SPE

Society of Petroleum Engineers

SPE 24442

Data Management: A Case Study

J. Meek and P.H. Lowry*, K&A Energy Consultants Inc.

*SPE Member

This paper was prepared for presentation at the SPE Petroleum Computer Conference held in Houston, TX, July 19-22, 1992.

ABSTRACT

A successful field evaluation and enhanced oil recovery (EOR) program design was accomplished with the aid of a field database system. Rigorous data handling, data screening and a data entry front-end were concluded to strongly influence the utility of the working database. Versatile manipulation of the database allowed the extraction of data in a variety of forms. Mapping of the data revealed production patterns and indicated reservoir dynamics. The results were an accurate assessment of the remaining reserves and the formulation of an appropriate EOR program.

INTRODUCTION

For field reserves studies, a properly designed and implemented database can facilitate data organization and analysis. This paper is a case study contrasting an idealized database structure for field reserves evaluations and an actual database used to evaluate reserves for an existing field. It emphasizes the practical uses and pitfalls associated with assimilating large quantities of existing field data. It also demonstrates the value of flexibility in the database environment.

An idealized database is presented as a comparison to the case study in order to evaluate its effectiveness. Concepts of an idealized database are discussed including its structure and analytical uses. The case study includes various aspects of data acquisition, manipulation, and analysis. As a result of the field study, procedures were developed to assimilate existing data into a usable format. Experience gained in the adaptation of the idealized database supports its usefulness for future field studies.

Bibliography and illustrations at end of paper.

IDEALIZED DATA MANAGEMENT CONCEPTS

In performing field studies, the assimilation of large amounts of varied data is essential to a successful evaluation. A database provides a flexible tool for establishing data consistency and for data manipulation and reporting. Organizing data into functional relationships is the primary domain of database management. Once the data is in the desired form, characteristics and relationships in the data can be examined.

The idealized database for field studies is a set of relational data files that include all available relevant field information. The individual files include a general well and cataloging data file, a drilling and completion data file, a production data file, and a geologic data file. This idealized database concept is graphically depicted in Figure 1 along with the data elements associated with each portion of the database structure.

No single database structure is suitable for every field study. The basic data available will dictate the modifications appropriate in similar field studies. This paper outlines a basic approach that was useful for this particular field study. The adjustments to the idealized structure in this study are typical of the adjustments that may be needed in other projects. The principal philosophy behind the organization of these files is a chronologic ordering of activities related to well operations and production.

Well Identification File

The general well and cataloging file contains information unique to each well. This should include a unique well identifier, location and notations of data availability. Of critical importance is a unique well identifier that can be used for searches and to establish relations between the other data files. Locations of the wells are necessary for the creation of maps. The notations of data availability expedite data access by reducing searches to wells with existing data. Additional information in this file may include well specific data that does not change over time such as a spud date, well head elevation, etc. A convenient way to create a well identification file is to digitize an existing field map into a CAD program and extract the well numbers and locations into a database format.

Operations File

The drilling and completion data file, also called an operations file, includes information on specific operations performed on individual wells. Drilling, shows, DSTs, casing, completions, stimulations, pressure tests, plugging and shut-in operations may be among the file data. Each record has the well identifier, date of operation, type of operation, depth, interval name, and notes pertaining to the operation. If this file is properly constructed, a comprehensive history of individual wells or specific zones can be analyzed including periods of production and periods of well or zonal shut-ins, whether permanent or temporary.

Production File

Production data is kept chronologically, by producing interval and by well. This file is simply a recording of all produced and injected phases. Each record contains the production data associated with one given date. Time increments between records may vary depending on the method of data collection.

Geologic File

Geologic data include formation and reservoir parameters including tops and bottoms, porosity, and water saturation. These data may come from cores, logs and seismic interpretations. These data allow definition of gross structural and stratigraphic trends of formations and net reservoir intervals.

Data Quality Control

The results of a field study may be improved by controlling and polarizing data at several stages in database development and use. Use of front-end menus and error-checking routines is essential for consistent, error-free data entry. Routines for grouping and comparing data and bulk editing help eliminate remaining inconsistencies.

At this point, it is possible to use interactive queries and sorting in order to gain familiarity with the basic characteristics of the field. Now the project becomes a true field study instead of an exercise in data gathering. As familiarity with the field is gained, queries pursuing more specific analyses may be performed. Data reflecting these analyses may be formatted for reporting or exporting to other software for external processing.

External processing includes mapping of geologic and well history parameters. Geologic structure, net sand isopach, trim surface and derivative maps and cross sections may be prepared. Production parameters such as initial production, cumulative recovery, phase ratios, and projected ultimate recovery make excellent maps for field analysis.

CASE STUDY

Data Entry and Validation

A reservoir study was performed on a large southeast Asian oilfield that had been producing for nearly one hundred years. The study was undertaken to determine remaining reserves and the potential for tertiary recovery (EOR) methods. A data-gathering team was sent to Asia with the task of assembling all available data for the field in question. Thousands of pages of handwritten and typewritten data were microfilmed. Core was also viewed on-site and selected samples were brought back.

Upon return, the microfilmed data were reproduced and each page was numbered sequentially for reference. The pages were then segregated into the types of data represented. Most of the data were entered into database files. However, some tabular form data were entered into a spreadsheet program.

Routine core analysis and petrographic examination of thin sections of the core samples were performed at K&A Laboratories. Data from these analyses were entered into the geologic database file.

Field maps were digitized into a CAD package to provide coordinates for well positions. Exact latitude and longitude were not provided by the operator, so a grid system was employed to give relative positions to each well. This grid system provided the base for all subsequent maps.

At this point, preliminary attempts to analyze the data proved to be unsuccessful. Emphasis had been placed on getting the information into computers, instead of preliminary examination and screening of the data. The inconsistencies in this approach began to surface. Data that had been entered into a spreadsheet format were needed in a database format for comparison. Spreadsheet data were imported into database files and analysis was restarted.

Problems of duplication of data and inconsistencies in data entry became apparent at this point. In reanalyzing the hard copies, it was discovered that there was much physical duplication of the data. Many of the sheets had been copied more than once. However, a more insidious problem was identified. In the process of gathering and reproducing the data, there had been numerous cases of identical data being distributed in several different reports.

For example:

1. The original handwritten field notes would contain production data from readings taken at the site.

2. This report would then be transcribed into an identical typewritten report.

3. These numbers would then be assimilated into a monthly field report showing individual well production from an area.

4. Another annual report would be made for each well, using the monthly reports and reorganizing them into well-by-well production reports.

5. Reports would also be generated monthly or annually for producing intervals, fault blocks and field-wide production.

6. Other various reports would regroup pieces of data from several reports into a single report.

A major effort was necessary to eliminate the duplicates. A program in the database language was written to handle the majority of the problems. However, a large number of man-hours were still required to manually search the database files and consolidate duplicate information.

Consistency was deemed a primary concern. In order to facilitate data organization, a set of abbreviations was established for well identification, producing intervals and fault blocks.

In the United States, API numbers are excellent for well identification. They are unique and tied to a specific well. In this case, there were no unique numbers to specify the wells creating confusion. There had been as many as six operators at any given time producing from the field. They had all started numbering their respective wells at one. This meant that there were as many as six wells with any given number. So locations, numbers and producing companies had to be cross-referenced to establish exact well identities. A system was established in which the companies were given a two-digit prefix from 10 to 60. These were added to the well numbers to give each well a unique identifier. For example, wells 100001 through 101105 were operated by the first operator. Wells 600001 through 600114 were operated by the sixth operator.

The productive intervals were given unique numbers roughly corresponding to their depths. The theory behind this was that sorting on zone numbers would present an ordering of the producing intervals by their relative positions in a given well. Thrusts, subthrusts, and A and B designations were handled with decimal additions to a given zone number, such that the designation 6.100 would mean the sixth zone (Zone F), the A member.

Fault block areas were given appropriate, unique abbreviations. For example, the letter S was used as an abbreviation for the South fault block. This facilitated sorting and made the files slightly more compact. Also, this simplified entering any new data by fault block areas. Small files were generated to cross-reference the abbreviations in the data files with the full designations of fault blocks and producing intervals. This facilitated the use of the full designations in the generation of reports.

The data review also found one more major inconsistency. Data that had been originally interpreted as one type (casing) was actually a completely different type (perforation). These data, previously not deemed critical to reserve estimates, was now considered to be a necessary part of the operations file. This necessitated a change in the structure of the operations file, which was modified to accept these new data without changing the way the data had been entered. This was accomplished by adding extra fields to accommodate the new data. The modification produced a large number of empty data fields within the file. This presented file size problems. However, this was considered preferable to reformatting or reentering the entire file to conform to the existing file structure.

<u>Data Analysis</u>

After making the data consistent, analysis of the data was conducted with the idea of consolidating it into discernable patterns. This was accomplished using the relational powers of the database in conjunction with the query and filter functions. With the available information and lack of any positive patterns from the operator, the initial queries were somewhat broad and even inefficient. Some of the preliminary search patterns included:

1. Initial production - by well, date and zone

2. Production - by well over time

3. Cumulative production - by zone and fault block

4. Shut-ins - by date and reasons

This broad approach was necessary to gain familiarity with the data. This led to the design of more specific queries that isolated data relevant to specific aspects of the field.

An example of a straightforward query is to index the database on well numbers and dates. Filters would then be set for the desired data parameters. For example, one way to find all the production in a given fault block, would be to set the index to the well numbers plus the date and then set the filter to the desired fault block. Output of the production data would then be by well, by date and by fault block.

Several observations came from these preliminary searches. Comparisons were made between cumulative production by fault block and a summation of all individual well production for the given fault block by date. By checking which block totals equaled the sum of the individual well production in the block, it became apparent which blocks had production information available on all producing wells. Also, since these block totals exactly matched the totals of individual wells, it was determined the most likely method of production reporting had been an actual measurement of total production for a fault block. Individual well production was then allocated by prorating cumulative totals to the individual wells. This method was verified by a consultant from the current operator.

Next, selected areas of the field were more closely scrutinized. Charts, graphs and maps of the data were prepared. Bubble maps were produced of initial production by date, most recent production, total liquid (oil and water) produced by well, shut-ins by date and wells shut-in due to high gas-oil ratio (GOR) or high water-oil ratio (WOR) by date.

To create the bubble maps, queries were set up in the database files that segregated the desired data. Text files of well identifiers, locations, dates and production information were output for use in a CAD program. The CAD program then took the data, referenced it to the well coordinates and drew a bubble of the appropriate color, representing the date, and size, representing the magnitude of the given data. Examples of bubble maps are given in Figure 2 and Figure 3.

Much of the data for the field production were output in a spreadsheet format for calculations and graphing. This was accomplished by first indexing the wells, and then setting filters to isolate the desired data. Using the built-in transfer capabilities of the database program, the data were exported directly into a spreadsheet format. The spreadsheet program could then be used to produce line graphs, bar charts or other reports.

Log and petrographic analyses resulted in the revision of net thickness and porosity values for most of the pay zones. The impact of this was the reduction of original-oil-in-place (OOIP) values. The operator-supplied recovery values had suggested that approximately 35% of the original-oil-in-place had been recovered. The revised values indicated that approximately 54% of OOIP had been recovered. Doubts began to be raised about the success of any EOR procedure.

Another major question answered during this process was why were certain wells not producing to their expected capacities. Several pay zones had been determined to have high potential recoveries prior to World War II. These zones had been left untouched as a reserve. Then after the war, when field production resumed, these untapped zones were reopened with disappointing results. Performance behavior in differing parts of the field required varying explanations.

In the central part of the field several wells, and especially one well, showed high initial production potential from one particular zone (Zone E). The next lower zone (Zone F) showed promise, but was not considered exceptional. Over the years Zone F was produced while Zone E was held in reserve. When Zone E was finally reopened for production, results were extraordinarily disappointing. No particular reason could be found by the operator, and it was attributed to poor testing of the initial producing wells.

The migration of oil between producing intervals, either along faults or wellbores, was identified by using the field database. A query was designed to isolate the two zones (E and F) and the two fault blocks in questions. Cumulative totals were acquired for the two zones and cross-referenced with the revised estimates for original-oil-in-place. It was determined that recovery in Zone E was approximately 47%, which was below the field average of 54% OOIP. Zone F had produced better than average, at about 61% of OOIP. Maps of initial production by zones and date and production at a later dates showed that Zone F was producing its oil from the same geographic location Zone E had been expected to produce. It was determined that oil was migrating from Zone E to Zone F. This resulted in depletion of Zone E with very few wells having actually been completed in that zone.

The field database provided a different conclusion for a similar problem on the downdip side of the field. In a lower producing zone the thickness of the oil column was found to be thinning with time due to encroaching water and an expanding gas cap. In this portion of the field, three en echelon sand bodies were identified as potentially productive and completions were made in each. These were treated as individual reservoirs. Similar to Zone E, the lower zone was shut-in as a reserve. When the zone was reopened, the water-oil ratio was too high for continued production. Queries similar to those performed for Zones E and F were performed for this portion of the field. However, unlike in Zones E and F, the adjacent zones did not show definitive patterns of oil migration between zones.

Migrations of the water-oil contact and the gas-oil contact were established relative to all three zones by performing queries designed to show the shut-ins for this portion of the field. These queries were further modified by limiting the shut-in wells to those known to have been shut-in for high GOR or high WOR. These data were imported into the CAD program and maps were plotted. The shut-in wells were color coded for date and whether they were shut-in for high GOR or high WOR. The maps showed that not only was the producible area of oil narrowing, but the oil column was also migrating updip away from the current well locations. This explained why the producing area was getting more difficult to locate with new wells.

Project Recommendations

Recommendations for future field development and EOR projects were now considered. Considering the depletion of the field, recommendations were made to modify recovery operations and reduce or eliminate new drilling in the field.

A thorough search of current literature was performed, in part by using an existing database listing many pertinent EOR articles published during recent years. Several potential techniques were studied in order to find one that met the conditions of the reservoir and the guidelines of the operator.

Based upon information from the operator, the database and a few reasonable assumptions about field conditions, simulations were run to determine the probable behavior of several EOR projects. These simulations indicated several behavioral problems such as a strong natural water drive and energy conservation.

It was determined early that costs were of primary concern. Contemporary injection projects, such as carbon dioxide and nitrogen injection, were ruled out due to costs. Any projects of a more elaborate nature were also ruled out for similar reasons.

It was decided that a hydrocarbon gas injection project was a workable solution. Using the information queried from the database, it was determined that a gas injection scheme had been used previously to offset the energy from the water drive. This had been successful at preserving the energy of the reservoir. A recommendation was made for two pilot injection projects, to be followed by field-wide injection contingent upon the success of the pilots.

Final decisions were made and a unified recommendation for an EOR project was developed. The final proposal was then printed and delivered to the client for consideration.

CONCLUSIONS

A field study project is usually a large and complex undertaking. The work required will depend greatly on the characteristics of a field and the complexity and availability of data. With experience gained from this project, several conclusions are provided to assist future projects of similar scope.

1. The need for pre-screening of data cannot be overemphasized in projects such as this. Many of the problems encountered in this study could have been minimized or eliminated by a more thorough pre-screening of hard-copy data. Duplication of data should be reduced as much as possible during this step.

2. Characterization of data during the pre-screening is also important. If the data can be better categorized during pre-screening, it serves to better define any modifications necessary to the initial database structure.

3. A rigorous and concise nomenclature should be developed early in a project. It can be the key to data consistency. Unique well identifiers are a must. Faults, fault blocks, productive intervals and any field parameters that may be used to segregate or analyze data should have a unique naming or numbering system.

4. A well thought out data file structure should be agreed upon and adhered to as closely as possible throughout the project.

5. Proper front-ending of the database can also help eliminate any residual duplication and facilitate data entry. A properly designed front-end is logical and similar in nature to the data to be entered. This allows a rhythm to be achieved and maintained during data entry. This not only speeds the entry process, but actually may help reduce errors. Another advantage to proper front-end design is that data may be entered by persons without the technical training to interpret the original data. It is much easier to train a person to enter the data when it follows a logical pattern. A proper front-end should allow for some flexibility in data entry, check for duplication of data and screen out anomalous data to a practical extent.

6. Using a database with advanced programming and querying facilities greatly improves the interactive search process. Queries may be refined or discarded as data patterns are discerned. This speeds data analysis and also affords more directions of potential investigation within the time allocated to the project. The logical and intelligent refinement of searches and queries is a key part in the success of a project.

7. Using graphical representation of data, such as bubble maps and graphs, improves the interpretation of the available data. Historical production trends may be visually interpreted using the maps. This may lead to more concise analysis in less time than traditional interpretation methods alone.

8. Efforts may be rewarded with mutually confirming geologic and engineering analyses. This agreement in analyses will usually lead to a more credible assessment of a given project.

Overall conclusions indicate a thoroughness in data handling at the beginning of a project will continue to be rewarded throughout the life of the project. Data screening, classification and removal of duplications can save hours of time and produce much cleaner results in the long term of a project. Time spent to manufacture a consistent and user-friendly front-end for data entry will be compensated for by the time saved by not having to refine the data after entry.

The usefulness of a database in a project of this type is difficult to overstate. Numerous investigations may be performed in a relatively short period of time. Consistency of available data leads to a more accurate evaluation of the field in question. A field study project of any size may benefit by using database analysis to supplement more traditional methods.

EQUIPMENT AND SOFTWARE

To offer insight into the magnitude of a field study such as this, the equipment and software used is listed below:

<u>Equipment used in this study:</u>

IBM XT, 640K RAM, 40M hard drive - data entry (4 machines)

286-6 PC, IBM Compatible, 640K RAM, 40M hard drive - data entry, spreadsheet, word processing

286-12 PC, IBM Compatible, 1M RAM, 40M hard drive - word processing

286-16 PC, IBM Compatible, 3M RAM, 40M hard drive - database analysis, spreadsheet, word processing

386SX-16 PC, IBM Compatible, 2M RAM, 40M hard drive - database analysis, spreadsheet, word processing

386SX-16 Laptop PC, IBM Compatible, 5M RAM, 40M hard drive - simulation studies, spreadsheet

386-33 PC, IBM Compatible, 16M RAM, 660M hard drive - for simulations studies

386-20 PC, IBM Compatible, 5M RAM, 130M hard drive - for database analysis, spreadsheet, word processing, CAD graphics, presentation graphics

VAX 11/780 Mainframe - database analysis, simulation studies

Calcomp Digitizing Tables - data entry (2 machines)

HP Laserjet II - report printing, black and white graphics printing (2 machines)

HP Paintjet XL - color graphics printing

Printronix Model 600 Lineprinter - bulk data dumps

Digital LA-120 Dot Matrix Printer - data printouts, report printing (3 machines)

Calcomp 1041 Plotter - black and white maps, isopach maps, contour maps

Versatec Color Plotter - color maps, stratigraphic cross-section

Software used in this study:

AutoCAD Version 10 - mapping, presentation graphics

BOAST II Black Oil Simulator - waterflood simulations, simulation of historical production parameters

dBase IV Version 1.1 - database analysis, report generation

CorelDraw Version 2.0 - presentation graphics

CPS-PC Version 4.11 - contour mapping, isopach mapping

Lotus 1-2-3 Versions 2.1 & 3.1 - spreadsheet analysis, presentation graphics

QEdit Version 2.1 - data cleanup, text editing

Pipeflo (in house software) - pipeline flow simulator

TerraStation Version 5.10 - log analysis, stratigraphic cross-sections

TETRAD Version 8.5 - reservoir simulation, steam injection simulations (huff 'n puff)

Windows Version 3.0 - operating system for CorelDraw

Wellflo (in house software) - well flow simulator

Word Perfect Version 5.1 - data cleanup, word processing, final report

ACKNOWLEDGEMENTS

The authors would like to thank K&A Energy Consultants Inc. for permission to publish this paper.

BIBLIOGRAPHY

Harvey, Greg. The Complete Lotus 1-2-3 Release 3 Handbook, Sybex, Inc., Alameda, CA 1990.

Harvey, Greg, and Kay Yarborough Nelson. Lotus 1-2-3 Instant Reference Release 2.2, Sybex, Inc., Alameda CA 1989

Liskin, Miriam. Liskin's dBase IV Programming Book, Osborne McGraw-Hill USA 1990

Omura, George: AutoCAD Instant Reference, Sybex, Inc., Alameda CA 1989

Rubel, Malcolm C. dBase IV Power Tools: Advanced Functions, Procedures and Utilities, Bantam Books 1989

Simpson, Alan. dBase III Plus Programmer's Reference Guide, Sybex, Inc., Alameda CA (1987)

Simpson, Alan. dBase IV Programmer's Instant Reference, Sybex, Inc., Alameda CA (1989)

Simpson, Alan. dBase IV User's Desktop Companion, Sybex, Inc., Alameda CA (1989)

Simpson, Alan. dBase IV User's Instant Reference, Sybex, Inc., Alameda CA (1989)

Thomas, Robert M. AutoCAD Desktop Companion, Sybex, Inc., Alameda CA (1989)

Townsend, Carl. Mastering dBase IV Programming, Sybex, Inc., Alameda CA (1989)

CONCEPTUAL DATABASE DESIGN

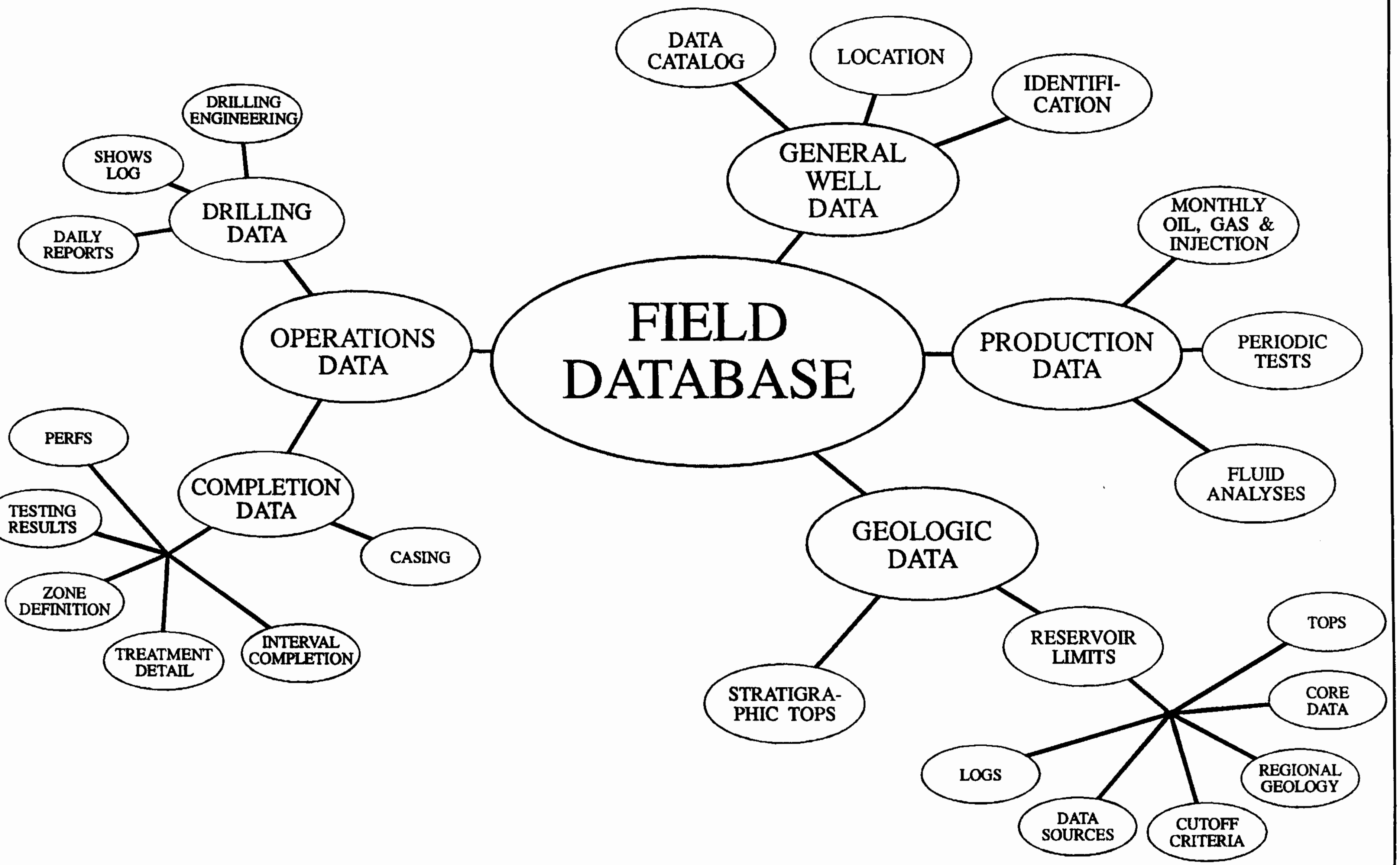

FIGURE 1

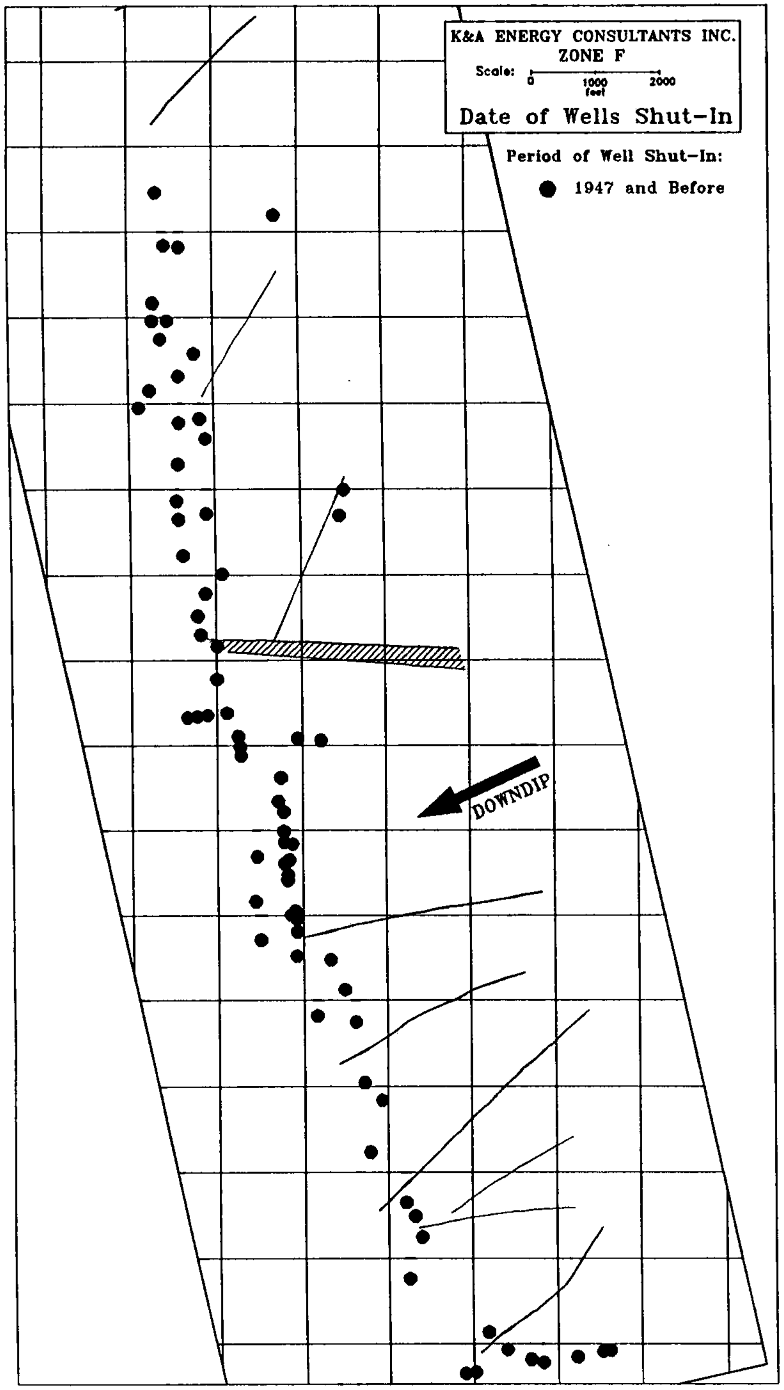

FIGURE 2

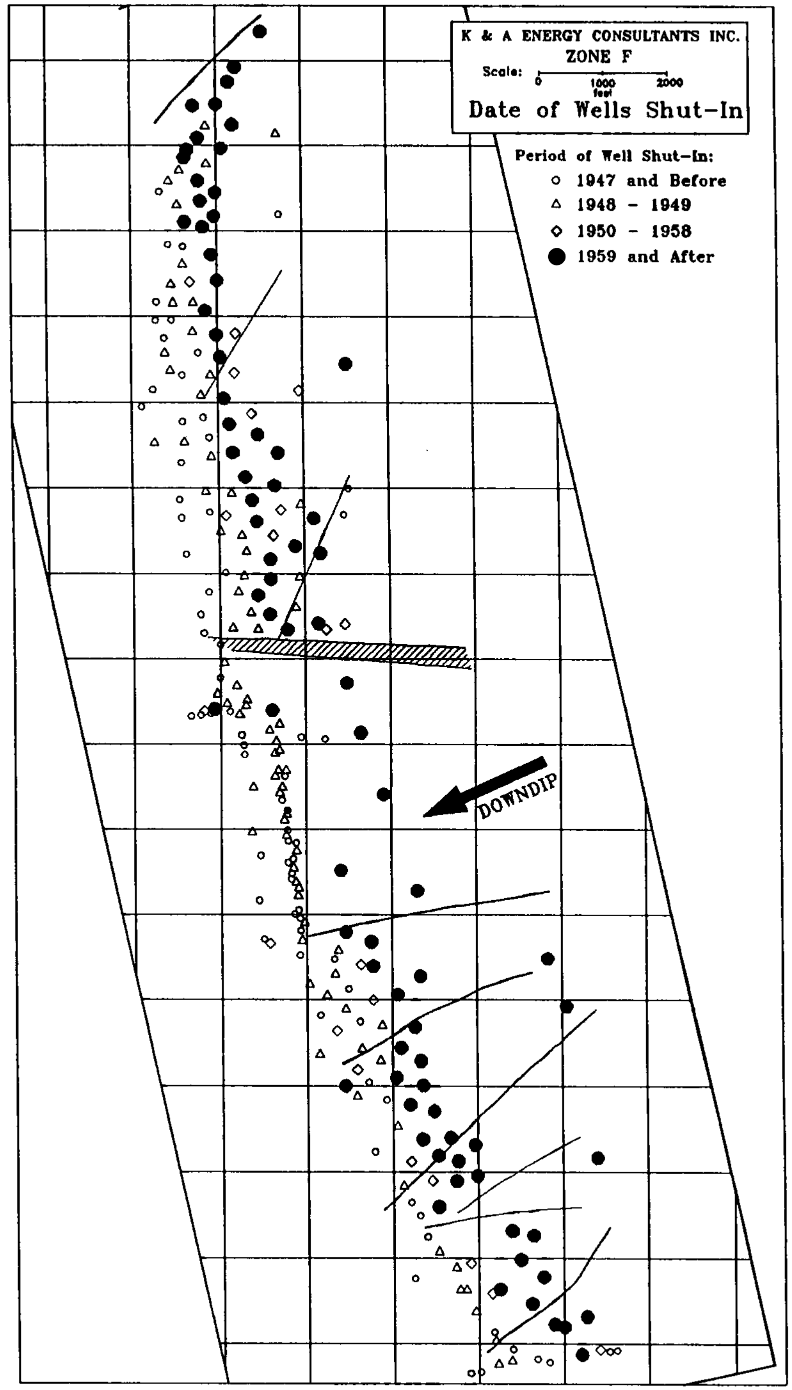

FIGURE 3

SPE 24442