Risk Reduction in Oil/Gas Exploration by Combining Geomorphological Analysis and Artificial Intelligence

S. Guberman, Vice President Technology, Digital Oil Technologies, CA, USA,*
Y. Pikovski, Head Geophysicist, Moscow State University, Russia, FSU
E. Rantsman, Head Scientist, Institute of Geography, Russian Academy of Sciences. Russia, FSU

Summary

This paper describes a technology developed for risk reduction in oil and gas exploration. The technology combines a morphological zoning technique (MSZ) and a computer pattern recognition system(CPRS). In 1985 this technology was applied to oil and gas exploration in the Andes Mountain range. At that point in time, 11 sites which had a high probability of containing giant oil/gas fields were identified (Guberman et al., 1986). Between 1986 and today, three new giant oil/gas fields were found in the Andes (South America) - Cano Limon, Cusiana-Cupiagua (Colombia) and Camisea (Peru). All three are located within the predicted sites. There are eight more sites in that region recognized as highly promising.

Introduction

Modern techniques of prospecting for new oil and gas fields are based mainly on models of the deep structure of the Earth's crust obtained by means of geological and geophysical methods. These models provide good results in the mapping of probable oil and gas traps, but the challenge remains in making accurate distinctions between productive and non-productive reservoirs. The geochemical and geophysical methods for direct detection of oil and gas pools do increase the success ratio of exploratory drilling, but the uncertainty in the interpretation of data obtained is still high. Therefore, selection of licenses for oil and gas exploration is inevitably associated with high risks of obtaining negative results.

This paper describes the application of a technology which forecasts geological phenomena, the field of oil and gas exploration. The technology combines a morphostructural zoning technique (MSZ), which defines an objective set of geological objects, and a computer pattern recognition system(CPRS) which classifies the objects by similarity of features. After 10 years of development in the USA and Russia this technology is now ready for introduction to the Geophysics and Geologic communities.

The goal of MSZ is to define the recent block structure of the Earth's crust. CPRS was developed using the "learning by examples" approach - the basic idea from F. Rosenblatt's work "Perceptron" (1958) on the first artificial neural network. A US-Russian team first applied the combination of these two techniques in the early 1970's, to predict new sites of strong earthquakes. The prediction was extremely successful: In California the location of 13 out of 14 strong earthquakes were correctly identified, including the Northridge earthquake. Subsequent work by Aminzadeh et al (1994) further proved the power of artificial neural networks in predicting earthquakes using the information from seismic precursors.

In 1985 this technology was applied to oil and gas exploration in the Andes Mountain range. A morphostructural map issued in 1981 was the primary source of information. It was found that of the 17 largest oil and gas fields in that region, 16 coincided with the intersections of lineaments*(nodes)*. The CPRS was then applied to all nodes in the region. The existence of different patterns, which differentiated nodes which contained giant oil fields from those which did not, were determined. As a result, 11 nodes which had a high probability of containing giant oil fields were identified but had not been explored at that point in time (1985). The total area of these 11 nodes is approximately 8% of the total basin area. That prognosis was published in 1986.

Between 1987 and today, three new giant oil/gas fields (more then 1 billion bbl of oil-in-place equivalent each) were found in the Andes (South America) - Cano Limon, Cusiana-Cupiagua (Colombia) and Camisea (Peru). All three are located within the predicted sites. The rate of success is estimated at over 75%. There are eight more sites in that region recognized as highly prospective.

The technique of formalized morphostructural zoning was developed in the late 1960s and 1970s. MSZ is based on regional and detailed hierarchical model of the recent block structure of the Earth's crust. Initial data used in the development of a regional model are topographic and hypsometric maps, satellite images, regional geologic, tectonic, and other special maps of scales ranging from 1:500,000 to 1:1,000,000.

A regional model of the recent block structure of the Earth's crust includes the following main components: 1) homogenous areas — *blocks*; 2) linear zones between blocks — morphostructural *lineaments;* 3) areas of *lineaments* intersections — morphostructural *nodes.* The lineaments and the nodes consist of small blocks, which are more active zones as compared with relatively stable large blocks. Activity of those zones increases from relatively stable large blocks, through the lineaments toward the nodes. All components of the hierarchical system comprise a single system.

The most important feature in the development of the regional model is the establishment of a hierarchy of blocks. In compliance with established parameters, the homogenous groups of blocks are to be united into mesoblocks and mesoblocks in turn grouped into macroblocks. The degree of homogeneity of blocks is determined by means of a formalized analysis concerning informative features of relief (height ratio, extension and density of linear forms of relief, erosion patterns). The hierarchy of blocks determines the hierarchy of lineaments, classifying them as 1st rank, 2nd rank, and 3rd rank. The higher the rank of a block boundary, the stratigraphically deeper the roots of that boundary in the Earth's crust.

The difference between regular morphostructural zoning and the technique described above must be stressed. In common morphostructural analysis, the notion of a lineament is well known and widely used. A lineament is determined as a linear element of the relief of sufficient length (for example, the linear part of the river, the linear border of the valley, the linear part of a shore) or a chain of linear elements. According to that definition the lineament is locally defined. This means that existence of the lineament does not depend on the surrounding areas. In that approach the lineament is a basic element of the morphological structure.

In the discussed approach, the lineament is a secondary element of morphostructure. The primary element is the block – a relatively homogeneous area. The borders of the blocks form the lineaments. Therefore, the lineaments are *secondary* to the blocks. This means that the existence and the position of the lineaments are determined not locally, but as a part of a broad pattern. This increases the reliability of the lineaments as treks of the tectonic life of the core. If certain linear morphological elements do not separate two areas with different morphologies, that element cannot be treated as a lineament in the analysis. Some linear but weakly expressed morphological elements will be treated as an evident lineament if it separates two evidently different areas.

Pattern recognition.

Pattern recognition is a recognized and widely used method of Artificial Intelligence. The problem of recognition was represented in the following form: a set of objects is given, with each object described by the answers to a standard form. Each object belongs to one of two classes. The goal is to find which class each object belongs to. The first step in recognition is the "learning phase" - to find the sets of characteristic features for each class using examples of objects of each class. The second step is the "recognition phase" which applies the characteristic features to each object and decide which class it belongs to.

The object of recognition was defined as the morphostructural node - the area of intersection of lineaments. It was observed[6] that in the Andes, 16 out of 17 giant oil/gas fields were located within the morphostructural nodes (i.e. within a radius of 45 miles). The choice of the object defined the formulation of the problem as follows. Highly prospective locations of large oil/gas fields were desired. The nodes would be the focus of the investigation. The goal is to separate the set of morphostructural nodes into two parts: nodes which contain large oil fields and ones which don't. To achieve this goal one needs to find a set of features (or combination of features) which are common to the nodes containing large oil/gas fields (the "oil set"), apply them to the remainder of the nodes and determine the nodes which are most similar to the nodes in "oil set". These nodes would be the most promising for future oil and gas discoveries.

Description

Each object was described by a number of parameters. These parameters were used by the pattern recognition algorithm for characterizing the classes of nodes with (class I) and without (class II) big oil/gas fields. The parameters are as follows:

1. Height in the center of the node (center of lineaments intersection) H_c as a measure of lifting (or sinking) of the node as a whole,
2. Maximum height difference in the node dH as a measure of differential vertical movements,
3. Number of node-forming lineaments N_L as a measure of faulting,
4. Highest rank of lineament in the node R_L as a measure of the node's depth,
5. Thickness of sediments,
6. Contact of relief types in the node (mountain-mountain, or mountain-foothill, or mountain-plain),

Most parameters describe past levels of tectonic activity of the node.

Learning

For the learning process one has to determine the representatives (examples) of objects of both classes (I and II). In this case the giant oil/gas fields in the Andes basins were used as the learning set. The set of examples for class I comprised nine morphostructural nodes containing large oil/gas fields (one contains two oil fields). Four nodes with large oil/gas fields were used in the analysis (as the control group). The set of examples for class II contained all the other 30 nodes, located in known oil and gas sedimentary basins. These included a number of class I nodes and the task and challenge was to identify them, as most nodes (over half) presumably belonged to class II, i.e. they lacked large oil or gas fields. For recognition Bongard's algorithm was used, described in details in[2,5]. As a result of learning process the program picked three class I and three class II criteria. That set of criteria forms the decision rule. The result of recognition using this decision rule was as follows.

Recognition

Eight nodes used as examples of class I were classified correctly as promising. One node, used as example of class I, was incorrectly recognized as belonging to class II (without large oil field). Four nodes with large oil/gas fields from the control group, which were not used in learning, were recognized correctly as promising. Of the 30 objects in class II, four were identified as the most promising. A set of nodes located in sedimentary basins without known large oil/gas fields (as at 1984) was also classified (37 nodes). Eleven out of them were recognized as promising great discoveries in future.

The reliability of the prognosis

One of the main dangers in pattern recognition arises when only small amount of data is available – inappropriate adjustment of the decision rule to conform to the known data. Special efforts to avoid self-deception were undertaken:

1. As the process of morphostructural zoning is not completely formalized, the adjustment of position of the nodes and location of large oil/gas fields is hypothetically possible. This possibility was excluded completely by using a morphostructural map issued two years before the analysis(1981) and prepared for different purposes[8].
2. Only one type of object was investigated - the node. Only one hypothesis about the relationship between the large oil fields and the type of object was checked and this hypothesis was accepted. In the event that several different types of objects were tested, the forecasting would be less reliable.
3. A small number of parameters were chosen for node's description (six). The larger the number of parameters the easier to find a rule which will can separate the entire set object, and also easier to adjust the decision rule to match the given data. However, it will be difficult to use such a rule in a different region as it relied on too many parameters specific to test region.
4. Only one run of the recognition algorithm was performed.
5. Only four objects for testing the decision rule were used. This data set was obviously too small. To increase the test set a recognized technique known as the "sliding test" was applied: a series of independent learning, eliminating from the learning data one of the objects at a time and recognizing the eliminated object using the others. Thus eight more test objects (twelve in total) were obtained and 11 were correctly recognized.

The results of all these tests were published in a 1986 forecasting map of the Andes[6].

Decision rule's transfer

The reliability of the technology was checked by applying the decision rule to other basins: West Siberia (plains), California (piedmont and intermountain depressions) and the North Sea (continental shelves)[10]. For each region a morphostructural map was prepared and the criteria for locating nodes with a high probability of giant oil/gas fields were applied to each morphostructural node. All parameters were expressed in relative units, so the difference in absolute heights between Siberia and the Andes disappears.

As one can see from Table 1 the test results are significant. The positive test results 1) increases the reliability of the technology; 2) estimates the level of success as equal greater than 75%; 3) demonstrates that the obtained decision rule is invariant to differences in geological conditions; and 4) excludes from the technology the learning process before recognition, replacing it with a universal decision rule applicable to any geological region.

The Field Test

The map published in 1986 contained predictions of areas with a high probability of large oil and gas fields in the basins of the Andean mountain belt(Figure 1). In all, 23 areas were identified by the computer. Each location is defined by a circle with a radius of 45 miles. Many of these predicted places (12) coincided with location of existing large fields. In addition, the computer identified 11 new areas as promising for large discoveries, but there were no large pools known in these areas at that time[2]. These 11 sites cover only 8 % of the total area of all Andes basins. Moreover, most of the future discoveries were located in basins with previously unknown oil or gas fields in the vicinity. For the small number of known large oil/gas fields one could not use the regular statistical methods to test the reliability of the prediction. As described above special logical and statistical methods of testing were developed and the decision rule passed these tests. In 10 years since publication of this forecast, only three giant oil/gas fields have been discovered: Cano-Limon and Cusiana in Llanos basin and Camisea in Ukayali basin [11—13]. All discoveries were made in places identified on the prognostic map as places of future great discoveries. The probability that all three discoveries will fall in 8 % of the predetermined area by chance is extremely low ($p = (0.08)(0.08)(0.08) = 0.0005$).

Discussions

How can one explain that oil deposits are located in nodes possessing special patterns, as described by the decision rule? The patterns can be interpreted as follows: 1) The junction of faults provide the ability for hydrocarbon migration, 2) The sinking pattern – low altitudes in the center of the node – reflects the divergence of the blocks and facilitates the migration of the organic material, 3) The high range of lineaments forming the node provides a high heat flow[14] which provides the ability for chemical transformation of the pre-oil material, 4) The increased number of lineaments forming the node reflects the existence of a large number of small blocks which increases the ability in creating structural traps, 5) The relatively large variation in altitudes in the node reflects the existence of vertical movements of small blocks which provides the appearance of the anticline structures in the above layers and structural. traps in the neighbor blocks, i.e. provide the ability for hydrocarbon accumulation, 6) The sufficient depth of the sediments in the node increases the probability of the needed combination of collectors and seals to the reasonable level (at least 75%).

Limitations

There are two limitations of the technology: 1) The technology cannot be applied to finding smaller sized oil/gas deposits, which is important in well explored oil basins. 2) The technology does not identify traps. Current seismic techniques are still required. This technology enhances and focuses current best practices.

Conclusions

The proposed technology complements traditional regional studies on promising basins and works on preparation of sites for exploratory drilling. The efficiency of the technology has been tested with a number of statistical, logical, geological tests as well as by field tests.

References

1. Gelfand I.M., Guberman Sh.A., et al. “Criteria of high seismicity determined by pattern recognition,” *Tectonophysics*, {1972) **13**, 415.
2. Gelfand I.M., Guberman Sh.A., et al. “Pattern recognition applied to earthquake epicenters in California,” *Earth Planet Inter*, 1976, No. 11, 227.
3. Rantsman E.Ya. *Sites of earthquakes and the morphostructure of mountainous lands.*, Nauka, Moskow, (1979).
4 . Glasko M.P., Rantsman E.Ya. “Geographical aspects of the block structure of the Earth’s crust, “ *Bulletin of the USSR Academy of Sciences, Geographical Series*, (1991) No. 1, 5
5. Guberman, Sh.A. *Informal data analysis in geology and geophysics,* Nedra, Moscow, (1987) (in Russian).
6. Guberman, Sh.A., Zhidkov M.P., Pikovsky Yu.I., Rantsman E.Ya. “On some criteria of oil and gas content of morphostructural nodes (the Andes),” *Transactions (Doclady) of the USSR Academy of Sciences, Earth Science Sections*, Scripta technica, INC, (1986) **291**, 1436.
7. Rantsman E.Ya. “Morphostructural subdivision of the West Siberian Plain according to formalized attributes with view to local prognosis of oil and gas fields,” — *Geomorphology*, (1989), No. 1, 30.
8. Zhidkov M.P. “Transversal morphostructural lineaments of the Andes,” *Geomorphology,* (1981), No. 2, 47.
9. Guberman, Sh.A., Pikovsky Yu.I. “Seismogenic disjunctive points with implications to the regularities in oil-and-gas field occurrence,”*Bulletin of the USSR Academy of Sciences, Earth Physics Series*, (1984) No 11, 10.
10. Vinogradov Ye.I., Guberman Sh.A., Dmitrievski A.N., Pikovski Yu.I., Rantsman E.Ya., Ulmasvay F.S. “Local prediction of very large oil and gas accumulations from morphostructural data,” *Transactions, (Doclady) of the USSR Academy of Sciences, Earth Science Sections*, Scripta technica, INC, (1989) **305**, 70
11. Rivero R.T., Domingues J.G., Slater J.A., Hearn C.L. Cano Limon field, “Colombia: the latest giant oil reservoir in South America,“ *Journal of Petroleum Technology*, (1988) No.7, 874.
12. Oil and gas exploration and production. South America: growth in E&P opportunities keep accelerating, *World Oil*, (March, 1993) **212**, 45.
13. South America, *World Oil*, (August, 1995) **214**, 62.
14. Pikovsky Yu.I. *Natural and technogenic flows of hydrocarbons in the environment,* Moscow University Press, Moscow, (1993).

Risk Reduction in Oil/Gas Exploration by Combining Geomorphological Analysis and Artificial Intelligence

Figure 1: Morphostructural map of Andes Mountain (northern part of S. America)

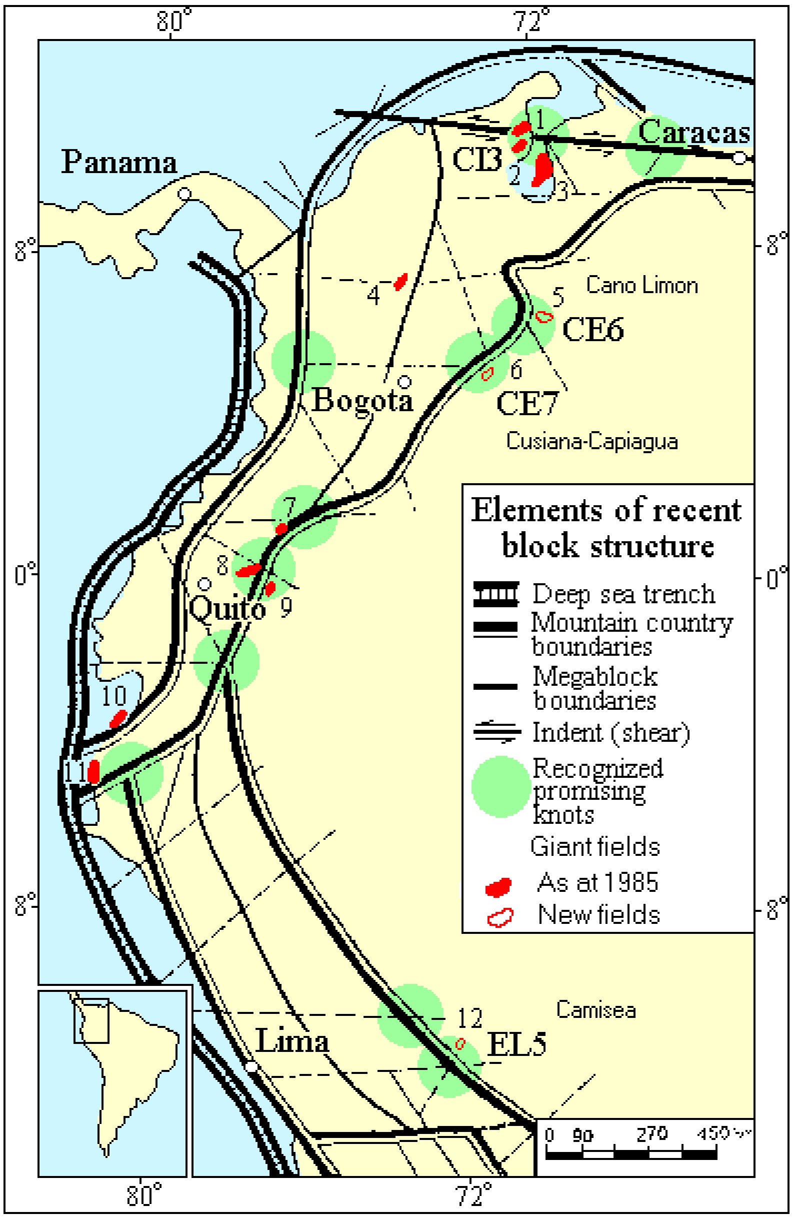

Table 1: Recognition of promising nodes in different geological areas

	West Siberia	The Andes	North Sea	California
Number of nodes, total	68	76	90	48
Number of promising nodes	23	23	13	21
Radius of promising areas, miles	45	45	30	30
Portion of basin area occupied by promising nodes	0.16	0.15	0.15	0.16
Percentage of 10 largest fields located within the promising nodes	100%	100%	100%	100%