Database for Assessment Unit-Scale Analogs (Exclusive of the United States)

Open-File Report 2007-1404

U.S. Department of the Interior
U.S. Geological Survey

Database for Assessment Unit-Scale Analogs (Exclusive of the United States)

By Ronald R. Charpentier, T.R. Klett, and Emil D. Attanasi

Open-File Report 2007-1404

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
DIRK KEMPTHORNE, Secretary

U.S. Geological Survey
Mark D. Myers, Director

U.S. Geological Survey, Reston, Virginia 2008

For product and ordering information:
World Wide Web: http://www.usgs.gov/pubprod
Telephone: 1-888-ASK-USGS

For more information on the USGS—the Federal source for science about the Earth,
its natural and living resources, natural hazards, and the environment:
World Wide Web: http://www.usgs.gov
Telephone: 1-888-ASK-USGS

Suggested citation:
Charpentier, R.R., Klett, T.R., and Attanasi, E.D., 2008, Database for assessment unit-scale analogs (exclusive of the United States): U.S. Geological Survey Open-File Report 2007-1404, 61 p.

Table of Contents

Abstract

This publication presents a database of geologic analogs useful for the assessment of undiscovered oil and gas resources. Particularly in frontier areas, where few oil and gas fields have been discovered, assessment methods such as discovery process models may not be usable. In such cases, comparison of the assessment area to geologically similar but more maturely explored areas may be more appropriate. This analog database consists of 246 assessment units, based on the U.S. Geological Survey 2000 World Petroleum Assessment. Besides geologic data to facilitate comparisons, the database includes data pertaining to numbers and sizes of oil and gas fields and the properties of their produced fluids.

Introduction

This publication presents a database of geologic analogs useful for the assessment of undiscovered oil and gas resources. Primary use of the database is to compare an immaturely explored area to a geologically analogous area that has been more maturely explored. Particularly in cases of assessing undiscovered resources in immature areas, sufficient discovery history data may not be available for other methodologies, such as discovery process models, to be employed.

Each of the 246 analogs in the database is at the assessment unit (AU) scale and is based on the assessments of undiscovered resources in the U.S. Geological Survey (USGS) 2000 World Petroleum Assessment (U.S. Geological Survey World Energy Assessment Team, 2000, referred to as the "WPA 2000"). In that publication, an assessment unit was defined as "A mappable volume of rock within the total petroleum system that encompasses fields (discovered and undiscovered) which share similar geologic traits and socio-economic factors" (Klett and others, 2000). An AU is equivalent to a "play" or a group of related plays. The horizontal and vertical sizes of assessment units (their areal extent and the volume of rocks involved) can vary from one that covers the entire province to one covering only a small portion (exploration-trend scale). Because the WPA 2000 provided estimates only for selected provinces outside the United States, the database does not include any analogs from U.S. areas.

The methodology used in most USGS assessments relies on estimates of numbers and sizes of undiscovered accumulations. An accumulation may consist of an entire field or only that part of a field that is within a particular AU. (Note: in the WPA 2000, fields were not subdivided, except that pools were used in the Western Canada Sedimentary Basin). The term "accumulation" is used herein when referring to the USGS methodology and the term "fields" is used when referring to the analog database. Thus, the database provides information primarily on numbers and sizes of discovered and undiscovered fields and the geologic characteristics by which analogs may be chosen.

The files containing the database are included in two formats. A Microsoft Excel version (WorldAnalogs.xls) can be read by users of that software. This version includes several software tools, described below, to assist the user in selecting sets of analogs and in creating interpretive charts. An alternate version, for non-users of Microsoft Excel, is presented as eight tab-delimited files (WA_geology.tab, WA_oil.tab, WA_gas.tab, WA_BOE.tab, WA_oilbins.tab, WA_gasbins.tab, WA_BOEbins.tab, and WA_ancillary.tab). This version can be used with other spreadsheet programs, but does not include any software tools.

Sources of Data

The basic data for our analog database are drawn from three sources: (1) sizes and numbers of discovered fields are mostly from the IHS Energy, Inc. Exploration and Production database (IHS Energy, Inc., 2004); (2) sizes and numbers of Canadian pools are from the NRG Associates (1995) Significant Oil and Gas Pools of Canada Data Base; and (3) sizes and numbers of undiscovered fields are from the WPA 2000 (U.S. Geological Survey World Energy Assessment Team, 2000).

Only those oil and gas accumulations considered discovered in the WPA 2000 were counted as discovered in our analog database. For the most part, those accumulations were discovered before January 1, 1996, except for the Central Congo Turbidites AU 72030302, for which 24 discoveries from 1996 to 1999 were also included in the WPA 2000.

In all cases, the basic data are presented as counts or statistical expectations (mean estimates) in the standard USGS field-size classes, as listed in table 1. Accumulation sizes are based on the reported recoverable quantities (volumes of cumulative production plus reserves) according to the primary product (oil in oil accumulations, gas in gas accumulations) rather than a total volume of all products. Oil accumulations are differentiated from gas accumulations in that they have a gas to oil ratio (based on cumulative production plus reserves) of less than twenty thousand cubic feet per barrel. Where the analog database relates sizes in barrels of oil equivalent (BOE), an equivalency of one barrel of oil to six thousand cubic feet of natural gas was used (1 BOE = 1 barrel of oil = 6,000 cubic feet of gas).

Table 1. Standard USGS field-size class definitions

Size Class	Oil Field Size * (Millions of barrels)	Gas Field Size * (Billions of cubic feet)
1	0.03125 to 0.0625	0.1875 to 0.375
2	0.0625 to 0.125	0.375 to 0.75
3	0.125 to 0.25	0.75 to 1.5
4	0.25 to 0.5	1.5 to 3
5	0.5 to 1	3 to 6
6	1 to 2	6 to 12
7	2 to 4	12 to 24
8	4 to 8	24 to 48
9	8 to 16	48 to 96
10	16 to 32	96 to 192

11	32 to 64	192 to 384
12	64 to 128	384 to 768
13	128 to 256	768 to 1,536
14	256 to 512	1,536 to 3,072
15	512 to 1,024	3,072 to 6,144
16	1,024 to 2,048	6,144 to 12,288
17	2,048 to 4,096	12,288 to 24,576
18	4,096 to 8,192	24,576 to 49,152
19	8,192 to 16,384	49,152 to 98,304
20	16, 384 to 32,768	98,304 to 196,608
21	32, 768 to 65,536	196,608 to 393,216
22	65,536 to 131,072	393,216 to 786,432
23	131,072 to 262,144	786,432 to 1,572,864

* Accumulations of sizes exactly on a class boundary are counted in the higher class.

Because the original data were from proprietary databases, some appropriate modifications were made for the version of the file released to the public. The field sizes were binned to obscure their precise sizes, but the sizes of the largest discovered fields were rounded to 1.38 times the lower limit of their bin (see Appendix 1). There were nine AUs (10080101, 10090102, 10160101, 11500301, 38170103, 60220101, 60340102, 60900103, and 80470201) that only had one discovered field. For these, ancillary data were not reported, and the totals for known oil, gas, natural gas liquids, and total petroleum (Geology worksheet columns W, Y, AA, and AC) are marked as "NA" (not available). None of the remaining data can thus be traced easily to any specific field.

Sizes of Discovered Fields

Sizes of discovered fields (except for Canada) were from the IHS Energy, Inc. Exploration and Production database (IHS Energy, Inc., 2004); production and reserve data used were as reported at the end of 2003. Sizes of fields were defined as the sum of cumulative production plus "2P" reserves (variables OIL_RECOVERABLE_PP_MMBBL, GAS_RECOVERABLE_PP_MMSCF, and COND_RECOVERABLE_PP_MMBBL). The reported "2P" (proved plus probable) reserves are generally considered estimates with a 50 percent probability of the actual values being larger. No estimates of reserve-growth potential were added, so the "2P" reserve estimates may be conservative. Because of the proprietary nature of the database, the raw oil or gas field sizes cannot be released but only summary field or pool counts per bin-size class.

For Canada, the data were from the Significant Oil and Gas Pools of Canada Data Base (NRG Associates, 1995). These data are given by oil and gas pool, rather than by field, but were considered more complete than the IHS Energy, Inc. data for Canada (IHS Energy, Inc., 2004). Otherwise, calculations were performed on the pool-size data for Canadian AUs in the same manner as those done to the field-size data. The Canadian pool-size data are included in our database for sake of completeness. It is the option of the user as to whether that data is appropriate to use as an analog along with field-size data.

The discovered oil and gas fields (pools for Canada) were allocated to those 246 AUs of the WPA 2000 that were quantitatively assessed. Each oil or gas field was allocated to only one AU, thus the counts per size class are all integers.

2000 World Petroleum Assessment

The WPA 2000 (U.S. Geological Survey World Energy Assessment Team, 2000) assessed oil and gas potential in 128 provinces outside the United States. However, no attempt was made in that effort to estimate potential in all areas; it included only those areas that already had significant production or that otherwise had some particular interest. Those assessments were based on geologic studies of 159 total petroleum systems (TPS), a TPS being defined as "a mappable entity encompassing genetically related petroleum that occurs in seeps, shows, and accumulations (discovered or undiscovered) that have been generated by a pod or by closely related pods of mature source rock, together with the essential mappable geologic elements (source, reservoir, seal, and overburden rocks) that controlled fundamental processes of generation, migration, entrapment, and preservation of petroleum" (Klett and others, 2000). A hierarchical structure was developed consisting of provinces that contained TPSs that, in turn, contained AUs. Many of the TPSs were contained within one province, but some crossed province boundaries and were assigned to the most appropriate province. Within each TPS, one or more AUs were identified, for a total of 270. Of these, 246 AUs (all conventional) in 150 of the TPSs were quantitatively assessed, and form the basis of our analog database.

The WPA 2000 assessment methodology included an estimation of numbers and sizes of undiscovered accumulations for each AU (Schmoker, and Klett, 2000). The methodology is defined in terms of an accumulation, which can include either an entire field or only those reservoirs in a field that lie within a certain stratigraphic interval. For the WPA 2000, entire fields were used (pools in western Canada). The uncertainty about the number of undiscovered fields was expressed as a triangular distribution; for our database, the mean value of that distribution was used as the total number of undiscovered fields. Because that is a statistical expectation, it is not necessarily an integer.

The size distribution of undiscovered fields was approximated by a shifted truncated lognormal distribution. By integration of this distribution, the proportions of undiscovered fields in each field-size class were determined (Attanasi and Charpentier, 2007). When those proportions were multiplied by the mean number of undiscovered fields, the result was the expected number of undiscovered fields per class. Again, this is a statistical expectation, and thus not necessarily an integer.

Calculations

The field-size distributions presented in the database are those of the entire natural population—that is, they include both discovered and undiscovered subpopulations. The basic counts of fields per size-class bin are recorded separately for discovered versus undiscovered fields in the database. Besides the basic counts of fields per size-class bin, a number of calculated parameters of the distributions are also given. Some parameters, such as the median size or the number of fields per 1,000 km^2, are dependent on the minimum size assessed. For convenience, these were calculated for both a minimum of 5 million barrels of oil equivalent (MMBOE) and for a minimum of 50 MMBOE.

The actual Visual Basic for Applications macros that were used in compiling the database are preserved as part of the WorldAnalogs.xls workbook. They are included in module 4, which can be accessed within Microsoft Excel by the Visual Basic editor. Modules 1 through 3 include the code for the tools (**Analog Search**, **Extend Selection**, **Analog Plot**, and **Analog Histogram**) programmed for use with the database.

Small-Field Extrapolation

In WPA 2000, estimates were made of numbers and sizes of undiscovered fields greater than a selected minimum field size that varied from AU to AU. This minimum field size ranged from 1 to 20 MMBOE, except for pool sizes within the Western Canada Sedimentary Basin, where minimum pool size was set at 0.5 MMBOE. Because some additional analyses require estimates of smaller field sizes, an extrapolation of the size distribution into smaller field sizes is helpful.

A similar analysis of small field sizes was performed for the 1995 National Assessment of U.S. oil and gas potential (Gautier and others, 1995). An estimate of “large” undiscovered fields (defined as 1 MMBOE or larger) was performed by an earlier version of the methodology used in the WPA 2000. An estimate of number and sizes of undiscovered fields between 31,250 BOE and 1 MMBOE was performed by a separate analysis (Root and Attanasi, 1993). This analog database used a methodology for small-field extrapolation similar to that used in 1995 (see Appendix 1).

Additional Calculations

The analog database was designed to be used as part of the USGS assessment procedures for conventional oil and gas resources (Schmoker and Klett, 2000). That methodology estimates numbers and sizes of undiscovered accumulations. The standard assessment input form is also documented in Schmoker and Klett (2000). Probability distributions are required for the numbers of undiscovered accumulations, the sizes of undiscovered accumulations, and for coproduct ratios (such as the gas-to-oil ratio in oil accumulations). Oil and gas accumulations are estimated separately.

The total number of fields and several parameters of the field-size distribution are dependent on the size of the smallest field considered for the assessment. Values dependent on the minimum field size are the median and maximum field sizes and the numbers of fields per 1,000 km^2. For purposes of future assessments, these values were calculated for minimum field sizes of 5 and 50 MMBOE. The variables calculated for minima of 5 and 50 MMBOE are marked in the analog dataset with names containing "> 5" and "> 50", respectively, even on those pages devoted to gas data in billions of cubic feet of gas (BCFG). Values for other minimum field sizes can be estimated based on the original binned data in the analog database.

Global Distributions

Except for the Geology worksheet, most columns of numerical variables have summary information at the bottom of the column. These data describe the "global" distribution of that variable; global, in this sense, refers to the entire set of 246 AUs in the analog database. Because the database is a biased set, preferentially containing the most successful provinces, this distribution should be characterized by larger field sizes and higher field densities than a true worldwide distribution that would also include non-productive and minimally productive areas. However, the amount of bias may be less for coproduct ratios and ancillary data, which may approximate more closely a true worldwide distribution. The global distribution is described by minimum, median, maximum, and mean values, as well as the number of AUs for which there is a value.

Use of Database

The analog database is meant to provide a source for analogs used in assessment of conventional oil and gas fields. Particularly in the case of frontier areas with few or no discovered fields, assessments must rely on comparisons to geologically similar areas that have been more maturely explored. In such frontier areas, other assessment methodologies, such as discovery process models, are unusable. Even in areas of moderately mature exploration, comparison to analogs can also be a useful check on estimates made by other methods. Although constructed with the USGS methodology in mind, the database can be used with other methodologies.

The foregoing is not to imply that resource assessment is the only use for our analog database. More than 95 percent of the world's discovered oil and gas resources as of the end of 1995 (excluding the United States) are in the 246 AUs included in the database, thus it offers additional opportunities for analysis of patterns of oil and gas occurrence.

The examples presented below are based on constructing sets of analogs, but analyses may be based on any number from one to all 246 AUs. Although some may prefer to search for a single "best" analog, the construction of a set of analogs better characterizes the range of uncertainty.

For a particular assessment, more than one set of analogs may be necessary. Numbers and sizes of fields are commonly constrained by the available trapping configurations, which in turn are related to the tectonic and sedimentologic characteristics of the area. Coproduct ratios, such as the gas-to-oil ratio in oil fields, relate more closely to source rock characteristics and thermal history. A different set of analogs may thus be appropriate for assessing the coproduct ratios.

Classification Variables

The first worksheet (Geology) of the analog database workbook includes 34 variables (referred to as classification variables) that allow the user to classify the AUs and thus select appropriate analogs (table 2). (Note: For those using the tab-delimited files, this is file "geology.tab.") The first six variables identify the assessment unit and place it into the hierarchical structure of province and total petroleum system. The other variables relate either to (1) province-scale geology, such as the tectonic regime or basin type; or (2) assessment-unit-scale geology, such as the source, reservoir, and trap characteristics. This geologic information was compiled as part of the WPA 2000 (U.S. Geological Survey World Energy Assessment Team, 2000), mostly from the text summaries describing each assessment unit.

Table 2. Classification variables in the world analog database and the allowed values for those variables with a limited range of values
[AU, assessment unit; TPS, total petroleum system; BBO, billion barrels of oil; TCFG, trillion cubic feet of gas; BBNGL, billion barrels of natural gas liquids; BBOE, billion barrels of oil equivalent]

Variable	Description	Values Allowed
AU_Code	USGS code for this assessment unit	
AU_Name	official USGS name for this assessment unit	
TPS_Code	USGS code for this total petroleum system	
TPS_Name	official USGS name for this total petroleum system	
Province Code	USGS code for this province	
Province Name	official USGS name for this province	
Structural Setting		Compressional
		Extensional
		Passive
		Unknown
Crustal System		Continental
		Oceanic
		Transitional
		Uncertain
Architecture		Backarc
		Craton interior
		Deltas
		Foreland
		Platform
		Rift, sag
		Rifted passive margin
		Small ocean basin
		Strike-slip systems
Trap System (Major)		Basement-involved block structures
		Compressional anticlines, folds, thrusts
		Diapir-related structures
		Extensional grabens and other structures related to normal faulting
		Gravity-induced growth faults
		Local uplifts of uncertain origin

	Paleogeomorphic Salt-induced structures Stratigraphic undeformed Transtensional and transpressional
Depositional System	Carbonate shelf Carbonate shelf margin, reefs Continental clastics Deepwater carbonates Paralic clastics Slope, clinoforms, turbidites
Source Rock Depositional Environment	Coaly Deep marine Lacustrine Shallow marine
Kerogen Type	No source rocks Type II Type III Types I and II Types II and III Types I, II, and III
Source Type	Anoxic/disoxic Carbonaceous mudstones, coal Disoxic/oxic Lacustrine anoxic/dioxic Marine anoxic Marine disoxic Terrestrial
Source Rock Qualifier	Assumed Known Probable
Status	Hypothetical Established Frontier
Specific Reservoir Rock Age	*various*
General Reservoir Rock Age	*various*
Reservoir Rock Lithology	Carbonates

		Siliclastics Siliciclastics and Carbonates Siliciclastics, Carbonates, and Other
Reservoir Rock Depositional Environment		Deep marine Nonmarine Nonmarine to Marine Shallow Marine Shallow to Deep Marine
Seal Rock Lithology		Evaporites Evaporites and Shale Shale
Trap Type	general trap type	Stratigraphic Structural Structural and Stratigraphic
Known Oil (BBO)	cumulative production plus reserves of oil	
Undiscovered Oil (Mean, BBO)	mean USGS estimate of oil in undiscovered fields	
Known Gas (TCFG)	cumulative production plus reserves of natural gas	
Undiscovered Gas (Mean, TCFG)	mean USGS estimate of natural gas in undiscovered fields	
Known NGL (BBNGL)	cumulative production plus reserves of NGL	
Undiscovered NGL (Mean, BBNGL)	mean USGS estimate of NGL in undiscovered fields	
Known Total Petroleum (BBOE)	cumulative production plus reserves of petroleum in BOE	
Undiscovered Total Petroleum (Mean, BBOE)	mean USGS estimate of petroleum in undiscovered fields (in BOE)	
Specific Source Rock Age		*various*
General Source Rock Age		*various*
Age of Maturation		*various*
Migration Distance		Short Short to long

Methods of Searching

For those using the Microsoft Excel version of the database, tools have been developed for easier use. These are located under the **Analog** menu in the menu bar. Two of these, **Analog Search** and **Extend Selection**, are aids to searching and building analog sets for further analysis.

Analog Search, found under the **Analog** menu in the menu bar, is a tool for simple searches on several of the most important geologic variables. An initial form asks whether to search within the current selection. This allows multi-level (Boolean AND) searches. Next, the tool provides a pop-up menu of the most important classification variables (fig. 1). Following selection of one of the variables, another pop-up menu provides a list of the possible values for that variable (fig. 2). The results of the search are highlighted on all the worksheets of the database.

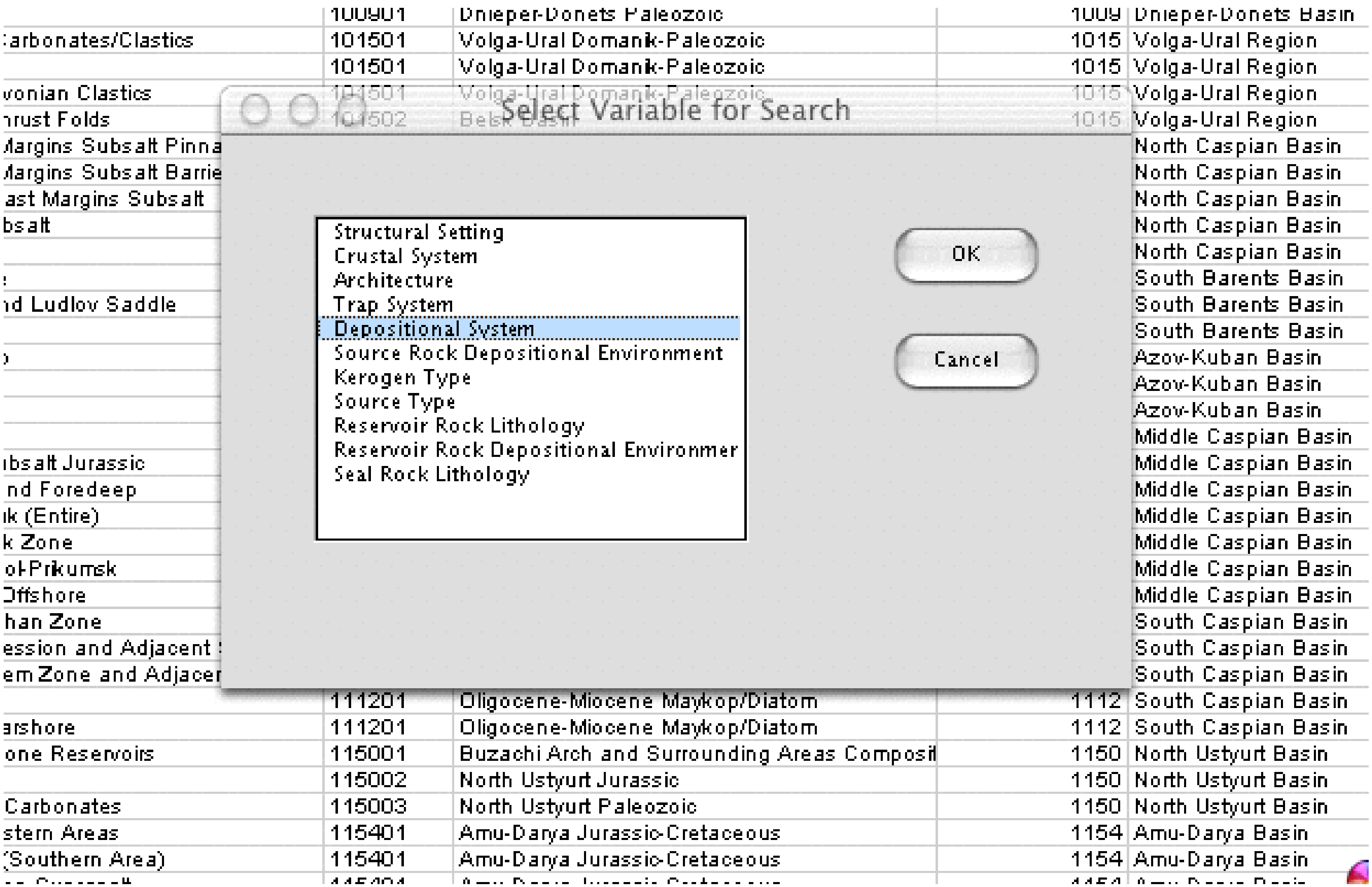

Figure 1. Select variables form for the **Analog Search** tool.

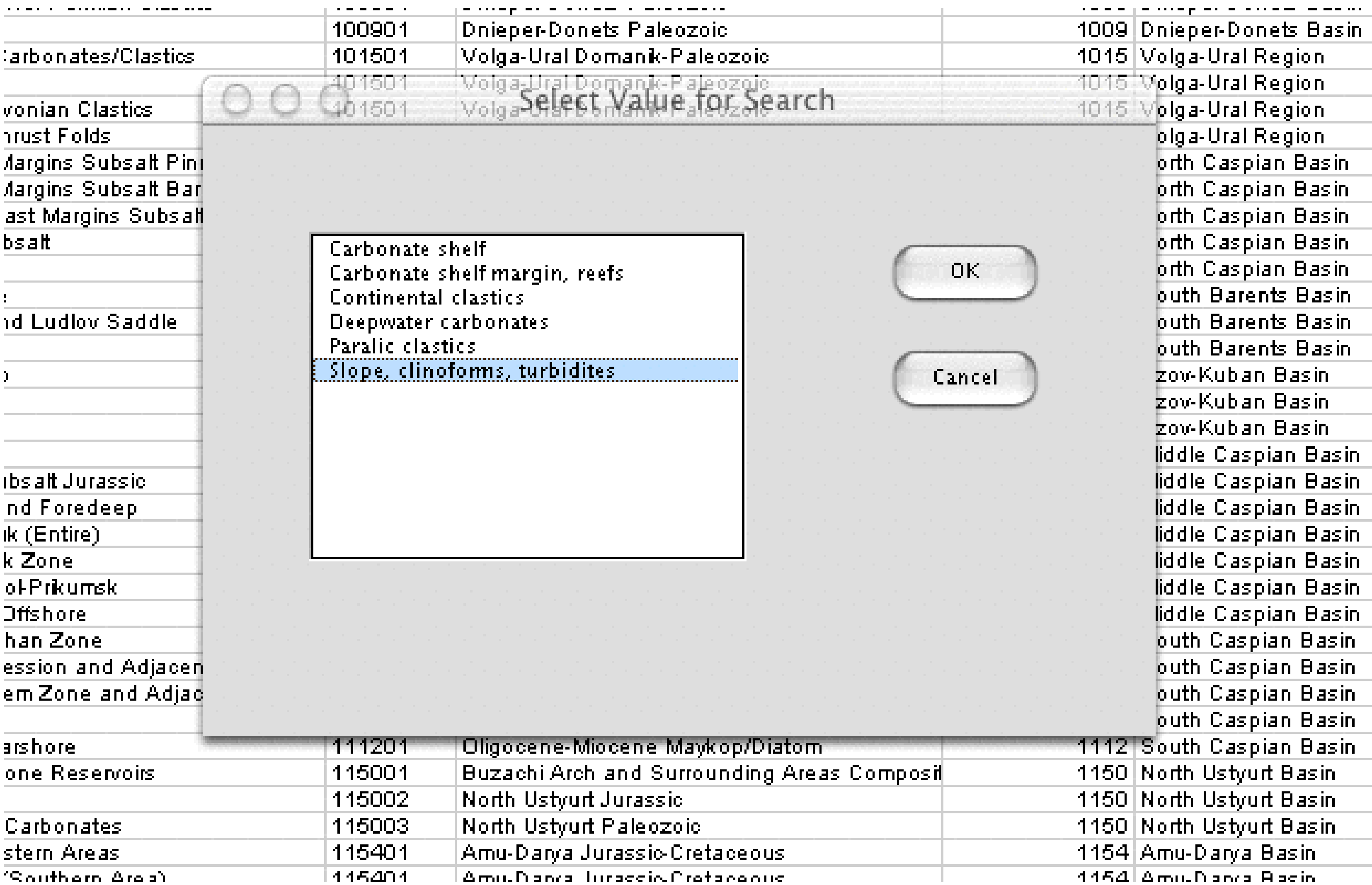

Figure 2. Select values form for the **Analog Search** tool.

Microsoft Excel users can use the other search capabilities built into the program, such as the **Find** or **Filter** commands or manually selecting with the mouse. The **Extend Selection** command can be used in conjunction with these search methods. **Extend Selection** finds all selected cells in the Geology worksheet and extends the selection such that the entire row is selected for each row with a selected cell. The selection is further extended to all the other worksheets in the database. As an example, one could manually select all the cells in column Specific Source Rock Age (column AE) on the Geology worksheet that had value "Devonian". Running **Extend Selection** would thus select the entire row for each of those chosen AUs on all the worksheets. Microsoft Excel users who create selections on the Geology worksheet other than by **Analog Search** and wish to use the additional graphing tools **Analog Plot** and **Analog Histogram** should run **Extend Selection** before using the graphing tools.

For users of other spreadsheet programs, searching for a set of analogs can be accomplished by use of tools specific to the user's spreadsheet program. The tools **Analog Search** and **Extend Selection** are not available outside the Microsoft Excel version of the analog database.

Utility Variables

Utility variables are so named because they can be used for analysis once an analog set has been selected. Most of the major utility variables are on the Oil, Gas, and BOE worksheets. Additional utility variables are on the Oil Bins, Gas Bins, BOE Bins, and Ancillary worksheets.

Variables Related to Number of Fields

The Oil Bins, Gas Bins, and BOE Bins worksheets contain the basic numbers of discovered and undiscovered fields in each size class. Most users will not want to use such basic data, but instead will use some of the summary data variables. These worksheets include totals of "Number of Discovered," "Number of Undiscovered," as well as "Number > 5," and "Number > 50." The latter two variables give the number of fields greater than the minima of 5 and 50 MMBOE respectively.

Because neither 5 nor 50 MMBOE are exact bin limits, the numbers of fields greater than these minimum sizes had to be estimated. For number of fields of at least 5 MMBOE, the counts from bin sizes 8 MMBOE and larger are added to two-thirds the number in the 4 to 8 MMBOE size class. For number of fields of at least 50 MMBOE, the counts from bin sizes 64 MMBOE and larger are added to one-third the number in the 32 to 64 MMBOE size class.

Perhaps more useful for calculating numbers of undiscovered fields in an area to be assessed are several variables listed on the Oil, Gas, and BOE worksheets. Here the variables "Number / 1000 km^2 for > 5," and "Number / 1000 km^2 for > 50" give field densities per 1,000 km^2 rather than raw numbers of fields.

Variables Related to Field Size Distributions

The raw field-size bin counts for each analog AU are given in the Oil Bins, Gas Bins, and BOE Bins worksheets. More useful to the assessor are several variables in the Oil, Gas, and BOE worksheets: “Median of > 5,” “Median of > 50,” “Maximum of > 5,” and “Maximum of > 50”, which give the medians and maxima of the size distributions of those fields greater than the minima of 5 and 50 MMBOE, respectively.

Median sizes are approximated from the bin count data. The bin counts were used to identify the bin that included the median field size, which was calculated within that bin by interpolation.

The variables “Maximum of > 5” and “Maximum of > 50” are defined as the maximum size of discovered fields (column BB in the bins worksheets) or the expected (mean) size of the largest undiscovered field (column BC in the bins worksheets), whichever is larger. In order to preserve the proprietary field-size data, the maximum size of discovered fields is rounded to 1.38 times the lower limit of the bin in which it resides. The variables “Maximum of > 5” and “Maximum of > 50” are different only in the case where the maximum size is between 5 and 50 MMBOE. In that case there would be a value for “Maximum of > 5” but not for “Maximum of > 50.”

Variables Related to Maturity of Exploration

Four utility variables are provided as indices of exploration maturity. These are “Discovered % by Number for > 5,” “Discovered % by Number for > 50,” “Discovered % by Volume for > 5,” and “Discovered % by Volume for > 50,” all of which are listed on the Oil, Gas, and BOE worksheets. These maturity measures are used by the **Analog Plot** tool described below.

Variables Related to Coproduct Ratios

The methodology used in the WPA 2000 treated fields as being either oil or gas (Schmoker and Klett, 2000). Oil fields are differentiated from gas fields in that they have a gas-to-oil ratio (based on cumulative production plus reserves) of less than 20 thousand cubic feet per barrel. Field sizes are expressed as volumes of oil in oil fields and volumes of gas in gas fields, a procedure that was followed in the analog database. Other petroleum products are handled by the use of coproduct ratios. In oil fields, associated and dissolved gas volumes are calculated using a gas-to-oil ratio and natural gas liquids (NGL) volumes are calculated using a NGL-to-gas ratio. In gas fields, total liquids are calculated using a liquids-to-gas ratio. These three coproduct ratios are provided in the analog database—the two coproduct ratios for oil fields are on the Oil worksheet, the coproduct ratio for gas fields is on the Gas worksheet, and all three coproduct ratios are on the BOE worksheet.

Data provided for coproduct ratios in the analog database is based on production and reserves data from discovered fields drawn from the IHS Energy, Inc.(2004), and NRG Associates (1995) databases. Coproduct ratios for individual fields were calculated from the cumulative production plus reserves of each product. Each field was allocated to a particular AU. The analog database contains five variables describing the distribution of each coproduct ratio—minimum, median, maximum, and mean, as well as the number of fields that had a calculated coproduct ratio of that type. The number of fields gives an indication of the confidence in the other measures. A small number of fields indicates poorer sampling, and the distribution is thus known with less confidence.

Variables Related to Ancillary Data

Ancillary data are not used to calculate volumes of undiscovered resources, but rather give information affecting the technical and economic recoverability of the resource. Some ancillary variables describe the quality of the oil and gas fluids: oil viscosity (in centipoises and in centistokes), API gravity of the oil, sulfur content of the oil in percent, and non-hydrocarbon fractions of the gas (percentages of hydrogen sulfide, carbon dioxide, and nitrogen). The ancillary data also include drilling depths and, for the offshore, water depths (in feet). As with the coproduct ratios, the data in the analog database comes from discovered fields only. The ancillary data are presented on worksheet Ancillary.

The ancillary data are presented in a manner similar to that of the coproduct data. There are five variables describing the distribution of each ancillary variable—minimum, median, maximum, and mean, as well as the number of fields that had a ancillary value of that type. The number of fields gives an indication of the confidence in the other measures. A small number of fields indicates poorer sampling, and the distribution is thus known with less confidence.

Data Analysis Tools

Several tools added to the Microsoft Excel version of the analog database can be accessed through the menu bar under **Analog**; to use, the analog set should first be selected on the Geology worksheet. If the selection is made manually, the program **Extend Selection** (under the **Analog** menu in the menu bar) should be run before using the data analysis tools.

Distributions for Small Analog Sets

For small analog sets (those with fewer than about 20 AUs), an additional graphing tool has been added to the Microsoft Excel version of the database. This tool, **Analog Plot**, can be found under the **Analog** menu in the menu bar. The resulting graph shows the distribution of a selected utility variable for the analog set selected in the Geology worksheet.

Each analog AU is represented on the graph by a vertical bar placed at the value of the utility variable (fig. 3). The length of the shaded portion of the bar is a maturity measure showing the percent of resource already discovered. The user may choose either percent by volume (suggested) or percent by number of fields to show maturity. Values of the utility variable are more credible for those AUs with higher maturity. Each bar is labeled by the AU name.

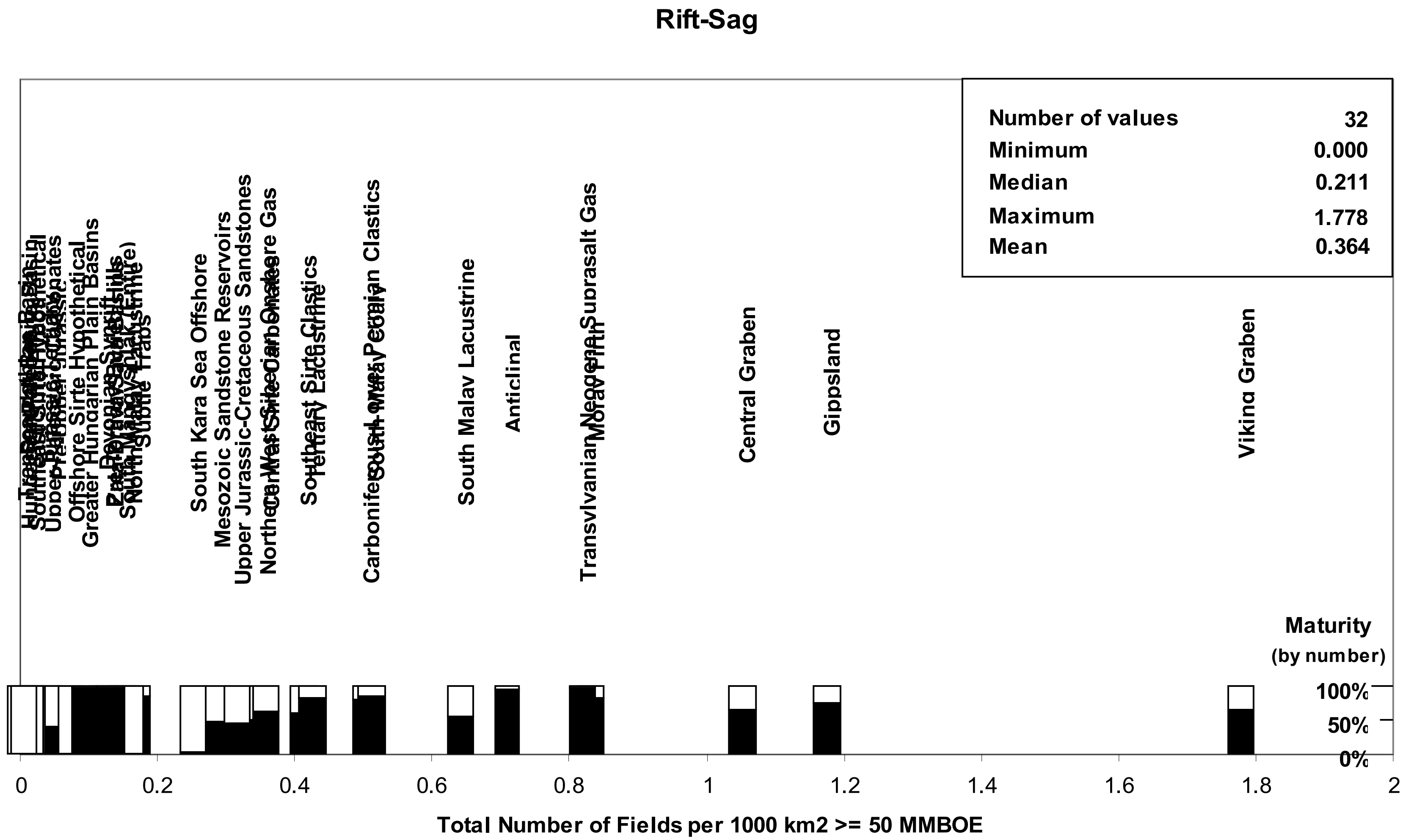

Figure 3. Example of chart generated by the **Analog Plot** tool. MMBOE, million barrels of oil equivalent.

A large number of analogs with similar values of the utility variable in question can lead to an unreadable chart. In such a case, one could remove AUs with extreme values from the analog set, thus changing the X-axis scale. Alternatively, one could use the **Analog Histogram** tool.

In constructing the plot (as in fig. 3), data are copied to worksheet Chart Data. Column A contains the values of the utility variable in ascending sorted order. Column B contains the maturity measures for the data in column A. The distribution of the utility variable data in column A is summarized in column K, where the number of values, minimum, median, maximum, and mean are given. These data are replaced when a new chart is constructed either by **Analog Plot** or by **Analog Histogram**.

Distributions for Large Analog Sets

For large analog sets (those with more than about 20 AUs), another graphing tool has been added to the Microsoft Excel version of the database. This tool, **Analog Histogram**, can also be found under the **Analog** menu in the menu bar. The resulting graph shows the distribution of a selected utility variable for the analog set selected in the Geology worksheet. With this tool a distribution histogram is constructed but no maturity measure or identification of AU names is given (fig. 4).

Bin	*Frequency*
0.0	96
0.9	43
1.8	43
2.7	23
3.6	9
4.5	11
5.4	8
6.3	4
7.2	0
8.1	2
9.0	0
9.9	2
10.8	1
11.7	0
12.6	2
13.5	1
14.4	0
15.3	0
16.2	0
17.1	1
18.0	0
18.9	0
More	0
Total	246

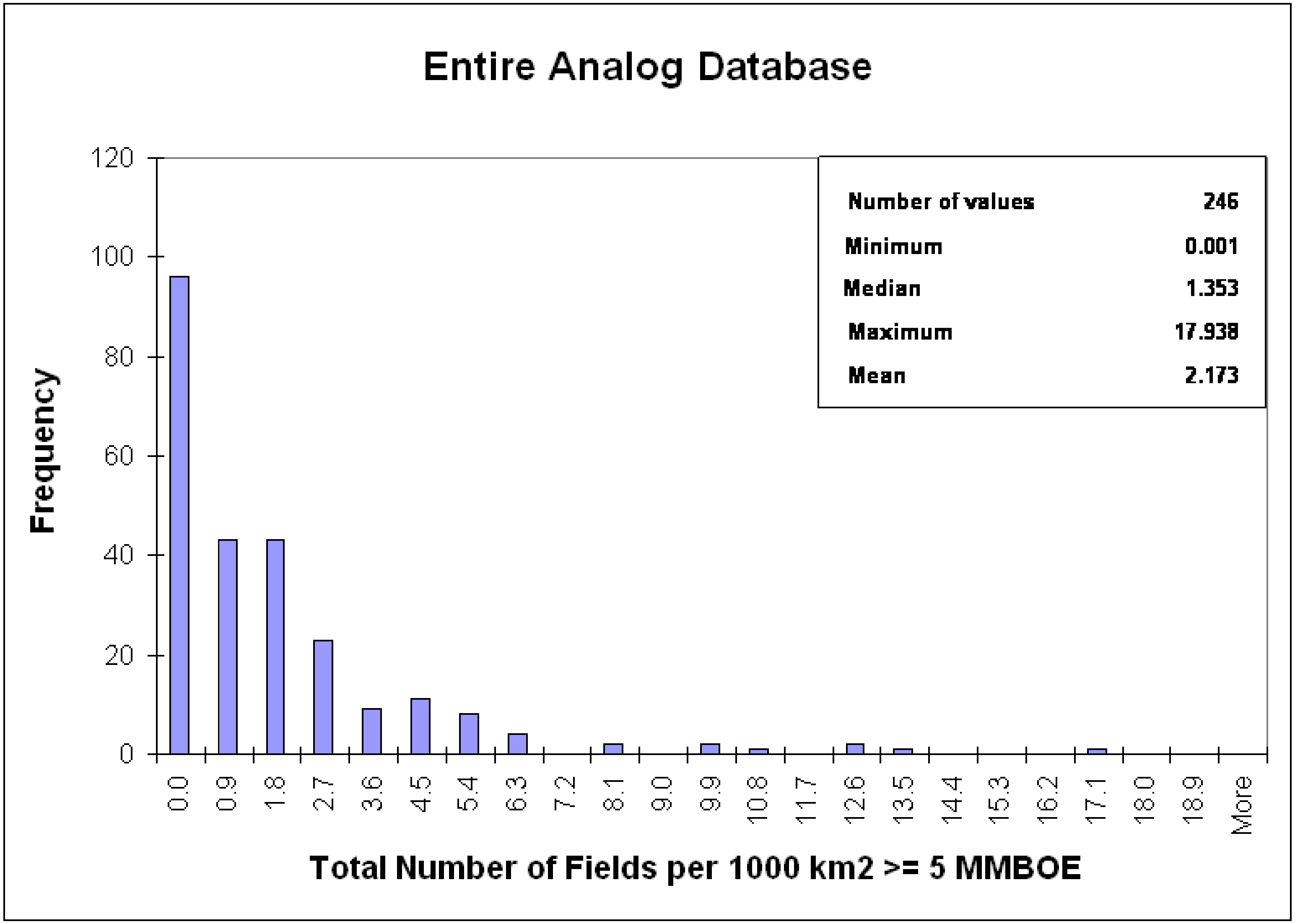

Figure 4. Example of chart generated by the **Analog Histogram** tool. MMBOE, million barrels of oil equivalent.

In constructing the histogram, data are copied to worksheet Chart Data. Column A contains the values of the utility variable in ascending sorted order. Column H contains the histogram class boundaries. The distribution of the utility variable data in column A is summarized in column K, where the number of values, minimum, median, maximum, and mean are given. These data are replaced when a new chart is constructed either by **Analog Plot** or by **Analog Histogram**.

Example of Use with USGS Assessment Methodology

The following example is based on the standard 2000 USGS methodology for assessing conventional oil and gas resources (Schmoker and Klett, 2000). The assessment input form (fig. 5) requires distributions for numbers and sizes of oil and gas accumulations and the associated coproduct ratios

Figure 5. Seventh approximation input data form, version 6.

SEVENTH APPROXIMATION
DATA FORM FOR CONVENTIONAL ASSESSMENT UNITS (Version 6, 9 April 2003)

IDENTIFICATION INFORMATION

Assessment Geologist: ____________ Date: ______
Region: ____________ Number: ______
Province: ____________ Number: ______
Total Petroleum System: ____________ Number: ______
Assessment Unit: ____________ Number: ______
Based on Data as of: ____________
Notes from Assessor: ____________

CHARACTERISTICS OF ASSESSMENT UNIT

Oil (<20,000 cfg/bo overall) **or** Gas (≥20,000 cfg/bo overall): ______

What is the minimum accumulation size? ______ mmboe grown
(the smallest accumulation that has potential to be added to reserves)

No. of discovered accumulations exceeding minimum size: Oil: ______ Gas: ______
Established (>13 accums.) ______ Frontier (1-13 accums.) ______ Hypothetical (no accums.) ______

Median size (grown) of discovered oil accumulations (mmbo):
1st 3rd ______ 2nd 3rd ______ 3rd 3rd ______
Median size (grown) of discovered gas accumulations (bcfg):
1st 3rd ______ 2nd 3rd ______ 3rd 3rd ______

Assessment-Unit Probabilities:

Attribute	Probability of occurrence (0-1.0)
1. **CHARGE:** Adequate petroleum charge for an undiscovered accum. ≥ minimum size:	______
2. **ROCKS:** Adequate reservoirs, traps, and seals for an undiscovered accum. ≥ minimum size:	______
3. **TIMING OF GEOLOGIC EVENTS:** Favorable timing for an undiscovered accum. ≥ minimum size:	______

Assessment-Unit GEOLOGIC Probability (Product of 1, 2, and 3): ______

4. **ACCESSIBILITY:** Adequate location to allow exploration for an undiscovered accumulation ≥ minimum size: ______

UNDISCOVERED ACCUMULATIONS

No. of Undiscovered Accumulations: How many undiscovered accums. exist that are ≥ min. size?:
(uncertainty of fixed but unknown values)

Oil Accumulations:	minimum (>0) ______	mode ______	maximum ______
Gas Accumulations:	minimum (>0) ______	mode ______	maximum ______

Sizes of Undiscovered Accumulations: What are the sizes (**grown**) of the above accums?:
(variations in the sizes of undiscovered accumulations)

Oil in Oil Accumulations (mmbo):	minimum ______	median ______	maximum ______
Gas in Gas Accumulations (bcfg):	minimum ______	median ______	maximum ______

Page 1

Assessment Unit (name, no.) ____________________

AVERAGE RATIOS FOR UNDISCOVERED ACCUMS., TO ASSESS COPRODUCTS

(uncertainty of fixed but unknown values)

Oil Accumulations:	minimum	mode	maximum
Gas/oil ratio (cfg/bo):			
NGL/gas ratio (bngl/mmcfg):			

Gas Accumulations:	minimum	mode	maximum
Liquids/gas ratio (bliq/mmcfg):			
Oil/gas ratio (bo/mmcfg):			

SELECTED ANCILLARY DATA FOR UNDISCOVERED ACCUMULATIONS

(variations in the properties of undiscovered accumulations)

Oil Accumulations:	minimum	mode	maximum
API gravity (degrees):			
Sulfur content of oil (%):			
Depth (m) of water (if applicable):			

	minimum	F75	mode	F25	maximum
Drilling Depth (m):					

Gas Accumulations:	minimum	mode	maximum
Inert gas content (%):			
CO_2 content (%):			
Hydrogen-sulfide content (%):			
Depth (m) of water (if applicable):			

	minimum	F75	mode	F25	maximum
Drilling Depth (m):					

Selection of Analogs

The example AU is in a deep-water environment with possible turbidite reservoirs; its area is 100,000 km^2. No fields have yet been discovered in this area. A minimum assessed size of 5 MMBOE will be used.

In order to choose a suitable set of analogs, a search was conducted using the **Analog Search** tool, within which the Depositional System variable (fig. 1) was searched for values that included “Slope, clinoforms, turbidites” (fig. 2). This yielded a set of 44 potential analogs (table 3). Further examination could refine this analog set to a smaller number of analogs that are geologically more similar to the area being assessed, but with some possible loss of insight into the uncertainty.

Table 3. Analog set for the numbers and sizes of fields used in text example (based on a search for the value "Slope, clinoforms, and turbidites" within the Depositional System variable)

Assessment Unit Number and Name	
11120104	Central Offshore
11120105	Iran Onshore/Nearshore
11740101	Upper Jurassic-Cretaceous Sandstones
20160201	Natih-Fiqa Structural/Stratigraphic
37010102	Brunei-Sabah Turbidites
38170102	Kutei Basin Turbidites
39100201	Petrel
39100301	Malita
39130101	Late Jurassic/Early Cretaceous-Mesozoic
39480101	Dingo-Mungaroo/Barrow
39480201	Locker-Mungaroo/Barrow
40250103	Central Graben
40470101	Foreland Basin
40470201	Deformed Belt
40480101	Greater Hungarian Plain Basins
40480201	Zala-Drava-Sava Basins
40480601	Hungarian Paleogene Basin
40600101	Neogene Flysch Gas
60210101	Late Cretaceous-Tertiary Turbidites
60220101	Amazon Delta and Submarine Fan
60290102	Late Cretaceous-Tertiary Deep-Water Sandstones
60340102	Late Cretaceous-Tertiary Slide Blocks and Turbidites
60340103	Abrolhos Sub-Volcanic Structures
60350101	Late Cretaceous-Tertiary Turbidites
60350103	Salt Dome Province Tertiary Sandstones
60360102	Salt-Structured Deep-Water Sandstones
60370101	Pelotas Platform and Basin
60450101	Sub-Andean Fold and Thrust Belt

60450102	Foreland Basins
60550103	Dorsal de Neuquen Structure
60980202	Orinoco Delta and Offshore
61030101	Carupano Basin Gas
61070101	Inner Forearc Deformation Belt
71920102	Akata Reservoirs
72030201	Gabon Suprasalt
72030302	Central Congo Turbidites
72030401	Cuanza-Namibe
73030101	Offshore
80420102	Indus Fan
80430102	Eocene-Miocene Cambay Deltaic
80470201	Western Shelf and Slope
80470301	Central Basin
80470302	Eastern Fold Belt
80480102	Irrawaddy-Andaman

Numbers of Fields

To capture the uncertainty in number of undiscovered fields, the density of fields per 1,000 km^2 was examined in the analog set. A minimum size of 5 MMBOE was used and the density of total fields was examined. Because of the relatively large number of analogs (more than 20), the **Analog Histogram** tool was used instead of the **Analog Plot** tool. Based on this set of 44 analogs, the field density ranges to about 10 fields per 1,000 km^2, but most values are from 0 to 2 fields per 1,000 km^2 (fig. 6).

Bin	*Frequency*
0.0	8
0.5	9
1.0	8
1.5	4
2.0	2
2.5	3
3.0	0
3.5	2
4.0	1
4.5	2
5.0	2
5.5	1
6.0	1
6.5	0
7.0	0
7.5	0
8.0	0
8.5	0
9.0	0
9.5	0
10.0	1
10.5	0
More	0
Total	44

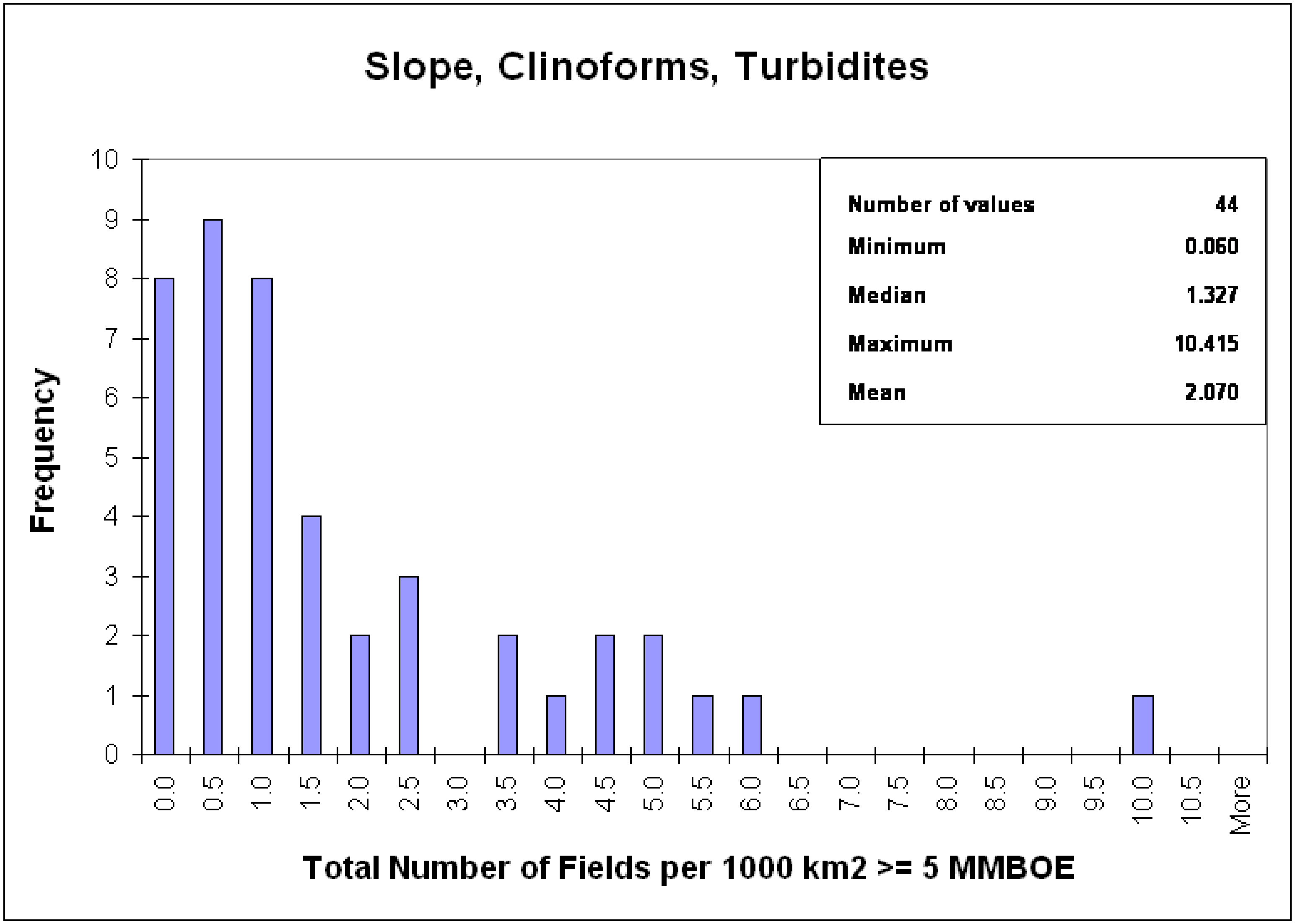

Figure 6. Plot of field density per 1,000 km^2 for example given in text. MMBOE, million barrels of oil equivalent.

Values appropriate to the input form (fig. 5) should reflect these densities. A possible modal number of oil plus gas fields could be approximately 1 field per 1,000 km^2 times the area of the AU, or 100 total fields. Dividing this number of fields between oil and gas fields is dependent on an understanding of source rocks and thermal history. Any previously discovered fields would need to be subtracted from these numbers. Similarly, the maximum number of oil plus gas fields could be as high as 10 fields per 1,000 km^2 times the area of the AU, or 1,000 total fields (to be divided between oil and gas fields).

Sizes of Fields

In a similar manner, uncertainty in sizes of undiscovered fields requires examination of the median and maximum field sizes in the analog set. A minimum size of 5 MMBOE was again used and the sizes of total fields were examined. Because of the relatively large number of analogs (more than 20), the **Analog Histogram** tool was used instead of the **Analog Plot** tool. The median field sizes range from 8 to more than 25 MMBOE (fig. 7), with most values about 14 MMBOE. One of the values for maximum field size exceeds 22,000 MMBOE (fig. 8), but all the others are less than 3 MMBOE.

Bin	Frequency
0	0
2	0
4	0
6	0
8	1
10	2
12	19
14	12
16	3
18	3
20	3
22	0
24	0
26	1
28	0
30	0
32	0
34	0
36	0
38	0
40	0
42	0
More	0
Total	44

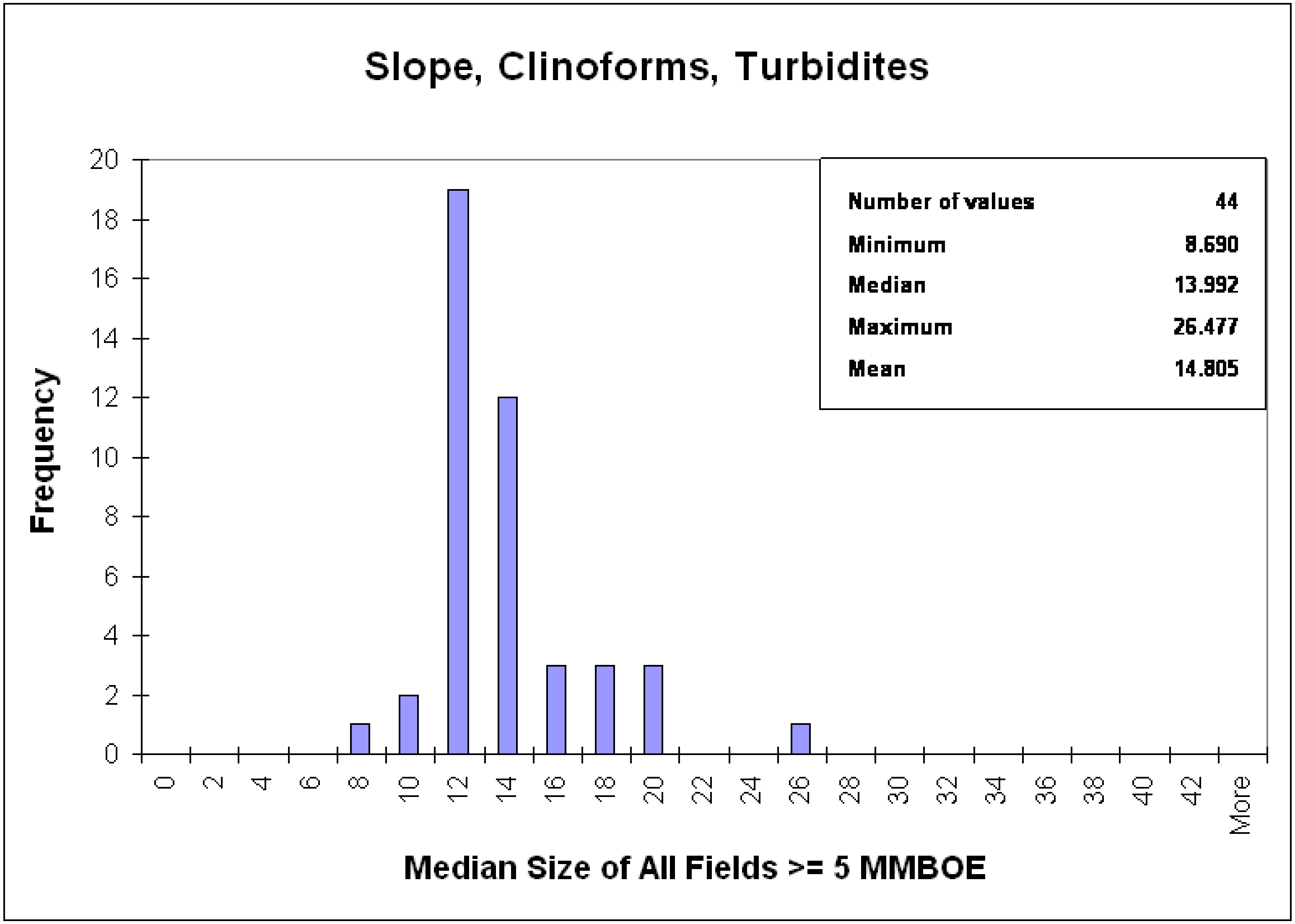

Figure 7. Plot of median field size for example given in text. MMBOE, million barrels of oil equivalent.

Bin	Frequency
0	30
1190	8
2380	5
3570	0
4760	0
5950	0
7140	0
8330	0
9520	0
10710	0
11900	0
13090	0
14280	0
15470	0
16660	0
17850	0
19040	0
20230	0
21420	1
22610	0
23800	0
24990	0
More	0
Total	44

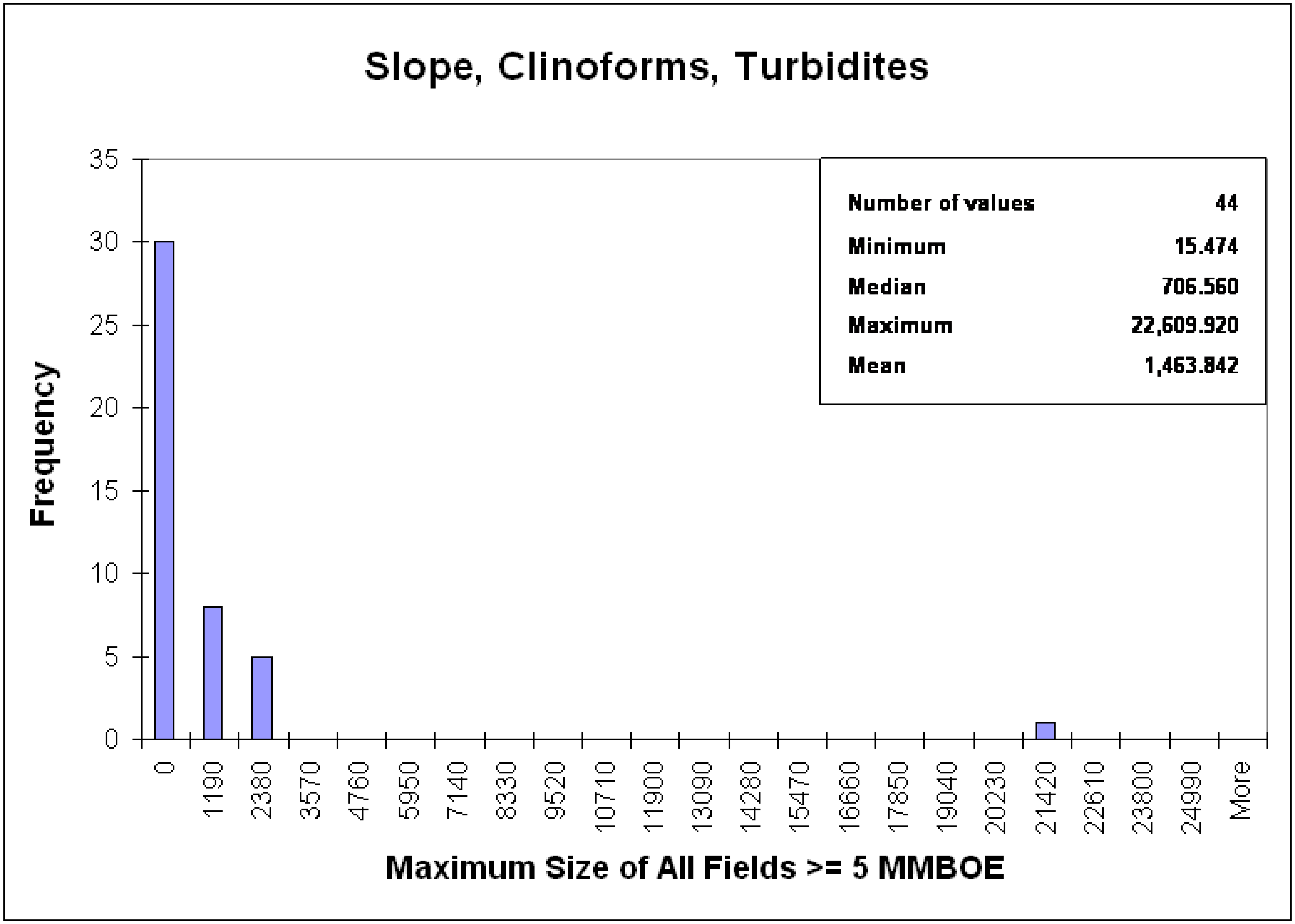

Figure 8. Plot of maximum field size for example given in text. MMBOE, million barrels of oil equivalent.

The values appropriate to the input form (fig. 5) should reflect these densities. The minimum size was defined in this example as 5 MMBOE**.** The median size could be about 14 MMBOE. Because the data for maximum field size include an outlier value much larger than any of the others, the analog AU containing that value should be closely examined for geologic similarity to the others in the analog set. If this AU cannot be reasonably eliminated from the analog set, the high uncertainty suggests a maximum size in the 25,000 MMBOE range. Also, any previously discovered fields could reduce the maximum undiscovered field size.

Coproduct Ratios

The coproduct ratios relate to a different set of geologic characteristics than the numbers and sizes of fields, so a different set of analogs is required. In this example case, we chose analogs that have lacustrine source rock depositional environments and types I and II kerogen (table 4). Distributions of the three coproduct ratios are given in figures 9, 10, and 11. The smaller (22) size of the analog set allows use of the **Analog Plot** tool. Although this is a small analog set, a distribution of values for each of the coproduct ratios is shown by the plots.

Table 4. Analog set for the coproduct ratios used in text example (based on a search for the value "Lacustrine" within the Source Rock Depositional System variable and the value "Types I and II" within the Kerogen Type variable)

Assessment Unit Number and Name	
11740201	Pre-Upper Jurassic
31150101	Upper Paleozoic/Lower Mesozoic Nonmarine Coarse Clastics
31270101	Tertiary Lacustrine
31270102	Pre-Tertiary Buried Hills
31280101	Jurassic/Triassic Fluvial and Lacustrine Sandstone
31420201	Jurassic Lacustrine
31440101	Subtle Traps
31440102	Anticlinal
31540102	Kuche (Northern) Foldbelt
37030101	South Malay Lacustrine
37030102	North Malay Lacustrine
38080101	Pematang/Sihapas Siliciclastics
38240101	Sunda/Asri
38280101	South Sumatra
60290101	Western Pre-Aptian Reservoirs
60350101	Late Cretaceous-Tertiary Turbidites
60350102	Cretaceous Carbonates
60350103	Salt Dome Province Tertiary Sandstones
60600101	North Falklands Basin
60600201	South Falklands Basin
72030101	Gabon Subsalt
72030301	Central Congo Delta and Carbonate Platform

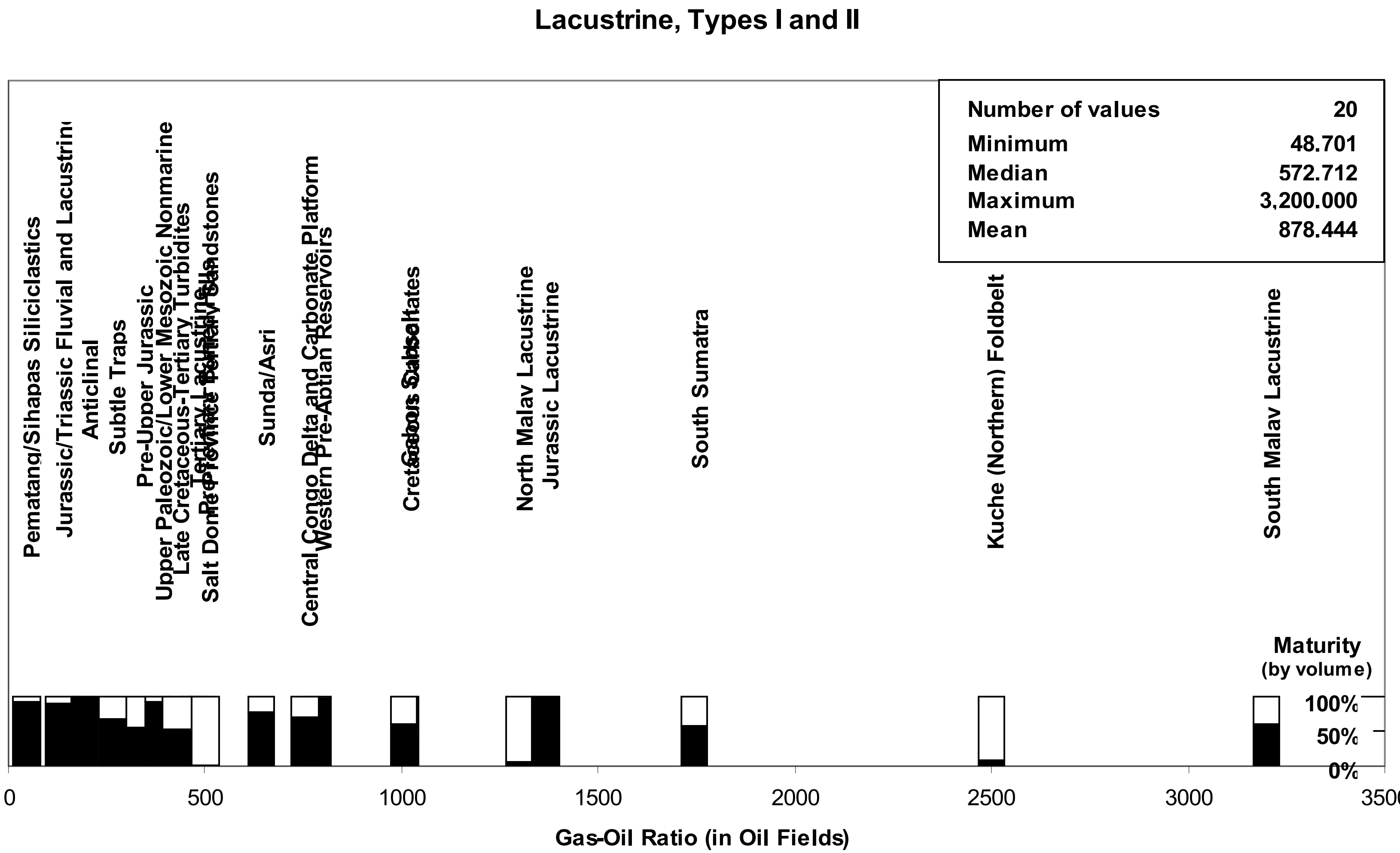

Figure 9. Plot of gas-oil ratio in oil fields for example given in text.

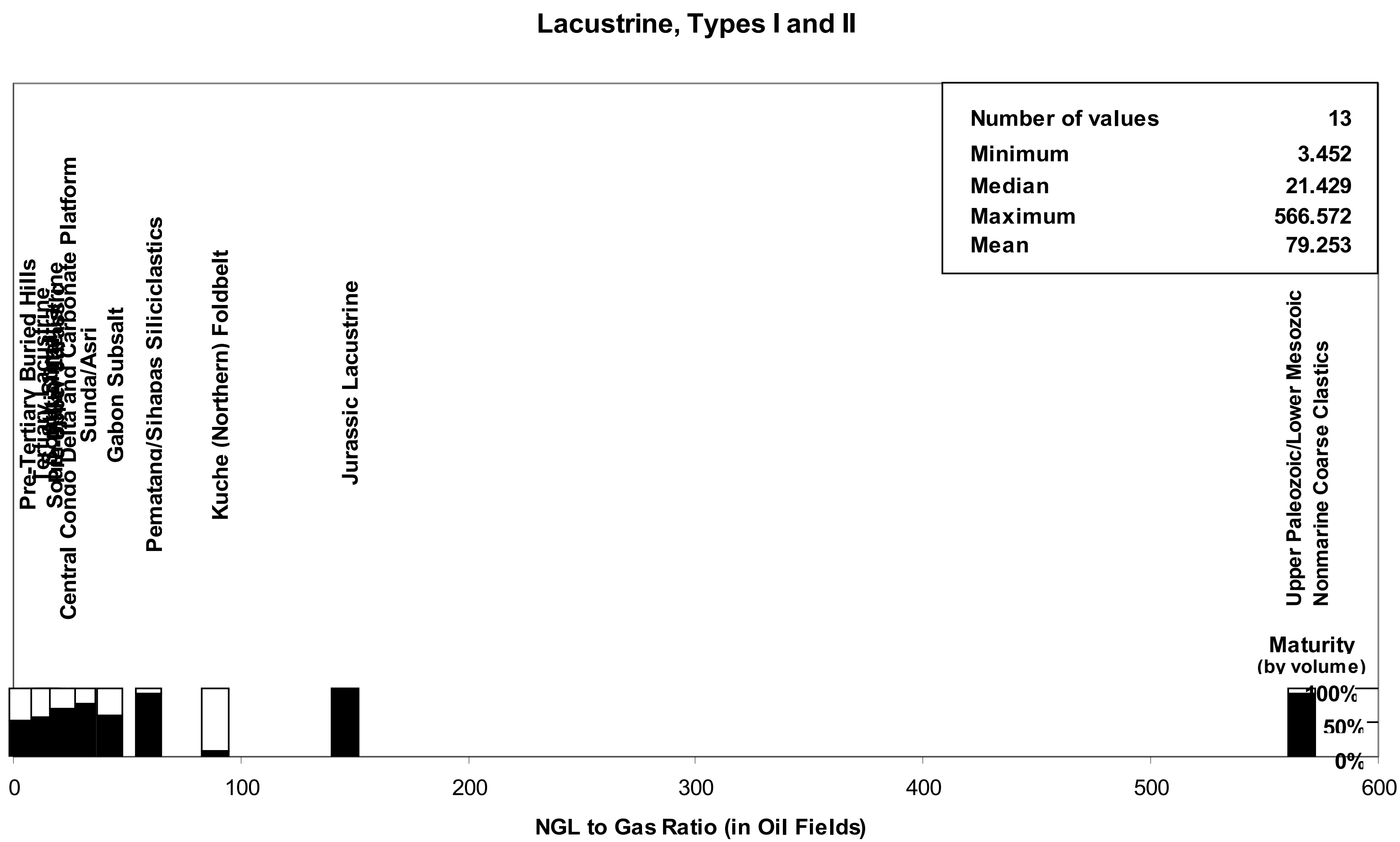

Figure 10. Plot of natural gas liquids (NGL) to gas ratio in oil fields for example given in text.

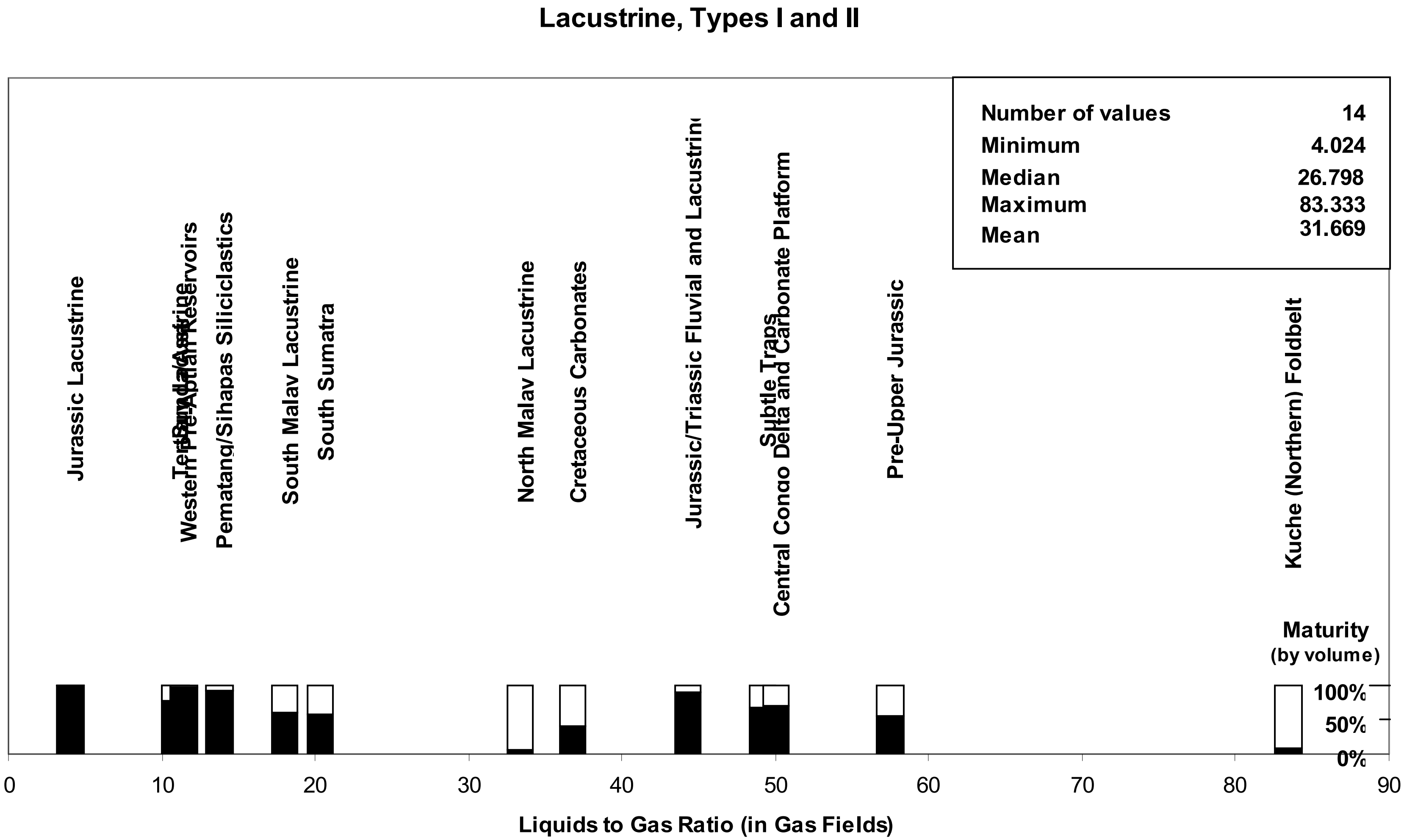

Figure 11. Plot of liquids to gas ratio in gas fields for example given in text.

The input form (fig. 5) requires an estimate of uncertainty of the mean values for each ratio. The calculated mean values for each coproduct ratio in this analog set are given in the text box. These means might be modified after considering whether to include some of the outlier values, such as the very high value for NGL to gas ratio (fig. 10).

References

Attanasi, E.D., and Charpentier, R.R., 2002, Comparison of two probability distributions used to model sizes of undiscovered oil and gas accumulations—Does the tail wag the assessment?: Mathematical Geology, v. 34, no. 6, p. 767-777.

Attanasi, E.D., and Charpentier, R.R., 2007, A program for partitioning shifted truncated lognormal distributions into size-class bins: U.S. Geological Survey Open-File Report 2007-1260, 16 p.

Drew, L.J., 1990, Oil and gas forecasting—Reflections of a petroleum geologist: New York, Oxford University Press, 252 p.

Drew, L.J., Schuenemeyer, J.H., and Bawiec, W.J., 1982, Estimation of the future rates of oil and gas discoveries in the western Gulf of Mexico: U.S. Geological Survey Professional Paper 1252, 26 p.

Gautier, D.L., Dolton, G.L., Takahashi, K.I., and Varnes, K.L., eds., 1995 National assessment of United States oil and gas resources—Results, methodology, and supporting data: U.S. Geological Survey Digital Data Series 30, one CD-ROM.

Houghton, J.C., 1988, Use of the truncated shifted Pareto distribution in assessing size distribution of oil and gas fields: Mathematical Geology, v. 20, no. 8, p. 907-937.

IHS Energy, Inc, 2004, IHS Energy exploration and production database: Houston, Texas, IHS Energy, Inc; database available from IHS Energy, Inc, P.O. Box 740619, Houston, Texas 77274-0619 U.S.A.

Kaufman, G.M., 1993, Statistical issues in the assessment of undiscovered oil and gas resources: Energy Journal, v. 14, no. 1, p. 183-215.

Klett, T.R., Schmoker, J.W., Charpentier, R.R., Ahlbrandt, T.S., and Ulmishek, G.F., 2000, Glossary, chap. GL *of* U.S. Geological Survey World Energy Assessment Team, U.S. Geological Survey World Petroleum Assessment 2000—Description and Results: U.S. Geological Survey Digital Data Series 60, 8 p.

NRG Associates, 1995, The significant oil and gas pools of Canada data base: Colorado Springs, Colorado, NRG Associates, Inc; database available from NRG Associates, Inc, P.O. Box 1655, Colorado Springs, Colorado 80901 U.S.A.

Root, D.H., and Attanasi, E.D., 1993, Small fields in the National Oil and Gas Assessment: American Association of Petroleum Geologists Bulletin, v. 77, no. 3, p. 485-490.

Rose, P.R., 1994, Chance of success and its use in petroleum exploration, *in* Steinmetz, R., ed., The business of petroleum exploration: American Association of Petroleum Geologists, Treatise of Petroleum Geology, Handbook of Petroleum Geology, v. 2, p. 71-85.

Schmoker, J.W., and Klett, T.R., 2000, U.S. Geological Survey assessment model for undiscovered conventional oil, gas, and NGL resources—The seventh approximation, chap. AM *of* U.S. Geological Survey World Energy Assessment Team, ed., U.S. Geological Survey World Petroleum Assessment 2000—Description and results: U.S. Geological Survey Digital Data Series DDS-60, p. 1-18.

Schuenemeyer, J.H., and Drew, L.J., 1983, A procedure to estimate the parent population of the size of oil and gas fields as revealed by a study of economic truncation: Journal of the International Association for Mathematical Geology, v. 15, no. 1, p. 145-161.

U.S. Geological Survey World Energy Assessment Team, 2000, U.S. Geological Survey World Petroleum Assessment 2000—Description and Results: U.S. Geological Survey Digital Data Series 60, four CD-ROMs.

Appendix 1. 1995 Methodology for Small-Field Extrapolation

In the 1995 U.S. National Oil and Gas Assessment (Gautier and others, 1995), a Pareto distribution was used to describe the size distribution of the set of undiscovered fields larger than 1 million barrels of oil equivalent (MMBOE) in each play. After aggregating to the province level and adding the set of discovered fields, a separate Pareto distribution was used to describe the population of large fields (discovered plus undiscovered) plus small fields. This second Pareto distribution was estimated using binned field sizes (table 1) and the log-geometric distribution (the binned equivalent of the Pareto distribution).

For the Pareto Distribution:

$$f(x) = \frac{\beta L^{\beta}}{x^{(\beta+1)}} = \beta L^{\beta} x^{(-\beta-1)}$$

where:

L = location parameter (minimum field size) and

β= shape factor ($\beta > 0$)

The proportion of the population between size L (the minimum) and size X is:

$$F(X) = \int_{L}^{X} f(x)dx = \left[\beta L^{\beta} \frac{X^{(-\beta-1+1)}}{-\beta-1+1}\right] - \left[\beta L^{\beta} \frac{L^{(-\beta-1+1)}}{-\beta-1+1}\right]$$

$$F(X) = \left[-\frac{L^{\beta}}{X^{\beta}}\right] - \left[-\frac{L^{\beta}}{L^{\beta}}\right]$$

$$F(X) = 1 - \frac{L^{\beta}}{X^{\beta}}$$

Thus, appropriately, F(X) equals 0 where X = L and approaches 1 as X goes to infinity.

The proportion of fields larger than size X would be:

$$1 - F(X) = 1 - \left[1 - \frac{L^{\beta}}{X^{\beta}}\right] = \frac{L^{\beta}}{X^{\beta}}$$

Relation Between Pareto and Log-Geometric Distributions

Distribute a Pareto distribution among bins that have limits of X, PX, P^2X, P^3X, and so on. Traditionally, the USGS uses a P of 2 with a starting value of X = 1 MMBOE, giving bin limits of 1, 2, 4, 8, and so on (table 1).

The proportion of the population between X and PX is:

$$F(X) = \int_{X}^{PX} f(x)dx = \left[-\frac{L^{\beta}}{(PX)^{\beta}}\right] - \left[-\frac{L^{\beta}}{(X)^{\beta}}\right]$$

$$F(X) = \frac{L^{\beta}}{X^{\beta}}\left[1 - \frac{1}{P^{\beta}}\right]$$

Similarly, the proportion of the population between PX and P^2X is:

$$F(X) = \int_{PX}^{P^2X} f(x)dx = \left[-\frac{L^{\beta}}{(P^2X)^{\beta}}\right] - \left[-\frac{L^{\beta}}{(PX)^{\beta}}\right]$$

$$F(X) = \frac{L^{\beta}}{X^{\beta}}\left[\frac{1}{P^{\beta}} - \frac{1}{P^{2\beta}}\right]$$

Also, the proportion of the population between P^2X and P^3X is:

$$F(X) = \int_{P^2X}^{P^3X} f(x)dx = \left[-\frac{L^{\beta}}{(P^3X)^{\beta}}\right] - \left[-\frac{L^{\beta}}{(P^2X)^{\beta}}\right]$$

$$F(X)=\frac{L^{\beta}}{X^{\beta}}\left[\frac{1}{P^{2\beta}}-\frac{1}{P^{3\beta}}\right]$$

For a binned distribution to be log-geometric, the ratios of frequencies (or proportions) between adjacent bins should be constant.

Let r_1 be the ratio between the proportion in bin X-PX to that in bin PX-P^2X:

$$r_1=\frac{\frac{L^{\beta}}{X^{\beta}}\left[1-\frac{1}{P^{\beta}}\right]}{\frac{L^{\beta}}{X^{\beta}}\left[\frac{1}{P^{\beta}}-\frac{1}{P^{2\beta}}\right]}=\frac{1-\frac{1}{P^{\beta}}}{\frac{1}{P^{\beta}}-\frac{1}{P^{2\beta}}}$$

Multiplying numerator and denominator by P^{β}:

$$r_1=\frac{P^{\beta}-1}{1-\frac{1}{P^{\beta}}}$$

Let r_2 be the ratio between the proportion in bin PX-P^2X to that in bin P^2X-P^3X:

$$r_2=\frac{\frac{L^{\beta}}{X^{\beta}}\left[\frac{1}{P^{\beta}}-\frac{1}{P^{2\beta}}\right]}{\frac{L^{\beta}}{X^{\beta}}\left[\frac{1}{P^{2\beta}}-\frac{1}{P^{3\beta}}\right]}=\frac{\frac{1}{P^{\beta}}-\frac{1}{P^{2\beta}}}{\frac{1}{P^{2\beta}}-\frac{1}{P^{3\beta}}}$$

Multiplying numerator and denominator by $P^{2\beta}$:

$$r_2=\frac{P^{\beta}-1}{1-\frac{1}{P^{\beta}}}$$

Because $r_1 = r_2$, the binned distribution is log-geometric.

Upon examining these r values for actual data sets in the Permian Basin of the United States, Drew and others (1982) and Drew (1990) concluded that field-size distributions were Pareto. They noted a pattern such as that shown in figure 12, drawn from data for the Supra-Domanik Carbonates/Clastics Assessment Unit in the analog database. For field-size classes larger than about 200 MMBOE, the numbers of fields in each class are small and the ratios of adjacent classes are unstable. For classes smaller than 200 MMBOE, however, the ratio becomes more stable and approaches a constant of approximately 1.4. This implies a log-geometric distribution for the binned data and thus a Pareto distribution for the field sizes themselves.

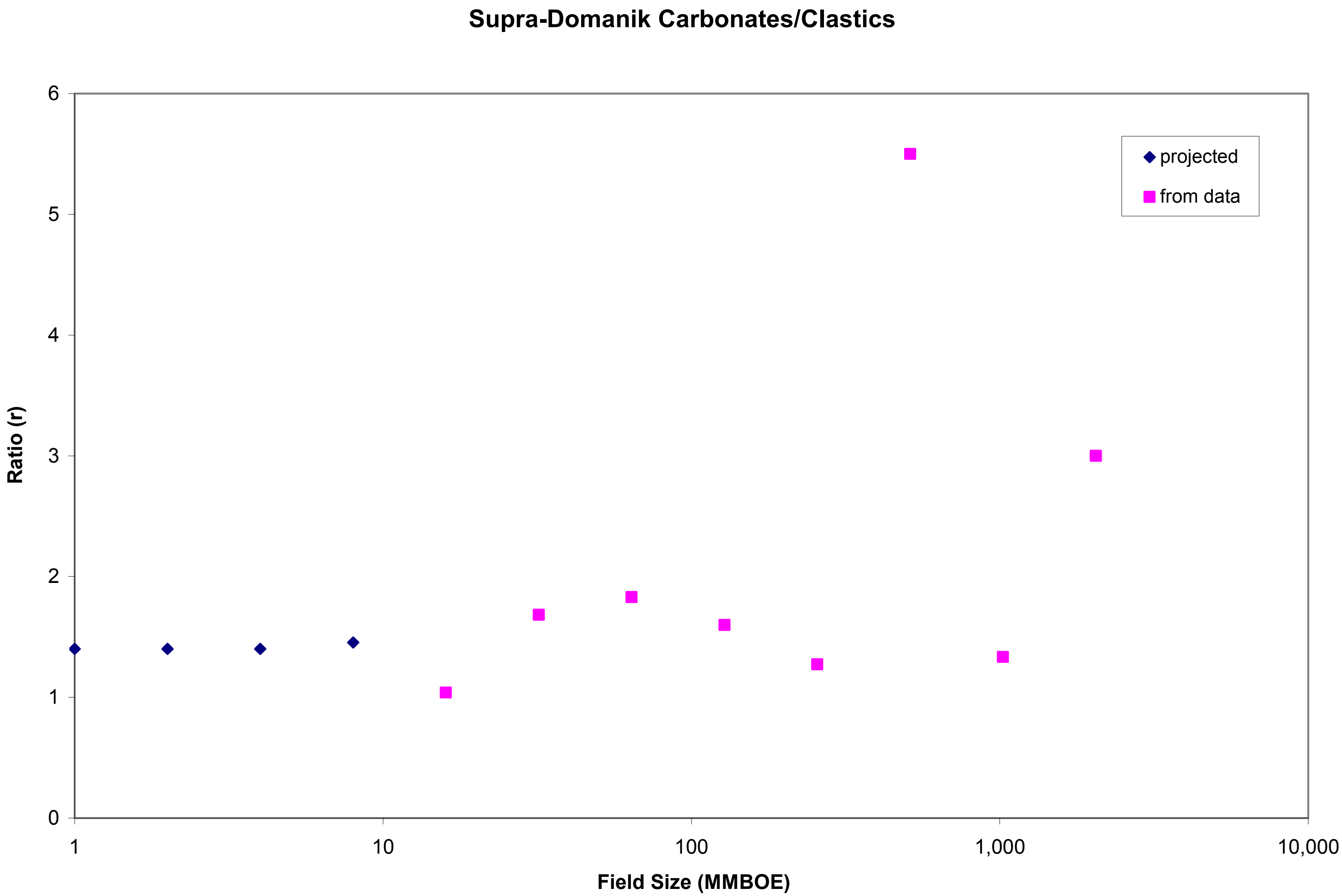

Figure 12. Ratio (r) between the numbers of discovered plus undiscovered fields in adjacent size classes for the Supra-Domanik Carbonates/Clastics Assessment Unit (10150101) of the Volga-Ural Basin, Russia. MMBOE, million barrels of oil equivalent.

Closer examination of these patterns using the analog database shows that the pattern is not always so clear (fig. 13). In some cases there is little convergence and in others the r's tend to converge on a value greater than 2. This is problematic because an r value greater than 2 implies that the resource approaches infinity as smaller and smaller size bins are considered. At an r value of 2, each bin has the same volume of resource.

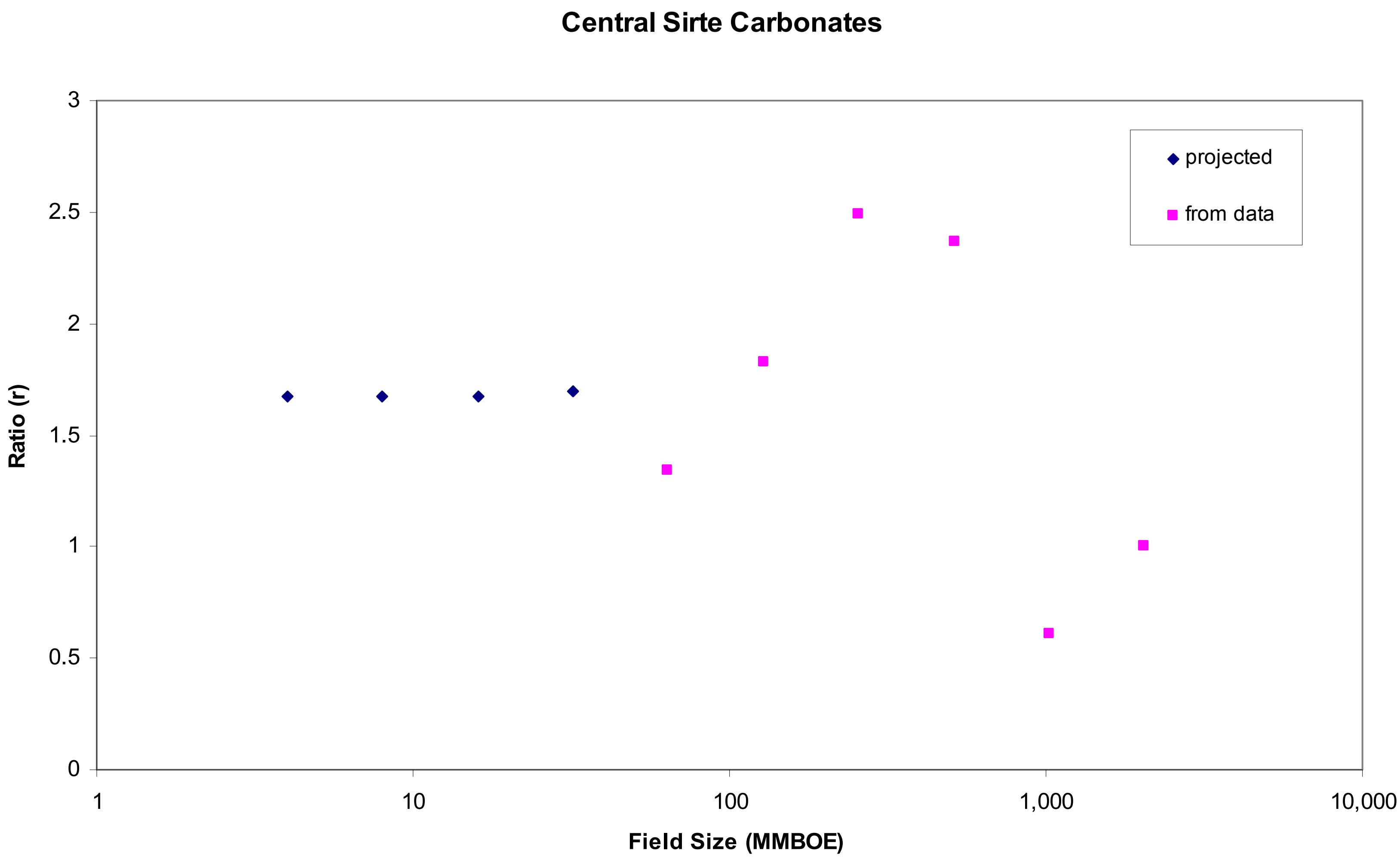

Figure 13. Ratio (r) between the numbers of discovered plus undiscovered fields in adjacent size classes for the Central Sirte Carbonates Assessment Unit (20430102) of the Sirte Basin, Libya. MMBOE, million barrels of oil equivalent.

In the 1995 U.S. National Oil and Gas Assessment, r was calculated for each of the 58 U.S. provinces with resource potential. Data on discovered oil and gas fields were binned by barrels of oil equivalent (BOE) (table 1). The Pareto distributions for undiscovered fields were divided into bins. This was the raw data set used for the calculation. The five size classes from class 7 (2 to 4 MMBOE) to class 11 (32 to 64 MMBOE) were used in the calculation. Class 6 (1 to 2 MMBOE) was not used because of partial economic truncation at that size range.

A least squares procedure was used to calculate a best-fit value for r. The number of fields in size classes 7 through 11 is:

$$N = \sum_{c=7}^{c=11} n_c$$

where n_c is the number of discovered fields plus the estimated mean value of number of undiscovered fields in class c.

If the distribution were log-geometric with ratio r:

Class 11 would have some number n_{11} of fields,
Class 10 would have r n_{11} of fields,
Class 9 would have r^2 n_{11} of fields,
Class 8 would have r^3 n_{11} of fields, and
Class 7 would have r^4 n_{11} of fields.

Thus, the total number of fields in classes 7 to 11 would be:

$$N = \left(r^4 n\right) + \left(r^3 n\right) + \left(r^2 n\right) + \left(rn\right) + n$$

Given a trial value for r = r_{trial}, one can calculate expected values of numbers in each size class by:

$$n_{11\exp} = \frac{N}{\left(r_{trial}^4 + r_{trial}^3 + r_{trial}^2 + r_{trial} + 1\right)}$$

$$n_{10\exp} = \frac{r_{trial} N}{\left(r_{trial}^4 + r_{trial}^3 + r_{trial}^2 + r_{trial} + 1\right)}$$

$$n_{9\exp} = \frac{r_{trial}^2 N}{\left(r_{trial}^4 + r_{trial}^3 + r_{trial}^2 + r_{trial} + 1\right)}$$

$$n_{8\exp} = \frac{r_{trial}^{3} N}{\left(r_{trial}^{4} + r_{trial}^{3} + r_{trial}^{2} + r_{trial} + 1\right)}$$

$$n_{7\exp} = \frac{r_{trial}^{4} N}{\left(r_{trial}^{4} + r_{trial}^{3} + r_{trial}^{2} + r_{trial} + 1\right)}$$

The expected value (given some trial r) of number of fields in each of the five size classes was compared to the number of discovered plus estimated undiscovered from the assessment and the sum of squared deviations calculated:

$$SSQ = \left(n_{11} - n_{11\exp}\right)^2 + \left(n_{10} - n_{10\exp}\right)^2 + \left(n_{9} - n_{9\exp}\right)^2 + \left(n_{8} - n_{8\exp}\right)^2 + \left(n_{7} - n_{7\exp}\right)^2$$

A sum of squared deviations from the expected was calculated for each value from 1.4 to 1.7 at 0.01 intervals (1.40, 1.41, 1.42 … 1.69, 1.70). The trial value of r with the least sum of square deviations was chosen as r for the province.

Checks were conducted for data that did not fit the extrapolated field size. In some cases the number of discovered fields in one of classes 1 to 6 was larger than that estimated from the small-field extrapolation procedure. In this case, the extrapolated number of fields was used rather than the actual number of discovered fields.

Changes to Small-Field Extrapolation Methodology

For the 1995 U.S. National Oil and Gas Assessment, small oil and gas fields were estimated at the **province** level only. For the analog database a similar methodology was applied at the **assessment unit** level. Because of the smaller numbers of fields for each assessment unit there was an increased chance of poor distribution fit. Checks were thus made for cases where there were more discovered fields than the extrapolation estimated in the small field-size classes. In such cases, the additional discovered fields were eliminated and the database cell was colored gold in the Oil Bins, Gas Bins, and BOE Bins worksheets.

The WPA 2000 used a shifted truncated lognormal distribution rather than a Pareto distribution for assessing large undiscovered fields. It also used a variable minimum assessed field size. Both of these changes made modification to the procedure necessary. In the 1995 methodology, the minimum size assessed was 1 MMBOE. Because a Pareto distribution was used, the modal bin was always the 1 to 2 MMBOE class (class 6). The five bins used for calculation of r were the five starting with the 2 to 4 MMBOE class (class 7) up through the 32 to 64 MMBOE class (class 11). In the present procedure, the modal bin may be somewhere above the minimum assessed size (fig. 14). The five bins above the modal bin are used for the calculation of r (fig. 15).

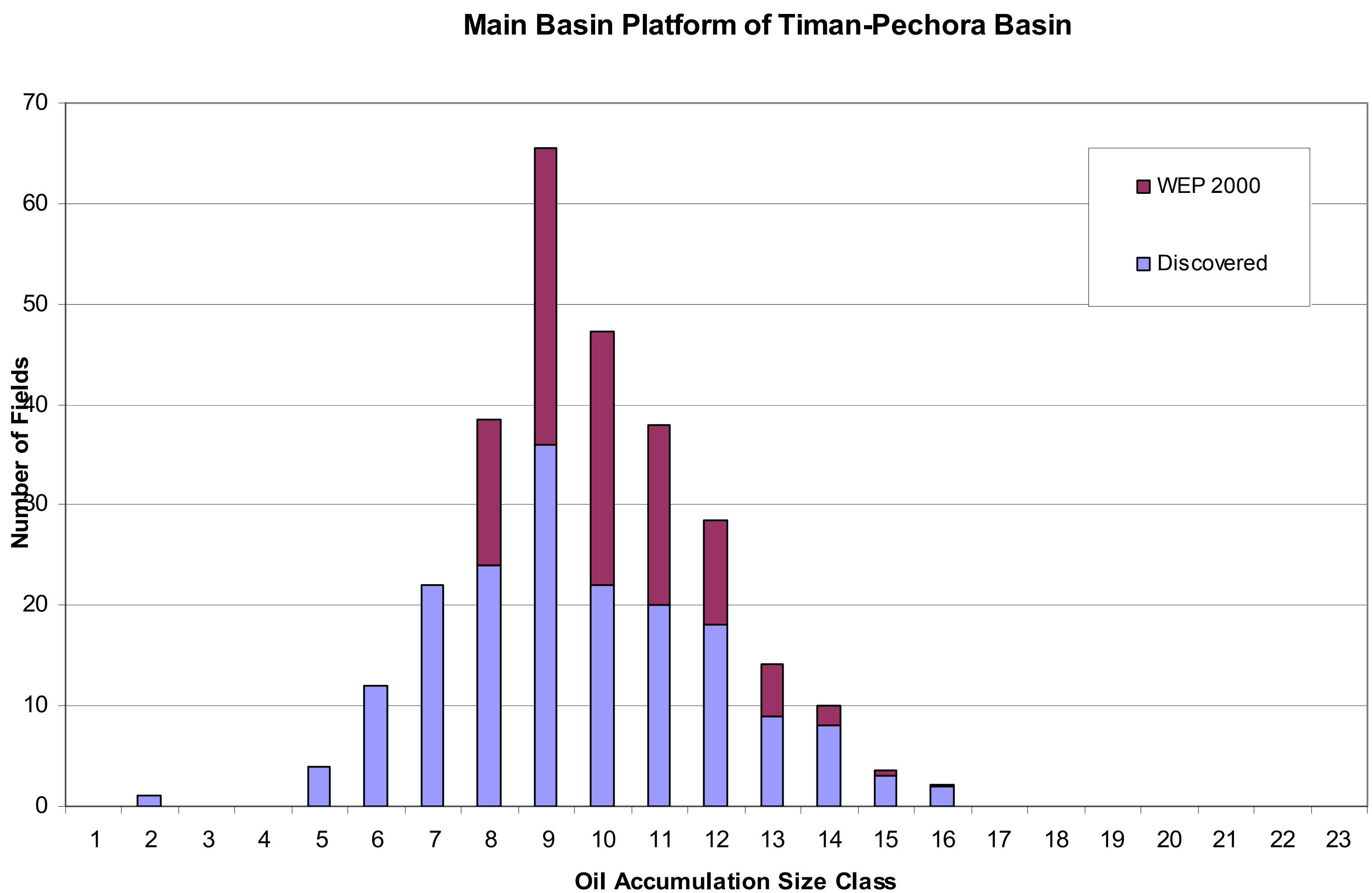

Figure 14. Numbers of discovered and undiscovered (from the 2000 USGS assessment) oil fields by size class in the Main Basin Platform Assessment Unit (10080102) of the Timan-Pechora Basin, Russia. Note the linear vertical scale.

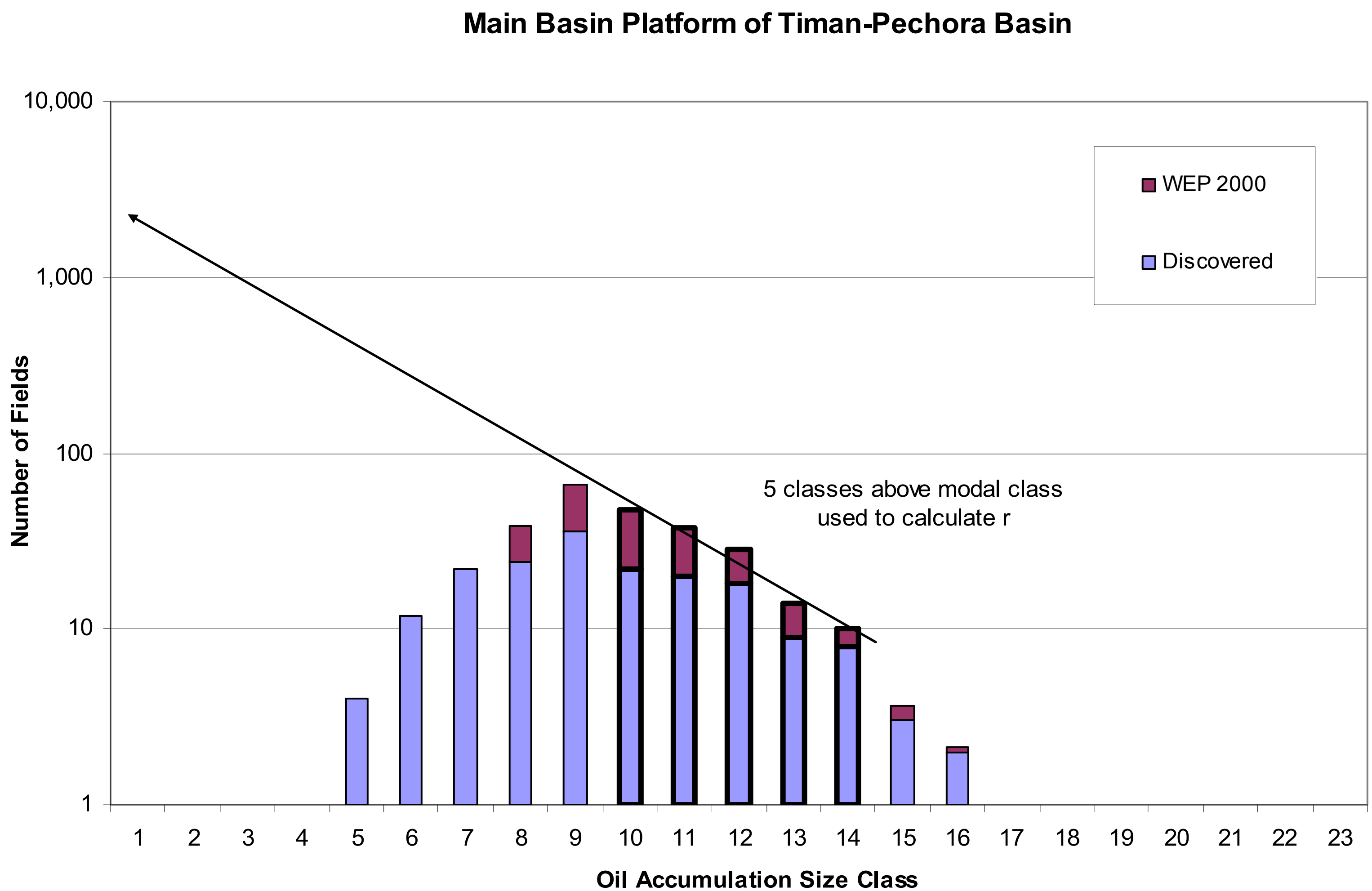

Figure 15. Numbers of discovered and undiscovered (from the 2000 USGS assessment) oil fields by size class in the Main Basin Platform Assessment Unit (10080102) of the Timan-Pechora Basin, Russia, showing an extrapolation of small-field sizes using the data from classes 10 to 14. Note the logarithmic vertical scale.

The mean size of fields within a size class can be calculated for a lognormal or a Pareto distribution. The size distributions in the analog database, however, are combinations of a lognormal distribution and additional discrete (discovered) fields. As a first approximation, the 1995 methodology used a mean size of 1.5 times the lower size-class boundary (half-way between the two bounds of the size class).

In experiments with lognormal distributions, it can be shown that the mean size in each field size class varies in location from class to class (table 5). Below the mode of the field-size distribution, the mean size of a class is closer to the upper boundary of the size class. At the mode, the mean size of a class is approximately at the center of the size class. Above the mode, the mean size of a class is closer to the lower boundary and approaches the lower boundary even more closely as one goes to larger size classes. One can specify the mean size of a class by using a multiplier between 1 and 2 times the lower size-class boundary. Use of different values for that multiplier per size class would have greatly complicated the calculations for the database because the multiplier values depend not only on the size class but also on the distribution itself. Experiments showed that a multiplier value of 1.38 times the lower size class boundary gave the best approximation if only one value was to be used. The new methodology thus uses the value of 1.38, rather than the 1.5 that was used in 1995.

Table 5. Simulation of a lognormal distribution with mean of 50 million barrels of oil equivalent (MMBOE) and standard deviation of 100 MMBOE divided into USGS field-size classes. Right column shows relation of the mean in each class relative to the class boundaries.

Class	Minimum size	Number	Volume	Mean size	Mean size/Minimum
1	0.03125	1	0.06	0.06	1.88
2	0.0625	27	2.93	0.11	1.74
3	0.125	151	31.04	0.21	1.64
4	0.25	1,173	466.55	0.40	1.59
5	0.5	5,665	4,422.19	0.78	1.56
6	1	21,401	32,905.38	1.54	1.54
7	2	59,212	179,567.86	3.03	1.52
8	4	121,337	724,043.50	5.97	1.49
9	8	186,682	2,192,422.75	11.74	1.47
10	16	215,607	4,976,658.55	23.08	1.44
11	32	185,660	8,425,796.98	45.38	1.42
12	64	118,833	10,603,253.55	89.23	1.39
13	128	56,751	9,961,970.76	175.54	1.37
14	256	20,621	7,111,576.20	344.87	1.35
15	512	5,545	3,768,982.60	679.71	1.33
16	1024	1,148	1,535,004.65	1,337.11	1.31
17	2048	166	443,442.94	2,671.34	1.30
18	4096	19	106,505.41	5,605.55	1.37
19	8192	0	0.00	0.00	0.00
20	16384	0	0.00	0.00	0.00
21	32768	0	0.00	0.00	0.00
Totals:		999,999	50,067,054	50.07	

Sizes and volumes in millions of barrels equivalent

Sensitivity Tests

A series of sensitivity tests was conducted to determine the effect of different assumptions on the calculation of r. The main assumptions are listed in table 6 and the results of the sensitivity analysis on several of these assumptions are given in table 7. Given the basic framework of accepting assumption 1, the tests determined that most of the assumptions were not particularly sensitive. The main exception was assumption 2, the constraint that r is constrained to lie between 1.4 and 1.7.

Table 6. Assumptions for the 1995 small-field assessment methodology.

1. The small fields have a log-geometric distribution.
2. r can vary from 1.4 to 1.7.
3. A least-squares procedure should be used to calculate r.
4. Mean size in a bin is 1.5 times the lower bin limit.
5. The least-squares procedure should not use bins with value less than 1.0.
6. r should be fit to BOE totals, not separately for oil and gas fields.
7. The least-squares procedure uses the five bins starting with the one larger than the modal bin.
8. Oil and gas should be proportioned similar to the modal bin plus the next two larger bins.
9. The least-squares fit should not use discovered fields of a type (oil/gas) not assessed.
10. The oil or gas proportion should not use discovered fields of a type (oil/gas) not assessed.
11. "Excess" discovered fields in a bin should be eliminated.
12. Least squares should be based on numbers of accumulations rather than volumes of resources.
13. r should be calculated at the province scale.

Table 7. Sensitivity studies of the calculation of r given different assumptions.

assessment unit number and name		base case r	extended base case	multi-nomial r	ratio 1/2	ratio 2/3	ratio 3/4	ratio 4/5	mean ratio	1.4 mean	1.38 mean	no <1 constraint
10080101	Northwest Izhma-Pechora Depression	1.70	3.00	3.00	2.56	3.54	4.86	16.83	6.95	3.00	3.00	3.00
10080102	Main Basin Platform	1.45	1.45	1.44	1.20	1.33	1.98	1.43	1.49	1.45	1.45	1.45
10080103	Foredeep Basins	1.70	2.22	2.20	2.17	1.99	2.44	3.03	2.41	2.22	2.22	2.22
10090101	Carboniferous-Lower Permian Clastics	1.60	1.60	1.60	1.46	1.64	1.65	1.77	1.63	1.60	1.60	1.60
10090102	Devonian Synrift	1.70	2.46	2.43	1.93	2.55	3.39	5.53	3.35	2.46	2.46	2.46
10150101	Supra-Domanik Carbonates/Clastics	1.40	1.40	1.39	1.45	1.04	1.68	1.83	1.50	1.40	1.40	1.40
10150102	Lower Volga	1.40	1.35	1.35	1.00	1.48	1.51	1.50	1.37	1.35	1.35	1.35
10150103	Sub-Domanik Devonian Clastics	1.45	1.45	1.43	1.24	1.64	0.98	3.31	1.80	1.45	1.45	1.45
10150201	Permian Reefs/Thrust Folds	1.70	1.75	1.71	1.43	1.44	1.56	134.98	34.85	1.75	1.75	1.76
10160101	North and West Margins Subsalt Pinnacle Reefs	1.70	2.36	2.33	1.93	2.33	3.13	4.69	3.02	2.36	2.36	2.36
10160102	North and West Margins Subsalt Barrier Reefs	1.70	3.00	3.00	3.26	32.61	0.00	0.00	17.93	3.00	3.00	3.00
10160103	East and Southeast Margins Subsalt	1.70	1.79	1.76	1.80	1.59	1.43	4.67	2.37	1.79	1.79	1.79
10160104	South Margin Subsalt	1.40	1.39	1.38	1.24	1.16	1.75	1.72	1.47	1.39	1.39	1.39
10160106	Suprasalt	1.70	1.82	1.80	1.46	1.68	2.50	2.44	2.02	1.82	1.82	1.82
10500101	Kolguyev Terrace	1.70	2.21	2.20	1.77	2.48	2.53	3.10	2.47	2.21	2.21	2.21
10500102	South Barents and Ludlov Saddle	1.54	1.54	1.53	1.42	1.31	1.75	2.22	1.67	1.54	1.54	1.54
10500103	North Barents	1.70	1.79	1.78	1.57	1.74	1.99	2.30	1.90	1.79	1.79	1.79
11080101	Tertiary Foredeep	1.68	1.68	1.71	1.45	2.50	1.67	1.01	1.66	1.68	1.68	1.68
11080102	Foreland Slope	1.50	1.50	1.53	1.38	2.03	1.75	0.78	1.49	1.50	1.50	1.50
11080103	Subsalt Jurassic	1.70	2.44	2.37	2.70	1.20	7.56	4.49	3.99	2.44	2.44	2.45
11090101	Foldbelt-Foothills	1.70	1.94	1.88	1.64	1.65	1.69	420.72	106.43	1.94	1.94	1.94
11090102	Terek-Sunzha Subsalt Jurassic	1.70	3.00	3.00	2.45	3.76	5.80	0.00	4.00	3.00	3.00	3.00
11090103	Foreland Slope and Foredeep	1.70	3.00	3.00	6.42	11.54	0.00	0.00	8.98	3.00	3.00	3.00

11090201	South Mangyshlak (Entire)	1.70	1.74	1.74	1.25	3.82	0.74	3.21	2.26	1.74	1.74	1.74
11090301	Offshore Prikumsk Zone	1.70	3.00	3.00	2.81	3.27	4.04	6.91	4.26	3.00	3.00	3.00
11090302	Onshore Stavropol-Prikumsk	1.70	1.91	1.90	1.70	1.40	5.68	1.02	2.45	1.91	1.91	1.91
11090303	Central Caspian Offshore	1.70	3.00	3.00	6.07	22.60	0.00	0.00	14.34	3.00	3.00	3.00
11120101	Apsheron-Pribalkhan Zone	1.45	1.45	1.46	1.14	1.88	1.85	0.89	1.44	1.45	1.45	1.45
11120102	Lower Kura Depression and Adjacent Shelf	1.46	1.46	1.45	1.38	1.33	1.31	2.62	1.66	1.46	1.46	1.46
11120103	Gograndag-Okarem Zone and Adjacent Shelf	1.64	1.64	1.65	1.34	1.91	1.84	1.49	1.65	1.64	1.64	1.64
11120104	Central Offshore	1.70	1.75	1.74	1.41	1.71	2.09	2.56	1.94	1.75	1.75	1.75
11120105	Iran Onshore/Nearshore	1.70	1.99	1.97	1.52	1.99	2.60	3.44	2.39	1.99	1.99	1.99
11500101	Mesozoic Sandstone Reservoirs	1.70	1.87	1.87	1.51	2.20	1.89	2.28	1.97	1.87	1.87	1.87
11500201	Jurassic-Tertiary Reservoirs	1.69	1.69	1.70	1.55	1.68	3.22	0.74	1.80	1.69	1.69	1.69
11500301	Upper Paleozoic Carbonates	1.70	2.08	2.06	1.79	2.05	2.43	2.92	2.30	2.08	2.08	2.08
11540101	Northern and Western Areas	1.70	1.88	1.87	1.60	1.78	2.75	1.75	1.97	1.88	1.88	1.88
11540102	Karabil-Badkhyz (Southern Area)	1.70	1.92	1.90	1.80	1.84	1.76	3.38	2.20	1.92	1.92	1.92
11540103	Murgab Depression Suprasalt	1.55	1.55	1.54	0.94	1.33	3.81	1.42	1.87	1.55	1.55	1.55
11540104	Murgab Depression Subsalt	1.40	1.24	1.25	1.08	2.02	0.82	1.24	1.29	1.24	1.24	1.24
11740101	Upper Jurassic-Cretaceous Sandstones	1.46	1.46	1.47	1.51	1.39	1.59	1.33	1.46	1.46	1.46	1.46
11740201	Pre-Upper Jurassic	1.70	1.74	1.74	1.64	1.78	1.85	1.71	1.74	1.74	1.74	1.74
11740301	Northern West Siberian Onshore Gas	1.40	1.37	1.37	1.13	1.32	1.82	1.26	1.38	1.37	1.37	1.37
11740302	South Kara Sea Offshore	1.40	1.40	1.39	1.15	1.47	1.32	1.99	1.48	1.40	1.40	1.40
12070101	Yenisey Foldbelt Riphean-Craton Margin Riphean	1.61	1.61	1.62	1.31	1.94	1.87	1.28	1.60	1.61	1.61	1.61
12100101	Baikal-Patom Foldbelt Riphean-Craton Margin Vendian	1.55	1.55	1.55	1.29	1.43	2.48	1.22	1.61	1.55	1.55	1.55
13220101	Onshore and Offshore Northeastern Shelf	1.64	1.64	1.65	1.16	2.30	1.79	1.40	1.66	1.64	1.64	1.64
20040101	Ma'Rib-Al Jawf/Shabwah/Masila	1.70	1.80	1.78	1.47	1.46	3.38	1.88	2.05	1.80	1.80	1.80
20140101	Ghaba-Makarem Combined Structural	1.54	1.54	1.52	1.29	1.10	2.45	2.50	1.83	1.54	1.54	1.54
20160101	Fahud-Huqf Combined Structural	1.70	2.15	2.19	2.50	2.95	1.31	1.47	2.06	2.15	2.15	2.15
20160201	Natih-Fiqa Structural/Stratigraphic	1.40	1.28	1.31	1.02	2.24	2.13	0.37	1.44	1.28	1.28	1.28
20190101	Cretaceous Reservoirs in Northwest Desert Anticlines	1.40	1.28	1.28	1.33	1.08	1.70	1.07	1.30	1.28	1.28	1.28

20190102	Cretaceous Reservoirs in South Gulf Suprasalt Structural	1.50	1.50	1.47	1.26	1.00	2.19	3.31	1.94	1.50	1.50	1.50
20190103	Mesozoic/Tertiary Foredeep Fold and Thrust	1.69	1.69	1.72	1.33	2.05	4.44	0.46	2.07	1.69	1.69	1.69
20190201	Jurassic Reservoirs in Northwest Desert Anticlines	1.52	1.52	1.51	0.97	1.70	1.65	2.79	1.78	1.52	1.52	1.52
20190202	Jurassic Reservoirs in South Gulf Suprasalt/Qatar Arch Structural	1.60	1.60	1.61	1.54	1.78	1.69	1.28	1.57	1.60	1.60	1.60
20190301	Khuff Carbonates in Salt Structures	1.70	1.81	1.79	1.59	1.45	2.70	2.56	2.07	1.81	1.81	1.81
20190302	Paleozoic Reservoirs	1.40	1.36	1.36	1.27	1.30	1.42	1.57	1.39	1.36	1.36	1.36
20210101	Central Arch Horst-Block Anticlinal Oil and Gas	1.70	1.93	1.92	1.45	2.11	2.18	3.42	2.29	1.93	1.93	1.93
20210102	North Gulf Salt Basin Structural Gas	1.51	1.51	1.52	1.37	1.69	1.61	1.32	1.50	1.51	1.51	1.51
20210201	Horst-Block Anticlinal Oil	1.66	1.66	1.63	1.31	1.08	2.92	6.29	2.90	1.66	1.66	1.67
20210202	Salt-Involved Structural Oil	1.56	1.56	1.61	1.63	1.74	7.08	0.18	2.66	1.56	1.56	1.57
20230101	Horst/Graben-Related Oil and Gas	1.70	1.90	1.87	1.62	1.81	1.90	4.43	2.44	1.90	1.90	1.90
20230201	Platform Horst/Graben-Related Oil	1.70	1.89	1.86	1.42	2.07	1.85	4.63	2.49	1.89	1.89	1.89
20230202	Basinal Oil and Gas	1.70	2.86	2.79	2.27	2.43	8.40	0.00	4.37	2.86	2.86	2.86
20300101	Cretaceous Reservoirs	1.47	1.47	1.46	1.22	1.55	1.38	2.18	1.58	1.47	1.47	1.47
20300102	Tertiary Reservoirs	1.66	1.66	1.66	1.89	2.59	0.71	2.14	1.83	1.66	1.66	1.66
20300201	Northern Qatar Arch Extension	1.46	1.46	1.44	1.16	1.56	1.20	3.01	1.74	1.46	1.46	1.46
20430101	Southeast Sirte Clastics	1.55	1.55	1.57	1.63	3.16	0.63	1.60	1.75	1.55	1.55	1.55
20430102	Central Sirte Carbonates	1.67	1.67	1.66	1.70	1.35	1.83	2.49	1.84	1.67	1.67	1.67
20430103	Offshore Sirte Hypothetical	1.70	2.34	2.32	1.87	2.41	3.12	4.21	2.90	2.34	2.34	2.34
20430104	Southeast Sirte Hypothetical	1.70	2.22	2.19	1.62	2.31	3.25	4.65	2.96	2.22	2.22	2.22
20480101	Bou Dabbous-Tertiary Structural/Stratigraphic	1.64	1.64	1.67	1.51	2.56	1.32	1.10	1.62	1.64	1.64	1.64
20480201	Jurassic-Cretaceous Structural/Stratigraphic	1.70	1.77	1.74	1.18	2.52	1.13	8.94	3.44	1.77	1.77	1.77
20540101	Tanezzuft-Oued Mya Structural/Stratigraphic	1.70	1.75	1.77	3.43	1.74	0.83	1.45	1.86	1.75	1.75	1.75
20540201	Tanezzuft-Melrhir Structural/Stratigraphic	1.70	1.70	1.69	1.30	1.86	1.89	2.19	1.81	1.70	1.70	1.70
20540301	Tanezzuft-Ghadames Structural/Stratigraphic	1.40	1.26	1.26	1.08	1.23	1.68	1.05	1.26	1.26	1.26	1.26

20560101	Tanezzuft-Illizi Structural/Stratigraphic	1.70	2.00	2.02	1.56	2.79	2.34	1.33	2.00	2.00	2.00	2.00
20580101	Tanezzuft-Timimoun Structural/Stratigraphic	1.40	1.21	1.18	1.21	7.17	0.13	2.06	2.64	1.21	1.21	1.17
20580201	Tanezzuft-Ahnet Structural/Stratigraphic	1.70	1.97	1.95	1.01	3.48	1.50	43.51	12.37	1.97	1.97	1.97
20580301	Tanezzuft-Sbaa Structural/Stratigraphic	1.70	1.84	1.80	1.80	5.32	0.25	168.11	43.87	1.84	1.84	1.84
20580401	Tanezzuft-Mouydir Structural/Stratigraphic	1.40	1.00	3.00	2.75	3.45	4.59	0.00	3.60	1.00	1.00	3.00
20580501	Tanezzuft-Benoud Structural/Stratigraphic	1.70	1.95	1.94	1.90	1.74	2.13	2.61	2.10	1.95	1.95	1.96
20580601	Tanezzuft-Bechar/Abadla Structural/Stratigraphic	1.70	1.93	1.91	1.46	1.94	2.50	3.17	2.27	1.93	1.93	1.93
20710101	Gulf of Suez Block-Fault Fairway	1.40	1.36	1.35	1.01	1.26	2.09	1.22	1.40	1.36	1.36	1.36
20710102	Gulf of Suez Qaa Plain	1.40	1.00	2.17	1.71	2.24	2.93	3.95	2.71	1.00	1.00	2.19
20710103	Southern Gulf of Suez	1.65	1.65	1.63	1.32	1.32	2.48	2.72	1.96	1.65	1.65	1.65
20710201	Red Sea Coastal Block Faults	1.57	1.57	1.57	1.29	1.55	1.98	1.68	1.62	1.57	1.57	1.57
20710202	Red Sea Salt Basin	1.70	3.00	3.00	2.92	2.93	3.33	4.42	3.40	3.00	3.00	3.00
31150101	Upper Paleozoic/Lower Mesozoic Nonmarine Coarse Clastic	1.40	1.39	1.39	1.27	1.53	1.26	1.60	1.41	1.39	1.39	1.39
31150201	Jurassic/Tertiary Fluvial and Lacustrine Sandstone	1.70	1.98	2.08	1.33	6.04	3.87	0.24	2.87	1.98	1.98	1.99
31270101	Tertiary Lacustrine	1.70	1.78	1.78	1.67	1.47	3.41	1.12	1.92	1.78	1.78	1.78
31270102	Pre-Tertiary Buried Hills	1.70	1.79	1.81	1.56	2.61	1.65	1.19	1.75	1.79	1.79	1.79
31280101	Jurassic/Triassic Fluvial and Lacustrine Sandstone	1.70	1.82	1.80	0.72	3.50	1.54	0.00	1.92	1.82	1.82	1.82
31420101	Southeastern Fold Belt	1.48	1.48	1.46	0.86	1.60	1.74	2.98	1.79	1.48	1.48	1.48
31420102	Northwestern Depression/Foldbelt	1.70	1.72	1.69	1.62	2.05	0.83	10.37	3.72	1.72	1.72	1.73
31420201	Jurassic Lacustrine	1.62	1.62	1.62	1.15	2.79	1.09	2.01	1.76	1.62	1.62	1.62
31420401	Leshan-Longnusi Paleohigh	1.70	1.76	1.73	1.78	1.81	1.06	5.65	2.58	1.76	1.76	1.76
31420402	Lower Paleozoic of Southeastern Fold Belt	1.70	2.37	2.35	1.89	2.44	3.19	4.29	2.95	2.37	2.37	2.37
31440101	Subtle Traps	1.70	1.79	1.76	1.33	2.52	0.84	0.00	1.56	1.79	1.79	1.80
31440102	Anticlinal	1.40	1.16	1.17	4.01	0.49	1.09	1.00	1.65	1.16	1.16	1.16
31440201	Structural Traps	1.70	3.00	3.00	3.55	4.84	11.79	0.00	6.72	3.00	3.00	3.00
31540101	Tarim Basin Excluding Marginal Foldbelts	1.70	2.21	2.18	1.53	2.40	3.61	4.05	2.90	2.21	2.21	2.21
31540102	Kuche (Northern) Foldbelt	1.70	3.00	3.00	2.89	3.03	4.41	17.20	6.88	3.00	3.00	3.00
31540103	Southwest Foldbelt	1.70	2.48	2.45	2.01	2.48	3.41	5.22	3.28	2.48	2.48	2.48

37010101	Brunei-Sabah Deltaics	1.40	1.38	1.37	1.13	1.70	1.12	1.78	1.43	1.38	1.38	1.38
37010102	Brunei-Sabah Turbidites	1.66	1.66	1.65	1.14	1.89	2.20	1.76	1.75	1.66	1.66	1.66
37020101	Central Luconia	1.52	1.52	1.52	1.46	1.28	1.99	1.51	1.56	1.52	1.52	1.52
37020102	Balingian	1.70	2.04	2.04	1.97	3.66	0.74	4.19	2.64	2.04	2.04	2.04
37020201	East Natuna	1.70	2.60	2.57	2.33	2.45	3.24	4.57	3.15	2.60	2.60	2.60
37030101	South Malay Lacustrine	1.62	1.62	1.61	1.09	1.68	2.38	1.91	1.76	1.62	1.62	1.62
37030102	North Malay Lacustrine	1.70	3.00	3.00	2.21	3.69	6.27	29.56	10.43	3.00	3.00	3.00
37030201	South Malay Coaly	1.40	1.38	1.39	1.29	4.89	0.34	1.56	2.02	1.38	1.38	1.38
38080101	Pematang/Sihapas Siliciclastics	1.52	1.52	1.51	1.51	1.03	2.23	2.05	1.70	1.52	1.52	1.52
38170101	Kutei Basin Deltaics	1.44	1.44	1.43	1.33	1.16	1.83	1.75	1.52	1.44	1.44	1.44
38170102	Kutei Basin Turbidites	1.64	1.64	1.62	1.28	1.58	1.96	2.44	1.81	1.64	1.64	1.64
38170103	Kutei Basin Fold and Thrust Belt	1.70	2.31	2.30	1.53	3.12	3.05	3.96	2.92	2.31	2.31	2.31
38220101	North Sumatra	1.70	1.93	1.92	1.82	1.76	2.23	2.36	2.04	1.93	1.93	1.93
38220102	Mergui	1.70	1.89	1.88	1.63	1.86	2.16	2.60	2.06	1.89	1.89	1.89
38240101	Sunda/Asri	1.70	1.74	1.72	1.42	1.66	1.91	3.12	2.03	1.74	1.74	1.74
38240201	Ardjuna	1.62	1.62	1.63	1.11	2.18	2.19	1.12	1.65	1.62	1.62	1.62
38280101	South Sumatra	1.40	1.33	1.32	1.11	1.09	1.76	1.64	1.40	1.33	1.33	1.33
39100101	Barnett	1.70	3.00	3.00	2.26	4.43	9.33	0.00	5.34	3.00	3.00	3.00
39100201	Petrel	1.70	2.27	2.25	1.33	3.08	3.71	6.35	3.62	2.27	2.27	2.27
39100202	Vulcan Graben	1.70	3.00	3.00	2.50	3.71	6.35	0.00	4.19	3.00	3.00	3.00
39100301	Malita	1.70	1.79	1.81	1.17	2.38	3.66	0.81	2.01	1.79	1.79	1.79
39130101	Late Jurassic/Early Cretaceous-Mesozoic	1.51	1.51	1.54	1.71	1.78	1.97	0.62	1.52	1.51	1.51	1.51
39300101	Gippsland	1.70	1.85	1.87	1.51	2.90	1.50	1.51	1.85	1.85	1.85	1.85
39480101	Dingo-Mungaroo/Barrow	1.70	1.84	1.83	1.57	1.95	1.94	2.30	1.94	1.84	1.84	1.84
39480201	Locker-Mungaroo/Barrow	1.67	1.67	1.73	2.25	2.12	2.20	0.44	1.75	1.67	1.67	1.68
40170101	Halten Terrace-Trondelag Platform	1.70	2.31	2.24	2.11	1.61	3.50	28.75	8.99	2.31	2.31	2.31
40170102	Mid-Norway Continental Margin	1.50	1.50	1.49	1.28	1.44	1.66	1.94	1.58	1.50	1.50	1.50
40250101	Viking Graben	1.43	1.43	1.41	1.59	1.07	1.46	2.13	1.56	1.43	1.43	1.43
40250102	Moray Firth	1.62	1.62	1.60	1.33	1.82	1.22	3.86	2.06	1.62	1.62	1.62
40250103	Central Graben	1.67	1.67	1.65	1.26	1.76	1.79	2.70	1.88	1.67	1.67	1.67
40360101	Southern Permian Basin-U.K. Onshore	1.70	2.60	2.58	7.91	0.00	0.00	0.00	3.95	2.60	2.60	2.62

40360102	Southern Permian Basin-Europe Onshore	1.40	1.23	1.21	1.06	0.98	1.28	2.24	1.39	1.23	1.23	1.23
40360103	Southern Permian Basin-Offshore	1.70	1.73	1.71	1.45	1.58	2.17	2.39	1.90	1.73	1.73	1.73
40470101	Foreland Basin	1.70	1.73	1.74	2.07	1.56	1.62	1.63	1.72	1.73	1.73	1.73
40470201	Deformed Belt	1.70	1.85	1.79	2.16	1.00	2.34	7.06	3.14	1.85	1.85	1.85
40470301	Paleozoic Reservoirs	1.70	3.00	3.00	2.69	4.01	7.03	0.00	4.58	3.00	3.00	3.00
40480101	Greater Hungarian Plain Basins	1.70	1.92	1.91	1.83	1.54	3.17	1.68	2.05	1.92	1.92	1.92
40480201	Zala-Drava-Sava Basins	1.70	1.91	1.93	1.65	3.35	1.07	2.11	2.04	1.91	1.91	1.91
40480301	Danube Basin	1.55	1.55	1.62	2.25	2.59	4.09	0.13	2.27	1.55	1.55	1.55
40480401	Transcarpathian Basin	1.70	3.00	3.00	1228.79	0.00	0.00	0.00	1228.79	3.00	3.00	3.00
40480601	Hungarian Paleogene Basin	1.70	2.77	2.76	1.60	5.10	4.94	13.44	6.27	2.77	2.77	2.77
40570101	Transylvanian Neogene Suprasalt Gas	1.40	1.31	1.28	1.16	1.11	1.04	3.83	1.79	1.31	1.31	1.31
40600101	Neogene Flysch Gas	1.60	1.60	1.60	1.09	1.82	1.94	2.06	1.73	1.60	1.60	1.60
40600201	Thermal Triassic	1.40	1.19	1.16	0.80	0.89	1.37	2.82	1.47	1.19	1.19	1.19
40610101	Moesian Platform	1.49	1.49	1.48	1.50	1.28	1.50	2.05	1.58	1.49	1.49	1.49
40610201	Romania Flysch Zone	1.70	2.11	2.07	2.10	1.39	3.26	5.22	2.99	2.11	2.11	2.12
40610202	Romania Ploiesti Zone	1.40	1.26	1.25	0.92	1.35	1.40	1.44	1.28	1.26	1.26	1.26
40680101	Subsalt	1.67	1.67	1.66	1.32	1.62	2.01	2.52	1.87	1.67	1.67	1.68
52000101	Northeast Greenland Shelf Rift Systems	1.60	1.60	1.59	1.34	1.55	1.79	2.07	1.69	1.60	1.60	1.60
52150101	Jeanne d'Arc	1.70	1.83	1.83	1.91	3.30	0.63	3.31	2.29	1.83	1.83	1.83
52430101	Keg River Gas	1.70	1.72	1.71	1.28	1.68	2.71	1.75	1.86	1.72	1.72	1.72
52430102	Keg River Oil and Gas	1.70	1.91	1.89	2.31	1.35	2.27	2.00	1.98	1.91	1.91	1.91
52430201	Leduc Gas	1.70	1.86	1.83	1.79	1.75	1.57	3.89	2.25	1.86	1.86	1.86
52430202	Leduc Oil and Gas	1.51	1.51	1.51	1.37	1.58	1.48	1.73	1.54	1.51	1.51	1.51
52430301	Exshaw-Rundle Gas	1.62	1.62	1.61	1.00	1.79	2.18	2.43	1.85	1.62	1.62	1.62
52430302	Exshaw-Rundle Oil and Gas	1.70	1.91	1.90	1.76	1.98	1.69	2.93	2.09	1.91	1.91	1.91
52430401	Combined Triassic/Jurassic Gas	1.70	1.87	1.88	1.26	3.01	1.77	1.85	1.97	1.87	1.87	1.87
52430402	Combined Triassic/Jurassic Oil and Gas	1.70	2.03	2.03	2.06	2.31	1.49	2.54	2.10	2.03	2.03	2.03
52430501	Manville Gas	1.70	2.34	2.33	2.14	2.36	2.84	2.24	2.40	2.34	2.34	2.34
52430601	Second White Specks-Cardium Gas	1.70	2.76	2.67	3.79	1.51	2.17	0.00	2.49	2.76	2.76	2.76
52430602	Second White Specks-Cardium Oil and Gas	1.70	1.95	1.97	1.96	2.22	2.04	1.27	1.87	1.95	1.95	1.95

52440101	Yeoman Oil	1.70	3.00	3.00	3.61	5.68	15.25	0.00	8.18	3.00	3.00	3.00
52440201	Brightholme Oil	1.70	3.00	3.00	1.47	75.14	0.00	0.00	38.31	3.00	3.00	3.00
52440301	Bakken Sandstone	1.70	3.00	3.00	1.84	23.50	0.00	0.00	12.67	3.00	3.00	3.00
52440401	Lodgepole Oil	1.40	1.31	1.32	1.22	1.29	2.26	0.67	1.36	1.31	1.31	1.31
53050101	El Abra-Like Reef and Backreef Limestone	1.70	1.70	1.70	2.59	1.01	2.09	1.53	1.80	1.70	1.70	1.70
53050102	Tamabra-Like Debris-Flow-Breccia Limestone Overlying Evaporites	1.40	1.38	1.38	1.13	1.35	1.86	1.24	1.39	1.38	1.38	1.38
53050103	Tamabra-Like Debris-Flow-Breccia Limestone and Overlying Strata	1.70	1.88	1.85	1.73	1.39	3.09	2.46	2.17	1.88	1.88	1.88
53050104	Tamabra-Like Debris-Flow-Breccia Limestone of the Golden Lane	1.40	1.20	1.18	0.55	1.40	1.43	1.72	1.27	1.20	1.20	1.20
53050105	Tamaulipas-Like Basinal Limestone and Tertiary Strata Without Underlying Evaporites	1.40	1.01	1.04	1.94	1.31	0.51	1.03	1.20	1.01	1.01	1.01
53050106	Tamaulipas-Like Basinal Limestone and Tertiary Strata Overlying Evaporites	1.70	2.10	2.09	1.28	2.63	3.81	2.73	2.61	2.10	2.10	2.10
53050107	Tamabra-Like Debris-Flow-Breccia Limestone North of Campeche	1.70	1.82	1.81	1.51	1.78	2.14	2.61	2.01	1.82	1.82	1.82
60210101	Late Cretaceous-Tertiary Turbidites	1.44	1.44	1.43	1.21	1.38	1.60	1.86	1.51	1.44	1.44	1.44
60210102	Cretaceous Carbonates	1.70	1.78	1.77	1.45	1.75	2.11	2.57	1.97	1.78	1.78	1.78
60210103	Late Cretaceous-Tertiary Nearshore Sandstones	1.70	1.83	1.81	1.49	1.79	2.17	2.66	2.03	1.83	1.83	1.83
60220101	Amazon Delta and Submarine Fan	1.70	2.10	2.07	1.76	1.82	3.21	3.39	2.54	2.10	2.10	2.10
60290101	Western Pre-Aptian Reservoirs	1.40	1.24	1.26	1.68	1.26	1.74	0.53	1.30	1.24	1.24	1.24
60290102	Late Cretaceous-Tertiary Deep-Water Sandstones	1.40	1.37	1.36	1.01	1.47	1.30	2.32	1.52	1.37	1.37	1.37
60340101	Espirito Santo Shelf	1.70	1.88	1.88	0.57	7.20	1.03	0.00	2.93	1.88	1.88	1.88
60340102	Late Cretaceous-Tertiary Slide Blocks and Turbidites	1.45	1.45	1.44	1.20	1.39	1.62	1.91	1.53	1.45	1.45	1.45
60340103	Abrolhos Sub-Volcanic Structures	1.70	1.75	1.74	1.52	1.70	1.95	2.27	1.86	1.75	1.75	1.75
60350101	Late Cretaceous-Tertiary Turbidites	1.40	1.30	1.28	1.08	1.05	1.44	2.28	1.46	1.30	1.30	1.30
60350102	Cretaceous Carbonates	1.56	1.56	1.58	1.08	1.97	3.52	0.58	1.79	1.56	1.56	1.56
60350103	Salt Dome Province Tertiary Sandstones	1.59	1.59	1.58	1.28	1.53	1.86	2.26	1.73	1.59	1.59	1.60
60360101	Santos Shelf	1.63	1.63	1.62	1.46	1.53	1.80	2.12	1.73	1.63	1.63	1.63

60360102	Salt-Structured Deep-Water Sandstones	1.70	1.76	1.75	1.38	1.73	2.16	2.73	2.00	1.76	1.76	1.77
60370101	Pelotas Platform and Basin	1.70	1.95	1.93	1.55	1.94	2.44	3.06	2.25	1.95	1.95	1.95
60410101	Hollin-Napo	1.40	1.24	1.23	1.05	1.27	1.18	1.58	1.27	1.24	1.24	1.24
60410201	Ene	1.70	2.75	2.71	1.98	3.04	5.26	13.18	5.87	2.75	2.75	2.75
60450101	Sub-Andean Fold and Thrust Belt	1.40	1.35	1.34	1.67	1.02	1.48	1.40	1.39	1.35	1.35	1.35
60450102	Foreland Basins	1.66	1.66	1.65	1.25	1.74	1.93	2.32	1.81	1.66	1.66	1.66
60450103	Foreland Central Chaco High	1.68	1.68	1.67	1.37	1.64	1.96	2.35	1.83	1.68	1.68	1.68
60550101	Neuquen Extensional Structures	1.63	1.63	1.63	1.23	1.49	2.78	1.58	1.77	1.63	1.63	1.63
60550102	Neuquen Foothills Structure	1.52	1.52	1.50	1.13	1.40	1.85	2.68	1.76	1.52	1.52	1.52
60550103	Dorsal de Neuquen Structure	1.70	1.84	1.86	3.31	0.87	5.13	0.53	2.46	1.84	1.84	1.84
60580101	San Jorge Extensional Structures	1.40	1.22	1.23	1.20	1.65	0.89	1.24	1.25	1.22	1.22	1.22
60580102	San Bernardo Fold Belt Structures	1.70	1.76	1.78	2.32	2.80	0.73	1.63	1.87	1.76	1.76	1.76
60590101	Magallanes Extensional Structures	1.59	1.59	1.60	1.36	1.93	1.58	1.47	1.59	1.59	1.59	1.59
60590102	Andean Fold Belt Structures	1.70	2.28	2.26	1.85	2.33	2.95	3.78	2.73	2.28	2.28	2.28
60600101	North Falklands Basin	1.70	1.83	1.82	1.63	1.80	2.02	2.30	1.94	1.83	1.83	1.83
60600201	South Falklands Basin	1.70	1.83	1.82	1.63	1.80	2.02	2.30	1.94	1.83	1.83	1.83
60630101	Malvinas Extensional Structures	1.70	1.98	1.96	1.63	1.96	2.39	2.96	2.24	1.98	1.98	1.98
60810101	Cretaceous-Paleogene Basin	1.44	1.44	1.43	0.91	1.69	1.59	2.06	1.56	1.44	1.44	1.44
60830101	Neogene Pull-Apart Basin	1.45	1.45	1.46	1.13	1.47	3.12	0.67	1.60	1.45	1.45	1.45
60830201	Cretaceous-Paleogene Santa Elena Block	1.62	1.62	1.65	1.84	1.83	2.20	0.63	1.63	1.62	1.62	1.62
60900101	Northern	1.40	1.35	1.38	1.43	1.82	2.25	0.39	1.47	1.35	1.35	1.35
60900102	Southern	1.40	1.17	1.18	1.01	1.70	0.91	1.12	1.19	1.17	1.17	1.17
60900103	Eastern	1.70	2.22	2.21	1.28	3.37	3.16	4.14	2.99	2.22	2.22	2.22
60900104	La Luna and Older	1.70	2.82	2.79	2.38	2.83	3.88	6.70	3.95	2.82	2.82	2.83
60960101	Central	1.40	1.19	1.18	1.01	0.80	1.88	1.41	1.28	1.19	1.19	1.19
60960102	Peripheral	1.40	1.28	1.28	1.09	1.40	1.35	1.26	1.27	1.28	1.28	1.28
60980101	East Venezuela Fold and Thrust Belt	1.40	1.18	1.18	1.10	1.23	1.31	1.04	1.17	1.18	1.18	1.18
60980102	Guarico Sub-Basin	1.40	1.37	1.36	1.19	1.17	1.70	1.66	1.43	1.37	1.37	1.37
60980103	Maturin Sub-Basin	1.40	1.21	1.22	1.20	1.23	1.48	0.90	1.20	1.21	1.21	1.21
60980201	Trinidad Basins	1.43	1.43	1.42	1.82	1.02	1.78	1.28	1.47	1.43	1.43	1.43
60980202	Orinoco Delta and Offshore	1.62	1.62	1.61	1.42	1.57	1.78	2.05	1.71	1.62	1.62	1.62

60990101	Main Maracaibo Basin	1.40	1.38	1.37	1.00	1.45	1.50	1.95	1.47	1.38	1.38	1.38
60990102	Southwest Maracaibo Basin Fold Belt	1.41	1.41	1.40	1.28	1.12	1.69	2.13	1.56	1.41	1.41	1.41
61030101	Carupano Basin Gas	1.40	1.39	1.40	1.38	1.61	1.18	1.42	1.40	1.39	1.39	1.39
61070101	Inner Forearc Deformation Belt	1.70	1.72	1.71	1.38	1.68	2.05	2.52	1.90	1.72	1.72	1.72
61170101	North Cuba Fold and Thrust Belt	1.64	1.64	1.63	1.61	1.47	1.87	1.73	1.67	1.64	1.64	1.64
70130101	Coastal Plain and Offshore	1.52	1.52	1.51	1.09	1.38	2.26	2.01	1.68	1.52	1.52	1.52
71830101	Coastal Plain and Offshore	1.70	1.71	1.68	1.37	1.49	2.04	3.78	2.17	1.71	1.71	1.71
71920101	Agbada Reservoirs	1.40	1.39	1.38	1.19	1.34	1.37	2.10	1.50	1.39	1.39	1.39
71920102	Akata Reservoirs	1.51	1.51	1.50	1.23	1.46	1.72	2.03	1.61	1.51	1.51	1.51
72030101	Gabon Subsalt	1.68	1.68	1.64	2.17	1.12	1.09	15.96	5.09	1.68	1.68	1.69
72030201	Gabon Suprasalt	1.61	1.61	1.59	1.53	1.37	1.45	3.79	2.04	1.61	1.61	1.61
72030301	Central Congo Delta and Carbonate Platform	1.40	1.32	1.32	1.08	1.51	1.31	1.39	1.32	1.32	1.32	1.32
72030302	Central Congo Turbidites	1.70	2.09	2.06	1.64	1.86	3.71	2.83	2.51	2.09	2.09	2.09
72030401	Cuanza-Namibe	1.40	1.37	1.37	1.34	1.71	1.09	1.36	1.37	1.37	1.37	1.37
73030101	Offshore	1.44	1.44	1.45	1.53	1.10	3.68	0.55	1.72	1.44	1.44	1.44
80260101	Kohat-Potwar Intrathrust Basin	1.70	1.83	1.77	1.22	0.87	52.85	0.00	18.31	1.83	1.83	1.84
80340101	Sylhet-Kopili/Barail-Tipam Composite	1.53	1.53	1.56	1.25	3.37	1.11	0.83	1.64	1.53	1.53	1.53
80420101	Greater Indus Foreland and Foldbelt	1.57	1.57	1.57	1.18	1.90	1.69	1.61	1.60	1.57	1.57	1.57
80420102	Indus Fan	1.70	2.08	2.06	1.79	2.05	2.43	2.92	2.30	2.08	2.08	2.08
80430101	Eocene-Miocene Bombay Shelf	1.70	1.72	1.71	1.22	1.95	2.34	1.73	1.81	1.72	1.72	1.72
80430102	Eocene-Miocene Cambay Deltaic	1.60	1.60	1.60	1.37	2.00	1.35	1.87	1.65	1.60	1.60	1.60
80470201	Western Shelf and Slope	1.70	1.74	1.73	1.57	1.71	1.87	2.08	1.81	1.74	1.74	1.74
80470301	Central Basin	1.46	1.46	1.46	1.26	1.40	1.60	1.87	1.53	1.46	1.46	1.46
80470302	Eastern Fold Belt	1.40	1.38	1.38	1.13	1.17	2.25	1.18	1.44	1.38	1.38	1.38
80480101	Central Burma Basin	1.66	1.66	1.64	1.17	1.62	2.02	3.33	2.03	1.66	1.66	1.66
80480102	Irrawaddy-Andaman	1.40	1.35	1.34	1.21	1.30	1.21	2.09	1.45	1.35	1.35	1.35

The base case r given in table 7 is that calculated by algorithms as close as possible to those used in 1995. The differences between the 1995 methodology and that used in the base case for the analog database were that the calculations were done at an assessment unit scale rather than at a province scale (table 6, assumption 13) and the distinction that the modal class was not always class 6 (because of use of the lognormal distribution).

Assumption 2 restricts r to a range from 1.4 to 1.7. Tests were run allowing the broader range of 1.0 to 3.0; table 7 refers to this as the extended base case. Even more extreme values of r would have resulted if these limits were further extended. The extended base case was the case used for all the other sensitivity tests. That is, the other sensitivity tests used r's limited to the range from 1.0 to 3.0.

Assumption 3 refers to the use of the original 1995 least squares method to calculate r. To examine the sensitivity of assumption 3, a maximum likelihood measure was used to choose r. The maximum likelihood calculation was based on the fit to a multinomial distribution ("multinomial r" in table 7); the results (table 7) show no significant change in r.

An additional examination of assumption 3 was done with calculations of r based on individual ratios of counts in pairs of classes above the mean. The column in table 7 entitled "ratio 1 / 2" gives the ratio of the count of the class one step above the modal class to that of the count of that class two steps above the modal class. The columns "ratio 2 / 3," "ratio 3 / 4," and "ratio 4 / 5" are defined similarly. The "mean ratio" is the mean of these four values. As expected, "ratio 1 / 2" is closest to the extended base case value. Because of small numbers of fields in some of the higher classes, "ratio 4 / 5" can have extreme values, which affects the "mean ratio" also.

The columns "1.4 mean" and "1.38 mean" in table 7 examine the sensitivity of assumption 4 of table 6. They show minimal changes from the extended base case.

The "no <1 constraint" column in table 7 relates to assumption 5 of table 6. In the original 1995 methodology, any bins with a count less than 1 were not used in the calculations. By dropping this constraint, several assessment units with low field counts showed changes in r.

Unresolved Issues

The main unresolved issue for the methodology used in constructing our analog database is the form of field-size population distributions. An extensive literature argues that the underlying distribution of field sizes is either lognormal (Kaufman, 1993; Rose, 1994) or Pareto (Schuenemeyer and Drew, 1983; Houghton, 1988). Attanasi and Charpentier (2002) studied the effects of distribution type on the 1995 U.S. National Oil and Gas Assessment distributions. One alternative to the Pareto procedure used in our database may be to fit the small-field end of the distribution using robust regression estimation procedures.

Another important unresolved issue is the constraining of r to values of 1.4 to 1.7. In the original 58 assessed provinces of the 1995 National Assessment, 16 values were calculated as 1.4 and 29 values as 1.7 (table 8). Only 13 provinces had values of r between 1.4 and 1.7. Thus, the constraint significantly affected the calculation of r.

Table 8. Calculated r values by province from the 1995 USGS National Assessment (Gautier and others, 1995).

	Province *	Calculated r
1	Northern Alaska	1.40
2	Central Alaska	1.58
3	Southern Alaska	1.40
4	Western Oregon-Washington	1.70
5	Eastern Oregon-Washington	1.70
7	Northern Coastal	1.70
8	Sonoma-Livermore Basin	1.70
9	Sacramento Basin	1.55
10	San Joaquin Basin	1.40
11	Central Coastal	1.70
12	Santa Maria Basin	1.40
13	Ventura Basin	1.40
14	Los Angeles Basin	1.40
17	Idaho-Snake River Downwarp	1.70
18	Western Great Basin	1.68
19	Eastern Great Basin	1.70
20	Uinta-Piceance Basin	1.70
21	Paradox Basin	1.70
22	San Juan Basin	1.52
23	Albuquerque-Santa Fe Rift	1.70
24	Northern Arizona	1.70
25	Southern Arizona-South West New Mexico	1.40
27	Montana Thrust Belt	1.41
28	Central Montana	1.70
29	Southwest Montana	1.70
31	Williston Basin	1.70
33	Powder River Basin	1.69
34	Big Horn Basin	1.40
35	Wind River Basin	1.40
36	Wyoming Thrust Belt	1.40
37	Southwest Wyoming	1.40
38	Park Basins	1.70
39	Denver Basin	1.70
40	Las Animas Arch	1.40
41	Raton Basin-Sierra Grande Uplift	1.70
43	Palo Duro Basin	1.70
44	Permian Basin	1.42
45	Bend Arch-Fort Worth Basin	1.70

46	Marathon Thrust Belt	1.70
47	Western Gulf	1.40
49	Louisiana-Mississippi Salt Basins	1.58
50	Florida Peninsula	1.40
51	Superior	1.70
53	Cambridge Arch-Central Kansas Uplift	1.70
55	Nemaha Uplift	1.70
56	Forest City Basin	1.70
58	Anadarko Basin	1.51
59	Sedgwick Basin	1.57
60	Cherokee Basin	1.46
61	Southern Oklahoma	1.49
62	Arkoma Basin	1.40
63	Michigan Basin	1.70
64	Illinois Basin	1.65
65	Black Warrior Basin	1.70
66	Cincinnati Arch	1.70
67	Appalachian Basin	1.70
68	Blue Ridge Thrust Belt	1.70
69	Piedmont	1.40

* Only those provinces with small-field assessments are included in this table.

The present USGS assessment methodology uses a truncated shifted lognormal distribution to approximate the numbers and sizes of fields larger than some minimum field size. If a small-field extrapolation is applied to this distribution, additional fields are added to the size class or classes just above the minimum field size. This leads to two different numbers of fields larger than the minimum: the number of fields larger than size X (as originally assessed) versus the number of fields larger than size X (with small field extrapolation). The median and mean field sizes above the minimum field size are similarly affected. Additional methodological work is needed to connect the values in the analog database (which include the small-field correction) with the values needed for the assessment input form, which not only do not include the small-field correction but also are for the undiscovered portion of the field population.

Appendix 2. Database Files

For those users with Microsoft Excel, the following workbook includes the complete analog database as well as several additional tools to aid use of the database.

WorldAnalogs.xls

LINK TO DOWNLOAD FILE HERE

For those users without Microsoft Excel, the database is distributed among eight files. Each file is tab-delimited to make it easy to import into the spreadsheet program of your choice. The additional tools are not included. The eight files are:

WA_geology.tab the classification variables
WA_oil.tab the utility variables for oil fields
WA_gas.tab the utility variables for gas fields
WA_BOE.tab the utility variables for all fields (in BOE)
WA_oilbins.tab the counts of discovered and undiscovered oil fields by size class
WA_gasbins.tab the counts of discovered and undiscovered gas fields by size class
WA_BOEbins.tab the counts of discovered and undiscovered all fields (in BOE) by size class
WA_ancillary.tab the utility variables for ancillary data

LINKS TO DOWNLOAD FILES HERE