approach is well defined in the context of the entropy law, economic processes, and systems analysis. It couples the sequence of exploration history for a region with exploration economics. Concepts are developed using three broad themes: (1) evolution of estimated oil and gas reserve accumulations; (2) succession of regional petroleum discoveries; and (3) use of systems guidelines in the management of exploration activities.

The bonding of strategic decisions with geological knowledge, historical quantitative reserve appraisal analysis, and sequential field discovery patterns within a producing region is the key to applying the methodology. Different exploration strategies will invoke various reserve outcomes with time. The accumulation of exploration results over time is quantified using the sequence and tempo of economic discovery, reserve forecasting, and reserve estimation data. Useful techniques are thoroughly presented in the book for selecting the optimum exploration-control strategy, which is predicated on the known petroleum accumulation patterns.

I recommend two books that provide information strongly supporting the foundation of the authors' methodology. Nicholas Georgescu-Roegen's book on *The Entropy Law and the Economic Process* explores the development of the concept of entropy used by the authors to analyze petroleum accumulations. *Systems Analysis and Project Management* by David Cleland and William King furnishes the bases of applying systems concepts to decision making and managing projects. The authors' methodology is a forward thinking approach that will impact our way in evaluating exploration priorities at the beginning of a new millennium.

Herman H. Rieke, Ph.D.
Professor and Head
Petroleum Engineering Department
University of Southwestern Louisiana
Lafayette, Louisiana
USA

Preface

This book deals with the exploration and appraisal strategies of oil and gas deposits as applied to a particular region, such as a petroleum province or petroleum region, an administrative province, or a tectonic unit. An exploration strategy applied in a region is reflected in the results of exploration (the history of the regional reserve accumulation or of the oil and gas fields discoveries).

For a detailed description of a specific method of exploration, the generic term "technique" is frequently applied in geologic literature. We apply the term "exploration strategy" to describe the development method of the oil and gas resources for the entire region in historical perspective.

This book focuses on the strategy of the petroleum resource development in a region and on the ways to improve this strategy. For many years, exploration technique has been successfully studied by numerous scientific institutions. However, development strategy did not attract the same attention from the experts and, therefore, studies on development strategy are limited. Despite the fact that strategic decisions are much more important in the development of the region's resource than tactical decisions when exploring the individual fields, experts only recently became interested in exploration strategies. The reserve accumulation history was the only issue related to the reserve development and evaluation process which was purposefully studied in recent years. Other important strategic questions, such as sequential patterns of oil and gas field discoveries, were not even properly formulated. The important issues of the optimization of the reserve discovery and evaluation process were completely ignored until now.

The following partial list of the applied problems addressed in this book verifies its practical value: (1) description and analysis of the internal patterns intrinsic in the deposit discovery and the reserve evaluation process; (2) optimization of this process based on the intrinsic patterns; and (3) long-term resource accrual forecast and evaluation. There are no books available which treat the deposit discovery and reserve evaluation process in its entirety, as a single regulated system. This system has defined properties and intrinsic evaluation patterns affected by certain system factors, including exploration strategies. This book will create and substantiate a theoretical concept of the reserve development strategy and reserve

evaluation process and, most importantly, help solve the practical problems of regional reserve development.

A particular feature of this book is that the theoretical findings that are expressed in the form of mathematical models. The role of modeling here is to refine the main concepts for better understanding of the nature of the exploration process.

Mathematical modeling of the deposit discovery and reserve evaluation process forced us to address situations beyond the traditional petroleum geology. The material presented requires continuous comprehension and systematization. Thus, the results of this study reflect the essence of the exploration process and how it is reflected in the reserve accrual and field discovery history.

Modeling in this book is applied at different stages, beginning with the statement of a general problem and ending with its specific solutions. The time-spatial patterns in the reserve accrual and in the sequence of field discoveries are the results of specific conditions having occurred in the implementation of the regional exploration strategy.

The modeling provided a scientific basis for the solution of applied problems. This made it possible to achieve higher quality, strict solutions to these problems, and to improve the reliability and validity of the estimates and forecast of the exploration process parameters. The main task of the study was to analyze the exploration process in Russia and the former Soviet Republics. It included practical problems associated with the exploration strategies optimization, resource forecast, long-term reserve accrual and drilling extent planning, and recommendations for future exploration. It was based on the data analysis from all major petroleum-producing regions. The exploration model was derived from the system of plausible verified conditions covering the process of oil and gas resource development. The core concept of the exploratory system is the regional exploration strategy, which is usually described, optimized, and tested in the process of modeling.

The book includes three parts covering the central issues of reserve development and evaluation. The first two parts are devoted to finding the intrinsic patterns of the process and its forecast based on these patterns. The third part deals with the control strategies for the process of regional resource development.

Dr. Leonid F. Khilyuk, Consultant
Former Department Chairman of Applied Mathematics
Kiev Technological University, Kiev, Ukraine

Table of Contents

PART I

Evolution of the Oil and Gas Reserves Discovery and Appraisal

This section deals with the history, or changes in time, of the annual and total oil and gas reserve accrual (growth) within individual regions. Hereafter, the term "reserve accrual" rather than "annual reserve accrual" is used, and the sum of annual reserve accruals from the beginning of development to the present is called the "initial evaluated reserves" or "accumulated reserves."

An analysis of reserve accrual patterns may determine the strengths and weaknesses of the exploration strategies for a given region. The following should be kept in mind, however: the reserve accrual is a major target parameter of exploration. For this reason, it may be reasonable to determine strategies based on the results, that is, the reserve accrual. The exploration* process may be deemed strategically sound if it provides for rapid growth of the reserve accruals to its maximum, and if the zone around the maximum (a few years prior and after the maximum) includes a substantial portion of the ultimate potential. The strategies used to develop a region may be considered satisfactory if the bulk of its reserves has been discovered and appraised rapidly, within a few years since the start of exploration. (A particular case may involve certain plays in the region that are accessible using the existing technology rather than the entire region). It is obvious that such an approach is only reasonable when discovery and appraisal of the reserves, with all the economic consequences, is the main purpose of exploration. If, for the sake of argument, we assume that the main purpose is

*Includes appraisal process.

conservation of the reserves for future generations who will use petroleum as a chemical raw material rather than as an energy source, then the strategies would be different.

In recent years, the subject was made much clearer when certain patterns were established in reserve accruals for regions with long petroleum exploration histories. The main pattern that emerged was that reserve accruals grow until they reach maximum. Afterwards, the increments decrease, and the decline is first rapid and then gradually slows. This pattern is inherent in the exploration process and in the natural reserve distribution. A substantial portion of the ultimate potential is concentrated within a few relatively large and giant fields that are usually discovered early. These major discoveries cause major reserve accrual.

The reserve accruals are studied from different aspects, such as their evolution in time and the amount of necessary exploratory drilling. A frequently used term is "the specific reserve accrual," or exploratory drilling efficiency, which is the amount of the reserve accrual per unit of work (such as one meter of drilling, one well, unit of expenditures, etc.). Patterns discovered in the reserve accrual give rise to corresponding techniques for evaluating the ultimate petroleum potential and for forecasting future reserve accruals. In American publications, such evaluation and forecast techniques, based on studies of the reserve accrual history ("dynamics") and exploratory drilling efficiency, are called "geologic-economic" or "historical-statistic" techniques. In Russian publications, these techniques are sometimes called "the development history ("development curve") method," and the graphs used to reflect the history of these parameters (reserve accrual or efficiency) as a function of time or exploration volumes are called "development curves." Forecasts are based on the extrapolation of these curves. Therefore, a forecast's reliability depends on the extrapolation reliability.

Using the traditional approach, smoothing and extrapolation of these curves was performed intuitively. That led to arbitrary evaluations that required a mathematical description of the reserve accumulation process. Indeed, forecast evaluations required modeling of the process. American geologists were the first to use mathematical equations for commercial reserves analysis and potential resources evaluation. A logistic function has been applied for this purpose by M. K. Hubbert [59]. Later, N. D. Uri also evaluated the U.S. petroleum potential using logistic curves [67]. S. L. Moore (1971) used the Gomperz-Makeham function as an equation describing the history of production accumulation for the evaluation of the U.S. initial proved reserves [63]. A. D. Zapp in his historical ("dynamic") model postulated that the rate at the end of each equal time interval represents a certain percentage of the reserve accumulation rate at the end of the previous interval [68]. He also noted that this rate is a linear

function of the accumulated production. M. K. Hubbert also believed that the reserve accumulation rate linearly declined as a function of the drilling volume [59]. In this case, the reserves accumulation is described by the so-called "modified exponent function."

The same function was later used by V. A. Balasanov and S. G. Kamensky to express the accrued reserves as a function of the total drilling volume [3]. Another group of Russian scientists suggested the use of the above equations (the logistic curve and the Gomperz-Makeham equation) for a long-term forecast of the appraised oil and gas reserves [6]. These curves may be used to characterize the historical behavior of the current reserves and current production. Later, one more model was introduced in which the logistic function parameters vary in time and depend on the prospective (possible) reserves [43]. The model described the relations between the reserve accrual and oil production.

CHAPTER 1

Modeling Principles

The modeling principles described in this chapter cover those facets of the reserve discovery and appraisal process that are studied in this book. The modeling is specific in that it depicts the process of the reserve accumulation. The modeling principles, however, are universal. In this sense, these principles are basic for all sections of the book. In other chapters, they are only slightly modified to accommodate specific features in the modeling of the phenomena considered in each case.

Mathematical Model

Purpose of the Study

The idea of a "model" is widely used in geological literature. However, there is no consensus of its interpretation. Various scientists understand the term differently. In order to make the following understandable, it is necessary to explain the concept of "model" as it is used in this book.

An important factor in understanding a phenomenon or an object is its abstract description, or general characterization, which builds to a theory. Such a characterization is always relatively incomplete, and only partially reflects the specifics of the subject. This kind of subject description is called a model. Geology, as in any scientific discipline, has its own language to describe such models, which is characterized by using everyday language. However, the persuasiveness of mathematics in geology resulted in a change from everyday language to the language of formal description.

A description of an object through the language of mathematics is called a mathematical model. Any model (including mathematical models) is phenomenologically based. Its construction begins with the study of a phenomenon and accumulation of related facts and observations. The facts are processed, systematized, and explained, which results in a model. This

model is not always mathematical. A mathematical model occurs only when a system of primary concepts is formulated in mathematical terms, and when the model describes not only the object itself but also a domain of rules determining acceptable operations with the object. In turn, an analysis of the model may provide valuable information about the properties of the object. The model itself begins serving as a source of information. It may indicate yet unknown areas of the process under study and show the direction of further studies. However, mathematical models are capable of this only when a question is asked and a problem is formulated. In such a case, mathematics provides a formal means to arrive at an answer to the question. Mathematics is a logical tool for decoding information contained in the model. A formal description and the abstract presentation of information usually makes it simpler to arrive at a conclusion and makes the conclusion more reliable and accurate.

An important problem with formalized descriptions is determining the point beyond which the model cannot be applied to adequately reflect reality. In other words, a mathematical model should adequately reflect the real phenomenon, and accurately describe properties of the phenomena under study. The applicability of a model is inseparable from the goals of the scientist. If a chosen model leads to the achievement of these goals, it is accurate for all practical purposes. Because words are used to describe a subject in geological studies, the goals are also defined in words. In mathematical modeling, the goals should be specified in precise mathematical terms. If the modeling goal is not specified, the adequacy of the model to the real object loses its meaning.

This book discusses the following goals for the modeling of the regional oil and gas initial appraised reserves accumulation process.

1. To develop a concept of the process, understand its nature, and connect the evolution of the reserve accumulation in time with certain motivating factors. This is one of the major goals that is not usually achieved using conventional analysis of the development history curves. Solving these problems helps determine major directions for study and analysis of possible options.
2. To forecast future reserve accruals and evaluate the ultimate petroleum potential. This is the most commonly pursued goal in the analysis of reserve accrual evolution.
3. To design a strategy for development.

As mentioned, mathematical models are phenomenologically based. The model should translate the patterns in the reserve accrual into proper mathematical hypotheses. In order to construct a model or a system of models, we will consider the specific features in the reserve accrual evolution and the causes that determine this evolution.

General Characterization of the Reserve Development Process

Oil and gas reserve accrual is a major parameter of the exploration process. The reserve accrual is a function of the exploration process as conducted in real time and under defined conditions. Therefore, the reserve accumulation evolution cannot be analyzed separately from this process.

For our purposes, the most important feature of the exploration process (hence, the reserve accumulation process), is that it is a controllable process. A number of features must be taken into account in order to model this process. We will analyze these features against the background of those typical of control processes.

Control is aimed at achieving a certain goal; the very idea of control does not make sense without a goal. There are many goals pursued by the process of exploration. Although the goal of fulfilling the reserve accrual plan is important, it is not the only one—there are short- and long-term goals. In the process of exploration, there is always a hierarchy of interrelated and inter-subordinated goals. It would be a mistake to give preference to one particular goal to the detriment of the general goals of exploration. For instance, one common goal is the total development of the region's resources, that is, maintaining a continuous and stable exploration process. However, evaluating the process relative to this long-term goal is insufficient for solving the problems associated with short-term goals. This is because these goals are widely separated in time. At the same time, the short-term goals should not obscure the long-term goals.

When discussing the hierarchy of goals, the long-term goals are determined by the need to provide resources for future oil production, and the medium-term goals are dictated by the internal evolution of the exploration domain. This evolution can be independent of the resource development. Even short-term goals are defined by the problems intrinsic in particular exploration phases, and these problems are not necessarily directly connected with the goals of the exploration process as a whole. Therefore, at any level of evaluation, whether from the standpoint of development as a whole (which includes the process of exploration as a component) or from the standpoint of individual exploration phases, there are "self-interest" goals and goals imposed by the "interests" of a higher level. It is important to note that "self-interests" and their respective goals may be different from other important but more remote goals, which form the individual rationales for the various exploration phases and the exploration process as a whole.

On the other hand, the following should be taken into account. Exploration is governed by organizations at different levels, from the management to field parties and even individuals. They form a hierarchical chain with each individual link having its own system of goals.

The interests of different levels usually do not coincide and may even be contradictory. For instance, the interests associated with the development of a modest field delay the discovery of larger fields, and, thus, may be contrary to the interests in the development of the region as a whole. At the same time, the goals of each link are determined by its own goals and those of higher levels. The actions of each link are to some degree affected by goals of higher levels. Thus, even with contradictions in goals, the system as a whole and its elements have some goals in common. In reality, these goals are not always clearly understood and defined. Short-term goals should not overshadow the long-term goals that change with time.

Exploration is conducted by different organizations, within various regions, within different areas of a particular region, and at different depths within a region. The goals change with time; currently, there are projects that require drilling to great depths, which was not feasible earlier. On the one hand, these goals arise as the need for mineral resources grows with the evolution of oil and gas production in a region. On the other hand, the goals are determined by the entire exploration process, by the depletion of large discoveries, or by their total absence within a study region, which jeopardizes the entire process.

The reserve accrual is achieved by field appraisal. A field must be discovered, and before that, the structure it is associated with must be identified and delineated. Therefore, goals associated with the continuity of exploration must comply with the phase structure of the exploration process. The complex, contradictory, and ever-changing structure of these goals affects the mathematical model. It is obvious that the model should address goal selection as well as the most efficient means of achieving this goal.

Studies associated with economic evaluations or the planning of exploration usually consider the reserve accrual or the extent of exploratory and appraisal drilling as their goals. Once reserve accrual is established, the drilling volume necessary for the achievement of this goal must be determined. Consequently, if the drilling footage is first assigned, then the amount of the reserve accrual achievable with this drilling must be determined. In reality, the interrelationship among the goals of resource development is more complex.

A goal is a precondition of control. To achieve the goal, however, certain means and resources of control are needed. Control is understood as the way these means are used (or controlling actions selected) to achieve this goal. Means of control include targets of exploration and appraisal, drilling rigs, and so forth, which, combined, determine the exploration play, the extent of drilling, and the funds expended.

Control is the selection of a particular action from some set of possible actions. This selection is determined by the available information; a control decision occurs as a result of information acquisition and processing. Information is needed for decision making and is inseparable from it. The quality of information is determined by the quality of the solution it provides. The quality of the solution is determined by the goal and the degree of its achievement. The process of exploration cannot be described apart from the information acquisition and processing that shape the decision-making process. As a decision-making process, the resource exploration process has a number of unique properties.

An analysis of exploration history and resource development can be considered a typical example. In the absence of substantial discoveries, drilling volume grows slowly. The first large discoveries cause a significant increase in drilling, which, in turn, results in more discoveries, leading to an even greater increase in drilling. These increases in drilling continue until it is determined that the growth in drilling volume will not lead to new substantial discoveries. Thus, the means of control are not set from the beginning but change in response to the results of the exploration process. An increase or decrease in the amount of exploration in a region depends on the results obtained (the scope of the discoveries); in turn, the amount of exploration affects subsequent results. It turns out that reserve accrual depends on our actions, which depend on information accumulated, or on the accumulated reserves. A type of relation where the control action is a function of the state of the system is called "feedback." Therefore, the process we are now considering is a process with feedback.

In exploration, the feedback structure cannot be accurately established in advance as in the case of control over a technical system. For example, the speed, position in space, flight direction, and other parameters pertinent to controlling an aircraft can be evaluated with a certain amount of accuracy. However, exploration control is accomplished under uncertain conditions. In such a situation, the so-called feedback reflectivity disappears, that is, an act of decision making does not necessarily lead to a single result. This type of result cannot be represented as a simple function of external actions (expenditures, drilling extent, etc.) and the state of the system (the status of exploration, the degree of knowledge of an area, etc.). In this respect, exploration may be considered a probabilistic process.

When examining the probabilistic nature of exploration, the following observations are necessary. As a result of the exploration process the maximum size field will be discovered. The time of such discovery, however, cannot be predicted exactly, if at all. The same goes for predicting

the reserve size of the largest field since it is also unknown and can only be predicted based on certain correlations.

Likewise, it is known that a maximum in the reserve accrual will be reached in the process of exploration. However, the timing of the maximum accrual and size are not known. The probability of these events also cannot be determined because we are not dealing with a mass of random trials, but with a single controlled process. The pattern of discoveries is affected not only by natural conditions but also by the extent to which the evolution of the exploration process (which is determined by man and not by nature) is in agreement with the natural conditions. Depending on the degree of this agreement, or on the exploration strategy, these uncertainties may rapidly decrease or may stay the same as exploration continues. The very system of this agreement can be perfected in the process of exploration. All of these factors cause continual changes, rapid or slow, in the respective uncertainties.

While discussing the probabilistic, random nature of the exploration process, it makes sense to address information quality. It is often evaluated statistically, in terms of statistical reliability, confidence interval, the probability of a variable reaching a certain value, and so forth. Information in the process of exploration is used for decision making; therefore, the evaluation of information is inseparable from decision making. Decision making is associated with individual actions, not mass phenomena. Thus, as a general rule, decision making deals with single selections. For this reason, information on the parameter's statistical properties, which characterize a partially known environment, is practically useless. A probabilistic description is not of much help in this situation because the information quality is described not only in terms of a confidence interval with guaranteed upper and lower limits, but also in terms of the uncertainty in achieving a goal depending on the length of this interval.

We have described the nature of exploration as a process in which the feedback structure cannot be precisely determined and where control is accomplished under uncertain conditions. Therefore, a preprogrammed method of achieving goals cannot be used. A program is a way to use the control resources for the achievement of a goal. The program includes a list of actions (jobs) along with the requisite evaluation of resources needed to accomplish these actions (jobs). If the process trajectory deviates from that programmed, the process is corrected and returned to the conditions specified by the program. Exploration differs in that as it evolves, the concepts of the geology and petroleum potential in the region that originally laid the groundwork for exploration are changing, and new, unforeseen situations occur. It is possible that rather than leading to the achievement of the goal, the programmed actions will now obstruct it because the goal has changed. So correction now consists, not of returning

to the planned program, but of adapting to the new situation because it turns out to be preferable. Therefore, in order to make control more efficient, exploration must be considered a multi-phase process. In this process the achieved results determine our actions. This process may be called a self-organizing process.

As with any controlled process, the concept of quality control may be applied to exploration. Better quality is determined by decisions that provide for better achievement of the same goal. The structure of goals has been mentioned previously. Depending on the goal, additional quality control criteria may be introduced. If, for instance, the goal is to fulfill a planned reserve accrual, then expenditures (drilling footage, etc.) are usually used as such a criterion. When developing a region as a whole, the timing of transferring the total ultimate potential into the appraised reserve category, or the time duration until the largest discovery, may be considered as such a criterion. The selection of such criteria has not been adequately studied. As a rule, these criteria are economic, but there are definitely other possibilities. Social, political, and other considerations may be important. A drive to provide better quality in exploration control opened new directions for optimizing exploration.

All this helps establish the preconditions for a foundation for modeling the history (evolution in time) of oil and gas reserve accumulation. Currently, exploration modeling is based on a variety of concepts. An approach in which the probability of a deposit's discovery is evaluated, depending on its parameters, is quite common. Sometimes this probability is associated with the extent of exploration (expenditures, drilling footage, etc.). In this case, the assignment of a multi-dimensional probability function to the target discovery requires a large amount of information that is difficult to provide.

It should not be forgotten that the probabilistic characteristics of parameters for a situation which is not completely known make sense only for multiple, repetitive events. If a single choice must be made, such information is practically useless. This approach ignores the main feature of the exploration; namely, that exploration is a controlled, intentionally organized process, albeit in an unclear environment. The exploration process is directed by the achieved results; a consideration of the achieved result is precisely what makes it an organized process. In the probabilistic approach, it appears that discoveries are not the results of our actions, but random events that depend on our actions only to the extent of the applied exploration efforts. The information accumulation and processing that are part of exploration are completely ignored. Strategy selection should always be based on information, which is crucial for decision making. Decision making is an individual action that derives little benefit from using probabilistic descriptions.

We believe that the approach based on the Shannon information theory is also misplaced. The Shannon theory deals with the processing of certain meaningful information. Entropy, as a measure of uncertainty, makes sense for processing information already gathered, as in the case where we use this information to characterize the resource parameters and structure at certain moments of time. Classical information theory does not explain what is most important to us: the process of information creation. It does not address the fact that information is gathered, changed, and acquires value in the process of exploration. The Shannon theory describes the process of communication in a fuzzy environment, but does not provide for information that is suitable for decision making. The evaluation of information cannot be separated from the problem of choice. Only the degree of goal achievement may be used as a measure of information quality. Entropy cannot be used here.

For this reason, a formal description of reserve accumulation should be based on a concept of exploration as a controlled process with feedback in which control is accomplished based on information gathered and processed during the system evolution. Feedback is a selection mechanism for deciding action based on the state of the object; that is, the control action, or the control function, is defined as a function of the system state. Therefore, reserve accrual should be considered a function of the control action. Likewise, the control function should be considered a function of the accumulated reserves (the reserves determine the state of the process from the information standpoint, among the other possible perspectives).

If the accumulated reserves are R, then dR/dT is the rate of accumulation. Our statement can then be expressed as follows:

$$dR/dT = f(E); \quad E = \varphi(R) \tag{1-1}$$

where E is the control function.

The selection of specific functions will be considered later; however, the following should be emphasized. In exploration, decisions are based on incomplete, unclear, or distorted information. Decision making is performed by people at different levels of the control system, people with different intellectual capabilities, different ideas, and sometimes different goals. As a result, the reserve accrual cannot be represented as a single function of external actions and the state of the resources. This distortion causes fluctuations in the reserve accrual, that is, spontaneous deviations from some average value. This change in the average value represents the general trend with which a deterministic reserve accrual description is associated. This description reflects patterns in the reserve accrual changes. Fluctuations may be considered random events governed by probabilistic laws. Therefore, a description of the reserve accumulation

process should incorporate both deterministic and random components that relate to each other in some way.

The following is a discussion of this issue.

Substantiation of Functional Relations

Regardless of the specific selection mechanism mentioned earlier, the reserve accrual can be described by some balance equations of the following form:

$$A = R + U \tag{1-2}$$

where A is the ultimate potential resource (or ultimate potential), R is the initial accrued reserve, and U is a "forecast," or undiscovered, resource.

This equation is always true, regardless of possible different evaluations of its components at different times. The other equation results from the fact that a certain part of a field's reserves is transferred into the appraised category. Let us suppose that all the oil and gas fields are subdivided into classes by reserve size, with a class of reserves being Q_i, where i is the number of a particular class; let us further assume that the portion of the reserves appraised during a given year (k_i) and the number of fields being appraised during a given year (m_i) depend only on the class. Then the balance equation can be written as follows:

$$\Delta R = \sum_i m_i k_i Q_i, \quad \text{or} \quad \Delta R = \sum_i (N_i - p_i) k_i Q_i \tag{1-3}$$

where N_i is the number of i-class fields discovered since the beginning of exploration, p_i is the number of i-class fields appraised since the beginning of exploration, and $\Delta R = R_{i+1} - R_i$ and R_i are the reserves accrued by the year i.

Balance equations provide some information about the studied process but do not give its complete description. They do not provide a means for determining a dynamic series of its sequential states based on its initial state. The incomplete nature of the model is a result of the fact that neither the structure of interrelations between the variables or the trend of change of this structure is taken into consideration. We need to identify a mode of inductive analysis and visualize the structure of hypotheses required to complete the model. In order to do this, we need to understand the motivations behind the actions. We also need to take into account factors causing changes in the variables of the balance equation. In other

words, we need to describe the process of the information inflow and, most importantly, the process of decision making.

Earlier we mentioned the structure of resources. Currently, there are various hypotheses concerning types of oil and gas field size distribution functions. Regardless of the type, ultimate potential is distributed among the fields in an extremely non-uniform manner, with a disparity of several orders of the magnitude between the smallest and largest fields. Many believe that the bulk of the ultimate potential is concentrated in large fields. Others believe that the resources concentrated in large and small fields are comparable. For instance, one study concluded that the large oil field reserves in mature North African oil and gas basins do not exceed 50% of the ultimate potential and are usually less than 30% [42]. The same patterns of resource distribution are observed for gas. At the same time, 44.8% of the oil and 39.4% of the gas are derived from small fields. The important issue is that the number of small fields is great and the number of large and giant fields is small.

Obviously, it is better to achieve a substantial reserve accrual by discovering a few large fields than a great number of small fields. Large fields are usually associated with large structures; they are major exploration targets and are practically impossible to miss wherever exploration is conducted. This is not so with small fields. Although their number is great, their contribution to the reserve accrual is small. When there are no more large fields left in a region, the reserve accrual declines. Thus, the conditions of reserve accrual, the conditions of field "selection," or the selection of methods of action changes in time. They change with the process of field discovery and appraisal. This is why the reserve accrual must be a function of the selection conditions.

The idea of "selection conditions" should also take into account that the number of fields on which selection is conducted, monotonously decreases beginning with the first discovery. The selection conditions change in the process of exploration, especially when large fields are discovered. In turn, large fields not only provide a high level of reserve accrual but also a drastic increase in the accrued reserves R. While only small fields are discovered, the reserve accruals and the R value are small. This is why the function of the selection conditions should be a function of R. A single large field makes a significant contribution to the reserve accrual, whereas a single small field makes an insignificant one. However, the number of small fields is greater. Therefore, the selection conditions are indicated by the size of the reserve accrual per unit of the accrued reserves, or $\Delta R/R$.

Exploration is a controlled process, that is, controls must be adjusted to conform to changing selection conditions. Therefore, we introduce one more function, namely, control of the selection. Control is based on the information obtained; information is used to focus the process in a

required direction. This direction is defined by discoveries or the reserve accrual. Indeed, exploration is conducted in such a way that control of exploration is a reaction to fluctuations that occur in the reserve accrual. Let us assume that exploration has begun in a region with many small fields and a few medium fields. The reserve accrual and the value of the accrued reserves in such an area will be low, and the potential of the region will be deemed low. Consequently, the capital investments will be low and the region's development (exploratory drilling in particular) will be slow. This will limit the flow of information about the region, which, in turn, will make the potential forecast and the region's evaluation (hence, exploration control) more difficult. The reserve accrual fluctuations will be small; but this state is unstable. The discovery of an average or a large field would generate a drastic change in the reserve accrual. The estimate of the region's potential would change. Exploration of the new, more important plays would be conducted. The investment in the region's development, including exploratory drilling, would increase. Consequently, the amount of information on reserve exploration would increase and control would improve.

This state, however, is also unstable. The depletion of large discoveries would create a new concept of the region's potential with all the ensuing consequences. The reserve accrual would decline, consequently, drilling extent would drop. At the same time, knowledge of the region would increase, the amount of accumulated information would grow, and the acquired experience in exploration (control) would also increase.

Thus, each new state occurs as a result of the reserve accrual fluctuations around its average value. These fluctuations are what directs the control system: they change the concept of the systems and, consequentially, the control action and control resources.

This indicates that a function of selection control should also be a function of the accrued reserves R.

Therefore,

$$\Delta R = \varphi(R)f(R) \tag{1-4}$$

where $\varphi(R)$ is the function of selection control and $f(R)$ is the function of selection conditions.

Before discussing a specific form of the functions $\varphi(R)$ and $f(R)$, it is important to note the following. We identified the factors of the reserve accrual and Equation 1-4 summarized them. As written, they do not explicitly include the accumulated drilling extent as a parameter. At the same time, the reserve accrual dynamics (evolution) is usually considered a function of the drilling amount. The drilling volume is not an independent, arbitrarily selected value. As previously shown, it depends on the

results of exploration. The drilling volume itself, to a certain extent, determines the exploration results (the reserve accrual). Thus, these parameters are interconnected, and the amount of drilling was taken into account through the R value.

Specifics of the Mathematical Description in this Study

The following must be taken into account when constructing the models. The process is usually modeled by a differential equation. A solution of this equation is a continuous deterministic function, which describes the flow of the process in time given the mechanism of its formation. Based on this, instead of the reserve accrual ΔR, we actually should have applied the first derivative, $r = dR/dt$, or the instantaneous velocity of accrual of the initial appraised reserves R.

The solution of any differential equation is a continuous function, whereas any detailed study inevitably requires a discrete description, which is convenient for computer simulation.

As for the reserve accruals, they may only be represented discretely because they are specified over equal time intervals (one year). In this case, the multitude of observations form a time series where the observations are not independent. In this situation, the analysis of continuous processes is replaced by the techniques used for describing discrete structures. One such technique is the analysis of time series. Models of time series describe the systems under study through the use of difference equations rather than differential equations. In these equations, a difference operator replaces a differential operator. Therefore, one can introduce a differential equation (continuous process) and its discrete equivalent (for the same process in discrete time). Discrete process models are often written in the form of equations reflecting dependence of a subsequent term of a time series on the preceding terms.

Most cases in this study use finite difference equations to describe the reserve appraisal process. It appears that difference models have not been used in geological studies before, although due to the discreteness of geologic observations, such models should have been a major mathematical tool.

We will designate $R(t)$ as a continuous deterministic function describing the reserve accumulation in time; its derivative will be designated $r(t)$. $R(t)$ values at discrete time moments i separated by the time intervals $\Delta t = 1$ will be designated R_i, and the reserve accrual during the $i + 1$ year will be designated $\Delta R = R_{i+1} - R_i$. Evolution of the reserve accumulation process is written as an equation describing the change in ΔR, or as a connection equation $R_{i+1} = f(R_i)$. This equation makes it possible

to determine subsequent R_i for discrete time moments (i = 1, 2, 3, . . .) beginning with some R_o (corresponding to i = 0).

We will now discuss the form of $\varphi(R)$ and $f(R)$ functions in Equation 1-4. The above discussion of the selection control function $\varphi(R)$ indicates that it should be a function growing with an increase in R. It may be suggested that it grows linearly with R:

$$\varphi(R) = kR \tag{1-5}$$

To better elucidate the selection conditions of the function $f(R)$ form, we will introduce the specific rate of the reserve accumulation and the relative reserve accrual. Specific rate r/R is an instantaneous reserve accrual per unit of the accumulated reserves, or the accrual "created" by a unit of the accumulated reserves. We will call $\Delta R/R$ the relative reserve accrual. It is similar to the specific rate of the reserve accumulation r/R. The difference between them is that the former represents ΔR, which reflects the average rate of the reserve accumulation over one year (Δt = 1); the latter represents instantaneous velocity $r(t)$, whose value does not have a real meaning because the reserves are integrated over the respective time interval (one year).

Actually, the assumption of the conditions of selection function $f(R)$ form is reduced to the assumption of the type of change for the relative reserve accrual $\Delta R/R$.

The following must be kept in mind. The selection is made among chosen exploration targets. Exploration targets in all regions analyzed in this book are oil and gas accumulations associated mainly with anticlinal traps. In the past, the entire system of exploration ignored non-anticlinal traps, as if they did not exist. By the same token, exploration is always directed toward technologically accessible depths or toward a certain stratigraphic interval with which the region's potential is believed to be associated. Therefore, the selection is conducted only among the targets within these particular intervals.

These elements determine the given situation for the selection. Let us assume that different types of targets are beginning to be explored and evaluated. For instance, they may be non-anticlinal traps that require different exploration techniques. Or let us assume that, after some interruption of the exploration process, deeper, older stratigraphic intervals are involved in exploration. These intervals may not have been exploration targets for many years and their resource development requires different technologies. This would create a different, autonomous selection situation.

For instance, early in this century, exploration was conducted in Checheno-Ingushetia for Middle-Miocene sediments within the Tersk and Sunzha anticlinal zones, and later (in the 1950s) for the Cretaceous within

the same zones. We have two different selection situations here. The concept of a "selection situation" is similar to that of an "exploration play" [41]. Here we are discussing a target selection out of a certain set of possible targets. This set includes only those targets that are being explored. Some of these targets are accepted and others are purposefully discarded (i.e., not due to objective causes). The goal and the means of control in this case are inseparable.

The spacial position of these targets (including spacial boundaries) is mostly determined by the exploration region to be controlled, that is, again, by the goal. If the intent is to control exploration within a certain region, then the space is limited by the region's boundaries. The term "exploration play" has a different meaning.

> An oil and gas exploration play [direction of exploration] is a combination of similar fields (discovered, undiscovered and postulated), which are being explored and appraised using common techniques (and technologies), which are located within a single petroleum basin and within a single tectonic zone (which may include several adjacent structural elements). [41, p. 25]

The quoted authors add that:

> . . . the criterion of whether the volume of an exploration play has been correctly defined is the applicability to a given set of accumulations (or to a geologic space containing this accumulation set) of the M. K. Hubbert models. [41, p. 26]

Prior to this statement, while explaining the concept of the exploration play, the authors state a need to

> . . . introduce an additional concept, which would limit the three-dimensional oil-gas potential space in such a way that the function
>
> $$y = cxe^{-(ax^2+bx)}$$
>
> describing exploratory drilling efficiency would be applicable to this space. [41, p. 25]

These three quotations are incompatible. The reader is left guessing what an exploration play really means and whether the dynamics are secondary (i.e., reflecting the resource exploration system) or primary (i.e., defining the set of targets).

The introduction of the play concept had the same goal we are pursuing when discussing the selection situation. This goal is to outline the boundaries of the situation under consideration. Throughout Part I of this book we will be modeling the reserve accrual in a given selection

situation. The case of a combination and/or superposition (or rather partial overlap) of different selection situations in time is not considered. Under the specific selection situation, the accumulated reserve graph has an inflection point. Consequently, the reserve accrual curve has a single maximum. All the considered cases of two or more maxima are due to the occurrence in those regions of a different selection situation; that is, deeper horizons began to be explored at the time when exploration of shallower horizons was already declining. Such a new selection situation in the region causes a new increase in exploration and new discoveries (including large ones), but already under different conditions. For this reason, we will be considering the reserve accrual curves with only one maximum.

Exploration in a region of interest stops when the ultimate potential has been transferred to the appraised category. This means that the $R(t)$ function has an upper limit or asymptote:

$$A = \lim_{t \to \infty} R(T)$$

The evaluation of A correlates with the evaluation of the ultimate potential (R_{init}). We believe, however, that the presence of an asymptote is not necessary. If the $R(t)$ function has an asymptote, ultimate potential resources can be evaluated based on observations. Otherwise, an additional criterion will be needed for its evaluation. This criterion should define the time of voluntary termination of the process. Such a criterion may be based on economic considerations such as the minimal commercial size of a single field and the minimal commercially feasible reserve accrual. We will discuss this in detail while determining specific models.

All this indicates that the selection function $f(R)$ does not have to converge to zero at some cut-off value of R. In fact, it may approach R asymptotically.

Thus, the selection condition function $f(R)$ must satisfy the following requirements:

- It may not monotonously increase with the increase of R and it may not be identically equal to a constant.
- It must be a monotonously decreasing function, at least beginning with some sufficiently small R value.
- It must asymptotically approach the R axis or intersect this axis at the point where R reaches its maximum.

Consequently, the product of two functions, $\varphi(R)$ and $f(R)$ which represent opposite tendencies (one is decreasing when the other is increasing),

defines a deterministic trend in the reserve accrual ΔR accrual with a typical maximum (Figure 1-1).

This maximum is reached at different R values. For instance, based on an analysis of world data, three groups of countries and regions were identified, with efficiency maximum reached, respectively, at 10 to 20%, 30 to 45% and 55 to 75% of the resource development [15]. If we take into account what was previously mentioned regarding the deterministic reserve accrual through fluctuations, this data would most likely indicate that the type of selection condition function $f(R)$ is different for different regions and that there is no single, universal $f(R)$ function. The development history of any region is specific and, to a certain extent, affected by the spontaneous mechanisms inherent in any system of exploration control.

Exploration strategy is of great importance under these circumstances. If from the very beginning exploration is conducted over the main plays, it will result in the rapid discovery of large fields. Otherwise, the achievement of large reserve accrual may be postponed for a long time. This means that the model is never capable of explaining the precise historical mode of a region's development. "Never" is a result of the informational and stochastic nature of the studied process and of the extreme variety of available selections. It also covers the prediction of further development beyond certain time limits. It would not be an overstatement to say that all innovations in exploration (new geophysical techniques, integration of geophysics and drilling, optimization, accelerated exploration methods, etc.) are directed toward limiting the available selections

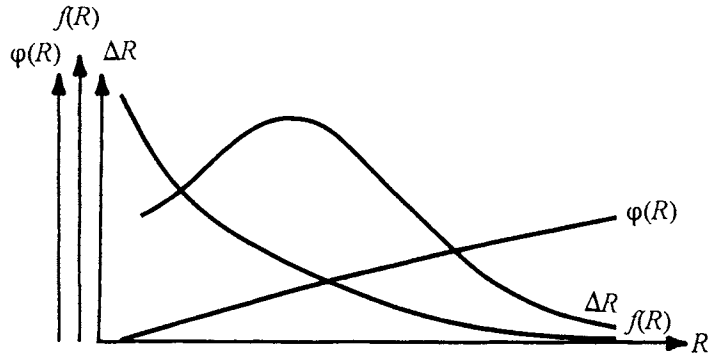


Figure 1-1. Representation of the accrual of reserves as a product of two functions [$\Delta R = \varphi(R)f(R)$], which show opposite tendencies with the growth of R.

and toward accelerated acquisition of more complete information for reserve evaluation.

The above considerations are still insufficient for determining the specific form of the $f(R)$ function. It can only be hypothesized. Changing the structure of hypothetical equations, we can develop an entire family of possible $f(R)$ functions. The multitude of these functions correlates with the multitude of possible models. Therefore, modeling in this environment can be reduced to the search for the best description, that is, the selection of the best model out of the set of possible models. As previously discussed, the set should be sufficiently complete to reflect the multifaceted nature of the process. If so, then the model will serve in various regions as an efficient means of typification and classification of the evolution of the reserve accumulation curves, and will serve as a means of understanding how a particular curve type evolved. These results will be a means to uncover the specific nature of the reserve accumulation process in a particular region as associated with the exploration strategy and the regional resource structure.

We will be analyzing the reserve accumulation evolution from different viewpoints: in time as well as comparing the expended exploratory drilling footage. The reserve discovery and appraisal process is considered, by its very nature, a process that combines deterministic and stochastic features. Consequently, the models combining the properties of these features will be constructed and discussed.

CHAPTER 2

Evolution of the Reserve Accumulation: Deterministic Models

The deterministic models describing the evolution of the oil and gas reserve accumulation process developed in a step-by-step manner. At each consecutive stage new models were introduced, prompted by the fact that the previously constructed models for one reason or the other were unsatisfactory or not completely satisfactory. Models created at each new stage were not always radically different from those constructed in the preceding stage.

As a result, three groups of deterministic models describing evolution of the the oil and gas reserve accumulation process were constructed:

1. Models interpreting the function of the selection conditions;
2. Models based on the transfer intensity of the undiscovered potential into appraised reserves;
3. Models of field selection quotas.

All three groups of models give different interpretations of the reserve discovery and appraisal strategy. Earlier we introduced a concept of relative reserve accrual that was associated with the function of selection conditions. The models in this chapter are based on this concept.

Study of the Reserve Discovery and Appraisal Strategy as a System of Selection

Construction of Models

There are different ways to describe the reserve discovery and appraisal strategy and its reflection in the reserve accumulation evolution. One

possible way is to describe the strategy as a system of selection where the form of the function of selection conditions serves as a system attribute. For the construction of these models, the well-known Gomperz and logistical functions were used. Earlier it was indicated that these functions have often been used for studies of oil and gas reserve evolution. This analysis led to the construction of models that visualize the reserve discovery and appraisal strategy as a system of selection. Common to all these models is the idea that the selection conditions function (or the relative reserve accrual) is a decreasing function of R. Three radically different cases are considered: the selection conditions function linearly decreases with the growth in R; it first rapidly, and then slowly decreases; and the other way around, it first slowly, and then rapidly diminishes. The reserve accrual maximum correspondingly shifts from the center toward low or great R values (Figure 2-1).

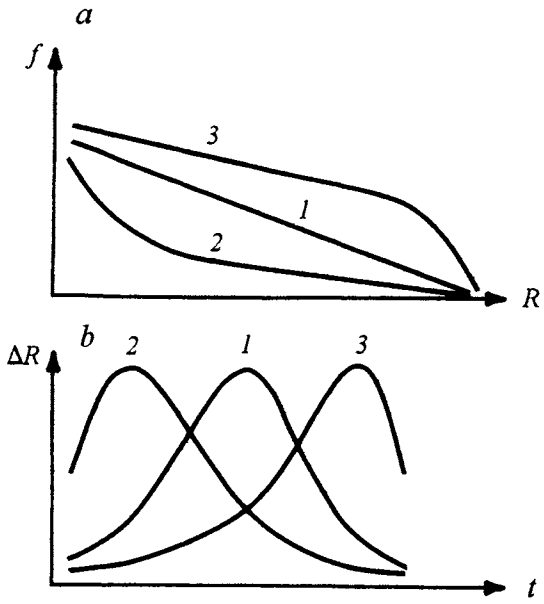


Figure 2-1. Change in the function of selection conditions [$f(R)$] (a) and corresponding accrual of reserves [ΔR] (b) at a constant rate (1), with deceleration (2), and with acceleration (3).

Model 1. The simplest assumption for changes in the relative reserve accrual (the selection conditions function) is that this function linearly decreases with the growth in R:

$$\Delta R/R_{i+1} = a - bR_i \tag{2-1}$$

This equation correlates, in a unique way, the relative reserve accrual during a year with the reserves accumulated by the beginning of the year (or the end of the previous year). From Equation 2-1:

$$R_{i+1} = R_i/(1 - a + bR_i) \tag{2-2}$$

This equation expresses the same relationships and describes the reserve accumulation as a discrete process.

The continuous function $R(t)$ is a logistic function in which the values of equidistant points separated by the time interval Δt (for instance, $\Delta t = 1$) form a discrete time series complying with Equation 2-2. Its form is as follows:

$$R(t) = \frac{A}{1 + \left(\frac{A}{R_o} - 1\right)e^{-kt}} \tag{2-3}$$

where

$$A = \lim_{t \to \infty} R(t)$$

is the upper limit of the accumulated reserves, or the asymptote of the $R(t)$ curve, and R_o is the accumulated reserves at the time $t = 0$.

The logistic function is a general solution of the differential equation

$$r/R = k - \varepsilon R \tag{2-4}$$

which is the Verhuhlst-Pearl's equation, frequently used to describe the growth of biological populations and usually called the logistic equation. The εR is inserted here to reflect the influence of the limiting factors. In the absence of the limitation, an integration of the equation $r = kR$ would produce the exponential (base e) growth $R(t) = R_o e^{kt}$. In the environment of exponential (base e) growth (i.e., without limiting factors), the specific rate is constant. In the process of growth, limiting factors begin to work at some moment in time, and the rate begins to decrease. It may be demonstrated, as application to the reserve accumulation evolution, that

if there is any exponential growth, it can only occur at the first stage of exploration, when R is small. As the reserves are accumulated, the conditions for new discoveries and for maintaining a high, constant rate of the reserve growth (the selection conditions), become less favorable. As a result, the reserve accumulation decreases and deviates from the exponential curve. This slowdown may be associated with various limiting factors. In the logistic model these factors are proportional to the amount of accumulated reserves. Thus, the logistic function is based on the assumption that the specific rate of the reserve accumulation decreases linearly with the growth in accumulated reserves.

The differential Equation 2-4 shows that the logistic curve may be considered an exponent with a variable parameter k: the value of k gradually decreases by a value proportional to the reserves accumulated by a given time. The k and ε parameters in the equation are constants (for a given region) and have certain physical characteristics: k is a specific rate with no limitations; ε is the limiting coefficient. Interconnections between the parameters of the differential Equation 2-4 and the parameters of its solutions (Equation 2-3) are presented in Table 2-1.

Therefore, Equation 2-1 shows that the logistic dependence assumes a linear decrease of the relative reserve accrual with an increase in reserve accumulation. This is exactly the equation describing the discrete process of the reserve accumulation in correspondence with the logistic function. Equation 2-1 is a discrete equivalent of the differential Equation 2-4. Parameters of Equations 2-1 and 2-4 are connected with each other and with the parameters of the logistic curve. At $\Delta t = 1$ this connection is presented in Table 2-1. The logistic curve is a symmetric S-shaped curve. Maximum rate of the reserve accumulation (maximum reserve accrual) is achieved at the time $t = t_{\max}$ as defined by the following equation:

$$t_{\max} = \frac{\ln(A/R_o - 1)}{k} \tag{2-5}$$

When a value of $t_{\max}$ changes, the $R(t)$ or $r(t)$ curves move to the right or left. Therefore, R_o determines the position of the curve relative to the time axis.

The maximum rate of the reserve accumulation (maximum reserve accrual) is:

$$r_{\max} = r(t_{\max}) = kA/4 \tag{2-6}$$

And the accumulated reserves are:

$$R_{infl} = R(t_{\max}) = A/2 \tag{2-7}$$

Thus, by the moment t_{max} when the maximum reserve accrual is achieved, one-half of the ultimate potential will be appraised; that is:

$$\frac{R_{infl}}{A} = \frac{R(t_{max})}{A} = \frac{1}{2} \tag{2-8}$$

This equation shows that the logistic curve $R(t)$ is symmetrical with respect to the inflection point, and its derivative $r(t)$, or the reserve accrual curve, is symmetrical with respect to its own maximum.

The relative rate at the moment t_{max} when the maximum is reached is equal to:

$$\frac{r_{max}}{A} = \frac{r(t_{max})}{R(t_{max})} = \frac{k}{2} \tag{2-9}$$

If one switches from the logistic function $R(t)$ to the logistic distribution $F(t)$ using $F(t) = R(t)/A$, then one will be able to consider the expectation, mode, median, and variance of the distribution. Due to the symmetry of the distribution, the expectation, mode, and median coincide and are equal to t_{max}. The variance is evaluated by:

$$\sigma^2 = \pi^2/(3k^2) \tag{2-10}$$

This shows the meaning of the k parameter. When it increases, the standard deviation decreases, the probability distribution becomes more concentrated, and the probability density in the expectation point increases.

Model 2. This model is also based on the assumption that the relative reserve accrual decreases in time:

$$\Delta R/R_{i+1} = 1 - a' R_i^b \tag{2-11}$$

As the accumulated reserves grow, the relative accrual decreases nonlinearly and the intensity of the limiting factors (the deterioration in the selection conditions) increases exponentially. Therefore,

$$R_{i+1} = a' R_i^b \tag{2-12}$$

where a and b parameters are related to a' and b' parameters in a certain manner (see Table 2-1). A continuous function $R(t)$ corresponding to Equation 2-12 is the Gomperz function, which has been previously used for describing of the reserve accumulation evolution. Equation 2-12

indicates that the reserves are accumulating according to the Gomperz law when the reserves accumulated by the end of a year form an exponential function of the reserves accumulated by the beginning of that year (i.e., by the end of the preceding year). The Gomperz function represents a solution of the differential equation:

$$r/R = \varepsilon e^{-kt} \tag{2-13}$$

Here, the decrease in the specific rate of the reserve accumulation is associated not with the amount of accumulated reserves, but with time. It is assumed that the decrease is exponential (base e) rather than linear.

The Gomperz function has the following form:

$$R(t) = Ae^{-me^{-kt}} \tag{2-14}$$

where A is the upper limit of the accumulated reserves (asymptote) and m is a parameter associated with R_o (see Table 2-1).

In geological publications this function is called the Gomperz-Makeham function. In other disciplines it is simply known as the Gomperz function, although the function derived by Gomperz had a somewhat different form, as did the differential equation he used to model life expectancy [57]. However, similarity with his mortality intensity expression made it possible to write the reserve accumulation equation in the form of Equation 2-13.

Despite this, we will call the above $R(t)$ function the "Gomperz" function. It represents a non-symmetrical S-shaped curve with initial exponential growth, as in the logistic curve. Its derivative $r(t)$ reaches its maximum (in a discrete situation, the maximum accrual) at the time $t = t_{max}$, which depends on the parameters of $R(t)$ and is determined based on the following condition:

$$e^{kt_{max}} = m, \quad \text{i.e.,} \quad t_{max} = (\ln m)/k \tag{2-15}$$

Therefore, the m parameter determines the curve position in the time dimension (a change in the m value leads to an offset of the $R(T)$ or $r(t)$ curves to the right or to the left).

The maximum rate of the reserve accumulation (maximum accrual) is determined from:

$$r_{max} = r(t_{max}) = kA/e \tag{2-16}$$

and the reserves accumulated by the time $t = t_{max}$ are:

$$R_{infl} = R(t_{max}) = A/e \tag{2-17}$$

Table 2-1
Different Presentations of the Studied Deterministic Models for the Evolution of the Reserve Accumulation in Time

Model	Analytical Expression	Recurrent Relation	Differential Equation	Relative Accrual	Connection of Parameters
1	$(t) = \dfrac{A}{1 + (A/R_o - 1)e^{i}}$	$R_{i+1} = \dfrac{R_i}{1 - a + bR_i}$	$\dfrac{r}{R} = k - \varepsilon R$	$\dfrac{\Delta R}{R_{i+1}} = a - bR_i$	$a = 1 - e^{-k}$ $b = \dfrac{1 - e^{-k}}{A} = \dfrac{\varepsilon(1 - e^{-k})}{k}$ $A = a/b = k/\varepsilon$
2	$R(t) = Ae^{-me^{-kt}}$	$R_{i+1} = aR_i^b$	$\dfrac{r}{R} = \varepsilon e^{-kt}$ $r = -kR\ln(R/A)$	$\dfrac{\Delta R}{R_{i+1}} = 1 - a'R_i^{b'}$	$a = 1/a'; b = 1 - b'$ $a = A^{1-e^{-k}}; b = e^{-k}$ $A = a^{\frac{1}{1-b}} = \left(\dfrac{1}{a'}\right)^{1/b'}$ $\varepsilon = mk; m = \ln(A/R_o)$

3	Not obtained in an explicit form	$R_{i+1} = \frac{R_i}{1 - e^{a-bR_i}}$	—		$\frac{\Delta R}{R_{i+1}} = e^{a-bR_i}$	$R_{\lim}$ value satisfies the following equation: $\frac{R_{\lim}}{e^{bR_{\lim}}} = \frac{\Delta R_{\lim}}{e^{a+b\Delta R_{\lim}}}$
4	$R(t) = \frac{A}{1 + \left(\frac{1}{(R_o/A)^{1/b}} 1\right) e^{k_i/b}}$	$R_{i+1} = \frac{R_i}{(1 - c + aR_i^{1/b})^b}$	$\frac{r}{R} = k - \varepsilon R^{1/b}$		$\frac{\Delta R}{R_{i+1}} = 1 - (1 - c + aR_i^{1/b})^b$	$c = 1 - e^{-k/b}$ $a = \frac{1 - e^{-k/b}}{A^{1/b}}$ $A = \left(\frac{c}{a}\right)^b = \left(\frac{k}{\varepsilon}\right)^b$
5	Not obtained in an explicit form	$R_{i+1} = \frac{R_i}{1 - c + aR_i^b}$			$\frac{\Delta R}{R_{i+1}} = c - aR_i^b$	$A = \left(\frac{c}{a}\right)^{1/b}$

The last two equations show that the specific rate of the reserve accumulation at the moment of the maximum rate ($t = t_{max}$) equals:

$$\frac{r_{max}}{R_{infl}} = \frac{r(t_{max})}{R(t_{max})} = k \tag{2-18}$$

which suggests one possible interpretation of the parameter k. Studies of the reserve accumulation process have usually paid great attention to the reserve share accumulated by the time the maximum reserve accrual is reached. This share can be calculated from:

$$\frac{R_{infl}}{A} = \frac{R(t_{max})}{A} = \frac{1}{e} \cong 0.37 \tag{2-19}$$

Thus, in the model under consideration, this share does not depend on the parameters of the $R(t)$ curve, but is a constant. It is a characteristic of the curve's $r(t)$ asymmetry.

Equation 2-12 is a discrete equivalent of Equation 2-13. Their parameters are interconnected and also connected with the Gomperz function parameters (at $\Delta t = 1$ this connection is shown in Table 2-1).

Model 2 may also be connected with another mechanism that may cause the above Equations 2-11 through 2-13. To understand this relationship, we will examine the function $F(t) = R(t)/A$, the values of which range between 0 and 1. In this case, $F(t)$ may be taken as a function of the probability distribution. In probability theory, the derivative dF/dt is called the distribution density and is denoted $f(t)$. Thus, model 2-13 can be written as follows:

$$f(t)/F(t) = \varepsilon e^{-kt} \tag{2-20}$$

which yields a distribution function $F(t)$ in the following form:

$$F(t) = e^{-me^{-kt}} \tag{2-21}$$

This is a distribution function of the maximum values from the collection of probability distributions of the exponential type [12]. Among these exponential distributions are normal, log-normal, logistic, χ^2, and γ distributions.

Gomperz himself used the following equation:

$$f(t)/[1 - F(t)] = \varepsilon e^{-kt} \tag{2-22}$$

Gomperz called the expression on the left the "mortality intensity." Today it is called the "intensity function," or the "risk function," and the function $1 - F(t)$ is called the "survival function." The distribution function derived here represents a function of minimum values distribution in the collection of distributions of the exponential type. It is symmetric to the maximum value distribution function.

Therefore, reserve accumulation according to the Gomperz law may be considered a distribution of the random variable t from a probabilistic standpoint. This requires some interpretation.

Consider an abstract unit of reserves—a unitary reserve concentration. Let us introduce a concept of the "easiness of discovery" of this singular accumulation (we will associate the reserve accumulation process with the field discoveries). The degree of "easiness" will be measured by the time elapsed from the beginning of exploration in the region to the discovery of that singular accumulation (the easier it is to discover the field, the sooner it will be discovered). Obviously, the degree of "easiness" for a determined exploration strategy will depend on how the singular accumulations are located. They are easier to discover if they are combined in large groups (or fields), are located at accessible depths, and are associated with the conditions for which exploration is targeted (for instance, anticlinal traps). The most significant factor is a combination into fields. The larger the field, the less time t will take to discover it, with all other conditions equal, and, therefore, to discover singular accumulations of which it is comprised. The larger the field reserves, the greater the number of singular accumulations of which it is comprised. The probability of singular accumulation discovery with the easiness less than a given t is defined by their share of the total number of singular accumulations. This is the reserves (their share of the limit A) accumulated by the time t.

Let us assume that the initial (natural) distribution of singular accumulations by the degree of easiness (i.e., discovery time t) is a distribution of the exponential type. If the exploration process is organized in such a way that only the maximum easiness singular accumulations are discovered (selected into a sampling), we will obtain the Gomperz distribution (distribution of maximum values).

Therefore, a mechanism for providing the reserve accumulation according to the Gomperz law is such an exploration strategy, which results in discovery of only the easiest accumulations. This is essentially the same as discoveries of the largest fields.

It is interesting here to return to the logistic function. It was shown that the logistic distribution is a distribution of the half-sum of the largest and smallest values for a symmetric exponential type distribution [57]. It would be difficult to organize an exploration process in such a manner

that singular accumulation of maximum and minimum degree of easiness were equally discovered. This makes the applicability of Model 1 questionable for describing of the reserve accumulation process. Such a conclusion, however, needs to be verified, which will be discussed below.

Distribution of maximum values (the Gomperz distribution) is defined by two parameters, m and k. Their meanings were discussed earlier. In terms of the distribution function, $t_{max} = (\ln m)/k$ is the distribution mode. The median μ is equal to:

$$\mu = t_{max} - \frac{\ln(\ln 2)}{k} = t_{max} + 0.36651\frac{1}{k} \tag{2-23}$$

and the variance σ^2 is defined by:

$$\sigma^2 = \pi^2/(6k^2) \tag{2-24}$$

These expressions clarify the meaning of parameter k. As it increases, the standard deviation decreases, the distribution becomes more concentrated, the probability density in the point of mode increases, and the distance between the mode and the median decreases.

Model 3. It was assumed in Model 1 that the relative reserve accrual and the specific rate of the reserve accumulation versus the accumulated reserves were linear. There is equal reason to assume that the logarithm of the relative reserve accrual versus the reserve accumulation is linear. In such a situation, for a discrete case:

$$\ln(\Delta R/R_{i+1}) = a - bR_i \tag{2-25}$$

or

$$\Delta R/R_{i+1} = e^{a-bR_i} = se^{-bR_i} \tag{2-26}$$

where $s = e^a$.

From Equation 2-26 it follows that the relative reserve accrual first rapidly decreases with the growth of R, then the rate of decline slows, and afterwards the relative accrual decline becomes so small as to be considered constant (i.e., the selection conditions function and the limitations lose their significance). This may be associated with a change in the structure of unappraised reserves, that is, the reserves which will be accrued in the future. With the increase in R, this structure, which was drastically non-uniform in the beginning, becomes more and more uniform.

As large field discoveries become exhausted, the conditions for new discoveries (selection conditions) first drastically deteriorate and then, as medium and small fields are discovered, the rate of deterioration drops. Eventually, when only small fields remain to be discovered, the conditions for new discoveries remain at the same low level.

It was not possible to obtain a solution for Equation 2-26 in an explicit form of $R(t)$. However, sequential values of the desired function R_i (for $i = 1, 2, 3 \ldots$) can be found from Equation 2-26 for given values of the parameters a, b, and R_o. This is exactly what we are interested in: actual values of the accumulated reserves are available only for points i, where $i = 1, 2, 3 \ldots$ (as was previously mentioned, this is caused by the very nature of the data recording). For this reason, we wrote Model 3 only for a discrete process: it is impossible to analyze a differential equation similar to Equation 2-26 due to the unavailability of instantaneous values of the specific reserve accumulation rate. The values of R_{i+1} may be found from:

$$R_{i+1} = R_i/(1 - e^{a-bR_i}) \tag{2-27}$$

The function $R(t)$, which represents a solution of Equation 2-26 or 2-27, has never been used for an analysis of the reserve accumulation history. It was first introduced by one of the writers [22]. It represents a three-parameter function (a, b, R_o) and forms an asymmetric S-shaped curve with initial exponential growth (at $a < 0$). As in the above analyzed models, a change in the R_o variable causes the $R(t)$ or $r(t)$ curve to shift to the right or left.

At the moment $i = t_{\max}$ when the maximum accrual $\Delta R_{\max}$ is obtained, the accumulated reserves are:

$$R_{infl} = R(t_{\max}) = 1/b \tag{2-28}$$

and the maximum accrual per se may be determined from the following equation:

$$\frac{\Delta R_{\max}}{e^{b\Delta R_{\max}}} = \frac{\Delta R(t_{\max})}{e^{b\Delta R(t_{\max})}} = \frac{1}{b} e^{a-1} \tag{2-29}$$

Function $R(t)$ given by Equation 2-26 or 2-27 does not have the upper limit (asymptote) A. This differs from the functions in Models 1 and 2. If one takes into account oil and gas accumulations of any size, up to the largest one, then the absence of asymptote is natural. Total volume of such accumulations, including dispersed hydrocarbons, may be indefinitely

large compared with the reserves of commercial interest. Commercial reserves are always limited by economic considerations. That is why ultimate potential is an economic category rather than the upper limit of the growth function. Apparently, there is a small reserve accrual value ΔR_{lim}, which can be justified based on economic criteria. The reserves R_{lim} accumulated by the moment when the reserve accrual rate of ΔR_{lim} is reached, represent the limit of the commercial reserves. Obviously, the value R_{lim} better describes the ultimate potential than A. It may be determined from the following equation:

$$R_{\text{lim}}/e^{bR_{\text{lim}}} = \Delta R_{\text{lim}}/e^{a+b\Delta R_{\text{lim}}} \tag{2-30}$$

The ratio $R_{infl}/R_{\text{lim}} = R(t_{\max})/R_{\text{lim}}$ (the degree of asymmetry) in Model 3 is no longer a constant and is dependent on the a and b parameters.

There may be another approach to Model 3. First, Equations 2-13 and 2-26 are similar in appearance. The difference is that the specific rate (relative reserve accrual) exponentially decreases with time for the former and with the accumulation of the appraised reserves, for the latter. An association with the reserves appears to be more convenient because exploration is constantly controlled and corrected by the results, that is, by the accumulated reserves. This was analyzed in detail in Chapter 1. The exponential nature of the decline in the specific reserve accumulation rate and the relative reserve accrual causes an interconnection between Models 2 and 3.

On the other hand, Model 3 may be considered a development of Model 1. This becomes clear if one compares Equations 2-1 and 2-25. The transition from Model 1 to Model 3 can be based on the following. Parameters a and b in Equation 2-1 are constants representing general conditions of exploration in a given region with given geology and given initial reserve structure. These conditions are, to a great degree, determined by the drilling extent, by the exploration strategy, and by the progress in science and technology. This is why the a and b parameters, most likely, will not be constant but will change with time or with the reserve accumulation (because information, control, and experience accumulated in the process of reserve development are of great significance for the rate of exploration process).

Equation 2-26 shows that the a and b parameters of Equation 2-1 (which we now denote as α and β) behave in the following manner with the growth in R:

$$\alpha = (1 + bR)e^{a-bR}$$

$$\beta = be^{a-bR} \tag{2-31}$$

The solution curve of Equation 2-1 has the asymptote $A = \alpha/\beta$ (see Table 2-1). It illustrates that the asymptote linearly increases with the growth of R:

$$A = 1/b + R \tag{2-32}$$

Therefore, the growth defined by Equation 2-26 may be interpreted as the growth of Equation 2-1 having variable parameters, in particular with the asymptote increasing with the growth of R.

Model 4. This model is a further development of the model first introduced for the formal description of the reserve accumulation process [43]. The process of accumulation is described as a continuous process.

The logistic Equation 2-4 is based on the idea that the specific reserve accumulation rate linearly decreases with the reserve accumulation. However, that behavior may have a non-linear nature as described by the following equation:

$$r/R = k - \varepsilon R^{1/b} \tag{2-33}$$

This is a differential equation of the reserve accumulation evolution. Its general solution is:

$$R(t) = A \Big/ \left[1 + \left(\frac{1}{(R_o/A)^{1/b}} - 1 \right) e^{-\left(\frac{k}{b}t\right)} \right]^b \tag{2-34}$$

where:

$$A = \lim_{t \to \infty} R(t)$$

is the asymptote, R_o is the accumulated reserves at $t = 0$, and k, ε and b are parameters ($\varepsilon = k/A^{1/b}$, see Table 2-1). This curve is defined by the four parameters, and the corresponding differential equation is a triple-parameter one, which makes it different from Models 1, 2, and 3.

Equation 2-34 shows that the relative reserve accrual in this case depends on R as follows:

$$\Delta R/R_{i+1} = 1 - (1 - c + aR_i^{1/b})^b \tag{2-35}$$

where the parameters c and a are related to the parameters of the $R(t)$ curve and of the Equation 2-33 model. At $\Delta t = 1$, this relationship is given

in Table 2-1. It is important to note the difference between the expressions for r/R and $\Delta R/R$ because they are sometimes mixed. A discrete equivalent of the Equation 2-33 model can be written as follows:

$$R_{i+1} = R_i/(1 - c + aR_i^{1/b})^b \tag{2-36}$$

If we now switch from R_i to a new function, $W_i = R_i^{1/b}$, then the relative accrual for this new function has the following simple form:

$$\Delta W/W_{i+1} = c - aW_i \tag{2-37}$$

This expression fully corresponds to Equation 2-1, which provides an opportunity to make one more interpretation of this model.

The $R(t)$ curve is also an asymmetric S-shaped curve. The $r(t)$ curve reaches its maximum at the time $t = t_{max}$, which may be found from the following equation:

$$\left(\frac{1}{(R_o/A)^{1/b}} - 1\right)e^{-(k/b)t_{max}} = \frac{1}{b} \tag{2-38}$$

The maximum is:

$$r_{max} = r(t_{max}) = A\frac{k}{b}\left(\frac{b}{b+1}\right)^{b+1} \tag{2-39}$$

The amount of accumulated reserves at the inflection point is:

$$R_{infl} = R(t_{max}) = A\left(\frac{b}{b+1}\right)^{b} \tag{2-40}$$

The specific reserve accumulation rate at this moment is:

$$\frac{r_{max}}{R_{infl}} = \frac{r(t_{max})}{R(t_{max})} = \frac{k}{(b+1)} \tag{2-41}$$

The ratio of the accumulated reserves in the inflection point to the maximum value (degree of asymmetry) is:

$$\frac{R_{infl}}{A} = \frac{R(t_{max})}{A} = \left(\frac{b}{b+1}\right)^{b} \tag{2-42}$$

Therefore, the asymmetry is not a constant but depends on the parameter b of the $R(t)$ curve.

Equation 2-33 is similar to a logistic equation with a variable parameter β, which depends on R:

$$r/R = \alpha - \beta R \tag{2-43}$$

therefore, with a variable asymptote:

$$B = \frac{\alpha}{\beta} = \frac{k}{\varepsilon} R^{(1-1/b)} \tag{2-44}$$

where $\alpha = k$, $\beta = \varepsilon R^{1/b-1}$.

As we can see, the asymptote behavior in Model 3 with increasing R was described by a straight line whereas here it is an exponential function.

Equation 2-33 may be inferred through another line of reasoning [43]. It is known that commercial reserves are appraised based on a preliminary evaluation of reserves (C_2 category)[1]. These, in turn, require a certain volume of initial exploration. A field appraisal crowns the evaluation of the commercial reserves and is preceded by exploratory drilling over the structure. This is preceded by the transfer of the structure for exploratory drilling, which is preceded by the delineation and mapping of the structure, which, in turn, is preceded by the discovery of the structure. The discovery of structures is based on regional studies, and so forth. In other words, the process of discovery and appraisal of commercial reserves consists of a number of steps forming a sequential chain. At each stage (step), the progress in exploration may be delayed by exploration at the preceding stage. There may be a large number of steps because each of the techniques used may be subdivided into a series of sequential activities, and the failure of any of them will delay performing the following one.

Let us assume that the number of such stages (steps) is b. We will designate these steps $X_1(t)$, $X_2(t)$, . . . $X_b(t)$, with $X_1(t) = R(t)$. Each of these steps may be considered as the reserve evaluation process of different degrees of reliability (after all, different reserve categories are being evaluated as a result of different exploration stages). The $X_b(t)$ process in a given region will be completed first and will provide the least reliable reserves with the maximum value A. The $X_{b-1}(t)$ process in the region will be completed next and will provide a more reliable reserve evaluation because the $X_{b-1}(t)$ process is directly associated with the transfer of less reliable reserves into more reliable reserves. Their maximum value will remain the same, A. For instance, the discovery of traps

represents the $X_i(t)$ process. The process is complete when all traps are discovered and the process reaches its maximum value of A (because the reserves with an A value may be contained only in the traps). Delineation and mapping of the traps for exploratory drilling represent the $X_j(t)$ process. When it has covered all (prospective) traps, it will reach its maximum value of A. Exploratory drilling for these traps represents the $X_k(t)$ process. If exploratory drilling has been conducted for all targets, which can include all reserves (of the value A), then the $X_k(t)$ process is completed and has reached its maximum value of A. The final process, $X_1(t)$ (or, using our old designations, the studied $R(t)$ process), will provide the most reliable reserves belonging to the commercial categories with the same maximum value of A.

If each consecutive process had begun only after the preceding one was completed, then there would be no restraining effect of the preceding process on the following one. Let us assume that in such a case the evolution of each process is described by the logistic Equation 2-4. We will rewrite this equation in a slightly different form:

$$\frac{x}{X} = k\left(1 - \frac{X}{k/\varepsilon}\right) = k\left(1 - \frac{X}{A}\right) \tag{2-45}$$

In reality, however, all these processes are occurring at the same time. If at the moment t the activities associated with the exploration step X_2 are stopped and never resumed, then its value of $X_2(t)$ will be the upper limit for the X_1 process (cessation of exploratory stage will limit the appraisal stage and, therefore, the reserve accrual). Likewise, if the activities associated with the X_3 process are stopped and never resumed, then its value of $X_3(t)$ will be the upper limit for the X_2 process, and so forth. Thus, when all processes occur simultaneously, the following system of equations is valid:

$$\begin{aligned} x_1/X_1 &= k_1(1 - X_1/X_2); \\ x_1/X_1 &= k_1(1 - X_1/X_2); \\ &\cdots \\ x_b/X_b &= k_b(1 - X_b/A); \end{aligned} \tag{2-46}$$

The X_b process does not have a preceding process and, therefore, is not restrained by it—its upper limit is A. The latter equation is a regular logistic equation with a constant asymptote. All the previous equations represent equations of logistic growth with a variable asymptote.

In order for the X_i process to occur, each X_{i+1} process must develop prior to the X_i process. The constraints imposed by this condition are:

$$\frac{x_1}{k_1X_1} \leq \frac{x_1}{k_2X_2} \leq \frac{x_3}{k_3X_3} \leq \ldots \leq \frac{x_b}{k_bX_b} \tag{2-47}$$

Then, from Equation 2-46:

$$\frac{X_1}{X_2} \geq \frac{X_2}{X_3} \geq \frac{X_3}{X_4} \geq \ldots \geq \frac{X_b}{A} \tag{2-48}$$

In particular, equality:

$$\frac{X_1}{X_2} = \frac{X_2}{X_3} = \frac{X_3}{X_4} = \ldots = \frac{X_b}{A} \tag{2-49}$$

leads to the optimum result in terms of the maximum reserve accumulation $X_1(T)$ by a given time T [43].

It is obvious that:

$$\frac{X_1}{X_2}\frac{X_2}{X_3}\frac{X_3}{X_4} \ldots \frac{X_b}{A} = \frac{X_1}{A} \tag{2-50}$$

Considering Equation 2-49, it follows that:

$$\frac{X_1}{X_2} = \left(\frac{X_1}{A}\right)^{1/b} \tag{2-51}$$

Taking this into account, the first expression in the Equation 2-46 system assumes the following form:

$$\frac{x_1}{X_1} = k_1\left[1 - \left(\frac{X_1}{A}\right)^{1/b}\right] = k_1 - \varepsilon X_1^{1/b} \tag{2-52}$$

where $\varepsilon = k_1/A^{1/b}$.

Thus, Equation 2-33 may be interpreted as a model of the reserve accumulation history under the condition of optimum evolution of all sequential exploration stages (i.e., the accumulation of sequentially increasing reliability of information). This optimum means reaching the maximum at the specified time. This is true under the condition that free and independent growth at each exploration stage is described by a logistic equation. In this case, b represents the number of such stages (or links in the chain), leading to the final result, which is the reserve accrual.

According to the initial assumption, if the reserve accumulation process were to occur under the condition that all preceding stages ($X_i(t)$, where $i = 1, 2, 3, \ldots b$) were completed, its model would be a logistic curve. Otherwise, Equation 2-33 would be its model. In reality, all preceding stages delay the reserve accumulation and it is normal to expect that the limiting factor (the deterioration of the selection conditions) of the Equation 2-33 model is greater than that of model 2-40. This is exactly what occurs in real life, which can be easily explained by comparing the models.

Let us write the models in the following form:

$$\begin{aligned} \frac{r}{R} &= k(1 - R/A) \\ r/R &= k[1 - (R/A)^{1/b}] \end{aligned} \tag{2-53}$$

This illustrates that the limiting factors for the models are respectively equal to:

$$kR/A \quad \text{and} \quad k(R/A)^{1/b} \tag{2-54}$$

If $R/A < 1$ and $b > 1$, one obtains:

$$R/A < (R/A)^{1/b} \tag{2-55}$$

Thus, dependence of the reserve accumulation on the preceding exploration stages results in a limitation of the accrual. This limitation increases with b, which is the number of steps or stages in the continuous chain of exploration activities. At the same time, however, the limitation itself is less and less determined by the value of R ($R^{1/b} < R$ when $R > 1$). Other factors begin to play a role and are associated with the need to maintain an advanced state of each stage versus the following stage.

Under the condition of certain limitations imposed on the parameters, the above reasoning explains model 2-33. In particular, b must be an integer greater than 1. If $b < 1$, this reasoning does not apply. Therefore, model 2-33 is essentially broader than these concepts, which explains the reserve accumulation history by the coordination of different exploration stages. The value of the b parameter is particularly important because it accounts for the nature of the specific rate change with the growth in R. When $b = 1$, the curve is a straight line and the decline rate is constant. When $b > 1$, the curve is concave and the decline is decreasing. When $b < 1$, the curve is convex and the decline is accelerating. These cases were shown in Figure 2-1. As a consequence, differential Equation 2-33

has several solutions. When $b = 1$, the solution is a logistic function and the $r(t)$ curve is symmetric. When $b > 1$, the $r(t)$ curve has a positive asymmetry (the tail is a longer part of the curve); when $b < 1$, the curve has a negative asymmetry (the tail is a shorter part of the curve).

Model 5. This model is constructed for the discrete process only in Equation 2-1, which is a discrete equivalent of the Verhuhlst-Pearl's logistic equation. It is assumed that the relative reserve accrual is a linear function of the accumulated reserves. However, a non-linear type of this dependence may also be considered:

$$\Delta R/R_{i+1} = c - aR_i^b \tag{2-56}$$

The same type of dependence can be explained based on the reasoning that was applied to the construction of Model 4. The difference is that now we are talking about the relative reserve accrual $\Delta X/X$ rather than the specific rate x/X. Consequently, the speed of the processes (stages) and their advancement versus the following stages should be associated with the relative accruals, exactly as it occurs in real life. Such an approach to interrelations between the $X_i(t)$ processes (their number is equal to $1/b$) better reflects the substance of the study: only annual reserve accruals, not their instanteneous values, can be measured. Respectively, coordination of the processes $X_i(t)$ refers to their coordination at discrete moments in time. All these make the discrete model 2-56 preferable compared to the continuous model 2-33, and the system 2-46 is applied to the discrete case.

When dealing with a discrete system, it is necessary to remember that it only applies to certain parameter values, as was mentioned in the discussion of Model 4.

It was not possible to find an explicit solution for Equation 2-56. However, as in the case of Model 3, it is possible to arrive at sequential R_i values for the assigned parameters a, b, c, and R_o:

$$R_{i+1} = \frac{R_i}{1 - c + aR_i^b} \tag{2-57}$$

where $i = 0, 1, 2, 3 \ldots$

Function $R(t)$ also represents an asymmetric S-shaped curve. As for the previously discussed models, changes in the R_o parameter cause the $R(t)$ or $r(t)$ curve to shift to the right or left.

By the time $i = t_{max}$ when the maximum reserve accrual $\Delta R_{max} = \Delta R(t_{max})$ is reached, the accumulated reserves may be determined from the following

approximate equation (it is true for the small b values, which is what normally occurs):

$$R_{infl} = R(t_{\max}) = \left(\frac{c}{a(b+1)}\right)^{1/b} \tag{2-58}$$

The approximate value of ΔR is determined by the following:

$$\Delta R_{\max} = \Delta R(t_{\max}) = \frac{cb}{1+(1-c)b}\left(\frac{c}{a(b+1)}\right)^{1/b} \tag{2-59}$$

The asymptote of the $R(t)$ function is:

$$A = (c/a)^{1/b} \tag{2-60}$$

Accordingly, the R_{infl}/A ratio (the degree of asymmetry) is:

$$\frac{R_{infl}}{A} = \frac{R(t_{\max})}{A} = \left(\frac{1}{b+1}\right)^{1/b} \tag{2-61}$$

Equation 2-56 describes a system of functions rather than a single function. The differences in functions are associated with the values of c and b parameters. When $b = 1$ (and $c < 1$), Equation 2-56 converts to Equation 2-1. Its solution is the logistic function. If $b < 1$, the relative reserve accrual declines with R in a decreasing fashion; respectively, the ΔR curve is positively asymmetric. If $b > 1$, the decline of the relative reserve accrual accelerates with R and its trajectory is no longer a concave (as in the case of $b < 1$) but, rather, a convex curve; respectively, the ΔR curve is negatively asymmetric.

When $c = 1$ (and $b < 1$), Equation 2-56 converts to Equation 2-11. Its solution is represented by the Gomperz function, which has a completely different meaning. Respectively, different functions are the solutions of Equation 2-56 when of $c < 1$ and $c > 1$.

Above, we analyzed the five different models and their forms are presented in Table 2-1. There is a transition from one model to the next. Models 4 and 5 are the most general in the considered set of models. Their interconnection can be observed by comparing the following rates:

$$r/R = k - \varepsilon R^{1/b} \tag{2-33}$$

$$\Delta R/R_{i+1} = c - aR_i^b \tag{2-56}$$

Model 4 uses the instantaneous rate r whereas Model 5 uses annually averaged ΔR. When $b = 1$, Models 4 and 5 (with $c < 1$) convert to logistic Model 1 (Equations 2-4 and 2-1). On the other hand, if $c = 1$, Model 5 converts to Model 2 (Equation 2-11). In turn, Model 2 is related to Model 3, which may be seen from the following comparison:

$$r/R = \varepsilon e^{-kt} \tag{2-13}$$

$$\Delta R/R_{i+1} = e^{a-bR_i} = se^{-bR_i} \tag{2-26}$$

After a certain transformation, Model 3 may be compared with Model 1:

$$\Delta R/R_{i+1} = a - bR_i \tag{2-1}$$

$$\ln(\Delta R/R_{i+1}) = a - bR_i \tag{2-25}$$

For the discrete case, Model 4, after transformation, is reduced to Model 1:

$$\Delta R/R_{i+1} = a - bR_i \tag{2-1}$$

$$\Delta W/W_{i+1} = c - aW_i \tag{2-37}$$

Two out of the above five models (the logistic function and the Gomperz function) are commonly used in science and technology for describing growth in knowledge or information [57]. All of the models may be considered for this case as describing growth in information rather than growth in reserves. It is well known that with the increase in the appraised reserves the degree of knowledge concerning a region grows, that is, the amount of geologic information grows. In this sense, the growth in reserves is equivalent to the growth in information. Thus, r/R may be interpreted as the accrual of information per unit of time and per unit of already accumulated information—it is the accrual of information provided by the previously accumulated information. It is clear that growth in information depends on the previously accumulated information; which is why models are based on the association between r/R and R.

Study of Intensity of Undiscovered Resource Transfer into the Appraised Reserve Category

Construction of Models

The oil and gas reserve accumulation process may be analyzed from a different viewpoint, which compels us to increase the number of models

under review. The diversity of the models is associated with diverse exploration environments and the diversity of the decisions made in order to control exploration. This is the reason for the introduction of the second group of models.

As indicated earlier, the process of exploration may be considered a process of transferring part of the undiscovered resources into the appraised category.

The entire exploration play situation, including the search for the highest potential zones and large fields (i.e., exploration strategy), is determined by the undiscovered resources or future potential. If undiscovered resources are significant (concentrated in large fields), and high potential zones or unexplored intervals are available, the respective reserve accruals may be large. If large discoveries are made early in the process of exploration and if, additionally, the bulk of the ultimate potential is concentrated in one or two giant fields, then the undiscovered resources will substantially decline after discovery and their structure may also drastically change. These circumstances may substantially influence all subsequent reserve accumulation history. In other words, undiscovered resources shape the concept of the region's potential. They determine the control actions, affect expenditure of funds, and affect the extent of exploratory drilling.

The reserve accrual is that part of the undiscovered resources transferred to the appraised reserves. Therefore, the following is true:

$$\Delta R = \mu U_i \tag{2-62}$$

where $U_i = A - R_i$ is undiscovered resources, μ is some function, and i is the number of time intervals in the exploration process.

It makes sense to call μ an intensity function because it determines the transfer intensity of the undiscovered resources in the appraised reserves. The intensity function is related to the control function and the selection conditions function so that the controlled process of the reserve accumulation may be described using this function.

If the intensity function is equal to a constant, $\mu = a$ ($a < 1$), then the following function is a solution of Equation 2-62:

$$R(t) = A - ce^{-kt} \tag{2-63}$$

where $k = -\ln(1 - a)$ (at $\Delta t = 1$). Function 2-63 is sometimes used as an approximation function for describing the reserve accumulation history and is called the "modified exponential (base e) function."

If the intensity function varies directly with the value of the accumulated reserves (i.e., $\mu = kR_{i+1}$), then Equation 2-62 may be written as Equation 2-1:

$$\Delta R/R_{i+1} = a - bR_i \tag{2-64}$$

where $a = kA$ and $b = k$.

The solution of Equation 2-64 is a logistic function, which provides another interpretation of the initial appraised reserve logistic increase.

Generally speaking, all aforementioned models may be interpreted based on Equation 2-62, which we will show later. At this time, it is important to emphasize that the results may be described in terms of their effect on the exploration process of undiscovered resources.

When constructing new models, we imposed on them identical requirements based on the finite nature of the resources and on the typical shape of the reserve accrual curves (the presence of a maximum). These requirements can be reduced to the following: an analytical function (a solution of the equation describing a model) must have an asymptote and an inflection point. The selection conditions function, as before, must be a decreasing function. Most often, only a share of the ultimate potential, $(A - R_i)/A$, is considered and not the undiscovered potential resources. This parameter, which is similar to the degree of the resources maturity, resource knowledge, and resource exploration, may be called a degree of resources immaturity.

Model 6. The substance of Model 6 is defined by the decline of the relative reserve accrual with an increase in exploration maturity by a logarithmic rule:

$$\frac{\Delta R}{R_{i+1}} = \ln\left(\frac{A}{R_i}\right)^b = b \ln \frac{A}{R_i} \tag{2-65}$$

where

$$\Delta R = R_{i+1} - R_i$$

Hence,

$$R_{i+1} = \frac{R_i}{1 - b \ln(A/R_i)} \tag{2-66}$$

This model is bi-parametric (if one discounts the R_o parameter). In this respect, it is similar to Models 1 through 3. Of particular interest is the similarity between Equation 2-65 and the expression for the specific rate in Model 2. As one may remember, the following was implemented (see Table 2-1):

$$r/R = k\ln(A/R) \tag{2-67}$$

This defines a link between Models 2 and 6. By the time $i = t_{max}$, when the reserve accrual reaches its maximum value $\Delta R_{max} = \Delta R(t_{max})$, the accumulated reserves are:

$$R_{infl} = R(t_{max}) = A/e \tag{2-68}$$

and the maximum reserve accrual is:

$$\Delta R_{max} = \Delta R(t_{max}) = (b/e)A \tag{2-69}$$

At the time when the accrual maximum ΔR_{max} is reached, the share of the accumulated reserves (the curve's asymmetry) is:

$$\frac{R_{infl}}{A} = \frac{R(t_{max})}{A} = \frac{1}{e} = 0.37 \tag{2-70}$$

This is a constant value not dependent on any parameter.

It is interesting to note that relations similar to those above but for a continuous case, occur in the Gomperz function (Model 2, Equations 2-16 through 2-19).

Model 7. In this model, the relative reserve accrual decreases according to the exponential law, with the independent variable being the degree of resources immaturity:

$$\frac{\Delta R}{R_{i+1}} = a\left(\frac{A - R_i}{A}\right)^b \tag{2-71}$$

Therefore,

$$R_{i+1} = \frac{R_i}{1 - a[(A - R_i)/A]^b} \tag{2-72}$$

After some transformations,

$$R_{i+1} = \frac{R_i}{1 - c(A - R_i)^b} \tag{2-73}$$

where $c = a/A^b$.

Similarities with and distinctions from Model 5 are apparent. In both cases, the decline of the relative reserve accrual is defined by an exponential function. However, in Model 5 that function depends on the appraised reserves, whereas it is undiscovered resources in this case. Model 7, like Model 5, is a tri-parameter model (discounting R_o). Here, as in Model 5, b determines the different natures of the change in the relative reserve accrual. At $b = 1$, Model 7 becomes Model 1 (where the rate of relative accrual decline is constant). At $b > 1$ and $b < 1$, the decline occurs with a decrease and an acceleration, respectively.

By the time $i = t_{max}$, when the maximum reserve accrual $\Delta R_{max} = \Delta R(t_{max})$ is reached, the accumulated reserves are:

$$R_{infl} = R(t_{max}) = \frac{A}{1+b} \tag{2-74}$$

The maximum accrual is determined from:

$$\Delta R_{max} = \Delta R(t_{max}) = \frac{a}{1+b} A\left(\frac{b}{1+b}\right)^b \tag{2-75}$$

and the respective share of the accrued reserves is defined by:

$$\frac{R_{infl}}{A} = \frac{R(t_{max})}{A} = \frac{1}{1+b} \tag{2-76}$$

The asymmetry of this curve is not a constant but depends on parameter b.

Model 8. The independent variable in this case is also the degree of the resource immaturity, that is, variable $(A - R_i)/A$. This time, an exponential (base e) function of the relative accrual of this variable is introduced. In order to satisfy the aforementioned requirement of the model, a constant c is also introduced. As a result, the model has the following form:

$$\frac{\Delta R}{R_{i+1}} = ce^{p(A-R_i)/A} - c \tag{2-77}$$

and

$$R_{i+1} = \frac{R_i}{1 - c[e^{p(A-R_i)/A} - 1]} \tag{2-78}$$

By the time $i = t_{max}$ when the maximum reserve accrual $\Delta R_{max} = \Delta R(t_{max})$ is reached, the accumulated reserves can be found from the following equation:

$$e^{p(A-R_i)/A} = \frac{A}{A - pR(t_{max})} \tag{2-79}$$

It is clear from this equation that the maximum accrual as well as the curve asymmetry are variables depending on the model's parameters.

Equation 2-77 may be re-written as follows:

$$\frac{\Delta R}{R_{i+1}} = e^{a-bR_i} - c \tag{2-80}$$

where $a = p + \ln c$; $b = p/A$.

Comparing Equation 2-80 with Equation 2-26, it is easy to identify the difference between Models 8 and 3. It can be said that Model 8 is Model 3 with an asymptote. On the other hand, in Model 8 the sum of the relative reserve accrual and the constant c exponentially (base e) decreases.

It appears that these newly added models were derived based on purely formal reasoning: all available sets of simple monotonously decreasing functions were used to describe the evolution of the relative reserve accrual R/R_{i+1}. This, however, is just an illusion. First, the concept of monotonous decline reflects an increase in the exploration maturity of an area, a progressive decline in the undiscovered potential (undiscovered resources), and an increase in the appraised reserves. Second, the transfer of the undiscovered potential into the appraised reserves may be performed at a different rate, which is reflected in the models. In the case of a successful exploration strategy and the right set of applied exploration techniques, the rate may be especially high at the beginning but will rapidly decline later (as in the case of a rapid discovery of large fields). Afterwards, the rate may be maintained at about the same level for a long period of time because the remaining undiscovered potential is associated with small fields.

In addition, other situations may arise. In the beginning, exploration is not conducted in the most promising areas. Large fields are discovered only after a long time delay; and, due to an unfavorable selection of exploration plays, there is a significant time lapse between discoveries. This case has a totally different transfer rate of undiscovered potential into the appraised reserves and, correspondingly, a different decline rate of the relative reserve accrual. As demonstrated, modeling is a means to describe these different situations and is a formal description that can be conveniently interpreted in geologic terms.

We have already discussed the substance of the models. Our model interpretations were based on the amount of accumulated reserves, however, they may also be associated with the amount of undiscovered resources. Table 2-2 displays all models in both forms and the transition from one form to another. The new presentation form also affords respective interpretations and illustrates a different system showing the interconnections between the models on a new basis, through the undiscovered potential. Table 2-2 shows that all model solutions may be considered functions of the undiscovered potential. Thereby, the reserve accruals are defined by the amount of undiscovered potential.

It appears that this explanation is at least as acceptable as the previous one.

Study of the Reserve Discovery and Appraisal Strategy from a Viewpoint of Field Shortage Quota

Construction of Models

Until now, we analyzed the models where the relative reserve accrual R/R_{i+1} (or the selection conditions function) is a monotonously decreasing function of R. It is associated with a decrease in the number of fields that can be selected for further development, after the first discovery. A deterioration in the selection conditions should be particularly significant after a discovery of large fields.

On the other hand, the ultimate potential is non-uniformly distributed over the fields. The frequency of occurrence of different reserve size fields varies. The number of fields with certain reserve sizes is a characteristic of the ultimate potential structure. Each new discovery, depending on its size, affects the structure of the undiscovered potential. Thus, the petroleum exploration process may be considered a process of destroying the initial potential structure. The type of destruction depends on the exploration strategy. If exploration is conducted from the beginning for the best plays with the largest fields, the structure of undiscovered resources is subject to rapid changes. If, on the other hand, exploration involves only the areas where large fields are absent, then a substantial change in the resource structure will be delayed for a long time. It may be expected that in such a case the selection conditions function may even grow at the beginning (prior to the beginning of exploration in the areas with large fields). It will peak at some relatively small R values (because small discoveries do not cause substantial growth in R) and then will begin to decline.

Table 2-2
Reserve Accumulation History:
Different Forms of Deterministic Models

Model Number	Association with R_i	Association with $\cup_i = A - R_i$
1	$\frac{\Delta R}{R_{i+1}} = a - bR_i$	$\frac{\Delta R}{R_{i+1}} = a\left(\frac{A - R_i}{A}\right)$
2	$\frac{\Delta R}{R_{i+1}} = 1 - aR_i^b$	$\frac{\Delta R}{R_{i+1}} = \frac{A^b - R_i^b}{A^b}$
3	$\frac{\Delta R}{R_{i+1}} = e^{a-bR_i}$	—
4	$\frac{R_{i+1}^b - R_i^b}{R_{i+1}^b} = c - aR_i^b$	$\frac{R_{i+1}^b - R_i^b}{R_{i+1}^b} = c\left(\frac{A^b - R_i^b}{A^b}\right)$
5	$\frac{\Delta R}{R_{i+1}} = c - aR_i^b$	$\frac{\Delta R}{R_{i+1}} = c\left(\frac{A^b - R_i^b}{A^b}\right)$
6	$\frac{\Delta R}{R_{i+1}} = a - b \ln R_i$	$\frac{\Delta R}{R_{i+1}} = \ln\left(\frac{A}{R_i}^b\right) = b(\ln A - \ln R_i)$
7	$\frac{\Delta R}{R_{i+1}} = (c - mR_i)^b$	$\frac{\Delta R}{R_{i+1}} = a\left(\frac{A - R_i}{A}\right)^b$
8	$\frac{\Delta R}{R_{i+1}} = e^{a-bR_i} - c$	$\frac{\Delta R}{R_{i+1}} = c\left(e^{p\frac{A-R_i}{A}} - 1\right)$

Model 9. For this model, it is assumed that the relative reserve accrual (or the selection conditions function) first increases and later declines with a growth in R. We will still honor the precondition stipulating that the solution of the new model is an $R(t)$ function with the asymptote A. We will also will add one detail which has not been previously mentioned. In many regions, the very first reserve accrual value ΔR manifests a substantial hike, with subsequent much smaller increases. The first accrual is not based on the general regional reserve accrual trend, but reflects the entire multitude of activities before the first discovery. If we extend the reserve accrual change curve not forward but back in time, zero would not be reached even over a long period of time. The origin of the first

accrual is different in principle; it manifests a new exploration stage in an intermittent manner. This means that the $R(t)$ function at $t \to -\infty$ has as its limit not zero, as it has been in all the above cases, but some value:

$$A_{lower} = \lim_{t \to -\infty} R(t)$$

Based on this, the new model may be presented in the following form:

$$\Delta R / R_{i+1} = e^{a - bR_i} - c/R_i \tag{2-81}$$

To better understand the meaning of Model 9, the analogy of the evolution of biological populations is appropriate. A population can evolve following its own internal patterns, which may be described by the equation of regular reproduction with the competition taken into account. However, if a part of the population is artificially eliminated (for instance, if fish are being caught from a pond) at a certain rate, then a differential equation of the population growth (differential equation of catch) represents the regular reproduction rate minus the rate of catch. A value characterizing the rate of catch is called the catch quota.

If we apply this to the reserve accumulation, we may be discussing its natural evolution (internal laws of an unconstrained evolution) and its artificial limitation ("catch," or shortage of part of the reserve accrual). This limitation is due to the fact that exploration was artificially delayed (conducted for a period of time in a low potential area). Equation 2-81 is a representation of the reserve accumulation rate ΔR as an equivalent of the differential equation of catch (i.e., it has two components). The second may be considered as a quota of catch, that is, the rate of the reserve accrual "shortage" caused by the limitations associated with prior strategic decisions. Model 8 may be interpreted similarly. However, the quota cR_{i+1} in this model is small at the beginning and increases with a growth in R_{i+1} (which appears natural: as the large discoveries become exhausted, it is increasingly difficult to accumulate reserves). At the same time, the quota cR_{i+1}/R_i in Model 9 behaves in the opposite way—it is large at the beginning and later turns into a constant approximately equal to c (which is typical for a totally different exploration strategy).

Function $R(t)$, which is a solution of Equation 2-81, has one inflection point and upper A and lower A_{lower} limits. The function at the right-hand part of Equation 2-81 is a function of R and has an inflection point and maximum. It crosses the R axis at the points A_{lower} and A, respectively.

From Equation 2-81 it follows that:

$$R_{i+1} = \frac{R_i}{1 - (e^{a - bR_i} - c/R_i)} \tag{2-82}$$

It is not possible to solve Equation 2-82 explicitly with respect to $R(t)$. As noted earlier, however, it is possible to obtain consecutive R_i values in discrete points i, $i = 1, 2, 3 \ldots$ As with the functions in Models 4, 5, 7, and 8, this function is a quadri-parameter. As in the above models, a change in the R_o parameter leads to a shift of the curve $R(t)$ or $r(t)$ to the right or left.

Values of the asymptotes A_{lower} and A satisfy the following:

$$\begin{aligned} e^{a-bA} &= c/A \\ e^{a-bA_{lower}} &= c/A_{lower} \end{aligned} \tag{2-83}$$

By the time $i = t_{\max}$, when the maximum reserve accrual $\Delta R_{\max} = \Delta R(t_{\max})$ is reached, the accumulated reserves are:

$$R_{infl} = R(t_{\max}) \cong 1/b \tag{2-84}$$

and the maximum accrual itself is:

$$\Delta R_{\max} = \Delta R(t_{\max}) \cong \frac{e^{a-1} - cb}{b[1 - (e^{a-1} - cb)]} \tag{2-85}$$

The ratio of R_{infl}/A (degree of asymmetry) is also a variable depending on the model's parameters a, b, and c.

The R_n value at which the conditions selection function $f(R) = e^{a-bR} - c/R$ reaches its maximum $f(R_n)$ is determined from the equation:

$$c/R_H^2 - be^{a-b/R_H} = 0 \tag{2-86}$$

The maximum itself is determined as follows:

$$f(R_H) = e^{a-bR_H} - c/R_H \tag{2-87}$$

Evaluation and Diagnostic Review of Models

Theoretical concepts included in a mathematical model can only be accepted by testing how well the model fits the observations.

Material for the Study

Data on two well-studied regions with long exploration histories was used to check how the proposed models fit the real reserve accumulation process. Two regions were selected in order to compare the modeling

results. Based on this comparative analysis, we can judge the effect of exploration strategy on the reserve accrual pattern. The test regions are opposite in their exploration strategies.

Region A. From the beginning, exploration was concentrated in the most promising area with large fields, that is, the strategy choice was correct. The strategy was in progressive consecutive studies of the shallow, and later deeper, horizons in the most promising area with simultaneous expansion of studies in other, less attractive areas.

As a result, the first discoveries produced relatively high reserve accruals. The most significant fields in the area were discovered relatively fast and almost simultaneously. In other words, large discoveries were exhausted rapidly and, therefore, reserve accruals quickly reached a maximum. Thereafter, it was progressively lower because the large fields were already discovered in the most promising area and were not present in the other areas. Lower and lower numbers of less substantial discoveries were made as exploration continued. Currently, mostly small fields are being discovered, and exploration maturity is high.

Region B. Exploration effort at the first stage was very slow and was concentrated in the least promising area. It took many years to study this area and to turn to the adjacent areas. The reserve accrual during this first stage was very low because exploration efforts were concentrated on the small fields. The reserve accrual was further lowered by the fact that poorly prepared (i.e., delineated and mapped) targets were drilled.

The discovery of a significant field changed this situation. Major exploration effort shifted to different areas, after which a number of large fields were discovered almost simultaneously. Later, more such discoveries were made. The discovery and appraisal of high potential areas, where 70% of the total exploratory drilling was concentrated, resulted in maximum reserve accrual. Albeit with ups and downs, high accruals in the area were observed for a substantial period of time. A progressive expansion into different areas became constrained due to concentration of exploratory drilling for the appraisal of already discovered fields. This, in turn, resulted in the decline of the exploration target number. Afterwards, the situation began to change. Total drilling footage (including appraisal drilling) declined, but the share of exploratory drilling in the total drilling footage increased. However, fields discovered during this period were mostly small. The reserve accrual substantially decreased. The current exploration maturity of the region is not high and is growing very slowly. A substantial potential is associated with poorly studied areas.

It is expected that different types of development in these two regions and differences in the system and methodology of exploration should be

reflected in the reserve accrual patterns. This should lead to different formal descriptions or different models.

Testing the Models Based on Observations

The best selection of a model should be based on its ties with field data. A two-step procedure of evaluation and diagnostic checks was used to establish these ties. The evaluation implies an efficient evaluation of the parameters of constructed models.

All the models considered were non-linear in terms of their parameters. In the process of non-linear evaluation, the models' parameters could not be analytically determined from observations. Different techniques were used that may be called algorithmic rather than analytical.

In this case, as well as later in this book, while analyzing the other models, parameters were determined by the iterative search of the values, which ensures the minimum of summed squared deviations of the observations from the corresponding model's outputs.

First, we will explain the search procedure using an example of a single parameter with the notation c. The sum of squared deviations $SS(c_1)$ was determined for the arbitrary initial value of parameter $c = c_1$. Then the parameter value was increased by a certain amount Δc (i.e., $c_2 = c_1 + \Delta c$). If the sum of squared deviations $SS(c_2)$, corresponding to c_2, was smaller than $SS(c_1)$, then c_3 was determined by further increasing c_2 (i.e., $c_3 = c_2 + \Delta c$). Otherwise $c_3 = c_1 - \Delta c$. This explains how the direction was determined for the consecutive sorting of c_i values toward the increase or decrease of c. Moving in the chosen direction with a certain step Δc, after a number of steps a value $\hat{c} = c_n$ is found, which corresponds to the minimum of summed squared deviations observed from model $SS(c_n) = \min SS(c_i)$. The value $\hat{c} = c_n$ is an evaluation of parameter c.

This is a practical implementation of the least squares method. This procedure was repeated several times for different initial c_1 values, which was necessary in order to find a global rather than local minimum $SS(c_i)$. While evaluating parameters, different Δc steps were also tested. In the beginning the step was relatively large but was subsequently decreased. This provided for a rough evaluation of the c area where $SS(c_i)$ minimum may be expected and then, within a narrowed-down interval, for a more precise estimation of parameter c.

The following should be kept in mind when evaluating b and c parameters. The set of possible parameter values forms a plane in the parameter space. A given value of, let us say, b provides a cross-section of the plane. Let us select b_1 as the initial b value. Performing the above described procedure for this cross-section b, we will find the best evaluation $\hat{c}(b_1)$ in this cross-section and the corresponding squared deviation sum $SS(b_1)$.

Then we will change the b_1 value by a certain Δb (i.e., $b_2 = b_1 + \Delta b$) and find the best estimation $\hat{c}(b_2)$ in this new cross-section and, respectively, $SS(b_2)$. If $SS(b_2) < SS(b_1)$, the next cross-section will be selected as $b_3 = b_2 + \Delta b$. In this new cross-section we will find the best evaluation $\hat{c}(b_3)$. Otherwise, the next cross-section will be selected $b_3 = b_1 - \Delta b$ and $\hat{c}(b_3)$ and $SS(b_3)$ found in this cross-section. Having determined the direction of the cross-section b_i selection and moving in this direction, we will eventually find a cross-section $\hat{b} = b_n$ where the minimum of summed squared deviations $SS(b_n)$ corresponds to the best $\hat{c}(b_n)$ estimate. These $\hat{b}$ and $\hat{c}(b_n)$ values are the best (in the least square sense) estimates of the b and c parameters.

This procedure is applied several times for different initial values of parameters and for different steps Δb and Δc.

The following helps reduce computing time. As the initial value of $c = c(b_i)$ in each new cross-section b_i, a value corresponding to the best $\hat{c}(b_{i-1})$ value of the previous cross-section was selected. This helps substantially reduce screening different c values for each b cross-section of the parameter plane.

The substance of the procedure remains the same when evaluating three parameters. The difference is, as in the previous case, that the best estimates of the $\hat{b}$ and $\hat{c}$ parameters are searched, this time on a parameter plane corresponding to a given value of the third parameter (let us call it a). Subsequently, the direction of selecting new cross-sections is determined. This time, the cross-sections are represented by a plane. Moving in this direction and determining for each step $a_i = a_{i-1} \pm \Delta a$ the best estimates $\hat{b}(a_i)$ and $\hat{c}(a_i)$, it is possible to find such a plane (cross-section) $\hat{a} = a_n$, where the minimum of summed squared deviations for the model $SS(a_n)$ among all screened planes corresponds to the best $\hat{b}(a_n)$ and $\hat{c}(a_n)$ estimates. These values a_n, $b(a_n)$, and $c(a_n)$ are the best evaluation of a, b, and c parameters. Here, as well, time savings are achieved by taking as initial approximations at each new step a_i (on each new plane) the best estimates $\hat{b}(a_{i-1})$ and $\hat{c}(a_{i-1})$ from the preceding step (preceding plane).

A similar procedure is employed when evaluating four or more parameters. The difference is that the dimension of the parameter space and the dimension of the respective cross-section, or hyperplane, in this space is greater.

Thus, the evaluations used correspond to the least squares method as our best estimates. They were found by sorting all possible values in the parameter space. The sorting was not conducted along a certain grid but, instead, a more efficient procedure was applied. This procedure provided for a drastic reduction of the possible parameter value area by selecting the right direction of movement in the parameter space and by using the results from preceding steps for subsequent steps. Variance σ^2 (or the

respective summed squared deviations *SS*) was used as the measure of adequacy of models to the available data. Sometimes this measure was improved by adding the correlation coefficient between the observed values and those evaluated from models.[2]

When studying our models, a great deal of attention was paid to the relative reserve accrual. This relative accrual (or its close equivalent, the specific rate) determined the meaning of a model and helped in its interpretation. The behavior of the relative accrual and the possibility of describing it using the proposed models was previously discussed [22]. In this chapter we analyzed the models in the $R_{i+1} = f(R_i)$ representation, which directly reflects the process of reserve accumulation.

Classification of Regions by Specific Exploration Strategies and Reserve Discovery and Appraisal Evolution

We will use the two selected test regions A and B to judge the extent to which the models fit the observations. We will also use these regions to determine the extent to which they reflect differences in the strategy of the oil and gas reserve discovery and appraisal, and to judge the evaluation of potential obtained through their application.

Study Results

Region A. Tentative evaluation of the model parameters and σ^2 are shown in Table 2-3. An approximation of the initial series of the accumulated reserves and, respectively, the annual reserve accrual series using various $R(t)$ functions, leads to the following conclusion: although the accumulated reserves and reserve accrual curves in general follow the observed series, they do not coincide. In order to compare the curves as described by all models, Table 2-4 gives their relative characteristics. The table shows that, with the exception of Model 3, the model curves rise at different rates, reach the inflection point at different moments in time, and in later years fall below the observed values. Model curves 2 and 5 are very close. Reserve accrual maxima ΔR_{max} are reached first on the Model 3 curve and last on the Model 1 curve. Model curves 2, 4, 5, 6, and 8 reach reserve accrual maxima almost simultaneously. The value of the ΔR_{max} differs in various models; the lowest value is demonstrated in Model 6 and the largest in Model 9. Model curves 3 and 9 are most asymmetric. Model curves 6, 7, and 9 show only a small portion of the ultimate potential concentrated around the maxima.

Table 2-3
Evaluation of Adequacy of the Deterministic Models to the Reserve Accumulation History and Field Data (Region A)

Model Number	Model	A_{lower}	a	b	c	A	R_o	SS	σ^2
1	$R_{i+1} = R_i/(1 - a + bR_i)$	—	0.23	0.00021	—	10.95	35	37,588.22	1,212.52
2	$R_{i+1} = aR_i^b$	—	2.86	0.851	—	11.56	3.5	15,840.25	510.9758
3	$R_{i+1} = R_i/(1 - e^{a-bR_i})$	—	–0.73	0.00297	—	23.71	3.1	23,512.10	758.4583
4	$R_{i+1} = R_i/(1 - c + aR_i^{1/b})^b$	—	0.166	900	0.1673	11.19	0.9	23,639.86	762.5761
5	$R_{i+1} = R_i/(1 - c + aR_i^b)$	—	0.7668	0.095	1.5	11.68	1.8	15,541.48	501.338
6	$R_{i+1} = R_i/[1 - b\ln(A/R_i)]$	—	—	0.122	—	12.14	0.9	20,737.72	668.9587
7	$R_{i+1} = R_i/[1 - c(A - R_i)^b]$	—	—	1.826	$6.3*10^{-7}$	12.93	17.7	15,614.32	503.6877
8	$R_{i+1} = R_i/(1 + c - e^{a-bR_i})$	—	–0.87	0.002	0.035	12.41	8.3	13,948.91	449.9648
9	$R_{i+1} = R_i[1 - (e^{a-bR_i} - c/R_i)]$	21.65	–0.54	$2.8*10^{-3}$	13	15.04	26.9	17,211.18	555.2

Table 2-4
Relative Characterization of the Deterministic Curves for the Reserve Accumulation Evolution as Described by Different Models (Region A)

Model Number	Position in a Series of Increasing Maximum Reserve Accrual ΔR_{max} Values	The Year $i = t_{max}$ when Maximum Accrual ΔR_{max} was Reached	Relative Value of the Maximum Accrual $\Delta R_{max}/A$, %	Curve Asymmetry (Ratio of Reserves Accumulated by the Time the Maximum Accrual was Reached, to the Ultimate Potential, R_{inft}/A, %)	Degree of Stretch of the Reserve Accrual Curve	
					Number of Years around ΔR_{max} during which the Sum of ΔR Reaches 1/2 of the Ultimate Potential A	Number of Years around ΔR_{max} during which the Sum of ΔR Reaches 2/3 of the Ultimate Potential A
1	4	14	6.5	50	8	12
2	3	11	5.9	37	8	13
3	7	9	3.1[3]	13.3[2]	—	—
4	8	11	6.8	39	7	11
5	2	11	5.8	37	9	14
6	1	11	5.2	36	10	16
7	5	12	5.6	39	10	15
8	6	11	5.9	35	9	14
9	9	10	5.2	25	12	21

As previously mentioned, model selection is predetermined by the goal. Assuming that the goal is to better understand the nature of the reserve accumulation process and to relate this process to certain significant factors, preference should be given to the model that best agrees with the field. In this respect, Model 8 is the best ($\sigma^2 \approx 450$), followed by the competing Models 5, 7, and 2, respectively ($\sigma^2 \approx 501$, 504, and 511). The latter three models are similar. Their competitiveness with Model 8 is based on the fact that their basic concepts substantially reflect the real flow of the process. Let us analyze the curves to better understand the difference between the models.

First, we see that the model curves 2 and 5 almost coincide. It was mentioned earlier that in some respect Model 2 may be considered a special case of Model 5 (at $c = 1$). On the other hand, Model 5 has some features pertaining to Model 2. It is important to remember that the Gomperz distribution is a distribution of maximum values. The transition from Model 2 to Model 5 causes a change in c values. Growth in c is associated with a decrease in b parameter influence. Taking into account that $\Delta R/R_{i+1} = c - aR_i^b$, one can note that a transformation from Model 2 to Model 5 is caused by a weakening in the effect of the achieved accumulated reserve level R on the relative reserve accruals. The smaller b is, the less R's influence. When $b \to 0$, $R^b \to const.$

In other words, in Model 5 relative reserve accruals are much stronger than in the Gomperz model and are associated with some constant factors. One such factor is a restraining influence on the discovery and appraisal of fields at preceding stages of the exploration steps, as was already mentioned when discussing Model 5. The number of such steps, or links in the chain, of consecutive jobs was associated with $1/b$. In this case the number turned out to be 10 or 11.

Model 7 was introduced in part as a counterbalance to Model 5 where the exponential function with the base R_i is replaced by a function with the base $\cup_i = A - R_i$. Does the solution of Model 7 differ from the solution of Model 5? In other words, can either model be reduced to the other? A comparison of variances (respectively, 503.69 and 501.34) ostensibly indicates that there is no significant difference between the models. In this case, it may be assumed that both Models 7 and 5 do not yield a new function and are close to the Gomperz model. The difference in variances for these models is small (respectively, 503.69 and 510.98). If, however, we look at the reserve accrual curves ΔR as described by Models 5 and 7, we will see that the assumption is erroneous. A similar approximation occurred not due to the similarity of curves. Comparison shows that the curves for Models 7 and 5 are different. There is nothing even remotely similar to the case of Models 5 and 2. The reserve accrual maximum of curve 7 is substantially higher than that of

Model 5 and is offset toward higher t values by one year. Model 7 curve is more compressed but more stretched around the maximum. During recent years it is continuously above curve 5 and their divergence increases with time. Thus, Model 7 may be considered as a model giving rise to a new kind of function. What was previously mentioned about Model 5 (that it provides solutions similar to the Gomperz model) is not applicable to Model 7.

Model 8 yields a totally different class of functions. Its difference from the Gomperz model is associated with a substantial difference in the degree of approximation (respective σ^2 values are 511 and 450). R and ΔR functions as provided by the respective models are quite different. The ΔR curves coincide only at the very beginning. Afterwards, the Model 8 curve runs above the Model 2 curve and the diversion increases with time. Although the maxima are reached simultaneously, the discrepancy between the curves is strongest at this time. After the maxima, the curves become somewhat closer but, then again, curve 8 separates and rises above curve 2. The divergence between the curves progressively increases. During the beginning years of exploration curve 8 is much higher than curve 2.

Another goal achieved by a selected model is forecast. A model should be a tool in determining the future reserve accruals and in the evaluation of the ultimate potential. To evaluate this, let us review the ΔR and R curves versus time. As we can see, curves 1, 2, 4, 5, and 6 during recent years run below the actual values, and the lag increases with time. Table 2-3 provides tentative evaluations of the asymptote A. They are compared with a tentative evaluation of the ultimate potential. The values of A given by Models 2 and 5 are only slightly higher than (tentative) the actual level during the last analyzed year. Consequently, Models 2 and 5 cannot be used for forecast purposes. The same is true for Models 1 and 4, if only because their evaluations of A are below the level actually reached.

There is one more argument for rejecting Model 4. As previously indicated, the b parameter in the model represents the number of steps (stages) included in the reserve discovery and appraisal process. In this case, b turned out to be too large (b = 900). It is hard to imagine, as Model 4 would require, an exploration process composed of that many coordinated steps. The optimization requirement of such a great number of stages would be exceptionally burdensome.

The evaluations of A as given by Models 7 and 8 are trustworthy. The evaluation obtained by using Model 7 is slightly higher than that obtained from Model 8, but they are still close. A tentative evaluation of the ultimate potential from Model 9 is the largest, except for Model 3 where we have R_{lim} instead of A. Based on the current concepts, A

evaluation as given by Model 9 is realistic. If we take into account the evaluation based on direct geologic techniques, the evaluation of A from Model 9 is not worse, and may be even slightly better, than that from Model 8. Therefore, the forecast capabilities of this model are no worse than those of Model 8.

Let us now review the R_{lim} evaluation from Model 3. If we take $\Delta R_{\text{lim}} = 1$, the R_{lim} evaluation from Model 3 is more than twice the accumulated reserves for the most recent analyzed year. Taking into consideration a high degree of exploration maturity in the region (i.e., the reserves are substantially depleted), such an evaluation appears to be too high.

As mentioned in Chapter 1, there is a third goal that can be served by an adequate model. Models may be used to direct the study process along the desired path. In this case, a model is constructed with the purpose of simulating the flow of a real process and of understanding the mechanism driving the process. In other words, we are now in the position of an observer trying to understand the substance of the process. Active interference with the process is a totally different matter and is beyond the scope of these models, which attempt only to describe the situation.

As previously discussed, control generates a new type of selection mechanism for actions, and, in a general case, such a mechanism is called feedback. Strictly speaking, control implies a certain structure of feedback.

We also deal with systems with feedback. However, because these systems are very complex, we are only interested in the consequences of control as stipulated by the feedback (i.e., what may be called a reflection of the behavior function). The interconnections determining the behavior functions are exceptionally complex and indirect (caused by the diversity and hierarchy of goals, by the presence of different levels, etc.). Thus, the only sensible method of study appears to be an experimental one. In combination with the conservation laws, these functions provide for closing the models and form a tool for studying the evolution of very complex processes through their approximate description. Construction of the relations that provide an approximate description of complex processes in the form of finite expressions is called "parametrization." Therefore, in this case the feedback is parametrized by the behavior function. It is not clear at this point whether it is possible to force the process to run along the desired path by active interference. (A desired trajectory is determined by an arbitrarily selected model with parameters, which are arbitrarily changed in time). As previously discussed, the reserve accumulation process is controlled by a number of natural patterns and laws so that the extent of arbitrary changes is very limited. At any rate, to control the process, various optimization problems must be solved (see Part III). The models discussed here do not provide for such solutions because they

are not control models. This is why it is not feasible (as previously attempted), for the purpose of active planning, to model the process of the reserve discovery and appraisal in the form of piecewise developmental curves $R(t)$ with arbitrarily changed parameters or, more specifically, arbitrarily changed asymptote [5]. The asymptote does not change in an orderly manner depending on certain conditions, as in Model 4 or 5, but is assigned arbitrarily: it is assumed to be equal to the estimate of possible prospective resources as known at the time of planning. Therefore, the goal of active planning is not reached.

Let us summarize the analysis of all models for region A. Model 8 appears to fit the goal of describing the reserve accumulation history in the region. It allows for the successful achievement of the two goals of modeling, thereby spanning the gap that is observed in the other models, between the quality of a process description and the forecast of its further evolution. The concepts of this model, as opposed to the others, appear to better imitate the real specifics of the reserve accumulation process. The mechanism of the reserve accumulation history is, in this case, described by Equation 2-80, which shows an exponential decline in the relative reserve accrual with the increase of reserves. The manifestation of this mechanism is associated with a change in the process of the reserve accumulation of the unappraised reserve structure, which is becoming increasingly uniform from an initially non-uniform structure. Thus, as large discoveries are made, the conditions for new discoveries drastically deteriorate. Later, as the medium- and small-size fields are explored, the rate of deterioration declines and, finally, when only small fields remain undiscovered, the conditions for new discoveries remain relatively constant and poor.

The logistic function (Model 1) is most frequently used for modeling the reserve accumulation history. This function is the worst in its agreement with observations and it does not satisfy the goals. Used for a similar purpose, the Gomperz function (Model 2) is also unsatisfactory, although it can to some extent compete with Model 8 (similar to Models 5 and 7). It is expected, however, that with time the reserve accumulation trajectory will significantly deviate from the curve in Model 2. A good approximation is testimony to the respective model agreeing with the observations over the analyzed time interval, but it does not mean that this fit will continue in the future. The trend observed over recent years indicates that most likely it will not.

In theory, a deterioration in the discrepancy between Model 2 and the observations appears to be inevitable if the model is interpreted as a result of the "selection" of fields with the largest reserves. The important point to remember is that the non-uniformity in the reserve size of undiscovered fields is significant in the initial exploration stages. In the

future, these undiscovered fields will become progressively more uniform. As a result, the initial concept of the "selection" of maximum reserve values out of the initial natural distribution is no longer applicable. Therefore, Model 2 (and Model 5) can be used for describing of the reserve accumulation process only during the initial exploration period, when undiscovered field reserve non-uniformity is still close to the undisturbed status.

In the initial exploration stages, therefore, the mechanism of the reserve accumulation is affected by the multi-step nature of exploration, which restrains the influence of the preceding stages over subsequent stages. In other words, this aspect of the mechanism is informational in its substance. Most important, however, for at least the first 20 to 30 years of exploration, is that this mechanism is formed by directing the exploration system toward the discovery and appraisal of the largest fields. This aspect of the mechanism is strategic in its substance.

As a result, the Gomperz curve cannot be used for forecasting. The situation is not improved by the recently proposed model similar to Model 4 [43]. The premises accepted in this model impose $b = 2$. However, in such a case, the model does not fit the observations. It may be applicable only in the modification (Model 4), where the reserve accumulation is limited not only by the low-category reserve accumulations.

In terms of predictive capabilities, Model 9 and, to some extent, Model 3 can compete with Model 8. However, the concept on which the model is based does not fit the observations. In particular, it was not confirmed that the conditions selection function is not a monotonously decreasing function but has maximum (as in Model 9). The value of this maximum is $f(R_n) = 0.3105$, and it was reached at $R_n = 103$, in other words, relatively quickly ($i = 6$). The idea of the initial ΔR jump (therefore, the presence of the lower asymptote A_{lower} in the $R(t)$ function) did not result in a better approximation. Of course, in this case, it may be explained by a rapid increase in reserves in the region due to the initial large discoveries. Consequently, the "jump" fit the general trend in increase of reserves without distorting it. This is also indicated by the substantial difference between A_{lower} and R_o.

Region B. As before, we will use σ^2 to judge how a model fits the observations. Tentative evaluation of the models' parameters and the value of the variance are listed in Table 2-5. The approximation of the input accumulated reserve series and of the reserve accrual series by different $R(t)$, and, correspondingly, by ΔR functions, gives us some idea of what degree the respective curves are different (Table 2-6). Table 2-5 indicates that Model 9 fits the observations the best. Its variance is almost 1.5 times smaller than the variance from Model 5, which is ranked second. This

Table 2-5
Adequacy Evaluation between the Deterministic Models of the Reserve Accumulation History and the Field Data (Region B)

Model Number	Model	A_{lower}	a	b	c	A	R_o	SS	σ^2
1	$R_{i+1} = R_i/(1 - a + bR_i)$	—	0.211	$9.3*10^{-5}$	—	22.69	3.1	58,429.50	1,579.18
2	$R_{i+1} = aR_i^b$	—	2.78	0.87	—	26.05	$3*10^{-14}$	93,114.41	2,516.61
3	$R_{i+1} = R_i/(1 - e^{a-bR_i})$	—	−1.26	$9.9*10^{-4}$	—	77.75	0.54	81,836.85	2,211.81
4	$R_{i+1} = R_i/(1 - c + aR_i^{1/b})^b$	—	not analyzed	not analyzed	not analyzed	not analyzed	not analyzed	not analyzed	not analyzed
5	$R_{i+1} = R_i/(1 - c + aR_i^b)$	—	$4.97*10^{-4}$	0.8	0.244	23.11	1.4	54,829.83	1,481.89
6	$R_{i+1} = R_i/[1 - b\ln(A/R_i)]$	—	—	0.061	—	51	0.001	219,339.60	5,928.10
7	$R_{i+1} = R_i/[1 - c(A - R_i)^b]$	—	—	0.972	$1.19*10^{-4}$	22.4	2.7	56,415.76	1,524.75
8	$R_{i+1} = R_i/(1 + c - e^{a-bR_i})$	—	−1.13	$6.8*10^{-4}$	0.049	27.73	0.58	57,752.35	1,560.87
9	$R_{i+1} = R_i[1 - (e^{a-bR_i} - c/R_i)]$	45.01	−0.91	$1.1*10^{-3}$	18	41.1	45.404	33,778.83	912.94

Table 2-6
Relative Characterization of the Deterministic Curves for the Reserve Accumulation History as Described by Different Models (Region B)

Model Number	Position in a Series of Increasing Maximum Reserve Accrual ΔR_{max} Values	The Year $i = t_{max}$ when Maximum Accrual ΔR_{max} was Reached	Relative Value of the Maximum Accrual $\Delta R_{max}/A$, %	Curve Asymmetry (Ratio of Reserve Accumulated by the Time the Maximum Accrual was Reached, to Ultimate Potential, R_{inf}/A, %)	Degree of Stretch of the Reserve Accrual Curve	
					Number of Years around ΔR_{max} during which the Sum of ΔR Reaches 1/2 of the Ultimate Potential A	Number of Years around ΔR_{max} during which the Sum of ΔR Reaches 2/3 of the Ultimate Potential A
1	5	28	5.9	50	9	13
2	4	27	5.1	40	10	14
3	1	27	1.5[3]	13[4]	—	—
4	not analyzed	not analyzed	not analyzed	not analyzed	not analyzed	not analyzed
5	6–7	28	5.9	50	9	14
6	2	35	2.4	36	—	—
7	8	28	6.2	52	8	12
8	3	27	4.7	38	11	17
9	6–7	26	3.3	23	24	—

makes Model 9 preferable over any other model for describing the mechanism of oil and gas reserve accrual. Observations confirmed the concepts input into the model. In particular, exploration in the region was conducted in such a way that the conditions selection function increased in the beginning, reached maximum, and then began to gradually decline. As previously mentioned, this was caused by specific features in the region's exploration process. The conditions selection function $f(R_H)$ reached a maximum of 0.23 at $R_H = 228$. It took 18 years ($i = 18$) to achieve this amount of initial appraised reserves.

The initial ΔR "jump," therefore, and the presence in the $R(t)$ function of the lower asymptote A_{lower}, has been confirmed. In this case, A_{lower} is very close to R_o. The concept of the lower limit A_{lower} resulted in a better approximation of the accumulated reserve curve R, especially during the first years. When following the $R(t)$ curve backwards, the curve does not tend to the t axis but rather runs parallel to it. The jump in ΔR is apparent because the subsequent reserve accruals ΔR do not rapidly grow but, rather, are close in value. This is also confirmed by the close A_{lower} and R_o values.

As indicated, Model 5 provides the best approximation of the accumulated reserve series after Model 9. Model 8 is of particular interest because it is the best for region A. Reserve accrual curve ΔR of Model 9 differs from these models as described below.

During the initial period, Model 9 runs lower than the curves of Models 5 and 8, reaches maximum prior to these curves, and the maximum is higher. Thereafter, over a long period of time, the curve coincides with the Model 8 curve and is lower than the Model 5 curve. Afterwards, it rises above both curves, and the separation from the Model 8 curve is observed only during the last two or three years. In general, the Model 9 curve is extended upwards more than the Model 5 and 8 curves, and its longer branch approaches the t axis more slowly. The same is clear from the relative characterization of the curves as indicated in Table 2-6. This table shows that the Model 9 curve is also the most asymmetric (except Model 3) and a relatively small portion of the ultimate potential is concentrated around the maximum.

The predictive capabilities of Model 9 may again be judged from its A evaluation. Its tentative value is 41.10 (see Table 2-5). As mentioned, after Model 9, Models 1, 5, 7, and 8 best fit the observations. Among all the models, Model 9 provides the largest A evaluation, which is 1.5 to 1.7 times greater than for the other models. This makes the A evaluation trustworthy although it is lower than the evaluations obtained using a conventional geologic approach.

A tentative conclusion, therefore, is that Model 9 is most preferable from any position. It is the most successful in describing the history of

the reserve accumulation and in its forecast. Thus, it provides the means for achieving both description and forecast goals.

Comparative Analysis of Exploration Strategies and Reserve Discovery and Appraisal Evolution for the Studied Regions

Models reflect exploration strategy. Therefore, comparative analysis of modeling results represents an analysis of a selection system and ensuing specifics in the reserve accumulation history within the compared regions.

The first thing to consider is the fact that variance σ^2 for region B is three to four times that for region A (see Tables 2-3 and 2-5). Is it possible that neither model is appropriate for the reserve accumulation history description in region B? The answer is a resounding no. By comparing the actual reserve accrual curves ΔR, we can see that this large variance is not a consequence of poor approximation, but of the large reserve accrual oscillations in B compared to A. The increase in the reserve accrual fluctuations in region B may be due to the substantial nonuniformity in the reserve size for new discoveries and to certain specific features of the exploration process as previously discussed. Therefore, the models are as representative for region B as they are for region A and may be legitimately compared with one another.

One difference in modeling results is obtained by ranking the models by the σ^2 size. The first model in region A is 8 and in region B, 9. At the same time, the $R(t)$ function of Model 9 has the same features in both regions. The corresponding maximum reserve accrual ΔR_{max} is higher than the same parameter given by the other models,[4] it is reached earlier than the other models (i.e., the t_{max} value is the smallest with the exception of Model 3 for region A), and only a small portion of the ultimate potential is contained in the ΔR_{max} zone. In general, the ΔR curve extends upward and its longer branch shows a slower tendency to approach the t axis. The greatest distinction between the Model 9 $R(t)$ curves for both regions is the size of t_{max}, that is, the time duration until the maximum reserve accrual is reached. This duration is much longer for region B. For region A, Models 5, 7, and 2 (the Gomperz curve) provide the closest result to that of Model 8. At the same time, Model 2 is ranked last in region B (except for Model 6). It provides the worst approximation and is significantly different from Model 5, which it was very close to in region A. Model 1 (logistic curve) is similar to Model 5 in region B but very distinct from it in region A, and approximates the observations worse than any other model.

Thus, although Model 5 is ranked second for both regions in terms of agreeing with the observations, the curves it describes are very

different. In region A this curve is similar to the Gomperz curve and in region B, to the logistic curve.

Therefore, identifying the differences in the curves may be achieved without analyzing their parameters but, instead, based on their sequential order in a variational series by the σ^2 value.

Let us analyze the curve described by Model 5. This model is of great interest as it is ranked second for both regions in terms of the agreement with observations. Remember that Model 5 has different solutions. When $b = 1$ and $c < 1$, it turns into Model 1; when $b < 1$ and $c = 1$, it turns into Model 2. The solution is different for $b < 1$ and $c < 1$ than for $b < 1$ and $c > 1$. These solutions correspond to different $R(t)$ functions. A comparison of Model 5 parameters in regions A and B (Tables 2-3 and 2-5) shows that the solutions differ for each of these regions. Namely, the reserve accumulation in these regions is described by different functions representing different solutions of the same Model 5. The $R(t)$ function in region B corresponds to the first condition ($b < 1$ and $c < 1$) and in region A, to the second condition ($b < 1$ and $c > 1$).

The difference in the Model 5 solutions for regions A and B is significant and is not associated with natural parameters (richness of the resources, field distribution by the reserve size, etc.) but, instead, with significant differences in the exploration strategies for both regions. To obtain additional confirmation of this fact, we will compare these curves in relative, not absolute, terms. This will eliminate the effect of richness of the resources and the largest fields. Usually the concept of difference in curves is explained by comparing these curves and their derivatives. In this case, additional comparison is made by analyzing some characteristic relations (see Tables 2-4 and 2-6).

According to Model 5, maximum reserve accrual in region A is almost twice as small as in region B. Its size, as evaluated from parameters of this model, relative to the ultimate potential is, respectively, 5.8% and 5.9% for regions A and B. Reaching this maximum in region A took half the time spent in region B. The share of the appraised reserves, evaluated using similar techniques, at the time of maximum was, respectively, 37% and 50%. In the maximum area within region A, it took 9 years to accrue the reserves one-half of the ultimate potential (same evaluation) and 14 years to accrue two-thirds of the ultimate potential. The respective results in region B were the same. These data illustrate the degree of the reserve accrual ΔR curve's vertical stretch and of its compression, or stretch, around the maximum. The variance of the distribution functions serves the same purpose.

Therefore, relative characteristics of the solution curves for Model 5 around the maxima in regions A and B are identical. The major distinction between them is the difference in speed and the amount of time it takes

to reach the reserve accrual maximum. This difference is no doubt a consequence of differences in the exploration strategies; in particular, the selection of major exploration plays combined with the smaller variances.

Thus, the main conclusion from the analysis of Model 5 is that the differences in its solutions reflect differences in the exploration strategies without any changes to the substance of Model 5. The model is quite successful in describing the reserve accumulation process in both regions. The best way to understand and interpret this process is to use the concepts input into Model 5 (which were discussed in detail earlier).

On the other hand, comparing the results of the reserve accrual analysis in regions A and B, it appears that the solution provided by Model 5 is only slightly different from the solutions provided by the already known functions—the logistic function (for region B) and the Gomperz function (for region A). Model 5 does not appear to produce functions that are substantially different. If this is true, Model 5 is simply a synthesis of these two functions and of no interest as a model integrating a new class, or classes, of functions.

It is interesting to review Model 7 from the same viewpoint. In both regions, the model has a small variance and is ranked third in this respect. Remember that at $b > 1$ the model has a solution different from the one at $b < 1$. In region B, $b < 1$ and in region A, $b > 1$ (Tables 2-5 and 2-3). Thus, the solutions for Model 7 (the same as for Model 5) are different in both regions. This is another confirmation of cardinal differences in the reserve accumulation process evolution for both regions. A distinction of the Model 5 solution, as was mentioned, is that for region A it provides a solution similar to the Gomperz function (Model 2) and for region B, to the logistic function (Model 1). When $b = 1$, both Model 5 and Model 7 are converted to Model 1. This leads to the assumption that Model 7 in region B will also yield a solution close to the logistic function. This supposition should be verified in regards to Model 8 because its variance ($\sigma^2 = 1561$) is closer to the variance produced by the logistic function ($\sigma^2 = 1579$) than the variance produced by Model 7 ($\sigma^2 = 1525$).

Let us review with the same purpose the curves of Models 1, 7, and 8, approximating the series of the accumulated reserves R and the series of the reserve accruals ΔR. We can see that the curves for Model 7 and for Model 1 are almost identical. Model 7's curves are even closer to the logistic curve, and to its respective reserve accrual curve, than the curves of Model 5. Therefore, Model 7 did not provide a new solution for region B.

This is not true regarding Model 8. The accumulated reserve curve (as described by Model 8) during the latest years lies above the actual R values, whereas in Models 1 and 7 it is below. As for the ΔR curve, the following is true. The maximum value of the ΔR_{max} of Model 8 is lower

than the ΔR_{max} given for Model 1, and lower yet than the ΔR_{max} given for Model 7. The reserve accrual curve of Model 8 rapidly reaches the maximum point and runs lower than the maximum ΔR_{max} of Model 1. It then crosses the curve of Model 1 and runs consistently above it. After this crossover, the deviation of the Model 8 curve from the Model 1 curve increases (i.e., the Model 8 curve is significantly different from all other curves). The distinctions are seen in all parameters (see Table 2-6). The curve is less stretched vertically and more asymmetric. The other curves basically look alike. Thus, in region B, Model 8 provides a distinct solution from the others.

We can conclude that the construction and analysis of the models confirmed a radical difference in the reserve accumulation history for the two regions under consideration. These differences are multi-faceted. Not only are the solutions different (which are given by Models 5 and 7), but the models differ to the extent they approximate the accumulated reserve series. In one region, these solutions are closer to the Gomperz function, whereas in another, to the logistic function. Finally, the model that combines a good process description and forecast is Model 8 for region A and Model 9 for region B.

Two model pairs should be set aside in both regions: Models 5 and 7 and Models 8 and 9. These models describe a diverse number of situations and compete with each other. Models 5 and 7 are constructed differently. In one case, the relative reserve accrual is plotted against the value of accumulated reserves (exponential curve), and in another the relative reserve accrual is plotted against the value of the undiscovered potential. This difference does not appear to be significant—the μ function is the derivative of the selection conditions function and increases the interpretations of the models. Models 5 and 7 are not expected to provide any substantial new information, whereas Models 8 and 9 yield radical new results. Both Models 8 and 9 take into account the field shortage quota, although the nature of the quota changes differs. At some point in time, the quota in Model 8 begins to grow in proportion to the value of the accumulated reserves, whereas in Model 9 it remains constant.

Deterministic Patterns of the Reserve Accumulation in Various Regions

When we began modeling the oil and gas reserve accumulation process, we based our study on the assumption that the evolution of this process is defined by the evolution of exploration. Exploration strategy is based on information obtained in the process of its realization. The

same information is used for the forecast, which, in turn, is used to develop subsequent control decisions. A given system of exploration control results in using a certain strategy to develop a region's resources. This strategy is defined by an agreement between the selection of the exploration plays and the location of the largest fields. One purpose of this study was to investigate to what extent the differences in exploration strategies require different formal descriptions of the reserve accumulation history. Two analyzed series of the accumulated reserves in regions A and B are standard. As was shown, using reserve accumulation history as an example, these series may serve as models reflecting in the reserve accrual two opposite exploration strategies, and, correspondingly, two different exploration control systems on a regional scale.

Modeling results confirmed assumptions of a drastically different evolution in the reserve accumulation in these regions. It was not possible to develop a single model for a formal description of both initially appraised reserve series. Model 8 had to be used in region A and Model 9 in region B. The descriptions using the same model were different not only in terms of parameters. In all cases where the models had different solutions (5 and 7), these solutions with respect to the accumulated reserve series in regions A and B did not coincide and were actually different.

There were two goals in modeling the initial appraised reserve accumulation process. The first goal was to understand the nature of the process—its mechanism. The second goal was to forecast its future course. In particular, this would allow us to estimate the ultimate potential and undiscovered resources. A number of trials to describe the process of the reserve accumulation were conducted which allowed for introducing and investigating various concepts of the process. The concepts were expressed as hypothetical functions describing the reserve control process. As a result, a model was found for each region that satisfied both goals. These goals are satisfied by Model 8 for the accumulated reserve series in region A and Model 9 for the similar series in region B.

When modeling the oil and gas accumulation process, we introduced a concept of the control function E, having thereby represented the reserve accrual as a result of control $dR/dt = f(E)$. This, in turn, required the determination of the model structure such that the reserve accumulation rate dR/dt was a function of the accumulated reserves R and geological conditions as described by a certain parameter λ:

$$dR/dt = \Phi(R, \lambda)$$

The dependence on R was due to the fact that the amount of knowledge concerning a region and the amount of information regarding

regional resources, increases with a growth in R. Based on that information, control is applied: exploration is molded to fit the changing environment, which is reflected in a certain method of drilling volume placement, and so forth. The increase in reserves is also a basis for the allocation of funds. In other words, the amount of R determines adjustments in exploration and this is exactly what affects the intensity in the reserve accrual. This adjustment was described by a function $\varphi(R)$, the function of selection control.

In a special case,

$$dR/dt = kR\lambda$$

The λ parameter is not a constant but, naturally, depends on R: $\lambda = f(R)$. A change in R causes changes in the process control system and also reflects changes in the geological conditions of exploration. As the fields are discovered and the reserves R accumulated, a parameter is changing that can be described as nature's ability to produce reserves. If the reserves are still concentrated in large fields, this capability is high; after large discoveries, it drastically drops. Thus, beginning at some point, the λ parameter must start decreasing with a growth in R.

Different types of λ decrease with an increase in R were tested (i.e., different variations of the $\lambda = f(R)$ function). The $f(R)$ function was named the selection conditions function. In the constructed models, where difference equivalents of the differential equations were used, a few possible forms of the $f(R)$ function were suggested. These models reflect the appearance in the process of exploration of a mechanism constraining the reserve accrual. This is the simplest self-inhibition mechanism, which becomes influential when the accumulated reserves R_i become large enough. The models differ in the R cut-off value, which "triggers" this mechanism, and in the λ decline rate with a growth in R.

The conditions selection function was assigned proportionately to the accumulated reserves (i.e., $N(R) = kR$). Therefore, the question remains as to the type of conditions selection function $f(R)$ resulting from different strategies. These different strategies lead to a different intensity of destruction of the original structure of ultimate potential. If a discrete description form is used, this question relates to changes in the relative reserve accrual. Strictly speaking, the purpose of modeling was to find a function that would best describe changes in the relative reserve accrual. A hypothetical function was proposed based on the typical features of exploration discussed in Chapter 1. Additionally, a concept of the relative reserve accrual was applied in a form that follows the logistic and Gomperz functions. These two functions were previously used for analyzing the reserve evolution.

Four classes of such functions were analyzed:

$$\Delta R/R_{i+1} = c - aR_i^b \tag{2-88}$$

$$\Delta R/R_{i+1} = c(A - R_i)^b \tag{2-89}$$

$$\Delta R/R_{i+1} = e^{a-bR_i} - n_i \tag{2-90}$$

$$\Delta R/R_{i+1} = \ln(A/R_i)^b \tag{2-91}$$

Of these, only one class (Equation 2-90) includes the functions that were preferred at different stages of this study and that were eventually chosen for both regions. These include Models 3 ($n_i = 0$), 8 ($n_i = c$), and 9 ($n_i = c/R_i$). These models were first introduced earlier and were not used in the other disciplines where growth models were used. This is a specific feature of the reserve accumulation that cannot be compared, for instance, with the growth in biologic populations or in the amount of information gained (for these, the logistic or Gomperz function are used). In the case of reserve accumulation, exponential equations of the Equation 2-90 type play a special role. This is associated with a drastic non-uniformity in the field's ultimate potential distribution (a small number of fields contain a substantial portion of the ultimate potential). The discovery of such fields manifests a rapid growth in the reserve accrual and its rapid decline after such discoveries are exhausted. Subsequently, the reserve accrual slowly dies out. The models belonging to the Equation 2-90 type are best equipped to describe this particular feature.

Another important feature of Model 3 is that its solution, the $R(t)$ function, does not have an asymptote. When discussing this model, we explained why the reserve accumulation series did not require an asymptote. This feature, as well as others, is not encountered in other scientific disciplines.

Model 9 is even more specific. This is the only model, according to the shortage quota, that allows the selection conditions to improve in the beginning. Its other distinctive feature is the concept of an intermittent appearance of some initial $\Delta R = A_{lower}$ after which a pattern in the reserve accumulation is observed. The initial accrual $\Delta R = A_{lower}$ does not fit this pattern—this occurs for totally different reasons and opens a new stage in a region's development.

A_{lower} may be described as a starting experiment as if conducted by someone in charge of exploration. This experiment includes all exploration efforts preceding the first discovery, after which a new stage begins in the region's exploration.

Thus, we have shown that variations of the Equation 2-90 functions, at different types of n_i function, describes numerous diverse situations for the control decisions leading to different resource development strategies.

Notes

1. "Preliminary" evaluated reserves, or C_2 category, encompass reserves that include the lower portion of the probable, and the upper portion of the possible, U.S. category.
2. Of course, any other procedure of non-linear minimization can be applied to determine the best estimates of the model parameters.
3. R_{lim} is evaluated at $R_{\text{lim}} = 1$.
4. Although in region B, the curve from Model 7 has a slightly higher ΔR_{max} value.

CHAPTER 3

Evolution of the Reserve Accumulation: Stochastic Models

As previously indicated, the reserve accumulation process is a reflection of the exploration strategy. When controlling exploration the problem of selection is always present. Each decision reflects a large number of degrees of freedom. Numerous possible solutions lead to fluctuations, that is, spontaneous variations in the accumulated reserves around some average behavior. A deterministic description is insufficient when dealing with problems of exploration control and, therefore, with the oil and gas reserve accumulation. In most cases, fluctuations may be considered random events governed by the law of probabilities. This leads us to consider the elements of chance and deterministic law which are both present in the reserve accumulation process. These elements interrelate in such a manner that they cooperate rather than contradict or conflict with each other. In particular, the regular can be created through the random, or the random may be a complication of the regular.

For this reason, this chapter includes a discussion of the interaction between determinism and chance in reserve accumulation evolution.

Study of the Reserve Discovery and Appraisal Strategy from the Information Viewpoint

The reserve accumulation process may be considered an information process. This leads to certain connections between subsequent and preceding results. These connections are affected by random (probabilistic) variations in the process evolution. In this chapter we will analyze the role of these random variations. Modeling of the reserve accumulation process

will be based on these differing concepts. What follows are the particulars of model construction in this situation.

The solution of a differential equation that depicts the mechanism of a certain phenomenon is a continuous deterministic function. This function describes the process with the certain mechanism of formation. As in Chapter 2, the idea of the selection function was described in the form of differential equations (or, rather, their discrete equivalents). Correspondingly, the process description was in the form of a continuous (deterministic) function. However, a deterministic function corresponds to a differential equation without a random term. When such a term is present (and has a fixed normal distribution with a zero average and nonzero variance), the differential equation is called "stochastic."

A distinctive feature of models constructed in the form of stochastic differential equations is that their solution is not a deterministic time function but, rather, a random function that changes as the process evolves. Process descriptions using the corresponding discrete equivalents of stochastic equations also result in stochastic (discrete) processes (stochastic approximation). In this section, the process of the initial oil and gas appraised reserve accumulation will be analyzed using discrete equivalents of the stochastic differential equations.

It should be noted that approximating the observation series (or the decomposition of the series into a regular component and random component) using deterministic continuous functions has for many years been used successfully in geologic studies. To the best of our knowledge, no modeling of processes was performed using stochastic differential equations or the corresponding stochastic discrete equivalents. This study appears to be the first of its kind [21]. The use of such models for analysis of the reserve accrual process may be very important.

Discrete Stochastic Models

Stochastic models were constructed similar to Models 1 through 5 (Chapter 2) by adding a random term to the respective equations. The reserves accumulated by the end of year i were designated as P_i. The reserve accrual during the $i + 1$ year was designated ΔP (i.e., $\Delta P = P_{i+1} - P_i$). Then, based on Equations 2.2, 2.12, 2.27, 2.36, and 2.57, the following five stochastic models were obtained:

$$P_{i+1} = \frac{P_i}{1 - a + bP_i} + Z_{i+1} \tag{3-1}$$

$$P_{i+1} = aP_i^b + Z_{i+1} \tag{3-2}$$

$$P_{i+1} = \frac{P_i}{1 - e^{a-bP_i}} + Z_{i+1} \tag{3-3}$$

$$P_{i+1} = \frac{P_i}{(1 - c + aP_i^{1/b})^b} + Z_{i+1} \tag{3-4}$$

$$P_{i+1} = \frac{P_i}{1 - c + aP_i^b} + Z_{i+1} \tag{3-5}$$

where Z_{i+1} is a random process (white noise).

The solutions are not the corresponding deterministic functions. The trajectory of the reserve accumulation described by each model will change from one year to the next depending on the actual P_i values and not on the "theoretical" values cleaned of all random noise. The trajectory may be complex or jagged, unlike the corresponding deterministic curve. In other words, the solution will be a stochastic process. It is described by Equations 3-1 through 3-5 and cannot be presented analytically as a function of t. At the same time, these equations show which deterministic function may approximate a respective dependence of P_{i+1} on P_i (in the absence of random disturbances). It may also be said that these equations describe respective deterministic curves with a random R_o value that changes each time depending on the actual P_i value at a given stage of the process (in a given year). In other words, this stochastic process is a combination of segments of the respective deterministic curve with constant a, b, and c parameters and a random R_o parameter.

From a probabilistic viewpoint this means that the distribution mode changes randomly; that is, the distribution curve shifts right or left but its shape is not altered.

As indicated, the models analyzed in Chapter 2 may be compared with accumulation of information. From an informational standpoint, the difference between the models, or their discrete equivalents, is clearer when written in the form of a differential equation without the random term and the stochastic differential equation. At each new stage of study the researchers use knowledge that has previously been obtained. Each subsequent level will be determined only by the preceding level (with some random variations). This very connection is exactly what is reflected by a stochastic equation. This situation will only occur in the case of ideal communications between scientists, when new information immediately becomes common knowledge and is incorporated in subsequent study (realization of a new opportunity). If this condition is not realized, the state of knowledge that affects the subsequent level will not be determined by the achieved level, but by some averaged, smoothed level as described

by the deterministic trend. In this case, one may imagine some inertia in the process of knowledge accumulation or an absence of timely reaction to new knowledge. This difference is especially important when analyzing physical systems.

Let us review oscillations of a pendulum damped by the air resistance. If we disturb its equilibrium with a single impulse Z, it will oscillate for a while and then stop. Its trajectory will be a fading sinusoid. The pendulum movements in such a case are described by a certain differential equation or its discrete equivalent with no random term; its solution is the fading sinusoid (a deterministic function).

Let us now imagine a situation where the pendulum is subjected to random impulses Z_i over equal time intervals. After the first impulse, before the pendulum has stopped, it is subjected to a second impulse, then a third, and so forth. The impulses are separated by equal time intervals but differ in their intensity and direction (i.e., they may go with or against the pendulum's movement). Thus, within each time interval between impulses, the pendulum's motion will behave as a segment of different fading sinusoids. Each time the amplitude and initial phase of these sinusoids will change randomly. At the same time, the frequency of oscillations and the damping factor will remain constant because they are determined solely by the pendulum's physical parameters. Therefore, instead of fading oscillation, the pendulum will now be performing a disturbed periodic movement. This movement will be described by a distorted periodic function, that is, a series of empiric observations will reveal a pseudo-periodic behavior. The trajectory will not represent a smooth fading sinusoid. The pendulum movement in this case will be described by the same differential equation but now with the random term distorting its trajectory. At the same time, the parameters of both equations are the same because they are associated with the sinusoid frequency and the damping factor, which, in turn, are associated with the pendulum's physical parameters (e.g., mass, length). This is because it is assumed that the pendulum immediately reacts to new impulses and begins the new movement from the position it was in at the moment of the impulse (i.e., the pendulum does not have inertia).

Thus, the stochastic trend may be interpreted as an indicator of efficiency in processing the information obtained in the reserve accumulation process. It can indicate a strong association between the obtained result and decisions on the direction of future work. These decisions are made immediately after obtaining the result, so the next reserve accrual depends on the reserve level accumulated at that moment. Conversely, if deterministic models better fit the observations, it will mean that the process under study has inertia.

Evaluation and Diagnostic Review of Models

The best of the constructed models may be selected after evaluating their agreement with field data in terms of variance σ^2, as obtained through a stochastic approximation of the observations.

Stochastic models were analyzed using the same region A. Model parameters were evaluated with the technique described previously. An interesting feature of this evaluation is that the R_o parameter is absent.

The stochastic process P_i occurs in time. The dependence of P_{i+1} on P_i is described by a deterministic function. The P_{i+1} versus P_i curves are similar for different models. The differences between them are their dissimilar inflections. As a result, their positions relative to each other change with a growth in P. The strongest curvature occurs in Model 3-1, whereas the least curvature occurs in Model 3-3. It is indicative that over the large P values, all curves run below the observed data except for curve 3-3, which runs above it.

In general, the stochastic process P curves 3-1 through 3-5 fit the observed reserve accumulation curves. The strongest discrepancies between them occur over the time interval (i) of 8 to 12 years. It is important that for the last observed years the model P values are lower than the observed ones. Again, the only exception is curve 3-3.

The stochastic trend ΔP of the reserve accrual for all models is not much different from the deterministic trend ΔR of Models 1 through 5. It is interesting to note, that the best approximation for some models, such as 3-1 and 3-4, leads to negative ΔP values over the last several i values. This imposes certain limitations on the use of these models.

Table 3-1 provides tentative estimates of the model parameters and variance σ^2. It shows that all models, except Model 1, yield a similar degree of approximation, with variance values within the 755 to 789 range. Still, Models 3 and 5 display somewhat smaller dispersion values. The F-criterion estimates of dispersion difference are insignificant (i.e., strictly speaking, all models should be considered equal in terms of their fit to observations). This table shows that the stochastic approximation does not reach the same degree of proximity as the deterministic approximations for Models 5 and 2 where the variance was, respectively, 501.34 and 510.98.

The fact that the reserve accumulation history for the period under consideration is better approximated by a deterministic function is an indication of the inertial nature of the exploration process. Significantly "smoothed" or "washed-out" information is used while exploration is being conducted. The most recent data are incomplete and are taken into account after a delay. This is another characteristic of the mechanism, which provides for a pattern in the reserve accumulation history. If,

Table 3-1
Adequacy Estimate of the Stochastic Trend Models of the Reserve Accumulation Evolution to Observations (Region A)

Model Number	Model	*a*	*b*	*c*	σ^2	*A**
1	$\frac{P_i}{1 - a + bP_i} + Z_{i+1}$	0.22	0.0002	—	912.05	11
2	$aP_i^b + Z_{i+1}$	2.86	0.851	—	765.1	11.56
3	$\frac{P_i}{1 - e^{a-bP_i}} + Z_{i+1}$	–0.79	0.0029	—	755.33	24.14[1]
4	$\frac{P_i}{(1 - c + aP_i^{1/b})^b} + Z_{i+1}$	0.1663	900	0.1676	789.37	11.04
5	$\frac{P_i}{1 - c + aP_i^b} + Z_{i+1}$	1.3167	0.066	2.1	759.57	11.8

*See Equation 2.3, p. 24.

however, in choosing a model we use σ^2 as well as other factors (as shown in Chapter 2), Model 3 should be considered. This model was given a certain preference for the deterministic approximation.

Let us conduct a comparative analysis. The variance for the deterministic approximation with Model 3 was 758.46. Based on this, it would be wrong to maintain that the deterministic model better describes the observations than the stochastic model. This cannot be asserted with certainty because the variances of the deterministic and stochastic Models 3 are close. A comparison of these two models indicates that the stochastic approximation curves for the accumulated reserves P and for the reserve accruals ΔP fluctuate around the corresponding deterministic curves. The deterministic curve appears to be a smoothing-out of the stochastic curve. The deviations between the curves are insignificant.

The similarity in the approximations provided by the stochastic and deterministic versions of Model 3, forces us to reconsider the earlier suggestion that the reserve accumulation process and, therefore, the process of exploration, are heavily inertial processes. Most likely, the nature of the process is mixed: it is affected by both the actual observed level P and the deterministic level R. The latter may be interpreted as an averaged result, which is not subject to significant fluctuations.

It may be concluded that the stochastic model does not provide clear results. Apparently, the deterministic trend better reveals the nature of the

reserve accumulation process. If any randomness is present, it is applicable to random components remaining after the removal of the deterministic trend. This aspect of stochastic approximation will be discussed later in this chapter.

Linearization of Models

Another important aspect of this stochastic approximation is its linear presentation of models. Linear models are simple; an evaluation concerning their fit with observations may be made graphically. For this reason, linear presentation is a convenient method of model identification.

Identification is a rather crude procedure for screening models and leaves only those which justify further analysis using more formal and efficient techniques. Identification and evaluation overlap, but the evaluation in this case is preliminary and is used only as an initial parameter approximation for further studies. In other words, the evaluation procedure discussed here assumes the role of identification. Linear evaluation of parameters is not difficult. It is performed using the least squares technique. Model reduction to a linear form, however, causes unavoidable distortions; thus, one must keep in mind that identification is always inaccurate. At a preliminary stage of study, model selection and the evaluation of its parameters are important. Quite often, analysis of the model is reduced to analysis of its linear representation. Some scientists go as far as to suggest that the model parameters should be evaluated from their linear form. This makes it important to determine to what extent the results based on linear representation differ from the final results.

Models 1 through 5 may be reduced to the following linear representations:

$$\Delta P/P_{i+1} = \alpha - \beta P_i + X_{i+1} \tag{3-6}$$

$$\ln P_{i+1} = \alpha + \beta \ln P_i + X_{i+1} \text{ or}$$

$$\ln(1 - \Delta P/P_{i+1}) = \alpha + \beta \ln P_i + X_{i+1} \tag{3-7}$$

$$\ln(\Delta P/P_{i+1}) = \alpha - \beta P_i + X_{i+1} \tag{3-8}$$

$$(P_{i+1}^{1/\beta} - P_i^{1/\beta})/P_{i+1}^{1/\beta} = \gamma - \alpha P_i^{1/\beta} + X_{i+1} \tag{3-9}$$

$$\ln(c - \Delta P/P_{i+1}) = \alpha + \beta \ln P_i + X_{i+1} \text{ or}$$

$$\Delta P/P_{i+1} = \gamma - \alpha P_i^{\beta} + X_{i+1} \tag{3-10}$$

The relationship between parameters in Equations 3-1 through 3-5 and 3-6 through 3-10 is indicated in Table 3-2.

Model 4 acquires a linear format only with a fixed parameter $b = \beta$. So, to obtain parameter estimates, different values of the β parameter were used. For each fixed value of β_i, γ_i and α_i parameters were estimated as well as σ^2 of the stochastic trend with these parameters. The value β^* at which the minimum of σ^2 was reached served as an estimate of the β parameter. Respectively, γ^* and α^* were estimates of the γ and α parameters. Model 5 becomes linear in two cases: at the fixed c parameter and at the fixed β parameter. The same technique was applied to determining the model's parameters. In this case, however, c was the arbitrarily varied parameter. In some regions c was small; consequently, for some P_i the difference $c - \Delta P/P_{i+1}$ became negative, which did not allow calculating logarithm. In such cases, b was the arbitrarily varied parameter.

Region A. Tentative estimates of Model 3-6 through 3-10 parameters (estimates of the linear connection) are listed in Table 3-3. For comparison purposes, these estimates are recalculated (see Table 3-2) as representation parameters 3-1 through 3-5 (linear estimates of the parameters). Due to different functions representing the left parts for different models, variances σ_l^2 are not comparable and cannot be used to evaluate which model better fits observations. The correlation coefficient r, which is applied for this purpose, is a dimensionless measure of the linear connection that does not depend on the format of the model's left parts. The stochastic approximation variance σ_s^2, obtained while linearly estimating the parameters, may be used for the same purpose.

Table 3-3 shows that the correlation between the models and the respective observations is strong. Model 1 is no exception. Its connections have a rather high correlation coefficient $|r| \approx 0.8$, although it is lower than in the other cases where $|r| \approx 0.9$. The correlation coefficient is especially high for Model 2 (on the first version of its linear representation) where the connection is almost functional. Judging by the correlation coefficient, preference should be given to Models 2, 3, and 5 in terms of fitting the actual observation.

Linear estimates of the model parameters are different from estimates obtained in the deterministic and stochastic approximation of the observations. The degree of such disparity is illustrated by variances of the stochastic and deterministic approximations with the linear evaluations. This data is listed in Table 3-3, which also shows the values of the R_o parameter corresponding to the linear estimates of the other parameters (without it, variance of the deterministic approximation cannot be calculated). When the model is linearized, however, it is impossible to obtain an estimate of the R_o parameter. Consequently, it was obtained by

Table 3-2
Linear Representation of Models

Sequential Number	Model	Linear Representation	Parameter Relations
1	$P_{i+1} = \dfrac{P_i}{1 - a + bP_i} + Z_{i+1}$	$\dfrac{\Delta P}{P_{i+1}} = \alpha - \beta P_i + X_{i+1}$	$\alpha = a;\ \beta = b$
2	$P_{i+1} = aP_i^b + Z_{i+1}$	(a) $\ln P_{i+1} = \alpha + \beta \ln P_i + X_{i+1}$ or (b) $1(\Delta P/P_{i+1}) = \alpha + \beta \ln P_i + X_{i+1}$	$\alpha = \ln a;\ \beta = \ln b$ $\alpha = \ln a';\ \beta = \ln b'$*
3	$P_{i+1} = \dfrac{P_i}{1 - e^{a-bP_i}} + Z_{i+1}$	$\ln(P/P_{i+1}) = \alpha - \beta P_i + X_{i+1}$	$\alpha = a;\ \beta = b$
4	$P_{i+1} = \dfrac{P_i}{(1 - c + aP_i^{1/b})^b} + Z_{i+1}$	$(P_{i+1}^{1/\beta} - P_i^{1/\beta})/P_{i+1}^{1/\beta} = \gamma - \alpha P_i^{1/\beta} + X_{i+1}$	$\alpha = a;\ \beta = b;\ \gamma = c$
5	$P_{i+1} = \dfrac{P_i}{1 - c + aP_i^b} + Z_{i+1}$	$(\Delta P/P_{i+1}) = \alpha + \beta \ln P_i + X_{i+1}$ or $\Delta P/P_{i+1} = \gamma - \alpha P_i^{\beta} + X^{i+1}$	$\alpha = \ln a;\ \beta = \ln b$ $\alpha = a;\ \beta = b;\ \gamma = c$

* $a = 1/a';\ b = 1 - b'$

Table 3-3
Parameters of Linear Models (Region A)

Model Number	Tentative Estimate of the Linear Dependence				
	α	β	γ	r	σ_l^2
1	0.367	0.0003624	—	–0.8044	0.01144
2a	1.284	0.8144	—	0.9945	0.01405
2b	–1.284	0.1856	—	0.9079	0.01402
3	–0.8716	0.003034	—	–0.9274	0.2391
4	0.1753	100	0.1879	–0.9059	$4.25*10^{-5}$
5	0.3394	0.0639	2.2	0.9226	$1.3610*10^{-3}$

Model Number	Tentative Linear Estimates of Parameters			Stochastic Approximation Variance σ_s^2 for Linear Estimates of Parameters	R_0 Estimate Corresponding to Linear Estimates of the Other Parameters	Deterministic Approximation Variance σ_d^2 for Linear Estimates of Parameters
	a	*b*	*c*			
1	0.367	0.0003624	—	1767.63	3.8	5972.48
2a	0.3611	0.8144	—	1026.3	0.5	3821.95
2b	—	—	—	—	—	—
3	–0.8716	0.003034	—	892.69	13.5	3583.09
4	0.1753	100	0.1879	949.41	0.4	2568.1
5	1.4041	0.0639	2.2	781	1.3	979.28

minimizing the deterministic approximation variance using linear estimates of the other parameters. Table 3-3 shows that the linear estimates result in substantially higher variances than the estimates obtained from directly forcing the parameters to fit the observations. The best stochastic approximation of the linear parameter estimates occurs in Models 5 and 3 and the worst in Model 1. This is consistent with the results obtained when the stochastic models were directly fitted to the observations (see Table 3-1).

This type of clear similarity is not observed in the case of deterministic approximation (see Table 2-3). It is important to note, that the worst

and the best results do coincide. The best approximation is reached in Model 5 and the worst in Model 1. With the parameters obtained by directly forcing Model 2 to fit the observations (see Table 2-3), the result (the value of variance) was similar to Model 5 (i.e., Model 2 was ranked second, after Model 5, in terms of fitting the observations). With linear estimates, it moved to fourth place and Model 4 became second. Earlier with direct estimates its variance was close to that of Model 3. At present, with linear estimates, it is preferable to Model 3 in this respect.

A comparison of the dispersions for the stochastic and deterministic approximations for linear estimates (see Table 3-3) indicates that the former is much lower, whereas the situation was reversed for direct estimates (compare Tables 2-3 and 3-1).

Thus, our analysis shows that the linear model can be used for determining which model better fits the observations. With the deterministic and stochastic approximations the indicator will be the minimum value of variance obtained for linear estimates. Although deviations do occur, the worst models can be screened out.

The linear representation, however, is not suitable for even rough parameter estimates or for evaluating which trend (stochastic or deterministic) better approximates the observations.

Continuing the parameter comparison for different model formats, it can be seen that the parameter estimates for the deterministic and stochastic models do not coincide (except for Model 2; compare Tables 2-3 and 3-1). This means that the stochastic model estimates may not be used for the deterministic model, although this appears attractive due to the smaller (by one) number of parameters in the stochastic approximations which makes the calculations simpler. Afterwards, R_o parameter can be estimated.

To illustrate the consequences of such a substitution, Table 3-4 displays variances for the deterministic approximations when the stochastic model estimates are used. This table also gives R_o estimates which were computed based on variance minimization using the stochastic approximation estimates. Comparing Tables 3-4 and 2-3, it becomes clear that the substitution does not result in deterioration of the deterministic approximations, except for Models 3 and 4. The best approximations are still given by Models 5 and 2, and the worst by Model 1. The same result is obtained by directly fitting the models to the observations.

Thus, unless rigorous standards are necessary, stochastic approximation estimates and corresponding R_o estimates are sufficient for the deterministic approximation. Correspondingly, these estimates may be used for evaluating different deterministic models with observations and determining which approximation (stochastic or deterministic) better fits the observations.

Table 3-4
Deterministic Approximation Variance with the Parameters Taken from Stochastic Approximation (Region A)

Model	R_o Estimate Corresponding to Estimates of Stochastic Approximation	Variance of Deterministic Approximation with Estimates Obtained with Stochastic Approximation
1	41	1,272.14
2	3.5	510.98
3	5	832.251
4	1	810.2926
5	1	509.0874

Region B. The results for the parameter estimates are displayed in Table 3-5. This table indicates that the correlation of the corresponding independent variables is very low (apparently, Model 2 in its first representation should not be used). Judging by the σ_s^2 and r value, Model 5 appears to be the most preferable. Model 1 can to some extent compete with Model 5. The worst, in terms of agreement with observations, is Model 2. Thus, an analysis of models in region B using their linear approximations provides results similar to those obtained under rigorous conditions.

It is interesting to examine to what extent the earlier established differences in the reserve accumulation evolution for regions A and B can be identified under linear model approximation (Tables 3-3 and 3-5). The first difference between the regions is in the value of the correlation coefficient. It is much lower for region B than for region A. The sequential order of the models by the size of stochastic approximation variance σ_s^2 is also different. Models 5 and 3 have the lowest variance in region A, whereas Models 5 and 1 are in corresponding positions in region B, and the largest variance is demonstrated by Models 1 (region A) and 2 (region B). In terms of identifying the best and worst models, the order is the same as that determined by a rigorous analysis. As shown previously, the solutions obtained by using Model 5 are different, which is reflected in the considerable parameter difference. In region A, Model 5 has $c > 1$, $a > 1$, and $b < 1$; whereas in region B, $c < 1$, $a << 1$, and $b > 1$. The difference in the b size is especially important. As we indicated earlier, there are drastic differences in the conditions selection curves for situations with $b > 1$ and $b < 1$.

Therefore, the linear representation of Models 1 through 5 may be used in both regions for judging which models best fit the observations.

Table 3-5
Parameters of Linear Models (Region B)

Model No.	Tentative Estimates of Linear Connection				
	α	β	γ	r	σ_l^2
1	0.1271	$3.5630*10^{-5}$	—	–0.2977	$6.5226*10^{-3}$
2a	0.1416	0.9953	—	0.9978	0.01001
2b	–0.1416	0.0047	—	0.07162	0.01001
3	–2.61	$3.716*10^{-4}$	—	–0.28	0.7578
4	Not estimated	Not estimated	Not estimated	Not estimated	Not estimated
5	$5.6372*10^{-10}$	2.5	0.1247	–0.3857	$6.09*10^{-3}$

Model No.	Linear Estimates of Parameters (Tentative)			Stochastic Approximation Variance σ_s^2
	a	*b*	*c*	
1	0.1271	$3.5630*10^{-5}$	—	2,332.48
2a	1.1521	0.9953	—	5,362.74
2b	0.868	0.0047	—	—
3	–2.61	$3.715*10^{-4}$	—	3,270.48
4	Not estimated	Not estimated	Not estimated	Not estimated
5	$5.6372*10^{-10}$	2.5	0.1247	1,788.56

However, it cannot be applied for parameter evaluation or for determining whether stochastic or deterministic trends occur in the reserve accumulation evolution.

At the same time, it may be used to identify the major differences in the reserve accumulation evolution within different regions. In other words, it is a suitable means to classify (to separate into groups) different reserve accumulation curves obtained for different regions.

Classification of Regions by Specific Reserve Discovery and Appraisal Strategies

Six regions were selected as study areas: regions C through H. Together, they represent the diverse situations that may occur in development histories and in exploration strategies as they fit the regions' geology. The main strategic issue is the selection of the best exploration

play. Whenever the right selection was made (even if by accident), large fields were soon discovered, exploration evolved rapidly, and the region as a whole was expeditiously developed. Among such regions are regions C, D, G, and the already mentioned region A. In the development of these regions the best plays were quickly identified, and exploration was concentrated there. Regional studies were in correct proportion to exploration as a whole. The prospects were efficiently identified and mapped using techniques best suited to the regional geology (e.g., map drilling, seismic). The entire sediment cover was quickly studied by exploratory drilling.

The expectation is that these regions will display similar reserve accrual patterns. The only exception might be region D, where a major field was rapidly discovered. Region C may also be somewhat specific because high reserve accrual was supported by the constant flow of large discoveries over an extended time period. Region F may also fall into this category; although exploration was not conducted in the best possible manner, we analyzed the region after the best exploration play was already identified.

Region E is the antithesis of all these regions. It does not have high potential and has been developed less intensively. The initial exploration resulted in an underestimation of the region's potential. Exploration work was reduced before it began to grow. This delayed the discovery of the best exploration play. Large discoveries were made at the closing stage of the region's development rather than at the beginning. Apparently, the reserve accrual pattern in this region will be described by different models or possibly by their corresponding solutions, as for the previously analyzed region B.

The position of region H in this series is not clear. On one hand, a major discovery was rapidly made. For a long time, however, only the upper 3.5 to 4 km of the sediment cover was relatively well studied. Putting aside for the moment the development of deeper horizons (which is a different selection situation) and taking into consideration only these upper sediments, the history of the region is reminiscent of regions C, D, and G. A major field was quickly discovered, exploration proceeded as planned, and development was evolving in well-planned phases. Techniques used for the identification and mapping of prospects was in accordance with the geology of the upper section. Correspondingly, the reserve accrual should be similar. A peculiarity may arise, however, due to the fact that the entire reserve accruals, including those made later at great depths, are being analyzed here.

The following is an examination of these suppositions and a review of how the features of the exploration strategies mentioned above for various regions affected the reserve accumulation evolution models.

We will try to identify the similarities and differences in the reserve accumulation curves as they are affected by the similarities and differences in

exploration strategies. To this end, all five models, both stochastic and deterministic versions, will be analyzed using linear model representations. Based on the results presented earlier, the following criteria will be used to compare the reserve accumulation curves: (1) differences or similarities in the solution of Model 5 (and 4); (2) the size of the correlation coefficient (based on the average for all models; Model 2 is only considered in variant 2 of its linear representation); and (3) the sequential order of models by the size of their deterministic and stochastic variance. When analyzing the sequential order, we will be most interested in models with the largest or smallest variance.

The results and the regions' classification are presented in Table 3-6. According to the differences in the Model 5 (and Model 4) solutions (namely, in the b parameter), two groups of regions were identified: those with $b < 1$ and $b > 1$. The first group includes regions C, D, F, G, H, and the earlier analyzed region A. The second group includes region E and the earlier analyzed region B. These groups encompass two totally different types of the exploration history and strategy, which were discussed in detail earlier. It is clear that the development history and exploration strategies are generally similar for the regions belonging to the same group and are totally different in regions belonging to different groups. On the other hand, a subdivision into two groups shows that regions A and B may be considered typical representatives of different regions. Region D in the first group may be identified as a special type; its reserve accumulation evolution is characterized by a special solution of Model 5. This type occurs because the largest discovery in the region, along with a substantial portion of the ultimate potential, was made rapidly, leading to a specific mode of change in the reserve accrual.

Based on the size of the correlation coefficient, the regions may be subdivided into two groups. The first group has high absolute values of the coefficient r, whereas the second group has low values. Region D in this respect is in an intermediate position. Using the same parameter, the regions are differentiated within the first group where region C in this respect is indistinguishable from region D. The specificity of the reserve accumulation curve in region C is due to the fact that a high reserve accrual level was maintained because of the number of large discoveries during an entire decade. This was determined not so much by the exploration strategy as by the resource structure in the region; hence, the specificity in region C may also be associated with the resource structure. The same may be true for region D. The presence of a unique[2] field makes the resource structure specific. This results in a specific type of reserve accumulation evolution.

Based on the size of the correlation coefficient, region H occupies a special position among the first group (absolute value of the coefficient

Table 3-6
Typical Features of Linear Models of the Reserve Accumulation Evolution

			Absolute Value of the Correlation Coefficient							
			≈0.9	0.6 – 0.8	≈0.5	0.3 – 0.4				
Model 5 (Model 4 in parentheses) parameters	b < 1 (b > 1)	c > 1 $a \approx n*10^{-1}$	G				3 and 1	Models Giving Minimum and Maximum Variance of the Deterministic Approximation	3 and 1	Models Giving Minimum and Maximum Variance of the Stochastic Approximation
			A				5 and 1		5 and 1	
			F		H					
				C			4 and 3		5 and 2	
		c < 1 $a \approx n*10^{-2}$		D					4 and 1	
	b > 1 (b < 1)	c < 1 $a \approx n*10^{-6}$				B			5 and 2	
						E			5 and 3	

ranges from 0.37 and 0.56). It appears quite natural. What was mentioned earlier regarding its development history and exploration techniques indicates that the region belongs to the first group only with respect to the upper portion of the sediment cover. If deeper horizons are considered, then its characteristics are those of the second group. This particular feature caused the correlation coefficient r to be close to that for the second group.

These properties allow for further subdivision within the first group. Region G is very special in this regard. This specificity is in the particular sequential order of the models. With stochastic approximation, Model 5 is more or less number one in each region (first or second group). Model 1 is usually ranked last among the first group of regions, and Models 2 or 3 among the second group of regions. With the deterministic approximation, Model 1 is not always last in the first group. In regions C and D, as in the second group, Model 3 ranks last. Results of stochastic and deterministic approximation (for models with the largest and smallest variance) coincide only for regions A, F, G, and H. Region G differs in that Model 3 yields the best results. This is a large region with a long development history. Currently, it is at a stage when only small discoveries are anticipated. This is why it is quite possible that the reserve accrual pattern, identified in region G, is more indicative. Thus, it is important to emphasize that Model 3 was the best for forecasting in regions A and B (which was one of the goals), but was not the best in approximating the observations. Namely, the reserve accumulation process (which was another goal) was better explained with another model. In region G, Model 3 is also preferable from this second standpoint (i.e., the standpoint of the reserve accumulation mechanism).

In general, Table 3-6 indicates a strong association between parameters: the curves that are different for one parameter are different for all others. Regions A, F, and G are clearly distinct from regions B and E in all parameters. Regions C, D, and H occupy an intermediate and unclear position. In some parameters, they are indistinguishable from regions A, F, and G, and in others from regions B and E. It is important for further analysis that regions A and B, which are most important in terms of the reserve accumulation evolution, are clearly different from each other in this respect and are in a certain sense typical. This provides a reason for a broad interpretation of the reserve accrual patterns established in these regions.

It may be concluded that the oil and gas reserve accrual patterns are less dependent on the resource structure than they are on exploration strategies. Sometimes, however, the resource structure may be decisive, as in regions C and D. Based on the development history, exploration strategy, and resource structure, at least two groups of regions are

identified. Respectively, there are at least two groups of reserve accumulation curves. Constructed models reflect specific features of these groups. In other words, modeling is an efficient means of classifying the reserve accumulation curves and of understanding how different types of reserve accrual patterns are formed. Thereby, modeling helps unravel the nature of the reserve accumulation process, which is related in particular to the exploration strategies and the resource structure. The latter two factors determine the type of relative reserve accrual (conditions selection function) from which the respective models were inferred. In turn, the reserve accumulation models provide a generalized, integral description of the development strategies, and they may be used for determining strategy. In particular, the best strategy provides the fastest reserve accrual. The reserves accrued over the maximum zone within the least amount of time, comprise the largest possible percentage of the ultimate potential.

These features are reflected in the type of model and the values of its parameters.

Study of Random Components of the Oil and Gas Reserve Accumulation Process

In this section, we will approach the study of the deterministic reserve accumulation patterns from a different angle. Random components in the reserve accumulation will be viewed as components complicating the deterministic accumulation. The purpose of this investigation is to evaluate to what extent the behaviors of the random and regular components are interconnected.

Rhythmic Nature of Random Components

A random component is a deviation of the actual value from that established by the deterministic model. For the accumulated accrued reserves, the random component is equal to:

$$Z_i = R_i - P_i \tag{3-11}$$

where R_i is the reserves accumulated during the year i given by a deterministic model, and P_i is the actual value accumulated in the year i.

For random reserve accrual components:

$$Z_i' = \Delta R - \Delta P \tag{3-12}$$

where $\Delta R = R_i - R_{i-1}$ and $\Delta P = P_i - P_{i-1}$.

The behavior of the Z and Z' series in regions A and B is shown in Figure 3-1. The models that produced the best description were used as the models that provided the deterministic function for approximating the observations (thereby, determining the R_i and ΔR values). These included Model 8 for region A and Model 9 for region B. Figure 3-1 indicates that random components Z and Z' behave similarly. In both regions they begin small and then increase, reaching a maximum amplitude before gradually weakening. They go through a swing (oscillation) before reaching a maximum and then fade. Looking at the curve, one may get the impression that there is a certain repetitiveness (a rhythmic nature) in the Z and Z' behavior. There is a vague similarity in shape over some intervals, which is explained as a simple repetition with some change. What we are most concerned with is the stochastic connection of separate curve segments. If S is the length, or period, of the rhythm, then the association between a series Z_i, Z_{i+1}, . . . Z_{i+m} and a series Z_{i+S}, Z_{i+1+S}, . . . Z_{i+m+S} must be determined. For studying rhythmicity, it is sufficient to examine only the case of linear association:

$$Z_{i+S} = a + bZ_i + \xi_{i+S} \tag{3-13}$$

where ξ_{i+S} is a random value with a mean value of zero.

The same relations are of interest for the Z' series:

$$Z'_{i+S} = a + bZ'_i + \xi_{i+S} \tag{3-14}$$

Tightness, or closeness of the association in this case, may be measured by the correlation coefficient r.

Study of the rhythmicity can be considered a preliminary analysis of the random components. Equations 3-13 and 3-14 are stochastic equations of the Z_{i+S} (or Z'_{i+S}) series over the Z_i (or Z'_i) series. The study of the random component series is provided using the Z' series in region B as an example.

The curve shapes (Figure 3-1) suggests that the left part of the curve is a mirror image of its right part. In other words, there is a linear correlation between random components corresponding to the years $k - i$ and $k + i$, where k is the year with one of the highest random deviations (the boundary between the left and right series) and $i = 1, 2, \ldots n$. The results of analyzing this correlation are illustrated in Table 3-7. The correlation coefficient $r = 0.46$ does not indicate a tight association in this case. One may conclude, therefore, that fading of the random components forward and backward in time from the maximum are somewhat similar. There is evidence that the left and right series are constructed quite differently. For this reason, the rhythmic nature of the left and right series were studied separately. The search for rhythms yielded the result

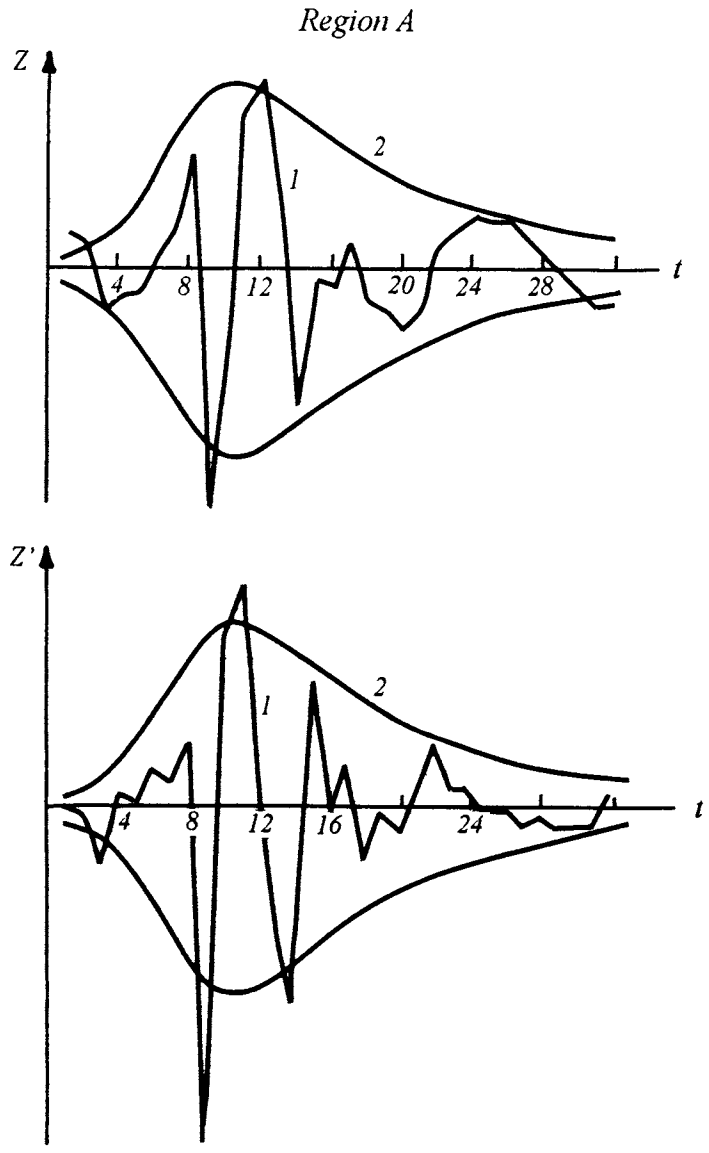


Figure 3-1. Sets of random components $Z = R - P$ and $Z' = \Delta R - \Delta P$ for regions A and B; (1) Fluctuation; (2) Envelope of ΔR

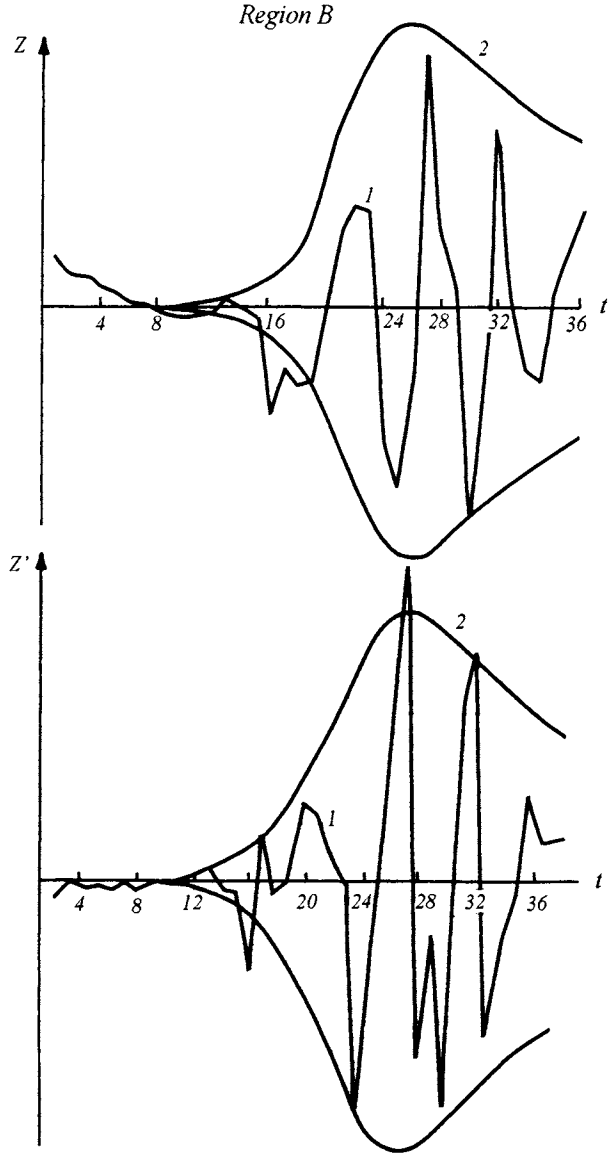


Figure 3-1 ***(continued)***

Table 3–7
Correlation between Members of the Series
$Z' = \Delta R - \Delta P$ (Region B)

Regressive Series	Period S	n	r	a	b	$SS\xi$	$\sigma\xi^2$
Left over right	—	12	0.459	3.4503	0.3911	20,525	1,865.94
Inside the left	8	15	0.8139	0.8607	0.2082	266	19.02
Inside the right	7	7	–0.5156	–2.3697	–0.4529	14,946	2,490.96
	5	11	0.6294	–4.64	0.4463	14,041	1,404.14

represented in Table 3-7. Based on these results, it may be stated with substantial certainty that random deviations in the left series repeat themselves, with the modifications described by Equation 3-14, every 8 years. A high degree of association between these random components is indicated by the correlation coefficient $r = 0.81$. The rhythmic character is somewhat weaker for the right series. In this case, for rhythms with different periods $r = 0.52$ and $r = 0.63$. The main rhythm appears to be 5 years and the 7-year rhythm is secondary. Its presence may be explained by the coincidence of the random components' opposite phases in the middle of the 5-year period (which caused a negative correlation coefficient). Interestingly, the periodicity in the left and right series is different.

Thus, the preliminary study of random component demonstrates some regularity in their behavior. This is expressed as a certain repetitiveness in these components with changes (they may weaken or strengthen) over certain time intervals. These intervals are different during the initial and later stages of the region's development. The change in the repetitive period occurs at the time of the greatest random component amplitude.

Interconnection of Random Components

The rhythmic nature of random components calls for a more general examination of this phenomenon, which takes into account the mechanism of its emergence. It is known that the reasons for periodic phenomena may be internal. In particular, the periodic nature of random deviations may be a result of their interaction and interconnection. Earlier we examined the stochastic models where the correlation of the subsequent accumulated reserve value P_{i+1} with the preceding accumulated reserve value P_i was determined. Here, again, we will be dealing with the stochastic approximation of periodic process models. But this time we

will be interested in models of periodic processes. We will concentrate our efforts on the second-order autoregression processes, which are most pertinent in this respect.

Let us examine a decaying sinusoid:

$$z(t) - Ae^{-ct}\sin(\omega t - \varphi) \tag{3-15}$$

where Ae^{-ct} is the amplitude of fading oscillations; ω is the oscillation frequency; φ is the phase shift; and c is the decay rate.

It is easy to demonstrate that the sinusoid ordinates in three equidistanced positions, t_o, $t_o + \Delta t$, and $t_o + 2\Delta t$, are connected through the following equation:

$$z(t_o + 2\Delta t) = \beta_1 z(t_o + \Delta t) + \beta_2 z(t_o) \tag{3-16}$$

where $\beta_1 = 2e^{-c\Delta t}\cos(\omega\Delta t)$ and $\beta_2 = -e^{-2c\Delta t}$.

After applying indexation to discrete z_i values:

$$z_{i+1} = \beta_1 z_i + \beta_2 z_{i-1} \tag{3-17}$$

If the sinusoid oscillates around the level η, then:

$$z_{i+1} - \eta = \beta_1(z_i - \eta) + \beta_2(z_{i-1} - \eta) \tag{3-18}$$

Equation 3-18 is a discrete equivalent of the following differential equation:

$$\ddot{z} + 2c\dot{z} + \omega_o^2 z = 0 \tag{3-19}$$

The solution of Equation 3-18, is a decaying sinusoid (Equation 3-15). Equations 3-18 and 3-19 describe the functioning of an oscillating system under impact of the initial push and with friction. Parameter ω_o is a parameter of the system's oscillating properties. It is called the natural oscillation frequency of the conservative system. The ω is called the natural oscillation frequency of the dissipative (nonconservative) system. The parameters of Equation 3-19 and those of its solution (Equation 3-15) are related as follows: $\omega^2 = \omega_o^2 - c^2$.

Equation 3-18 shows that when the value of a subsequent element is a linear combination of the two preceding elements, the interconnection of such elements causes a periodic process. β_1 and β_2 coefficients are dependent on the decay rate c, oscillation frequency ω, and time interval Δt. Therefore, this relationship (Equation 3-18) between ordinates is typical for any sinusoid with a given frequency ω decaying at a rate c. When the frequency or decay rate change, β_1 and β_2 coefficients also change.

The inverse problem also may be solved, that is, finding β_1 and β_2 coefficients from interconnection of subsequent values z_{i+1}, z_i, and z_{i-1} and using those coefficients (in particular, at $\Delta t = 1$) to determine the sinusoid frequency:

$$\cos\omega = \frac{\beta_1}{\sqrt{2(-\beta_2)}} \tag{3-20}$$

or its period $T = 2\pi/\omega$ and the decay rate:

$$c = \frac{\ln(-\beta_2)}{2} \tag{3-21}$$

Until now, we were examining differential Equation 3-19 and its discrete equivalent 3-18. Let us now consider the stochastic differential equation and its discrete equivalent (as we have already done earlier) with the additional random term in the equations' right parts. In this case, the stochastic equivalent of Equation 3-18 will be:

$$Z_{i+1} - \eta = \beta_1(z_i - \eta) + \beta_2(z_{i-1} - \eta) + \xi_{i+1} \tag{3-22}$$

where ξ_{i+1} is a random value with a zero mean value (expectation) and variance σ_ξ^2.

Equation 3-22 is indeed the second-order autoregression equation. The second-order autoregression process describes the evolution of a system, of which the status at a given moment in time is determined by its state at the two preceding moments in time. The solution of the difference Equation 3-22 is no longer a sinusoid. In the interval between i and $i + 1$, this equation describes a segment of a decaying sinusoid with a random amplitude A and a random phase φ. Frequency ω and the decay rate c remain constant. Earlier we analyzed how such a description may be connected with an oscillating system (for instance, a pendulum). As a result, Equation 3-22 is describing a distorted periodic function. Instead of the deterministic harmonic curve (Equation 3-15) we now have a stochastic process of a periodic nature. This is the basis we will use when studying random deviations.

It should be mentioned that a second-order stochastic differential equation and the respective second-order autoregression equation describe oscillations with random amplitude and phase only under certain limitations imposed on the equations' parameters. A condition $c < \omega_o$ should be observed for a continuous case and a condition $\beta_1^2 + 4\beta_2 < 0$ for a discrete case. Otherwise, the solution of Equations 3-18 and 3-19 is a combination of decaying exponential curves.

Equation 3-22 is called a second-order autoregression equation because independent variable Z_{i+1} is regressed over its preceding values. Examination of the autoregression process is in effect reduced to regression analysis. A second-order autoregression process equation is a typical regression equation describing the connection of a variable with two other variables. Its customary form (without a random component) is as follows:

$$z = \beta_1 y + \beta_2 x + \beta_3 \tag{3-23}$$

To study a second-order autoregression model, one needs to examine how it fits the observations. In the process, the model's parameters must be evaluated. To identify the model, the paired correlation coefficients r_{zy}, r_{zx}, and r_{xy} and the share of residual variance are used. It is important to note that $r_{xy} = r_{11}$ or $r_{yz} = r_{12}$ are the first terms of the autocorrelation sequence, whereas $r_{zx} = r_2$ is the second term. Their values may be used to determine how the correlation between the random components changes as the distance between them grows (as the original observation series is shifted by the value $\tau = 1, 2$).

Equation 3-23 parameters are estimated from observations based on the following expressions:

$$\beta_1 = \frac{\sigma_z}{\sigma_y} k_1; \quad \beta_2 = \frac{\sigma_z}{\sigma_x} k_2; \quad \beta_3 = \bar{z} - \beta_1 \bar{y} - \beta_2 \bar{x} \tag{3-24}$$

where

$$k_1 = \frac{r_{12} - r_2 r_{11}}{1 - r_{11}^2}; \quad k_2 = \frac{r_2 - r_{12} r_{11}}{1 - r_{11}^2}$$

σ_z, σ_y, and σ_x are the standard deviations of the respective variables and $\bar{z}$, $\bar{y}$, and $\bar{x}$ are the mean values. Correspondingly, in Equation 3-22:

$$\eta = \frac{\beta_3}{1 - (\beta_1 + \beta_2)} \tag{3-25}$$

The part of residual regression not exhausted by autoregression is:

$$D = 1 - (k_1^2 + k_2^2 + 2 r_{11} k_1 k_2) \tag{3-26}$$

In order to check the supposition that the behavior of random components may be related to the second-order autoregression processes, the Z' series in both regions A and B were analyzed. As a result of the preliminary

examination, there were some reservations about whether the random deviations behave similarly at the initial and subsequent stages. In order to clarify these, in addition to the complete series, the left and right series were analyzed separately. As before, the boundary between them was set in the area of maximum random deviations. The sequential order of the series terms was assumed different. For the complete and right series, the dependence of subsequent observations on the preceding observations was evaluated. This reflects the natural flow of time. In the left series, as earlier, the sequential order of observations was backwards in time. There, Z'_{i+1} observation in time precedes Z'_i observation. The reason for this is that in the left series random components increase in time, whereas the autoregression models describe decaying fluctuations. If the movement is backwards in time, the fluctuations are decaying. The results are presented in Table 3-8.

Region A. The complete Z' series is weakly related to the terms differing by one year ($r_{11} \approx r_{12} \approx -0.13$). For any practical purpose, the nearest random deviations may be considered as having no connection between them. The connection between the random components separated by a two-year interval is noticeably stronger ($r_2 = -0.37$). It is important to note that the connection is negative. β_1 and β_2 coefficients are also negative. They satisfy a relation $\beta_1^2 + 4\beta_2 < 0$. Therefore, the autoregression equation is describing an oscillating process (i.e., random components display a pseudo-periodic behavior). If the corresponding connections were deterministic rather than randomly distorted (i.e., if there was a deterministic harmonic curve instead of a stochastic one), then the period would be $T = 3.67$ years. This value is a significant characteristic of the Z' series and shows how rapidly random components change from positive to negative values. The share of the residual variance is $D = 0.82$. This means that the second-order autoregression process (a periodic nature of random components caused by the respective connections between them) is responsible only for 18% of the Z' series variance.

The structure of the right Z' series is similar to that of the complete series. The connections between the terms are about the same for a period of 1 to 2 years; β_1 and β_2 coefficients have approximately the same values and also satisfy the relation $\beta_1^2 + 4\beta_2 < 0$. This, again, is an indication that the periodic function with a period T reflects random components, and the period is practically indistinguishable from the T value of the complete series. The difference is a higher correlation coefficient r_2 and, as a consequence, a lower share of the residual variance D. Here, the second-order autoregression process exhausts about 25% of the variance. This means that the right series Z' was, to a considerable degree, generated by this process.

Table 3-8
Second-Order Autoregression (Series $Z' = \Delta R - \Delta P$)

Region	Series	n	r_{11}/r_{12}	r_2	β_1	β_2	η	T	D	$SS\xi$	$\sigma\xi^2$
A	Complete	30	–0.1267/–0.1270	–0.3741	–0.1773	–0.3968	–0.55	3.67	0.8291	18,388	634.07
	Right	23	–0.1400/–0.0697	–0.4784	–0.1057	–0.3742	1.34	3.79	0.7521	9,207	418.51
	Left	6	–0.2998/–0.1264	0.387	0.01077	0.2912	–2.81	—	0.8501	359	71.87
B	Complete	35	0.0538/0.0547	–0.3401	0.073	–0.34	0.77	4.17	0.88	44,110	1,297.35
	Right	12	–0.0729/–0.1000	–0.4097	–0.1295	–0.4086	4.45	3.76	0.8152	29,201	2,654.64
	Left	22	0.3941/0.3421	0.1805	0.1562	0.0254	4.01	—	0.8805	2,520	119.98

All characteristics of the left Z' series are different from the previous two series. The first terms of the autocorrelation sequence, r_{11} and r_{12}, are positive, and their absolute values are greater than those for the complete and right series. The second term r_2 is also positive, which is also different from the above series where $r_2 < 0$. β_1 and β_2 coefficients are also positive. They no longer satisfy the relation $\beta_1^2 + 4\beta_2 < 0$. This means that the random components of this series do not change in a periodic manner. Their behavior is described by a combination of distorted decaying exponential curves. The residual variance share not exhausted by the second-order autoregression model is the highest ($D = 0.85$) here.

Therefore, the structure of the left series is different from that of the right series. Random components occurring at the initial stage of the reserve accrual are different in their behavior from random components occurring after the initial stage has ended. In other words, random components on the ascending branch of the ΔR deterministic curve behave differently from those on the descending branch.

The right series dominates the left one. For this reason, the structure of the complete series is similar to that of its right part.

Region B. The r_{11} and r_{12} values for the complete Z' series in this case are almost equal to zero, that is, there is no connection between the adjacent random deviations. The autocorrelation sequence second term $r_2 = -0.34$ indicates a negative connection between the random components separated by a two-year period (when one increases, the others decrease). The β_1 coefficient is positive; however, this is not significant because it is very close to zero and its deviation from zero in any direction is simply a matter of chance. The β_2 coefficient is negative and is considerably different from zero ($\beta_2 = -0.34$). β_1 and β_2 coefficients satisfy the relation $\beta_1^2 + 4\beta_2 < 0$ (i.e., random components perform a disturbed periodic motion). The oscillation period is $T = 4.17$. The share of residual variance $D = 0.88$ indicates that the complete series Z' is generated, to a small degree, by a second-order autoregression process. This process is responsible for approximately 12% of the complete series Z' variance.

The right series Z' has a similar structure. The first terms of the correlation sequence r_{11} and r_{12} are almost equal to zero. The second term is negative and its absolute value is somewhat greater than that for the complete series ($r_2 = -0.41$). Random components assume the characteristics of a decaying sinusoid with random phase and amplitude and with period $T = 3.76$. This is indicated by the fact that β_1 and β_2 coefficients satisfy a relation $\beta_1^2 + 4\beta_2 < 0$. The share of the residual variance for this series is somewhat lower than that of the complete series. The distinctions between the right and complete series do not appear significant. Besides a slightly different D, r_2 is slightly higher, β_1 is negative (but also close to zero), and T is slightly lower.

The structure of the left series Z' is considerably different from the previous two, which was also true in region A. The connection between the series of random deviations separated from each other by one or two years is positive in this case. In addition, r_{11} and r_{12} are substantially different from zero, with the absolute value of r_2 being smaller than those of r_{11} and r_{12} ($r_{11} = 0.39$; $r_{12} = 0.34$; $r_2 = 0.18$). β_1 and β_2 coefficients are positive (with $\beta_2 \approx 0$) and do not satisfy the relation $\beta_1^2 + 4\beta_2 < 0$. This indicates that this Z' series does not have a pseudo-periodic structure, but is composed of a combination of distorted decaying exponential curves. The residual variance D share is close to that of the complete series and larger than that of the right series. For this reason, only a small portion of these series variance may be explained due to the second-order autoregression process.

Thus, fluctuations in the ascending branch of the deterministic ΔR curve in region B are different from those of the descending branch. Here as well, the right series is dominant over the left one (although its observations comprise only one-third of the complete series observations as opposed to almost five-sixths for region A).

The following is a comparison of the results obtained in regions A and B. The structure of the Z' series is identical for both regions. Inasmuch as the deterministic components of the accumulated reserves in both regions were different, different models had to be constructed for their descriptions. The same could not be said about the behavior of random components. The similarity of the Z' series is multi-faceted. First, the left and right parts differ from one another for both regions. The right series is represented by a distorted periodic function with the same period in the both regions. The share of residual variance not accounted for by the second-order autoregression is very small for random components of the right series compared with the complete and left series for both regions. The left series' random components in both regions are not pseudo-periodic, but are described by a combination of decaying exponential curves. In both regions, the right series dominates the left one. As a consequence, the complete series is constructed similar to the right series. In both regions, the second term of the autocorrelation sequence in the right series is greater in its absolute value than the same term in the complete series.

The differences between the Z' series in regions A and B are insignificant. The largest difference is that the second-order autoregression process affected the random components in region A more than it did in region B (the D value for all series is smaller in region A than in the corresponding series for region B). Another difference is that the first term of the autocorrelation sequence (r_{11} and r_{12}) and the β_1 coefficient for the complete series in region B are almost equal to zero, which is not true in region A. There are also some differences between the left series. In

region A, the connection between observations is stronger at the two-year shift than at the 1-year shift ($r_2 > r_{11}$, $r_2 > r_{12}$); the situation is reversed in region B ($r_2 < r_{11}$, $r_2 < r_{12}$). In region A, $\beta_1 \approx 0$, $\beta_2 > 0$, whereas in region B, $\beta_1 > 0$ and $\beta_2 \approx 0$. These facts, however, are not reflected by any significant differences in the structure of the Z' series for both regions.

Modulation of Random Components and the Causes of their Regular Behavior

Differences in the random components' behavior for the initial and subsequent periods (the left and right Z' series) must be explained. To achieve this, a different method of analyzing the random components (based on the modulation phenomenon) was applied.

Modulation is the regular change (exhibiting a pattern) in values through time, which characterize a regular process under certain external influences. Besides random components Z and Z', Figure 3-1 shows $\Delta R(t)$ curves or, rather, the curves with ordinates proportionate to ΔR_i (i.e., $\Delta R(t)$ with a precision to within the constant). The ΔR_i values obtained from the $R(t)$ function corresponding to Model 8 in region A and Model 9 in region B, were used as $\Delta R(t)$. As noted before, these models were shown because they best described the reserve accumulation process. Figure 3-1 shows that the $\Delta R(t)$ curve is an envelope curve for the reserve accumulation process. During the i years when ΔR values are small, random components are also small. Conversely, at times corresponding to maximum ΔR values, random components increase and also reach their maxima. This random component behavior is typical for Z as well as Z'. It is interesting that in region B, where ΔR maximum values are higher, the random component amplitudes are also larger.

It is important that the value of ultimate potential resources in region B is greater than that in region A. This value determines the maximum of reserve accrual and random components. It may be concluded, therefore, that the random component values change in time according to a pattern given in time by the ΔR value. Because this is similar to the modulation of oscillations, in particular to its amplitude, it may be concluded that the random components under study are modulated. The $\Delta R(t)$ function is a modulating function that sets the pattern for the random component change in time. Likewise, the term of original (unmodulated) random components may be introduced.

Thus, in the case of modulating the size of original random components by a function $\Delta R(t)$, a modulated random component may be presented as:

$$Z_{i+1} = \varepsilon_{i+1}\Delta R$$
$$Z'_{i+1} = \varepsilon'_{i+1}\Delta R \qquad (3\text{-}27)$$

where ε_{i+1} and ε'_{i+1} are original random components.

As we can see, modulation is proportional to ΔR. The relationship between Z_{i+1} and ΔR cannot be linear. In such a case, if ΔR tended to zero, the Z (or Z') value would tend to be some constant other than zero. In other words, modulated random components would remain at a certain constant level even if the reserve accumulation process had not begun, or had already ended. This situation is highly unlikely. When $\Delta R = 0$, Z (or Z') must be equal to zero. Equation 3-27 corresponds to this condition.

Therefore, there is a certain connection between random components and the reserve accrual. The reserve accrual controls random deviations or, conversely, changes in random components. Their change from one level to the next causes determinism in the reserve accrual.

This interrelationship needs explaining because, generally speaking, it is not obvious. In fact, it would be reasonable to expect just the opposite (i.e., a decrease in the random deviations around the maximum ΔR values). Remember, the reserve accumulation evolution is a reflection of the exploration evolution in the region. The ΔR curve can be subdivided into three segments: prior to maximum accruals, maximum accruals, and following maximum accruals. Depending on the type of ΔR curve, exploration evolution in a region is accordingly subdivided into three respective periods.

Relatively small oil and gas discoveries and relatively significant expenditures in effort and funds are typical for the initial period of exploration. The reason for this is poor knowledge of the field distribution pattern. Knowledge of the region in area and depth increases during the intermediate period because there is an opportunity to select the best exploration plays, which results in large discoveries. During the late period, the region is relatively mature, the size of discoveries declines, and large discoveries are unlikely. At the same time, the reserve discovery becomes more difficult and their appraisal increases due to greater exploration depths and more complex targets. It is obvious that the stability of the exploration process and its results must change from one period to the next. Random components are expressions of this aspect of exploration.

In this case, one could expect that further exploration of the region, and the resulting possibility of selecting the most efficient approach and best exploration plays, would stabilize the achieved results to a great degree. In reality, the opposite occurs, and the contrasting results are most obvious during this intermediate period. This occurs because, during the

period of large discoveries, the contrast is greatest between the new large discoveries and the fields currently being appraised. The appraisal targets are not only the large fields but also medium and small fields. Thus, the intermediate period is the time of greatest diversity of fields in terms of reserve size. During the initial and late periods the discoveries are more uniform and mostly small, therefore, random components are also more uniform. Equally responsible for the increase in random deviation during the intermediate period is the contrast in decision making. Large discoveries cause changes in estimates. Over-optimistic and over-pessimistic estimates may alternate. It is believed that the nature of control decisions becomes most unstable during this tempestuous period—the amount of exploration fluctuates, the drilling footage in various areas and on various structures fluctuates, and the proportions of exploratory and appraisal drilling change. Thus, the strategy, changing with the exploration stages, leads to respective changes in control actions.

For example, T. D. Vesnina who analyzed data on a number of oil-gas provinces (Yakutia, Tatarstan, Bashkortostan, Chechen-Ingush region, Kuybyshev, Tyumen regions, Krasnodar Kray, Azerbayjan, etc.) indicates the following actual trends in the change of proportions of exploratory and appraisal drilling [7]. During the initial exploration stage, exploratory drilling in oil-gas regions is 6 to 8 times greater than appraisal drilling. At the intermediate stage associated with the appraisal of the largest fields, the appraisal drilling footage grows 10 to 15 times, whereas exploratory drilling declines 3 to 4 times. At the late stage, a great number of small structures are being explored, so the amount of exploration work increases again by a factor of 3 to 4. Due to the need to discover a much larger number of small fields, the exploratory drilling footage relative to the appraisal footage increases 5 to 6 times.

In the former USSR, exploration efforts declined immediately following the first commercial discoveries. Later, this resulted in a conflicting situation with the fulfillment of the commercial reserve accrual plan. This leads to a peculiar "swing" effect: either the exploration increases together with a decline in the appraisal effort, or the appraisal efforts increase and the exploration work declines. This causes an instability in the exploration process that plays an important role in increasing the random component amplitude. Random components behave likewise: they "swing."

Thus, modulation of the initial random components using the $\Delta R(t)$ modeling function helps explain the behavior of random components, in particular their features discussed earlier. The $\Delta R(t)$ curve is asymmetric, thus, it is not surprising that the left and right series of the random component are different.

Finally, it is necessary to determine whether the initial (unmodulated) random components are purely stochastic elements or if there are some

connections between them. In other words, are they generated by a random process (white noise) or by a process with memory and with afteraction (a type of second-order regression process)? If the latter is true, then what are the relations between randomness and correlation in the emergence of the random components? In this case, the randomness can be considered a result of miscellaneous, poorly controlled factors that are difficult to foresee and to account for in the process of exploration. We have already discussed the importance of the non-singularity of decisions and the great degree of freedom while making control decisions.

As expression 3-27 shows, unmodulated random components form two series, the ε series and the ε′ series:

$$\varepsilon_{i+1} = \frac{Z_{i+1}}{\Delta R} = \frac{R_{i+1} - P_{i+1}}{\Delta R}$$

$$\varepsilon'_{i+1} = \frac{Z'_{i+1}}{\Delta R} = \frac{\Delta R - \Delta P}{\Delta R} \qquad (3\text{-}28)$$

We evaluated the possibility of describing these series using a second-order autoregression model, which may be presented as follows:

$$\varepsilon_{i+1} - \eta = \beta_1(\varepsilon_i - \eta) + \beta_2(\varepsilon_{i-1} - \eta) + \xi_{i+1}$$

$$\varepsilon'_{i+1} - \eta = \beta_1(\varepsilon'_i - \eta) + \beta_2(\varepsilon_{i-1} - \eta) + \xi_{i+1} \qquad (3\text{-}29)$$

The results of this evaluation are listed in Table 3-9.

In region A, terms of this series are closely connected for a one-year offset ($r_{11} = r_{12} = 0.43$). When the distance between the series members equals two years, the connection between them weakens and becomes negative ($r_2 = -0.13$). The β_1 and β_2 parameters satisfy the relation $\beta_1^2 + 4\beta_2 < 0$. This is an indication of the pseudo-periodic structure of the series. The random deviations' behavior is described by a decaying sinusoid with a random amplitude and a period of 6.05. The second-order autoregression model exhausts about 32% of the observed variance (the share of the residual variance $D = 0.68$). This is higher than earlier results. Thus, the random components of series ε were formed under significant influence of the second-order autoregression. Their emergence was not totally random: it was one-third stochastically predetermined by the two preceding random components. The "swing" just mentioned has a period of about 6 years.

Incomplete series ε and ε′ were studied in region B. The first nine terms have been discarded because ΔR values for the first nine years are

Table 3-9
Second-Order Autoregression [Series $(R_{i1}\ P_{i1})/R$]

Region	n	r_{11}/r_{12}	r_2	β_1	β_2	η	T	D	$SS\xi$	σ_ξ^2
A	30	0.4379/0.4344	–0.1307	0.6188	–0.3722	–0.038	6.05	0.6839	5.75	0.1983
B	27	0.4652/0.4242	0.0205	0.5195	–0.2077	–0.1496	6.52	0.7802	8.43	0.3242

less than 1, and the ε and ε' values ($i = 1, 2, \ldots 9$) for this time interval are determined with significant distortions from Equation 3-28.

The random components' behavior in this region is similar. Almost all analyzed parameters in regions A and B for the ε series, especially the major parameters of the series, are identical. In region B, this series also has an oscillating nature with a period $T = 6.52$, which, for evaluation purposes, is the same as $T = 6.05$ for ε in region A.

Values of the ε series in region B are not as closely approximated by the second-order regression model—the residual variance share is $D = 0.78$. The series was formed under a considerable degree of random influence; its generation by the second-order autoregression process is much weaker. In other words, the connection between the observations are not as close. Table 3-9 indicates the cause for this is r_2. There is no significant connection in region B between the observations separated by two years. The value of ε_{i+1} is stochastically predetermined not by the two preceding random deviations ε_i and ε_{i-1} but, rather, by only one, ε_i. That is, any "memory" of consequence covers only one year. With equal success, the ε series in region B may be described by a first-order autoregression model.

The results of the ε' series evaluation are listed in Table 3-10. The residual variance value is close to 1. This indicates that the terms of the ε' series in both regions A and B are random values with no correlation between them. They are mutually independent and have similar distributions. These series are generated by a totally random process (white noise). This is also supported by the low values for the first two members of the autocorrelation sequence; with the exception of r_{12} and r_2 in the ε' series for region A, they are close to zero.

This is not an unexpected result. It is based on the aforementioned features of the ε series. From Equations 3-11 and 3-12:

$$Z'_{i+1} = \Delta R - \Delta P = (R_{i+1} - R_i) - (P_{i+1} - P_i) = Z_{i+1} - Z_i \tag{3-30}$$

Therefore,

$$\varepsilon' = \frac{Z_{i+1}}{\Delta R} - \frac{\Delta R^*}{\Delta R}\frac{Z_i}{\Delta R^*} = \varepsilon_{i+1} - \frac{\Delta R^*}{\Delta R}\varepsilon_i \tag{3-31}$$

where $\Delta R^* = R_i - P_i$.

In the first approximation,

$$\varepsilon' \approx \varepsilon_{i+1} - \varepsilon_i \tag{3-32}$$

When a series is discretely assigned, taking the first differences is analogous to differentiating a continuous function. Indeed, if we switch

Table 3-10
Second-Order Autoregression [Series (*R P*)/*R*]

Region	*n*	r_{11}/r_{12}	r_2	β_1	β_2	η	*T*	*D*	$SS\xi$	$\sigma\xi^2$
A	30	–0.0100/–0.176	–0.18	–0.0195	–0.1816	0.068	3.94	0.9668	7.88	0.2717
B	27	–0.0617/–0.0658	–0.0851	–0.0714	0	–0.0071	3.71	0.9877	10.54	0.3906

from a discrete series ε_{i+1} to its differences $\Delta\varepsilon = \varepsilon_{i+1} - \varepsilon_i$, we come up with the rate of change for the sequence ε_{i+1}. Both operations eliminate the linear trend. In this case, a natural result of this operation is the disappearance of correlation between the adjacent terms of the series. If there is a spectrum $S(\omega)$ corresponding to a function $y(t)$, and if the frequency harmonics composing this function are known, then the derivative $y'(t)$ has a corresponding spectrum $\omega^2 S(\omega)$. This means that the lower frequency harmonic amplitude in the spectrum will decrease considerably after differentiation because for lower frequency components $\omega << 1$. This is why we can consider the ε' series as a residual series obtained from the ε series by subtracting low-frequency, long-period components. The structure of this series is formed mainly of short- and medium-period components (it is not surprising that the T value decreased by almost a factor of 2). In this series, the effects caused by short- and medium-period harmonics are more substantial (these harmonics were masked in the ε series by long-period components). This is the reason for the increased correlation coefficient r_2 in the ε' series compared to the ε series (region A).

Thus, we have found the patterns of change for the random components. The initial random deviations do not emerge as purely random (only random components ε_{i+1} are of interest in this regard). On the contrary, their emergence, to a significant extent, is dependent on the random components for the preceding two years; that is, they are generated by the second-order autoregression process. This is why they display a complex, distorted periodic behavior. Their oscillation period is about 6 years. This time duration between the rise and fall of random components is valid for both regions. Having emerged in the manner described, the random components are modulated by the $\Delta R(t)$ function. They strengthen or weaken as this function strengthens or weakens at different stages of the regional resource development. In other words, modulation of random components is connected to the rate of reserve accumulation. Random components strengthen when the reserve accrual accelerates. Such interpretation of the random deviations rise and evolution helps explain most of their variability. Once the deterministic component was removed, the random components variance was 450 in region A (see Table 2-3) and 913 in region B (see Table 2-5). By contrast, we found that the variance of a purely random process is just 0.20 for region A and 0.32 for region B (see Table 3-9).

As previously illustrated, the exploration process was historically different in regions A and B. As a result, the reserve accumulation occurred differently and is described by different models. Together, they reflect the different changes in the target selection policy and selection conditions. These differences in the reserve accumulation process are natural considering the different histories and strategies for the regions.

With all these differences, the purely random components still behave in an almost identical manner. Unmodulated random components, devoid of the influence by the reserve accrual $\Delta R(t)$ process, do not appear to be affected by the differences just mentioned, and in particular, the differences in intrinsic factors of nature. Most likely, they reflect some factors of control (e.g., psychological factors, future planning based on the already achieved levels, a certain degree of risk-taking when making decisions, etc.). These factors are not determined by the evolution of exploration but, rather, by the general approach to exploration or by some sort of psychology in its conduct. This is similar in all management decisions concerning exploration strategies. This may be the key to the common mechanics of random component emergence during reserve discovery and appraisal in these regions.

Patterns in the Evolution of the Reserve Accumulation Process

Chapters 2 and 3 show that a deterministic description of the reserve accumulation process using the $R(t)$ function is insufficient. The reserve accumulation evolution is significantly affected by random components that emerge, to some extent, depending on random events. Thus, the final formal description of the process under study is not a single model but, rather, a system of models reflecting the behavior of both deterministic and stochastic elements. The system can be expressed as follows:

$$P_{i+1} = R_{i+1} + Z_{i+1}$$

$$R_{i+1} = f(R_i)$$

$$Z_{i+1} = \varepsilon_{i+1}\Delta R$$

$$\varepsilon_{i+1} - \eta = \beta_1(\varepsilon_i - \eta) + \beta_2(\varepsilon_{i-1} - \eta) + \xi_{i+1} \qquad (3\text{-}33)$$

The main effort was directed toward finding the form of equation $R_{i+1} = f(R_i)$ in the above system (3-33). This form determines the change in time of the accumulated reserves, or $R(t)$. It is not coincidental that the above equation was used to determine the form of the $R(t)$ function. Dependence of the R_{i+1} reserves accumulated by the following year on the R_i reserves accumulated by the preceding year is determined by the very essence of exploration. Exploration is directed by the results achieved or, rather, by information contained in these results. The result is that exploration is basically an information process. The information used to

move exploration forward is unknown in the beginning and is only obtained through the process of exploration. Discoveries, which affect increases in exploration, funding, and drilling volumes and lead to the reserve accrual, reflect the dependent nature of each result. How the reserve accumulation evolves depends on the discovery results (fields of different sizes); that is, on the region's resource development history.

The discovery history inherently depends on the exploration strategy in a region. The history-dependent nature of the process is one where it is not possible to define a single universal function $f(R_i)$ that would describe the reserve accumulation evolution in any region. The $f(R_i)$ functions, however, may be similar for regions with similar resource development histories. We analyzed regions with different, and in some ways typical, exploration evolution and discovery sequences. For this reason, we believe that the equations we found:

$$R_{i+1} = \frac{R_i}{1 - (e^{a-bR_i} - c)} \text{ (region A)}$$

$$R_{i+1} = \frac{R_i}{1 - (e^{a-bR_i} - c/R_i)} \text{ (region B)} \qquad (3\text{-}34)$$

are also typical and may be applied in other regions with relatively similar exploration histories. The study results lead us to conclude that even if the $R_{i+1} = f(R_i)$ equations for all regions are not totally and exclusively described by Equation 3-34, they are still within the same class of functions, such as the following equation:

$$R_{i+1} = \frac{R_i}{1 - (e^{a-bR_i} - n_i)} \qquad (3\text{-}35)$$

where n_i = const, or $n_i = f(R_i)$.

Apparently, this class of functions reflects some typical "set of exploration histories" for different regions, and these histories affect the oil and gas reserve accumulation evolution. The important factor, again, is that the large and giant fields in any region contain a substantial portion of the ultimate potential resources.

A drastic non-uniformity in the structure of the ultimate potential resources is the factor predetermining a relatively limited diversity among regions in terms of their exploration history.

Equations that provide a description of the deterministic reserve accumulation process are usually called "evolution equations." These equations characterize the adaptation (adjustment) mechanisms, which play an important role in evolution processes.

Another type of mechanism, called a bifurcation, has a totally different nature. For this mechanism, the process result is unpredictable in principle because at certain times the state of stable equilibrium becomes destabilized. As a result, the evolving system tends to find a new state of equilibrium. The choice of this new state of equilibrium, however, is predetermined by the random state of the system at the moment it loses its stability (because new possible states of stability are close to each other). Instabilities occur as a result of spontaneous fluctuations generated by the system itself. In the process of evolution, these instabilities tend to be small and are easily extinguished. The large fluctuations, however, occur in the proximity of an unstable state. When stability is lost, the process of evolution jumps to a different state. As a result, the system changes its status, and it does so in an unpredictable manner.

Clearly, every new state of the reserve accumulation process occurs as a result of a reserve accrual fluctuation. Accordingly, these new states later change to yet another temporarily stable state. Thus, it is apparent that the trend in the reserve accrual occurs through fluctuations that direct the system of control (i.e., changes in concepts, therefore, changes in controlling actions and resource control). The reserve accrual goes through a number of unstable states caused by fluctuations and, through subsequent deterministic evolution, arrives at a new stable state. In other words, deterministic patterns are created through fluctuations, after fluctuations, and as a response to fluctuations. This demonstrates a unity, or interdependence, between the random and the regular component rather than a contradiction between them.

A more detailed description of this aspect of the process of discovery and appraisal of reserves is presented in Chapter 5.

Notes

1. R_{lim} value is given for $\Delta R_{\text{lim}} = 1$.
2. Oil and gas fields are classified in Russia as small, medium, large, giant, and unique.

CHAPTER 4

Evolution of the Reserve Accumulation as a Function of Drilling Volumes: Deterministic Models

The important aspects of the process of discovery and appraisal of reserves, such as the oil and gas reserve accumulation evolution, can be viewed from another angle. This method is also based on modeling for the purpose of resource forecasting, long-term planning of the reserve accrual, and determining the evolution of the exploration process. Evolutionary models of development are commonly used for this purpose; however, the crucial connection is between drilling footage and the reserve accrual. This relationship is used in this chapter as the basis for the modeling.

Modeling of the oil and gas reserve discovery and appraisal process, within the constraints of evolutionary models of the resource development (historic curves or development curves) is usually reduced to *a priori* assignment (fitting) of a function $R(L)$. This describes the accumulated reserves R with sufficient approximation in relation to the total drilling footage L since the beginning of exploration. As a rule, this function $R(L)$ is not inferred based on concepts of the oil and gas resource development environment; it is simply assigned as an axiom. Its selection is predetermined by some characteristic features of the development curve (historic curve), such as the inflexion point and the asymptote. The study usually includes only evaluating the model parameter. The adequacy of the model to the actual data is usually not examined.

Various functions similar to the appraised reserve accumulation curves were proposed for use as the $R(L)$ function. Among them are the logistic function, the Gomperz function, the modified exponential (base e) function, and so forth [1, 3, 6, 26]. It is important to note that these publications

do not offer alternative models. Apparently, each author believes that their function is universal enough to describe the process under consideration in different environments of its implementation. Such an approach to modeling has been justly criticized [30].

Construction of Models

The approach to modeling in this book is different. We do not assign an axiomatic model but, rather, infer it inductively from a system of initial concepts. These concepts relate to the nature of the exploration process and that of factors affecting the relationship between drilling footage and accumulation of reserves. They also make this relationship the focus in forecasting the extent of exploration and reserve accrual. Naturally, they are initially translated into the language of mathematics and were specified earlier. The following briefly reiterates the concepts we will need for the subject of Chapter 4.

1. The exploration process is a process of transferring the resources into the appraised reserves. This is a controlled process that is realized in a complex controlled system. The main variables of this system are the reserve and drilling footage. These and other variables are mutually connected due to certain control actions undertaken for specific purposes, which are based on certain geologic concepts and methodological principles of exploration and appraisal. If the process is not realized in a controlled system, such a connection might not occur. More specifically, the connection between the accumulated reserve and drilling footage is a result of the existence of feedback—that is, discoveries and appraisal (i.e., accrual of reserves) cause an increase in drilling footage, which, in turn, leads to new discoveries. The equilibrium between them is not coincidental but is a result of the feedback shaping the control policy.
2. The control results are affected by numerous factors, some natural, some man-made (e.g., diversity in goals, concepts and means of control, completeness of information, etc.). It is difficult to take all factors into account. For this reason a purposeful action (control) may not lead to a unique result. Therefore, the connection between the factors is stochastic. The deterministic component describing the trends in this connection, which are objective and long-term in nature, is determined by the resource development strategy (i.e., to what extent the resource development control fits the natural conditions). This means that it is impossible to construct

a single universal model that would describe the studied process with equal success in any region.

3. Modeling of the reserve discovery and appraisal process as a function of reserves from exploratory drilling is not associated with the construction of control models. These are different types of models. They describe the relations between the elements of the system without trying to change these relations. Despite statements to the contrary, these models cannot be used for active planning of the accrual of reserves. These are response-type models that reflect the response (extent of drilling) to the result (accrual of reserves) and vice versa. In this sense these models are adaptation (adjustment) models of the oil and gas resources development process.

The following is a more detailed description of these concepts at a level sufficient for model construction.

The speed of the appraised reserve accumulation R is expressed as the derivative dR/dt and is a change in R within a small time interval. Similarly, the speed of increase in drilling footage is dL/dt. The relative speeds of the reserve accumulation and of the increase in drilling footage may be represented as $(1/R)(dR/dt)$ and $(1/L)(dL/dt)$. The relative speed of change of a given parameter may be called its rate of change.[1] According to the concept described above, within the evolving system of exploration and reserve accrual, the ratio of their growth rates is an all-important characteristic of the process:

$$\frac{(1/R)(dR/dt)}{(1/L)(dL/dt)} = k \tag{4-1}$$

When $k > 1$, the rate of accumulation of reserves is greater than the rate of increase in the extent of drilling. When $k < 1$, the situation is reversed. Therefore, the k value may be used as an indicator of how efficiently the reserves are developed in a region (it is, of course, desirable for the reserve accumulation rate to be greater than that of the increase in drilling). Thus, k is a very important parameter.

Equation 4-1 may be re-written as follows:

$$dR/dL = kR/L \tag{4-2}$$

The ratio dR/dL, or, more precisely, the ratio of oil and gas reserve accrual ΔR over a certain time period to the drilling extent ΔL over the same

period, is called "exploration efficiency," "current efficiency," or "specific reserve accrual." The ratio of accumulated reserves to the total drilling footage from the start of exploration is called the "accumulated efficiency." Therefore, k is a proportionality coefficient between the current and accumulated efficiency. The L in Equations 4-1, 4-2, and thereafter, is not the drilling footage from the start of exploration, but from the start of the first reserve accrual.

The function k is not a constant. In accordance with the concepts of the resources development process as explained earlier, k should change with the passage of time. Therefore, $k = k(t)$. The specific form of its change is determined by the particulars of the region's development, the exploration strategy, and the region's geology (initially by the resource's structure, or field size distribution). The $k(t)$ function is a reflection of the adjustment and adaptation of the exploration process to changing geological conditions in the process of the region's development—mainly the average size of discovered fields. This is why k may be called the adaptation (adjustment) function. Naturally, adjustment is connected with the strategy of the region's development so that the $k(t)$ function reflects the exploration strategy. However, it does so after the fact and cannot be used for the development of the strategy, which was discussed earlier in Chapter 3.

As follows from Equation 4-2,

$$dR/R = k(t)dL/L \tag{4-3}$$

We are interested in the dependence R on L. Thus, in order to solve this differential equation, we will switch from $k(t)$ to $k(L)$ (parameter k changes in time, therefore, it changes with increasing L). A general solution of differential Equation 4-3 is the following function:

$$R = \mu \exp\left(\int \frac{1}{L} k(L)dL\right) \tag{4-4}$$

where μ is the equation parameter (a constant).

Thus, we have obtained a function that shows the dependence of the accumulated reserves R on drilling footage L. The specific form of the function depends on the function $k(L)$, which is determined by the reserve discovery and appraisal process through control of the extent of drilling. In order to determine the specific form of $k(L)$, it is necessary to identify the main factors in the exploration process that affect $k(L)$. These factors must be associated specifically with the region's resources development (see Chapter 2) and with the exploration strategy. One such important

factor is the goal: the goal of exploration is to first discover the largest fields that have the greatest economic significance. Such a strategy is subordinated first and foremost to the economic interests. This is why, from the viewpoint of $k(L)$ behavior, the most important factors in the history of a region's development are at what stage of the exploration process this goal is achieved and how it changes with the ongoing process of the region's development. For instance, it is important whether or not the largest undiscovered fields are rapidly found.

In line with this concept, the following typical situations (strategies) may be identified in the transfer of potential resources into appraised reserves (and the respective types of adjustment function $k(L)$ format):

1. Most typical (and most efficient) is the situation where large fields are discovered at early stages of the region's resource development. From a strategic viewpoint this means that the main, most efficient exploration plays have been determined at the early stage, and the main exploration effort has been concentrated on the exploration of these plays. During this period, the reserve accumulation rate is significantly greater than the rate of growth in drilling volume. As large discoveries are exhausted, the rates begin to equalize. At later stages, when mainly small discoveries are made (at a lower efficiency), the increase in the rate of drilling outstrips the reserve accumulation rate, and the amount of outstripping increases with time. This is the same as a progressive decrease in k with the growth of L (i.e., the value of the drilling footage adjustment function $k(L)$ continuously decreases). Exploration rapidly adjusts to geologic conditions.

The adjustment function speed of change may be altered in various ways through the process of exploration. Three different cases can be identified: (a) the speed of change of the adjustment function $k(L)$ at first declines rapidly, then very slowly declines with an increase in L, and eventually does not change significantly (the case of decreasing speed); (b) the speed remains constant; (c) the speed of change of the adjustment function $k(L)$ at first declines slowly, then declines more rapidly with an increase in (L) (the case of increasing speed). The first case is characterized by a convex $k(L)$ curve, the second case by a straight line, and the third case by a concave curve. The causes for the differences in the appearance of the adjustment function are discussed below.

1a) A slowing of the adjustment function decline speed is associated with an exploration strategy that led to the largest discoveries at the earliest stages, with a relatively small drilling effort expended. All large fields were discovered with a small drilling footage. Subsequently, even a substantial increase in drilling was accompanied by a significant decline in the reserve accumulation rate due to the exhaustion of large fields. In general, this is the most successful strategy. It indicates a rapid and correct adjustment of exploration to the specific geologic environment,

where correct concepts of a region's geology and potential were formed at the earliest exploration stages and drilling footage was properly allocated. It is obvious that the sooner the major discoveries are made, (1) the larger their reserves, which depend on the region's ultimate potential (or ultimate potential resources); (2) the smaller the drilling effort expended; and (3) the more drastic the adjustment function decline rate. This decline at the later stages will depend on the structure of undiscovered resources (in terms of field sizes) and on the rational distribution of drilling footage. Their different combinations may lead to different appearances of the $k(L)$ curve: a more rapid $k(L)$ decline with a subsequent slow decline or a smoother declining behavior.

Thus, the adjustment curve is comprised of two parts: a steep branch and a gentle one. Depending on the combination of the conditions stated above, the transition from one part to another will be different. Accordingly, the adjustment function may be approximated by one of the following functions:

$$k = a/\ln L \tag{4-5}$$

$$k = aL^{-b} \tag{4-6}$$

$$k = ae^{-bL} \tag{4-7}$$

The functions in this succession are in order of decreasing contrast in the transition from the steep branch of the curves to the gentle branch (each preceding function is an envelope for the following one in this succession). For instance, it follows from Equations 4-5 and 4-6 that:

$$k' = -[1/(L\ln L)]k \tag{4-8}$$

$$k' = -(b/L)k \tag{4-9}$$

where k' is the first derivation of k.

This illustrates that, over small L segments and for small changes of $\ln L$, function 4-5 can be approximated by function 4-6 with accuracy to within the constant. In turn, it follows from Equation 4-7 that

$$k' = -bk \tag{4-10}$$

By comparing Equations 4-9 and 4-10, we can see that over small segments, when L may be considered almost constant, function 4-6 may be approximated by function 4-7. Therefore, function 4-5 is the envelope of function 4-6, which, in turn, is the envelope of function 4-7.

The rate of change of $k(L)$, that is, k' to within the constant, is successively determined (at a given k) first by $L\ln L$, then by just L, and

eventually no longer depends on the variable L. Consequently, the transition contrast from one branch of the $k(L)$ curve to the other decreases.

1b) If the speed of decline of the adjustment function is constant, then

$$k = a - bL \tag{4-11}$$

This is a degenerated case of the $k(L)$ curve presented above, as if only one steep branch of the curve is represented. The absence of the gentle branch may be due to one of two reasons.

The first reason is a very successful exploration, such that only a small amount of drilling was sufficient to transfer a substantial portion of the ultimate potential into the appraised reserves. This can occur if most of the resources are concentrated in a single giant field discovered at an early stage of exploration. This is why further drilling would not lead to $k(L)$ curve stabilization (slow decline). The curve continues its rapid and steady decline due to the lack of discovery of any significant fields. The $k(L)$ straight line in such a case should have a substantial slope (a relatively large b parameter).

The second reason is related to a totally different aspect of exploration and resource distribution. The gentle branch is not present because it is simply too early; it will show up later, when a significant amount of drilling is performed. This occurs when the large discoveries are somewhat delayed in time due to an unsuccessful selection of exploration plays compared to (1a). This unsuccessful selection may be compounded if large fields are not present in the region. Such a case would cause a very smooth transition from the steep to the gentle branch, to the extent that both branches look like a single straight line.

A straight line may represent the gentle branch of the curve (i.e., a degenerated case). A situation like this occurs when the play selection is even less successful, which substantially delays large discoveries, if present. The slope of such a curve is insignificant. Later, as discussed in the following case, the steep branch may appear.

(1c) An increased speed of decline of the adjustment function is possible in a situation where large fields are discovered after a substantial amount of drilling. These discoveries are preceded by numerous smaller (small and medium) discoveries (i.e., the adjustment process is stretched). Due to a gradual addition of more substantial discoveries, it is possible to maintain a weak change in the rate of the reserve accrual compared with the rate of increase in the extent of drilling. Due to the fact that at the time of large discoveries a significant portion of the medium-size (and small) fields have already been discovered, the reserve accrual rate after the large discoveries drastically drops compared to the rate of increase

in drilling. As a result, the curve is convex. From the viewpoint of the "goal" factor, such a curve represents a less successful exploration strategy (a less successful play selection) than the other curves.

This type of function may be well approximated by the following equation:

$$k = c - aL^b \quad (b > 1) \tag{4-12}$$

(2) A less typical situation exists when the adjustment function does not change with the growth in L and remains constant:

$$k = a \tag{4-13}$$

This means that the rate of the reserve accrual is proportional to the rate of the increase in drilling. This would seem natural (as in living organisms, proportionality is maintained between the growth of different organs) if discovered fields were not disparate in their size. In the process of oil and gas reserve discovery and appraisal, maintaining constant ratios between the elements of the system under consideration is only possible when the structure of the ultimate potential is poorly differentiated. In other words, this can occur only if large discoveries are made at a relatively late stage in the resource's development, provided that the ultimate potential and drilling footage are relatively small. The development strategy is inferior compared to the previous cases. The most important targets are being explored late, after a substantial amount of drilling elsewhere.

(3) Publications analyzing the oil and gas reserve discovery and appraisal process emphasize that large discoveries are usually made at the initial stages of exploration, medium discoveries at a later stage, and small discoveries are made both simultaneously with large and medium discoveries, and also later. Evidently, this is caused by a certain exploration strategy.

It is possible, however, that in reality a different situation exists. There are regions where exploration was long conducted in low-potential areas with no large fields. This was done not because of erroneous selection of exploration plays, but because it was impossible to develop major exploration plays or those at great depths. Only later, after extensive drilling in the low-potential areas without large or medium fields, was exploration concentrated in the best areas. At the initial development stages in such regions, the increase in the drilling rate was greater than the rate of reserve accumulation. Only when the major plays began developing, did the ratio change—the rate of increase of reserve accumulation became higher than the rate of increase in drilling. As a result, the adjustment function $k(L)$ in such regions grows, reaches its maximum (associated with the most significant discoveries), and then begins to

decline. Adjustment to geologic conditions in these regions is slow and is accompanied by extensive drilling. The adjustment function of this kind may be approximated by the following equations:

$$k = ae^{-bL}L \tag{4-14}$$

$$k = (a - bL)L \tag{4-15}$$

When the adjustment function, despite extensive drilling, has not yet reached its maximum (which may be associated with a substantial delay in large discoveries), the ascending branch of the adjustment function may be approximated by the following curve:

$$k = a + bL \tag{4-16}$$

Thus, we believe that the connection between the reserve accumulation and drilling footage is the result of a certain type of exploration policy. Because we consider this policy a system of strategic decision making in the reserve discovery and appraisal within an entire region, we will discuss the exploration strategy. To characterize the strategy, we use the most important features: how fast (in terms of time and drilling expenditures) large discoveries are made, and how numerous were the preceding small and medium discoveries. The following pertain to the entire exploration system in a region: the selection of exploration plays; distribution of the extent of exploration among the plays; concentration of exploration along the major and subordinated plays; quality of geologic forecast; subordination of exploration to the goal of developing the whole region or its separate parts, and so forth. Quantitatively, strategy is characterized through the respective form of the adjustment function $k(L)$. The adjustment functions we derived form a series reflecting the deterioration in development strategies. A sequential evolution of the adjustment function form, as the large discoveries are either delayed or increased, is shown in Figure 4-1. A transition from one function to the next as strategies change is clearly illustrated. A concave $k(L)$ curve progressively (1) flattens, (2) becomes a straight line, (3) becomes convex, (4) becomes horizontal, and (5) acquires a "hump." If the "hump" is not reached yet, the curve may look like an ascending straight line.

As reflected in Equation 4-1, all adjustment function values must be positive. The form of the proposed adjustment functions 4-5 through 4-7, Equation 4-14, and Figure 4-1 indicates that as L increases, the curves asymptotically approach the X axis; therefore, these functions are always positive. Functions 4-11, 4-12, and 4-15 are exceptions. When L is greater than a/b, $(c/a)^{1/b}$, and a/b, the functions become negative. Thus, we can

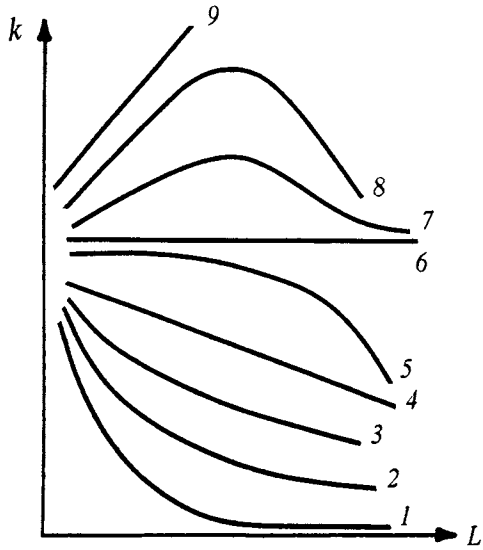


Figure 4-1. Directional evolution of the adaptation function $k(L)$ in connection with sequential differences in the oil and gas exploration strategy. 1. $k = a/\ln L$; 2. $k = aL^{-b}$; 3. $k = ae^{-bL}$; 4. $k = a - bL$; 5. $k = (b > 1)$; 6. $k = a$; 7. $k = ae^{-bL}L$; 8. $k = (a - bL)$; 9. $k = a + bL$.

only approximate adjustment function $k(L)$ using segments of these functions over the area where they do not exceed L.

According to the differences in the adjustment function, the solution of Equation 4-3 will be different; hence, the specific form of the $R(L)$ function in Equation 4-4 will be different. Previously, we related the form of $k(L)$ to the respective exploration strategy. Therefore, different forms of $R(L)$ corresponding to different forms of $k(L)$ are also related differences in strategies. This proves that there may not be a universal dependence between the reserve accumulation and drilling footage. The form of such dependence is diverse and is related to the specifics of exploration in a particular region. Defining the types of dependence is only possible to the degree that the specifics are similar. This also covers the oil and gas reserve discovery and appraisal models.

It follows from the aforementioned that the adjustment function is a suitable means for (1) finding similar and dissimilar features in the development histories of different regions, (2) classification of regions by these histories, and (3) constructing the reserve discovery and appraisal process models.

As already indicated, the foundation of the process under study is finding a dependence of the reserve accrual on the extent of drilling. Thus, the solution 4-4 of the Equation 4-3 for each specific adjustment function $k(L)$ will represent the model we are searching for. Table 4-1 illustrates the integrated results of Equation 4-3 when corresponding $k(L)$ functions are substituted. This table shows nine models of the oil and gas reserve discovery and appraisal process corresponding to different stages of this process. The obtained solutions $R(L)$ of Equation 4-3 are self-evident with the exception of the $R(L)$ function presented on line 3 of Table 4-1. The reason for this is when substituting the function $k = ae^{-bL}$ into Equation 4-3, after separation of variables, the following equation is obtained:

$$\ln R = a\int \frac{e^{-bL}}{L} dL + \eta \tag{4-17}$$

where η is a constant of integration.

The integral on the right side of Equation 4-17 is not solvable as a function. For numerical integration, the e^{-bL} function was expanded into a Fourier series. After certain transformations, we obtain an expression for $R(L)$ as shown on line 3 of Table 4-1. For practical computation of the P value, the infinite series was approximated by a finite series. The P series was terminated when its subsequent term was smaller than some small value, assigned in advance, which characterizes the error in the P determination. We accepted $1*10^{-4}$ as this measure of error.

It is worth mentioning that the models we arrived at do not coincide with the models that are usually used to describe the connection between the accumulated reserves and the extent of drilling. The model on line 7 of Table 4-1 is the only exception. This function is known as the Gomperz function. As mentioned earlier, it is often recommended for use for the stated purpose. This particular model describes a rare situation and, in most cases, this model will not adequately represent reality. Unfortunately, applying it as a universal model for all exploration regions appears to be a common but inadequate solution.

The $R(L)$ functions' behavior is presented in Table 4-1 through their characteristic features: the presence of the maximum, asymptote, inflection point, and so forth. This provides an opportunity to visualize the types of curves described by the functions. For instance, not all $R(L)$ functions increase with increasing L. Some of them grow, reach their maximum, and then begin to decline. The decline is caused by the $k(L)$ values becoming negative after having reached the L cut-off value. In a case like this, only the parts of the represented functions that correspond to their increasing values serve as the model of relation between R and L—the accumulated reserves cannot decline.[2]

Table 4–1
Typical Features of Adaptation Functions *K(L)* and their Respective Functions *R(L)* Describing the Relationship between the Accumulated Reserves *R* and Total Drilling Footage *L*

Model Number	Characteristics of *k(L)*							
	Form of *k(L)*	Asymptote *A*	Cut-Off *L* Value at which $k \geq 0$ Condition is Satisfied	Maximum Point		Inflexion Point		Form of *R(L)*
				k_{max}	L_{max}	k_{infl}	L_{infl}	
1	$k = a/\ln L$	0	≥ 1	—	—	—	—	$R = \mu(\ln L)^a$
2	$k = aL^{-b}$	0	∞	—	—	—	—	$R = \mu e^{-(a/b)L^{-b}}$
3	$k = ae^{-bL}$	0	∞	—	—	—	—	$R = \mu L^a e^{aP}$, where $P = \sum_{m-1}^{\infty} \frac{(-bL)^m}{m * m!}$
4	$k = a - bL$	—	$\leq a/b$	—	—	—	—	$R = \mu L^a e^{-bL}$
5	$k = c - aL^a (b > 1)$	—	$\leq (c/a)^{1/b}$	—	—	—	—	$R = \mu L^c e^{-(a/b)L^b}$
6	$k = a$	—	—	—	—	—	—	$R = \mu L^a$
7	$k = ae^{-bL}L$	0	∞	$ae^{-1}(1/b)$	$1/b$	$ae^{-2}(2/b)$	$2/b$	$R = \mu e^{-(a/b)e^{-bL}}$
8	$k = (a - bL)L$	—	$\leq a/b$	$a^2(4b)$	$a/(2b)$	—	—	$R = \mu e^{\left(a-\frac{1}{2}bL\right)L}$
9	$k = a + bL$	—	—	—	—	—	—	$R = \mu L^a e^{bL}$

Table 4–1
(continued)

	Characteristics of $k(L)$					
	Minimum Point		Asymptote	Inflection Point		
	R_{max}	L_{max}	A	R_{infl}	L_{infl}	Maximum Efficiency max (dR/dL)
1	—	—	—	$\mu(a-1)^a$	e^{a-1}	$\mu a(a-1)^{a-1}e^{1-a}$
2	—	—	μ	$\mu e^{-(b+1)/b}$	$[(b+1)/a]^{-1/b}$	$\mu e^{-(b+1)/b}\left(\frac{b+1}{a}\right)^{1/2}(b+1)$
3	—	—	—	$\mu L_{infl}e^{bL_{infl}}$	Satisfies condition: $e^{-bL_{infl}} = (bL_{infl}+1)/a$	$\mu L_{infl}e^{bL_{infl}}(b+1/L_{infl})$
4	$\mu(a/b)^a e^{-a}$	a/b	—	$\mu e^{-a(1-a^{-1/2})}[a/b(1-a^{1/2})]^a$	$(a/b)(1-a^{-1/2})$	$\mu e^{-a(1-a)^{-1/2}}[(a/b)(1-a^{-1/2})]^{a-1}a^{1/2}$
5	$\mu(c/a)^{a/b}e^{-c/b}$	$(c/a)^{-1/b}$	—	$\mu L_{infl}^c e^{-(a/b)L_{infl}^b}$	$[1/(2a)]\{(2c+b-1) \pm [4cb+(b-1)^2]^{1/2}\}^{1/b}$	$\mu L_{infl}^{(c-1)}e^{-\frac{a}{b}L_{infl}^b(c-aL_{infl}^b)}$
6	—	—	—	—	—	—
7	—	—	μ	μe^{-1}	$-[\ln(b/a)]/b$	$\mu e^{-1}(a^2/b)$
8	$e^{a^2/2}$	a/b	—	$\mu e^{\frac{1}{2b}(a^2-b)/(2b)}$	$a/b - b^{-1/2}$	$\mu e^{\frac{1}{2b}(a^2-b)/(2b)}b^{1/2}$
9	—	—	—	—	—	—

Note: All L values are taken from the time of the first reserve accrual. All parameters μ, a, b ,and c are positive constants.

Not all of the functions have an asymptote. Many scientists believe that the presence of an asymptote is a precondition. They associate the size of ultimate potential resources with the asymptote. In reality, this is not true. The appraised reserves will reach the value of the ultimate potential after some finite amount of exploration has been achieved. There will be a point on the $R(L)$ curve corresponding to this amount of exploration. Therefore, the entire history of the accumulation of reserves will be represented by some segment of the $R(L)$ curve. There may not be an asymptote, but if present, it does not mean it can be used to evaluate ultimate potential. This is not a crucial precondition. An asymptote can reflect the estimate of ultimate potential or may be higher than this estimate.

Another important fact is that the models we have inferred are quite flexible. The location of the characteristic points is not singularly and rigidly fixed on the $R(L)$ curve. It varies with the change in the model's parameters. Some rigidity occurs only with the Gomperz function (Model 7). This feature has been used to argue that the typical functions (Gomperz, logistic, etc.) do not allow for variations in the development history of each region. The "rigidity" of the Gomperz function is evidenced by the following: when the maximum reserve accrual is reached (the inflexion point $R(L)$) it uniquely corresponds to a certain percentage of the transfer of the initial potential into appraised reserves. This fraction equals $1/e$. For the rest of the models these percentages are not predefined and depend on the parameters (see Table 4-1).

The absence of a universal dependence of R on L, and the relationship between the $R(L)$ and the exploration strategy, requires us to use a different approach to the resource forecast and long-term planning depending on a region's development model. Previously discussed results indicate that the reserve accrual forecast reliability strongly depends on the model selection. For long-term planning, strategic issues of the exploration specifics in a region should be taken into consideration, which determines the model selection.

This is important when the requirements for the forecast reliability are relatively rigid. For a rough forecast with a relatively short period in the change of L, the models may be used, which approximate the corresponding segments $k(L)$ by a simple function such as a straight line or even a constant. As an example, for a stage that is expressed by the "tail" of the $k(L)$ adjustment function asymptotically approaching the L axis, Models 3 or 4 may be used, which approximate the "tail" by an exponential curve and a straight line, respectively.

The following is no less important for accurate forecasting. There is a unanimity in the belief that it is possible to use $R(L)$ for determining reserve accrual ΔR from a planned extent of exploration ΔL over the planned time interval. This is difficult to substantiate because an increase in the extent

of exploration will lead to an accelerated reserve accrual rate only at a certain rate of increase in drilling footage. Otherwise, due to the replacement of independent wells for dependent ones, the increase in the amount of drilling will not produce the expected result [18, 19]. Thus, the amount of drilling has a certain tendency to change in time, which causes the transition from $k(t)$ to $k(L)$.

The forecast in time using an $R(L)$ model is only possible if the assigned drilling footage is implemented according to that tendency. In other words, drilling is expended not over an arbitrary time interval, but over the time interval defined by the connection between $k(t)$ and $k(L)$. Correspondingly, the forecasted reserve accrual may only be attributed to that time interval.

A common method of forecasting uses the absence of an explicit time presence in the model to determine the connection between R and L. As Equation 4-1 shows, however, time is implicitly present. Therefore, a time forecast from the $R(L)$ model should be supplemented either by a transition function from $k(t)$ to $k(L)$, by a model of the drilling time evolution $L(t)$, or by a model of the reserve accumulation evolution $R(t)$ (if the requirement is to determine the amount of drilling based on the reserve accrual). Models of the $R(t)$ type also correspond to certain specifics of exploration strategy in a region, as was discussed in Chapter 2.

All of the above should be kept in mind if an improvement in forecast reliability is the goal.

Effect of the Exploration Strategy on the Relations between the Evolution of Oil and Gas Reserve Accumulation and Drilling Volume

The purpose of examining the degree to which the models describe real-life processes was to determine the validity of the assumptions regarding the form of the adjustment function $k(L)$ and, correspondingly, the type of dependence of R on L based on the particulars of the development process in a region. We wanted to find out if the differences in the degree of approximation using different models were statistically significant. It was possible that all models approximated equally well (or equally poorly) the real dependence of R on L. If this were the case, any of these models could be considered a universal model. If this were not the case, it would require ensuring that the differences were indeed connected with the differences in the resource development strategies and with the particulars of the exploration process in the regions. Finally, it was necessary to prove that if this were the case, then the proposed progressive

evolution of the adjustment function $k(L)$ reflected progressive change in the strategies (toward their deterioration).

We selected three regions for the examination: regions A, B, and E. As we have already indicated, the development histories in these regions are different. As a reminder, regions A and B are drastically different in their exploration techniques and systems, and region E occupies an intermediate position.

When evaluating the model parameters, we substituded $y = \ln R(L)$ for $R(L)$. This reduces Models 1, 4, 6, 8, and 9 to a linear form (of one or two variables). The models' parameters were evaluated using the least squares method. The balance of the models are linear versus the variable, which itself is a non-linear function of b. Sorting of b was performed, and its subsequent evaluation was conducted by the minimum of summed squared deviations $SS_y = \Sigma(y_i - \hat{y}_i)^2$. Whether or not the model fits the observations may be judged by the value SS_y or variance σ_y^2. This value is obtained as an approximation result of $y = \ln R(L)$ values by the corresponding $\hat{y}$ expressions from the models, and by the correlation coefficient $r_{y\hat{y}}$ between y and $\hat{y}$. These two variables are related. The correlation coefficient $r_{y\hat{y}}$ as opposed to σ_y^2, may be more convenient for comparing results in different regions because it does not depend on the absolute value of R; that is, on the magnitude of the region's ultimate potential. At the same time, the F-criterion (SS_y or σ_y^2) may be used for testing (within a single region) the statistical significance of the differences in describing the connection through different models.

Table 4-2 shows the results of the models' examination against the actual data. A review of this table shows that there is a core group of models for each of the three regions. The model at the core's center exhibits the smallest SS_y value. As the distance from the center grows (i.e., as we switch to the other models according to their order in Table 4-1), the SS_y value increases. The last in this series, in any direction from the core's center, is the model with the worst approximations of the observations. Its distinctions from the central model become significant. The F-criterion was used to examine the statistical significance of these differences. Such a model defines the core boundary.

Table 4-2 indicates that each region contains a specific set of core models. From region A to E and then to B, the locations of the center, as well as of the actual core, progressively shift: the models on the upper line of Table 4-2 are replaced by the models located on the lower lines. Regions A and B have a clearly-defined core of two models each, whereas the region E core is nebulous and includes seven models. This is easily explained by the position of region E between the other two in terms of exploration strategy, which gives it certain similarities to both regions.

Table 4-2
Models of the Reserve Accumulation Evolution vs. Drilling Footage Adequacy Evaluation against Observations

Model Number	Region A		Region E		Region B	
	$r_{y,\hat{y}}$	SS_y	$r_{y,\hat{y}}$	SS_y	$r_{y,\hat{y}}$	SS_y
1	0.994	0.7038	**0.98**	**1.2577**	0.9109	14.9212
2	0.9947	0.6226	**0.9911**	**1.0177**	0.9395	10.2833
3	**0.9973**	**0.319**	**0.9913**	**0.9979**	0.9394	10.3047
4	**0.9973**	**0.3111**	**0.9913**	**0.996**	—	—
5	—	—	**0.9914**	**0.9825**	—	—
6	0.9838	1.8802	**0.9904**	**1.0979**	0.9394	10.309
7	0.9861	1.6091	**0.9846**	**1.7508**	**0.9865**	2.3563
8	0.9105	9.9796	0.9664	3.7875	**0.9916**	**1.4668**
9	—	—	—	—	0.9672	5.6492

Note: Bold data pertains to core models.

These similarities are reflected in some commonality of models between regions A and E, on one hand, and B and E, on the other.

The aforementioned confirms the dependency of the $R(L)$ function form on the specifics of the exploration history in the regions. It also confirms the validity of the preconditions assumed when constructing the models; in particular, the validity of the assumption that the proposed evolutionary sequence of the adjustment function $k(L)$ reflects the progressive change in the development strategy. A good fit of the core models with reality is supported by high values of the correlation coefficients and low values of variances.

Thus, we developed a specific technique for solving practical tasks, which includes an evaluation of the ultimate oil and gas potential from the development history curves; the reserve accrual and resource status forecast; the proposed extent of drilling and corresponding funding for long-term planning; etc. Reliable solutions for these tasks may be obtained by different treatments in each region and by considering historical specifics of the region's resource development, rejection of a single universal model approach, and the transition to the more specific, higher-quality models proposed above.

At the same time, the proposed models provide a convenient means for the quantitative characterization of exploration strategies in the

region's resource development through the adjustment function $k(L)$ and its parameters. Based on this, comparisons between regions and their classification may be conducted, which is important for an analysis of exploration techniques and results.

Notes

1. For the considered case we define the terms as follows: speed is dR/dt or dL/dt; rate of change is $1/R*dR/dt$ or $1/L*dL/dt$.
2. Of course, it is not a matter of the reserve decline for the reason of production.

CHAPTER 5

Reserve History and Evolution of the Undiscovered Resource Structure: Physical Models

As mentioned previously, reserve evolution is the result of a certain order in field discoveries, and is closely related to the structure of discoveries. Using a physical analogy, one may consider two levels of description: macro-level and micro-level. The first level represents a dynamic description of movement (change) and trajectories (curves). The second level represents a description in terms of distribution functions, their evolution in time, and thermodynamics. These levels reflect different facets of the same exploration process.

The possibility of unifying these two levels of description and of better understanding the substance of the observed patterns from a physical viewpoint, occurs in connection with the concepts developed in physics for describing the self-organizing process

We will briefly characterize this concept and will present the possibility of describing the oil and gas reserve accumulation process from a completely different perspective. Generally speaking, we will examine the process of reserve accumulation through the physical techniques of comparing the phenomena observed in the process of reserve accumulation with the laws of physics. Thus, we can use a more general language for the description of this process.

An analysis of complex systems through physical models is quite common in different scientific disciplines (biology, economics, ecology, etc.). The similarity between physical concepts and techniques and the tasks of cybernetics, informatics, and so forth, resulted in the birth of such disciplines as physical cybernetics, and informational thermodynamics (or physical information theory and informo-dynamics).

The process we are analyzing is a controlled process—a process based on certain principles. At the same time, concepts and techniques have been developed in physics to study the self-organization processes. Consequently, the analogy is quite appropriate and may help enrich our ideas and open new avenues of study.

Self-organization is the emergence of order in a system that results in a system of consecutively more-ordered states. In a self-organizing system, new more-or-less stable structures constantly appear and continuously change. Self-organization, therefore, is the evolution of these structures. What is the connection between order and chaos? Why and under what circumstances does one turn into the other? What sort of mechanism turns chaos into order? What are the laws governing the birth and growth of structures in the system where previously there were no structures? The answers to these questions are provided by synergetics, which is a theory of self-organization and the development of structures. Currently, the concept of self-organization has already penetrated physics, chemistry, hydrodynamics, biology, ecology, and some other disciplines. Our interest in structural evolution is restricted to the evolution in the structure of undiscovered resources based on a certain structure of discoveries.

Let us analyze this analogy from several different angles.

Dynamic and Thermodynamic Descriptions

When attempting to apply physical techniques, the first question to ask is whether a description of the reserve accumulation evolution (dynamics) using the techniques of dynamics, is legitimate. Dynamics studies behavior that is reversible in time. The equations of classic dynamics are invariant with respect to the $t \rightarrow -t$ substitution. In other words, they are symmetric with respect to the time reversal $t \rightarrow -t$: each process has a corresponding process that travels "backwards." Thus, there is no difference between past and future dynamics. Consider the following principle in dynamics: if a system's state is known with sufficient accuracy at a certain point in time, the future state may be predicted, and the past state may be restored (at least in principle) because the system is controlled by the equations determining its motions. In other words, a distinctive feature of classic dynamics is its strict determinism. The only ambiguity is in the initial conditions. An incomplete description is determined only by a scatter in the initial conditions, however, it does not affect the reversibility of distribution function equations. To a significant extent, dynamics emphasizes the regularity, stability, and invariability of motions. The world of classic dynamics is a stable, static world, with no place for the emergence of the new (that is, a world that is devoid of evolution).

The form of the $R(t)$ function is not always explicitly defined. Nevertheless, we arrived at it as a deterministic function; that is, we assigned the trajectory of the reserve accumulation based on deterministic principles, which is reflected in the equation $R_{i+1} = f(R_i)$. This means that the trajectory is stable. A natural question arises: Is this determinism prescribed by the nature of the process under scrutiny, or was the function only made to fit the observations? In other words, is the $R(t)$ function simply a mathematical approximation of the historical evolution, which does not reflect the substance of evolution and its changes?

In this respect, the oil and gas reserve accumulation process reflects the adaptation of exploration to changing conditions. This adaptation may have an adjustment characteristic, just as living organisms adjust to the changing environment. The process evolves in such a way that adaptation to the changed environment is the best response. The organization of the process assumes a form that is most stable in a given environment.

This evolution of the process is predictable to a certain degree. It may be described by an equation reflecting the connection between the preceding and the following state. It is exactly this aspect of the process that is described by the $R(t)$ function.

The stability of the process, however, may be shattered by a drastic change in the environment. The adaptation to the changed environment requires a different form of organization. The continuation of the process is no longer possible with the existing form of organization. Ambiguity, uncertainty, diversity, and complexity arise in its continuation because its trajectory is no longer defined by the previous dependence $R_{i+1} = f(R_i)$. The process is transformed to a new state where the present does not contain information about the future. The invariancy of a regular dynamic description is destroyed and the process becomes irreversible.

Thus, the future is no longer included in the present, and, as a result, the time becomes uni-directional and irreversible. Techniques for describing a wide range of processes with uni-directional time are developed in thermodynamics. They address simple, irreversible processes as well as complex processes accompanied by self-organization. The processes studied in thermodynamics travel spontaneously in only one direction. This is called the evolution process. The concept of evolution is introduced into physics by the second law of thermodynamics. The second law introduces a physical value called entropy, which gives a certain direction to the time. It is impossible to describe evolutionary processes without introducing the direction of time. The second law establishes the direction of movement of irreversible processes. According to this law, all irreversible processes in closed systems occur in such a manner that the system's entropy increases until it reaches the maximum state of thermodynamic equilibrium. All irreversible processes are non-equilibrium

processes. Equilibrium for these processes is the final state of their thermodynamic evolution. The second law of thermodynamics associates the positive direction of time with the increase in entropy.

In some cases, the external conditions of a system are already determined. For instance, for a system with a given volume at a set temperature, the law of increasing entropy (valid for isolated systems) is replaced by the law of decreasing free energy. Under these conditions the direction of real processes is determined by the decline in free energy, and the state of equilibrium is reached when free energy is at a minimum.

Open systems can exchange energy and/or matter with the surrounding medium. Irreversible processes in these systems may run at a constant, or even decreasing, entropy because of the entropy exchange with the surrounding medium. The increment of entropy includes two definitions: one describes the entropy transfer over the system's boundaries, or the entropy exchange between the system and the surrounding medium; and the second describes entropy generated within the system as a result of irreversible processes. This second term is called the entropy generation. The second law of thermodynamics states that the entropy generation within the system is always positive. This is the main distinction between the reversible and irreversible processes: only the irreversible processes contribute to the entropy generation. Accordingly, the state of equilibrium is characterized by zero entropy generation.

Therefore, if the elements of exploration that are associated with irreversible processes are emphasized, instead of the stability and harmony, a dynamic description is insufficient. The significance of non-dynamic elements is even greater if we switch from deterministic dynamic processes to those evolving from a more-ordered state to a less-ordered state, and, conversely, to those leading to increased diversity and complexity in the system.

Two Directions of Evolution

Certain physical processes in nature travel in one direction only: order deteriorates with the passage of time and is replaced by the lack of order (chaos). For instance, if the gas concentration in a closed vessel is non-uniform, it will become uniform with time. This evolution complies completely with the second law of thermodynamics. As discovered by Boltzmann, entropy is a measure of molecular chaos, meaning that increasing entropy reflects a growing disorganization in the process of evolution. Thus, he introduced the microscopic meaning of entropy. He associated entropy with the probability of state and indicated that a growth in entropy corresponds to evolution toward the most probable state. Thus,

the concept of probability was first introduced to theoretical physics through the molecular interpretation of irreversibility. Thermodynamic equilibrium corresponds to the state of total disorder, or chaos, which is the most probable state. A system then tends to achieve its most probable state in the process of evolution. These results led to the association of classic thermodynamics with the destruction of structures. However, there is another domain within thermodynamics that addresses the spontaneous creation of structures, which will be discussed later.

There are also numerous processes in nature that travel in the opposite direction, which are associated with the emergence of order (with the creation of structures). These processes are especially conspicuous in the biological and social domains. The Darwinian theory of biological evolution does not agree with Boltzmann's physical theory of evolution; the essence of evolution in biology is diametrically opposite. Biological organization is not a result of evolution toward molecular disorder, but appears as a series of structures and mutually related functions that become even more complex and hierarchically ordered. The Darwinian theory describes the evolution from primitive living systems to the highly-ordered structure of a human organism (i.e., biological evolution reflects the transition to ever-higher levels of complexity). Biological systems are extremely non-uniform (heterogenous).

It is the same for other phenomena associated with self-organization processes. Increasing complexity of organization and growth in the diversity of organizational structures are among the most important characteristics of the evolution process.

Thus, evolution is directed differently in inanimate matter than in the living world (and also in the inanimate systems analogous to biological systems). This led us to question whether or not it is possible to explain phenomena of the living domain through physical (chemical) laws, and whether or not it is possible to apply to living matter the power of acting law in inanimate matter. Many difficulties were encountered in applying the second law of thermodynamics to living systems. These problems were solved by Ludwig Bertalanffi who was the first to show that living organisms are open systems. The second direction of evolution may only occur in open systems that can exchange energy or matter with the surrounding medium. This explains the decrease in entropy observed in such systems, and was the basis for Erwin Schrödinger's theory that a living organism feeds on negative entropy in supporting its organizational level, in spite of the second law of thermodynamics [54].

Among the most important characteristics of the self-organization process is the growing complexity we call evolution. Even more important may be the ever-increasing ability to assimilate energy. Living forms are continuously emerging on this planet that are capable of assimilating

ever-increasing amounts of external energy. The more energy and matter they can assimilate, the more complex their organization may be.

Among the most important tenets in the evolution of self-organization processes is the principle of minimum dissipation of energy, or the principle of minimum growth (or maximum decrease) of entropy. Its substance is as follows.

If there is more than one form of the process realization under the laws of physics, then the form that is accompanied by the slowest growth in the system's entropy will be implemented. In other words, only the most economic processes, in terms of energy, are implemented in nature. As soon as the efficiency in the utilization of exogenous energy stops growing, the system stops developing. This supposition is apparently supported by the example of termites and ants. An excessive stability in their evolution (associated with the total loss of individuality by the members of society) results in an evolutionary dead-end.

Based on this theory and on the established laws of thermodynamics, thermodynamics was expanded to include a natural physical explanation to the evolution theory of self-organizing (including the living) systems. The multitude of systems with both types of behavior (systems behaving chaotically in the thermodynamic sense in some situations, and an organized, heterogenous, and coherent manner in the other situations) corresponds to different physical situations, but to only one type of physical law. For the former, entropy sufficiently describes the degree of organization. It ascribes to such an organization the architectural image of a contradictory unity of the conservation and destruction processes. For the latter, entropy is no longer sufficient to describe the degree of organization.

This variance in the behavior of systems in different situations and, consequently, the difference in their evolution, is of great interest. As shown later, at different stages of the regions' resource development, the process being analyzed may be compared with (1) the evolution of living (self-organizing) systems, and (2) the evolution of inanimate (evolving toward chaos) systems. Physical explanation of evolution for these two types of systems helps us to better understand the significance of the reserve accumulation process.

States of Equilibrium and Non-equilibrium

Systems have different types of behavior depending on whether or not, from a thermodynamic standpoint, they are in a state of equilibrium. Although the law of equilibrium is universal, when a system is far from equilibrium, its behavior can be very specific. Forms of matter that emerge and continuously change are non-equilibrium structures. In this

case, thermodynamic equilibrium is simply not possible. A permanent exchange of matter, momentum, and energy occurs between such a system and its surrounding medium.

In a state of thermodynamic equilibrium, the speeds of irreversible processes (chemical reactions, diffusion, etc.) occurring in the system, and the corresponding generalized forces (temperature gradients, gradients of chemical potentials), are all equal to zero. Similarly, the final state for the reserve accumulation process (when all ultimate potential is transferred to the appraised reserves) may be considered the state of equilibrium. In this stage (the equivalent of speed $\Delta R = 0$) all exploration is completed, and the expenditures of funds and drilling footage at that moment in time are also equal to zero.

Resources are finite and nonrenewable, so no matter how the reserve accumulation process is progressing, eventually it reaches a state of equilibrium; after which exploration of the region is completed and the maximum information is obtained. The stage corresponding to the long, descending branch of the ΔR curve at its end segment, may be called the near-equilibrium state. At this stage, the process is degrading to its degenerative status. Conversely, all preceding stages may be called the far-from-equilibrium states. At this stage, the process is not degrading but, on the contrary, is flourishg. It appears that in regions A, B, and others the process is still at this stage. They have not yet crossed the critical cut-off point after which their status may be considered near-equilibrium. The process we are studying behaves differently depending on the systems' behavior far from and near equilibrium. Far from equilibrium, the process has a great deal in common with self-organizing systems; near equilibrium, it behaves as if it were a thermodynamic process.

Emergence and Destruction of Structures

Destruction of structures is prevalent near thermodynamic equilibrium. On the contrary, emergence of structures may occur, according to certain non-linear laws, outside of the area of stable states, which corresponds to regular thermodynamic behavior. This does not mean that structures cannot exist near equilibrium. The structure of equilibrium states is ruled by the major principle of Gibbs' canonical distribution, or a probability distribution of a given microscopic state of the system in thermal equilibrium with the surrounding medium where a constant temperature is maintained (as if with thermostat). This is called the Boltzmann order principle. It explains the existence of equilibrium structures and provides a means for describing a great number of such structures. Among such structures, for example, are snowflakes. In a

general case, the processes of dissipation eliminate any vestiges of order and result in thermodynamic equilibrium.

Far from equilibrium, the thermodynamic description acquires a different form. An important feature of this form is that far from equilibrium the systems may develop a new type of structure: dissipative structures. These structures are a result of the lack of equilibrium, which is important in understanding the coherence and organization of biological systems. The concept of the Boltzmann order principle does not include non-equilibrium situations. The appearance of such ordered structures at thermodynamic equilibrium would be unlikely. It may mean, for instance, that a macroscopically observed number of molecules would spontaneously organize into an ordered flow. The very reason for a system to form ordered, highly cooperative structures is that external limitations keep the system significantly far from equilibrium. For such situations, a new principle of order appears, which cannot be reduced to the Boltzmann order principle. This second principle of order is called "order through fluctuations."

The specific role of fluctuations is discussed later. In the meantime, it needs to be emphasized that organization pertaining to biological systems is only possible because living systems are non-equilibrium systems. They are never in a state of equilibrium. Using their free energy to work against equilibrium, they are able to avoid degradating to equilibrium, and thereby avoiding the inert, "dead" state of equilibrium.

For exploration, a state of equilibrium is also an inert, "dead" state. Unfortunately, it is unavoidable due to the finite nature of resources. While the size of yet undiscovered fields allows for it, however, all efforts are directed toward maintaining the process far from the state of equilibrium. According to the non-equilibrium or equilibrium state of the process, we deal with different degrees of structural non-uniformity equivalent to either chaos or high heterogeneity of biological or physical systems. Near equilibrium, the increase in "chaos" (destruction, decrease in order) is reflected in the increase of uniformity (in terms of the reserve size) of the fields to be discovered. Field sizes hardly differ (they are all small) and there is no drastic heterogeneity, non-uniformity, or inequality. Such uniformity is similar to the uniform distribution of gas density in a closed vessel, which is a typical example of the evolution toward a lack of order.

This similarity is only increased by the fact that small fields, due to their great number and the absence of specific limitations for the conditions of their formation, are distributed over the regions in a more uniform, "disordered" manner. They do not create resource concentrations in certain plays but, on the contrary, "smear" resources over different plays. Accordingly, exploration is no longer concentrated on certain plays. As the advantages of different plays smooth-out, random elements become

prevalent in decision making regarding the selection of plays for exploration. This is where the Boltzmann order principle comes into play. It is no longer possible to maintain the system far from equilibrium by increasing the extent of exploration (well drilling), because exploration of small fields requires large expenses per ton of reserves.

The far-from-equilibrium situation is substantially different. The resource structure is drastically non-uniform, heterogeneous, and ordered to the highest degree. A significant portion of the resources is concentrated in a few large or giant fields. These fields attract most of the exploration work. Functionally organized, hierarchical structures appear—from exploration in the region as a whole to exploration and appraisal of individual fields and accumulations. Each level has its own decision-making principles, which are far from being random (concentration on certain plays is quite purposeful), and exploration efficiency is high. This is a realization of the minimum dissipated energy principle; exploration targets do not lose their individuality due to drastic differences in the size of reserves for different fields.

Fluctuation and Instability

Ordered structures emerge as a result of complex, non-linear interactions in systems far from the state of equilibrium. Whether chaos or order prevails depends on the existence and nature of instabilities in the system. Instabilities can occur in any system. They are generated by fluctuations (deviations from movement) spontaneously occurring in the system. If they are small and rapidly decay upon emergence, they do not lead to significant deviations from the initial trajectory in the system's movement—the movement remains stable. In such a case, instabilities play a secondary role of negligibly small deviations, which may be disregarded. Fluctuations spontaneously created by the system itself tend to be small. The environment tends to dampen the fluctuations through interactions at the boundaries of fluctuating areas and stabilizes the movement through energy exchange. When fluctuations are small, boundary effects dominate and cause them to dampen and disappear.

Fluctuations play a decisive role in the emergence of dissipative structures. When external disturbances are greater than a certain cut-off value, fluctuations do not subside, and the reaction to these fluctuations does not return the system to its initial state. Instead, these fluctuations amplify and may destroy the existing structure giving rise to a new structure. In such a case, the fluctuations cannot be disregarded and must be taken into consideration. The causes of fluctuations are complex interactions within the systems and correlations (statistical connections)

between the elements of the systems. These interactions are non-linear; that is, a reaction to an external action is not proportional to the size of this action. This is why large fluctuations, critical for the existing regime representing a certain structure, can occur as a result of small changes in external influence.

In connection with fluctuations, problems of movement and structural stability arise. Movement is stable if small deviations lead to trajectories close to the initial one over indefinite time period. States corresponding to thermodynamic equilibrium, and states near equilibrium in linear non-equilibrium thermodynamics, are automatically stable. In these stable regimes, fluctuations, or deviations from equilibrium or some other stationary state close to equilibrium, decay exponentially. The fact that instabilities do not occur is a result of linear relations between the generalized fluxes and generalized forces.

These relations are not applicable to non-equilibrium systems. Having once emerged, fluctuations, due to non-linear physical laws, can exceed the stability cut-off level. The emergence of a new structure in an open system is always the result of instability. Evolution of systems obeying non-linear laws can progress through a number of instabilities. As a result of non-linearity, instabilities and multiple transitions to progressively more-ordered structures can occur.

All organized structures, including living organisms, maintain only temporary stability. Any organization will eventually be destroyed and any living organism will unavoidably perish. A state of a given complexity may possess a "memory" of prior instabilities. Each of these can contribute to the emergence of specific features that are essential for the stability and preservation of a given state. In living organisms this defines information, which needs to be transmitted. Instabilities are preconditions for selective growth and evolution. The existence of different organizational levels in living matter correlates with a sequence of instabilities. This is how order emerges through fluctuations.

From a mathematical point of view, instability is associated with a loss of stability in solutions of corresponding non-linear equations: small deviations in the system's parameter(s) lead to big changes in the system's output. The emergence of a certain regime is determined by the stability of the corresponding solution under conditions imposed on the system.

The process of reserve accumulation in a state close to equilibrium is stable. At this stage of development, fluctuations are small and gradually fade away. Most likely, the reserve accrual under these circumstances linearly correlates with the drilling footage, which is no longer concentrated on certain dominant plays. Also, such a situation does not pose a problem of choice: all prospective unexplored targets are of interest. This results in the absence of multiple solutions.

In a state far from equilibrium the process evolves according to non-linear laws. It progresses through a number of instabilities, undergoing transitions to new states. These transitions result from the decision to shift exploration work to certain areas and plays. Here, the problems of choice and selection are constantly present, with many possible solutions and numerous degrees of freedom. The decisions are made based on fluctuations in the reserve accrual process or, rather, on information concerning the most significant discoveries reflected in these fluctuations. They create preconditions for new exploration, new discoveries, and new instabilities, based on the record of prior instabilities. The process is directed by information acquired in the course of exploration. This creates a feedback mechanism and the preconditions for instability. All efforts in controlling exploration are focused on not allowing the process to stabilize. A temporary stability at a stage far from equilibrium is a result of poor play selection. An example is region B, where only small reserves were accrued over many years, thereby resulting in only small fluctuations.

According to the thermodynamics of non-equilibrium processes, fluctuations far from equilibrium are not the only ones that increase. Similar phenomena are observed in the reserve accrual. This is associated with the modulation of random components of ΔR functions. The interconnection of fluctuations, described by the second-order autoregression, may be considered an analog to periodic solutions of non-linear equations.

Stability Criteria

Stability may be roughly defined as the internal capacity of a system to resist disturbing factors and to suppress fluctuations. As already mentioned, the mathematical definition of stability is that the system's trajectory in phase space is within a certain assigned area for certain finite disturbances of a wide spectrum. The investigation of stability can be performed using Lyapunov's functions. The same approach is used when determining thermodynamic stability. The existence of the Lyapunov function enables all fluctuations to subside. The second law of thermodynamics makes it possible for the Lyapunov function to exist in isolated systems, and thermodynamic equilibrium is the state to which the system will return after being disturbed.

In the domain of linear non-equilibrium thermodynamics, generation of entropy corresponds to the Lyapunov function. Let us consider the area of the state of equilibrium. If a fluctuation forces the system to deviate from its initial state, the system will evolve in response and return to the state of thermodynamic equilibrium (i.e., zero entropy generation). If this is prevented by the assigned boundary conditions, the system will transition

to a stationary non-equilibrium state with minimum entropy generation (the state where the matter transfer ceases, whereas the energy transfer continues, otherwise known as the state of "minimum dissipation"). These states are unstable. Far from equilibrium, the stability of a stationary state is judged by the so-called excess in entropy generation. If the values of the excess in entropy generation are not negative, then the excess is associated with the Lyapunov function. Therefore, any fluctuation near the stable equilibrium causes a negative entropy generation. Any fluctuation near a stable stationary state is always expressed by a positive entropy generation. In the case of an excess of negative entropy generation, the stationary state is unstable.

As previously indicated, only irreversible processes contribute to entropy generation. Irreversibility and instability are closely related. Irreversibility can result from a lack of knowledge and an incomplete description due to the complex nature of movement, especially when instability is present. It can also be the result of a certain mechanism that disrupts the invariant nature of a regular dynamic description with respect to time reversibility. A probabilistic interpretation of the law of increasing entropy can be related to an incomplete theoretical description. On the other hand, the second law of thermodynamics may be considered a selection criterion that provides an opportunity to exclude some physically non-implementable states that is the cause of irreversibility. This means that the second law of thermodynamics is valid only for systems where there are states, which turn into prohibited states when time is reversed. Allowable states are implemented with certain probability. In any case, information concerning irreversibility is measured by entropy or, in general, by the Lyapunov function. Statistical expressions for these functions depend on the statistical description of a system, this system must take into account not only probabilities (distribution functions), but also correlations between the elements of the system because the correlations are often connected with time reversion. Interactions within the system result in fluctuations.

The above mentioned theories attempt to use information as a measure of stability and as a structural characteristic of a system's organizational level. Some propose to use informational entropy (or its equivalent) as a measure of stability. Informational entropy, however, only applies to a system with weak interactions. As soon as one considers systems with strong interactions, the informational entropy measure is unsatisfactory. The concept of entropy in classical information theory (the theory of message transmission) corresponds to the Boltzmannian statistical concept of entropy. Entropy characterizes the amount of "ignorance." The amount of information is complementary to entropy. Therefore, the theory of information, which uses this concept, cannot provide more than thermodynamics in its

classical form, which does not address kinetics of dissipative structures. Information theory, analyzing only the amount of information expressed in bytes and not taking into consideration the possibility of new information emerging and its value, is of no consequence for the study of self-organizing processes.

Again, the value of information is determined only by the quality of the decision made. The quality of a decision, associated with understanding the goal, may be evaluated only to the degree of goal achievement. On the other hand, there is a source of irreversibility in the fact that the future is not included in the present. This allows for ascribing corresponding content to information. The future is not predetermined by the initial conditions, or by the final goals, because it is created within the temporal flow through a chain of free, mutually related choices (selections). The system performs a free selection. It is worthwhile to note that, strictly speaking, there is no choice (no selection) for a deterministic or stochastic system. Each selection has certain informational content. Complete freedom is associated with the possibility of non-trivial choice.

A characteristic feature of biological systems is their ability to assimilate and process information. The emergence of information (the instructive—teaching/learning—programming action of molecular and above-molecular information) is a determining condition in biological systems. In natural selection, the value of an informational, instructive (teaching/learning) program is of prime importance. Self-organization is impossible in the absence of instructive properties. In essence, the theory of information, including the emergence of information and evaluation of its value, was constructed during mathematical analysis of natural selection. The value of information attributes a quantitative meaning to the expression "survival of the fittest." Physically, it is reflected in the concepts of the "selection value." The emergence of a new species may be interpreted as an ecological fluctuation. The new species can either disappear (the fluctuation subsides) or displace the previous inhabitants (the fluctuation increases and instability appears). The selection criterion that is directly associated with the imposition of external conditions, which create selection in the system, is exactly what measures the degree of selection advantage for a new species—its chance to survive and dominate.

At the near-equilibrium reserve accumulation stage of a region's development, exploration does not introduce substantial changes in the knowledge of the region's geology and petroleum potential. At this stage, accumulated information has reached its maximum value and the opportunities to acquire new information become very restricted. Applying the theory of information to the description of the reserve accumulation process is justified. At this point, it is appropriate to mention Shannon's entropy and its growth in connection with (1) the degradation of the field

size distribution toward that which is more uniform, (2) the loss of individuality by the fields, (3) an increase in uncertainty of the play selection, (4) the difficulty (ignorance) in the search for small fields, and (5) the lack of significant changes in the knowledge of the regional geology. At this stage, information loses its value and ceases to be a worthwhile control feedback mechanism.

At the stage of reserve accumulation far from equilibrium, the situation is different. Here, the process is controlled based on information; that is, information renders instructive (teaching/learning) programming action. Information arises, accumulates, and acquires value in the process of exploration. Because of this information, the state of process in the future (R_{i+1}) depends on its state in the past (R_i). Based on this information, major exploration plays (exploration and appraisal targets) are selected. The acting feedback mechanism at this stage forces us to make exploration decisions that avoid process degradation to a degenerated state and that maintain a high reserve accrual level. For this reason, large and giant discoveries are most valuable from not only an economic standpoint. They have the highest "selection value" because they determine the entire course of exploration by separating the major plays from the secondary plays. They also determine the highest value of information since their discoveries cause the most drastic changes in the understanding of the region's geology and its potential.

Large discoveries lead to an increase in funding, drilling footage, and, as a result, a general increase in exploration. Also, large fields give the highest reserve accruals. This is due to the fact that they implement the most efficient method of reserve accrual. It is interesting to note that economic profitability is associated with the highest degree of destruction of the initial reserve structure; whereas in inanimate nature these processes are implemented among the processes of conservation and destruction of matter organization, which disturb this organization the least (Le Chatelier principle).

There are usually only a few large fields in each region. They are the exception rather than the rule among the most typical field sizes. Nevertheless, these exceptional targets determine the system of exploration (play selection) since they set the rule of actions, the rule of "selection," and have control of the process. Thus, a high degree of reserve accrual is implemented through the most improbable events in the exploration system.

A similar situation is observed in biological systems. Darwin explained irreversible evolution to higher organized structures in living systems through a process of natural selection. In other words, because natural selection favors exceptional situations (new species that correspond to some type of an ecological fluctuation emerge in small numbers), these exceptions (i.e., new organizational structures) eventually become the rule.

The reason they become the rule is only because a biological system remains far from equilibrium, which is not the case of the process we are studying. However, as long as this process is far from the state of equilibrium, the same situation applies.

Determinism and Chance

The loss of stability is followed by the destruction of the old structure and the creation of a new one. During the period between the loss of stability and the emergence of the new structure the old structure is maintained. The system evolves, adjusting to changing external conditions and preserving stability in the best way possible. In such a situation there is no thermodynamic equilibrium. The system evolution at this stage is predictable. Between the two points of stability loss, a system evolves according to deterministic laws. Fluctuations do not play a significant role and may be disregarded. Evolution mechanisms at this stage belong to the adjustment type.

In the periods of stability loss, the system is governed by different laws. As mentioned earlier, a cause of instability is fluctuations (i.e., a stochastic element). Instability of this type is described by the theory of bifurcation. Bifurcation, or branching, is a change in the number and stability of solutions for an equation describing the movement of the system. Bifurcation occurs the moment a new equation solution emerges after a characteristic parameter of the equation reaches some cut-off value. At this moment, the previous solution becomes unstable and the new one becomes stable. Due to the influence of non-linear effects, which totally alters the situation, the system changes to a new state in an intermittent fashion.

The new solution may be unique. Other situations, however, are also possible when the uniqueness gives way to multiple solutions. This happens when there are several simultaneously stable states near the bifurcation point. The system may gravitate to this stable state after losing stability. If these states are close to one another and are equally likely, it is impossible to predict which one will be selected by the system. This is caused not so much by the closeness of the new stable states, as by the random state of the system at the moment of stability loss. This is associated with the random nature of emergence of a corresponding fluctuation.

Thus, the "selection" among the branches emerging at bifurcation is made by fluctuations. They determine the branch that will be followed by the system in its continuing evolution. This is precisely the cause of similarity in the continuing evolution, which becomes unpredictable. The existence of fluctuations near points of instability makes traditional techniques of the theory of probability inapplicable. The system makes a

free choice, as mentioned above. This choice cannot be predicted even in probabilistic terms.

A series of consecutive instabilities is similar to a series of consecutive bifurcations. The final result depends on the mutual effects of deterministic laws and a probabilistic sequence of fluctuations. The various options for the evolution path at the points of bifurcation form a stochastic temporal sequence. Therefore, the details of any particular model depend on the accurate history of the evolution. Thus, bifurcation in some sense introduces the historic element into physics. An interpretation of the system's current state presumes an analysis of the system's history as it proceeds through the points of previous bifurcations.

Each bifurcation is followed by a stable solution. This means the fluctuation stabilizes a new regime representing the structure that emerged following the instability. As a result, consecutive bifurcations lead to an increase in the diversity and complexity of systems. Thus, the bifurcation mechanisms are responsible for these characteristics of evolution.

The above pertains to the states far from equilibrium. Near the state of equilibrium the solutions are stable, and bifurcations do not arise. Fluctuations are small and do not play any significant role. In this case, a deterministic description is sufficient.

For the reserve accumulation process it is also possible to distinguish bifurcation points. They are associated with the transition of the process to an unstable state and the appearance of multiple options for the process to continue in a new stable state. Such points are associated with discontinuity moments when decisions, or choices, are made. These moments are when new information is acquired that changes our understanding of the region and requires a re-direction of the exploration process. New information is associated with new discoveries, which are reflected in fluctuations of the reserve accrual. A selection of new exploration organization structure in bifurcation points is never singular. However, after the selection has been made, the exploration system is stable until the point of obtaining new substantial information: it is adjusting to the new conditions. Until the next bifurcation the process evolves in a predictable manner. Again, as soon as exploration leads to a new situation that allows a non-singular continuation of the process, further evolution turns unpredictable. The final result (the reserve accumulation evolution) acquires the appearance of one or the other depending on the reserve accrual fluctuations. In other words, it depends on the history of the decision selections at the moments when new data is obtained (exploration strategy). This explains why the oil and gas reserve accumulation process in any new region is unpredictable in principle.

A model, or a particular form of the $R(t)$ function, is selected by fitting it to the results obtained over a significant length of time. The

deterministic form of the model is created under the influence of chance (fluctuations) as well as of determinism (evolution toward a new selection). In this case, chance and determinism do not conflict but, rather, complement each other.

Evolution and Selection

The instabilities mentioned above are preconditions for the selective growth and evolution of living systems. We have already referred to the role of information in this process and to the fact that the value (in terms of being valuable, not in terms of its amount) of information is of prime importance. Self-organization requires an instructive (teaching/learning) capability at the molecular level. However, the self-organization process includes numerous random events without any instructive (teaching/learning) functional significance. The question is to what extent these random events are able to provide feedback for the source of their emergence, thereby serving as a cause of some amplified action. Unless information acquires value, the selections of self-organization will not result in evolution. The value of information, combined with the emergence of this information, is the result of selection. Selection occurs among special substances and under special circumstances. Special substances carry information and are capable of instructing their own synthesis. One of the special circumstances is deviation from equilibrium. Selection and evolution cannot occur in systems that are in a state of equilibrium or near equilibrium, even if the necessary special substances are available.

As mentioned, the reserve accumulation process at a stage far from equilibrium also develops based on incoming information. Here, also, information and selection (selection of major plays and targets) are mutually related. Information is precisely what is needed for such a selection. The availability of choice (the presence of large fields against a background of less-desirable, smaller fields) requires information that otherwise would not be needed. This is the value of information, which is a conduit for the implementation of feedback.

Conclusions

Thus, we can clearly identify two stages in the reserve accumulation process. The first stage is evolution. It corresponds to a far-from-equilibrium state, with all the aforementioned characteristic features that pertain to this stage. The second stage is the degradation process; the stage of its relaxation toward the degenerated state. This stage is typical for a

near-equilibrium state. At this stage, the process acquires completely different features and proceeds under completely different conditions than in the evolutionary stage.

Degradation occurs when there are no preferable choices. Whereas at the evolutionary stage exploration is purposefully organized, at the degradation stage, using an exaggeration, it is "disorganized."

In concluding this chapter, we would like to define where the exploration process fits into the system of living and inanimate nature. These inanimate systems within self-organization processes should also be assigned to living nature.

There are selection principles of implementable movements in nature (movements being defined, in a broad sense, as the change in or interaction with material objects) out of numerous "conceivable," "virtual" movements. Among the first group of principles are laws of conservation (of momentum, of energy, of mass). If movements do not comply with these laws, they do not make physical sense. The next principle is that of stability. According to this principle, only stable movements are implemented out of all possible movements. Unstable movement forms, if they emerge at all, rapidly disappear. The third principle is introduced by the second law of thermodynamics: out of numerous trajectories, only those are realized for which entropy does not decline. As already discussed, self-organization processes play an important role in the minimum energy dissipation principle. Only those with virtual movements are realized, for which energy dissipation is at a minimum. This principle distinguishes the most economic method of movement in terms of energy expenditure.

These combined principles result in the following: in inanimate nature, only those processes that are the least destructive with respect to the existing organization are realized. This creates an image of a stable physical world governed by the deterministic laws of dynamics. Such a world is a world without evolution.

On the other hand, a characteristic property of living systems is instability, demonstrated by the tendency to not only conserve, but also destroy, the existing organization. They are also characterized by the property of self-organization and the emergence of coherent dissipative structures. Live nature evolves by way of sequential bifurcations. The state of equilibrium for living systems is death. These systems evolve in time, where evolution is understood as a process of the emergence of structures and a continuous increase in the complexity of systems.

The reserve accumulation is also an evolutionary process (at the first evolutionary stage). Its similarity with the development of biological systems is in that it is self-organizing, goes through bifurcations, is directed by the information that is generated, and acquires value in the process of exploration. As with biological systems, a concept of goal is applicable to the reserve accumulation process. It is evolving in order to

achieve a certain goal. The most economical, efficient method of achieving the goal is preferable among all possible avenues. The laws defining the development of the oil and gas reserve accumulation process at the evolutionary stage are similar in nature and are opposite to those of living systems. Evolution of the reserve accumulation process can be considered a reversed biological evolution, or "antibiological" evolution.

Biologic evolution results in a more complex structure and a greater diversity of objects. Conversely, exploration is conducted so that the largest fields are discovered as soon as possible. This simplifies the structure of undiscovered potential to the maximum and impoverishes the diversity of targets. The final goal of exploration is the total uniformity of targets and the structure of undiscovered potential as opposed to their initial drastic non-uniformity. In this sense, evolution of the exploration process is (1) the evolution from order to chaos, (2) the evolution resulting in the destruction of structures, and (3) the evolution in compliance with the second law of thermodynamics, which leads to the most probable state of equilibrium.

Biologic evolution also results in a more complex structure and functions of organisms. Structure and functions are related. Functional organization of exploration is also related to the structure of resources. However, as structural uniformity grows with the discovery of large fields, organization of exploration loses its goal, the effort "dissipates," functional associations weaken, and control becomes progressively more difficult. In other words, the main function of control (the selection of the first priority plays and targets) becomes more simple (or does not make sense). Organization of the process changes to "disorganization."

Biological evolution does not have a limit. This is possible because biological systems maintain themselves far from the fatal state of equilibrium. In order to achieve this, they work continually. The process we are studying also evolves at the expense of continuous exploration work. The goal of this work, however, is quite opposite: to drive the process to the state of equilibrium (the most important fields are discovered first). Inasmuch as the resources are finite, the reserve accumulation process is also finite.

The evolution of the reserve accumulation process at the degradation stage is different. In this case, a similarity with inanimate nature is appropriate, even with the movement along deterministic trajectories according to the aforementioned principles of path selection.

Thus, if in nature the animate emerges from the inanimate, then in exploration, evolution procedes in the opposite direction: the traits of the "animate" are replaced in time by the traits of the "inanimate."

The discussion in this chapter shows that the opportunities to describe and study the reserve accumulation process are far from being exhausted. The similarities with physical phenomena present a hope that in the future

it will be possible to construct a physical model of this process, which will allow us to use physics to study its most subtle facets. The purpose of this chapter was to outline this perspective.

Evolution of the reserve accumulation process as described is due to a clear non-uniformity of the initial resource structure and incomplete information on which the decisions are based. If all the fields and their locations were known from the very beginning of exploration, it would be conducted based on different principles, and the reserve accumulation evolution would be governed by different laws. Therefore, the development and improvement of information-gathering techniques will guarantee making exploration and, especially the most expensive tool, drilling, more cost-effective. This implies a development and improvement of forecast techniques in the broadest sense, including geophysical methods. This is the major factor in refining the exploration process organization.

Discrepancies between the deterministic and stochastic approximation may be explained by the inertia of exploration. The use of information acquired in the process of exploration, begins to be significantly delayed when it is no longer new (when it is possible to figure out its average level over a certain time interval). This delayed reaction in using the obtained information negatively affects the efficiency of exploration. This may be associated with an overcautious attitude and an unwillingness to take risk until the exploration situation is completely clear. This also may be associated with the delays connected with processing and perceiving the information. This can be remedied by an acceleration in the processing of information and in a timely use of the obtained results.

PART II

Evolution of Oil and Gas Field Discoveries

In Chapter 5, we emphasized the interconnections between the history (evolution) of the reserve accrual and the sequence of discoveries. Both of these phenomena reflect intrinsic patterns of the reserve accumulation process. The patterns of the reserve accrual are made obvious through a dynamic description (trajectories, or curves). The evolution of discoveries is described using different techniques. The reason for this is that in the process of exploration not only the structure of already discovered fields, but also of those yet to be discovered, changes. This is why it is necessary to describe and study the evolution of the structure of discoveries in time.

Apparently, patterns in the evolution of discoveries depend on the total number of fields present in a region and on the initial resources distribution among the fields of different sizes. For this reason, a proper examination of discoveries should precede the study of initial concentration of resources in the fields.

CHAPTER 6

Size Distribution of Oil and Gas Field Reserves: Its Formation Mechanism

This chapter addresses the patterns governing the total number of fields in an oil and gas region and the distribution of the ultimate potential among fields of different sizes. This is important because it relates to the discovery evolution study of different sized fields, which is presented in Chapter 7. It is obvious that the structure of discoveries can be correctly identified and interpreted only if we have a clear idea of how many and what size fields are expected in the region at the outset. Only then can we judge how effective the real system is in the sequence of discoveries.

The subject of this chapter is of great importance for solving problems of forecasting and resources evaluation for any region and, in particular, for forecasting the number and sizes of yet undiscovered fields. This forecast, in turn, is needed for the selection of major objectives in the region's development and the long-term planning of exploration: the number and size of forecast fields determine exploration efficiency and strategy and, hence, the efficiency of the control[1] decisions.

Besides having a purely practical application, this problem has an independent theoretical significance related to the problem of the formation and geographical distribution of oil and gas fields. At the center of this problem are the geological conditions leading to the formation of hydrocarbon accumulations. The proposed formation concepts make it possible to recreate the formation conditions for a particular field, or group of fields, and for the hydrocarbon accumulation process in these fields.

The issue of the formation of hydrocarbon accumulations has a long history of investigation with substantial achievements. The other side of the problem, namely, the distribution of reserves among different size fields, is much less studied and its solution is much less successful. There is no single theoretical concept of the field size distribution, although

theoretical research into this domain could substantially enrich the knowledge already accumulated and lead to unexpected results. Methodological mathematical study then becomes inseparable from the subject study by way of preceding, justifying, and directing it. It should result in the creation of a formation concept for an aggregate of fields as a whole system with its own properties. The non-uniformity in the field size distribution should be a natural outcome of the formation environment for such a system and not of the separate fields. A system approach is the only way to investigate the field size distribution over a region of interest.

Numerous publications are devoted to the study of the reserve concentration in different size fields. These publications analyze the data but are almost entirely lacking in theoretical support. Scientists all agree that the number of small fields is much greater than the number of large ones, and that a substantial percentage of oil and gas resources is concentrated in the large fields [53]. It is established that the field size range is within 4 to 5 orders of magnitude [46].

Scientists disagree on the amount of resources associated with small and large fields. Some believe that the majority of resources are concentrated in giant and large fields, whereas others maintain that a substantial share of the resources is present in small fields [41, 42]. The latter is currently the majority viewpoint. The problems of the relations between small fields, their minimum size, and their number are also debatable [24]. This is understandable given the lack of reliable information concerning these fields, which are always poorly studied.

Finally, there is no agreement on whether the field size distributions are similar or different in various regions. Some believe that each region possesses its own characteristic size distribution [41]; others hypothesize that there are no substantial differences [24]. Clearly, due to the fact that not all fields are appraised and information concerning small fields is minuscule, these questions cannot be answered based only on the available data. For this reason, it was necessary to model field size distributions with different assumptions concerning the nature of the formation process.

It has become traditional to use mathematics for describing field size distributions. Some models have been proposed, analyzed, and tested. These models, however, were constructed (or rather borrowed) without any interpretation of their substance, which would describe the geological concept of distribution. A theoretical foundation is important because without it, based on practical data only, the problem of field size distribution cannot be solved. The controversy surrounding this problem and the scattering of proposed solutions are, indeed, a result of insufficient attention to its objective, substantive analysis.

Currently, there are several different approaches to the modeling of field size distribution. Our study is different in that the models are all based on the concept of the process forming a pattern in this distribution.

The numerous possible descriptions for patterns of oil and gas field distribution by size can be encompassed using three approaches. The first approach is a probabilistic description using a distribution function. The share of fields of a particular size within a size range is interpreted as a frequency. The probability function density is associated with sets of fields. Correspondingly, the difference in fields is explained by some generating probabilistic mechanism. This is the most common approach and is associated with statistical studies of the field assemblages within a certain region.

The second approach is based on the relationship between the sizes of adjacent fields in their sequential sets. Correspondingly, the patterns in the structure of these sets are studied.

The third approach substitutes the rank concept for the frequency concept. It analyzes the dependence of the field size on its location (number) in a sequential set. An analysis of this kind may be called structural, as opposed to statistical.

Both the second and third approaches are based on the concept of a regular intermittent change in the field sizes. All three approaches are presented in this book.

Formulation of the Problem

Typical features of empirical field distributions are currently well known, although they are not fully understood. These features are described here. The curve of the field reserve distribution has a long tail. This is caused by the presence of large and giant fields that have reserves significantly exceeding the reserves of other fields. Although the number of such fields is low, they contain a significant part of the ultimate potential. This could be called the reserve concentration phenomenon, which is characterized by the tail portion of the field size distribution. Another feature, pertaining to the part of the distribution to the left of maximum point, is its significant asymmetry. Most fields are not concentrated around some average size with a certain scatter at the edges. The curve has a peak sharply biased toward the small and very small fields. It is quite possible that this distribution is amodal (devoid of mode).

This statement may be disputable because all of our observations are based on already discovered fields. They do not, however, constitute a representative selection. They are not observations chosen at random, and there is no means of obtaining any other observations. Thus, the presence of a mode in the discovered field size distribution (which happens quite often) is not significant in this situation. This may easily be the result of the exploration strategy and organization, which provide for a non-random

selection of observations, and by the discovery of large and giant fields at early stages of exploration.

With the progress of exploration, the share of small and very small fields grows and that of the large fields declines. Many scientists believe that this is confirmation of the amodal nature of the field size distribution. It must be noted, however, that this evidence is poorly supported. The mode may be strongly offset to the left, toward the small and very small fields, so that its presence (or absence) may be proven only after the discovery of a great number of small fields. Such extended data, however, is so far unavailable in any petroleum region of the world. However, there are those that believe that for the well-studied U.S. basins "more that 30 years ago the amodal nature of the field size distribution was established with a significant confidence. This trend became even stronger over the recent 30 years" [24, p. 29]. If we assume that the number of the smallest, abandoned fields (F-class fields) is smaller than that of the next larger E-class "then the amodal distribution nature of the targets bearing commercial interest is recorded with confidence" [24, p. 29]. This may be called the reserve dispersion over a great number of small and very small fields.

The bulk of the fields are small and a significant (or major, according to some scientists [50]) part of reserves is concentrated in a small number of fields. Thus, there is: (1) "size" growth in the tail part, (2) "number" increase in the left portion of the distribution, (3) reserve dispersion over a great number of small fields, and (4) reserve concentration in the large and very large fields. This phenomenon may be called the concentration and dispersion phenomenon (after Yablonsky), which is a qualitative characterization of the oil and gas field size distribution [56].

The work by Kontorovich and Demin represents the most typical philosophy in field size distribution modeling [24]. It appears that all scientists adhere to this philosophy, which can be formulated as follows: the form of the distribution "could have been theoretically derived by constructing the theoretical model of the hydrocarbon generation and destruction processes. This solution, albeit promising, is so far unavailable. Therefore, one can only check an actual distribution, the factual data against different theoretical distributions" [24, p. 28]. Based on this philosophy, scientists approximate the empiric field size distribution using one of the known probability distributions. The problem of the distribution function format has long been discussed, so far with no consensus. The arguments in favor of a particular probability distribution function format are based exclusively on empiric verifications of the studied actual (or perceived) field size distribution in a certain region.

Currently, two formats of the distribution function are most commonly accepted. The first one maintains that the reserve size is distributed log-normally [2, 16, 46]. The second one contends that the correct distribution

is the Pareto distribution [24, 42, 53]. The latter became especially popular during the last decade. The major difference between these distributions (the pivot of the arguments) is that the log-normal distribution has a mode, whereas the Pareto distribution is amodal, with the probability density decreasing continuously with the increasing volume of reserves. The proponents of the former concept maintain that it is satisfactory in describing the field size distribution in mature regions. The proponents of the latter assert that a good fit of the field size distributions (log-normal) in the individual regions is caused by a distortion of the actual distribution due to the non-random nature of observations in the process of exploration.

Attempts were made to approximate the field size distribution using the exponential and the Weibull distributions [1, 35]. The exponential distribution was rejected after checks against the well-studied oil and gas provinces or parts thereof. Nalivkin et al. used the gamma-distribution as a model for the field size distribution [36]. Thus, we see a wide range of distributions have been used and were believed to adequately describe the actual distributions, or at least able to describe them after the discovery of all small fields.

The following is worth mentioning regarding the fit between the theoretical and the observed distributions. Observations used for the construction of a frequency distribution must be extracted through a randomized procedure covering the target under study (for instance, a province) in its entirety. If this is accomplished, the conclusions based on selective observations may be attributed to the entire target. As mentioned earlier, when studying field size distributions, the observations do not constitute a representative sample. Under these circumstances, testing a hypothesis of the field size distribution function format requires a procedure different from the conventional one. No available publication indicates how the testing procedure was developed under given circumstances; at best, hypothesis testing is warranted. We believe that the particulars specified above were not taken into account.

An important problem is the coverage of the entire target by observations, and whether or not the fields under analysis belong to the same distribution. What is the area (the regional geologic unit) over which the same distribution can be used? In other words, what is the area over which the field formation follows the same patterns? The answer to this question depends on the purpose of the modeling and, hence, on the degree of detail required in the corresponding system description.

As an example, let us assume we are interested in the life expectancy of humans. If our interest lies in the entire human race and is associated with the study of the biologic nature of humans, all people may be considered within the same distribution in the belief that nature (biologically)

predetermined the same longevity for all of them. If, however, the study is of the longevity in different geographic or social environments, it is natural to attribute the same distribution to only those people living in a similar environment in terms of the chosen criterion.

Most interesting from a geologic viewpoint is an oil and gas basin or province (see Eremenko, 1991 [69]). It represents a unique natural system with certain characteristic features determining the basin type. The origin may distinguish it from other basins. The distinguishing characteristics include: the areal extent, the sedimentary volume, the thickness and number of oil-saturated formations, the total sedimentary thickness, the depositional speed, the number of major regional sedimentation gaps, and so forth. These are the most general conditions typical of the entire basin (as a singular system). This is why it makes sense to apply the field distribution to a hydrocarbon basin or province. This determines the degree of complexity of the description. The other, more detailed factors that determine the subdivision of the province into areas, sub-areas, and the like, are not taken into account.

It appears that the observations in specific studies by various authors cover only part of the target (in cases where it is indicated that the hypothesis was examined using the fields only over a certain part of the basin). In a case like this, an additional condition should be imposed specifying which target (province, oil and gas area, etc.) is considered universal (general). Unfortunately, crucial reservations of this sort are usually not presented. The result is the oil and gas field size distribution function format is an open question even in the cases where it is simply a matter of an approximating function. This is very important because even the strongest experimental confirmation of a hypothesis is never sufficient since hypotheses are always open to further examination. Any statistical testing technique can only confirm that a hypothesis does not contradict the observations, and a single negative test is a sufficient cause to discredit the hypothesis. A simple test of whether a hypothesis is adequate or not does not solve the problem of model selection, because different theoretical distributions may be forced to fit the observations equally well.

It is important to note that a description of the reserve relations among fields using a language of distribution functions is appropriate only in the case where the distribution under investigation is probabilistic. This is possible if the field reserves R are considered independent random variables governed by the same distribution pattern. In such a case, the R value is distributed continuously and its distribution function $F(Q)$ (this notation for the variable of integration must be different from the upper limit) is expressed through the corresponding probability density $f(Q)$ as follows:

$$P(R < Q) = F(Q) = \int_{-\infty}^{Q} f(u)du$$

where R = reserves (random variable); Q = argument of the distribution function $F(Q)$; u = variable of integration; and $P(*)$ = probability of the event (*).

In reference to the above discussion, the field size variations must be associated with the presence of factors classified as random. Only in this case does the field randomly acquire a certain size. This is not self-evident if applied to oil and gas fields. Therefore, the legitimacy of this approach must be proven first. Unfortunately, there are no publications in which this has been attempted.

The concentration and dispersion phenomenon mentioned earlier does not allow one to assume that the difference in field sizes is determined by some random "uniformly infinitesimal" factors. If the field size was determined by purely random factors, their scatter would be more uniform. In such a case, the distribution under investigation would not have stood apart from the conventional domain of probabilistic distributions called "normal" or "Gaussian": the binomial, the Poisson, and all others that may be approximated by the normal Gaussian law. The field distribution is non-Gaussian because the random value numbers are more scattered than under the "Gaussian" law. Also, the probability of larger than average deviations is greater, and large groups of small values and, conversely, single values of the same order as the sum of the rest, are possible. The concentration and dispersion phenomenon forces us to rule out a Gaussian-type field size distribution and conclude that the distribution is based on a totally different law.

In the search for such a law, we will deviate from the conventional phenomenological approach to the construction of a mathematical model of the field size distribution. Such a conventional approach consists of an analysis of empirical information and "forcing" a corresponding distribution function to fit it. Most important is that the frequency distribution has a genetic meaning. The goal is not to postulate a theoretical distribution law as an empiric fact but to infer this law theoretically, based on simple and well-known geologic concepts. In other words, a distribution function must be derived, inferred from a certain stochastic model describing field formation of certain sizes, and thereby assigning a substantial interpretation to the distribution function. Such theoretical consideration of the field size distribution laws may result in an expansion of geologic studies to encompass the regularities in the formation patterns of different field size populations.

Three different approaches to the construction of the required mathematical model may be identified. All of them consider field populations

of different sizes belonging to one distribution as a complete natural system, which possesses either some specific properties or some specific behavior. The study of such systems is predicated on the selection of a cognitive philosophy. One such philosophy is striving to find the cause-and-effect relation between phenomena. The second is searching for the meaning of the events as expressed in the goals. The third is based on the idea that a system's behavior cannot be reduced to causes and goals, but has its own internal meaning.

The first philosophy calls for an analysis of the formation process of the population of different size fields (i.e., an analysis of the system's dynamic properties).

The second philosophy accepts the proposal that some integral harmony of form can be attributed to the system of fields. This is a result of the proposal that the system's behavior is an optimum implementation of a set goal. This harmony of form is reflected in the quantitative relations between field reserves. These relations reflect the final result which is achieved when the goal is reached (and not in the process of movement toward the goal). Thus, they elucidate statistical properties of the system.

The third approach is also statistical. It analyzes the function describing the system's final status through the characteristics of its intrinsic properties.

Therefore, the first approach calls for the construction of a process that models the distribution function desired. The second and third approaches can be reduced to the study of principles, which the function is based on. Each of these approaches reveals only one facet of the phenomenon. Combined, they complement each other and provide a more detailed characterization of the phenomenon.

It is convenient to begin with the third approach.

Collection (Population) of Fields as a System Possessing Specific Statistical Properties

In this section we will infer the probability distribution in the event that a separate field, in the process of its formation, reaches size Q (i.e., with reserves Q). For obvious reasons, this distribution may be considered a distribution over the field population. Our modeling is based on a principle we call the field's accumulating capacity growth principle. This is a selection principle that allows us to narrow-down the multitude of possible solutions and to choose only one. On one hand, this principle is a consequence of the aforementioned concentration and dispersion phenomenon and, on the other, of a phenomenon not yet discussed.

The following phenomenon is observed in all oil and gas basins. Small fields are uniform in their size, with very little difference between

them being insignificant. As the field size increases and the fields become medium and large, the disparity in the size of reserves increases and the fields become more non-uniform. Small fields in the population may differ by a fraction of one million tons. At the same time, two large individual fields next to one another may differ by hundreds of millions (or even billions) of tons of reserves. For small fields (the left part of the curve, or the beginning of the distribution), the presence of fields similar or close in magnitude of reserves is the rule. For large fields (the right part of the curve, or the tail of the distribution) this is the exception. This phenomenon is similar to the relationship among the average mineral concentrations in a rock with variances in these concentrations, and each component with its own concentration variance. This is a common explanation for asymmetry in many frequency distributions observed in geochemistry. A basic law of geochemical processes was also formulated that is similar to the concentration and dispersion phenomenon and represents the commonality of low concentrations and rarity of high concentrations of dispersed unstable elements.

These phenomena reflect specific statistical properties of a field population belonging to the same distribution. Based on these phenomena, the following principle may be formulated. If in the process of field formation a field's size reaches the value Q, then the greater the Q the lower the probability of its further increase by a small value dQ (i.e., the reserves of the fully formed field are not greater than $Q + dQ$).

The concentration of hydrocarbons within a trap grows much like a snowball rolling downhill. The farther the ball rolls, the greater its size. The greater its size, the less sensitive it is to any obstacles it may encounter along the way. Thus, there is a great chance that the ball will continue to roll and grow in size until it meets with a significant (rare) obstacle.

The same is true for a field. The larger the field, the less its growth (accumulation of hydrocarbons) is subjected to various instabilities. The site of the field becomes a focus of major formative processes. The field is in the path of major energy and hydrocarbon flows. The larger the field, the higher the hydrocarbon accumulation rate. For example, the population of cities grows in the same manner. The larger the population, the more attractive the cities become with a larger flow of people toward them. The same occurs with capital accumulation. The larger the capital, the smaller the probability that it will stop growing. A large capital is more stable with respect to various "disturbances." The rate of capital accumulation increases with capital growth.

This explains the peculiarity of oil and gas field size growth and, from the probabilistic viewpoint, what separates the process we are examining from the other processes. From this viewpoint, two types of processes may be

identified. One type includes the processes we have just examined. They may be called the accelerating processes. We refer to this principle as the "field's accumulating capacity growth principle."

The second type includes the processes occurring with deceleration. Mechanisms impeding their progress appear during their progression. This type includes, for example, the processes enabling vital activity of living organisms. As life proceeds, unfavorable elements accumulate in the organism which lowers its activity. For this reason, a totally different phenomenon is observed regarding life expectancy: the longer a person lives, the higher the probability that this person will die soon. This appears to be a result of an opposing principle that may be called "the principle of lowering viability."

The accumulating capacity growth principle may be expressed as follows. The probability of a random field size value R being equal or larger than Q is $1 - F(Q)$, and the probability of the field having a size between Q and $Q + dQ$ (where dQ is a small increment of Q) is $f(Q)dQ$. The conditional probability of a field reaching a size between Q and $Q + dQ$ (after which its formation ceases), provided that the field has already reached a size equal or larger than Q, is:

$$\frac{f(Q)dQ}{1 - F(Q)} \tag{6-1}$$

Let us suppose that this conditional probability can be expressed as $\mu(Q)dQ$. The $\mu(Q)$ is called the intensity function in probability theory. Clearly, the selection principle we are currently analyzing must be formulated as follows: the intensity function corresponding to the oil and gas field distribution function must be a function of the reserve size, and this function must continuously decrease with the growth of Q. For normal, log-normal, or exponential distribution functions, the intensity function $\mu(Q)$ does not satisfy this condition. One mandatory condition is a monotonous decrease of the intensity function $\mu(Q)$ with the growth of Q. Another condition is associated with determining the $\mu(Q)$ value decline limit. It is natural to assume that the intensity function asymptotically tends to zero. Otherwise, the conditional probability under discussion would be limited by a certain value, which contradicts the very substance of the process we are analyzing. When the Q value is great, the probability of encountering two similar size fields Q and $Q + dQ$ is negligible.

The distribution function for the assigned intensity function may be obtained from the following equation:

$$1 - F(Q) = [1 - F(Q_o)]\exp\left[-\int_{Q_o}^{Q} \mu(z)dz\right] \tag{6-2}$$

where Q_o is an arbitrary value of Q.

According to this expression, the limit of possible intensity function values decreases with the growth in Q and is defined by the condition that the integral $\int_{Q_o}^{\infty} \mu(z)dz$ diverges. This condition imposes limitations on the format of the intensity function $\mu(Q)$ and, hence, narrows the number of possible classes of functions.

Thus, the distribution function of interest must belong to the distribution function class with the intensity function $\mu(Q)$ satisfying the following conditions:

1. $\mu(Q)$ is a monotonously decreasing function;
2. $\mu(Q) \to 0$ when $Q \to \infty$;
3. $\int_{Q_o}^{\infty} \mu(z)dz = \infty$

Out of the simple and most common functions, only the exponential function satisfies these conditions. These conditions exclude the law of linear, exponential (base e) decline or other similar laws.

Two solutions of $F(Q)$ correspond to the selected format of the intensity function:

$$\mu(Q) = aQ^{-b} \quad (a > 0,\ b > 0) \tag{6-3}$$

These solutions correspond to different limits of the exponent b change. The limits of change for the exponent b are defined by the above three conditions and by the fact that $F(Q)$ is a distribution function (i.e., $0 \leq F(Q) \leq 1$). Thus, the limits are restricted; namely, $0 < b \leq 1$. Exponent b may not assume any other value.

At $0 < b < 1$, a distribution function

$$F(Q) = 1 - \exp\{-[a/(1-b)]Q^{1-b}\} \quad (Q > 0,\ a > 0,\ 0 < b < 1) \tag{6-4}$$

where $\exp\{x\} = e^x$ corresponds to the intensity function format $\mu(Q) = aQ^{-b}$.

At b = 1, the distribution function

$$F(Q) = 1 - (Q_o/Q)^a \quad (Q \geq Q_o,\ a > 0) \tag{6-5}$$

corresponds to the intensity function $\mu(Q) = a/Q$.

Therefore, theoretical considerations based on the principle of growth of the accumulating capacity of a deposit result in two field size distribution functions. The following is an analysis of their properties.

The probability density $f(Q)$ of the Equation 6-4 distribution is expressed by the following equation:

$$f(Q) = aQ^{-b}\exp\{-[a/(1-b)]Q^{1-b}\} \tag{6-6}$$

The distribution is a Weibull's distribution with certain limitations imposed as its parameters. The Weibull's distribution can be presented as follows:

$$F(x) = 1 - e^{-\lambda x^n} \quad (x > 0,\ n > 0,\ \lambda > 0)$$

and its density is expressed as $f(x) = n\lambda x^{n-1}e^{-\lambda x^n}$.

Inasmuch as $n = 1 - b$ and $\lambda = a/(1-b)$, the distribution we derived is the Weibull's distribution, with the n parameter within a $0 < n < 1$ range. This shows a relationship between the Weibull's distribution parameters and the intensity function $\mu(Q)$ parameters. The Weibull's distribution median using these parameters is expressed as $\tilde{x} = [(n-1)/(\lambda n)]^{1/(n-1)}$. Thus, the distribution 6-4 does not have median. Mathematical expectation $M(\chi)$ and variance $D(\chi)$ of a random variable χ are determined as follows using the Weibull's distribution:

$$M(\chi) = \lambda^{-1/n}\Gamma\left(\frac{1}{n}+1\right); M(Q) = \left(\frac{a}{1-b}\right)^{-1/(1-b)} * \Gamma\left(\frac{1}{1-b}+1\right) \tag{6-7}$$

$$D(\chi) = \lambda^{-2/n}\left\{\frac{2}{n}\Gamma\left(\frac{2}{n}\right) - \frac{1}{n^2}\left[\Gamma\left(\frac{1}{n}\right)\right]^2\right\}$$

$$D(Q) = \left(\frac{a}{1-b}\right)^{-2/(1-b)}\left\{\frac{2}{1-b}\Gamma\left(\frac{2}{1-b}\right) - \frac{1}{(1-b)^2}\left[\Gamma\left(\frac{1}{1-b}\right)\right]^2\right\}$$

where $\Gamma(t)$ is a gamma-function defined by $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx$.

As mentioned before, the Weibull's distribution has been used by some scientists for describing field size distribution. We can now maintain that this distribution's properties satisfy the specifics of formation of deposits using our principle only at a certain value of one of the distribution's parameters (namely, the n parameter). Thus, we are no longer discussing the Weibull's distribution in general, but a special case where the distribution does not have a mode.

The second distribution we derived (Equation 6-5) is the Pareto's distribution, which is commonly used for describing field size distribution. Probability density of this distribution has the following format:

$$f(Q) = \frac{a}{Q_o}(Q_o/Q)^{a+1} \tag{6-9}$$

The distribution's characteristic parameter a is the proportionality coefficient between the intensity $\mu(Q)$ and $1/Q$. The distribution is amodal. The distribution moments are given by the following expressions:

$$M(R) = \frac{a}{a-1}Q_o \quad (a > 1); \; M(Q) = \infty \quad (a \leq 1) \tag{6-10}$$

$$D(R) = \frac{a}{(a-1)(a-2)}Q_o^2 \quad (a > 2); \; D(R) = \infty \quad (a \leq 2) \tag{6-11}$$

Therefore, at certain a values the first two moments of distribution diverge. Their infinity value technically reflects the non-Gaussian features of the Pareto distribution. Infinity is associated with the very slow convergence of this distribution. Random values corresponding to this distribution do not concentrate around the mathematical expectation and may be located so far from the mean that determination of the mean becomes meaningless. As shown later, often $a = 1$ and rarely $a > 2$. Thus, the infinity of mean values and variances, or a pathological phenomenon from the viewpoint of those accustomed to "normal" distributions, is an important and frequently occurring feature of oil field distribution.

In the absence of moments, the description of a distribution using empirical values for the mean and variance becomes unreliable in precisely the same way as the estimate of the moments using the least square technique. This technique is unreliable due to the infinite variance. In a case like this, the only reliable method is to describe the entire distribution, using the median and other quantiles instead of means and variances. The estimate of the a parameter becomes especially important here. This explains the difficulty of forecasting the size of undiscovered fields.

Thus, we have shown that an intensity function of the format $(Q) = aQ^{-b}$ (at $0 < b \leq 1$) corresponds to the conditions of formation of a field expressed in terms of its accumulating capacity growth principle. An intensity function of this format may have two distributions: the Weibull's type (Equation 6-4) and the Pareto type. The Pareto distribution corresponds to the $b = 1$ case. In this sense, the Pareto distribution may be considered a limiting case for the Weibull's distribution of the

Equation 6-4 format when $b \approx 1$. When $b \approx 1$, the Equation 6-4 distribution becomes indistinguishable from selected data of the Pareto distribution. It is worth mentioning that the Pareto distribution is the distribution with the most rapidly decreasing intensity function $\mu(Q) = aQ^{-b}$ (at $0 < b \leq 1$). This intensity function tends to its zero asymptotic value at a maximum speed proportional to $1/Q^2$. This is precisely what causes an extremely slow convergence of the distribution which, in turn, causes the infinity of the variance and, often, of the mean.

It is important to note that at $b = 0$ the distribution becomes exponential. This means that when b is small (close to 0)—when the intensity function $\mu(Q)$ declines very slowly—the Weibull's distribution is indistinguishable from the exponential distribution. In such a case, it is quite possible when working with empirical data, to "establish" the field size distribution as exponential.

Thus, the full range of the b parameter ($0 \leq b \leq 1$) covers a chain of distributions: exponential, the Equation 6-4-type Weibull, and the Pareto. The first and last distributions in this chain correspond to the lower and upper limits of possible b values. The balance is described by the Equation 6-4-type Weibull's distribution. At $b \sim 0$ and $b \sim 1$ it is impossible to distinguish the Weibull distribution from its neighbors on the left and right, respectively. This is the cause of the controversial results in the studies based on the simple examination of the oil and gas field size distribution adequacy hypothesis using borrowed theoretical distributions. Until now, the meaning of these results was not clear.

Thus, based on this principle we established a family of distributions connected by some transitions and forming an orderly (ranked) set. In particular, this establishes similarities and distinctions among the well-known distributions (the exponential, Weibull, and Pareto) used by scientists as competing means for the description of oil and gas field size distribution. It is clear that due to their intrinsic properties, some of these distributions may be distinguished from one another only for certain values of Q. These values are usually small, over areas where there is an insufficient amount of observations. (Of course, this is true only if one does not take into account the process generating a distribution and relies only on observations.) In this sense, all current debates concerning the type of distribution function are pointless.

Once again, a simple test of the model-to-observations adequacy hypothesis (especially under specific circumstances of observation gathering, as discussed previously) does not solve the problem. However, such an examination is needed to determine whether or not the input concepts used for model construction fit the real world. A little later we will investigate a fit of the inferred models to observations. At this point, it

is sufficient to indicate that a qualitative peculiarity of all inferred distributions is their lack of mode. This feature does not contradict the data; to our knowledge, not a single petroleum province in the former USSR, or anywhere else in the world, has a firmly established mode in its field size distribution.

Process of Forming a Field Population (Set) as a Randomized Process of an Individual Deterministic Growth

The following sections suggest a different approach to the derivation of the field size distribution function. This approach is based on the construction of a formation model for a field set. Examination of the field formation process means that our attention is now concentrated directly on a causal description of a field set. We have already mentioned that the frequency distribution has a genetic nature. It is natural for a stochastic model of the field formation process to be based on concepts associated with a stochastic scheme (skeleton of the process). Such schemes are developed, for instance, based on the concept of geochemical processes when studying concentrations of chemical elements in rocks. They form the basis for the stochastic models from which the distribution functions are derived [8]. However, no schemes have been developed with respect to oil and gas field formation.

Based on the oil and gas formation specifics, two totally different schemes may be proposed. One such scheme is based on deterministic concepts and another one, on probabilistic concepts.

Deterministic Process of a Deposit (Field) Formation

In the process of its formation, each individual field progresses through certain evolutionary stages. The field size, or the volume of accumulated hydrocarbons, changes as it evolves. The purpose of this section is to construct a mathematical model of the hydrocarbon accumulation process during field evolution. Such a model should satisfy the natural hypotheses of the field formation and should also result in a law of field size (size of the reserves)^2 distribution. Obviously, we will be dealing with simplified models of the formation process that will provide a visual concept of the field evolution as it moves from one stage to the next.

The field formation processes include sequential (or simultaneous) processes of creation and destruction. These processes interact with each

other and determine the growth rate of the volume of hydrocarbons accumulating in a trap. Let us assume that the mechanism of this growth is described by a differential equation with constant coefficients (within a region or area in which the fields are attributed to a single distribution). Let us also assume that the hydrocarbon accumulation rate accelerates with time and that the growth process is non-linear. Non-linearity is a general rule in modern chemistry, hydrodynamics, and biology. This rule may be extended to geologic processes that, in large measure, develop spontaneously. There is no place for linear interrelations here. Non-linearity of the growth models is a matter of principle. Non-linear equations describe qualitatively different phenomena than linear ones.

In practice, non-linearity may show up in an endless variety of ways. In the future, we will limit our examination of differential equations to only the most common type. Therefore, we can assume that the growth of the hydrocarbon volume accumulating in a trap is controlled by deterministic rules. These rules are described by a corresponding differential equation and by its solution, which is the equation of hydrocarbon accumulation in time. Field reserves are determined by the effective pore volume with conditions favorable for the hydrocarbon accumulation. They are also determined by the volume of oil generated and by the volume which migrated, reached the trap, and remained there. Thus, for each field we will consider a hypothetical system called the "trap forming-generation-migration-accumulation system," which evolves in time non-linearly. An indicator of the system's evolution is the size of reserves $Q(t)$.

The hydrocarbon mass dQ accumulating in traps during a small time interval dt depends on two antithetical processes: hydrocarbon accumulation and dispersion. If we consider the formation of hydrocarbon accumulations as a single, complete, and continuous process, then a balance equation can be formulated:

$$dQ/dt = A(Q) - B(Q) \tag{6-12}$$

where $A(Q)$ is the hydrocarbon accumulation operator and $B(Q)$ is the hydrocarbon dispersion operator.

The format of the second term does not present any problems. It is natural to assume that a hydrocarbon mass dispersed from a trap(s) is proportional to the amount of accumulated hydrocarbons:

$$B(Q) = d(t)Q(t) \tag{6-13}$$

where $d(t)$ is the dispersion coefficient determined by a number of conditions such as the lithology of seals, the hydrogeological regime, and the tectonic activity in the basin.

The hydrocarbon accumulation operator may also be determined by the affecting factors. Apparently, the hydrocarbon mass in a trap may be increased from three hydrocarbon sources: (1) dispersed hydrocarbons available in the basin's sediments by the time t; (2) hydrocarbons being expelled by the source rocks or input from outside the basin over the time interval dt; and (3) hydrocarbons resulting from the destruction of the previously existing accumulations. Thus, the $A(Q)$ operator may be subdivided into three terms corresponding to each of the above sources:

$$A(Q) = A_1(Q) + A_2(Q) + A_3(Q) \tag{6-14}$$

Assuming that the hydrocarbon mass accumulated in the trap from the first source is proportional to the mass of dispersed hydrocarbons $M_1(t)$, then $A_1(Q) = aM_1(t)$, where it is reasonable to consider the proportionality coefficient a as a function of the accumulated hydrocarbons: $a = a(Q)$. Likewise, $A_2(Q) = b(Q)M_2(t)$, where $M_2(t)$ is the hydrocarbon mass from the second source. It is logical to assume that the hydrocarbon mass from the third source is proportional to $\sum_i B_i(Q)$. Part of these hydrocarbons fills up the other fields. Thus:

$$A_3 = c(Q)\sum_i B_i(Q)$$

Summarizing the above, the field evolution can be represented as follows:

$$dQ/dt = a(Q)M_1(t) + b(Q)M_2(t) + c(Q)\sum_i B_i(Q) - d(t)Q(t) \tag{6-15}$$

This equation provides a generalized description of the formation process of a field. In order to use it for a study of the distribution structure, the specific relations—components of the equation—need to be elucidated. As first approximation, it may be assumed that all Q-dependent proportionality coefficients are proportional to Q. In this case, $a(Q) = aQ(t)$, $b(Q) = bQ(t)$, and $c(Q) = cQ(t)$, and

$$dQ/dt = [aM_1(t) + bM_2(t) + c\sum_i B_i(Q(t)) - d(t)]Q(t)$$

The terms of the above equation (the first cofactor in the equation's right part) are functions of t and characterize the entire oil and gas basin and not a separate field. We will designate the $\varphi(t)$ function, which we call "the capability of the medium to support the hydrocarbon accumulation," as:

$$\varphi(t) = aM_1(t) + bM_2(t) + c\sum_i B_i(Q(t)) - d(t) \tag{6-16}$$

Then the basic equation may be written as follows:

$$dQ/dt = \varphi(t)Q(t) \tag{6-17}$$

Thus, the rate of change in the hydrocarbon mass in the traps of a deposit (field) is determined as a product of the hydrocarbon mass (accumulated in the traps by the time t; that is, $Q(t)$), and the factor we call "the capability of the medium to form fields or to support the hydrocarbon accumulation." Depending on the specific conditions, this factor's evolution in time may be governed by different laws. These laws will determine the specific hydrocarbon mass growth equation in the process of a field formation. In other words, the solution of Equation 6-17 is determined by the specific function $\varphi(t)$.

Specific Mathematic Models of the Hydrocarbon Accumulation Process

Changes in the volume of accumulating hydrocarbons in a field depend on many factors. These factors have been meticulously analyzed in many publications. In regional terms, oil and gas accumulation is greatly affected by the position of oil and gas accumulation zones relative to the regional highs and depressions; by timing and history of the structural evolution; by lithology of the sequences into which the hydrocarbons are expelled and those through which they migrate; by the selective retention of hydrocarbons while they are migrating, and so forth. These factors change and combine in time and determine the changes in the capability of medium to support the hydrocarbon accumulation. They are materialized through this capability as different patterns (laws) of the hydrocarbon mass growth in individual deposits (fields). This transition between the regional and local factors, and the reflection of the regional environment in the evolution patterns of individual fields, are still poorly studied. For this reason, we will limit ourselves to general considerations.

The simplest assumption is that regional processes are balanced in nature and so interconnected that $\varphi(t)$ does not change in time (i.e., $\varphi(t)$ = const). In this case, Equation 6-17 for the mass growth of accumulating hydrocarbons has the following format:

$$dQ/dt = \beta Q \tag{6-18}$$

and the hydrocarbon mass in the field at the time t is:

$$Q(t) = \gamma e^{\beta t} \tag{6-19}$$

We obtained an exponential equation for the growth of the accumulating hydrocarbon mass. In this respect, field evolution is no different than many natural processes, which are also characterized by an exponential development. The function $\varphi(t) = \beta$ is constant in time. Equation 6-18 shows that:

$$\frac{1}{Q}\frac{dQ}{dT} = \frac{Q'}{Q} = \beta \tag{6-20}$$

The Q'/Q ratio is the relative hydrocarbon mass increment per unit of time, or the reserve growth rate. Therefore, balancing the regional environment is expressed as a constant growth rate for individual fields. The condition $\varphi(t) = \text{const}$ results in the conservation of local conditions.

These are not the only considerations. It is believed that the capability of the medium to form fields and to support hydrocarbon accumulation (i.e., $\varphi(t)$), must decline with time. This is facilitated (1) by a natural (due to field formation) decline in the mass of organic matter accumulated in the basin and, possibly, the associated decline in the oil and gas generating potential; (2) by a decrease in migration due to rock deformations (and the resulting deterioration of the reservoir properties); and (3) by a decrease in the accumulation processes due to an increase in the avenues of migration as a result of faulting. These ideas appear to be supported by circumstances associated with local conditions. As mentioned before, the transition from regional to local conditions can be represented by the following equation:

$$\varphi(t) = \frac{1}{Q}\frac{dQ}{dt}$$

The right part of this equation is a description of time changes of the field reserve growth rate. Field reserve growth rate cannot increase with time. Otherwise, it would overtake the generation, migration, and accumulation capabilities of the medium. Most likely, growth rate is constant (as in the case examined above) or is decreasing with time.

Let us assume for the sake of simplicity that $\varphi(t) = \beta/t$ and

$$dQ/dt = (\beta/t)Q \tag{6-21}$$

The solution of this equation is given by an exponential function

$$Q(t) = \gamma t^{\beta} \tag{6-22}$$

The γ parameter depends on the scale of the Q and t axes ($\gamma = Q_o$ when $t = 1$). The exponent β does not depend on scale and represents a dimensionless characteristic of the growth process. Its size should be limited based on the following considerations. It follows from Equations 6-21 and 6-22 that $dQ/dt = \gamma\beta t^{\beta-1}$. This means that, depending on β, the field reserve accumulation speed may grow and then remain constant or decline with time. Let us assume that the hydrocarbon mass increase in the field's traps occurs at an increasing speed. This is one of the basic concepts of the field reserve non-linear growth. Consequently, $\beta > 1$.

Thus, we have established that the hydrocarbon reserve accumulation process in the traps of a single field may be described by an exponential function with base e, or by an exponential function with a different base and the exponent of $\beta > 1$. Both growth models describe accelerated evolution. The case of the exponential function with the base e corresponds to a situation of constant growth rate. The growth according to the exponential function with a different base corresponds to a situation when the growth rate varies inversely with time.

Randomization of the Individual Deterministic Growth Process

The equations obtained above describe the size growth of a single, individual field. It can be assumed that the equation parameters and the final time t_{fin} of field formation are unique and vary for different fields. Their specific values depend on many different factors. Their combination allows us to consider these parameters and the finite time t_{fin} as random values with corresponding distribution functions determined over the field set. In this case, the general Equation 6-17 and the specific Equations 6-19 and 6-21 describe the evolution of the field size distribution structure. The random nature of the mentioned quantities presents an opportunity to juxtapose a certain field size distribution function for each growth model. This is the purpose of the randomization process.

Random Parameter: The Time of Completion of Hydrocarbon Deposit Formation. The evolution of a field as a complex natural system depends on the joint influence of many factors. If even a single factor ceases to act due to random circumstances (i.e., at a random moment in time), the evolution of the entire system is interrupted. Therefore, the laws of probability (random elements, in this case) manifest themselves in that the field reserve size growth ceases at a random time. As a result, the field size variation is determined by the joint influence of deterministic (the laws of growth) and random (duration of the formation process) causes.

Let us introduce the notation $F(t)$ for the distribution function of the random time of completion τ. The specific format of the eventual field size distribution depends on the $F(t)$ distribution function, which is unknown. The information necessary to arrive at a specific proposal as to its format is unavailable. The only way to propose a possible format for the function is to apply the analogies between the field formation process and functioning of technological or biological systems. Based on such analogies, three different suggestions may be given regarding the format of the unknown distribution function $F(t)$.

1. Random variables are often distributed according to the exponential (with the base e) law. Such variables include the random operation time of radio-equipment before the first failure; random time intervals between the sequential occurrences of rare events; duration of busy intervals of a telephone line; the waiting time before the first order in the flow of orders is submitted to a service system; and the lifespan of some materials, devices, and equipment. Using this analogy, one can assume that the same pattern is observed in the distribution of the time of completion for the processes leading to field formation (time before the first failure). In this sense, τ is the duration before the first failure of a complex, natural system forming the field.

The exponential distribution has the following format:

$$F(t) = 1 - e^{-\alpha t} \quad (t > 0) \tag{6-23}$$

and its distribution density is:

$$f(t) = \alpha e^{-\alpha t} \tag{6-24}$$

Parameter α is called the failure intensity. It is associated with the average value of the random quantity τ. The average duration of functioning before the first failure varies inversely with the failure intensity ($t_{avg} = 1/\alpha$). The introduced intensity function is also associated with the failure intensity. In fact, it is identical to it: $\mu(t) = \alpha$. Earlier we subdivided all processes into two groups: those with a positive acceleration and those with a negative acceleration. Accordingly, the intensity function for random time should be a nondecreasing function. This condition is fulfilled for the exponential distribution.

The exponential distribution function of a random time-before-failure for the field-forming system has a corresponding field size distribution function. The latter depends on the type of field size (reserve) growth in time and, hence, on the equation describing this growth. This follows from the fact that the probability density $f_\eta(x)$ of a random value $\eta = \phi(\xi)$ (where ξ is a random variable with the probability density $f_\xi(x)$ and $y = \phi(x)$ is a monotonously changing function) is defined by the equation:

$$f_\eta(x) = f_\xi[\psi(x)]|\psi'(x)| \tag{6-25}$$

where $\psi(x)$ is the inverse with respect to the $\phi(x)$ function.

The derived field size distribution functions corresponding to the reviewed growth equations are represented on line 1 of Table 6-1. They are the same as the earlier-derived Pareto and Weibull's distributions. The latter is in a format 6-4 with a corresponding intensity function $\mu(Q)$, which tends to zero with the growth of reserves Q. Therefore, each field size distribution may be explained by a certain type of growth in time of field reserves, which is caused by certain hydrocarbon generation and accumulation processes. The meaning of the distribution parameters also becomes clear. They are related to the growth equation parameter and to the random time distribution parameter, allowing for a simple and descriptive interpretation.

For instance, the Q_o parameter in the Pareto distribution is the scale parameter γ in the growth equation $Q(t) = \gamma e^t$. Q_o is not the size of the smallest field but the field size at time $t_o = 0$. This value has a genetic rather than an economic (commercial importance of the field) meaning. It is associated with the scale of hydrocarbon accumulation in a trap, which is reflected by the parameter in the corresponding growth equation. This parameter determines the time t_o (and the corresponding minimum field size Q_o) after which the Pareto's distribution becomes valid for the description of hydrocarbon accumulations.

Smaller accumulations, whether or not they have any commercial value, must comply with a different distribution function. It may be assumed that their emergence was caused by a different mechanism of hydrocarbon accumulation. This means that the regional system of accumulations has a dual organization. Hydrocarbon fields are no exception, and the same type of duality is found in different applications as well. It is important to remember that Pareto was an economist, and the distribution named after him is valid only for the population of taxed individuals (i.e., for the persons with income above a certain threshold imposed by the tax law). Q_o is just such a threshold (although not artificial but natural), emerging as a result of natural causes.

The second parameter (a) in Pareto's distribution equals the ratio of failure intensity α to the growth intensity β. Although this is interesting in itself, it implies another important result that can shed light on the problem of the duration of the oil and gas field formation process. As already mentioned, α is characteristic of the average duration of field formation: $t_{avg} = 1/\alpha$. In turn, as Equation 6-19 shows, $1/\beta$ is the time t_e during which the field size increases by a factor e (from a size Q_o to a size eQ_o). Therefore, a characteristic parameter a of the Pareto's distribution measures the relative duration of the increase of the initial field

Table 6-1
Field Size Distribution Functions Corresponding to the Selected Distribution Functions of Field Formation Duration t

Distribution Function of Field Formation Duration	Equation of Hydrocarbon Accumulation Growth in Time during the Field Formation		Comment
	$Q(t) = \gamma e\beta^t$	$Q(t) = \gamma t\beta$	
1. Exponential distribution $F(t) = 1 - e^{-\alpha t}\ (t \geq 0)$	1.1 $F(Q) = 1 - \dfrac{1}{(Q/\gamma)^{\alpha\beta}}$ $(Q \geq \gamma)$	1.2 $F(Q) = 1 - e^{-\left(\frac{\alpha}{\gamma^{1/\beta}}\right)Q^{1/\beta}}$ $(Q \geq 0, \beta > 1)$	1.1 is the Pareto distribution $F(Q) = 1 - (Q_o/Q)^a$, $Q_o = \gamma,\ a = \alpha/\beta$
			1.2 is the Weibull distribution $F(Q) = 1 - e^{-\gamma Q^n}$, $\gamma = \alpha/\gamma^{1/\beta},\ n = 1/\beta$
2. Weibull distribution $F(t) = 1 - e^{-\alpha t^m}\ (t \geq 0, m > 1)$	2.1 $F(Q) = 1 - e^{-\frac{\alpha}{\beta^m}\left(\ln\frac{Q}{\gamma}\right)^m}$ $(Q \geq \gamma)$	2.2 $F(Q) = 1 - e^{\frac{\alpha}{\gamma^{m/\beta}}Q^{m/\beta}}$ $(Q \geq 0, \beta > m)$	2.1 distribution at $m = 2$ is similar to the log-normal distribution $(Q > e\gamma)$. 2.2 is the Weibull distribution $F(Q) = 1 - e^{-\gamma Q^n}$, $\gamma = \alpha/\gamma^{m/\beta},\ n = m/\beta$
3. Gompertz distribution $F(t) = 1 - e^{-\alpha t^{mc}}\ (-\infty < t < \infty)$	3.1 $F(Q) = 1 - e^{\frac{\alpha}{\gamma^{m/\beta}}Q^{m/\beta}}$ $(Q \geq 0, \beta > m)$	3.2 $F(Q) = 1 - e^{\alpha\left(1 - e^{\frac{m}{\gamma^{1/\beta}}Q^{1/\beta}}\right)}$ $(Q > 0)$	3.1 is the Weibull distribution $F(Q) = 1 - e^{-\gamma Q^n}$, $\gamma = \alpha/\gamma^{m/\beta},\ n = m/\beta$

size by a factor e. The unit measure relative to which this duration is measured is the average field formation time $a = t_e/t_{avg}$. In the case when $a \approx 1$ (which is maintained in a number of publications) the average field formation time is an interval during which the field size (reserve) will increase by a factor e.

If one considers t_{avg} and Q_o as regional characteristics pertaining to the entire field set, then the individual field characteristics (formation duration t and reserves Q accumulated during this time) are also connected with the regional characteristics through parameter a. Indeed, $t/t_{avg} = a\ln(Q/Q_o)$. If one introduces dimensionless parameters $\tilde{t} = t/t_{avg}$ and $\tilde{Q} = Q/Q_o$, then $1/a = (\ln \tilde{Q})/\tilde{t}$. Therefore, a is a peculiar indicator of the process speed. And this peculiar growth speed $1/a$ is identical for all fields.

An interpretation of the Weibull's distribution parameters also yields certain information. If the Weibull's distribution is presented as follows:

$$F(Q) = 1 - e^{-\gamma Q^n}$$

then $n = 1/\beta$. Therefore, this distribution parameter varies inversely with the growth intensity parameter β. As mentioned above, the Weibull's distribution special format corresponds to the $0 < n < 1$ condition. This means that the condition $\beta > 1$ must be true. We arrived at the same condition based on different assumptions associated with the derivation of growth Equation 6-22. Thereby, the limitations of the Weibull's distribution parameter n are associated with specific features of the field size growth process; namely, with the increase in the speed of this process ($\beta > 1$).

To interpret the parameter λ, let us denote by t_1 the time when the field reaches a unit size: $Q(t_1) = 1$. It follows from Equation 6-22 that $t_1 = 1/\gamma^{1/\beta}$. Thus, the meaning of the λ parameter in the Weibull's distribution is clear. It describes the ratio of the time for the field to evolve to the unit size to the average field formation duration: $\lambda = t_1/t_{avg}$. As in the previous case, the λ parameter characterizes a dependence between the individual and regional parameters: $t/t_{avg} = \lambda Q^{1/\beta}$. Thus, λ plays the role of a rate, or proportionality coefficient.

2. In certain situations, the lifespan of electronic parts is described by the Weibull's distribution function. For this reason, we can also assume that this function describes the random duration of field formation. In this case, the Weibull's distribution function looks as follows:

$$F(Q) = 1 - e^{-\alpha t^m} \qquad (6\text{-}26)$$

As indicated above, the duration of work is a process with deceleration. Therefore, in this case the Weibull's distribution must have a mode:

$$\tilde{t} = \left(\frac{m-1}{\alpha m}\right)^{1/(m-1)}$$

The intensity function $(t) = \alpha m t^{m-1}$ must be an increasing function of t, which is true if $m > 1$.

The second line of Table 6-1 contains the distribution functions corresponding to the random time distribution we are analyzing and the two considered field size growth equations. Exponential growth is described by a distribution function, which was first derived by the writers. In contrast to the cases examined above, this distribution has the mode $\tilde{Q}$, which is determined by the solution of the following equation:

$$\frac{\alpha m}{\beta^m}\left(\ln\frac{\tilde{Q}}{Q_o}\right)^m + \ln\frac{\tilde{Q}}{Q_o} - (m-1) = 0 \qquad (Q_o = \gamma)$$

The intensity function $\mu(Q)$, corresponding to this distribution, has the following format:

$$\mu(Q) = \frac{\alpha m}{\beta^m}\left(\ln\frac{Q}{Q_o}\right)^{m-1}\frac{1}{Q} \tag{6-27}$$

This function does not monotonously decrease with increasing Q. At the point where $Q_o = \gamma$, this function is equal to zero and then it increases to a maximum at the point $\tilde{Q} = Q_o e^{m-1}$. Afterwards, it decreases again. This distribution is of interest to us at $m = 2$. The intensity function of the log-normal distribution (which is frequently used to describe field size distribution) has a maximum strongly biased toward small z values; when the values are large, it is described by the following equation:

$$\mu(z) = (\ln z)/z$$

Thus, the examined distribution function with the intensity function of a similar format is similar to the log-normal distribution at large Q values (larger than eQ_o) and at $m = 2$.

At $m = 1$, the examined distribution converts to the Pareto's distribution because

$$e^{-\frac{\alpha}{\beta}\ln\frac{Q}{Q_o}} = \left(\frac{Q}{Q_o}\right)^{-\alpha/\beta} = \left(\frac{Q_o}{Q}\right)^{\alpha/\beta}$$

Thus, we have derived a more general distribution function. The Pareto's distribution represents a particular case of this more general distribution at $m = 1$. In another case, at $m = 2$, the distribution coincides with the log-normal distribution within the domain where Q is greater than some relatively small limit. Most of the discovered fields, in terms of their reserve sizes, are usually within this domain: small and, especially, very small fields are discovered at the latest stages of exploration. This is why if the field distribution fits the log-normal distribution, it is the same as if it fits the derived distribution at $m = 2$.

A growth according to the exponential law (see Table 6-1) is described by the Weibull's distribution function, which was already discussed in detail. Additionally, its parameters are linked to the growth Equation 6-22 parameters and the random time distribution function parameters. From what was mentioned earlier concerning the only form of Weibull's distribution suitable for field size distribution, one can conclude that $(m/\beta) < 1$. Therefore, the condition $\beta > m$ must be true. This model shows that the growth intensity parameter is not just greater than one, but also greater than m.

3. In the previous case, we selected the distribution function for the duration of the deposit (field) formation by analogy with the distribution of the lifespan for technological systems. We will now be considering more complex and, at the same time, more reliable systems—namely, live, biologic systems. The life expectancy distribution is illustrated by the Gompertz distribution as follows:

$$F(t) = 1 - e^{-\alpha e^{mt}} \qquad (-\infty < t < \infty) \tag{6-28}$$

with the probability density of

$$f(t) = \alpha m^{mt} e^{-\alpha e^{mt}} \tag{6-29}$$

The distribution mode is $\tilde{t} = -(\ln\alpha)/m$ and the mortality intensity grows as the exponential function (with base e) of life expectancy t:

$$\mu(t) = \alpha m e^{mt}$$

It is interesting that the Weibull's distribution intensity function declines with increasing t as an exponential function (Equation 6.3), whereas the Gompertz distribution intensity function increases exponentially (with base e). This helps us understand the difference in these distributions, one of which describes the life expectancy of live organisms and the other, the life of technological systems.

The discussed distribution of the field formation duration (with the exponential (base e) growth of the field size) has a corresponding field size distribution as described by the Weibull's distribution (3-1 in Table 6-1). This distribution has the same format as distribution 2-2 of the same table. Thus, the same format of the Weibull's distribution may be obtained based on different premises. According to Equation 6-25, exponential growth of the field size has a correspondent function:

$$f(Q) = \frac{\alpha m}{\beta \gamma^{1/\beta}} e^{-\alpha e^{\frac{m}{\gamma^{1/\beta}} Q^{1/\beta}}} e^{\frac{m}{\gamma^{1/\beta}} Q^{1/\beta}} Q^{1/\beta - 1} \tag{6-30}$$

This function, however, is not a distribution density because its integral is taken from zero to infinity and is equal to $e^{-\alpha}$ (i.e., it is less than one, if $\alpha > 0$). Therefore, in order to convert Equation 6-30 into a density distribution, it should be written in the following format:

$$f(Q) = e^{\alpha} \frac{\alpha m}{\beta \gamma^{1/\beta}} e^{-\alpha e^{\frac{m}{\gamma^{1/\beta}} Q^{1/\beta}}} e^{\frac{m}{\gamma^{1/\beta}} Q^{1/\beta}} Q^{1/\beta - 1}$$

in which case the distribution function $F(Q)$ will be equal to:

$$F(Q) = 1 - e^{\alpha\left(1 - e^{\frac{m}{\gamma^{1/\beta}} Q^{1/\beta}}\right)} \quad (Q > 0) \tag{6-31}$$

This distribution is of no interest because at $\beta > 1$ its intensity function

$$\mu(Q) = \frac{\alpha m}{\beta \gamma^{1/\beta}} e^{\frac{m}{\gamma^{1/\beta}} Q^{1/\beta}} Q^{1/\beta - 1}$$

declines with increasing Q only over a limited interval and begins to grow with increasing Q after having reached a maximum at the point $\tilde{Q} = \gamma[(\beta - 1)/m]^{\beta}$.

Thus, we have derived four distributions. They include the already known Pareto and Weibull's distributions and two new distributions.

In the case of exponential growth (base e, Equation 6-19), at the moment $t_o = 0$, $\gamma = Q_o$; in the case of power growth (Equation 6-22), at the moment $t_1 = 1$, correspondingly, $\gamma = Q_1$. Taking this into account, distribution densities of the Q value corresponding to the four derived distributions may be represented as follows:

$$f(Q) = \frac{\alpha}{\beta Q_o}\left(\frac{Q_o}{Q}\right)^{\alpha/\beta+1} \tag{6-32}$$

$$f(Q) = \frac{\alpha m}{\beta Q_1^{m/\beta}} e^{-\frac{\alpha}{Q_1^{m/\beta}}Q^{m/\beta}} Q^{m/\beta-2} \tag{6-33}$$

$$f(Q) = \frac{\alpha m}{\beta^m} e^{-\frac{\alpha}{\beta^m}\left(\ln\frac{Q}{Q_o}\right)^m}\left(\ln\frac{Q}{Q_o}\right)^{m-1} Q^{-1} \tag{6-34}$$

$$f(Q) = e^{\alpha}\frac{\alpha m}{\beta Q_1^{(1/\beta)}} e^{-\alpha e^{-\frac{m}{Q_1^{1/\beta}}Q^{1/\beta}}} e^{\frac{m}{Q_1^{1/\beta}}Q^{1/\beta}} Q^{1/\beta-1} \tag{6-35}$$

Data on reserves of small and very small fields are not available. In view of this, it is important to examine the matching, or divergence, of these distributions at sufficiently large Q values (i.e., as it applies to the existing fields which can be analyzed). The tail portions of the distribution under examination are as follows:

$$p(Q) = e^{A-(\alpha/\beta+1)\ln Q} \rightarrow_{(Q\to\infty)} e^{-\lambda \ln Q} \qquad (\lambda > 1)$$

$$p(Q) = e^{B-\left[\frac{\alpha}{Q_1^{m/\beta}}Q_1^{m/\beta}+\left(1-\frac{m}{\beta}\right)\ln Q\right]} \rightarrow_{(Q\to\infty)} e^{-q}Q^{m/\beta} \qquad (1 < m < \beta)$$

$$p(Q) = e^{C-\frac{\alpha}{\beta^m}\left(\ln\frac{Q}{Q_o}\right)^m+\left(\ln\frac{Q}{Q_o}\right)^{m-1}-\ln Q} \rightarrow_{(Q\to\infty)} e^{-k(\ln Q)^m} \qquad (m > 1)$$

$$p(Q) = e^{\alpha + D - \alpha e^{\frac{m}{Q_1^{1/\beta}} Q_1^{1/\beta}}} + \frac{m}{Q_1^{1/\beta}} Q^{1/\beta}$$

$$- \frac{\beta - 1}{\beta} \ln Q \to_{(Q \to \infty)} e^{-\alpha e^r Q^{1/\beta}} \qquad (\beta > 1)$$

The right parts of these expressions (to the right of the sign "→") represent the functions describing the distribution's behavior when $Q \to \infty$.

As one can see, the rate of convergence is different for these distributions. As expected, the distribution with density (6-35) converges most rapidly. The Weibull's distribution (its density is given by Equation 6-33) converges much slower. In general, its tail is longer than that of the exponential distribution because the latter's conversion order is not

$$e^{-qQ^{m/\beta}} \qquad (m < \beta)$$

but e^{-qQ}. The distribution with density (Equation 6-34) converges even more slowly. It is interesting that the log-normal distribution has the same order of convergence as this distribution at $m = 2$. This follows from the description of the log-normal distribution's tail

$$p(Q) = e^{R - \ln Q - k(\ln Q)^2} - \to_{Q \to \infty} e^{-k(\ln Q)^2}$$

We indicated above that these distributions coincide at sufficiently large field reserves. The Pareto distribution (its density is given by Equation 6-32) converges most slowly. Its convergence order is not $e^{-k(\ln Q)^m}$ $(m > 1)$ as previously indicated, but $e^{-\ln Q}$, which leads to infinite moments on the order of $n < \lambda - 1$. A slow convergence of the distribution and the presence of a long tail are associated with the presence of very large and giant fields, which contain a substantial part of the ultimate potential. Depending on how significantly the reserves are concentrated in a small number of fields, the tail length of the field reserve distribution curve will differ.

Random Parameter: Growth Intensity Parameter β. Individual differences in the reserve size of each particular field may be associated with random fluctuations of the parameter β in the field size growth Equations 6-19 and 6-22. Until now we neglected these differences assuming that they were too small. Thereby, we operated with some average value, β_{avg}. We will now assume that the duration of formation of individual fields fluctuates so insignificantly that these fluctuations can be neglected. Thus, t becomes constant, equal to some average value (t_{avg}).

It is natural to assume that the intensity parameter β is different for different fields. Its specific value is determined by a complex combination of many natural factors that allow us to consider the intensity as a random value. In order to study the structure of the field size distribution, it is necessary to know the distribution $F(\beta)$ determined using data for all of the numerous fields. The format of $F(\beta)$ is not known. However, it may be determined based on the following. It is a natural concept that a hydrocarbon accumulation becomes a field (i.e., its size becomes commercial) only at a certain minimum β value. In other words, there is a certain threshold in the field formation, and the β parameter corresponds to this threshold. This is why the field distribution is defined by the β values that exceed the threshold. Considering the concentration and dispersion phenomena, it may be assumed that the probability density $f(\beta)$ is not an ascending function. More likely, any β values within the range from its minimum to its maximum value are equally probable. We cannot exclude the possibility that the probability decreases with increasing β (the larger the β, the less probable it is). Taking all this into consideration, it can be hypothesized that the distribution of parameter β is either exponential or uniform, namely:

$$F(\beta) = 1 - e^{-\alpha\beta} \tag{6-36}$$

$$F(\beta) = \frac{1}{\beta_{max} - \beta_{min}}(\beta - \beta_{min}) \quad (\beta_{min} < \beta < \beta_{max}) \tag{6-37}$$

Thus, the field size distribution is either the Pareto's distribution or another new distribution, namely, the ln$Q$ distribution (Table 6-2). Earlier we arrived at the Pareto's distribution based on completely different considerations. It is worth emphasizing again that many different primary concepts of the field formation laws lead to this distribution. Another, new distribution, contrary to the Pareto distribution, corresponds to a uniform, rather than exponential, parameter distribution. Similar to the Pareto's distribution, it corresponds to both evolution equations: exponential (base e) and general exponential. Density of this distribution has the following format:

$$f(Q) = \frac{1}{\ln(Q_{max}/Q_{min})}\frac{1}{Q} \tag{6-38}$$

A comparison with the Pareto's distribution density $[f(Q) = a(Q_o^a/Q^{a+1})]$ shows that this distribution represents a limiting case of the Pareto's distribution at $a \sim 0$. These two distributions are practically identical at

Table 6-2
Field Size Distribution Functions Corresponding to the Selected Distribution Functions of the Growth Intensity Parameter β

Distribution Function of Growth Intensity Parameter β	Equation of Growth of HC Accumulation in Time in the Process of Field Evolution		Comment
	$Q(t) = \gamma e^{\beta t}$	$Q(t) = \gamma t^{\beta}$	
1. Exponential (base e) distribution $F(\beta) = 1 - e^{-\alpha\beta} (\beta \geq 0)$	1.1 $F(Q) = 1 - \left(\frac{Q_o}{Q}\right)^{\alpha/t}$ $(Q \geq Q_o)$	1.2 $F(Q) = 1 - \left(\frac{Q_o}{Q}\right)^{\alpha/\ln t}$ $(Q \geq Q_o)$	Pareto distribution; correspondingly, $Q_o = \gamma$
2. Uniform distribution $(\beta) = \frac{1}{\beta_{max} - \beta_{min}} (\beta - \beta_{min})$, $\beta_{min} < \beta < \beta_m$	2.1 $F(Q) = \frac{1}{\ln(Q_{max}/Q_o)} \ln \frac{Q}{Q_o}$ $(Q_o \leq Q \leq Q_{max})$	2.2 $F(Q) = \frac{1}{\ln(Q_{max}/Q_o)} \ln \frac{Q}{Q_o}$ $(Q_o \leq Q \leq Q_{max})$	In distribution 2.1, $Q_o = \gamma e^{\beta_{min} t}$; $Q_{max} = \gamma e^{\beta_{max} t}$ In distribution 2.2, $Q_o = \gamma t^{\beta_{min}}$; $Q_{max} = \gamma t^{\beta_{max}}$

small a values. It is important to note that at $a < 1$, the Pareto's distribution does not have finite mean and variance.

Equations 6-19 and 6-22 show that the time it takes for a field to double in size from γ to 2γ is determined, respectively, by $t_2 = (1/\beta)\ln 2$ and $\ln t_2 = (1/\beta)\ln 2$. The value $\alpha = 1/\beta_{avg}$ determines the average time it takes for a field to double in size. Thus, the characteristic parameter a of the Pareto's distribution is proportional to the average doubling time interval. In the former case, $a = \alpha/t = 1/(\beta_{avg}t) = [1/(t\ln 2)]\bar{t}_2$ and its logarithm. In the latter case, $a = \alpha/\ln t = 1/(\beta_{avg}\ln t) = [1/(\ln t \ln 2)]\ln \bar{t}_2$; where $\bar{t}_2$ is the average time of doubling of reserves.

It must be remembered that the Q_o parameter of the Pareto's distribution and of the $\ln Q$ distribution have different meanings. In the former case, it represents the field size at the times $t = 0$ and $t = 1$ in the respective growth equations. In the latter case, depending on the growth equation, it represents either $Q_o = \gamma e^{\beta_{min} t}$ or $Q_o = \gamma t^{\beta_{min}}$. In the latter case, the Q_o is defined by the minimum (threshold) value of the growth intensity parameter β.

The $\ln Q$ distribution does not leave the derived distribution series. It may also be close to the Pareto's distribution if the Pareto's distribution characteristic parameter a is very small. This distribution converges even more slowly than the Pareto's distribution because its convergence order is not $1/Q^{a+1}$ but, rather, $1/Q$. The common use of the Pareto's distribution in field size distribution studies may be attributed to the fact that many other distributions are similar to it. Actually, some are so close that the differences are not noticeable on the empiric level.

Process of the Field Population Formation as a Randomized Markov-type Stochastic Process

It was assumed earlier that the formation process of an individual deposit (field) is governed by deterministic causes, so that the field size at any given moment is completely determined by corresponding evolution equations. The moment when the field reaches a certain reserve size is uniquely determined by these equations.

In this section field formation is considered a probabilistic process. In this case, there is no unique relation between the time and the field size. There are many possible results, each with a certain probability measure associated with it. The explanation for this is that the field formation process depends on an indefinitely large number of factors acting at random. Field evolution is no longer established deterministically, but, instead, has stochastic character.

Stochastic Process of Field Formation

The main concept used in this section is that of the state of a deposit (field), which is defined by the size of the accumulated hydrocarbons. An increase or decrease by some unit is interpreted as a transition to a new state. Thereby, the process of field formation (the change in the amount of hydrocarbons in the trap) is understood as a sequence of transitions from one state to another. We will assume that this process proceeds in such a manner that the probability of a field changing from one state to another at the next moment in time depends only on its state at a given moment and does not depend on the preceding states. In other words, each subsequent step in the change of deposit (field) size depends stochastically only on the state of a field at a given moment. This constitutes the condition of being a Markov process: future behavior of the process does not depend on its behavior until the present moment.

It is reasonable to assume that the field formation process has the following properties:

1. The probability of a change from the field state Q during the interval $(t,\ t + \Delta t)$ is proportional to Δt. The probability of an increase in the field size to $Q + 1$ and of the transition $Q \rightarrow Q + 1$ is $\lambda\Delta t + 0(\Delta t)$. The probability of a decrease in field size to $Q - 1$ and of the transition $Q \rightarrow Q - 1$ is $\mu\Delta t + 0(\Delta t)$. λ and μ are certain positive values called the "transition density" and $0(\Delta t)$ is an indefinitely small value of a higher order than Δt.
2. The probability of more than one transition during the time Δt is a value on the order of $0(\Delta t)$ (i.e., it is negligibly small).
3. The probability of no transition during the time Δt is correspondingly, $1 - (\lambda + \mu)\Delta t + 0(\Delta t)$.

We will also assume that the transition densities are not constant. Based on the previous discussion, it is reasonable to assume that the probability of an increase or decrease in the size of a deposit (field) during time interval Δt depends on the field size at time t. In other words, the probability of a field transition to a new state within the $(t,\ t + \Delta t)$ time interval must be a function of the state Q of the field at the moment in time t; that is, $\lambda = \lambda(Q)$ and $\mu = \mu(Q)$. Taking this into account and considering that for any practical purpose only one change in the field state may occur during the time Δt, the transition probability can be expressed as follows:

$$P_{Q,Q+1}(t,\ t + \Delta t) = \lambda(Q)\Delta t + 0(\Delta t)$$

$$P_{Q,Q-1}(t,\ t+\Delta t) = \mu(Q)\Delta t + 0(\Delta t)$$

$$P_{Q,Q}(t,\ t+\Delta t) = 1 - [\lambda(Q) + \mu(Q)]\Delta t + 0(\Delta t)$$

The Markov process under consideration is a regular process. Therefore, a system of differential equations for the probability of a field being in a particular state Q at the moment of time t, may be determined from the system of ordinary differential equations for a regular Markov process. It has the following format:

$$\frac{dP_Q(t)}{dt} = -[\lambda(Q) - \mu(Q)]P_Q(t) + \lambda(Q-1)P_{Q-1}(t) + \mu(Q+1)P_{Q+1}(t) \tag{6-39}$$

if $Q \geq 1$, and

$$\frac{dP_o(t)}{dt} = -[\lambda_o + \mu_o]P_o(t) + \mu_1 P_1(t)$$

Usually, a Markov process is analyzed beginning at a certain time $t = 0$. In this particular case, the initial conditions at which the system is solved are:

$$P_Q(0) = \begin{cases} 1 \text{ if } Q = 1 \\ 0 \text{ if } Q \neq 1 \end{cases}$$

Field Formation Process as a Linear Process of Reproduction and Extinction

The solution of system 6-39 is defined by specific functions $\lambda(Q)$ and $\mu(Q)$. Our assumption is that the functions $\lambda(Q)$ and $\mu(Q)$ can be presented in the form $\lambda(Q) = \lambda Q$ and $\mu(Q) = \mu Q$, where λ and μ are proportionality coefficients. Likewise, we can assume that the larger the field has become, the easier it is for the field to transition to a new, higher reserve level. Similarly, it is easier for this field to transfer to a lower level because the dispersion of hydrocarbons increases, as discussed earlier.

Under these assumptions, system 6-39 has the following format:

$$\frac{dP_o(t)}{dt} = \mu P_1(t);\ \frac{dP_Q(t)}{dt} = -(\lambda + \mu)QP_Q(t) + \lambda(Q-1)P_{Q-1}(t) + \mu(Q+1)P_{Q+1}(t) \tag{6-40}$$

The initial conditions remain the same. Solving the system using the generating function $G(t,u) = \sum_{Q=o}^{\infty} P_Q(t)u^Q$, we obtain a differential equation $\partial G(t,u)/\partial t = (\lambda u - \mu)(u - 1)\partial G(t,u)/\partial u$ with the initial condition $G(0, u) = u$, which has the following solution:

$$G(t,u) = \frac{\mu x + u[1 - (\lambda + \mu)x]}{1 - u\lambda x}$$

where:

$$x = \begin{cases} \dfrac{1 - e^{(\lambda-\mu)t}}{\mu - \lambda e^{(\lambda-\mu)t}}, & \text{if } \lambda \neq \mu \\ \dfrac{t}{1 + \mu t}, & \text{if } \lambda = \mu \end{cases}$$

It follows from this that the probability $P_Q(t)$ of finding a field (deposit) in state Q at the moment in time t (provided that at the initial moment in time $t = 0$, its reserves were equal to one—in some unit of measurement) is equal to:

$$P_o(t) = \mu x$$
$$P_Q(t) = (1 - \lambda x)(1 - \mu x)(\lambda x)^{Q-1} \quad (Q > 1) \tag{6-41}$$

The probabilistic model of field formation considered is usually applied to the description of randomly branching reproduction and extinction processes, where probabilities of birth and death at any given moment are proportional to the size of the population. Thus, the field formation process may be considered a linear process of reproduction and extinction. Each individual or particle in this process within the time interval $(t, t + \Delta t)$, may cause the emergence of a new individual or particle with the probability of $[\lambda\Delta t + 0(\Delta t)]$, or may disappear with the probability of $[\mu\Delta t + 0(\Delta t)]$.

The mathematical expectation of the process (i.e., the field's average size by the time t) is determined by the value $M(t) = \sum_{Q=0}^{\infty} QP_Q(t)$. If one differentiates the last equation with respect to t, then:

$$\frac{dM(t)}{dt} = \sum_{Q=0}^{\infty} Q\frac{dP_Q(t)}{dt}$$

$$= \sum_{Q=0}^{\infty} Q[-(\lambda + \mu)QP_Q(t) + \lambda(Q - 1)P_{Q-1}(t) + \mu(Q + 1)P_{Q+1}(t)]$$

$$= (\lambda - \mu)M(t)$$

After solving this differential equation provided that the field size at the initial moment in time was equal to one, the following is obtained:

$$M(t) = e^{(\lambda-\mu)t} \tag{6-42}$$

The average field size changes exponentially (base e) in time. This coincides with the exponential (base e) growth Equation 6.19 derived earlier. However, the field size will grow only if $\lambda > \mu$. If $\lambda = \mu$, the size will remain constant in time, and if $\lambda < \mu$, the average field size will exponentially (base e) decline in time manifesting the prevalence of the destructive processes over the generating processes. The coincidence of Equation 6-42 with the equation of deterministic exponential (base e) growth may by considered a reflection of only basic trends pertaining to the field growth process in the deterministic growth models. The deterministic evolution is the evolution of the average. The deterministic growth parameter b is associated with the transition densities through the relation $\beta = \lambda - \mu$.

The particular case of the analyzed process when $\lambda \gg \mu$ is of interest. In this case, we can assume $\mu \approx 0$. It is necessary to mention that we are not studying the decayed fields. Thus, it would be reasonable to assume that the hydrocarbon dispersion from the accumulations formed by the time t is small compared to the hydrocarbon accumulation. Therefore, the probability of a transition $Q \rightarrow Q - 1$ is small compared with the probability of a transition $Q \rightarrow Q + 1$. In general, it may be assumed that the dispersion is just delaying the accumulation. We then obtain a system with two states: within the time interval Δt it will either remain in its previous state Q or will transition into a new state $Q + 1$. The presence of dispersion will cause a decline of the λ value.

Thus, the process of deposit (field) formation is considered a linear process of pure reproduction (without extinction). Under these conditions, the solution 6-41 of the system of Equations 6-40 can be expressed as follows:

$$P_Q(t) = \begin{cases} e^{-\lambda t}(1 - e^{-\lambda t})^{Q-1} & Q = 1, 2, 3, \ldots \\ 0 & Q = 0 \end{cases} \tag{6-43}$$

where λ is the process intensity parameter.

This function represents the well-known Hule-Farri distribution. In our case, it describes the probability of a field size reaching Q at the time t, provided that $\mu = 0$, and that at the initial time $t = 0$ the size of field reserves was equal to one (in some units of measure). The mathematical expectation of the process (i.e., the field's average size by the time t)

changes in time exponentially (base e) and depends only on the parameter $\lambda M(t) = e^{\lambda t}$.

Randomization of a Branching Process

We assumed earlier that the intensity parameter $\beta = \gamma - \mu$ is different for different fields and represents a random value with the distribution density $f(\beta)$ defined over the general population of fields. The fields are not distinguished individually (depending on β), but only by the final result (i.e., the reserve size). Due to this, the probability density $f(Q)$ should have been obtained through the averaging of individual results (Equation 6-41), which are denoted $P_Q(t) = P(Q/\beta)$, over the set of fields with different β:

$$f(Q) = \int P(Q/\beta) f(\beta) d\beta \tag{6-44}$$

A random process (Equation 6-44) averaged with respect to parameter β is a final description of the field population formation. This process (named after Yablonsky) may be called a randomized branching process [56].

Equation 6-44 describes the final distribution of the field population $f(Q)$ by their field size Q. Its specific format would have been determined depending on the distribution function $f(\beta)$. The exponential or the uniform distribution could have been assumed as such a distribution.

In this case, however, this cannot be accomplished because, in the general case of Equation 6.41, the $P_Q(t)$ value is a function of λ and μ but not of their difference $(\lambda - \mu)$, which means it cannot be expressed as $P(Q/\beta)$. For this same reason, $f(Q)$ also cannot be determined.

A pure reproduction process (6-43), however, lends itself to randomization. In this particular case, the distribution $f(Q)$ was obtained under the exponential distribution law of the λ parameter (i.e., $f(\lambda) = \alpha e^{-\alpha\lambda}$). Its format is quite complex, but under some simple assumptions it asymptotically converges to the Pareto's distribution [56]:

$$f(Q) = \frac{\alpha}{t} \frac{1}{Q^{\alpha/t+1}} = \frac{a}{Q^{a+1}} \tag{6-45}$$

where $a = \alpha/t$ is the characteristic parameter of the distribution.

Therefore, the model of the field formation process, considered as a linearly branching pure reproduction process, also results in the Pareto's distribution under the hypothesis of an exponential distribution of the process intensity parameter λ.

This distribution is identical with the distribution listed in Table 6-2. The meaning of the characteristic parameter a was explained earlier. In this case, $Q_o = 1$. This is also clear from the initial conditions. Therefore, Q_o here serves as the unit measure which determines the difference between the states Q and $Q + 1$.

Another specific case of the Equation 6-40 process is a process with the condition $\lambda = \mu$. It is described by the following probability $P_Q(t)$ distribution of a field being in state Q at the time t:

$$P_Q(t) = \frac{1}{(1+\lambda t)^2}\left(\frac{\lambda t}{1+\lambda}\right)^{Q-1} \tag{6.46}$$

Equation 6-42 indicates that the mathematical expectation of the process is constant and is equal to one. As before, λ may be considered a random parameter driving the process. Averaging Equation 6-46 with respect to the λ parameter (randomizing the process), we obtain:

$$f(Q) = \int \frac{1}{(1+\lambda t)^2}\left(\frac{\lambda t}{1+\lambda t}\right)^{Q-1} f(\lambda)d\lambda \tag{6-47}$$

We showed above that the distribution of a λ-type parameter, namely, the β parameter, may be exponential or uniform. Assuming the uniform distribution for $f(\lambda)$,

$$f(\lambda) = \frac{1}{\lambda_{max} - \lambda_{min}} \tag{6-48}$$

and integrating Equation 6-47 (taking into consideration Equation 6-48), one obtains the distribution density, which is proportional to the beta-function:

$$f(Q) = \frac{1}{(\lambda_{max} - \lambda_{min})t} B(Q, 5) \tag{6-49}$$

where $B(Q, 5) = [\Gamma(5)\Gamma(Q)]/[\Gamma(Q + 5)]$ is the beta-function and $\Gamma(Q)$ is the gamma-function.

Based on the Stirling asymptotic expansion for $\Gamma(Q)$, with increasing Q we will asymptotically obtain:

$$f(Q) = \frac{1}{(\lambda_{max} - \lambda_{min})t}\Gamma(5)\frac{e^5}{Q^5} = \frac{4!e^5}{(\lambda_{max} - \lambda_{min})t}\frac{1}{Q^5} \tag{6-50}$$

Distribution 6-50 asymptotically converges to the Pareto's distribution with the characteristic parameter $a = 4$ when $Q \to \infty$. This implies that:

$$Q_o = \left[\frac{e^5 3!}{(\lambda_{max} - \lambda_{min})t} \right]^{1/4}$$

The distribution's stability condition is t = const. Therefore, Q_o turns out to be connected with the $(\lambda_{max} - \lambda_{min})$ (i.e., with the range of the λ parameter).

Field Population as an Equilibrium System

In this section, we will explore a different approach to the construction of the field size distribution function. The process under study seems to be a stochastic process but, strictly speaking, it is not probabilistic. It is assumed that the field size fluctuations, generally, are not random, but are subordinated to some common goal to which a field population as a complete natural system is tending. The fields are subordinated by a certain harmony of this system, and this harmony is expressed in quantitative relations among the field reserves. In short, it is assumed that a set of fields is integrated into a single system that is characterized by a purposeful behavior. Whereas the previous two sections described the fields from the cause-and-effect viewpoint, here we will consider the goal viewpoint.

Thermodynamic Model of Field Population

The purposeful behavior of systems is usually associated with processes that may be interpreted as transition processes to thermodynamic equilibrium in a physical system with a certain energy level spectrum, and when the system is in contact with a thermostat. When a physical system occurs in the state of equilibrium, an extremum of the corresponding thermodynamic potential is reached. Likewise, in these processes a function whose value determines its usefulness in the system's behavior, is minimized. In accordance with this, let us consider the field multitude as an equilibrium system where random field transitions from one state to another, do not disturb the system's equilibrium on average. This allows us to study the field distribution through the thermodynamic techniques used to analyze the equilibrium distribution of molecules in a gas.

As in thermodynamics, we will characterize each field by the energy $E(Q)$ depending on its reserves Q. Goal achievement occurs when field

distribution $f(Q)$ corresponds to the state of thermodynamic equilibrium. Correspondingly, we will consider the system's entropy H as the goal function. A field population, therefore, will have a distribution of states $f(Q)$, which will make entropy H maximum at an average energy E. Actually, this is Boltzmann's variational principle, which can be expressed in the following format:

$$-\int f(Q)\ln f(Q)dQ = H \rightarrow \max \tag{6-51}$$

under the following constraints:

$$\int E(Q)f(Q)dQ = E$$

$$\int f(Q)dQ = 1$$

Boltzmann's variational principle is used as the basis for the "thermodynamic approach." It is important to note that such an approach, introducing visual "physical" concepts, is widely used for the analyses of complex systems.

Let us interpret the external medium of an oil and gas basin as a thermostat with temperature T; let us further consider T a characteristic of an external effect (as the intensity of external effort). Then the field distribution (a statistic distribution for a minor subsystem that is part of a large closed system at a state of equilibrium that serves as a thermostat) is the Gibbs distribution:

$$f(Q) = \frac{1}{z}e^{-E(Q)/T} \tag{6-52}$$

where $z = \int e^{-E(Q)/T}dQ$ is the so-called statistical integral. This represents a general solution of the variational problem 6-51.

Geologic Interpretation of Energy E(Q)

Generally speaking, the Gibbs distribution describes the microscopic state of a system in thermal equilibrium with the surrounding medium, where a constant temperature is maintained (thermostat). $E(x)$ is the system's total energy. Kinetic energy is associated with the motions of molecules, and potential energy is caused by the fact that the system is in an external force field. Temperature T is connected with the average energy of the molecular thermal motion and serves as a measure of its intensity.

Various scientific disciplines using thermodynamic models ascribe different interpretations to the system's energy $E(x)$ of state x and average

energy E. For instance, in economic applications $E(x)$ represents the cost of manufacturing product x, and E represents the average resources. In linguistic applications, these are, respectively, the "costs" of words and letters and the average "price" of a character. The scientologic analysis identifies energy $E(x)$ with efforts (and respective time expenditures) necessary for a scientist to publish x papers. In this case, an equilibrium state of the scientific community is contemplated in the "external" (social) medium.

When analyzing a field population, it can be assumed that the equilibrium among the fields is a compromise between the contradictory trends of homeostasis preservation and variability. On the one hand, each field tends to maintain its status in the total system, which is a guarantee of the system's equilibrium. On the other hand, the fields are forced to change their state under external efforts in order for the system to maintain its equilibrium in a new environment. In this circumstance, each field tends to "spend" a minimum of effort. It may be assumed that nature works in an optimal manner: it reaches a maximum effect by expending minimum energy. It is believed that this principle is most successfully implemented for large fields: relaxation toward the equilibrium state occurs most often at the expense of the change in state of the large fields. Then, energy $E(Q)$ should increase with the growth of Q, but not in proportion to Q. Energy grows much slower than Q, which provides a high efficiency of energy expenditures (in terms of the achieved result, or change in the field size).

It is clear that the field size distribution $f(Q)$ is defined by a specific function $E(Q)$. In view of the aforementioned, this function may be defined as follows:

$$E(Q) = \rho \ln Q \tag{6-53}$$

where ρ is the proportionality coefficient.

It is important to emphasize that energy is associated with a specific field (individuum) and not with a system of coordinates. In this sense, it may be compared to its own time of field evolution. Contrary to the usual notion of time as a duration, its own time may be associated with the number of field transitions from one state to the next during the same time interval. The larger the field, the greater its own time.

In connection with Equation 6-53, it is also interesting to note that Price suggests estimating the elitism of a scientist by the logarithm of the number of his publications [40, p. 326]. Yablonsky, considering the energy $E(x)$ as a complexity characteristic for writing x number of articles, arrived at the similar expression $E(x) = \rho \ln x$ [56]. He indicated that $E(x)$ must be identified with the time expenditure, which can be considered a measure of effort for writing x articles. In principle, this time is also the

own time for each scientist. Such agreement in results increases the reliability of the proposed model.

Of course, such an interpretation of field "behavior" is an oversimplification. The geologic interpretation of energy $E(Q)$, however, may be simplified by understanding that the final result is a consequence of the field's potential to evolve (a consequence of combining the geologic conditions differently, which may be favorable for some fields and unfavorable for others). Energy is a measure of the field's capacity for moving to the next state under pressure from external conditions (a measure of its activity). This activity, in turn, depends on the local manifestations of different geologic factors affecting the evolution of traps, the availability of reservoir rocks, and the migration of hydrocarbons.

Field Population Distribution in the State of Equilibrium

The statistical integral in Gibbs distribution 6-52 can be determined from the boundary conditions and from considering Equation 6-53. Assuming a field with minimal reserves Q_o as the lower limit of the field population, one obtains:

$$\int_{Q_o}^{\infty} e^{-E(Q)/T} dQ = \frac{1}{(1 - \rho/T) Q_o^{1-\rho/T}}$$

Using Equation 6-53, the final distribution of the field population in the state of equilibrium is as follows:

$$f(Q) = \frac{\rho/T - 1}{Q_o} \left(\frac{Q_o}{Q} \right)^{\rho/T} \tag{6-54}$$

This is the already considered Pareto's distribution with the parameter $a = \rho/T - 1$. In this case, it was obtained based on the variational principle through the construction of a "thermodynamic" field model. Parameter T is a characteristic of the entire field system; the temperature is often used as the characteristic of thermostat. It may be considered a characteristic of external influence (such as the intensity of external conditions) affecting the field formation. These conditions are regional and pertain to a given oil and gas basin. Therefore, the parameter of Pareto's distribution determines the regional specifics of the field population formation in an oil and gas region.

The parameter a is also associated with the average energy E. Using the techniques of Lagrange multipliers for searching the $f(Q)$ function (realizing the maximum entropy), one obtains:

$$E\int_{Q_o}^{\infty} e^{-E(Q)/T}dQ = \int_{Q_o}^{\infty} E(Q)e^{-E(Q)/T}dQ$$

Consequently, assuming $\ln Q_o \approx 0$, one obtains: $\rho/T - 1 = \rho/E$. Thus, the distribution 6-54 can be represented in the following format:

$$f(Q) = \frac{\rho/E}{Q_o}\left(\frac{Q_o}{Q}\right)^{\rho/E+1} \tag{6-55}$$

This equation shows that the parameter $a = \rho/E$ depends on the average activity of fields. This activity is also a characteristic of the basin as a whole. The parameter a determines the non-uniformity of the field distribution by the reserves.

In addition, the system's entropy H can also be expressed through the distribution parameter a (Equation 6-54):

$$H = -\int_{Q_o}^{\infty} f(Q)\ln Q dQ = -(2a+1)\ln Q_o - \ln a + \frac{1}{a} + 1 \tag{6-56}$$

There is a unique relationship between the entropy H and the parameter a. Entropy is usually considered a measure of non-uniformity of a system (or measure of the scatter of the values of its variables). Taking this into account, it is reasonable to consider the distribution parameter a as a measure of the field size non-uniformity.

Therefore, the field size distribution under Pareto's law follows the variational principle that describes the equilibrium state of the field population within the external medium. The value of this approach is that it allows for the study of a system of fields and, therefore, of the processes forming this system, based on entropy and the associated concepts of structuring, value, and information ideas.

Study of the Field Size Distribution Based on Stable Distributions Different from a Normal Distribution

The field size distribution has a non-Gaussian nature and the patterns behind this distribution are radically different from the Gaussian. Probability theory has a special mathematical technique for examining these patterns. Let us explore the application of this technique to the study of field size distribution.

The technique is based on the examination of the sum of independent, similarly distributed random values. Under certain conditions, a distribution of this preliminarily normalized sum converges to a normal distribution. The normal distribution can be used for approximating many different distributions. We, however, are interested in the case where a limit distribution differs from the normal.

It is important at this point to discuss stable distributions. A distribution is considered stable if a linear combination of two such distributions produces a distribution belonging to the same type. Normal distribution is a stable distribution. Stable distributions may converge to a normal distribution (finite variance), as well as to distributions different from the normal (infinite variance). The convergence conditions for the distributions of normalized sums of random variables (with the same distribution) to the stable distributions (different from normal) when $x \to \infty$, are as follows:

$$F(-x) \approx \frac{c_1}{|x|^{\alpha}}; 1 - F(x) \approx \frac{c_2}{x^{\alpha}}; c_1 \geq 0, c_2 \geq 0; c_1 + c_2 > 0 \qquad (6\text{-}57)$$

The processes leading to a normal distribution are continuous (although they may be non-differentiable such as Wiener's process). At the same time, the processes leading to stable distributions different from the normal, are discontinuous (i.e., they have a discrete nature).

It appears that the field size distribution may be approximated by one of the limiting stable distributions different from normal. This is indicated by the concentration and dispersion phenomena, which may be associated with indefinite variance: divergence of moments of stable non-Gaussian distributions results in an increase in the random value scatter compared to "Gaussian" laws. This view is supported by the discrete nature of change in the field reserve, as indicated earlier. A discontinuous nature is typical for processes that converge accordingly to stable non-Gaussian laws.

The densities arising from stable laws are unimodal and are different from zero over the entire number line, or over the semi-number line. The problem with using stable distributions is that, as a rule, their explicit form is not known. Pareto's distribution, however, coincides with the asymptotic behavior of stable distributions different from normal (Equation 6-57). This fact convinced some scientists that Pareto's distribution, in some domains, plays almost the same universal role as the normal law for stochastic problems with finite variance [56]. This is supported by the diverse applications of Pareto's distribution in different disciplines, as discussed later.

Two stable non-Gaussian distributions are known in their explicit format as:

$$f(x) = \frac{1}{\sqrt{2\pi}} x^{-3/2} e^{-1/(2x)} \qquad x > 0 \tag{6-58}$$

Its asymptotic behavior when $x \to \infty$, is described by the following equation, the shape of which coincides with the Pareto's distribution:

$$f(x) \approx \frac{1}{\sqrt{2\pi}} x^{-3/2}$$

This equation represents the so-called "positive stable law with parameter $a = 1/2$." It describes numerous branching process models: the ruin of a gambler, the moment of the first passage of a given level for the Brownian process, chemical reactions, and so on. It also appears as the limit distribution for the return time of a symmetric random walk [45].

The time until the nth return grows approximately as n^2. The number of returns (threshold crossings) is frequently interpreted as a "result" ("outcome") pertaining to a complex system, and its square (the time of returns, or of threshold crossings), as "expenses" ("efforts," "input"). In this context, the process of random walking is used as a mathematical model for describing various square patterns with expenses growing in proportion to the squared results. This interpretation is interesting in view of the specific analogy between the "condensation" effect in the Brownian process for the number of threshold crossings (the effect consists of "excessively" long return times in the random walk problem), and the resource concentration effect in a small number of large and giant fields.

The second stable non-Gaussian distribution is the Cauchy distribution with the density:

$$f(x) = \frac{1}{\pi(1 + x^2)} \qquad (-\infty < x < \infty) \tag{6-59}$$

The asymptotic expression of the density when $x \to \infty$ is given by the following equation:

$$f(x) \sim \frac{1}{\pi} \frac{1}{x^2} \tag{6-60}$$

The Cauchy distribution coincides with the probability distribution of the ξ_1/ξ_2 ratio of independent random values ξ_1 and ξ_2, which have the same normal distribution with a mathematical expectation of 0 and the variance of 1. The same probability distribution is typical for a tangent

($tg\alpha$) of a random value α, uniformly distributed over the $[-\pi/2, \pi/2]$ interval. This distribution is of interest because the limit distribution (Equation 6-60) is identical in its form to Pareto's distribution when $a = 1$. Many scientists believe that the field size distribution is governed by Pareto's distribution with this parameter. At least, its value of $a = 1$ is the most commonly used in publications. The fact that the Cauchy distribution is symmetric, whereas Pareto's distribution is not, is not important because a truncated Cauchy law is considered for the purpose of field reserve description.

Thus, we have demonstrated a close analogy between the stable distributions different from normal and Pareto-type distributions commonly used for the field size distribution description. This analogy may indicate that processes resulting in stable non-Gaussian distributions may be applied to the model of the field size distribution. The theory of stable non-Gaussian distribution can be used as a reliable tool in the study of the field population distribution. This is not as much a point of the formal mathematicals as it is of qualitative conclusions.

Investigation of the Relationship between the Models and the Oil and Gas Field Size Distribution in Specific Regions

It was illustrated above that the processes of oil and gas field population formation is reflected in the field size distribution function. It is now necessary to examine whether or not the models adequately fit the observations. The model's accuracy can be verified by how well it fits the actual data. Verification is complicated by the fact that the discovered fields (or observations) are not a result of randomized selection from one population and, therefore, do not reflect the true proportions among the various size fields (in other words, the observations are not representative). A field discovery is not a result of the application of a randomized procedure but, rather, a result of the so-called exploration filter [53]. The observed distribution may be substantially different from the true field size distribution. At some stage of exploration, the discovered field size distribution is often similar to a log-normal distribution. This may explain the common belief that field sizes are log-normally distributed. It is not clear from the publications how, if at all, scientists account for the differences between the discovered and initial distributions and make corrections for distortions of the original distribution due to the type of exploration process.

Quite frequently, the proximity between the theoretical and empirical distributions is given only in a qualitative form—by comparing histograms

with the distribution curve. Kontorovich and Demin also presented a graphic comparison, but on a different qualitative examination of the exponential (base e) and Pareto's distributions [24]. Typically, the logarithm of the number of fields within a certain set of reserve range plotted against the average field reserves from that range (or their logarithm), results in a straight line. Thus, if the observations result in a straight line, it may be interpreted as a qualitative agreement between the empiric and the corresponding theoretical distributions. Of course, the inverted selection is of utmost importance.

When observations are used to verify the hypothesis that the size of Q follows a certain theoretical distribution (based on a certain goodness-of-fit criterium), usually only the criterion itself and the result are quoted (i.e., whether or not the hypothesis being tested has been rejected). As a rule, it is not indicated how, if at all, the problem of observation representativeness was solved. It appears that there are no stringent testing procedures for satisfactory observation fit that are not distorted by the exploration process.

Kaufman et al. avoided this problem by comparing the tested hypothesis (in their case, a log-normal field size distribution) with an alternative hypothesis (the hypothesis of gamma-distribution) [16]. It is important to mention that even when observations are representative, it is only possible to distinguish between the log-normal distribution and the other similar distributions by using a very large number of statistical samples [44].

Various authors have given a great deal of attention to the distribution parameters. Most frequently, the studies involve the evaluation of model parameters rather than the examination of model adequacy. This is especially typical for situations where the field size distribution is used as a forecast tool. A statistical testing of the distribution law is not even discussed, as the distribution function is considered given.

It is worth mentioning that even if the size distribution for a selected group of discovered fields coincided with the distribution for all fields in the oil and gas basin, the parameter evaluation for such distributions as Pareto, Weibull, and log-normal is not simple. This problem has been discussed in a large number of publications proposing precise and approximate evaluation techniques [58, 61].

A peculiar feature of Pareto's distribution is that it may not have mathematical expectation and/or variance. As indicated earlier, at large values of the independent variable some distributions behave like Pareto's distribution: their density decreases as the c/Q^{a+1} function. Pareto's distribution is the simplest representative of this distribution family. Because the discovered fields, as a rule, correspond to large Q values, it appears that the fields are distributed according to Pareto's distribution.

When $0 < a \leq 2$, these distributions do not have variance, and when $a \leq 1$, they also do not have a mean value. Infinite dispersion leads to a specific problem with the evaluation of the distribution parameters. The use of mean values may be unreliable. Taking into account the low reliability of a single-number Q forecast, the distribution is characterized not by mean values but, rather, by other probabilistic parameters, such as mode, median, and quantiles.

In order to avoid infinite mathematical expectations, some scientists resort to a truncated Pareto's distribution. In such a case, the reserves of the largest field in the region, or the independently determined value of the ultimate potential, serve as the upper limit. A more accurate forecast may be achieved only by knowing the upper limit, which is not always practically possible. A favorable aspect of this procedure is that the distortion of the sample of observations by the specifics of exploration may be taken into account.

Kontorovich and Demin proposed a technique to evaluate the Pareto's distribution parameters that uses only part of the observations [24]. They address a truncated Pareto's distribution. It is assumed that the size (reserves) of the smallest (Q_o) and the largest ($Q_{\max}$) fields are known, as well as the ratio ($\alpha = R_\alpha/R_{\text{init}}$) of the fields with reserves Q_α and greater (R) to the ultimate potential R_{init}. Of course, Q_α is assumed to be given. Instead of the maximum field size $Q_{\max}$, they suggest using the value $\gamma R_{\text{init}} = (2 \text{ to } 3)Q_{\max}$. They do not, however, explain the reason they selected this ratio. These five input parameters are sufficient to determine the only unknown distribution parameter a. In practice, this technique requires drawing the distribution curve $F(Q)$ through a single point corresponding to Q_α (additionally, there are two *a priori* assigned points: $F(Q_o) = 0$ at $Q = Q_o$ and $F(Q_{max}) = 1$ at $Q = Q_{\max}$, or $Q = \gamma R_{\text{init}}$). Naturally, the accuracy of the evaluation of the a parameter in this case is low; and the statistical examination of the distribution function format is out of the question.

Generally speaking, this technique may be of value in cases where observations are absent, and the input values of the five above mentioned parameters are determined based on *a priori* considerations. Of course, the authors themselves do not do this. They modify the values of all five input parameters in order to come up with a combination that produces the most accurate description of the field size distribution and the reserve distribution among the fields of different sizes. The authors do not explain what is understood as the description accuracy and how it is evaluated. It may be assumed that the actual results were taken into account. If so, it does not make sense to introduce these five input parameters. In this regard, Q_o, $Q_{\max}$, and a values could have been modified in order to determine the best theoretical and observed distribution fit, including reserve distribution among the fields of different sizes.

The procedure for examining whether or not the observed distribution density fits the theoretical density is considered next. Only a superficial approach allows for the comparison of these two distributions. In reality, there is no connection between them because an observed density does not exist. By selecting some class of interval partitioning, and by counting the number of observations within the corresponding intervals, one can obtain some observed distribution density. This density is arbitrary because the partitioning interval length and the selection of the initial partitioning point are arbitrary. The effect of the first step is well known, whereas the effect of the second is rarely mentioned.

The ambiguity in the notion of "observed" distribution results in the uncertainty of the χ^2 value when using this criterion for testing continuous distributions. Yet, with a shift of the initial partitioning point, and with everything else being the same, considerably different χ^2 values can be obtained for the same observations. These values may differ for the same level of significance. In order to avoid this ambiguity, it is necessary to use distribution comparison techniques that are based on the comparison of each individual observation with the respective theoretical value, and on the equal probability intervals instead of the equal length intervals.

In such a case a technique is applied that produces approximate straight lines for the observed frequencies. If the theory is true, the observations must concentrate around a straight line. Each individual observation is plotted. Conversion of the theoretical curve $F(Q)$ into a straight line provides an opportunity to obtain equal probability intervals. The hypothesis of the observation distribution uniformity in the equal probability intervals may be tested using conventional techniques, for instance, χ^2 criterion.

The following procedure may be suggested for the field size distribution comparison. Taking into account that the discovered fields do not reflect the true nature of the distribution, the comparison can be made only within a certain range of the reserve size Q. This assumes that the reserve size Q_α exists and all fields of the Q_α and larger size have already been discovered. This requirement is only satisfied in regions with a sufficiently high exploration maturity of the ultimate potential. In poorly explored regions, the examination is not appropriate. Consequently, the testing results allow us to judge whether or not the observations fit the theoretical distribution only within the area of large Q values. Contrary to the above discussed technique, this method of the distribution parameter evaluation uses not only one point, but the entire interval of the actual $Q \geq Q_\alpha$ curve.

Let us assume that there are N number of fields in a region (this number is unknown). Let us then arrange the fields in order of their increasing size. The m-th field will then have the cumulative frequency $F_m = m/N$, and the last (and the largest) field will have the cumulative

frequency of 1. This cumulative frequency of the m-th value of the F_m is itself a random value. Its average value (when the random value is not limited) is a corresponding value of the initial distribution function:

$$\overline{F}(Q_m) = \overline{F}_m = m/(N+1) \tag{6-61}$$

It is important to note that the mathematical expectation of the difference $F(Q_m) - F(Q_{m-1})$ is equal to $1/(N+1)$, which is constant (the N value, although unknown, is fixed). Therefore, Equation 6-61 provides the necessary solution to the plotting problem of the observed points. Indeed, if all fields are positioned (arranged) in a declining order, the m-th field will have a number $k = N - m + 1$. For the unlimited increase in the random Q value (the case under examination, with no truncated distribution), $1 - F(Q_1) \neq 0$ (but only slightly different from 0). It follows from Equation 6-61 and from the permanency of the mathematical expectation of the $F(Q_{k-1}) - F(Q_k)$ difference, that the values $1 - F(Q_k)$ are located on a straight line ($y = c + dk$, where $c \approx 0$).

Equal probability intervals are determined using a similar approach. They correspond to equal segments of the straight line. They may be determined as Q_i quantile values corresponding to the probability $F(Q_i) = p_i$ levels, which increase (or decrease) by the same value $p_i = p_{i-1} + \Delta p$.

The distribution parameters may be determined in the process of curve construction. The values are taken as estimates that result in a minimum deviation of observations from the straight line. The correlation coefficient, for example, may be used to measure the tightness of the grouping of observations around a straight line. In such a case, it is possible to compare different distributions and their fit to observations because the correlation coefficient is dimensionless. If, for instance, the variance was used, comparison would be difficult because different functions of Q_k, which may have significantly different values, would be examined. The proposed estimate technique may be considered a special case of the least squares method. Its benefit is that, besides taking each individual observation into account, only fields with the reserves $Q \geq Q_\alpha$ may be used for the estimate without any need to deal with truncated distributions. Another advantage is that all parameters (except the Q_o parameter in Pareto's distribution) can be estimated directly from observations. The Q_o parameter in Pareto's distribution may not be estimated for the simple reason that the linear equation in this case has the following format: $Q_o^a/Q_k^a = c + dk$ or $1/Q_k^a = c/Q_o^a + d/Q_o^a k$; therefore, the correlation coefficient does not depend on Q_o.

When some *a priori* information must be used to estimate the distribution parameters, regularization procedures are recommended.

The Bayes's solution to the parameter estimate problem is also based on the use of *a priori* information. It is assumed that the parameter has a random value with a known distribution. For instance, Beylin suggested the use of *a priori* information concerning the a parameter of the Pareto's distribution as a probability density function in the form of the gamma-distribution [4].

We now can demonstrate the fit of the field distribution to the inferred theoretical distributions using, as an example, a region composed of all the regions examined in Part I (except region B).

Correlation coefficient r (for estimating a linear correlation between $1 - F(Q_k)$ and k) and the χ^2 criterion (for the statistical testing of the number of fields of uniform distribution within intervals of equal probability) were used as the numerical fitting criteria. The region is highly explored and, therefore, it was reasonable to assume that all fields with the reserves in excess of 30 MMT have already been discovered. That is why the initial approximation selected was $Q_a = 30$ MMT. In the process of model adequacy testing, the boundary of the completely discovered fields (i.e., the Q_a value) was modified by a gradual shift toward smaller Q values. On the one hand, this resulted in an increase in the number of observations. On the other hand, it provided an opportunity to estimate the degree of change in the observation fit to different distributions under the suggested distortion of the true distribution (caused by the inclusion in the selection of fields not totally discovered within a given size class).

The fit of the field size distribution to Pareto's and Weibull's distributions was examined because they correspond to the most commonly used concepts. The test results on the first criterion are listed in Table 6-3. They indicate that the empiric values of the $1 - F(Q_k)$ distribution tightly group around the $1 - F(Q_k) = c + dk$ straight line, with the correlation coefficient $r = 0.9982$ for the Pareto's distribution and a somewhat lower coefficient $r = 0.9978$ for the Weibull's distribution. Test results on the second criterion are listed in Table 6-4.

Table 6-4 shows the observed (y) and calculated (y') numbers of fields for the equal probability intervals of the tested distributions. It also lists χ^2 values and the corresponding probabilities.

As can be seen from this table, the fit between the theoretical data and the factual data (at least within the domain of the studied Q values) is quite satisfactory: χ^2 values are such that the probability of observed differences between y and y' is rather high if the examined hypotheses are executed. Therefore, the examination results would suggest that the proposed patterns of the concentration of reserves in fields are supported by the actual data, and that the observed distributions may be explained based on the theoretical concepts elucidated above. These results do not prove that either of the two distributions is excluded by the observed data.

Table 6-3
Grouping of Observations near the Straight Line

	Evaluation Results				
		Distribution Parameters			
Distribution	**Straight Line Equation**	**Correlation Coefficient *r***	Q_o	***a***	***b***
$F(Q) = 1 - \left(\frac{Q_o}{Q}\right)^a$	$\frac{1}{Q_k^a} = c + dk$	0.9982	Evaluation impossible	1	—
(Pareto distribution)					
$F(Q) = 1 - e^{-bQ^a}$	$e^{-bQ_k^a} = c + dk$	0.9978	—	0.01	93
(specific case of Weibull distribution)					

Table 6-4
Observed (y) and Calculated (y') Number of Fields within Equal Probability Intervals

Distribution		Number of Fields										
$F(Q) = 1 - \left(\frac{Q_o}{Q}\right)^a$	y	8	14	9	9	13	14	13	13	14	12	$\xi^2 = 4.11$
	y'	11.9	11.9	11.9	11.9	11.9	11.9	11.9	11.9	11.9	11.9	$0.9 < p_5(4.11) < 0.95$
$(Q) = 1 - e^{-b}$	y	11	14	9	10	13	14	14	17	11	11	$\xi^2 = 4.23$
	y'	12.4	12.4	12.4	12.4	12.4	12.4	12.4	12.4	12.4	12.4	$0.7 < p_5\ (4.23) < 0.8$

Although the Pareto's distribution has higher r and $p(\chi^2)$ values, it is still not preferable to Weibull's distribution.

We will conclude with a brief discussion of the distribution parameters. The a parameter in the Weibull's distribution is of particular interest. We obtained $a < 1$ following evaluation. This agrees with the theory. Thus, in this respect, one may assert that the observations fit the theory. The a parameter in Pareto's distribution is also important. The a estimates for the U.S. oil and gas basins in this case were published by Kontorovich and Demin [24]. As we indicated, these estimates are not particularly reliable, especially for poorly explored basins. Our data indicates that at $a = 1$ the correlation coefficient is very close to 1; fine-tuning the a value does not make sense. For the well-explored Illinois Basin (exploration degree = 0.82) and Midcontinent Basin (exploration degree = 0.93) they list the values 0.90 and 0.97 for the a parameter, respectively (i.e., close to 1).

Shpillman indicated that the evaluation of a based on the discovered fields grows with the exploration degree and tends to 1 [53]. This particular case of $a = 1$ for Pareto's distribution was used for describing the field size distribution. Clearly, this case is the only one which is correct (in reference to Pareto's distribution). This will be further discussed when analyzing different approaches to describing the relations between fields of different sizes. In connection with the problems associated with using Pareto's distribution to approximate the distribution of fields, another interesting parameter is Q_o. In many publications this parameter is not evaluated based on the observations, but on other considerations that are mostly related to determining which minimum accumulation size is commercial. Generally speaking, these considerations may not have any relevance to the Q_o parameter, the physical meaning of which was explained. It is of value for the analysis and evaluation of undiscovered potential (forecast resources).

In conclusion, we will summarize the first approach to the description of relations among various size fields. This is a probabilistic approach using the language of probability theory and mathematical statistics, with a major emphasis on the distribution function. This approach was discussed in great detail due to the fact that it is rather old, very popular, and has become practically conventional. For this reason, we studied it carefully, paying particular attention to the theoretical problems. Strange as it may seem, they slipped the attention of previous researchers and remained totally unexplored. This gap must be eliminated. At the same time, if we remain within the constraints of probabilistic concepts, with their laws of random dispersion, it is difficult to find a satisfactory geologic concept that would explain some patterns in the field size distribution.

There is an extremely low probability of events such as the formation of fields of even a small size. Let us again consider Pareto's distribution

as an example. Kontorovich and Demin assumed Q_o = 0.13 MMT for the field size distribution of all North American oil and gas basins [24]. Under this assumption, at a = 1, the probability of formation of fields within the 0.13 to 1.6 MMT reserve range is 0.92. The probability of formation of the 1.6 MMT field is close to 0.05. Events with this probability are considered almost impossible. This means that the fields with greater reserves should not have formed. As for the very large and giant fields, the probability of their formation is quite negligible. Yet, there are quite a number of such fields in the world.

Remaining within the constraints of this approach, it is also impossible to explain the values obtained for the distribution parameter. For instance, why is the value of the a parameter of Pareto's distributions close to 1? This does not necessarily follow from the concepts we proposed, which also explained the meaning of this parameter. As an example, according to one of the concepts $a = \alpha/\beta$. It is not clear, however, why the failure intensity α and the growth intensity parameter β turn out to be commensurable values.

For the above reasons the description of the various size field relations using the language of distribution functions is far from being universal. It is quite possible that rather than the probabilistic, some other laws are at work here; laws that may better explain the fact that some field sizes are more frequent than others, and the observed patterns in the relations among the various sizes. The following is a discussion of this problem.

Field Population as an Ordered Self-Organizing System: Non-linear Relations in the System

This and the following sections are devoted to the investigation of the other processes resulting in certain relations among field sizes. These relations are no longer based on probabilistic assumptions. They are based on a concept of a regular step-wise (jump-like) change in the field size. For this purpose, we will analyze the structure of a ranked variational field series. This is why this approach may be called structural rather than probabilistic. It is structural because it studies the process of emergence of ranked structures. The language of probability theory is not suited to the description of such processes and, for this reason, poorly reflects their substance. From a structural perspective, the field association study technique based on the analysis of distribution functions cannot be stringently substantiated. A field set (population) may not necessarily reflect the frequently occurring phenomena resulting from random

dispersion. (Keep in mind that a field is formed in certain physicochemical and geologic conditions.)

We will consider a field set (population) an open natural system capable of self-organizing at the expense of external energy. Earlier, we discussed the self-organizing process, which is the process whereby an order emerges within a system. New, more-or-less stable, structures continuously emerge in a self-organizing system. In other words, in such a system order gradually emerges from chaos, and structures emerge where previously there were none. Self-organizing processes may occur only in open systems. Open systems are those that can exchange energy, matter, or momentum with the surrounding medium. The structures emerge far from the state of equilibrium according to certain non-linear laws. The system can form ordered, highly co-operative structures because external restrictions keep it substantially far from equilibrium. The emergence of a new structure in an open system is always a result of instability.

The reason stable states become unstable is that fluctuations spontaneously emerge in the system itself. They appear as a result of complex interconnections among elements of the system. These interactions are non-linear (i.e., a reaction to an external action not in proportion to the amount of this action). Instabilities may occur as a result of these non-linear actions. The loss of stability is followed by the destruction of the old structure and the formation of a new one: the system changes rapidly (jumps) to a new state.

The self-organizing process evolves in the direction of increasing complexity, from less-ordered to more-ordered structures. Another important characteristic of this process is a continuously increasing capacity for assimilating free energy more intensely.

When considering a field set (population) as a natural system capable of self-organizing, it is necessary to recognize that it represents an open system. This is important because deposits (fields) form in an environment with a constant influx of energy and matter which form hydrocarbons. Deposits (fields) constitute unstable structures by default. Any type of thermodynamic equilibrium in this situation is out of the question. To maintain a system far from equilibrium, work is permanently occurring in nature. The mechanism can be described as follows. External energy entering the sedimentary basin is received and transformed by the components of the system. Various types of internal energy are formed from external energy sources (tectonic motions, etc.). Formation of deposits (fields) is associated with the formation of structures, which adapt themselves to the transfer of energy from higher to lower levels. A new fluctuating energy wave arriving from outside forms a new movement, or a new organization. All of this determines the processes of ordering the structure of the deposits (fields) and of making the system self-organized.

The general direction of this process is known. Just as the evolution of a self-organizing process progresses from less-orderly to more-orderly structures with an increasing organizational complexity, so does the deposit (field) formation process, which is directed toward the concentration of initially dispersed hydrocarbons into ever larger accumulations. It is possible that the presence of very large and giant fields may be associated with their capacity for assimilating more external energy. Indeed, larger deposits (fields) require more energy for their formation.

The step-wise, jump-like evolution of a self-organizing process implies an important property that is pertinent to a population of fields. This property is discreteness. If we place the fields in order of their decreasing reserves, we will obtain a variational ordered (ranked) set, where the field sizes do not follow one after another continuously, but are separated by different intervals associated with the amount of jump. This set does not include only large and small fields; large fields are not positioned adjacent to small fields, but are separated from them by intermediate (medium size) fields. Thus, such a population has a certain order in the variation of field sizes, and this discreteness displays a regular nature. This means that field reserves Q_k and Q_{k+1}, with respective sequential numbers k and $k + 1$ in this set, are in some way associated. This connection reflects a general self-organizing process, or the formation process of ever-larger fields from a dispersed state.

Like the evolution of order in the system, this process does not have a limit. This is supported by the known phenomenon of the dependency of the largest field reserves Q_1 on the ultimate potential: the greater the ultimate potential, the greater the Q_1. With increasing ultimate potential (which is undoubtedly associated with the amount of energy), the ordered (ranked) set increases from the left, wherever larger fields appear at the beginning of the set (as if the increase in ultimate potential occurs on the basis of adding large fields). This is also an indication of a regular connection between the volume of reserves in the adjacent fields of this set. Accordingly, it can be said that the mechanism of self-organizing, which determines the achieved degree of hydrocarbon concentration within an individual field (size of reserves), is a mechanism with memory. This mechanism provides the connection between the reserve sizes of adjacent fields in this set.

The general "goal" of this self-organizing process is the transition from a dispersed state to the formation of ever-larger deposits (fields). With this in mind, let's look at the "purposeful" behavior of the field sizes in the ranked set toward achieving this goal. As an example, the behavior of water droplets within a stream coming out of a garden hose is subordinated to the target (goal) the stream is directed at. The bulk of droplets will merge into a single large stream and reach the goal. Some

smaller aggregates, however, will lose connection with the main stream and may not reach the goal. Their behavior in the stream is slightly dependent, or even completely independent, of the main stream. The smaller the aggregates the droplets merge into, the less they are subordinated to the main stream, and the greater the degree of freedom they possess.

It can be analogized that the large fields, to a greater degree, are subordinated to the general goal, and the self-organizing process (which led to their emergence) is better expressed in them. At the same time, small fields are less affected by this process. The connections between them in a ranked set (population) become weaker. As a result, the difference in the size of reserves becomes smaller and they become more uniform in size.

Thus, the field sizes in the ranked descending set are connected, or $Q_k = f(Q_{k+1})$. This connection is non-linear and the difference $Q_k - Q_{k+1}$ monotonously decreases with increasing k. The format of the $Q_k = f(Q_{k+1})$ function may be determined, as a first approximation, based on the following considerations. First, Q_k increases with increasing Q_{k+1}. Therefore, the equation connecting them must reflect this proportionality. Accordingly, the second term must take into account the aforementioned interaction between Q_k and Q_{k+1}, which results in a non-linear connection between them. Thus:

$$Q_k = aQ_{k+1} + bQ_kQ_{k+1} \tag{6-62}$$

In a special case of $b = 0$,

$$Q_k = aQ_{k+1} \tag{6-63}$$

Another equation of interest (discussed in detail in the next section) corresponds to the special case of $a = 1$. Then:

$$\Delta Q = Q_k - Q_{k+1} = bQ_kQ_{k+1} \tag{6-64}$$

This equation shows that the jump in the reserve size ΔQ is proportional to the interaction.

Thus, the analysis of a field set as an ordered (ranked) self-organizing system suggests the existence of a regular connection within the system, corresponding to the interaction of deposit (field) sizes.

The evaluation results of these parameters for models 6-62 and 6-64 are listed in Table 6-5. This table also provides the correlation coefficients between the left and right sides of the equations and the variances σ^2 of the observed deviations from the values given by the right side of the equations. These data enable us to evaluate the fit between the models

Table 6-5
Comparison of Results Obtained for Models 6-62 and 6-64, Describing Field Size Interrelations in the Ranked Set of Fields (Fit with Observations)

Model	Parameters		Correlation Coefficient r	Variance σ^2
	a	b		
$Q_k = aQ_{k+1} + bQ_kQ_{k+1}$	1.0248	$2.8134*10^{-4}$	0.99931	335.19
$Q_k - Q_{k+1} = bQ_kQ_{k-1}$	—	$2.8696*10^{-4}$	0.99765	343.94

and the observations. The same data used earlier were used as the observations. The results presented in Table 6-5 show that the models are adequate to observations. In other words, the above-mentioned connections are indeed in a ranked field set (at least within that part containing the largest fields). These connections are described well by model 6-64. The more general model 6-62 yields results practically indistinguishable from its special case; that is, Equation 6-64: correlation coefficients r and variances σ^2 differ only slightly. The reason for the identical results is that $a \approx 1$ in model 6-62. This phenomenon is explained in the next section.

Compensation Principle of the Emergence of Order (Ranking) in a System of Fields

In this section, another approach based on the self-organizing concept is presented. Here, again, the pattern in the field size variation is not associated with probabilistic causes, but with field order (ranking) within the general system. For ranking, in this case, let us place fields into a variational set in descending order of reserves and number the fields. Then the reserve size turns out to be a function of the sequential number k, or $Q_k = f(k)$. This concept is connected with the concept examined in the preceding section because it is easy to switch from $Q_k = f(k)$ to $Q_k = f(Q_{k+1})$. In this case, however, the dependence $Q_k = f(k)$ was obtained based on different considerations.

It is important to mention that American scientists do not analyze the field distribution. Instead, they analyze the connection between the field sequential number in a ranked set of reserves (for example, see Ivanhoe [60]). Accordingly, they analyze correlation graphs $Q_k = f(k)$ rather than frequency diagrams.

In order to explain the studied relation, keep in mind that large fields are situated selectively. Certain conditions are necessary for their emergence, which are only satisfied within limited areas. Small fields are encountered anywhere; they may be adjacent to the large fields, medium fields, etc. There are no specific restrictions on their emergence. Therefore, each field appears to have its own area of existence within which all factors necessary for the emergence of the particular field are combined. An analogous biological concept is habitat, referring to the field habitat (dispersion in space). Large field habitat is quite narrow (restricted). At the same time, small fields are dispersed throughout the prospective area of an oil and gas basin. It may be assumed that the larger the field, the smaller its habitat.

The process of purposeful hydrocarbon concentration into different size accumulations, as a self-organizing process, must be characterized by a certain correlation between the field size and the size of its habitat. Let us designate the latter as S_k. It may be assumed from the analogy with thermodynamic equilibrium that the field's size growth is compensated by the narrowing of its habitat. As a result, the product $Q_k S_k$ for the system is constant:

$$Q_k S_k = \text{const.} \tag{6-65}$$

This situation is similar to the equi-probable states (p = const) corresponding to maximum entropy. As previously mentioned, the state of order has the maximum entropy, just as a state of equilibrium between the isolated system and the surrounding medium has maximum entropy. This pattern is established for large systems which are optimally organized and is known in advance. Therefore, a constant product may be considered a characteristic indication of an optimally organized system. A deviation from this equality may be considered an indication that the process deviated from its optimal course.

The basic assumption, therefore, is that the result of the self-organizing process in the course of purposeful hydrocarbon concentration into an integral accumulation system is the compensation condition, so that:

$$Q_k S_k = Q_{k+1} S_{k+1} = C \quad (k = 1, 2, 3, \ldots) \tag{6-66}$$

The simplest assumption regarding the S_k is that the field habitat linearly contracts to a point with increasing field size (Figure 6-1): $S_k = \alpha k + \beta$. In other words, the habitat of a field numerated k is smaller than the habitat with a field number $k +$ by a constant; that is:

$$S_{k+1} - S_k = \alpha \quad (k = 1, 2, 3, \ldots) \tag{6-67}$$

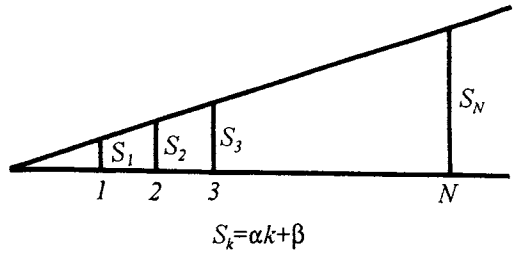


Figure 6-1. Narrowing of habitat S_k in the process of purposeful hydrocarbon concentration.

Thus:

$$Q_k - Q_{k+1} = \frac{C}{S_k} - \frac{C}{S_{k+1}} = C\frac{S_{k+1} - S_k}{S_k S_{k+1}} = \frac{\alpha}{C} Q_k Q_{k+1} \tag{6-68}$$

This equation has already been derived (see Equation 6-64). It follows from Equation 6-66 that:

$$Q_k = \frac{C}{\alpha k + \beta} \tag{6-69}$$

Equation 6-69 describes a connection between the field size Q_k and its sequential number k in the ranked series. It explains the circumstance discussed, namely, that the characteristic parameter a in Pareto's distribution is usually equal to one. This is discussed in more depth later. For now, we will show that Equation 6-69 may be derived based on similar properties that were used when studying the frequency of occurrence of different size fields. In the special case of $a = 1$, the following is true if the fields are distributed under Pareto's law. Let us subdivide all fields into classes with the boundaries Q_o, mQ_o, m^2Q_o, . . . (the upper boundary value in any class is m times higher than the lower boundary value). Let us further assume that the combined reserves of all fields within a class equal the product of a field number in the class and the mid-interval covered by each class, indicated by those upper and lower boundaries. Then, the combined reserves of these classes are the same for all classes. In particular, this property is often used when constructing a field distribution with m assumed equal to 2 ($m = 2$). In this case, starting with the largest class (where the field number is equal to one), the field number

in each class doubles, and the interval occupied by each class, correspondingly, becomes twice as small. This is how the field distribution by equivalent classes is determined.

Similarly, for a variational field set one can assume that the combined reserves of the subsequent m fields are equal to the reserves of the first field Q_1, and that the combined reserves of the subsequent m^2 fields are equal to the combined reserves of their preceding m fields and, respectively, the reserves of the first field. This equality is preserved for the m^3 fields that follow, for the m^4 fields following the previous m^3 fields, and so forth. To be more precise, the equality discussed here, as in the analysis of equivalent sets, is approximate. This means that the combined reserves for each set of m^i fields (where $i = 1, 2, 3, \ldots$) is defined as the product of reserves of the central field (from the m^i fields) by their number m^i, or $Q_{t_i} m^i$ (Figure 6-2). It follows from this condition of equality of the defined combined reserves in each set of m^i fields (where $i = 1, 2, 3, \ldots$) that:

$$Q_{t_i} m^i = Q_1 \tag{6-70}$$

The sequential numbers t_i for the central field of each set of m^i fields are equal to:

$$t_i = \frac{m^i - 1}{2} + 2 + m + m^2 + \ldots + m^{i-1} = (m^i - 1)\left(\frac{1}{2} + \frac{1}{m-1}\right) + 1$$

Thus,

$$m^1 = 2\frac{m-1}{m+1} t_i + \left(1 - 2\frac{m-1}{m+1}\right) \tag{6-71}$$

It follows from Equation 6-70 that:

$$Q_{t_i} = \frac{Q_1}{m^i} = \frac{Q_1}{\left[2\frac{m-1}{m+1} t_i + \left(1 - 2\frac{m-1}{m+1}\right)\right]} \tag{6-72}$$

Returning to the conventional $t_i = k$ and assigning $\gamma = 2(m - 1)/(m + 1)$:

$$Q_k = \frac{Q_1}{\gamma k + (1 - \gamma)} \tag{6-73}$$

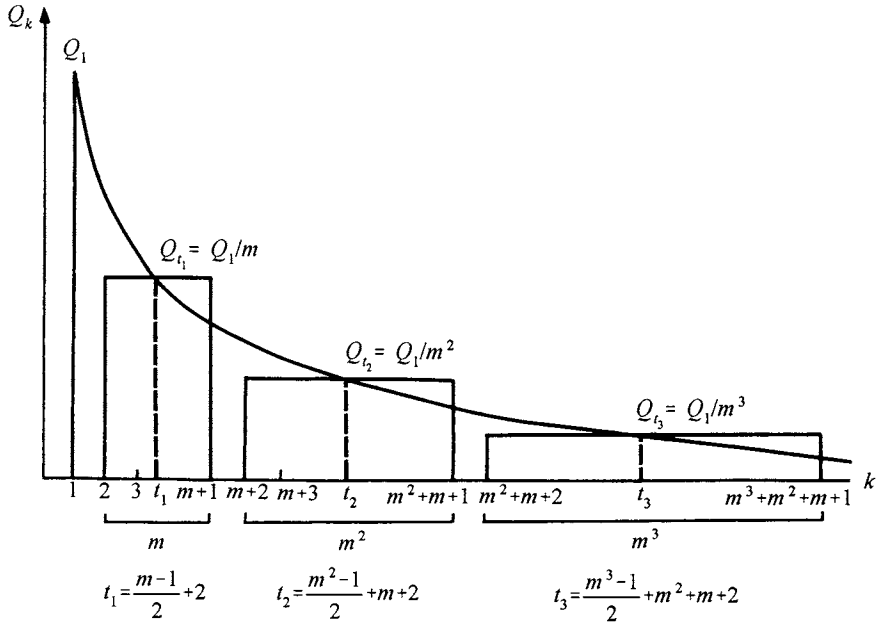


Figure 6-2. A schematic diagram clarifying field reserve relationships in their variational field set.

The above equation is analogous to Equation 6-69, Q.E.D. (which was to be demonstrated).

Thus, a field set satisfying Equation 6-73 possesses the property we are analyzing. The value of m is determined by the γ parameter in Equation 6-73. This property of sets makes it easier to evaluate the initial potential. Conversely, if the initial potential resources are known, this property allows us to easily determine the number of fields.

Equation 6-73 may be re-written as $Q_k = 1/(a + bk)$, where $a = (1 - \gamma)/Q_1$ and $b = \gamma/Q_1$. We evaluated the parameters of this model using the same fields discussed earlier, and obtained $a = -9.8 * 10^{-5}$, $b = 3 * 10^{-4}$, and $\sigma^2 = 499.5$. Judging by the σ^2 value, the model is in good agreement with the reality. Given the model parameters, it is not difficult to determine the Q_1 and γ values and, subsequently from $\gamma = 2(m - 1)/(m + 1)$, the m value. In our case, $\gamma = 1.485$ and $m = 6.7629 \approx 7$. This means that the seven fields after the first one have combined reserves approximately equal to the reserves of the first field. The same is true for the subsequent 49 fields, the subsequent 343 fields, and so on.

Equation 6-73, which illustrates the dependence of the volume of a field's reserves on the field's sequential number in a ranked set, was derived on the assumption of the linear nature of change in the distribution of deposits (fields). Since this is the most natural assumption, it may be assumed that the pattern for Equation 6-73 is the most common. This equation, as discussed later, is a direct analogue of Pareto's distribution with the characteristic parameter $a = 1$. In published literature, field distributions of this type are most commonly used.

The area of habitat, however, may change in size in a non-linear manner. It may be assumed that area contraction, to a degree, may follow according to different patterns that are described using simple functions. Table 6-6 lists the simplest of such non-linear functions and the respective relations between the field reserves and their sequential number in a ranked set. To make this table complete, the analyzed case is also included. The nature of the curves showing the change in S_k can be visualized from the included graphs, whereas the nature of the Q_k curves can be judged from the dQ/dk value. The table shows that sometimes dQ/dk decreases proportionally to Q (curve 3), and sometimes to Q^2 (curves 1 and 2), or to Q to a higher and lower power (curve 4). Sometimes it depends not only on Q, but also on the sequential number k (curves 5, 6, and 7).

Discussion of the Study Results of the Formation Mechanism for Different Field Size Distribution Types

We have examined the patterns of relations between the reserve volumes of different fields from two different viewpoints.

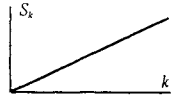
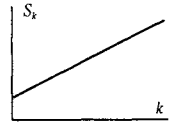
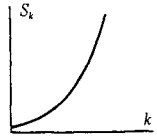
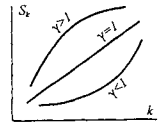
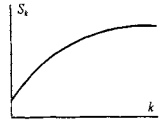
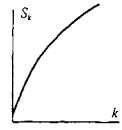
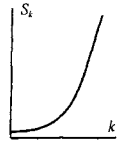
The first viewpoint treats the field formation process, for an interconnected set of fields, as a stochastic process where the differences in the size of reserves are associated with random causes. From this perspective, the field population is characterized by a certain distribution function, or by a certain frequency of occurrence of fields of certain size.

The second viewpoint considers the formation of fields in terms of certain goal-oriented, self-organizing processes. These processes result in particular relationships among the field sizes.

Thus, these two viewpoints are radically different.

The formal description can be considered independently of the content. In both cases, the analysis of field size patterns may be conducted using three forms of description: as a distribution function and as functions of $Q_k = f(Q_{k+1})$ and $Q_k = f(k)$ types. It is always possible to switch from one form of description to another. If fields form due to a process

Table 6-6
Functions Describing the Type of Change in the Field's Habitat S_k and Respective Formulas Determining Field Reserves Q_k as a Function of its Number in a Ranked Field Set

Sequential Number of Function and Curve	S_k	Characteristic Appearance of $S_k = f(k)$ Curve	$Q_k = f(k)$	Speed of Change for Curve $Q_k = dQ/dk$
1	αk		$\frac{C}{\alpha k}$	$-\frac{\alpha}{C}Q^2$
2	$\alpha k + \beta$		$\frac{C}{\alpha k + \beta}$	$-\frac{\alpha}{C}Q^2$
3	$\alpha e^{\beta k}$		$\frac{C}{\alpha e^{\beta k}}$	$-\beta Q$
4	$(\alpha k + \beta)^{\gamma}$		$\frac{C}{(\alpha k + \beta)^{\gamma}}$	$-\frac{\alpha\gamma}{C^{1/\gamma}}Q^{(1+1/\gamma)}$
5	$\gamma - \alpha e^{-\beta k}$		$\frac{C}{\gamma - \alpha e^{-\beta k}}$	$-\frac{\alpha\beta}{C}e^{-\beta k}Q^2$
6	$\gamma - \alpha k^{-\beta}$		$\frac{C}{\gamma - \alpha k^{-\beta}}$	$\frac{-\alpha\beta}{C}k^{-(\beta+1)}Q^2$
7	$\alpha e^{\beta k - \gamma}$		$\frac{C}{\alpha e^{\beta k} - \gamma}$	$-\frac{\alpha\beta}{C}e^{\beta k}Q^2$

of self-organization, the distribution function is not appropriate because change does not contribute to the goal-driven, self-organization process. In this case, the use of probability theory language is also not applicable because neither is the distribution function. Conversely, if one accepts the stochastic nature of field emergence, it does not make sense to discuss their regular connections in a ranked set, or the connection of their reserves and their sequential number in this set. In this case, only the distribution function should be applied.

It is not difficult to switch from one description format to another. Table 6-7 lists all three description formats for typical situations. From the derived distribution functions, we produced the other corresponding description formats; or, moving in the opposite direction, from the functions analyzed above, we produced the corresponding distribution functions.

Some of the results are discussed here. Column 2 of Table 6-7 begins with Pareto's distribution function with the characteristic parameter $a = 1$. It is difficult to explain why field distributions are usually in agreement with this specific case of Pareto's distribution, for it appears there is no definitive reason for this agreement. The situation becomes clearer if the explanation is based not on the premises used for the derivation of the distribution function, but on the premises used when deriving the $Q_k = f(k)$ equation, which corresponds in its format to the specific case under consideration (Pareto's distribution; line 1, column 4). The linear nature of the habitat contraction provides a plausible explanation of this fact. Thus, fields form as a result of the self-organizing process, and the distribution function is a purely formal description.

Pareto's distribution is a probabilistic form of Zipf's law. The equations in column 4, lines 1 and 2, of Table 6-7 provide a rank representation of Zipf's law. This particular form was used by Zipf, usually with the parameter $1/a = 1$ (as shown in line 1). The description format $f(Q)N = aQ_o^a N/Q^{a+1}$, where $f(Q)$ is the distribution density and N is the number of fields, is a frequency representation of Zipf's law ($\gamma = a + 1$ is the Zipf's law parameter). This format is widely used in scientology and is often called the Lotka law (usually for the case of $a = 1$).[3] The Zipf law is widely discussed in many publications. It describes a broad range of phenomena mostly related to human activities and natural and social systems (growth of cities, income distribution, taxonomy, the skill of playing golf, distribution of examination results, distribution of scientists by the number of their publications, etc.). Zipf's law was also used for studies of various types of deposits (copper, lead, gold, uranium, and oil and gas for which $1/a = 1$) [64, 65].

Semikhodski and Timoshin studying the rank distribution of fields in the Dnieper-Don Basin in the Ukraine, assumed that the compliance of a

system with Zipf's law is an indication of its optimum organization [48]. These authors believe that non-compliance of a large system with Zipf's law may be considered an indication of an incompleteness of partial formation or of partial degradation. This also may occur if the system has more than one independent source of formation [48, p. 10].

Functions $Q_k = f(Q_{k+1})$ (lines 4 and 5 of Table 6-7), which are also important, have the following format:

$$Q_k = \beta_1 Q_{k+1} + \beta_2 Q_k Q_{k+1} \tag{6-74}$$

The difference is that $\beta_1 > 1$ on line 4 and $\beta_1 < 1$ on line 5. When $\beta_1 = 1$, this equation converts into the expression on line 1, corresponding to a special case of Pareto's distribution. When $\beta_2 = 0$, it converts into the expression on line 3. Thus, one can observe the similarities and differences of the distribution functions entered on these lines and of the $Q_k = f(k)$ functions. They may almost coincide when β_1 and β_2 parameter values are close to the indicated boundaries and convert to one another when these parameters change.

The equation on line 6 also deserves attention:

$$\ln(\beta_o Q_k) = \ln(\beta_o Q_{k+1}) + \beta_2 \ln(Q_k \beta_o) \ln(Q_{k+1} \beta_o) \tag{6-75}$$

Often the logarithms of reserve volumes are used.

If logarithms are used, Equation 6-75 becomes indistinguishable in its format from Equation 6-74:

$$z_k = z_{k+1} + \beta_2 z_k z_{k+1}$$

where:

$$z_k = \ln\left(\frac{e^{\beta}}{Q_o} Q_k\right); \beta_2 = \frac{1}{N\beta}$$

Table 6-7 includes new distribution functions and new $Q_k = f(Q_{k+1})$ and $Q_k = f(k)$ functions, which we have not analyzed previously. It is important to examine how these new models fit the observations and to compare the convergence of all the obtained results.

As previously, the fit between the field distribution and these newly derived theoretical distributions was tested using two criteria. The first one evaluated the concentration of the observations around the $1 - F(Q_k)$

Table 6-7
Different Description Formats for the Relationships among Sizes of Deposits (Fields)

Line	Distribution Function $F(Q)$	$Q_k = f(Q_{k+1})$ Function (Interaction)	Function $Q_k = f(k)$ (Habitat Contraction)	Comments
1	2	3	4	5
1	$F(Q) = 1 - \frac{Q_o}{Q}$	$Q_k = Q_{k+1} + \frac{1}{NQ_o} Q_k Q_{k+1}$	$Q_k = \frac{NQ_o}{k}$	Special case of Pareto's distribution, Lotka's law
2	$F(Q) = 1 - \left(\frac{Q_o}{Q}\right)^a$	$Q_k^a = Q_{k+1}^a + \frac{1}{NQ_o^a} Q_k^a Q_{k+1}^a$	$Q_k = \frac{N^{1/a} Q_o}{k^{1/a}}$	Pareto's distribution, Zipf's law
3	$F(Q) = 1 - e^{-\lambda Q^n}$	$e^{\lambda Q_k^n} = e^{\lambda Q_{k+1}^n} + \frac{1}{N} e^{\lambda Q_k^n} e^{\lambda Q_{k+1}^n}$	$Q_k = \left(\frac{1}{\lambda}\right)^{1/n} (\ln N - \ln k)^{1/n}$	Weibull distribution
4	$F(Q) = 1 + \ln\left[1 - \frac{Q_o(1 - e^{-1})}{Q}\right]$	$Q_k = e^{-1/N} Q_{k+1} + \frac{1 - e^{-1/N}}{Q_o(1 - e^{-1})} Q_k Q_{k+1}$	$Q_k = Q_o(1 - e^{-1}) \frac{1}{1 - e^{-k/N}}$	
5	$F(Q) = 1 - \ln\left[\frac{(e-1)Q_o}{Q} + 1\right]$	$Q_k = e^{1/N} Q_{k+1} + \frac{e^{1/N} - 1}{Q_o(e-1)} Q_k Q_{k+1}$	$Q_k = Q_o(e-1) \frac{1}{e^{k/N} - 1}$	

6	$F(Q) = 1 - \dfrac{\beta}{\ln(e^{\beta} Q/Q_o)}$	$\ln\left(\dfrac{e^{\beta}}{Q_o} Q_k\right) = \ln\left(\dfrac{e^{\beta}}{Q_o} Q_{k+1}\right) + \dfrac{1}{N\beta} \ln\left(\dfrac{e^{\beta}}{Q_o} Q_k\right) \ln\left(\dfrac{e^{\beta}}{Q_o} Q_{k+1}\right)$	$Q_k = Q_o e^{(N/k-1)\beta}$	At $m = 2$, the distribution is similar to the log-normal; at $m = 1$, it converts into Pareto's distribution
7	$F(Q) = a \ln \dfrac{Q}{Q_o}$	$Q_k = e^{1/(aN)} Q_{k+1}$	$Q_k = e^{1/a} Q_o e^{-k/(aN)}$	
8	$F(Q) = 1 - e^{-1\left(\ln \frac{Q}{Q_o}\right)^m}$	$Q_k = Q_o e^{\left[\ln\left(e^{-a\left[\ln \frac{Q_{k+1}}{Q_o}\right]^m} - \frac{1}{N}\right)^{-1/a}\right]^{1/m}}$	$Q_k = Q_o e^{\left[\ln\left(\frac{N}{k}\right)^{1/a}\right]^{1/m}}$	

$= c + dk$ straight line. The corresponding correlation coefficient r was used as a criterion. The second criterion was the 0^2 value obtained through statistical testing of the uniform distribution of the number of fields using the equi-probable intervals.

Table 6-8 shows the test results corresponding to the first criterion. All three correlation coefficients are high. A value of $r = 0.9982$ was previously obtained for Pareto's distribution and $r = 0.9978$ for Weibull's distribution. Although all five distributions have very high r values (> 0.99), of greatest interest are the three distributions with the highest r values (all equal to 0.9982). The lowest r value belongs to the distribution on line 6 of Table 6-7.

Table 6-9 shows the test results corresponding to the second criterion. Although the test results do not refute any one of the derived distributions, it is important to note that, here again, the probability of χ^2 values for the distribution on line 6 of Table 6-7 is much lower than for the remaining two. As previously discussed, the test results for Pareto's and Weibull's distributions showed that Pareto's distribution and the distribution on line 5 of Table 6-7 have the highest probabilities $p(\chi^2)$. Thus, not a single distribution contradicts the available data. Nevertheless, Pareto's distribution and the distribution on line 5 of Table 6-7, appear to be preferable.

As shown before, in the $Q_k = f(Q_{k+1})$ description format (corresponding to Pareto's distribution and the distribution on line 5 of Table 6-7; see Table 6-5) the correlation coefficients were equal to 0.9976 and 0.9993, respectively. The second function has a slightly (insignificantly) higher coefficient of correlation. We discussed the distribution of line 5 rather than line 4 of Table 6-7 because $\beta_1 > 1$ ($\beta_1 = 1.0248$). The difference from 1 is very small, so that one distribution is almost indistinguishable from another. This conclusion is supported by all the criteria. Thus, Pareto's distribution does not appear to be preferable and the distributions on lines 4 and 5 of Table 6-7 successfully compete with it. They also represent a more general case because β_1 in the function 6-74

$$Q_k = \beta_1 Q_{k+1} + \beta_2 Q_k Q_{k+}$$

is arbitrary, whereas Pareto's distribution represents a special case of this function at $\beta_1 = 1$.

To complete the discussion, we will review the estimated coefficients of the distribution on line 6 of Table 6-7. These are the parameter estimates of the function corresponding to this distribution and the related Q_k and Q_{k+1}. This function is listed in column 3 of Table 6-7 and can be rewritten as follows:

Table 6-8
Grouping of Observations around the Straight Line

Line Number in Table 6-7	Distribution	Straight Line Equation	Estimates		
			Correlation Coefficient r	Distribution Parameters	
				Q_o	β
4	$(Q) = 1 + \ln\left[1 - \frac{Q_o(1-e^{-1})}{Q}\right]$	$\ln\left[1 - \frac{Q_o(1-e^{-1})}{Q_k}\right] = c + d$	0.9982	0.1	—
5	$(Q) = 1 - \ln\left[\frac{(e-1)Q_o}{Q} + 1\right]$	$\ln\left[\frac{(e-1)Q_o}{Q_k} + 1\right] = c + dk$	0.9982	0.1	—
6	$F(Q) = 1 - \frac{\beta}{\ln(e^{\beta}\, Q/Q_o)}$	$\frac{\beta}{\ln(e^{\beta}\, Q/Q_o)} = c + dk$	0.9960	11	0.01

Table 6-9
Observed (y) and Calculated (y') Number of Fields within Equal Probability Intervals

Line Number in Table 6-7	Distribution		Number of Fields										
4	$(Q) = 1 + \ln\left[1 - \frac{Q_o(1 - e^{-1})}{Q}\right]$	y	11	14	9	10	15	14	13	18	9	11	$\chi^2 = 6.15$
		y'	12.4	12.4	12.4	12.4	12.4	12.4	12.4	12.4	12.4	12.4	$0.5 < p_5(6.16) < 0.7$
5	$(Q) = 1 - \ln\left[\frac{(e-1)Q_o}{Q} + 1\right]$	y	11	14	9	10	13	14	14	16	10	12	$\chi^2 = 3.75$
		y'	12.3	12.3	12.3	12.3	12.3	12.3	12.3	12.3	12.3	12.3	$0.8 < p_5(3.75) < 0.9$
6	$(Q) = 1 - \frac{\beta}{\ln(e^{\beta}Q/Q_o)}$	y	5	12	11	11	16	18	16	16	8	11	$\chi^2 = 12.13$
		y'	12.4	12.4	12.4	12.4	12.4	12.4	12.4	12.4	12.4	12.4	$0.05 < p_5(12.13) < 0.10$

$$Q_k = \frac{1 + [1/(N\beta)]\ln(e^{\beta}/Q_o)}{1 - [1/(N\beta)]\ln(e^{\beta}/Q_o)} \ln Q_{k+1}$$

$$+ \frac{1/(N\beta)}{1 - [1/(N\beta)]\ln(e^{\beta}/Q_o)} \ln Q_k \ln Q_{k+1}$$

$$+ \frac{[1/(N\beta)][\ln(e^{\beta}/Q_o)]}{1 - [1/(N\beta)]\ln(e^{\beta}/Q_o)}$$

or in a shorter form:

$$\ln Q_k = \beta_1 \ln Q_{k+1} + \beta_2 \ln Q_k \ln Q_{k+1} + \beta_3$$

The following evaluations were obtained: $\beta_1 = 0.6142$, $\beta_2 = 0.0450$, and $\beta_3 = 0.8254$, with the correlation coefficient $r = 0.9994$ and variance $\sigma^2 = 0.00104$. As we can see, the correlation coefficient is very high and the corresponding variance is low.

As far as forecasting is concerned, the $Q_k = f(k)$ format is convenient for predicting specific values for the expected field reserves. In reality, it is used to forecast the senior members of the set; that is, to evaluate the probability of large discoveries in a region (the large discoveries are of major interest). It is important to know, in this case, the volume of reserves of yet undiscovered large fields. It is impossible to determine this by analyzing the distributions because they provide the forecast interval, which is too wide for the large fields. Thus, the forecast loses significance.

When forecasting is based on the field distribution function, it is important to keep in mind that we are using a distribution with no mathematical expectation and/or variance. In such a case, the reliability of a single-value forecast of a random variable is low. The estimates of averages are unreliable. The only reliable method is a description through a distribution described by such probabilistic parameters as mode, median, and characteristic quantiles, rather than with empirical values of mathematical expectation, variance, and so forth. For a random variable, it is necessary to assign a reliable interval within which its true value falls within the assigned degree of accuracy. Non-uniform intervals, which double in size with increasing field reserves, are used in practice. The intervals are relatively narrow for small fields and very wide for large fields. Under these conditions, the forecast of the number of fields falling into an interval is meaningful only for small fields—their number is high and the predicted range of volume of reserves is small.

Notes

1. Generally, the term "control" is used to describe a policy (or strategy) leading to a desired goal. The term is used throughout this book because it is assumed that a policy (or strategy) is well-structured and can be described in mathematical terms.
2. In the text, "field size" refers to the reserve size.
3. The same law described by these expressions of frequency and rank formats is referred to by different names (Zipf, Lotka, Estu, Mandelbrot, Pareto, etc.), depending on where it is used.

CHAPTER 7

Sequence, Structure, and Rate of Oil and Gas Field Discoveries

In the previous chapters we studied relations between the field sizes and the hydrocarbon reserve amounts within the fields belonging to various classes. This chapter deals directly with the study of field discovery evolution. Field discovery evolution was previously studied mainly for forecasting the size and sequence of discoveries. The size of discoveries affects the exploration efficiency and capital investments needed for oil and gas production. A more objective and more balanced planning of regional studies, seismic exploration, and exploratory drilling, as well as better long-term forecasts of their evolution, result from information about (1) the anticipated sequence of discoveries, (2) the number of large, medium, and small discoveries at the subsequent exploration stage, and (3) the reserves that each of these classes provides. The study of field discovery evolution occupies an important place in the preparation of long-term exploration forecasts and in perfecting the basis of quantitative oil and gas potential evaluation. Field discovery evolution also serves as the basis for the forecast of future discoveries.

Studies indicate that there is a certain order in oil and gas field discoveries. The sequence of discoveries follows its own internal patterns. The most general pattern is that large fields in any region are discovered mainly at the early stage of exploration. At a later stage, mostly medium-size fields are discovered, whereas at the final stages small discoveries prevail. This pattern of discoveries occurs practically without exception. All available publications emphasize the universal nature of this pattern.

Status of the Problem

There are two substantially different approaches to the study of this problem: probabilistic and deterministic. The first approach uses probability theory language for describing patterns pertaining to the oil and gas field discovery process. The second approach observes in time (or in the extent of exploratory drilling) the main trends in the sizes of discoveries. There are, however, only a few publications in which this problem is discussed.

Probabilistic studies of field discovery evolution are usually performed through stochastic modeling of the exploration process. Two classes of models (analytical and algorithmic) are applied for this purpose [4]. When constructing a model based on an algorithmic concept, the input and output data are not connected by obvious functional relations. If a model has several parameters that need to be identified, this may render it too complex to be solved. In such a case, the study must be conducted based on computer simulation from the very beginning. Analytical models provide an opportunity for more efficient parameter identification and interpretation of modeling.

Stochastic models of the exploration process constructed on an analytic basis use a number of probabilistic characteristics as input data known prior to the beginning of exploration. From these data, the forecast characteristics of the exploration results are obtained. The starting point of these studies is the assumption that the field (deposit) reserves are independent random variables with the same distribution density $f(Q)$. In addition, the probability $p(L/Q)$ of autonomous discovery of a target with reserves Q after implementing an exploration effort (extent of exploration drilling, monetary resources) of 0 to L, regardless of whether or not other targets have been discovered, is considered to be given. The events of different discoveries are considered mutually independent.

Based on these initial suppositions it is easy to derive the probabilistic characteristics of the forecast result [4]:

a. Probability distribution density for the reserves of the fields (deposits) discovered when L changed from L_1 to L_2 (in particular, at $L_1 = 0$),

$$\varphi^{L_1,L_2}(Q) = \frac{[p(L_2/Q) - p(L_1/Q)]f(Q)}{c^{L_1} - c^{L_2}}$$

b. Probability distribution density for the reserves of the fields (accumulations) undiscovered after implementing an exploration effort of L,

$$\varphi^L(Q) = \frac{[1 - p(L/Q)]f(Q)}{c^L}$$

where: $c_L = \int_Q [1 - p(L/Q)]f(Q)dQ$

probability of any discovery when exploration effort changed from L_1 to L_2;

$$p^{L_1,L_2} = \int_Q [p(L_2/Q) - p(L_1/Q)]f(Q)dQ = c^{L_1} - c^{L_2}$$

probability of no discovery with any reserves at the expended exploration effort L;

$$c^L = \int_Q [1 - p(L/Q)]f(Q)dQ$$

probability of discovering k fields out of the total number of N fields in the (L_1, L_2) exploration effort interval (considering a discovery as a "success" in N Bernoulli's trials);

$$p_k^{L_1,L_2} = C_N^k (p^{L_1,L_2})^k (1 - p^{L_1,L_2})^{N-k}$$

probability of failing to discover n fields out of the total number of N fields after expending an exploration effort of L

$$q_n^L = C_N^n (c^L)^n (1 - c^L)^{N-n}$$

The most complete model of this type was presented by Beylin [4]. In this publication the total number of fields N prior to the beginning of exploration is treated as a random variable distributed under the negative binomial law. This leads to changes in the equations of probabilities for the discovered number of fields and for the number of fields which remain undiscovered. In addition, Beylin derived the probabilistic characteristics describing the number of discovered and undiscovered fields provided that m fields with reserves $Q_1, Q_2, \ldots, Q_m$ were preliminarily discovered after expending an exploration effort L. If N is the number of fields prior to the exploration and M and N_L are, respectively, the numbers of discovered fields and undiscovered fields after an exploration effort L, then according to Bayes' formula:

$$p/[(N_L = k)/(M = m)] = \frac{p[(M = m)/(N = k + m)]p(N = k + m)}{p(M = m)}$$

The above relations were derived using formal procedures according to well-known laws of probability theory. In terms of content, the emphasis in modeling is not on the derivation of corresponding equations, but on substantiation of the initial distribution function format. In the preceding chapter we discussed in detail the probability distribution density $f(Q)$. It is usually assumed in stochastic models that $f(Q)$ corresponds to Pareto's density or to Weibull's density [4, 1]. Function $p(L/Q) = 1 - (Q_o/Q)^{\beta L} e^{-(\gamma L + \eta L^2)}$ is used as probability of autonomous discovery of a target with reserves Q. Here, L is exploration effort expended in order to discover a field with reserves Q and $p(L/Q) = Qn/A$, with n being the number of exploratory wells drilled in the region [4, 1]. In this case, the discovered field size distribution density within the region, being a unimodal function, changes in time depending on the exploration effort or the areal density of exploratory wells. In the process of exploration, the curve's maximum shifts to the left, toward the smaller fields. In addition, under the same initial conditions as Beylin's model, the reserve size distribution density for the fields that remain undiscovered after exploration effort L has been expended has the same format as the initial one (Pareto's density) except for the different numerical parameters:

$$\varphi^L(Q) = \frac{\alpha_L Q_o^{\alpha_L}}{Q^{\alpha_L + 1}} \qquad \alpha_L = \alpha + \beta L$$

The two reviewed publications are actually the only ones that address analytical models of the exploration process. Shpilman (1983) examined the same problem when he introduced the concept of an "exploration filter" [53]. An explicit expression of this filter is not provided, however, it does describes the probability of discovering an accumulation of size Q at a certain value of parameter. This parameter is tuned to the discovery of accumulations having a certain size. Thus, it may be assumed that the function $\Phi(\gamma)$ characterizes the probability of discovering the accumulations having a size Q at a certain exploration status γ, and is similar to $p(L/Q)$. In the above quoted publication, the problem is constructed differently: it is inverted compared to the problem discussed here. Shpilman analyzes the trend of change for the probability density of discovered fields as exploration evolves and arrives at the initial field (accumulation) distribution function and the corresponding density $f(Q)$.

Algorithmic models of the process of discovering oil and gas fields of various sizes use a somewhat different formulation of the problem. They also analyze events and probabilities associated with field discoveries. The central issue of the models, however, is the probability $p(Q)$ of the next discovered field containing the reserves Q. Kaufman et al.

determined this probability as the ratio of Q to the total reserves of yet undiscovered fields [16]. Belonin and Podolsky assigned $p(Q)$ in a similar manner [5]. The difference is that they use the reserve ratio of yet undiscovered fields (in the reserve class Q) to the total volume of undiscovered reserves. Thus, the issue is the probability of the next discovery of the same class field. Sometimes, the areal extent of the structures instead of the field reserves is analyzed. Ryzhik et al. maintains that the probability of discovering a structure belonging to a certain size class is proportional to the summed area of all undiscovered structures in that size class [46].

It is assumed in all these cases that the order of field discoveries is determined by the selection process, without replacement from a finite field set with an initial distribution density $f(Q)$. Obviously, discovered field reserves are neither independent nor similarly distributed values, as is the case of the initial distribution created by nature. Probabilities of new discoveries depend on the previous ones. In such a case, the field discovery process cannot be represented in analytic format. It may, however, be simulated using information on the number of undiscovered fields and their distribution among the size classes and the average field size in each class. The computer discovery simulation is usually performed using the Monte-Carlo technique.

When analyzing discovery probability, sometimes other probabilities are introduced, which are similar to the probabilities $p(L/Q)$ and $p(Q)$, but not identical to them. For example, Kaufman et al. studied the probability of discovering a field (accumulation) by drilling an exploration well [16]. This probability is assumed to be proportional to the ratio of the volume of undiscovered deposits (accumulations) to the volume of potentially petroliferous sediments. This probability is related to the success rate of oil and gas accumulation discoveries. Here, as well, one has to deal with the dependent events.

There are several types of studies on the succession of field discoveries based on the deterministic approach. The simplest method is to analyze the trend in the change of the average size of discoveries as it relates to the accumulated volume of the appraised reserves of hydrocarbons (or to the exploration ratio, or maturity rate) [34, 31]. Such an analysis demonstrates that the average discovery size increases, reaches a maximum, and then begins to decline. Similarly, the correlation between the discovered field size and the discovered succession was studied [52]. The curve presenting this correlation also has a maximum. After reaching this maximum, the curve rapidly flattens.

Kontorovich et al. performed a more thorough analysis using a wider array of parameters reflecting the trends of interest [25]. This makes the structure of the study more complex. Among such parameters, for instance,

are the distribution parameters of the fields discovered at each stage. This is a deterministic study of probabilistic characteristics. This analysis assumes that the size distribution for the fields discovered over a certain time interval, as well as by a certain time, is governed by the log-normal law. The authors determined the most plausible estimates of mean and variance, as well as the average of normalized reserve logarithms and their variances for the discovered fields. They then examined how these values change depending on the exploration ratio (maturity rate) of the ultimate hydrocarbon potential. These functions, for the fields discovered since the beginning of exploration and through a certain point in time, are described by mutually non-associated regression equations of different types. The mean values behave similarly to what was found in the other publications cited previously; namely the curve of the mean values plotted against the maturity rate has a maximum. The variances behave in a similar manner.

The above analysis also examines the relations between the normalized values of the total number and combined reserves of the discovered fields within individual classes, together with the maturity rate. The normalized value of the total number of fields is represented as a share of the total number of fields in a basin, and the combined reserves of the discovered fields are represented as a share of the ultimate potential. Altogether, seven such classes were identified. The authors did not mention how this classification was made. We can conclude, based on an example quoted in the publication, that the class boundaries were defined by equal intervals on a logarithmic scale. Relationships for each class were studied autonomously, isolated from possible connections between relationships for different classes. These relationships are approximated piecewise for separate intervals of the maturity rate. The third power polynomial was used for the approximation. The number of intervals ranged between two and three. The maturity rate at which all fields within a particular class will be discovered was determined from a linear regression of the maturity rate over the mean reserve logarithm for all fields in a particular class. This particular value was determined based on the accepted Pareto distribution. The correlation coefficient and mean square deviation estimates quoted in the publication may lead one to conclude that the piecewise approximation of the studied relationship is quite satisfactory.

Purpose of the Study

As can be seen from the above review of the status of studies in patterns of oil and gas field discovery, the factors affecting discovery evolution are considered to be either the volume of exploration effort

(mostly drilling volume or the number of wells drilled since the beginning of exploration) or the exploration maturity extent of a region. Generally speaking, these parameters are connected, so there is no difference in principle between them. It is tacitly assumed in this approach that the success and quality of discoveries are determined by the exploration and appraisal volume. The positioning of wells, or of exploratory and appraisal drilling volumes, is not taken into consideration. The fact that exploration is guided by information gathered in the process of exploration is also not taken into account.

In other words, the problem studied is not associated with the strategy. However, as we have already discussed, it is impossible to overestimate the effect of strategy on the sequence of discoveries. The strategy, to a significant degree, predetermines the evolution of the entire exploration process. Since we are reviewing the problem from a strategic viewpoint, the probabilistic approach that treats a discovery as a random phenomenon appears to be unacceptable. When studying the sequence of discoveries as a deterministic phenomenon resulting from a conscious selection of certain exploration plays, the presence of feedback is required, which was already mentioned. Feedback evolves in time, therefore, in order to account for it, the discovery evolution in time needs to be examined. The necessity of accounting for time is also associated with the fact that the planning and control of exploration requires time forecasts (predicting when a certain result will be obtained). These facts force us to study oil and gas discovery trends in time.

The simplest method is to obtain an idea of this trend from the change in the average size of the discovered fields. This type of study was rejected because the technique is only suitable for the simplest conclusions, when a coarse solution not intended for uncovering deep internal associations is satisfactory. We, however, are interested in exactly this. For this reason, the structure of the discoveries, and not just their mean value, is important. It is also important to study how many and what kind of fields are discovered every year and the trends in their number and size. This is what forced us to study the discovery evolution as a problem of describing changes in the structure of discoveries. In other words, it is necessary to analyze the evolution of discoveries within each class,[1] as well as the evolution of interrelations between field discoveries in each class. Only this formulation of the problem accounts for the particulars of the discovery evolution caused by a certain system of exploration. In this manner, one can see not only which fields are discovered earlier and which later, but also which fields are less important and which are discovered at the same time (i.e., how selective the exploration system is).

Thus, the problem may be formulated as the study of intra- and interclass patterns of discoveries and identifying trends of their change in time.

Intra-Class Patterns in the Oil and Gas Field Discovery Evolution

For any class of fields there is a moment in time for the first discovery and a moment in time for the last discovery. Correspondingly, there is a time interval within which all fields of a given class are (or will be) discovered. As practical experience demonstrates, the fields within this time interval are discovered non-uniformly. There are more discoveries in some years and less discoveries in other years. In some cases, no discoveries are made within a given class for a given year, or even a few consecutive years. This is a typical situation with no exceptions.

Discovery patterns appear by way of a certain increase or decrease in the number of discoveries. A convenient way to study them is to register the total (accumulated) number of discoveries at a certain point in time. The result is a curve showing the growth of the number of discoveries in time. The actual curve is a broken line with horizontal plateaus when there are no discoveries in a given class.

Analysis of many different regions has shown that the accumulated discoveries growth in time curve for any field class has an S-type shape. The number of annual discoveries (discovery rate) is first small, then increases, reaching its maximum, after which the discovery rate declines. This pattern is typical for all field classes, at least for those classes where the accumulated discovery curve is already beyond its inflection point. For those classes where the discovery rate has not yet reached its maximum, the curve behavior at the initial stages is similar. For this reason, it is believed that the field number growth type for these classes will run according to the same pattern. This pattern appears to be universal in that no region or field class displays a different pattern.

Therefore, there are at least three stages of discoveries in each class. At the first stage the discoveries are not the main exploration target and are made incidentally, or not on purpose, which results in a low initial discovery rate. Then they become the main exploration target and attract more attention. The current technology and techniques are directed specifically to such targets. As a result, the discovery number increases and reaches its maximum. Later, the interest in these targets declines due to the transition to different types of targets as well as the fact that undiscovered fields of this class are associated with geologic conditions unfavorable (or not typical) for the previous stage, which makes discoveries difficult. As a result, the discovery number declines, as they are less and less the result of purposeful activities.

Inter-Class Patterns in the Discovery Evolution of Oil and Gas Fields

As indicated, the inter-class discovery pattern is one of coordinated, interrelated growth in the total number of discoveries for all field classes. To establish such a pattern, the manner in which a field population is subdivided into classes is of paramount importance. If the classes are selected according to the commonly used[2] ranking (for instance, greater than 30 MMT, 10 to 30 MMT, 5 to 10 MMT, and less than 5 MMT), the following interesting example may be given regarding the discovery number increase in each class and their interrelation.

For many years, large discoveries (greater than 30 MMT) outnumber the discoveries in any other class. After a time, however, the rate of second-class (10 to 30 MMT) and fourth-class (less than 5 MMT) discoveries rapidly increased. As a result, their number became greater than that of the large discoveries. Over the entire studied interval, the number of third-class discoveries (5 to 10 MMT) was smaller than the number of discoveries in any other class. At the same time, in this particular example, the first discovery in the third class was made prior to the first discoveries of the second and fourth classes (i.e., during the time when only large discoveries were being made).

This example does not indicate a clear order in the increased number of discoveries in each class or of a definite connection between the discovery number in different classes. A forward increase of the first-class discoveries at the first stage is understandable. At the same time, similar growth in the discovery number of the second and fourth classes and the lagging of the third-class discoveries is perplexing, and is not conducive to defining any pattern of inter-class relations. However, this is not surprising. This example does not contain any useful information concerning the exploration system and the associated discovery system. The reason for this is that in any oil and gas region the number of fields in the indicated classes is unequal. This makes it impossible to understand the cause of the discoveries: it may be a certain exploration system or it may be that there are simply more fields of one class in nature than another. Naturally, the fields prevailing in nature have a better chance of being discovered.

This example demonstrates the importance of field class selection and class boundaries for studies of inter-class relations.

It is clear from the above that the fields must be subdivided into equiprobable classes where, theoretically, each class includes an equal number of fields. In such a case, a forward growth in the discoveries of one class compared to the others would be associated with the exploration

system—or the exploration strategy. An examination of the inter-class relationships would also acquire a clear meaning. The analysis would become more convenient because the growth curves must eventually reach the same value equal to the total initial number of fields within each class of an oil and gas basin. How the curves tend to this number is revealing.

Based on the findings of the previous chapter, the boundaries between the equiprobable classes were determined from the premise that the entire population of fields (discovered and undiscovered) in the oil and gas basin is governed by Pareto's distribution with the characteristic parameter $a = 1$ (a special case of Pareto's distribution that is much discussed in this book). It is only natural to assume that the subdivision into classes under this condition be performed in more than one manner. Everything depends on the width of the interval of each class and, correspondingly, the number of fields belonging to each class. This is a typical problem encountered when the domain of a studied value is being subdivided into a finite number of multitudes (groups, intervals) for the statistical examination of different hypotheses. The results of a subdivision that is too coarse or, vice versa, too fine, are well documented.

In this case, the following should be kept in mind. Let us assume we identify classes allowing for a more-or-less detailed study of large fields (i.e., large fields should be represented by several classes). The distinctions between the classes of small fields would become insignificant (they would measure in the hundredths or thousandths of a unit), and the number of classes would become too great. The opposite is also true: if we subdivide small fields into convenient classes, it would lead to a coarsening of classes among the large fields, which would naturally reflect the quality of their analysis. Some important details might be lost. Unfortunately, it is impossible to combine both goals. We must look for a golden mean.

Whichever system of class separation is selected, classes sooner or later appear in the small field domain where the interval they occupy will be below the accuracy used for the determination of reserves (i.e., 0.1 MMT). This leads to rounding, and, as a result, the entire class may be represented by fields of the same size—for instance, an even 1.0 MMT—and it does not end here. Within an even smaller domain the intervals will be so small that a fixed value, such as 0.2 MMT, will represent several intervals (i.e., the fields of this size will occupy several classes). Thus, under our chosen system of field subdivision into classes, fields with reserves of 0.2 MMT (the fields of this size are also entered into balance books) must be attributed to classes 401 through 667 (i.e., they will occupy 267 classes). When analyzing the actual data we included a corresponding share of the discovered fields in each class. This is why there are growth curves with discovery numbers of less than one.

We will now begin analyzing the inter-class patterns. The sum of in-place reserves for all accumulations of a field is used as the value of reserves for this particular field. The year when the first accumulation is discovered is considered the discovery year. The analysis includes comparison of the discovery growth curves within a certain region for each consecutive class.

A typical picture of the inter-class relations is shown in Figure 7-1, displaying smoothed growth curves for the first six classes. As previously, an increase in the class sequential number indicates a decrease in the reserves of the fields within that class. As Figure 7-1 shows, there is a clear tendency toward order in the position of the curves. As the classes become coarse, this order becomes obvious. The essence of this order is that the growth curves for field classes with smaller sequential numbers are located higher and reach their limit value earlier than the growth curves for field classes with greater sequential numbers. Each subsequent curve is located lower than the preceding one. Each preceding curve rises to its maximum value more steeply than the subsequent one. This means that the larger the fields within a given class, the earlier they begin to be discovered, the earlier the process of their discovery ends, the higher their accumulation rate, and the smaller the time interval between the first and last discovery. Also, the larger the fields within a class, the greater the number of accumulated fields at any given point in time. The only

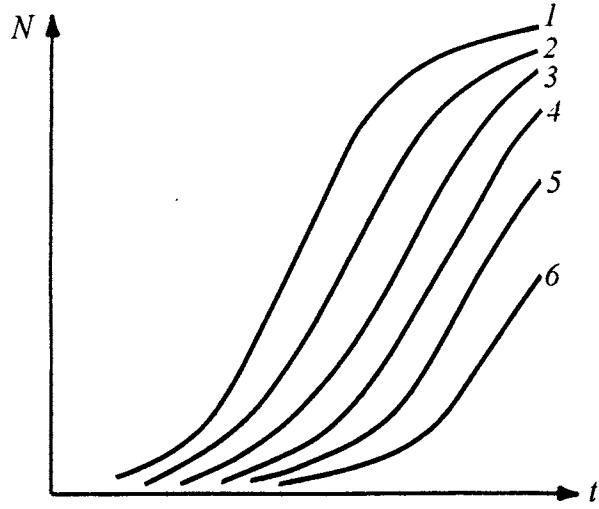


Figure 7-1. Smoothed curves of the increase in time t of the number of discoveries N for different field classes. 1 through 6: field classes.

exception is at the time when all fields in the compared classes have been discovered. At this time their number becomes equal. Of course, these patterns appear in an averaged form as a general trend.

Analyzed curves are a convenient format for presenting the discovery results. They provide information on more than just inter-class patterns. They also show the entire structure of discoveries: the number and quality of annual discoveries, the growth of the discovery number within each class and their total for all classes, which classes are fully discovered, and which are close to that status. They also allow a scientist to make an approximate evaluation of the discovery structure for future years: how many and which discoveries will be made in a particular year, when the discoveries within a class of interest will be exhausted, how many fields will be discovered annually, and so forth. However, in order to remove individual features from each curve and to leave only those features defined by the general discovery system, even this presentation format should be improved. A format is needed that would reflect the total discovery system and the totality of class relationships. For this purpose, we will examine the analyzed patterns from a somewhat different viewpoint.

As indicated earlier, based on studies conducted by the other scientists, the first discovered fields are not the largest ones. The time of the first discoveries among the large fields comes a little later. In other words, it is necessary to study the general trend in sequential discoveries of the first, second, and subsequent fields within each consecutive class.

The smoothed curves of this relation, for any number of discoveries, are shown in Figure 7-2. These curves may be considered contours of equal discovery numbers. Contours are horizontal cross-sections of some surface. It may be concuded that the total discovery system can be described by a surface in coordinates (k, t, N). This surface has S-shaped curves for its vertical cross-sections parallel to the t-N coordinate surface, and it is limited at the top by a horizontal plane at the level equal to the total (initial) number of fields within each class.

Other properties of the surface can be determined from an examination of the contour behavior (see Figure 7-2). The slope of the contours increases as their value increases. They have the appearance of a fan. The distance between the contours along the t axis increases as the class number grows. This indicates that the rate of increase in the number of discoveries declines as the field size decreases. Distinctions between the classes decline in the same fashion. Whereas these distinctions are significant for senior classes (contours are steep), they gradually become less substantial for subsequent classes (the contours become less steep and at a certain stage they are almost parallel to the k axis indicating that temporal distinctions between classes have almost disappeared). This means that the selective, purposeful nature of the process weakens for

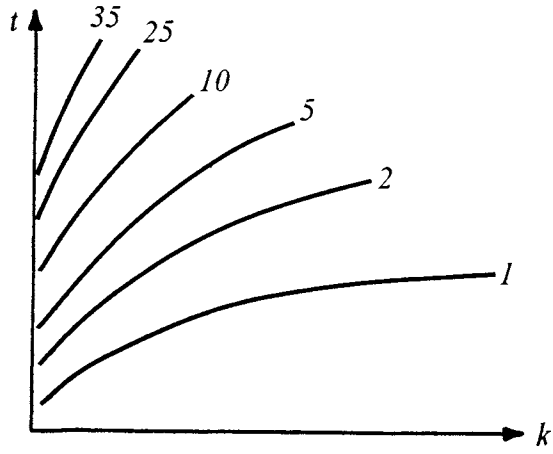


Figure 7-2. The relationship between the time t and the class number k for the different number (total) of discoveries (contours represent an equal number of discoveries).

small fields, and distinctions between the small fields do not affect the process during exploration.

Vertical cross-sections of the surface (corresponding to each class) display the accumulation character in time of the number of discoveries in each class under conditions where individual features of each growth curve are eliminated and only major features, associated with the discovery system, remain. Distinctions between the curves are associated with the differences in the growth rate. The smaller the class number, the higher the number of discovery accumulation rate. This is reflected in regular changes of the corresponding differential curves (i.e., curves showing the number of annual discoveries within each class). The maximum number of discoveries decreases in value from senior to junior classes and moves right on the t axis. The flat segment of the curve in the maximum area becomes wider, that is, the "top" of the curve (maximum) becomes lower and wider.

These are inter-class patterns in the oil and gas field discovery evolution.

Modeling of the Oil and Gas Field Discovery Evolution

The above patterns represent a qualitative level of study. All constructions were estimated, including the smoothing procedure. This was sufficient for the discovery of general patterns. The construction of an

evolutionary model would be impossible without this stage. Now that the general patterns are known, we can begin the modeling process for different size field discoveries. It is now easy to formulate the purpose of the modeling, which is to find the analytical type of a surface reflecting the above patterns in the interconnected growth of the total number of discoveries within different classes. One precondition is that the subdivision into classes be based on the principle of an equal field number within each class (equiprobable classes). We will assume that observation deviations from the surface are random. Correspondingly, deviations of actual growth curves from the surface cross-sections are also random (i.e., each growth curve is just one possible implementation of the regular growth process).

The investigated function $N = f(k,t)$ possesses a number of properties that were mentioned above. These properties, naturally, should determine the format of this function. The following should be added. At a given k we have a cross-section of the surface (i.e., a function of t: $N_k = f(t)$). It is reasonable to assume that for each region the format of the N_k function does not depend on k; with an increase in k only parameters of the N_k function change. This means that the characteristic points of the S-shaped N_k function (position of the inflection point, maximum rate, the degree of asymmetry, etc.) should depend on the parameters and fluctuate with the change in parameters rather than remaining rigidly fixed. Only a limited class of the S-shaped functions answers to this condition. It is also reasonable to suggest that this function should have an asymptote. The asymptote should depend on a single parameter of the model. The value of this parameter should be constant for all N_k functions (at any $k = 1, 2, 3 \ldots$) within a given region. This satisfies the condition of a theoretically equal initial field number for all field classes. It is possible to proceed from the $N_k(t)$ function to the surface model, assuming that with the change in k the $N_k(t)$ function parameters change. Let us assume these parameters are $a, b, c, t_o \ldots$ (parameter a is associated with the asymptote; for instance, it is a constant value for all $N_k(t)$). Then, based on the aforementioned, all parameters are functions of k, or:

$$b = b(k),\ c = c(k),\ t_o = t_o(k) \ldots$$

Thus, the surface model describing the entire discovery structure has the following format:

$$N(k,t) = N_k[t,\ a,\ b(k),\ c(k),\ t_o(k)]$$

Specific modeling of the discovery system in each region requires determining the format of the $N_k(t)$ function, and functions $b = (k)$,

$c = \psi(k)$, $t_o = f(k)$ . . . describing changes in the $N_k(t)$ curve parameters with the growth of k.

Patterns of the Oil and Gas Field Discovery Evolution in Various Regions

In order to determine how the exploration strategy affected patterns in this evolution, the field discovery evolution study was performed for two regions where exploration strategies were different. As you may recall, this comparative analysis is conducted throughout this book.

Region A. This region is mostly coincident with region G (see Chapter 3). It is practically the same except for slightly wider boundaries. The exploration strategy applied in the region was described in Chapter 3. This strategy has certain features common for all regions denoted "region A" throughout this book.

A selection of the $N_k(t)$ function in the region is determined by a number of previously mentioned conditions, as well as certain specific properties of the function's behavior. One specific property that has not yet been discussed is as follows. When analyzing patterns of the reserve accumulation we previously used an important parameter characterizing the behavior of the growth functions that we called the specific, or relative, rate. All derived models included a monotonous decline of this parameter in time. As our analysis showed, in this particular case the relative, or specific, growth rate of the number of discoveries $(dN_k/dt)/(1/N_k)$ changes according to a different law. It first increases, reaches a maximum, and then begins to decline. These conditions and specifics are satisfied by a function $N_k(t)$ of the following format:

$$N_K(t) = \exp\left\{\frac{a}{[1 + (a^{1/c} - 1)e^{-b(t-t_o)}]^{c-1}}\right\} \tag{7-1}$$

Parameter a is connected with the asymptote A of the N_k curve: $A = e^{a-1}$. At $t = t_o$, $N_k = 1$ (i.e., the t_o parameter records the time of the "first discovery"). To better understand the meaning of the other parameters, we introduce a variable:

$$z = \ln N_k + 1.$$

Function $z(t)$ is similar to the function $R(t)$ in Equation 2-34. Behavior of $R(t)$, as well as its derivative $dR/dt = r(t)$, was discussed in Chapter 2.

The derivative of z is nothing but the specific growth rate (speed) of the number of fields discovered at the moment in time t:

$$z' = \frac{dz}{dt} = \frac{dN_k}{dt}\frac{1}{N_k}$$

Remember, parameters of the $R(t)$ function determined the maximum value of the $r(t)$ function and the time moment when this maximum is reached. Similar to Equations 2-38 and 2-39:

$$z'_{\max} = ab\left(\frac{c}{c+1}\right)^{c+1}$$

$$t_{\max} = \frac{1}{b}[\ln c + \ln(a^{1/c} - 1)] + t_o$$

where $z'_{\max}$ and $t_{\max}$ are the maximum z' value and the time it has been reached.

Therefore, parameters a, b, c, and t_o define characteristic points of the $\dot{z}(t)$ function (which describes the behavior in time of the specific rate) or characteristic points of the $z = \ln N_k + 1$ function (which describes the growth of not the number of discoveries, but their logarithm). Similar to Equation 2-40, the $z(t)$ value at the inflection point $t_{\max}$ is:

$$z(t_{\max}) = a\left(\frac{c}{c+1}\right)^{c}$$

Thus, we see that specific features of the $N_k(t)$ function behavior are determined by its parameters. When the parameters change, the function value at the inflection point, the time to reach a maximum rate, and the maximum rate value change. The same is true for the specific rate. The model is sufficiently flexible and none of its characteristic points are fixed (i.e., independent of the parameters).

Studies, for the purpose of selecting the $N_k(t)$ function format, were conducted with individual discovery growth curves for different classes. The classes selected were those where the discoveries were practically exhausted or close to this state. The accumulated discovery growth curves for these classes reached their inflection points and almost reached their limit value (asymptote), which was an indication of their representativeness. The maximum analyzed time interval, beginning with the first

discovery, was 54 years for some classes. The maximum discovery number in one of the classes was 35. Different models were evaluated to determine how well they fit the actual data.

Tested models included different growth models examined in this book, including the above model 7-1. This last model provided the best result. For instance, correlation coefficients for classes 1, 3, and 5 between the actual number of the accumulated discoveries and the "theoretical" values from model 7-1 were, respectively, 0.9963, 0.9973, and 0.9971. Based on this, a function of the 7-1 format was selected as the $N_k(t)$ function. Thus, the model describing the entire discovery structure in the analyzed region has the following format:

$$N(k,t) = \exp\left\{\frac{a}{[1 + (a^{1/c(k)} - 1)e^{-b(k)(t-t_o(k))}]^{c(k)}} - 1\right\} \qquad (7\text{-}2)$$

Naturally, the format of the $b = (k)$, $c = \psi(k)$, and $t_o = f(k)$ functions describing the behavior of the parameters with the growth of k is now important.

The search for the formats of functions describing connections between parameters and the class number k was conducted using the actual data. At the first stage, the $N_k(t)$ function parameters for each class $k = 1, 2, 3 \ldots$ were determined with no limitations of their value. This included the a parameter which was not assigned a constant value. As a result, the a_k, b_k, c_k, and t_{ok} parameters of the $N_k(t)$ function were determined ($k = 1, 2, 3 \ldots$). If we take the 0.1 MMT field as the smallest field (fields of this size are included in the balance books), we will have over 1,000 classes. However, for classes with great k number (and that is exactly where a field under 1 MMT would fit) no pattern is recorded for field discoveries, and they are impossible to describe through model 7-1. Discoveries in these classes are random. Purposeful, selective exploration has not yet reached fields of this size. Thus, a dualism exists in the discovery system: fields of a certain size are discovered under one law and fields of a different size under another. The separation occurs above the fields with reserves less than 1 MMT. Therefore, the discovery system analyzed here encompasses fields 1 MMT and larger.

As we indicated earlier, at the reserve evaluation accuracy of 0.1 MMT, the fields with a set reserve number (for instance, 1.9 MMT) occupy several classes (in this example, three). In a case like this, the observed number of discoveries are equally divided between such classes. In this particular example each class was ascribed one-third the number of 1.9 MMT discoveries. As a result, in this particular region we identified 105 classes. Interestingly enough, not a single one of the first 105 classes was missing (i.e., there were discoveries in each of these classes). This

was the database for our analysis. The observation results of these 105 classes were used at the second step of the final stages of analysis. When searching functions $b = (k)$, $c = \psi(k)$, and $t_o = f(k)$ only those classes were analyzed where there was a significant number of discoveries over a substantial, representative time interval.

It is easier to begin the analysis of the first stage with the a and t_o parameters because their meaning is clear (their values correspond to the asymptote and the timing of the "first discovery"). No pattern was observed in a_k values with an increase in k. Their behavior was systemless and apparently random. Therefore, we must recognize that, indeed, a_k = const (k = 1, 2, 3. . .). The general trend in t_{ok} is clear: when changing from one class to the next, it increases. It is possible that the few highest classes are an exception to this rule. It appears that the t_{ok} value may decrease up to k = 3 and begin to grow monotonously afterwards. Therefore, there is a minimum but it is "washed out." This suggestion corresponds to the observation that the first discoveries are usually not the largest ones. The large discoveries are slightly delayed and occur after some smaller discoveries. On the whole, nothing unexpected occurs in the behavior of t_{ok} with an increase in k. This behavior corresponds with the general ideas presented when discussing the "first discovery."

Taking into account the possibility of a minimum in the $t_o = f(k)$ function, the following first step was taken as an approximating function:

$$t_o(k) = \alpha_t - \beta_t k^{\gamma_t} e^{-\rho_t k} \qquad (\alpha_t > 0,\ \beta_t > 0,\ \gamma_t > 0,\ \rho_t > 0) \tag{7-3}$$

Based on the aforementioned, we assume that this function adequately describes the relation under investigation.

Function 7-3 has a minimum at the point $k_{\min} = \gamma_t/\rho_t$. Its value at this point (the minimum value) is equal to:

$$t_o(k_{\min}) = \alpha_t - \beta_t \left(\frac{\gamma_t}{\rho_t}\right)^{\gamma_t} e^{-\gamma_t}$$

In addition, this function has two symmetrical inflection points with respect to the minimum point:

$$k_{infl} = \frac{\gamma_t}{\rho_t}(1 \pm \gamma_t^{-1/2}) = k_{\min} \pm \frac{\gamma_t^{1/2}}{\rho_t}$$

The value of function at the inflection points is:

$$t_o(k_{infl}) = \alpha_t - \beta_t \left[\frac{\gamma_t}{\rho_t} (1 \pm \gamma_t^{-1/2}) \right]^{\gamma_t} e^{-\gamma_t (1 \pm \gamma_t^{-1/2})}$$

This function has an asymptote α_t.

The relation between the obtained b_k and the class number k is that at the beginning b_k values increase with increasing k; then, upon reaching their maximum, begin to decline. They do not asymptotically tend to zero, but to some positive value. The presence of a maximum (i.e., the absence of a monotonous decline in b_k with increasing k) is another indication that the discovery process is not necessarily moving in the direction from senior classes to junior ones. The fields in the senior classes are not associated with the maximum. The maximum covers the classes directly following the most senior classes. Accordingly, an approximating function similar to 7-3 is used that only differs from it in that instead of the difference, the sum of the two components is evaluated:

$$b(k) = \alpha_b + \beta_b k^{\gamma_b} e^{-\rho_b k} \qquad (\alpha_b > 0,\ \beta_b > 0,\ \gamma_b > 0,\ \rho_b > 0) \qquad (7\text{-}4)$$

Function 7-4 also has two inflection points: maximum and an asymptote, which are determined in a similar manner.

The c_k behavior as k grows displays the same features as those typical for t_{ok} but with a clearer minimum. The c_k values decrease with increasing k, reach their minimum, and then begin to increase tending to some limit. This pattern should be described by the same function as the t_{ok} behavior. Therefore,

$$c(k) = \alpha_c - \beta_c k^{\gamma_c} e^{-\rho_c k} \qquad (\alpha_c > 0,\ \beta_c > 0,\ \gamma_c > 0,\ \rho_c > 0) \qquad (7\text{-}5)$$

We can conclude that the discovery evolution in this region is represented by a function $N(k,t)$ of the format 7-2, where the functions $t_o(k)$, $b(k)$, and $c(k)$ are represented by Equations 7-3, 7-4, and 7-5. If we assume that the β parameter can have different signs, these equations can be represented in the same format:

$$m(k) = \alpha + \beta k^{\gamma} e^{-\rho k} \qquad (7\text{-}6)$$

Here, $m(k)$ may represent $t_o(k)$, $b(k)$, or $c(k)$.

Thus, the model of the surface is defined by the thirteen parameters. Based on Equation 7-6, one can assume later that eleven of them (including

the parameter β_b) are positive and two of them (β_t and β_c) are negative. These are the modeling results of the first stage.

The second step included the parameter evaluation for the obtained model and testing the model fit to the observations. Observations of the build-up in time of the discovered fields in all 105 classes, during the 54-year period beginning with the first discovery in the region, were used as the input data. The fields taken into account included those of 1 MMT and larger. Model adequacy was measured by the variance σ^2 and correlation coefficient r. The evaluation of all 13 parameters was conducted using the technique described in Chapter 2. The results are listed in Table 7-1.

Note, first, the high value of the correlation coefficient between the observations and the data calculated from 7-2 with the consideration of 7-3, 7-4, and 7-5. This indicates that the analyzed model fits the observations well; therefore, the derived functions $t_o(k)$, $b(k)$, and $c(k)$ adequately describe the nature of the parameter change for sections of the 7-1 surface when passing from one class to the next.

Reviewing the estimates of the surface equation parameters, one can observe that the $b(k)$ curve has a maximum in class 3, the $c(k)$ curve has a minimum in class 4, and the $t_o(k)$ curve has a weak minimum in class 2: the $t_o(2)$ value only slightly differs from the adjacent $t_o(1)$ and $t_o(3)$ values (Figure 7-3). The curves under review tend to their asymptotes at different

Table 7-1
Adequacy Evaluation of the Discovery Evolution Model in Regions A and B to Observations

	Region and the Number of Observations			Region and the Number of Observations	
Variables	**A (1,378)**	**B (620)**	**Variables**	**A (1,378)**	**B (620)**
a	4.55	2.8	ρ_c	0.0465	4.57
α_b	0.057	0	α_t	60.1	75.2
β_b	0.204	0.51	β_t	–47	–74.2
γ_b	0.98	–0.18	γ_t	0.038	–0.2
ρ_b	0.395	0	ρ_t	0.0189	0
α_c	10.55	0.058	σ^2	3.6	0.917
β_c	–9.41	0.048	r	0.9712	0.945
γ_c	0.192	12.5			

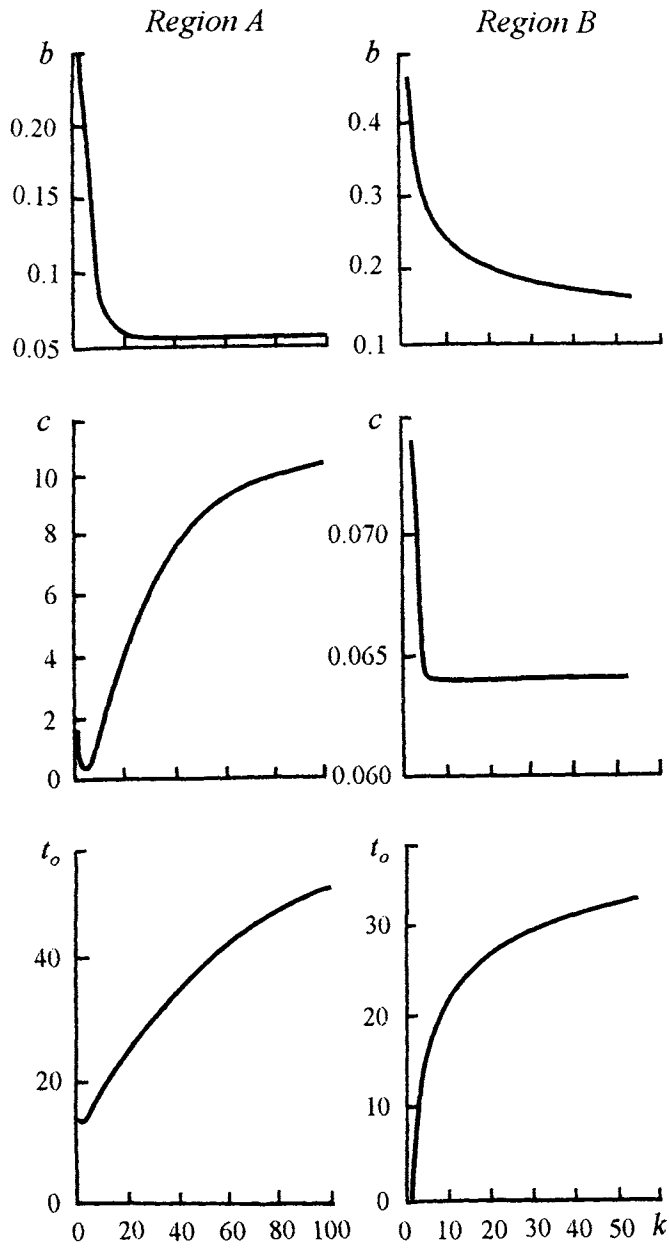


Figure 7-3. The $b(k)$, $c(k)$, and $t_o(k)$ curves in regions A and B.

rates. The $b(k)$ curves have the highest rate and the $c(k)$ curve have the lowest rate. It is important that the $b(k)$ asymptote is not equal to zero (Figure 7-3).

Region B. This is the same "region B" referred to throughout this book. The model selection and its evaluation for this region were based on the discovery system beginning with the first discovered field and continuing through the next 60 years. The largest number of fields in one class is 20. The discovery system consists of fields with individual reserves of 1.9 MMT and greater. This limitation has the same character as region A: the discovery evolution of the smaller fields does not display any pattern and the discoveries appear to be random. Thus, compared to region A, the boundary between the fields discovered under different laws is offset toward larger fields. The fields in this region fall into 53 classes. Some of these classes are empty (no discoveries yet). There are discoveries in the 20 senior classes, then in classes 26, 27, 31, 32, 43, 44, and 49 through 53. Altogether, data for 31 classes were analyzed.

Our discovery evolution analysis in various classes indicates that it is best described by model 7-1. Therefore, the $N_k(t)$ in this region has the same format as in region A. As for the other components of the discovery evolution model $N(k,t)$ (namely, functions $b(k)$, $c(k)$, and $t_o(k)$), attention should be paid to the following in Equation 7-6. As a special case, this equation represents many functions. For instance, at $\alpha = 0$ and $\rho = 0$, Equation 7-6 converts to an exponential function. Depending on the sign of γ ($\gamma > 0$ or $\gamma < 0$), this function is either increasing or decreasing. At $\beta < 0$ and $\gamma = 0$, Equation 7-6 converts to the modified exponential function (base e). Using an analogy, the other special case of Equation 7-6, at $\beta < 0$, $\gamma < 0$, and $\rho = 0$, is a modified exponential function (not the base e function). These special cases allow us to approach the determination of a format for components of the discovery evolution model in the region in an empirical manner. This approach consists of a model parameter estimate with components as described in Equation 7-6. The values obtained for the parameters characterized the format of the $b(k)$, $c(k)$, and $t_o(k)$ functions. The results of the parameter estimates and the model examination are listed in Table 7-1. These estimates present the complete model of the discovery evolution in the region (Table 7-2). A high correlation coefficient indicates satisfactory agreement between the model and the observations. The adequacy here, however, is somewhat lower than in region A where the correlation coefficient is higher. An irregular discovery evolution is also indicated by the fact that the discoveries are not made in all available 53 classes.

Comparing regions A and B, one can see that the differences in the exploration strategy are reflected in clear differences in the discovery

Table 7-2
Comparison of Models (Surface Equations) of the Oil and Gas Field Discovery Structure in Regions A and B

Model Components	Format of Components in Regions	
	A	B
$N_k(t)$	$exp\left\{\frac{a}{[1+(a^{1/c}-1)e^{-b(t-t_o)}]^c}-1\right\}$	$exp\left\{\frac{a}{[1+(a^{1/c}-1)e^{-b(t-t_o)}]^c}-1\right\}$
$b(k)$	$\alpha_b+\beta_b k^{\gamma_b}e^{-\rho_b k}$	$\beta_b k^{-\gamma_b}$
$c(k)$	$\alpha_c-\beta_c k^{\gamma_c}e^{-\rho_c k}$	$\alpha_c+\beta_c k^{\gamma_c}e^{-\rho_c k}$
$t_o(k)$	$\alpha_t-\beta_t k^{\gamma_t}e^{-\rho_t k}$	$\alpha_t-\beta_t k^{-\gamma_t}$

evolution patterns. These differences result in the evolution patterns being described by different models. The models differ in the format of the $b(k)$, $c(k)$, and $t_o(k)$ functions (see Table 7-2) but not in the format of the discovery surface cross-section (i.e., $N_k(t)$ function) format. Differences in the format of the $b(k)$, $c(k)$, and $t_o(k)$ functions results in differences of the respective curves (Figure 7-3). It is noteworthy that the $b(k)$ curve in this region does not have a maximum and smoothly approaches its asymptote, which is equal to zero in this case. The $c(k)$ curve displays a maximum instead of a minimum, after which it rapidly approaches an asymptote (not equal to zero). It is mostly similar to the $b(k)$ curve of region A. The $t_o(k)$ in region B rapidly ascends in the area of the senior classes, after which the growth decreases and proceeds without any drastic changes. This indicates that after the discovery of fields in the senior classes, a significant time interval elapses (drastic growth of the curve) before the first regular discoveries begin in the junior classes. Afterwards, the difference in the first regular discoveries among the subsequent classes gradually decreases. With the further growth of k, the regular difference in the time of the first field discovery in the preceding and subsequent classes gradually decreases. The asymptote is an indication that in the fields within the junior classes this difference becomes almost imperceptible. The "first fields" in these classes are discovered (i.e., the accumulated number reaches 1) almost simultaneously. Any orderliness in the discoveries disappears and random factors start to play a significant role.

Notes

1. The size classes mentioned here are identified based on certain rules that will be discussed later.
2. In the former USSR and, currently, in Russia.

PART III

Strategy of Control Over Oil and Gas Exploration

Parts I and II were devoted to the internal laws of the oil and gas reserve discovery and appraisal process. So far, we have observed this process from the sidelines. In Part III, we will be addressing active intervention in the process—its control. We will no longer be observers, but, rather, participants of the process.

Control is a purposeful (goal-oriented) change of a process. This is why control presumes a goal which is the first component of a control model. If there is no goal, it is meaningless to discuss control. Goal is a concept of the motivations behind controlling any process. Formation of a goal is an informal procedure; goal is an external factor relative to the control system.

The second component of a control model is a formalized description of the means to achieve the goal or of the strategies. They are referred to as "control laws" in control theory. Control laws in effect are principles for selecting the specific controlling actions, out of the many available, that provide for the achievement of the control goal. A function representing a formalized description of the control law may be a function of time, of phase coordinates, of external action, or it may have a more general format. This is an "open function" at our disposal, which we may select at our discretion. Control laws always define the value of a control action as a function of the system's position relative to the control goal. The goal may be achieved in different ways. This is why the model should include a description of the many available options or control laws.

The concept of information is closely associated with the concept of control. Any selection technique of the control option is based on the information regarding the correspondence of the control goal and state of the system.

The fact that the goal may be achieved in numerous ways leads to the problem of selecting a certain control out of the many available. Naturally, it should be the best option. A mathematical definition of the term "the best" requires an evaluation of the control quality based

on some additional criterion. The criterion of control quality is often attributed to the control goal as a goal function. In this manner, the goal may be achieved under certain conditions and with certain restrictions.

Thus, a mathematical model of the controlled system includes a model of the processes taking place in the system (i.e., the processes used to generate and implement the control actions) and a model of the rational selection of the control system. A mathematical model of the rational selection of the control system option includes (1) mathematical descriptions of the control goal, (2) information concerning the operating conditions of the system, (3) the criterion of control quality (or the evaluation criterion of a fit between a given control system option and the goal of control using the current extent of information about the system's operating conditions), and, (4) a description of possible control options and restrictions.

The model of rational selection of the control system represents a class of decision-making mathematical models. Its investigation results in creating selection rules for option screening and choosing the rational option. A decision-making model may be defined as the mathematical procedure for comparing options out of a chosen set of possible options. A general outline of the decision-making mathematical model may be described as follows. Provided a given extent of information, each control law (or a decision option) is in correspondence with a certain state of the controlled process. The presence of the goal and of the quality criterion technically means that the states of the process corresponding to different control laws and to different extents of information are equivalent (certain comparative relations are established between them). These relations provide an opportunity to eliminate all options that failed comparison. Therefore, a model is a formalization of the fact that the control goal establishes the comparative relations between the states of the process. In other words, it provides the rule for screening the decision options based on formal parameters.

Thus, the process status is uniquely determined by the selection of a control law, which is possible under conditions where information is complete. The rule of option selection schematically represents the criterion for evaluation of the fit between the control option and the control goal. If information is incomplete, an uncertainty appears which does not allow for comparison of the options. In this case, the selection of a control option may be given as a set of possible process states that correspond to various extents of information.

Optimization models are the most frequently used in solving control problems. When the decisions are made in an environment of complete information, these models can be presented as follows. Let us assume that E is an element in the set of decision options (in our case it is a particular control law), the information level is a and x is the state of the process

that depends on the control and the information level (i.e., $x = \mu(E,a)$). Then, the comparative relations established over the set of possible states x are such that the scalar function or functional $f(x)$ will be a quality criterion, or a criterion of fit between the chosen option and the control goal. The greater the value of $f(x)$, the better the fit between the respective system option and the control goal. In other words, the selection problem discussed here is formulated as an optimization problem: out of a set of E options, several (or one) must be found for which the condition $f(x) \rightarrow \max$ is fulfilled at $x = \mu(E,a)$. Therefore, the control goal is formulated in the optimization language, in terms of the maximization (minimization) of some function or functional. This is the situation with decision-making in an environment of complete information.

When information is incomplete, the $f(x)$ (goal function) value is not precisely assigned; instead, it depends on the uncertainty of the level a. Solving the problem $f(x,a) \rightarrow \max$, one can determine an optimum control only as a function of a:

$$E = E(a)$$

If there is no information concerning the uncertainty level a, the optimization result of $E(a)$ and $f(x,a)$ will be arbitrary. This means that the optimum option can only be found according to the information level. If the choice is multiple, then the average selection result makes sense. Averaging would be naturally performed over the set of a values. If there is only a single choice (which is the most typical situation in petroleum exploration control), averaging does not make sense. In this situation the uncertainty may be handled using a stage-by-stage optimization. Its essence is that control of a process, especially a long-term process, is subdivided into several stages. The stages differ from one another in that new information is obtained at each stage. Strategy selection at each stage is performed based on available information that is considered certain.

Thus, there are two sides to the decision-making model. One side is a formal procedure of comparing and screening-out the decision options. The other side is substance-oriented. Its formalization is very difficult. It deals with goal formation and the use of available information for the achievement of goal.

Part III of this book addresses modeling of the exploration process, optimization, and forecast of its further evolution.

CHAPTER 8

Model of Oil and Gas Exploration Control

This chapter addresses the control of oil and gas field exploration by means of exploratory drilling. We will limit our analysis to drilling because it is the most expensive. Its cost over recent years has increased relentlessly. Despite some success, the task of the discovery and appraisal of new oil and gas fields and of the accrual of reserves is not being fulfilled efficiently. For this reason, optimization of the exploration process is crucial.

Another limitation imposed on the process is that it must be considered within stationary conditions. The results of Parts I and II indicate that this process, in general, is not stationary (i.e., its parameters change in time). Still, time intervals may be identified in the evolution of the process when it may be considered stationary as a first approximation.

Indeed, an analysis of actual data indicates that annual average parameters of the exploration process over a certain time interval change insignificantly. For instance, the number of new field exploratory wells and the number of appraisal wells are fairly constant over these years. The number of new prospects drilled annually and the number of discovered fields in appraisal drilling also do not signifantly change from one year to the next. The average duration of well construction and the average depth of wells, as well as the discovery success rate and the number of exploratory and appraisal wells for each structure (or field), behave similarly.

Because the above parameters experience only insignificant fluctuations around the average (with no clear trend), and are obtained by averaging over large areas and over long time intervals, the exploration process may be considered stationary during limited time intervals. Stationary conditions substantially simplify modeling. Under these conditions the process parameters are not a function of time (within the specified time limits).

From the perspective of the completeness of information and the level of uncertainty, what is modeled is some averaged, aggregate control option.

Process Parameters: Problem Solving

The oil and gas exploration and appraisal process through well drilling is described by a number of parameters. From a problem-solving viewpoint, these parameters may be conveniently subdivided into three categories:

1. A goal parameter defining the purpose of exploration and appraisal. It is a planned accrual of reserves $Q = \tau\Delta Q$, where τ is the time interval during which Q must be obtained.
2. Methodological parameters defining exploration techniques (in a broad sense). They are considered here as variable parameters selected based on certain considerations. Strictly speaking, exploration control is reduced to the selection of these particular parameters. The following are the most important parameters:
   - Number of prospects $\overline{M}_{\text{exp}}$ and fields $\overline{M}_{app}$ completed by exploratory and appraisal drilling operations in a current year;
   - Number of prospects P_{exp} in new-field exploratory drilling
   - Number of fields P_{app} in appraisal drilling
   - Number of wells N_{exp} used for an evaluation (field/dry) of one prospect
   - Number of wells N_{app} expended for an appraisal of one field
   - Well footage expended for an evaluation of one prospect from the beginning of exploration (L_{exp}) and annually (H_{exp})
   - Well footage expended for an appraisal of one field from the beginning of appraisal (L_{app}) and annually (H_{app})
   - Duration of evaluating a single prospect by exploratory drilling (T_{exp})
   - Duration of appraisal stage of a single field (T_{app})
   - Number of rigs (n_{exp}) used for exploratory drilling in a current year
   - Number of rigs (n_{app}) used for appraisal drilling in a current year
   - Number of rigs (η_{exp}) used simultaneously in exploratory drilling for a single prospect in a current year

- Number of rigs (η_{app}) used simultaneously in appraisal drilling at a single field in a current year

3. Technological parameters that depend on the technology level and cannot be changed at will. This is why, from the viewpoint of exploration and appraisal process control, they are considered constant (at a given time and for a given region). They include:

   - Duration (t) of a single well construction in a current year
   - Average annual drilling footage per one rig (h) in a current year
   - Cost of a single well (c) in a current year

   The latter parameter, strictly speaking, is not a technological one (although it depends on the level of technology). It is convenient, however, to assign this parameter to this category because in this case it is also constant, which cannot be arbitrarily changed for the purpose of control.

Methodological parameters are not independent. They are closely interrelated and connected with the goal and technology parameters. In order for the exploration process to proceed in an efficient manner, there must be purposeful relations between these parameters. In the absence of this, the course of the process will sway from the optimum, which will significantly complicate control and, as a result, increase the cost of exploration.

Therefore, the problem is how to make the exploration and appraisal process optimally balanced.

Process Model: Controlling Parameters

Under stationary conditions, the exploration-appraisal process can be modeled as follows. If Q reserves needs to be accumulated over the time interval τ, then, at the average single field reserves of $\overline{Q}$, $M_{app} = Q/\overline{Q}$ fields should be appraised by the time τ. Respectively, $M_{\exp}$ prospects should also be explored by this time. Due to the stationary nature of this process, the appraisal should be completed annually on $\overline{M}_{app} = M_{app}/\tau$ fields. For the exploration and appraisal to be continuous, the number of fields where appraisal has been initiated annually should be equal to the number of fields where appraisal has been completed, which is $\overline{M}_{app}$. This means that exploratory drilling should be initiated annually for as many prospects as needed for the subsequent appraisal of $\overline{M}_{app}$ new fields (i.e., $\overline{M}_{\exp} = M_{\exp}/\tau$.) Therefore, exploratory and appraisal drilling

is initiated annually at $\overline{M}_{\text{exp}}$ and $\overline{M}_{app}$ structures. Annually, drilling is completed on same number of structures. Under these conditions, the number of prospects P_{exp} drilled with the purpose of discovering new fields and the number of fields P_{app} in appraisal drilling remain constant. This exploratory and appraisal procedure is schematically presented in Figure 8-1.

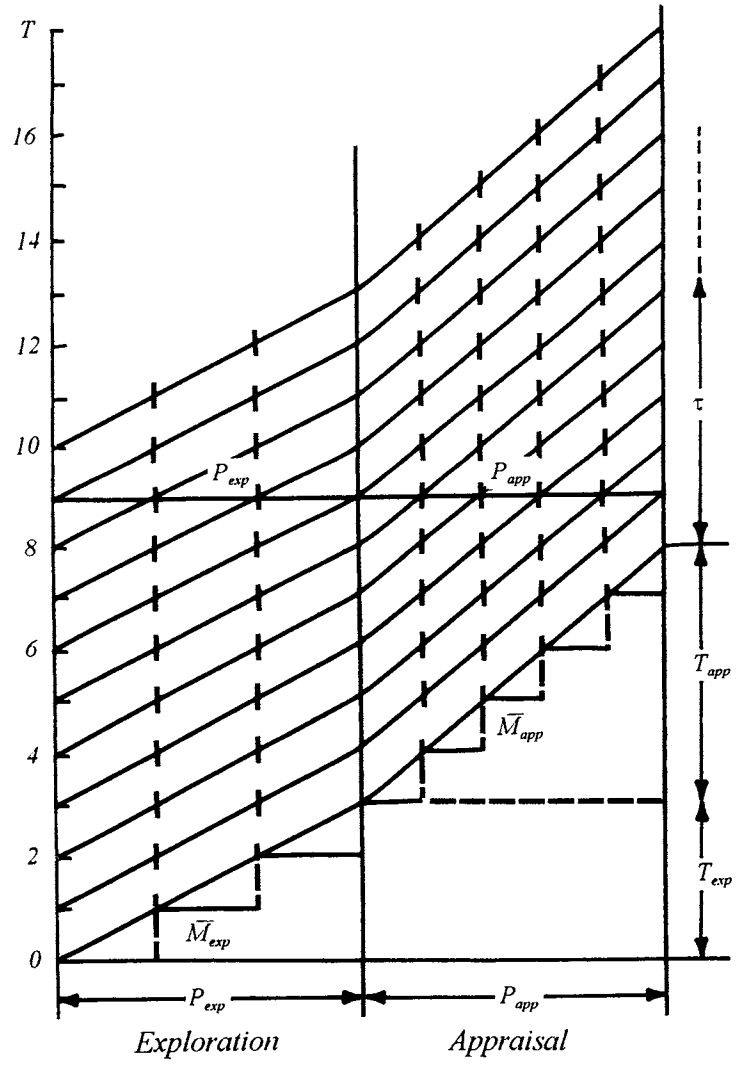


Figure 8-1. Schematic representation of the exploration-appraisal process.

General Relationships

If the average duration of the exploratory evaluation of a single prospect is T_{exp}, then, as shown in Figure 8-1, $P_{exp} = \overline{M}_{exp}T_{exp}$. Correspondingly, for the average duration of the appraisal T_{app}, $P_{app} = \overline{M}_{exp}T_{app}$.

On the other hand, let us assume that exploratory drilling at each prospect is conducted using on the average η_{exp} rigs, that appraisal drilling in each field is conducted using on the average η_{app} rigs, and that the average annual footage per rig (with all lost time accounted for) is h meters. Let us further assume that in order to complete the exploratory evaluation of a single prospect, N_{exp} wells are needed, and in order to complete the appraisal of a single field N_{app} wells must be drilled. If the average well depth is D, its drilling (including testing) will require time $t = D/h$ (h takes into account all lost time). Therefore,

$$T_{exp} = \frac{N_{exp}}{\eta_{exp}} t; \quad T_{app} = \frac{N_{app}}{\eta_{app}} t \tag{8-1}$$

Total duration of the exploratory and appraisal drilling at a single prospect/field will be equal to: $T = T_{exp} + T_{app}$.

Using Equation 8-1, P_{exp} and P_{app} can be represented as follows:

$$P_{exp} = \overline{M}_{exp} \frac{N_{exp}}{\eta_{exp}} t; \quad P_{app} = \overline{M}_{app} \frac{N_{app}}{\eta_{app}} t \tag{8-2}$$

If out of the total number of drilling rigs (n), n_{exp} are occupied in exploratory drilling, whereas n_{app} rigs are used in appraisal drilling, $P_{exp} = n_{exp}/\eta_{exp}$ and $P_{app} = n_{app}/\eta_{app}$. Then, using Equation 8-2:

$$n_{exp} = \overline{M}_{exp}N_{exp}t; \quad n_{app} = \overline{M}_{app}N_{app}t \tag{8-3}$$

If the annual exploratory drilling footage is H_{exp} and appraisal drilling footage is H_{app}, then $H_{exp} = P_{exp}\eta_{exp}h$ and $H_{app} = P_{app}\eta_{app}h$. From Equation 8-2:

$$H_{exp} = \overline{M}_{exp}N_{exp}th; \quad H_{app} = \overline{M}_{app}N_{app}th \tag{8-4}$$

Derived Equations 8-1 through 8-4 describe the interrelations among the parameters of the exploratory and appraisal process. They show that control of the process may be performed by the selection of three parameter groups: (1) $\overline{M}_{exp}$ and $\overline{M}_{app}$; (2) N_{exp} and N_{app}; and (3) η_{exp} and η_{app}.

Through these parameters, all the remaining methodological parameters and, therefore, the course of the exploratory-appraisal process, are defined. Compliance with Equations 8-1 through 8-4 in the course of the exploratory-appraisal process indicates a balanced process.

The parameters in these three groups are not independent. Inter-relationships among them are studied below.

Inter-target Relationships

Inasmuch as $\overline{M}_{\text{exp}}$ and $\overline{M}_{app}$ represent a set of objects (targets), the relationship between them can be considered as inter-target connections. Taking into account the fact that they also belong to different exploration stages (exploration and appraisal[1]), they also characterize the relations between these stages.

At any given time, the number of prospects where exploratory drilling has been completed should be greater than the number of fields where appraisal has been completed—$M_{\text{exp}} > M_{app}$. This is obvious because appraisal is initiated at the discovered fields. Only the M_{exp} targets found to be productive (kM_{exp}) will be appraised (where k is the field discovery success rate). Therefore, $kM_{\text{exp}} \geq M_{app}$. The difference $kM_{\text{exp}} - M_{app}$ determines the number of discovered fields (at each point in time) where appraisal has not yet been conducted. With an increasing number of such fields, they will begin to be appraised, and, hence, produced later (with everything else being equal). Accumulation of such fields results in the freezing of capital spent for their exploration. This is why the simplest case would be the one of minimum difference (i.e., $k\overline{M}_{\text{exp}} - M_{app} = 0$). This gives rise to the following basic equation:

$$M_{app} = kM_{\text{exp}}$$

The following method may also be applied. Let us assume not all the fields are equal. Under this assumption, it would not make sense to appraise all of them at this particular time. Thus, only some favorable fields would be appraised; namely, those that can fulfill the plan for the commercial accrual of the reserves at the lowest possible cost. Less-favorable fields would be suspended (temporarily abandoned), thereby not diverting funds needed for the appraisal of better fields. In other words, the accumulation of non-appraised fields in this case would be considered a positive factor, allowing us to concentrate efforts on targets of first priority. For instance, if there are a sufficient number of large and medium-size discoveries, there is no reason to appraise small fields, and, in fact, it would be best to postpone their appraisal.

Thus, not just the number of discoveries, but the part that will be appraised, is of importance. Obviously, in such a case the relations

between M_{exp} and M_{app} are no longer controlled by the discovery success rate k. Instead, it will be determined by the extent to which an increase in simultaneously explored prospects increases the probability of discovering ever-larger (i.e., more favorable) fields. Based on this, we would be able to appraise an ever smaller number of these fields, which nonetheless would provide the necessary accrual of reserves. In this case, the basic equation will convert to $M_{app} = k_1 M_{\text{exp}}$; thus:

$$\overline{M}_{app} = k_1 \overline{M}_{\text{exp}} \tag{8-5}$$

where k_1 is the success rate only for those fields that will be appraised.

In a general case of $k_1 \leq k$, k_1 is equal to the success rate ($k_1 = k$) only in the case when all discoveries are appraised. This may occur in two situations. The first situation is where all fields (discovered and yet to be discovered) have equal parameters and are equally unfavorable (i.e., there are no better, more-favorable fields to select for the appraisal). The second is where, due to insufficient drilling footage at the exploratory, pre-appraisal stage, a necessary backlog of discovered fields was not accumulated in order to select the best one for appraisal. This makes it clear that the k_1 parameter, to an extent, reflects natural conditions and indicates those fields for which an appraisal is justified. For instance, k_1 may be the portion of large and medium fields among all discoveries. In this sense, the value of k_1 is set by geological processes (nature). In addition, it also depends on the exploration techniques. Thus, the economics define the rules of field differentiation into classes (i.e., whether or not it is economical to appraise them), although they cannot dictate the changes in the ratios among the number of fields in different classes (ratios created by nature). At the same time, the quality of exploration (e.g., prospect delineation by seismic) and the reliability of the oil and gas forecast determines the share of "dry" targets among the total number of exploration targets, and, hence, the value of the discovery success rate k, related directly to k_1.

This leads us to believe that in a given natural environment and at a given discovery level, k_1 controls the average size of reserves $\overline{Q}$ accrued as a result of one field appraisal. This is discussed below in detail.

Whether to appraise a discovered field or to suspend it depends on many factors. Oil and/or gas reserves of the field are among the most important factors. For simplicity, we will assume that the screening-out of the discovered field is based exclusively on its size. In other words, the largest fields in a given region are appraised, whereas fields with a lower volume of reserves are suspended (temporarily abandoned). Then, at a given number of already discovered fields with increasing k_1 (up to and including k), an ever-increasing share of these fields is transferred to appraisal, which results in an increase in the number of fields being

appraised. This is performed at the expense of appraising fields with ever-decreasing reserves, which, in turn, leads to a decrease in the average volume of reserves per one appraised field. Therefore, provided that there is a set number of discovered fields, an increase in k_1 results in a decrease in the average volume of reserves per field for fields intended for appraisal.

In order to attain a balanced exploration process and find its optimum parameters (including optimum k_1 value), the following question must be answered: What would be the average reserves of a single field, not at the actual k_1 value, but at different values up to k? Obviously, they would be smaller than the actual volume of reserves, and their minimum would satisfy the condition $k_1 = k$ (a condition when all discovered fields are appraised). Inasmuch as the suspended fields have not been appraised, the estimate of their reserves will be only approximated (they must be calculated). For this purpose, we proposed and solved the following problem.

Let us assume we have a set A of all fields (appraised and suspended). Let us further assume that the field sizes in the A set are distributed according to the function $f(Q)$. The fields within the A set are subdivided into regional groups a, $a_1, \ldots, a_w$. In turn, each group a_i is subdivided into two subgroups (b_i and c_i) with the respective number of fields m_i and r_i. The fields in the subgroup b_i form a set $B = \{b_i\}$ and the fields in the subgroup c_i form a set $C = \{c_i\}$, where $A = B \cup C$. Correspondingly, the number of fields in sets A, B, and C equals $\Sigma(m_i + r_i)$, Σm_i, and Σr_i.

B is the set of appraised fields and C is the set of suspended fields. Let us assume that a function $g(Q)$ describes the distribution of fields sizes in set B. We need to restore an unknown function $f(Q)$ from the known function $g(Q)$. This problem was solved as follows. The average size of reserves $\overline{Q}(b_i)$ of a field was calculated and the number of fields m_i was counted in each region of the b_i set. Then the regions were positioned in a variational set in order of decreasing average size of reserves:

$$\overline{Q}(b_1) > \overline{Q}(b_2) > \overline{Q}(b_3) > \ldots > Q(b_w)$$

The values of

$$m_1, m_2, m_3, \ldots, m_w$$

$$r_1, r_2, r_3, \ldots, r_w$$

were positioned, respectively. No specific data was known concerning the reserves of the suspended fields so that the following simplest assumption was made: the average size of reserves $\overline{Q}(c_i)$ of a single suspended field in the region, having the sequential number i, with equal probability may be equal to the average size of reserves $\overline{Q}(b_q)$ of a single appraised field in any region with a greater sequential number ($q > i$); that is:

$$P(\overline{Q}(c_i) = \overline{Q}(b_q)) = \frac{1}{W-1} \qquad (q = i+1,\ i+2,\ \ldots,\ W-1)$$

This supposition appears to be quite natural because its alternative is excluded by the definition of the problem. According to this assumption, the r_i value was equally distributed between all b_q sets (regions) with ordinal numbers greater than i ($q > i$). Thus, inasmuch as the function $q(Q)$ determined the number of fields with reserves $\overline{Q}(b_i)$ (on the average) as m_i, then it will be determined by the function $f(Q)$ as:

$$n_i = m_i + \sum_{q=1}^{i=1} \frac{r_q}{W-q}$$

Respective probabilities can be assigned as follows:

$$P_{g_i} = \frac{m_i}{\sum m_i} \quad \text{and} \quad P_{f_i} = \frac{n_i}{\sum n_i}$$

These expressions can be rewritten in the following form:

$$P_{f_i} = \frac{n_i}{\sum n_i} = \frac{m_i + \sum_{q=1}^{i-1} \frac{r_q}{Q-q}}{\sum_{i=1}^{W} m_i + \sum_{i=1}^{n} r_i}$$

$$= \frac{\sum_{i=1}^{W} m_i \times \frac{m_i}{\sum_{i=1}^{W} m_i} + \frac{r_1}{W-1} + \frac{r_2}{W-2} + \ \ldots\ + \frac{r_{i-1}}{W-(i-1)}}{\sum_{i=1}^{W} (m_i + r_i)} =$$

$$P_{g_i} \sum_{i=1}^{W} m_i + \frac{r_1}{\sum_{i=1}^{W} r_i} \times \frac{1}{W-1} \sum_{i=1}^{W} r_i + \frac{r_2}{\sum_{i=1}^{W} r_i} \times$$

$$\times \frac{1}{W-2} \sum_{i=1}^{W} r_i + \ \ldots\ + \frac{r_{i-1}}{\sum_{i=1}^{W} r_i} + \frac{1}{W-(i-1)} \sum_{i=1}^{W} r_i$$

$$= \frac{}{\sum_{i=1}^{W} (m_i + r_i)}$$

Introducing

$$W_q = \frac{r_q}{\sum_{i=1}^{W} r_i}; \quad P_q = \frac{1}{W-q}; \quad P_1 = \frac{\sum_{i=1}^{W} m_i}{\sum_{i=1}^{W} (m_i + r_i)} \quad \text{and} \quad \frac{\sum_{i=1}^{W} r_i}{\sum_{i=1}^{W} (m_i + r_i)}$$

one obtains:

$$P_{f_i} = P_1 P_{g_i} + P_2 \sum_{q=1}^{i-1} W_q P_q$$

Thus, a conversion of the $g(Q)$ distribution into the $f(Q)$ distribution is accomplished by the following rule:

$$f(Q) = P_1 g(Q) + P_2 \int_{\Theta > Q} \dot{Q(\Theta)} q(\Theta) d\Theta$$

The last transformation can be interpreted as follows. The value

$$W_q = r_q \Big/ \sum_{i=1}^{W} r_i$$

is the share of suspended fields in the regions with ordinal number q, where the average reserves per appraised field are $\overline{Q}(b_q)$. The expression $P_q = 1/(W - q)$ is the probability that the average reserves for the suspended fields in the region numbered q will equal the average reserves $\overline{Q}(b_1)$ of the appraised fields in the region numbered l ($l > q$). Their combination (at $q = 1, 2, \ldots, i - 1$) represents the probabilitiy that the average size of reserves for a single appraised field $\overline{Q}(b_i)$ in a region numbered i will equal the average reserve size for a single suspended field in the regions numbered $1, 2, \ldots, i - 1$. Their sum weighted for the proportions of the suspended fields in these regions, that is:

$$\sum_{q=1}^{i-1} W_q P_q$$

is the average probability that a single, suspended field may contain reserves $\overline{Q}(b_i)$.

P_1 and P_2 are probabilities of appraising and suspending the discovered fields. Thus, the probability

$$P_{f_i}$$

which is given by the $f(Q)$ function, is the sum of two addenda. One of the addenda is the probability that the appraised fields will have the reserves $\overline{Q}(b_i)$. It is taken with weight P_1 representing the probability of the appearance of fields being appraised. The second addendum is the (average) probability that the suspended fields will have the same average reserves $\overline{Q}(b_i)$. It is summed with the weight P_2, which reflects the probability of the appearance of fields being suspended. In other words, the probability P_{f_i} of the average reserves of a single field being equal to $\overline{Q}(b_i)$ is the sum of weighted probabilities of the average reserves for a single appraised field and a single suspended field being equal to $\overline{Q}(b_i)$. Therefore, the transition from $g(Q)$ to $f(Q)$ may be given a real physical meaning.

The function $g(Q)$ is based on the reserves data of the actually appraised fields; thus, it corresponds to the actual value of $k_1 = k_o$. On the other hand, the $f(Q)$ function corresponds to the value $k_1 = k$ since it describes the size distribution for all discovered fields regardless of whether they have been suspended or transferred for appraisal.

Correspondingly, the average size of reserves of a single field is expressed by the following equations:

$$\overline{Q}_{k_o} = \frac{\sum_{i=1}^{W} \overline{Q}(b_i) m_i}{\sum_{i=1}^{W} m_i} \qquad (\text{at } k_1 = k_o)$$

$$\overline{Q}_k = \frac{\sum_{i=1}^{W} \overline{Q}(b_i) \left(m_i + \sum_{q=1}^{i-1} \frac{r_q}{W - q} \right)}{\sum_{i=1}^{W} (m_i + r_i)} \qquad (\text{at } k_1 = k)$$

Some intermediate k_1 values, lying between k_o and k, correspond to a special case when the total number of fields for which the size of reserves is calculated, does not include all of the appraised fields (when $k_1 = k_o$) and all of the appraised and suspended fields (when $k_1 = k$). Instead, it includes the appraised and some of the suspended fields, that is:

$$\sum_{i=1}^{W}(m_i + d_i)$$

fields ($\sum d_i < \sum r_i$). Therefore, in order to determine the average size of reserves of a single field corresponding to an arbitrary k_1 value in the interval from k_o to k, it is necessary to transform the initial distribution $g(Q)$ into the $f_1(Q)$ distribution. This transformation is performed according to the rule developed above. It is necessary to keep in mind that the specific probability values will change so that the average value of reserves will be determined from the following equation:

$$\overline{Q}_{k_1} = \frac{\sum_{i=1}^{W} \overline{Q}(b_i)\left(m_i + \sum_{q=1}^{i-1} \frac{d_q}{W-q}\right)}{\sum_{i=1}^{W}(m_i + d_i)}$$

Let us transform this equation. Inasmuch as $k_o/k = \sum m_i/\sum(m_i + r_i)$, $\sum m_i = (k_o \sum r_i)/(k - k_o)$. On the other hand, from equations $k_o/k = \sum m_i/\sum(m_i + r_i)$ and $k_1/k = \sum(m_i + d_i)/\sum(m_i + r_i)$, we obtain the equation $k_o/k_1 = \sum m_i/\sum(m_i + d_i)$ which gives equation $\sum d_i = [(k_i - k_o)/k_o]\sum m_i$. Inserting the expression found for $\sum m_i$, we will obtain the following equation:

$$\sum d_i = \frac{k_1 - k_o}{k - k_o} \sum r_i$$

Therefore, $\sum d_i$ linearly increases with increasing k_1. Variation of the specific d_i values is unknown. Taking into account that their sum increases linearly, however, it is reasonable to assume that:

$$d_i = \frac{k_1 - k_o}{k - k_o} r_i$$

Thus, one can obtain the following equation:

$$\sum_{q=1}^{i-1} \frac{d_q}{W-q} = \frac{k_1 - k_o}{k - k_o} \sum_{i=1}^{i-1} \frac{r_q}{W-q}$$

In this case, the equation for $\overline{Q}_{k_1}$ will be as follows:

$$\overline{Q}_{k_1} = \frac{\sum_{i=1}^{W} \overline{Q}(b_1)\left(m_1 + \frac{k_1 - k_o}{k - k_o} \sum_{q=1}^{i-1} \frac{r_q}{W - q} \right)}{\sum_{i=1}^{W} m_i + \frac{k_1 - k_o}{k - k_o} \sum_{i=1}^{W} r_i}$$

After transformation one will obtain:

$$\overline{Q}_{k_1} = \frac{(k - k_o)\sum_{i=1}^{W} \overline{Q}(b_i)m_i - k_o \sum_{i=1}^{W} \overline{Q}(b_i) \sum_{q=1}^{i-1} \frac{r_q}{W - q} + k_1 \sum_{i=1}^{W} \overline{Q}(b_i) \sum_{q=1}^{i-1} \frac{r_q}{W - q}}{(k - k_o)\sum_{i=1}^{W} m_i + (k_1 - k_o)\sum_{i=1}^{W} r_i}$$

After grouping the variables, which in this case are assumed to be constant and equal to their actual values, and introducing the parameters α, β, and γ as follows:

$$\alpha = (k - k_o)\sum_{i=1}^{W} \overline{Q}(b_i)m_i - k_o\beta; \quad \beta = \sum_{i=1}^{W} \overline{Q}(b_i) \sum_{q=1}^{i-1} \frac{r_q}{W - q}$$

$$\gamma = (k - k_o)\sum_{i=1}^{W} m_i - k_o\varepsilon; \quad \varepsilon = \sum_{i=1}^{W} r_i$$

the following equation can be obtained:

$$\overline{Q}(k_1) = \frac{\alpha + \beta k_1}{\gamma + \varepsilon k_1} \tag{8-6}$$

Equation 8-6 provides an answer to our question. It enables us to determine the average reserves $\overline{Q}(k_1)$ added after the appraisal of a single field as a function of k_1. Numerical values α, β, γ, and ε in Equation 8-6 were determined from the actual data accrued during the investigation period, and a graph of $\overline{Q}$ versus k_1 was constructed. The graph shows that the average size of reserves added after the appraisal of a single field $\overline{Q}$ smoothly declines as k_1 increases [20].

On average, the annual accrual of reserves in one appraised field will be $\overline{Q}(k_1)/T_{app}$, and in P_{app} fields will be $Q_{ann} = P_{app}\overline{Q}(k_1)/T_{app} = \overline{M}_{app}\overline{Q}(k_1)$. Thus, it is obvious that the number $\overline{M}_{app}$ of annually appraised fields must

be determined by the required annual accrual of reserves Q_{ann} and the average reserve accrual from the appraisal of a single field $\overline{Q}(k_1)$, namely:

$$\overline{M}_{app} = Q_{ann}/\overline{Q}(k_1) \tag{8-7}$$

where $Q_{ann} = Q/\tau$.

$\overline{Q}(k_1)$ is a function of k_1. The number of annually explored prospects $\overline{M}_{exp}$, according to Equation 8-5, is related through the same k_1 parameter to the number of annually appraised fields $\overline{M}_{app}$. Therefore, all of these parameters, at an assigned accrual of reserves Q, will depend on k_1. (The Q parameter as the goal parameter must be assigned.)

Thus, our analysis of the interconnection between $\overline{M}_{exp}$ and $\overline{M}_{app}$ indicates that their selection during exploration and appraisal process control is, in turn, determined by the selection of k_1. In this case, the number of targets in exploration and appraisal are linked to each other and to the accomplishment of the assigned task, which is to achieve the assigned accrual of reserves Q during the required time interval.

Intra-target Relations

The second (N_{exp} and N_{app}) and third (η_{exp} and η_{app}) groups of exploration-appraisal process control parameters describe operations on a local (within a single target) rather than regional scale.

The N_{exp} and N_{app} parameters, as well as the k_1 parameter, also reflect natural geologic conditions. Geology dictates that, even with the best exploration techniques, a minimum number of wells must be drilled in order to fulfill the exploration and appraisal task. This minimum cannot be decreased when certain information is required. On the other hand, the number of wells is directly related to the applied techniques. Economics also determines the amount of information required under specific conditions.

The η_{exp} and η_{app} parameters reflect the rate (tempo) of operations. They are not directly related to the geologic scenario and, thus, are purely technological.

The number of wells N_{exp} and N_{app} drilled at a single target for fulfilling the exploratory and appraisal tasks depends on the number of wells simultaneously started (η_{exp} and η_{app}). Determining how many wells should be started simultaneously (i.e., what should be the rate of operations and what would be gained or lost if the rate is increased or decreased) has not been precisely defined. Because existing techniques to determine the optimum number of wells do not take this factor into account, the results should be examined with caution.

What is gained from a change in the rate (tempo) of operations depending on the number of wells drilled? This question can be answered

by determining how the number of wells used for the exploration of one prospect and the appraisal of one field is related to the number of rigs simultaneously drilling at a single target (the number of simultaneously drilled, independent wells). A relationship exists based on the successive adaptation of the exploration-appraisal process to the information being accumulated. Indeed, the exploratory and appraisal process occurs in such a manner that drilling of subsequent wells is guided by the results obtained from the preceding wells. If many wells are started simultaneously, the most likely result is that some of them are not drilled in the best geologic environment. This causes an increase in the number of wells necessary to obtain the amount of information required for fulfilling the exploratory and appraisal tasks.

Let us assume that I_{exp} and I_{app} are the amounts of information needed to explore one prospect and to appraise one field, respectively. Let us further assume that only one well is being drilled in the field. In other words, exploration and appraisal are conducted using one rig ($\eta_{\text{exp}} = 1$ and $\eta_{app} = 1$), and all wells are dependent.[2] Let us further assume that in order to obtain the amount of information I_{exp} and I_{app}, we had to drill, respectively, $N^{(1)}_{\text{exp}}$ and $N^{(1)}_{app}$ wells. If we simultaneously start two exploratory and two appraisal wells ($\eta_{\text{exp}} = 2$ and $\eta_{app} = 2$), then in order to obtain the same amounts of information I_{exp} and I_{app}, a greater number of wells is needed: $N^{(2)}_{\text{exp}}$ and $N^{(2)}_{app}$. The $N^{(2)}_{\text{exp}} > N^{(1)}_{\text{exp}}$ and $N^{(2)}_{app} > N^{(1)}_{app}$ because the locations of two simultaneously drilled wells will generally be different from that of the two subsequently started wells. The former situation is less favorable because we cannot correct the location of the subsequent well based on the information obtained from drilling the preceding well. Similarly, if three wells are started simultaneously ($\eta_{\text{exp}} = 3$ and $\eta_{app} = 3$), then under the same conditions and for the same purpose, $N^{(3)}_{\text{exp}}$ and $N^{(3)}_{app}$ number of wells will be needed ($N^{(3)}_{\text{exp}} > N^{(2)}_{\text{exp}}$ and $N^{(3)}_{app} > N^{(2)}_{app}$).

Thus, generally speaking, increasing the number of simultaneously started wells (the number of simultaneously used rigs) will cause the number of wells to increase. As a first approximation and lacking other considerations and sufficient knowledge of the problem, this increase may be considered linear. In this case, the aforementioned correlations may be presented as follows:

$$N_{\text{exp}} = N^{(1)}_{\text{exp}} + a_{\text{exp}}(\eta_{\text{exp}} - 1) \tag{8-8}$$

$$N_{app} = N^{(1)}_{app} + a_{app}(\eta_{app} - 1)$$

where $N^{(1)}$ and a are the parameters.

Evaluation of the parameters led to the following equations [19]:

$$N_{exp} = 2.05 + 1.73(\eta_{exp} - 1) \tag{8-9}$$

$$N_{app} = 8.15 + 3.59(\eta_{app} - 1)$$

These equations should be used for determining optimal values of N_{exp} and N_{app}.

The derived Equations 8-9 (relationship between N and η) enables one to introduce a new parameter characterizing the efficiency of locating the wells (i.e., the $a/N^{(1)}$ ratio). When $a = 0$ and, therefore, $a/N^{(1)} = 0$, the N value does not depend on η and remains constant and equal to $N^{(1)}$. In other words, the final number of wells is the same whether the wells are drilled one after another, started in groups, or started all at once. Obviously, this will occur only when the selection of well locations (drilling of wells started simultaneously) was as successful as locations selected using sequential dependent drilling. Thus, the dependent selection of well locations is less justified the closer the $a/N^{(1)}$ ratio is to zero.

If, on the other hand, $a/N^{(1)} = 1$, then N increases in the same manner as η. If, for instance, the number of simultaneously started wells increases by a factor of 2, then the final number of wells will also increase by a factor of 2. This situation is possible only in a case when one of the two simultaneously started wells helps achieve the task (gives useful information), whereas the second well is useless in providing information due to an extremely unfavorable geologic location. The closer the $a/N^{(1)}$ ratio is to one, the more justified the sequential drilling of wells is. Even if several wells are started simultaneously, only one will be useful, and no rate (tempo) acceleration of operations will occur due to unfavorable locations of the other wells.

It is obvious that $0 \leq a/N^{(1)} \leq 1$. If $a/N^{(1)} > 1$, the increase in N should surpass that of η. This would be possible only in the case where, for example, one well provides more useful information than two simultaneously drilled wells. This is practically impossible because the location of one of the two simultaneously started wells coincides with the location of one of the two subsequently started wells. In our specific case, we obtained:

$$a_{exp}/N^{(1)}_{exp} = 0.84$$

$$a_{app}/N^{(1)}_{app} = 0.44$$

This shows that at the exploratory stage, starting the drilling of wells simultaneously did not provide any significant advantage because the wells turned out to be unfavorably located. For any practical purpose, only

one of the two simultaneously started wells was actually exploring. If so, it would be better to start drilling wells sequentially, one after another.

The situation was different at the field appraisal stage. In this case, the simultaneous initiation of drilling of several wells was advantageous—this actually accelerated the rate (tempo) of the process. This was due to a more favorable selection of the locations of a series of wells being drilled simultaneously. "More favorable" means that the locations for a substantial portion of wells coincided with those which would have been selected in the case of sequential drilling (one well after another).

The following conclusions can be made in summarizing this chapter:

1. The number of wells for exploring a single prospect (N_{exp}) and appraising a single field (N_{app}) depends on the number of simultaneously working rigs (the number of simultaneously started wells). (See Equation 8-9.)
2. At the exploratory stage, only "dependent" wells are justified (this confirms the recommendations made by Lindtrop and Mukhin [31] to explore prospects using a single well technique). At the appraisal stage, a certain number of "independent" wells is acceptable for the acceleration of process rate (tempo).
3. An increase in the rate of the exploratory-appraisal process comes with a price: an increase in the number of wells. This rate cannot increase without an increase in spending. This factor should be taken into account when determining the optimum number of exploratory and appraisal wells at a single target.

Equation 8-9 individually reflects only in-pair relationships for the exploratory and appraisal operations. They are not mutually related. Whereas inter-target relations, at the same time, reflected relations between the stages, the intra-target relations are only those within the individual stages.

The model we developed is mathematically described by Equations 8-1 through 8-8. This model provides an opportunity to condense the control over the exploratory-appraisal process to the selection of just two groups of parameters: (1) k_1; and (2) η_{exp} and η_{app}. These parameters represent the process control parameters as well. All methodological parameters, and, hence the course of the process, are determined from them.

Control is reduced to the selection of control parameters if Equations 8-6 and 8-8 are known. The first equation indicates the extent to which an increase in the number of simultaneously explored prospects improves the selection of the discovered fields for further appraisal.

Equation 8-8 shows how an increase in the operation rate (i.e., the number of simultaneously drilled wells) affects an increase in the number of wells used for the exploration of one prospect and appraisal of one field. Balanced character is achieved for the process through coordination of its parameters with the control parameters according to Equations 8-1 through 8-5 and Equation 8-7.

Notes

1. Strictly speaking, "exploration" encompasses geologic, geophysical, and other operations for the purpose of discovering deposits of economic minerals (e.g., hydrocarbons), whereas "appraisal" includes operations for evaluating the commercial value of the discoveries.
2. "Dependent" wells are those drilled sequentially (one after another) so that the information obtained from one well can be used to select the location of a subsequent well.

CHAPTER 9

Optimization of Oil and Gas Exploration-Appraisal Process

As discussed in Chapter 8, in order to optimize the exploration-appraisal process, optimum values of control parameters must be found. Optimum selection of the control parameters uniquely determines the course of the exploration-appraisal process and its balance. In turn, the search for optimum values of the control parameters requires a preliminary formulation of the optimization criterion or a determination of the goal function.

Optimum Parameters of the Exploration-Appraisal Process

Previously, we assumed the minimum of normalized spending as the optimization criterion. Thus, we were searching for the optimum solution that corresponds to the minimum normalized spending z required for the accrual of reserves by the actually achieved value Q:

$$z = (cN_{\text{exp}}M_{\text{exp}} + cN_{app}M_{app})(1 + 0.08)^T + 0.12(\chi n_{\text{exp}} + \chi n_{app})(1 + 0.08)^T \quad (9\text{-}1)$$

where c is the cost of a single well and χ is the price of one drill rig with all the necessary equipment.

Technological parameters were also assumed to be equal to their actual average values. All remaining variable parameters had to be evaluated.

The minimum of normalized spending was determined as follows. Assigning arbitrary values to the control parameters, the remaining process parameters were determined from Equations 8-1 through 8-8. Then expenditures z were calculated using all the parameters. Values of k_1 ranged from the actual k_o (0.178) to the field discovery success rate k (0.331). The latter describes the condition when all discovered fields are appraised. Thus, by exploring various control parameter values and determining the corresponding expenditure for each, it was possible to find a combination of control parameters corresponding to the minimum expenditures z.

The selected optimum parameters of the exploration-appraisal process are listed in Table 9-1. They represent the mean statistical parameters for the former USSR (entire regions of exploration) over the studied time interval. For this reason, the parameters which are integer-valued in their nature, are shown in Table 9-1 as fractional values.

Comparison of the actual parameters of the exploration-appraisal process with their optimum values (shown in Table 9-1) reveals that some are quite different, whereas others are almost identical. For instance, the duration of an exploratory evaluation of an individual prospect by drilling (T_{exp}) is close to the optimum, whereas the appraisal duration of an individual field (T_{app}) is significantly shorter. This means that the appraisal is performed faster than required. The number of prospects at which the exploratory stage is completed annually ($\overline{M}_{\text{exp}}$) and the number of fields in which the appraisal is completed annually ($\overline{M}_{app}$) are practically optimum. At the same time, the number of prospects in exploratory drilling (P_{exp}) and the number of fields in appraisal drilling (P_{app}) were substantially higher than the optimum. This means that the exploration and appraisal operations are spread over an unusually large number of targets. The reason for this is that the number of rigs conducting exploratory and appraisal drilling is substantially greater than the optimum. As a result, the number of rigs simultaneously operating at one prospect η_{exp} (exploratory drilling) and in one field η_{app} (appraisal drilling) is significantly greater than the optimum. Thus, the number of wells N_{exp} drilled for exploring an individual prospect and the number of wells N_{app} used for appraising an individual field are higher than the optimum.

Optimality Analysis of the Oil and Gas Exploration in the Former Soviet Union

An analysis of Table 9-1 provides answers to a number of important questions and unresolved issues concerning exploration methodology. These issues are reviewed below.

Table 9-1
Parameters of the Exploratory-Appraisal Process

Parameter	Optimum Value	Actual Value
Coefficient k_1	0.178	0.178
Portion of discoveries transferred to appraisal, k_1/k	0.54	0.54
Number of rigs simultaneously operating at an individual prospect in exploratory drilling, η_{exp}	1	1.37
Number of wells drilled for exploration of an individual prospect, N_{exp}	2.05	2.3
Duration of exploratory drilling evaluation for an individual prospect, T_{exp}	2.97	2.83
Number of prospects annually evaluated by exploratory drilling, $\overline{M}_{exp}$	232.29	237
Number of prospects in exploratory drilling, P_{exp}	689.9	762.4
Number of rigs conducting exploratory drilling in a current year, n_{exp}	689.9	1,043.40
Number of rigs simultaneously conducting appraisal drilling in an individual field, η_{app}	1.55	2.7
Number of wells expended for appraisal of an individual field, N_{app}	10.12	10.4
Appraisal duration of an individual field, T_{app}	9.47	8.27
Number of fields annually evaluated by appraisal drilling, $\overline{M}_{app}$	41.35	40.6
Number of fields in appraisal drilling, P_{app}	391.44	444.4
Number of rigs in appraisal drilling during a current year, n_{app}	607.25	1,204.80
Normalized expenditures, z (% of optimum expenditures)	100	104

Prospect Drilling Strategy. N. T. Lindtrop (1958) was the first in the former USSR to raise the question of the need for starting exploratory drilling using one rig (one or two wells) [31]. Several years later it was the recommendation of both Lintrop and V. V. Mukhin "to conduct exploratory drilling of positive structures using one to two wells (one rig) and to drill subsequent wells depending on the results from the preceding wells" [31, p. 403]. Currently, this concept is supported by most investigators and was substantiated by the fact that "when drilling anticlinal

structures, the first or the second well is usually a discovery well; if two first wells do not yield a commercial flow, and if further (third, fourth, etc.) wells are drilled without finding petroleum, it is only rarely that commercial oil or gas are discovered" [31, p. 395]. Bunin et al. offer a somewhat different, although unjustified in their publication, recommendation: "In order to shorten the time for new field discoveries and to increase the success rate in the regions with no known commercial oil and gas, it is recommended to drill the prospective structures using one rig and in the regions with known commercial fields, using two or three rigs simultaneously" [55, p. 20].

Table 9-1 indicates that exploratory drilling should be conducted using only one rig. Simultaneous drilling of several wells does not improve exploration efficiency; shortened evaluation time, in this case, results in an unjustifiably high cost.

Gabrielyants and Sorokin stated that "the procedure of a prospect drilling using one rig results in the improved location of each subsequent well and more efficient use of geologic information from the preceding well while drilling, logging and testing the subsequent wells. However, this slows down the evaluation of the prospect's commercial value. This affects in a negative way the exploration and appraisal process in the entire region . . . which justifies the economic efficiency of simultaneous drilling of several wells" [10, p. 6]. The latter statement is not justified. The results listed in Table 9-1 characterize the best course of exploration and appraisal in the former USSR and were obtained based on criteria of economic efficiency.

Number of Exploratory Wells Per Individual Prospect. Lindtrop and Mukhin discussed the advantages of drilling a prospect using one or two wells (one rig) [31]. These advantages are associated with the fact that most fields are discovered by one or two wells and that drilling three or more wells usually results in a discovery of small fields. Thus, one can raise the following question: Does it pay to spend additional funds if the first two wells turn out to be dry? [31, p. 400]. Lindtrop and Mukhin further concluded that in most cases it is sufficient to drill three wells. The recommended number of wells corresponds to discovering only part of the fields. It is assumed that the rest of the fields will be missed, but these fields, as a rule, will be insignificant. In other words, a certain acceptable risk is understood and, therefore, some fields will be classified as dry. It is important to recognize that a set of fields is the main concern, and not each individual field.

In this book, we will be discussing a different problem. Based on the results of exploratory drilling, the answer to the question of whether a drilled prospect contains a field or is dry must be unambiguously singular.

The element of acceptable risk is excluded here. The recommended number of wells should accomplish the task of exploring each individual prospect; however, the time factor is taken into consideration. Melik-Pashayev stated that "the number of exploratory wells depends on the rate (tempo) of exploration. The more exploratory wells are simultaneously drilled on an individual high, the greater number of wells may turn out to be dry. On the other hand, if drilling is conducted sequentially (depending on the results from a previously drilled exploratory well), exploration may be stretched over a long period of time. The right solution . . . depends on the potential of the region: the higher the potential, the higher should be the exploration tempo, and the greater the number of exploratory wells started" [33, p. 16].

As shown by our analysis (Table 9-1), the above considerations will be satisfied by drilling, on average, two wells at each prospect. This is the optimum number.

Duration of the Exploratory Drilling Evaluation for an Individual Prospect. We were not able to find any published recommendations with regard to this subject. The published information only relates to the time actually expended for specific prospects or for specific regions.

We arrived at an optimal estimate that is equal to approximately three years. It is important to note that this parameter depends on a number of factors, including the transportation time and the time for well construction and testing. While determining the optimum, we considered these factors constant. Their values were taken to be equal to the actual values. Obviously, if these time expenditures are shortened, the optimum duration of the exploratory drilling evaluation for an individual prospect will also be shortened.

Our evaluation is based on the total time balance. Kozlovsky discussed efficient time distribution at different stages [23]. In particular, he noted that as a result of worktime redistribution among the exploration stages, it is possible: (1) to decrease the total spending for the same amount of appraised reserves; (2) to increase the amount of the appraised reserves at the same level of expenditures; or (3) both (1) and (2). The duration of the exploratory phase as indicated above is based on the rational time distribution between the two stages under analysis (exploration and appraisal) and the decrease in total expenditures.

Rate (Tempo) of Appraisal. We reduced the problem of exploration rate to the number of appraisal wells drilled simultaneously or, in other words, to the number of rigs used in the appraisal of an individual field.

Although this problem is widely discussed in the literature, we were not able to find any direct suggestions regarding the appraisal rate. There

are, however, some indirect suggestions based on recommended well patterns (triangular, ring, profile, etc.) According to Melik-Pashayev, the triangular system is not very efficient due to the fact that drilling each new well depends on the positive results obtained from the previous one; hence, the appraisal is slowed down [33]. In his opinion, as well as that of other investigators, the profile system (a row of wells drilled simultaneously) is the most efficient.

Frolov et al. do not exclude the possibility of field appraisal using the "creeping" system, when mostly dependent wells are drilled sequentially or in small groups of 2 or 3 [39]. They prefer the "concentrated" appraisal system, however, when a significant number of independent wells are drilled and the dependent wells are started in groups. Other investigators indicated that, depending on the specific conditions, not only the profile system but also the uniform triangular system may be efficient. Another group of investigators is in favor of the preliminary delineation of the accumulation. They believe that such a delineation must be conducted by drilling individual "dependent" (sequential) wells.

As we have shown, there are a number of contradictory opinions concerning this problem. Most importantly, the question is not answered in an unequivocal manner.

As the above discussion indicates, our solution to the problem is based on the fact that the increase in the appraisal rate (tempo) results in an increase in the total number of wells. This situation, as well as the need to establish how many wells depend on the number of simultaneously working rigs, enables one to consider the appraisal rate as a control parameter that determins the funds and the time expenditures for exploration-appraisal.

The optimal number of simultaneously working rigs is 1.55. Thus, to appraise two fields requires an average of three rigs. In other words, half of the fields should be appraised by one rig and the other half by two rigs. This does not mean that the appraisal of some particular fields cannot be made using three to four, or even five rigs. It does, however, indicate that the number of fields appraised using one rig will be more than 50% of the total number.

This result casts some doubt on the universal efficiency of the profile system, or the simultaneous drilling of a well profile (or of several well profiles).

Appraisal Duration of an Individual Field. Although no literature could be found on this subject, the problem can be stated as follows: What should the appraisal duration be in order to achieve a certain goal or satisfy a certain criterion? Table 9-1 shows that the actual duration of

appraisal is shorter than the optimum. This leads us to wonder if the values in the table are prolonging the appraisal.

It is important to mention that the optimum estimate is obtained from minimizing normalized spending (i.e., the goal for the indicated duration is not to spend an unnecessary amount of funds). In practice a different tenet is pursued, that of the speediest initiation of appraisal regardless of the amount of money spent. This occurs frequently. It is believed that shortening the time by spending more money may be beneficial because the produced oil and gas will help pay extra expenses more quickly. In some cases this may be true, for instance, when there is a shortage of petroleum products. In such a case, the optimal decision should be sought as an option with a minimum of normalized spending, provided that the duration of appraisal, or of the total exploration, is limited (i.e., less than an assigned value ($T_{app} < \alpha_1$; $T_{\exp} + T_{app} < \alpha_2$)).

The economic evaluation of the time factor takes into account the freezing of the capital investment, or loaned capital, and requires the efficient use of funds to generate new production. New production occurs only as a result of field development. For this reason, it makes sense to evaluate the cost of time for exploring and appraising a field only when the field is transferred into development immediately following the appraisal. Otherwise the increment in new production will occur over a time interval different than that used when calculating total expenditures for the discovery and appraisal of oil and gas reserves. Shortening the appraisal duration may be economically justified only for fields that will be put on-line immediately. There may be other reasons, however, for shortening this time period such as the on-time fulfillment of the reserve accrual plan.

Number of Appraisal Wells. The problem of determining a sufficient number of appraisal wells is widely discussed in the literature. In particular, it was analyzed by Knoring [17]. Most often, the necessary number of appraisal wells is determined from the stabilization diagrams of the evaluation of reserves or the calculated parameters. It is erroneously believed that this technique provides the optimum number of wells. Karpushin et al. analyzed the evolution of stabilized parameter values for the area (effective pay thickness and porosity) in some accumulations of the Dnieper-Done Basin, Central Asia, and West Siberia [38]. They indicate that the number of wells necessary for the evaluation of the stabilized values of porosity was, respectively, 3 to 10 (average, 6), 5 to 16 (average, 9), and 4 to 11 (average, 6). Respective numbers for the evaluation of effective pay thickness were 3 to 18 (7), 3 to 16 (8), and 5 to 15 (10). The values of the productive area in the basin would stabilize

after drilling, depending on the calculated version, 7 to 22 (13 to 14), 4 to 18 (9), 6 to 23 (12), and 4 to 14 (7 to 8) wells. Similar studies are likewise limited to specific examples. None offers recommendations for the average number of wells per field, which is what we are specifically interested in comparing.

It is well documented that three wells are sufficient for a preliminary appraisal of small accumulations. If the prospect delineation is not accurate, this number may increase to four or five. For medium and large accumulations the number of wells within the oil/water contact outline should be no less than four to six and no more than eight to ten. On the other hand, in order to approximately appraise the massive gas accumulations,[1] one to four wells are sufficient. Three to six additional wells are needed to appraise the stratigraphic traps due to facies changes.

Currently, recognition is given to the method of determining an optimal number of wells in the field that minimizes the total appraisal expenditures and the anticipated loss during field development [13, 17, 39]. Attempts to determine the optimum number of wells using this criterion in some fields produced the following results: 6, 7, 17, 40, and 22 [13, 39]. Average numbers, even for a certain group of fields, were not determined by the investigators.

Our technique for determining the optimal number of wells is different, although somewhat related to the aforementioned methods. There is a certain optimum evaluation of parameters, or a certain amount of information I, corresponding to minimum appraisal expenditures and anticipated development losses. This value is used implicitly in our technique: it is assumed that the number of wells N_{app} determined from the equation for (N_{app}, η_{app}) provides the amount of information I. This should satisfy the condition of minimizing appraisal expenditures and anticipated development losses. In addition, our method minimizes expenditures based not only on receiving the optimum amount of information, but also on considering a number of other factors, such as the rate of appraisal (tempo). Expenditures are minimized not for an individual field, but for a large set of fields. For this reason, this technique is more general. In a special case, when an individual field is examined and an individual factor (evaluation accuracy or the amount of information) is considered, this technique may be reduced to the method of minimizing the total appraisal expenditures and development losses. Because the formulation of the problem is different in these two techniques, it is not surprising that the results are different.

We determined that the number of wells for the appraisal of an individual field should be 10 to 12 on average. It was impossible to compare these findings with those of other investigators because they did not obtain the average optimal data.

Strategy of Exploration-Appraisal. Two opposing exploration and appraisal systems are currently applied: "creeping" and "concentrating." The "creeping" strategy concentrates operations within one or two zones (areas) after commercial accumulations have been discovered in them. The purpose of this strategy is to rapidly accrue commercial[2] reserves. Operations are moved to the other targets after a significant decline in the exploration efficiency. The "concentrated" strategy consists of conducting operations throughout the region in order to determine all major oil and gas accumulation zones (areas). The appraisal is concentrated within the richest accumulation zones, with large fields discovered at the initial stage of exploration. Less prolific zones are covered by exploration and appraisal operations as oil and gas fields in the region are developed.

If the "creeping" and "concentrated" systems are not considered, the following three idealized exploration and appraisal strategies (modified "creeping" and "concentrated" systems) can be envisioned:

1. *Consecutive strategy.* First, all of the targets are explored and then those targets found commercially productive are appraised. In other words, at the beginning, the entire drilling footage is expended for the discovery of fields (exploration) and, later, for their appraisal. Thus, the problem of efficient relationships between exploratory and appraisal drilling at any moment in time does not exist.

   At the initial stages of exploration this strategy provides sufficient, reliable information concerning the geology and petroleum potential of a region. Once the exploratory stage is complete, the largest fields may be transferred to the appraisal stage, which provides high operation efficiency. At the same time, such a strategy requires (1) a unified long-term exploration program for the entire region, (2) concentrating funds for exploration only, over an extended period of time, at the expense of an appraisal, and (3) accrual of commercial reserves and field development for the duration of such time, which can only adversely affect spending. Additionally, the utilization of rigs is inefficient because at the appraisal stage, the rigs must be returned to the exploratory stage locations, resulting in a waste of funds and time.
2. *Sequential-stepwise strategy.* The exploratory-appraisal process is subdivided into stages (steps). During each step some part of the region (or some fraction of the total number of targets in a region) is first explored with expending the total drilling footage. Afterwards, all productive targets are appraised, again with expending the total drilling footage. Once the targets are appraised, the second step begins. The next area (or group of targets) is selected repeating the same procedure. Thus, under this strategy, appraisal

is also conducted only after completing the exploration of a group of targets. In this case, however, only some of the targets within the region are explored. Therefore, this strategy has the same advantages and drawbacks as the previous strategy, although on a smaller scale.

3. *Parallel-stepwise strategy.* Using this strategy, some of the targets at each step are explored, while others proven productive after exploration at a previous step, are appraised. Thus, exploration and appraisal are conducted simultaneously, or parallel to one another. Part of the drilling footage is expended for exploration and part for appraisal.

In reality, only the parallel-stepwise strategy is implemented. The relationship between the amounts of exploration and appraisal may change with time. For instance, the extent of exploration may be greater than that of the appraisal and spread over a larger number of prospects. This is a situation not unlike that found using the "concentration" system. It has the same typical features of expending almost the entire drilling footage during the initial phase of exploration and of accelerating evaluation of potential for most of the targets. At the same time, this is not done at the expense of appraisal and commercial accrual of reserves. On the other hand, only a small number of prospects may be explored so that the degree of exploration is no greater than that of the appraisal. In such a case, this system closely mimics the "creeping" system, by concentrating operations over limited areas and accelerating appraisal after the first commercial discoveries in order to rapidly accrue the commercial reserves.

Table 9-1 shows the efficiencies of various strategies.

As previously mentioned, as the k_1 coefficient increases, the average size of reserves of an individual field decreases. Minimum spending is achieved at a minimum possible k_1 value (equal to the actual k_o). This means that only the largest discovered fields are being appraised. In other words, operations are conducted according to the "skim the cream" principle. Such a strategy requires a significant spread of exploratory operations in order to provide a large inventory of discoveries, which can then be screened to select the best ones for the appraisal. In this case the appraisal is concentrated over a small number of targets. This is reflected in Table 9-1 which shows that the number of prospects in exploration is almost twice that of fields being appraised.

From an economic viewpoint, this "concentrating" strategy is beneficial. The optimum value of coefficient k_1 is equal to its actual value k_o. Therefore, it may be decreased if even larger fields are actually discovered. This confirms that it is more profitable to conduct operations

with the purpose of creating the largest possible inventory of discoveries in order to select the best ones for preferential development. This strategy, however, only works as long as the broader spread of exploration increases the probability of large discoveries (i.e., improves operations efficiency at the expense of better selection of fields for the development). So far, this is the case.

It is not clear whether or not the "concentration" strategy is efficient from other viewpoints. Using this strategy, only 54% of the fields are developed and the rest are suspended (Table 9-1). It is also important to consider the social factors: (1) petroleum will be needed by future generations; (2) an increase in the petroleum's future value as an important chemical feedstock, rather than just as an energy source; and so forth. These factors are not considered in this criterion, which only treats economical aspects, without regard to the future. The "concentration" strategy is more efficient in terms of spending over the period of planned accrual of reserves. It is doubtful, however, that it will be efficient over the entire period of the realization of potential resources.

Number of Prospects where Drilling is Started. Table 9-1 lists the number of prospects to be evaluated annually (positively or negatively) at the exploratory level. Under the modeling conditions, it is assumed that the number of prospects annually introduced into exploratory drilling is given. Some believe, however, that these two parameters are not the same.

The number of prospects where drilling is started is important to know not only in connection with the problems we are solving, but also for planning the number of prospects (structures) delineated by seismic for exploratory drilling. Determining the number of prospects prepared (delineated) annually for exploratory drilling is based on the selection of the so-called "replenishment coefficient," which is the ratio of the number of prospects prepared (delineated) to the number of prospects transferred to the exploratory drilling stage during the same period. Optimal value of the replenishment coefficient, as well as "the sufficient number of prepared (delineated) prospects" (the ratio of the total number of delineated prospects as of year-end to the number of prospects where exploratory drilling was started during that year) are discussed in other publications. However, we could not find any literature that discusses these parameters and also the optimal number of prospects where exploratory drilling was started. The latter parameter is as important as the former two for determining the amount of exploration (geologic and geophysical) needed for prospect preparation (delineation).

It is clear that the number of prospects introduced to exploratory drilling depends on the k_1 parameter. Its inverse value may be called the "coefficient of excess" of the prospects evaluated by exploratory drilling

over the prospects (fields) where appraisal has been completed. Thus, there are some "transition coefficients" from the number of appraised prospects to the number of explored ones, and from these to the number of prepared (delineated) structures. Therefore, the optimum parameters for planning the prospect preparation (delineation) may be determined from the same concepts presented in this book. In particular, the following analogy may be suggested: The k_1 value is determined by the ability to select the best fields for priority appraisal. The value of the "replenishment coefficient" should then be determined by the ability to select the best and largest prospects for their priority exploration evaluation.

Relations between the Reserves of Different Categories. Exploration and appraisal operations should be conducted in such a manner as to provide optimum relations among the regional studies, prospect delineation, and the exploratory and appraisal drilling. It is conventional to measure the proportions of different exploration phases based on different categories of reserves. These categories may be used for planning operations at different stages. When preparing such plans, the goal is to provide a justifiable increase in the reserves of commercial categories, while optimally replenishing the higher reserve categories at the expense of the lower ones. According to Leibson:

> Practical experience indicates that . . . the exploration success . . . usually does not exceed 20 to 25%. This means that in order to discover one commercial field . . . it is necessary to plan the exploratory evaluation of 4 to 5 local structures. In terms of the reserves it means that the ratio of the C_2 reserves to the planned average annual accrual of reserves C_1 should be 8 to 10. [29, p. 14]

As shown in Table 9-1, if the operation is conducted optimally, it is necessary to explore 3.02 prospects in order to discover one field (k_1 = 0.331), and it is necessary to explore 5.62 prospects in order to discover one field that is passed for appraisal (k_1 = 0.178). Based on an analogy with the above quotation, one can find the relationship between the total C_2 reserves and the planned average annual accrual of the C_1 category reserves. On the other hand, if a prospect is found to be favorable after exploratory evaluation, its reserves are attributed to the C_2 category. At the same time, the appraised field reserves are included in the $B + C_1$ (partially C_2) categories. In other words, field appraisal may be considered the transfer of reserves from the C_2 category to the $B + C_1$ (and partially C_2) category. Optimal values of exploratory and appraisal drilling in Table 9-1 satisfy the condition of their correspondence (i.e., the number of explored targets optimally replenishes the number of appraised targets). Thus, one can conclude that the extents of operation, presented in

Table 9-1, provide the optimum replenishment of reserves in higher categories ($B + C_1$ and, partially, C_2) at the expense of lower categories (C_2).

Given the average reserves for a discovered field ($k_1 = 0.331$) and for an appraised field ($k_1 = 0.178$), it is easy to calculate the relationship between the categories C_2 and $B + C_1$ (partially C_2). The ratio is: $C_2/(B + C_1) = kM_{\text{exp}}\bar{Q}_{\text{exp}}/(M_{app}\bar{Q}_{app})$. In this particular case, $C_2/(B + C_1) = 1.22$.

Relations between Exploratory and Appraisal Drilling. Among the tasks of exploration and appraisal process control optimization is determining rational proportions of operations at different stages, in particular those between exploratory and appraisal drilling.

It was observed that these proportions do not stay constant, but, rather, change as the exploration and appraisal process progresses. As indicated earlier, the entire evolution history of exploration and appraisal in petroliferous regions may be subdivided into three periods (stages): initial, intermediate, and late. Conditions of exploration and appraisal operations change from one stage to the next, as do the proportions of exploratory and appraisal drilling. The following actual proportions are quoted by Vesnina for a number of oil and gas provinces (Yakutia, Tatarstan, Bashkortostan, Chechen-Ingushi, Samara and Tyumen oblasts, Krasnodar Kray, Azerbaydzhan, etc) [7]. At the initial exploration stage, the exploratory drilling footage is 6 to 8 times that of appraisal drilling. At the intermediate stage, when the largest fields are being appraised, the relative footage of appraisal drilling increases by a factor of 10 to 15, and that of exploratory drilling declines by a factor of 3 to 4. At the late stage, due to the introduction of numerous small prospects, the exploratory extent again increases by a factor of 3 to 4. The need to discover a great number of small fields at this stage leads to a relative 5 to 6-fold increase in the extent of exploration compared to that of appraisal.

It is not clear how close these actual ratios are to the optimum. Many believe that the current structure and distribution of the extents of exploration versus appraisal need to be reconsidered. Bunin et al. stated:

> Allocated extents of oil and gas drilling should be used in a more efficient way. Currently 48% of the total footage is assigned to the exploratory drilling and 52%, to the appraisal drilling. Such distribution is inefficient and indicates an excessive spending of appraisal drilling in discovered fields at the expense of discovering new fields. Exploratory drilling footage should be increased by transferring the task of detailed appraisal to the production drilling. [55, p. 20]

On the other hand, it is quite clear that the change in proportions must affect the system of exploration and appraisal operations. For instance, in the U.S. where exploratory drilling substantially dominates appraisal

drilling, exploration (with a single well) is simultaneously conducted at a large number of new prospects. This provides an opportunity to rapidly explore significant areas of the potential regions (areas of sedimentary basins). However, in the former USSR, following the first commercial discoveries, attention was shifted to appraisal. This has caused difficulties fulfilling the plans of commercial accrual of reserves in future years. Thus, the problem arises of rational (optimal) relations between the exploration and appraisal footage.

The following basic relations between the parameters of exploration and appraisal drilling may be derived based on the discussion in Chapter 8. The following equation is obtained from Equations 8-4 and 8-5:

$$\frac{H_{exp}}{H_{app}} = \frac{\overline{M}_{exp} N_{exp}}{\overline{M}_{app} N_{app}} = \frac{1}{k_1} \frac{N_{exp}}{N_{app}}$$

The ratio of exploration to appraisal drilling footage is proportional to the ratio of exploratory to appraisal wells used in one field. The portions of exploratory and appraisal drilling are determined from the following equations:

$$V_{exp} = \frac{H_{exp}}{H_{exp} + H_{app}} = \frac{N_{exp}}{N_{exp} + k_1 N_{app}}$$

$$V_{app} = \frac{H_{app}}{H_{exp} + H_{app}} = \frac{k_1 N_{app}}{N_{exp} + k_1 N_{app}}$$

where $V_{exp} + V_{app} = 1$.

The relations for the other exploration and appraisal drilling parameters may be derived from Equations 8-1 through 8-8.

The number of rigs used for exploration and appraisal drilling is:

$$\frac{n_{exp}}{n_{app}} = \frac{\overline{M}_{exp} N_{exp} t}{\overline{M}_{app} N_{app} t} = \frac{H_{exp}}{H_{app}}$$

Annual expenditures for exploration and appraisal drilling are:

$$z_{exp}/z_{app} = H_{exp}/H_{app}$$

Cumulative footage (from the beginning of operations) in the exploratory wells at the prospects evaluated by exploration in a current year, and appraisal in the fields appraised in a current year, is:

$$\frac{L_{\text{exp}}}{L_{app}} = \frac{\overline{M}_{\text{exp}} N_{\text{exp}} D}{\overline{M}_{app} N_{app} D} = \frac{H_{\text{exp}}}{H_{app}}$$

Cumulative footage (from the beginning of operations) in the exploratory and appraisal wells in an individual field completed by appraisal is:

$$\frac{L'_{\text{exp}}}{L'_{app}} = \frac{N_{\text{exp}} D}{N_{app} D} = \frac{N_{\text{exp}}}{N_{app}} = k_1 \frac{H_{\text{exp}}}{H_{app}}$$

The number of prospects in exploratory drilling and the number of fields in appraisal drilling are:

$$\frac{P_{\text{exp}}}{P_{app}} = \frac{1}{k_1} \frac{N_{\text{exp}} t}{\eta_{\text{exp}}} \frac{\eta_{app}}{N_{app} t} = \frac{\eta_{app}}{\eta_{\text{exp}}} \frac{H_{\text{exp}}}{H_{app}}$$

Thus, the relations between the different parameters of exploratory and appraisal drilling are in some way expressed through the ratio of exploratory to appraisal footage (H_{exp}/H_{app}). This, in turn, depends on the coefficient k_1 and the number of drilled exploratory and appraisal wells (N_{exp}, N_{app}) (i.e., in the final analysis, on the control parameters). This explains the empiric patterns mentioned earlier. At the initial and late stages of the exploration and appraisal process the value of k_1 coefficient is low and, hence, the H_{exp}/H_{app} ratio is high. At the intermediate stage, the value of k_1 increases and, due to the appraisal of the largest fields, the value of N_{app} also increases, which results in a significant decline in the H_{exp}/H_{app} ratio.

Knowing the optimal values of the control parameters, from the above equations it is easy to determine optimum relations between exploratory and appraisal drilling (Table 9-2). The table indicates that the structure of the drilling volume distribution between exploratory and appraisal drilling should be reconsidered. The exploratory footage must be increased.

Generally speaking, the issue is not the drilling footage per se, but its distribution between the targets. Certainly, exploratory drilling should proceed ahead of appraisal drilling; however, this should be expressed, not in the amount of footage, but in the number of targets drilled. Table 9-2 shows that the number of prospects in exploratory drilling is greater by a factor of 1.76 than the number of fields in appraisal drilling. Even greater is the respective difference between the number of targets completed annually

Table 9-2
Optimum Relations between Parameters of Exploratory and Appraisal Drilling

Parameters	Parameter Ratio		Share of the Total	
	Exploratory/ Appraisal	Appraisal/ Exploratory	Exploratory/ (Exploratory+ Appraisal)	Appraisal (Exploratory+ Appraisal)
Annual drilling extent during current year (H_{exp}, H_{app})	1.14	0.88	0.53	0.47
Drilling extent (cumulative from the beginning of operations) at prospects and in fields completed by drilling during current year (L_{exp}, L_{app})	1.14	0.88	0.53	0.47
Drilling extent (cumulative from the beginning of operations) in an individual field completed by drilling during current year (L'_{exp}, L'_{app})	0.2	4.94	0.17	0.83
Number of prospects and fields in drilling (P_{exp}, P_{app})	1.76	0.57	0.64	0.36
Number of prospects and fields completed by drilling during current year ($\overline{M}_{exp}$, $\overline{M}_{app}$)	5.62	0.178	0.85	0.15
Total number of wells used per individual field from the beginning of operations (N_{exp}, N_{app})	0.2	4.94	0.17	0.83
Duration of drilling evaluation for an individual target (T_{exp}, T_{app})	0.31	3.19	0.24	0.76
Drilling rate (tempo): number of rigs simultaneously operating at an individual target (η_{exp}, η_{app})	0.65	1.55	0.39	0.61
Total number of rigs (n_{exp}, n_{app})	1.14	0.88	0.53	0.47
Current drilling expenditures (z_{exp}, z_{app})	1.14	0.88	0.53	0.47

by exploratory and appraisal drilling (the former is greater than the latter by a factor of 5.62). In other words, 85% of all annually completed targets must be prospects completed by exploratory drilling and 15% of the fields must be completed by appraisal drilling. Compare this with the number of targets in exploratory drilling (64% of the total number) and with the exploratory footage (53% of the total footage).

Spreading the exploratory footage among a greater number of targets is accomplished by simultaneously using 1.55 times fewer rigs at each exploratory target than at each appraisal target (field). Correspondingly, 53% of rigs are conducting exploratory drilling and 47% are conducting appraisal drilling.

In an individual field, after exploratory and appraisal drilling is completed, the appraisal footage and the number of appraisal wells are almost 5 times higher than the exploratory footage and number of wells. If, however, all targets completed annually by each type of drilling are considered, the total exploratory footage (since the beginning of operations) is 1.14 times higher than the total appraisal drilling.

Optimal time distribution is such that the duration of appraisal is 3.19 times higher than that of the exploratory evaluation for an individual prospect (76% of the total time spent for exploratory and appraisal drilling).

The optimal ratio of exploratory to appraisal drilling expenditures is 1.14:1.00. In other words, expenditures for exploratory drilling account for 53% of the total exploratory and appraisal drilling expenditures. This may seem surprising since it is believed that appraisal drilling requires higher expenditures. In reality, only 17.8% of all explored targets are transferred to appraisal. In those fields where appraisal operations have been completed, the appraisal footage is 4.94 times higher than the exploratory footage. In other words, in developed fields expenditures for appraisal drilling are 83% and 17% for exploratory drilling.

Notes

1. Massive accumulations are those with bottom water underneath the entire area of the accumulation. Most often they are associated with unbedded rocks (e.g., reefs).
2. "Commercial" reserves is a rather general term meaning appraised reserves (Russian categories $A + B + C_1$), corresponding to the proved reserves and a more reliable portion of the probable reserves. Preliminarily evaluated reserves (Russian category C_2) correspond to the less reliable portion of the probable reserves, and the more reliable portion of the possible reserves. The largest part of the next category ($C_3 + D_o$) corresponds to the less reliable portion of possible reserves and is classified as resources, not reserves.

CHAPTER 10

Forecast of Exploration Evolution

It is very important to predict the future evolution of the exploration-appraisal process and its optimal parameters. This is a precondition for exploration planning and determining optimal funding.

In order to forecast the exploration-apraisal process, it is necessary to forecast the evolution of the factors that determine the course of the process. This is a very difficult task. Several different approaches to solving this problem are analyzed below.

Forecast of Exploration Evolution Based on Changes in the Average Size of Discoveries

The first approach takes into account the average size of a field, which is an important factor.

Many believe that mostly small- and medium-size fields will be discovered in the future, and they expect that the average size of these fields will continue to decrease. Earlier it was shown that this decrease will result in an increased k_1 coefficient. This coefficient was used in forecasting.

Thus, the field reserves were used in forecasting the evolution of the exploration-appraisal process, with all other conditions equal. Optimum parameters of the exploration and appraisal process were determined for different values of the k_1 coefficient with (1) constant planned accrual of reserves Q, (2) duration of well construction (including testing) t, (3) well cost c, (4) field discovery success rate k, and (5) relations among N_{exp}, η_{exp}, N_{app}, and η_{app} and the parameters describing them. The results are listed in Table 10-1.

It is important to note that parameters of the exploration-appraisal process for a particular target do not change. This pertains to both the

exploration and the appraisal process. These parameters include the number of rigs working simultaneously at one prospect η_{exp} (in one field η_{app}), the number of wells used for the exploration of one prospect N_{exp} (for appraisal of one field N_{app}), and the duration of exploratory (appraisal) drilling at one prospect T_{exp} (in one field T_{app}). Optimum values of these parameters are, on the average, constant regardless of the size of field reserves.

This is a very interesting phenomenon. It might be expected that with a decrease in the size of reserves the number of wells, especially appraisal wells N_{app}, would decline. Apparently, the analyzed range in the variation of the size of reserves is not wide enough to affect the final number of wells. On the other hand, the operations rate (tempo) is analyzed and studied for field sets, not for a single field. Therefore, expenditures are affected not only by the operations methodology in one field, but also by the distribution and re-distribution of exploration and appraisal extents between the targets.

However, parameters characterizing the number of targets change significantly. The general trend is a decrease in exploratory operations and an increase in appraisal operations. This results in a decrease in the number of prospects completed annually by exploration ($\overline{M}_{exp}$) and the number of prospects in exploratory drilling (P_{exp}), and an increase in the number of fields appraised annually ($\overline{M}_{app}$) and the number of fields in appraisal drilling (P_{app}). The number of operating rigs changes, respectively. On the whole, the ratio of exploratory to appraisal drilling footage changes in favor of the latter. Although this shift toward appraisal drilling is unexpected, it occurs because there is no longer enough "cream" to fulfill the reserves accrual plan. Therefore, especially because of the decrease in reserves, it is necessary to appraise a higher percentage of the discovered fields. There is no longer a need to suspend a substantial portion of the discoveries, and as a result, the number of exploration targets declines and, correspondingly, the number of appraisal targets increases. Thus, exploratory operations are concentrated over a smaller number of targets, whereas appraisal operations are distributed over a greater number of targets, covering those that are not currently transferred into appraisal. We then begin to view suspended discoveries as fields. In effect, the field discovery success rate increases with decreasing exploration operations.

The next step is to examine how expenditures change. Appraisal of a field is much more expensive than its exploration; therefore, it is reasonable to expect an increase in future expenditures (see Table 10-1). For example, if currently 54% of the discoveries are transferred to appraisal, then in the future this number will increase. Correspondingly, if some targets are abandoned after exploratory drilling, then in the future

Table 10-1
Optimal Parameters of the Exploration-Appraisal Process for Different Values of k_1

	Coefficient k_1								
Parameters	**0.17**	**0.19**	**0.21**	**0.23**	**0.25**	**0.27**	**0.29**	**0.31**	**0.331**
Average reserve size for an individual field (% of the average reserve size in an individual field at the actual k_o value equal 0.178)	103	96	90	85	80	75	72	69	66
Fraction of fields transferred to appraisal	0.51	0.57	0.63	0.69	0.76	0.82	0.88	0.94	1
Number of rigs simultaneously operating at a single prospect (exploratory, η_{exp})	1	1	1	1	1	1	1	1	1
Number of wells used for exploration of a single prospect, N_{exp})	2.05	2.05	2.05	2.05	2.05	2.05	2.05	2.05	2.05
Duration of a single prospect exploratory evaluation, T_{exp}	2.97	2.97	2.97	2.97	2.97	2.97	2.97	2.97	2.97
Number of fields annually completed by exploration, $\overline{M}_{exp}$	236.28	226.72	218.95	212.35	206.72	201.9	197.63	193.79	190.33
Number of prospects in exploratory drilling, P_{exp}	701.75	673.36	650.28	630.68	613.96	599.64	586.96	575.65	565.28

Number of rigs conducting exploratory drilling, η_{exp}	701.75	673.36	650.28	630.68	613.96	599.64	586.96	575.65	565.28
Number of rigs simultaneously operating in one field being appraised, η_{app}	1.55	1.55	1.55	1.55	1.55	1.55	1.55	1.55	1.55
Number of wells used for appraisal of a single field, N_{app}	10.12	10.12	10.12	10.12	10.12	10.12	10.12	10.12	10.12
Duration of a single field appraisal, T_{app}	9.47	9.47	9.47	9.47	9.47	9.47	9.47	9.47	9.47
Number of fields completed annually by appraisal, $\overline{M}_{app}$	40.17	43.07	45.98	48.84	51.68	54.51	57.31	60.07	63
Number of fields in appraisal drilling, P_{app}	380.59	408.07	435.64	462.74	489.64	516.46	542.98	569.13	596.89
Number of rigs operating in appraisal drilling, η_{app}	589.92	632.5	675.24	717.24	758.85	800.51	841.63	882.16	925.19
Minimum nozmalized expenditures (% of optimal expenditures at actual $k_o = 0.178$)	99.6	100.7	102.2	103.9	105.8	107.9	110.1	112.4	114.9

they will have to be evaluated by appraisal drilling. Expenses will increase because a number of appraisal wells in a single field is five times that of exploratory wells. Thus, redistribution of relationships between exploratory and appraisal drilling will result in the exploration and appraisal process becoming more expensive.

There is a good reason for the continuous increase in exploration and appraisal expenditures, and the reason is well documented. It is caused not only by such factors as increased drilling depth and more complex field structure, but also by the redistribution between exploratory and appraisal drilling due to a decrease in the size of fields (average size of reserves).

Forecast of Exploration Based on Trends in Changes of the Parameters of Empiric Equations

Forecasting based on the trends in parameter changes is very common. It must be remembered, however, that the existing exploration control system for various reasons is most often not optimal. Therefore, if we directly use the observed trends for a forecast, the parameters will substantially deviate from their optimal values. Efforts in exploration control are directed toward making these deviations minimal, which should be taken into consideration in forecasting. If forecasting takes into consideration the optimality of control actions over the exploration process, then the actual trends in parameter changes are not the primary interest. Instead, the focus turns to the changes in relationships between parameters and how they determine the optimum solution.

Using the constructed model of the exploration and appraisal process (Equations 8-1 through 8-8), it is possible to determine the optimal parameter values of the process at a given point in time provided that the equation parameters and a number of technological parameters are known at that moment (obviously, technological parameters are changing). Thus, the evolution of the parameters of the exploration-appraisal process, with a correction for optimality, is determined by the evolution of the parameters for these equations and technological parameters. This should be studied.

The major parameters of the exploration-appraisal process are the number of exploratory wells (N_{exp}) and appraisal wells (N_{app}) completed with the purpose of exploring and appraising a single field (local operations target), depending on its size (reserves) and the k_1 coefficient. The number of wells is a local characteristic of the operations strategy and k_1 is a regional characteristic. The evolution of the number of exploratory and appraisal wells and of the success rate was thoroughly studied. This

study was applied, in particular, to forecasting the number of wells. This study, however, dealt with the actual observed trends. It is clear, therefore, that the most important relationships in the system of parameters of the exploration-appraisal operation are those presented by Equations 8-6 and 8-8. They reflect empirical relations that occur as a result of a certain exploration-appraisal operation. This is why certain empirical trends, requiring study using the actual data, may be applied only to the parameters of these equations. Only these empirical trends should be taken into account while forecasting. In the following section, we will discuss only these empirical trends, as trends of economic and technological parameters are beyond the scope of this book.

Table 10-2 lists the values of $N^{(1)}$ and a parameters in Equation 8-8 for exploratory (with the goal of field discovery) and appraisal drilling during certain time intervals in the former USSR as a whole. The most pronounced trend is in the changes of the $N^{(1)}_{exp}$ parameter. Its value consistently declines: from 2.7 to 3 during the first time intervals t_1, t_2, to 2.0 to 2.2 and, eventually, to 1.6 to 1.7 (and even lower) during the t_{11}–t_{14} time intervals. In other words, the number of dependent, sequentially started exploratory wells per field consistently decreases. The a_{exp} parameter also decreases, but much less than $N^{(1)}_{exp}$. As a result, the $a_{exp}/N^{(1)}_{exp}$ ratio does not decrease (i.e., the selection of locations of several simultaneously started wells does not change).

It is reasonable to maintain that changes in $a_{exp}/N^{(1)}_{exp}$, a_{exp}, and $N^{(1)}_{exp}$ are a quantitative reflection of the following trend in exploratory drilling strategy: prospects where exploratory drilling is initiated are more often drilled using a single rig. This resulted in a decrease in the specific (per unit) number of exploratory wells which, in turn, led to improved efficiency. Thus, the progress in strategy is associated with the introduction of the "single-well technique" [31].

The selection of locations for several simultaneously started wells (when the "single well technique" is not followed) also somewhat improved, which led to a decrease in the a_{exp} parameter.

In appraisal drilling, there is no noticeable decrease in the values of $N^{(1)}_{app}$ and a_{app} parameters. They remain, respectively, approximately 7 to 8 and 3 to 4. This means that when switching to a sequential drilling system, or better locations for simultaneously drilled wells, there is no decrease in the number of wells. This result is normal since locating wells dependently (sequentially) would result in a substantial increase in the appraisal duration. During appraisal, several wells are started simultaneously and, as a result, the $a_{app}/N^{(1)}_{app}$ ratio is lower than the $a_{exp}/N^{(1)}_{exp}$ ratio. In general, appraisal drilling possesses a greater inertia than exploratory drilling and the progress is slower. Evolution of the a_{exp}, $N^{(1)}_{exp}$, and a_{app}, $N^{(1)}_{app}$ parameters is different.

Table 10-2
Changes in Parameters of Equation 8-8, Demonstrating the Correlation between the Number of Wells and the Number of Simultaneously Started Wells

Time Intervals	Exploratory Stage			Appraisal Stage		
	$N_{exp}^{(1)}$	a_{exp}	$a_{exp}/N_{exp}^{(1)}$	$N_{app}^{(1)}$	a_{app}	$a_{exp}/N_{app}^{(1)}$
t_1	2.68	0.78	0.29			
t_2	3	0.6	0.2			
t_3	2.1	1.9	0.9			
t_4	1.9	1.2	0.63			
t_1–t_5	2.21	1.18	0.53			
t_6–t_8	1.9	1.77	0.93			
t_8	2.64	0.73	0.28			
t_9	2.05	1.13	0.53	7.83	4.07	0.52
t_{10}	2.49	1.39	0.56	8.91	1.49	0.17
t_6–t_{10}	2.05	1.73	0.84	—	—	—
t_8–t_{10}	—	—	—	8.22	3.82	0.46
t_{11}	1.9	0.49	0.26	14.83	0.45	0.03
t_{12}	1.72	1.33	0.77	8.03	4.22	0.53
t_{13}	1.41	1.01	0.72	7.49	2.53	0.34
t_{14}	1.6	1.13	0.71	3.35	3.2	0.96
t_{15}	1.52	0.62	0.41	—	—	—

These results cover a short time interval and, therefore, do not provide an opportunity to reliably extrapolate the noted trend. What is important, however, is that the parameters of Equation 8-8 very slowly decrease with time. This is especially true for appraisal drilling.

The parameters of Equation 8-6 determine the specific format of the $\overline{Q}(k_1)$ function. They also may change in time for various reasons. As they change, the graphic representation of the $\overline{Q}(k_1)$ function (type of the curve) will also change because the parameters of Equation 8-6 determine the rate at which it is declining.

As indicated earlier, the solution to the problem of forecasting the exploratory-appraisal process, taking optimization into consideration, is associated with forecasting the values of equation parameters at any moment in time based on their evolution. In this case, it is the evolution

of parameters α, β, γ, and ε and the corresponding shape of curve 8-6. Thus, it is important to determine the causes of these evolutions.

From t_8–t_{15} data we calculated the parameters of Equation 8-6 and correspondingly plotted the $\overline{Q}(k_1)$ graphs shown in Figure 10-1. The corresponding curves are situated one under the other. The upper curves exhibit a steeper decline than the lower curves. This trend is quite clear. In other words, the greater the values of the average actual reserves $\overline{Q}(k_o)$ of one field appraised for development, the steeper the curve, and the more rapidly it becomes flat. Inversely, flatter curves correspond to smaller $\overline{Q}(k_o)$. This indicates that when the average value of the actual reserve in one field appraised for development is low, all fields selected for this purpose are relatively uniform. They are not much different in terms of their reserves from the suspended fields. Therefore, a field selection in this case is difficult. Conversely, when the average actual reserve value of one field appraised for development is high, all selected fields are significantly non-uniform in terms of their reserves. For this reason, even if the largest suspended fields are appraised and ready for development, the average size of reserves of these fields will noticeably decrease. It is not difficult to select fields with a large $\overline{Q}(k_o)$ value: the largest discoveries are prepared for development.

Thus, it may be stated that the α, β, γ, and ε parameters in Equation 8-6 are related to the reserves of discovered fields. Their evolution is determined by that of the discoveries. The evolution of discoveries in most areas is such that the size of the discovered reserves tends to decrease with time. The shape of the curves corresponding to Equation 8-6 in these regions will change regularly with time: steep curves will be gradually replaced by flatter curves. Therefore, forecasting the evolution of the α, β, γ, and ε parameters in Equation 8-6, or the appearance of its corresponding curve, may be based on the forecast of the evolution for the average size of discoveries. Each average discovery size will have a corresponding curve and a certain set of α, β, γ, and ε parameters. Plotting the parameter averages against time will set up the scenario for the evolution of parameters.

We previously analyzed the forecast problem considering the optimization of the exploration-appraisal process control. In this case, the actual evolutionary trends of certain parameters were not important, but, rather, the trends of changing relations among the parameters, which determine the optimal solution. Most important for the exploration-appraisal process parameters are functions 8-6 and 8-8. We thereby analyzed evolutionary trends for these equation parameters. Equation 8-8 parameters evolve in time depending on the progress in well placement, decreasing very slowly, especially for appraisal wells. Equation 8-6 parameters evolve parallel to a decrease in the size of the reserves of discoveries. Identification and

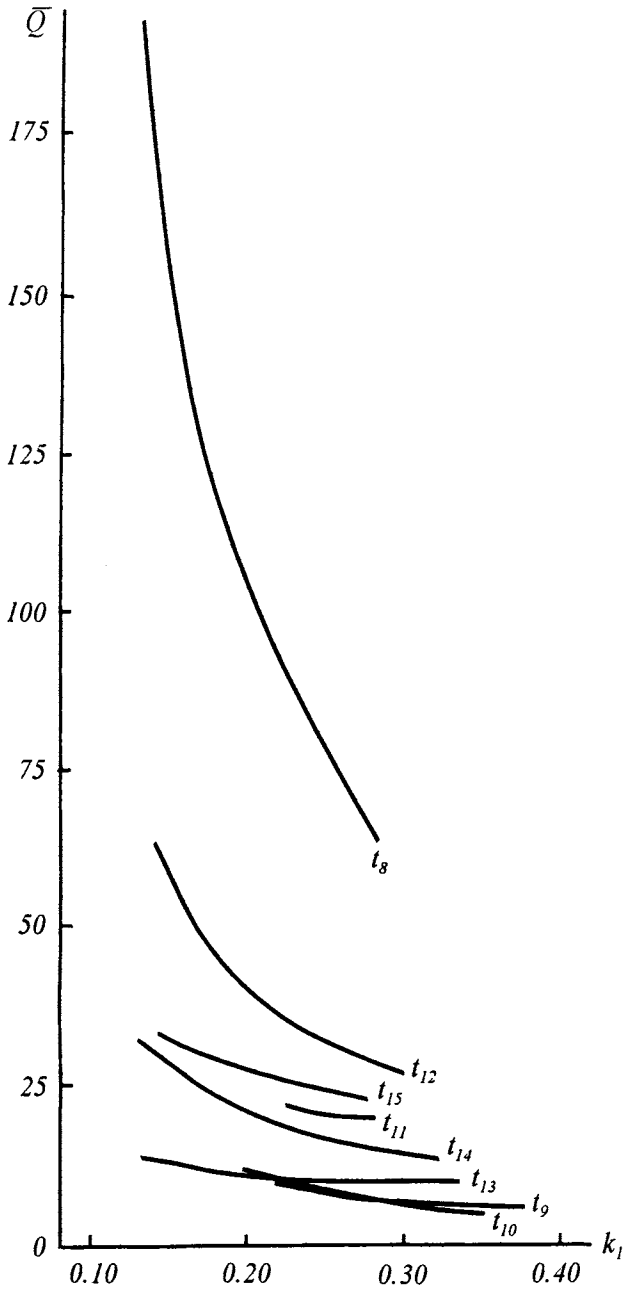


Figure 10-1. Average size of reserves of a field transferred to appraisal ($\overline{Q}$) versus k_1 at different time intervals.

extrapolation of these trends provide a tool for forecasting the optimal number of exploratory and appraisal wells at local targets belonging to different classes. They also provide a tool for forecasting the optimal number of fields transferred to appraisal and then appraised for development. In the final analysis, this results in the optimal solution of the purpose of the exploration-appraisal process.

Alternate Avenues of the Evolution of Oil and Gas Exploration Using the System Approach

The oil and gas exploration-appraisal process is a controlled system. Control efficiency significantly depends on whether separate elements of the system are considered independent (autonomous) control objects or the system principle predominates. With this in mind, we will examine the following situation.

A general structure of the system under consideration is presented in Figure 10-2. The targets that may become oil and gas fields are "run" through the system. Exploration and appraisal targets identified using various techniques, are fed into the input of the system. This is the main, but not the only, source of objects fed into the system. Sometimes, exploratory drilling is conducted on improperly or insufficiently delineated targets (e.g., after drilling a parametric well). If only the exploration-appraisal process is considered, the output of the system will be the appraised fields or fields ready for development. This output may also be measured by the commercial volume of reserves. Non-productive targets are taken out of the system and all other objects remain in the system. These objects either proceed consecutively through the main flow or remain as a reserve ("production waste" at each stage).

We represent the system as a screening system comprised of a chain of filters. This is the purpose of different exploratory and appraisal stages. The completed inventory at a certain stage is a source of replenishment for the completed inventory at each subsequent stage. Only those targets screened through certain criteria and requirements may enter the subsequent stage. The rest of the targets remain in the inventory and, sooner or later, they will enter the subsequent stage. Target selection for the inventory of targets completed at the subsequent stage is performed from the like inventory at the preceding stage. For some time the selected targets remain in the current inventory. Thus, each subsequent link in this chain is fed by the preceding link. Corresponding selection speed represents the annual number of targets processed through the corresponding stage.

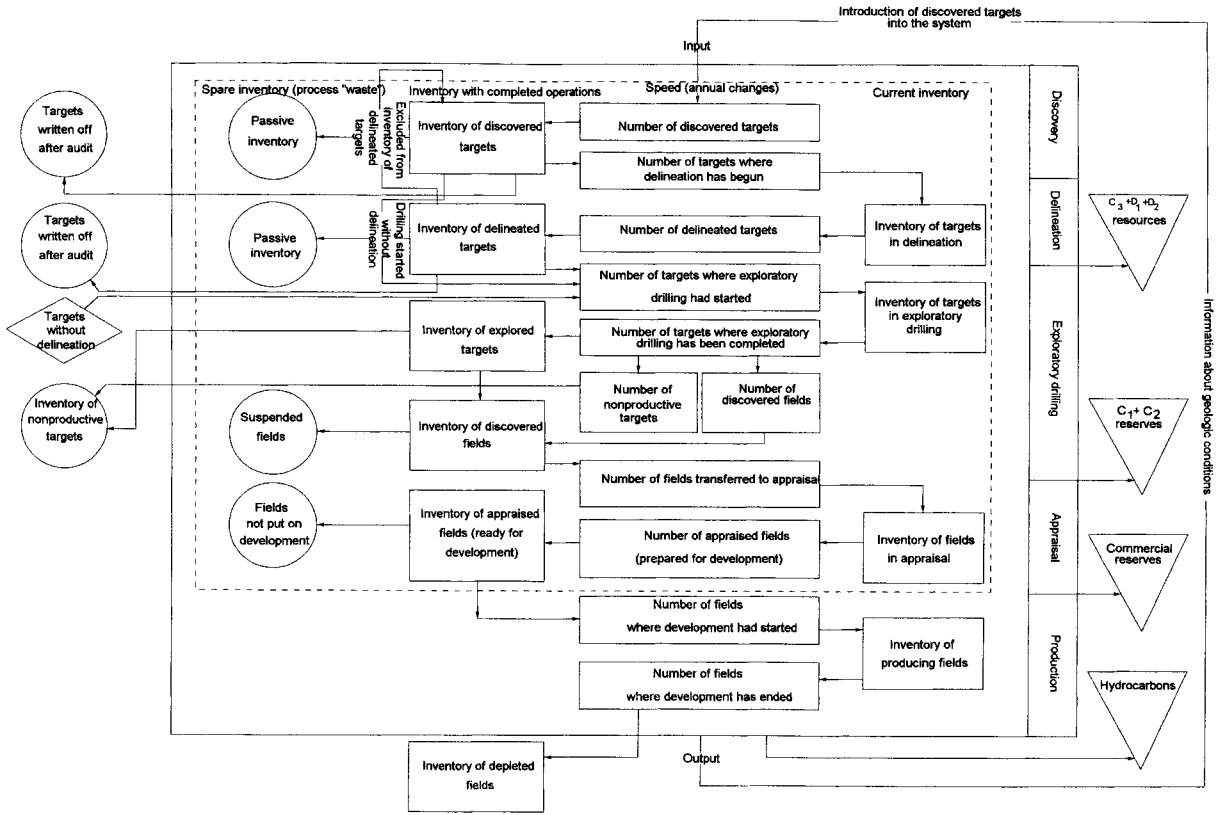


Figure 10-2. Resource development as a unified control system.

Consecutive selection of drilling targets by stage is not the only task accomplished in the system. Information is also accumulated, primarily concerning the targets in the system. Additionally, information is synthesized concerning the environment in which the system is operating—the geologic conditions, undiscovered potential and its size and structure, location of forecast fields, and so forth. Information concerning the environment may also be considered one of the system's outputs. C_2 reserves and C_3 resources may be considered intermediate outputs of the system.

The tasks of the described system are quite complex. A great deal depends on external factors related to the system; namely, which targets are input into the system. The system is conducting a selection among input targets and is incapable of replacing them. This is why, in the interest of efficiency, it is very important to correctly develop the strategies associated with selecting the richest and most prospective areas for conducting exploration and appraisal. As we can see, these conditions of general exploration and appraisal strategy concerning the optimum distribution of operations amounts in different plays, are external with respect to this system. However, their development must be based on information, and, in particular, information generated within the system. As already indicated, generation of such information is one of the system's tasks. In this sense, it defines its orientation within the domain of strategic solutions. The first priority targets, in a sense the best-of-the-best, should enter the system. On the whole, the system we are analyzing may be considered a small subsystem or part of a large, closed system. This large system includes the exploration and appraisal strategy, regional studies, and organizational methodology of the process, whereby the targets are discovered for further entrance into the small subsystem.

In connection with the efficiency of exploration and appraisal control, it is clear the small subsystem operation should be organized so that the best targets are the first to be output. They must be given preferential treatment. On the other hand, those targets of no immediate interest should be screened-out as early as possible and not fed into the system at all. A general requirement of the system as a whole is to achieve the desired output with the least possible expenditures. From this viewpoint the system is incomplete. It would be acceptable to lower the overall exploration and appraisal efficiency if this negative offset is compensated by a consequent gain incurred during field development. The accrual of reserves is not the only goal. For this reason, we included the field development stage in the design, correspondingly changing the system's output (Figure 10-2).

If the efficiency of the entire large system is considered, the result at each individual stage is not independently significant. What is important

is the final result. From this perspective, an increase in the per-unit costs, for instance, for structure delineation or exploratory drilling, should not be viewed as a decline in efficiency. There must be a subordination of interests at each individual stage to the common goal (i.e., a decrease in the per-unit cost at the output, or for the final product). From this perspective, the control efficiency problem is reduced to determining the rational balance between the links of a single chain in terms of: (1) the number of targets in corresponding inventories (a reasonable amount of backlog providing a wide, but necessary, target selection); (2) the removal rate and replenishment of inventories; (3) the amount of expenditures, and (4) the formulation of tasks. Only after this, can efficient operations at each stage be discussed. Many analyze efficiency parameters at each stage as if this were an isolated process. These parameters are often planned parameters used to measure performance, which does not create incentives for efficient control of the entire system. Narrow interests associated with operations dominate at each stage. To alleviate this situation, some consider planning not only the accrual of commercial reserves, but also of C_2 reserves and C_3 resources [27, 37].

Obviously, the above discussed relations cannot remain constant indefinitely. The system must adjust to a changing external environment. The following characteristic feature will determine the adaptation of the control system to the external environment. This feature may be called "branching of the targets (objects) in the system." When the targets are discovered and delineated for exploratory drilling, each target represents a whole unit; however, it may be delineated at different stratigraphic horizons. When exploratory drilling is started, this unity may remain intact, although there may be two (or more) levels of exploration due to a significant depth difference. Appraisal levels are identified during the appraisal phase, and each level is appraised using its own well grid. These are now independent targets. Appraisal may be completed at one level, whereas it may be only beginning for another one. When the field is being developed, there may be several production targets within a single appraisal level and they may be put on-line at different times.

The system under consideration is sufficiently complex that it is not always easy to find rational relations between its segments, even in a stationary environment. Our analysis concluded that the problem of optimum control was not solved for the entire multi-link chain, but only for two adjacent links. One link was represented by the field search through exploratory drilling and the other by their appraisal. Switching from a complex multi-link system to a two-link system is a frequent phenomenon. For instance, when analyzing the exploration results, the question always arises whether or not there are enough delineated structures to support the exploratory drilling footage. There are publications

that discuss the desirable ratios of different categories of oil and gas reserves and a number of others that address the reserve-to-production ratio [49, 14, 28]. In many ways these issues are associated with this system, but they do not offer any solutions to the problem. Additionally, considering the stationary state of the system, they do not answer the questions of choice and rational inventory (backlog). These problems are usually solved by using the product output from a previous stage at a subsequent stage.

A systemic study of the exploration-appraisal process (with the addition of the development stage) is needed for the solution of efficient control of this process. This approach may be advantageous. Currently, it is not possible to describe the entire system. We believe that this task should involve finding a system organization under which the system would be in a state of equilibrium, and then studying it using the thermodynamic methods usually applied to analyzing the equilibrium distribution of molecules in gas. This should be kept in mind when searching for the required ratios and evaluating their variation with changing external conditions. Even now, certain trends may be discerned that could possibly offer a key to forecasting the exploration-appraisal process. These trends are discussed below.

The major observed trend, in response to external conditions, is a steady decline in the size of the discovered reserves. This also indicates a constant deterioration of the geologic scenario for the oil and gas field exploration and appraisal. This trend is especially conspicuous in highly mature areas, where (1) mostly small fields are discovered, (2) the depth of wells increases, and (3) operations are conducted in increasingly complex geological settings.

The trends occurring in the exploration-appraisal process reflect the system's adjustment to these changing external conditions, and, thus, the evolution of the exploration-appraisal process. They are caused only by the deteriorating geologic scenario. In this sense, the trends related to different aspects of the process are interrelated because they have a common cause. This cause generates the primary change in the system from which all other changes are derived. This major change results in the redistribution of expenditures (as applied to a single field) and their transfer from the late stages to the initial stages. For example, major fields are characterized by (1) a high cost of development, (2) a much smaller, but still significant, cost of appraisal, and (3) a comparatively low cost of discovery and delineation of structures and exploratory drilling. The cost distribution among the stages is non-uniform and is concentrated at the latest stages of field development. In large measure, they are associated with the branching operations mentioned previously. This effect is much more typical for large fields, where development accounts for

96.9% of the total expenditures, exploratory and appraisal drilling for about 3%, and discovery and delineation for only 0.1%.

With the passage of time, ever smaller targets are supplied as the system input. In some regions, these are only small targets and expenditures within the system tend to level off. For example, for 1 to 5 MMT fields the discovery and delineation cost amounts to 1%, the exploratory and appraisal drilling cost for 10%, and the development cost for 89%. The respective numbers for 0.1 to 1 MMT fields are 11%, 21%, and 68%. Levelling of the cost is significantly affected by the increased load at the initial stages due to deterioration of conditions. The portion of targets transferred to subsequent stages drastically decreases, the discovery success rate drops, and the ratio of fields transferred to development to those appraised drops. For example, in the Timan-Pechora province (Russia) the fraction of appraised fields transferred to development for long intervals of time (1929–1945, 1946–1975, 1976–1980, and 1981–1985) decreased as follows: 0.5, 0.3, 0.16, and 0.06. Thus, the waste within the system is growing. This waste is especially significant at the earlier stages.

At the same time, the tasks of each stage change and boundaries within the system become obscurred. For instance, the appraisal tasks partially shift to the development stage: some appraisal wells are replaced by advanced production wells. In small fields, one or two wells drilled after the discovery well may be able to completely finalize the appraisal [51]. Thus, the stages of estimation and appraisal of reserves for development may practically coincide. Simultaneously, the field knowledge requirements of the appraisal stage are lowered, and the domain of appraisal becomes narrower: development "pushes" the appraisal. The appraisal ends with the evaluation of C_1 reserves, whereas when mostly large fields were being appraised, the requirements were higher. These trends in requirement changes and their dependence on the category of reserves are well studied. The ratio of exploratory-to-appraisal wells also changes in favor of exploratory wells because exploration in small fields is rarely completed by one or two wells, whereas only a few wells may be sufficient to appraise such a field.

Operation techniques change during the exploration stage, and exploration using a single well is becoming conventional. Although there is a risk of missing a field, this risk is justified due to the small size of the field, whereas it is not justified when searching for large fields. As a result, the average number of per-structure exploratory wells is gradually decreasing. For instance, in the Volga-Ural Province in Russia this number declined from 3 or 4 wells during the 1966–1970 period to 1.5 to 2 wells during the 1976–1980 [27]. The boundaries of the exploratory stage also become obscured which, in turn, "push" the appraisal. This is reflected in the requirement changes for different categories of reserves. In this

case, the C_2 reserves are evaluated at the exploratory stage. It appears that the appraisal stage is "pulled apart" and absorbed by the exploration and development stages, which become "closer" to one another. It is possible that the boundary between the exploratory and delineation stages will also change. If direct techniques of hydrocarbon detection are perfected and become commercially applicable, the separation of fields from dry structures will be achieved not through exploratory drilling, but through geophysical techniques at a stage similar to a delineation stage but with indistinct boundaries.

Therefore, the general trend is a shift in loading and expenditures from the final stages to the earlier stages. For example, in mature producing regions of Russia the amount of drilling declined and the budgets for the structure discovery and delineation increased. Their share of the total expenditures for exploration and appraisal during the 1966–1970 to 1976–1980 periods increased from 16% to 27% and from 25% to 40% in the Volga-Ural Province alone. In Tatarstan, the cost of structure delineation during this same period was almost equal to the cost of exploratory and appraisal drilling [27]. This was caused by economics, namely, the process "waste" became too large. Cost efficiency in small fields drastically decreased and attempts were made to replace drilling with less-expensive appraisal techniques. This, in turn, increased the demand for geophysical surveys. These are now assigned not only the delineation task but also the screening task (direct methods). At the appraisal and development stages, it is also necessary to obtain the information through geophysical techniques in combination with the geologic analysis of wells. This causes the redistribution of expenditures: their outflow from the later stages and inflow to the earlier stages. These trends are important for further improvement of the exploration-appraisal efficiency.

Currently, the oil and gas exploration-appraisal process is at a crossroads. Future techniques and the efficiency of operations will depend on which road turns out to be more successful and on the progress in geophysical methods.

Currently, the technological challenges in the field of geophysics are: (1) depth conversion, (2) non-linear seismic signal processing, (3) geostatistics, (4) effective media theory, and (5) unconventional wave progradation (fractals theory, cellular automata, stochastic wave equation, chaos theory). Geophysical technology is used to determine the geologic structures and to identify reservoir properties. Derivation of the 3D density, P-wave and S-wave fields, and various anisotropy parameters and inelastic attenuation factors (generalized inversion) is a challenging problem.

Many exploration and exploitation successes are attributed to 3D seismic data. The 3D seismic reduces the exploration costs and risk and improves success rates. With the tremendous level of activity and the vast

range of 3D seismic applications, appropriate application of this technology becomes an imporant consideration. The continued increase in the complexity of "geological" problems is focusing more attention on technological aspects of 3D seismic.

Sequence stratigraphy evolved from lithogenetic stratigraphy (1960s) and seismic stratigraphy (1970s), inspiring a conceptual geologic revolution during the middle and late 1980s. Sequence stratigraphy goes beyond the established methods of seismic stratigraphic analysis and interpretation of unconformity-bounded lithostratigraphic units. Based on the joint interpretation of seismic data, well logs, geochemistry, petrography, and outcrop data, sequence stratigraphy provides powerful tools for predicting spatial location, lithogenesis, and stratigraphic succession of depositional systems. For petroleum geologists, these tools may help predict reservoir, seal, and stratigraphic trap potentials.

The geophysical methods enable (1) better 3D description of producing reservoirs, (2) more accurate evaluation of reserves, and (3) more efficient development design [11]. Some investigators note a high potential of seismic for determining fluid contacts and reservoir lithology [51]. This reduces the number of appraisal wells by half without jeopardizing the reliability of reserve estimates and with practically no risk of having dry holes. Berman et al. stated that "it is necessary to increase the experimental methodological studies of wider seismic applications in the field appraisal and development in order to clarify the possibilities and conditions for their successful applications in solving geologic and operations-related field problems for reducing the drilling footage and simultaneous improvement of the quality of information about the field structure and of providing the opportunity to increase the hydrocarbon yield" [9].

The development of direct detection techniques, various resolution increasing techniques, and a means of finding and delineatiing low-amplitude anticlines and non-anticlinal traps holds great promise. The number of discovered non-anticlinal traps is growing. Krylov et al. noted that during the 1976–1980 period the portion of the targets delineated by seismic methods almost doubled [27].

Despite the fact that the mathematical analysis used in petroleum geophysics is as advanced as that applied in general relativity theory, there have been no significant improvements in oil and gas exploration [32]. Verification of structures, especially deep ones, by drilling decreased in highly mature regions, such as North Caucasus, Mangyshlak, Lower Volga Region, and so forth.

Tucker stated that the basics of seismic exploration do not require significant improvement [66]. Only small, but very expensive, corrections are possible. Despite this, the results are becoming more meager and are rapidly approaching zero.

Even if geophysical techniques properly solved the structural problems, they would not significantly improve the general situation. This situation is aggravated by a declining discovery success rate and an increase in the number of dry structures. The task of structural delineation is becoming more expensive, and large amounts are expended for delineating numerous and unnecessary empty prospects. According to some calculations, the discovery success rate resulting from locating exploratory wells at random rather than on delineated prospects, may reach 27%, with a more common figure of 12 to 20% [32]. Thus, the success rate of random drilling would be close to that actually achieved in some regions. It is interesting to compare these results with those of Menard, who analyzed the exploration and appraisal data in the U.S. between 1880 and 1985 and concluded that the per-foot accrual of reserves was no higher than that which would be obtained by totally random drilling [62].

This strongly suggests that only direct hydrocarbon detection techniques may improve discovery efficiency. In the area studied by Salmanov, only 6 out of 52 bright spots tested by drilling, produced oil and gas [47]. This does not discredit the direct detection techniques. The new is always created within the framework of the old. The question is whether these new techniques will be competitive with conventional methods and eventually displace them.

Thus, the future evolution of the exploration-appraisal process will be determined by the evolution of geophysical techniques. If the forecast of geologic section and bright spot identification techniques are effective and their cost is comparable to the amount saved by not drilling wells,[1] even greater pressure will be exerted on the initial stages of the system and the cost will be even more uniformly distributed among the stages. Part of the tasks currently accomplished by exploratory drilling will be performed at the stage of target delineation. The tasks of appraisal drilling will be considerably different and will not involve studying field geology, delineating accumulations, reservoir characterization, and so forth. These will be addressed by geophysics, and the boundaries between appraisal and development (developmental appraisal) will gradually disappear.

Under favorable circumstances, a development well may be drilled immediately after an exploratory well. The location of the next development (or appraisal) well will be determined only after well logging and testing, re-interpretation of geophysical results (seismic, gravimetric, magnetic), and detailed geophysical studies conducted for refining the structure and reservoir characterization and model construction. This may be called a geophysical evolution.

A different scenario will evolve if geophysical methods do not advance. Oil and gas field exploration will be conducted using random drilling. This is due to the low reliability of delineated structures. Thus,

the single-well exploration technique, which is almost random, will predominate. In this case, the system will lose its first links (the stages of target discovery and delineation will disappear), as exemplified in Bashkortostan and Tatarstan in Russia.

Another possibility is a change in the order of the system elements. Target delineation will be conducted after a discovery is made through random drilling. This will be a cost-savings technique because there will be no need to delineate non-commercial structures. With this technique, target delineation will become an element of field appraisal. These stages, which are currently far apart, will merge. This method of evolution can be called stochastic.

The above demonstrates that in order to improve the efficiency of the exploration-appraisal process it is not enough to improve the efficiency at each stage. Instead, it is necessary to visualize the system as a whole. In this respect, the analysis of the exploration-appraisal process is skewed in that only certain stages are studied in-depth. Investigators concentrate their efforts on a particular stage. Perfecting the process at every stage will result in a solution of the problem of control for the entire exploration-appraisal process. In reality, this situation is much more complex. Given this complexity, the capability of geophysical methods to detect hydrocarbons directly and the solution of field appraisal and development problems, have first priority. This will define the further evolution of the exploration-appraisal process which, in turn, will determine its efficiency. This is a pending problem that must be solved quickly. Once solved, important decisions must be made on linking the exploration-appraisal process with either direct detection techniques or some variation of random drilling. Experimental studies in this direction should be increased.

It must also be kept in mind that neither concept of evolution will appear in a pure form. Somehow they will combine with integrated subsurface geology techniques, which use geologic data to select location of exploratory wells.

A substantial number of fields are discovered using subsurface geology techniques in combination with surface methods, geophysics, and core drilling. The role of the subsurface geology techniques is especially important in identifying the non-anticlinal traps. For this reason, neither geophysics nor random drilling will be independent operation concepts, but will be used in combination with the subsurface geology techniques.

Note

1. Currently, the cost of delineating a structure is equal to the drilling cost of one or two wells.

References

1. Baymukhametov, K. S. and Sattarov, M. M. (1975). "On a Particular Probabilistic Statistical Model for the Forecasting of Oil and Gas Reserve Accrual." *Geologiya Nefti i Gaza,* No. 3: 27–32.
2. Bakirov, V. A. (1972). "A Statistical Model of Oil and Gas Field Size Distribution." *Geologiya Nefti i Gaza,* No. 2: 63–68.
3. Balasanov and V. A., Kamenetskiy, S. G. (1977). "Some Methodological Problems in Forecasting of Oil Field Appraised Reserves." *Economics of Oil Industry,* No. 9: 26–28.
4. Beylin, A. M. (1985). "Quantitative Techniques in Forecasting of Exploration and Appraisal Drilling Efficiency." In *Economics of Exploratory and Appraisal Drilling for Oil and Gas.* Moscow, Nedra, pp. 55–112.
5. Belonin, M. D. and Podolskiy, Yu. V. (1983). "Geologic-Mathematical Modeling of the Reserve Accrual and Production Evolution and the Order Determination in Oil and Gas Field Discoveries." In *Geologic and Economic Evaluation of Oil and Gas Resources.* Leningrad, pp. 172–188.
6. Buyalov, N. I., Korniyenko and V. N., Lavrushko, N. P. (1979). "Methodological Basics of the Forecast for a Quantity of the Discovered Oil and Gas." *Sovetskaya Geologiya,* No. 8: 3–13.
7. Vesnina, T. L. (1976). "Questions of Economic Evaluation in Oil and Gas Exploration." *Geologiya Nefti i Gaza,* No. 6: 31–39.
8. Vistelius, A. B. (1963). *On Probability Distribution Function of Phosphorus Concentration in Granitoids of Switzerland, Guyana and Equatorial Africa.* Proc. Acad. Sci. USSR, 152,(6): 1449–1452.
9. Berman, L. B., Neyman, V. S., Abrikosov, A.I. et al. (1987). "Capability of Seismic Surveys for a Solution of Geologic Field Problems." *Geologiya Nefti i Gaza,* 54–60.
10. Gabrielyants and G. A., Sorokin, Yu. V. (1976). "Basic Principles of Locating Wells at the Exploratory Stage." Scientific Review. Series: *Geology and Exploration of Gas and Condensate Fields.* Moscow, Vniigazprom, 52 pp.
11. Gogogenkov, G. N., Moiseyev, V. N. and Savostyanov, N. A. (1987). "Application of Geophysical Techniques in Oil Field Supplemental Appraisal and Development." *Geologiya Nefti i Gaza,* 11.
12. Gumbel, E. (1965). *Statistics of Extreme Values.* Mir, Moscow, 450 pp.

13. Dobrovolskiy, M. B. and Volodin, I. A. (1974). "Application of Mathematical Techniques for the Determination of Rational Appraisal Degree for Oil and Gas Fields when Transferring from Appraisal to Development." Review. Series: *Oil and Gas Geology and Geophysics.* Moscow, VNIIOENG, 62 pp.
14. Zheltov, Yu., P. (1985). "Mathematical Model of Useful Mineral Appraisal and Development and its Application in Long-Term Oil and Gas Industry Evolution Forecast." In *Forecast of Geologic and Economic Quality of Oil and Gas Resources,* Moscow, Nauka, pp. 64–72.
15. Modelevskiy, M. Sh., Vasiliyev, O. V., Feldman, S. L. and Shimanovskiy, V. V. (1977). "Exploration and Appraisal Efficiency versus Maturity of Oil and Gas Resources." *Oil and Gas Geology and Geophysics,* No. 11: 11–16.
16. Kaufman, G. M., Balker, I. and Cruit, D. (1978). "A Probabilistic Model of the Exploratory and Appraisal Process." In *Evaluation Techniques of Oil and Gas Undiscovered Resources.* Moscow, Nedra, pp. 137–163.
17. Knoring, L. D. (1980). *Basics of the Optimization Theory at the Appraisal of Oil and Gas Fields.* Leningrad, Nedra, 307 pp.
18. Knoring, L. D. (1981). "Optimization of Oil and Gas Field Exploration and Appraisal through Deep Drilling." *Sovetskaya Geologiya,* No. 9: 19–31.
19. Knoring, L. D. (1981). "Effect of Operations Tempo on the Well Number at Oil and Gas Field Exploration and Appraisal." *Geologiya Nefti i Gaza,* No. 3: 26–31.
20. Knoring, L .D. (1983). "Optimization of Oil and Gas Field Exploration and Appraisal Control at Deep Drilling." *Sovetskaya Geologiya,* No. 6: 23–32.
21. Knoring, L. D. (1986). "Modeling of the Oil and Gas Reserve Accumulation Evolution." *Sovetskaya Geologiya,* No. 9: 9–17.
22. Knoring, L. D. (1986). "On the Forecast of Oil and Gas Accumulation Evolution." *Geologiya Nefti i Gaza,* No. 12: 10–15.
23. Kozlovskiy, E. A. (1979). "Mineral Raw Materials Base and the Time Factor." *Sovetskaya Geologiya,* No. 3: 3–17.
24. Kontorovich, A. E. and Demin, V. I. (1979). "Forecast of the Oil and Gas Field Number and the Reserve Size Distribution." *Geologiya i Geofizika,* No. 3: 26–46.
25. Kontorovich, A. E., Demin and V. I., Strakhov, I. A. (1985). "Different Size Oil and Gas Field Discovery Patterns in Oil and Gas Basins." *Geologiya i Geofizika,* No. 11: 3–15.
26. Krylov, N. A. (1982). "On the Notion of Exploration Play in Exploration for Oil and Gas Fields. *Geologiya Nefti i Gaza,* No. 11: 17–22.
27. Krylov, N. A., Aleksin, A. G. and Baturin, Yu. N. (1986). "Tasks and Ways of Scientific and Technological Progress in Oil Exploration within Highly Mature Regions." *Geologiya Nefti i Gaza,* No. 7: 1–7.
28. Krylov, N. A., Baturin, Yu. N. and Ryzhik V. M. (1986). "Study of Evolution of the Oil Reserves to Production Ratio." *Geologiya Nefti i Gaza,* No. 12: 1–6.
29. Leybson, M. G. (1974). *Methodological Recommendation on Planning the Structure (Prospect) Delineation for Deep Oil and Gas Drilling.* Leningrad, VNIGRI, 29 pp.

30. Leybson, M. G. (1985). "A Technique for Geologic-Economic Resource Forecasting and Parameters of Oil and Gas Exploration Evolution." *Geologiya Nefti i Gaza,* No. 5: 18–23.
31. Lindtrop, N. T. and Mukhin, V. V. (1963). "On Methodology of Oil and Gas Exploration by Deep Drilling." Proc. VNIGRI, 220: 387–404.
32. Lossovskiy, E. K. (1986). "Seismic Surveys of Anticlinal Structures and Random Drilling in Oil and Gas Basins." *Geofizichesky Zhurnal,* 8 (6): 41–48.
33. Melik-Pashayev, V. S. (1968). *Methodology of Oil Field Appraisal.* Moscow, Nedra, 183 pp.
34. "Report" (1988). *Methodological Recommendation on the Application of "Provisional Methodology for Economic Evaluation of Oil and Gas Undiscovered Resources and Preliminarily Estimated Reserves."* Leningrad, VNIGRI, 142 pp.
35. Belonin, M. D., Buyalov, N. I., Zakharov, E. V., et al. (1979). (Editors: N. I. Buyalov and V. D. Nalivkin). *Techniques of Oil and Gas Potential Evaluation.* Moscow, Nedra, 332 pp.
36. Nalivkin, V. D., Kuzmin, V. I. and Lukyanova, V. G. (1982). "Natural Boundaries in the Series of Oil and Gas Field Size Distribution." *Proc. Acad. Sci. USSR,* 266 (4): 947–951.
37. Shcherbakov, V. P., Purkin, L. B., Slavkin, V. S. and Gilbershtein, A. P. (1982). "On Some Aspects of Perfecting the Planning of Oil and Gas Exploration and Appraisal Operations." *Geologiya Nefti i Gaza,* 7: 1–5.
38. Karpushin, V. Z., Osipova, G. A., Khalizova, E. V. and Mozgovaya, N. S. (1977). "Determination of Rational Information Volume for Evaluation of Calculation Parameters." In *Methodology of Appraisal and Reserve Calculation of Oil and Gas Fields.* Moscow, Nedra, 59–69 pp.
39. Frolov, E. F., Bykov, N. E., Egorov, R. A. and Fursov, A. Ya. (1976). *Appraisal Optimization for Oil and Gas Fields.* Moscow, Nedra, 303 pp.
40. Price, D. (1966). "Small Science, Big Science." In *Science on Science.* Moscow, pp. 18–31.
41. Krylov, N. A., Afanasyev, Yu. T., Baturin, Yu. N. and Ryzhik V. A. (1985). "Forecast of Dynamic Parameters in the Oil and Gas Resource Development." In *Forecast of Geologic and Economic Quality of Oil and Gas Resources.* Moscow, Nauka, pp. 19–32.
42. Kontorovich, A. E., Fotiadi, E. E., Demin, V. I. et al. (1981). *Forecast of Oil and Gas Fields.* Moscow, Nedra, 350 pp.
43. Pronina, E. N., Kuzmin, V. I. and Nalivkin, V. D. (1982). "Modeling of the Production and Reserve Evolution for Non-renewable Natural Resources." *Sovetskaya Geologiya,* No. 1: 36–47.
44. Prokhorov, Yu. V. (1965). "On a Question of Log-Normal Distribution in Geochemical Problems." *Probability Theory and its Applications,* 10 (1): 184–187.
45. Prokhorov, Yu. V. and Rozanov, Yu. A. (1967). *Probability Theory. Basic Concepts, Boundary Theorems, Random Processes.* Moscow, Nauka, 495 pp.
46. Ryzhik, V. M., Alenin, V. V. and Ivanova, M. N. (1985). "Forecast of the Field Size Distribution." In *Forecast of Oil and Gas Resources Geologic-Economic Quality.* Moscow, Nauka: 115–120.

47. Salmanov, F. K. (1988). "Main Tasks for Restructuring in Logistics and Conduct of Oil and Gas Exploration and Appraisal Operations." *Sovetskaya Geologiya* (2): 3–9.
48. Semikhodskiy, G. E. and Timoshin, Yu. V. (1982). "Forecast of Gas Potential for Dnieper-Don Depression based on Statistical Data." *Geologiya Nefti i Gaza,* 7: 9–14.
49. Trofimuk, A. A. (1964). "On Rational Ratio of Oil and Gas Reserves in Different Categories and a Technique for Planning of their Accrual." *Geologiya Nefti i Gaza,* 7: 5–12.
50. Uspenskaya, N. Yu. and Tabasaranskiy, Z. A. (1966). *Oil and Gas Provinces of the USSR.* Moscow, Nedra, 213 pp.
51. Fursov, A. Ya. and Khalimov, E. M. (1987). "Experience with Appraisal of Small Oil Fields." *Geologiya Nefti i Gaza,* 11: 40–46.
52. Shwemberger, Yu. N. (1978). "Forecast of the Oil and Gas Field Sizes." *Geologiya Nefti i Gaza,* 3: 58–62.
53. Shpilman, V. I. (1982). *Quantitative Forecast of Oil and Gas Potential.* Moscow, Nedra, 215 pp.
54. Schroedinger, E. (1948). *What Is Life from Physics Viewpoint?* Moscow, Fizmatgiz, 53 pp.
55. Bunin, O. I., Sakharov, V. V., Ladyzhansky, P. I. et al. (1971). "Efficiency of Oil and Gas Exploration and Appraisal in the USSR and the Ways of its Improvement." In *Status Of Oil and Gas Exploration and Appraisal in the USSR and the Ways to Increase Their Efficiency,* 7–22.
56. Yablonsky, A. I. (1975). "Stochastic Models of Scientific Activities." In *Systemic Studies. Year-Book.* Moscow, Nauka, pp. 5–42.
57. Yanch, E. (1970). *Forecast of Scientific-Technology Process.* Moscow, Progress, 568 pp.
58. Dubey, S. D. (1966). "Asymptotic Efficiencies of the Moment Estimates for the Parameters of the Weibull Laws." *Natnl. Res. Logist. Quart,* 13 (3): 265–288.
59. Hubbert, M. K. (1967). "Degree of Advancement of Petroleum Exploration in the United States." *Bull. Am. Assoc. Petrol. Geol.,* 51/11: 2207–2227.
60. Ivanhoe (1976). "Oil/gas Potential in Basins Estimated." *Oil and Gas Journal,* 6: 154–155.
61. Likes, I. (1972). "Estimation of Probabilities P_r $(A < x < b)$ for the Pareto Distribution." *Statistische Hefte,* 3: 291–297.
62. Menard, H. W. (1981). "Toward a Rational Strategy for Oil Exploration." *Sci. Amer.,* 244 (1): 47–57.
63. Moore, C. L. (1962). "Method for Evaluating the United States Crude Oil Resources and Projecting Domestic Crude Oil Availability." Washington, Dept. Interior, Office of Oil and Gas, 112 pp.
64. Rowlands, N. I. and Sampey, D. (1977). "Zipf's Law—An Aid to Resource Inventory Prediction in Partially Explored Areas." *Math. Geol.,* 9 (4): 383–391.
65. Schuenmeyer, S. H., and Drew, L. J. (1983). "A Procedure to Estimate the Parent Population of the Size of Oil and Gas Field as Revealed by a Study of Economic Truncation." *J. Inst. Assoc. Math. Geol.,* 15 (1): 145–161.

66. Tucker, L. R. (1981). "Future Direction of Petroleum Exploration." *Unconventional Meth. Expl. Petrol. and Natur. Gas.* 2nd Symp., Dallas, TX, pp. 249–256.
67. Uri, N. D. (1979). "New Look at U.S. Reserves Shows Higher Potential." *World Oil,* 188 (2): 61–66.
68. Zapp, A. D. (1962). "Future Petroleum Producing Capacity of the United States." Contribution to economic geology. *Geolog. Survey Bull.,* 1142–H. U.S. Government Printing Office, Washington, 36 pp.
69. Eremenko, N. A. and Chilingarian, G. V. (1991) *Petroleum Geology Handbook.* OSI Publications, Los Angeles, California, 600 pp.

Subject Index

F

G

H

I

L

M

Author Index