A Soft Computing-Based Method for the Identification of Best Practices, With Application in the Petroleum Industry

CM-5024

Shahab D. Mohaghegh & Razi Gaskari, West Virginia University

2005 IEEE International Conference on Computational Intelligence for Measurement Systems & Applications – 20-22 July 2005 – GIARDINI NAXOS - TAORMINA, SICILY, ITALY

ABSTRACT

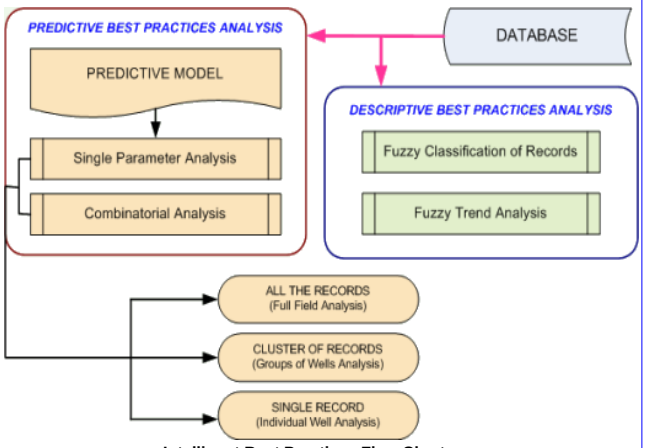
This paper introduces a new and novel methodology for fully data-driven best practices identification and analysis based on soft computing techniques. Using this new methodology "best practices" in any operation (industrial or otherwise) can be identified and recommendations can be made on how to conduct the operation in future in order to accomplish the best results. Since this is a fully data-driven process, no managerial or other biases will enter the process and it allows the data to speak for itself. In order for this method to be applicable it is required that sufficient amount of data for the process under investigation be available. The methodology uses neural network to build a representative predictive model of the process, fuzzy logic to provide analysis of the existing practices and genetic algorithms to identify the major trends. A recommendation matrix at the end of the process is developed that would serve as the foundation for making best practices recommendation. Application of this methodology to a problem in the petroleum industry is presented.

INTRODUCTION

Identification of best practices in many industrial operations is gaining unprecedented momentum. Companies that have gathered large amounts of data now realize that they own a valuable commodity that can play an important role in increasing efficiency in their day to day operations. The question is how this vast amount of data can be used in order to help the company's bottom-line. This paper attempts to address this question by introducing a newly developed methodology that enables companies to deduce information and knowledge from the existing data. The deduced information and knowledge can then be used in developing business rules and decision making.

As the volume of data increases, human cognition is no longer capable of deciphering important information from it by conventional techniques. Data mining and machine learning techniques based on soft computing must be used in order to deduce information and knowledge from the raw data that resides in the databases. The Intelligent Best Practices Analysis – IBPA that is introduced here incorporates the state of the art in data mining and machine learning to assist professionals in making the most of their existing data.

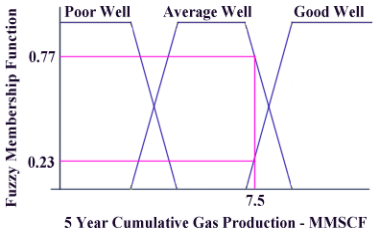
IBPA starts by identifying a process outcome. The process outcome is a dependent variable in the database that is used in order to measure the degree of success of a particular practice (an objective function). For example "return on investment" would be an appropriate outcome for identifying the best practices in a project evaluation process.



Intelligent Best Practices Flow Chart.

DESCRIPTIVE BEST PRACTICES ANALYSIS

Descriptive Best Practices Analysis tries to find and display patterns that exist in the database using a fuzzy averaging technique. It does not manipulate the data in any shape or form. It simply presents data in a new light that makes detection of existing trends and patterns possible. This is performed in two steps. First the process output (the dependent variable) is classified using fuzzy set theory.



Fuzzy classification of a particular well (record) in the database.

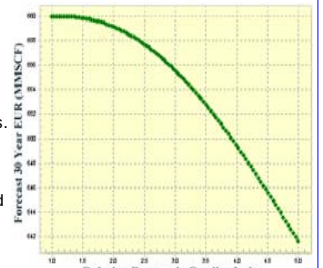
For example when "five-year cumulative production" is used as the outcome, a well can be classified as poor, average or good. If "return on investment" is used as the outcome the ROI of a project can be classified as low, average or high. Upon completion of this step all the records in the database are classified using fuzzy membership functions. Each of the independent variables in the database can be analyzed in order to see if a trend or pattern can be identified and then plotting them in the form of a bar chart. In the figure shown, while the average value of permeability for all of the wells in the database is about 0.09 md, poor wells have an average permeability of 0.04 md and good wells an average permeability of 0.33 md. This is an obvious trend and very well expected but for other parameters that are not necessarily so obvious, this kind of visualization of data can reveal important information.

Fuzzy averaging of reservoir permeability for all the wells (records) in the database.

PREDICTIVE BEST PRACTICES ANALYSIS

The main idea behind the predictive analysis is to fill the gaps in the solution space in order to make a comprehensive analysis possible. A predictive model is built based on the existing data. This model provides a continuous hyper-dimensional surface that is full of hills and valleys. This surface covers the entire solution space and the records in the database are discrete points on this surface. Goal of predictive analysis is first, to develop a predictive model that can accurately approximate this solution space and second to exhaustively search through, and query the solution space in order to identify patterns that can be used as guides in the decision making process. This step of the analysis is performed using genetic algorithms. The neural network model developed in the first step is used as the objective function of the genetic algorithms.

Once a predictive model is developed, single and multiple-parameter analyses are performed during which the effect of independent variables on the process outcome is studied and recorded. This figure shows a single-parameter analysis. From the recorded results of single and multiple-parameter analysis a recommendation matrix is generated. The recommendation matrix is then used in order to draw conclusions from the analysis and make operational recommendations pertaining to the best practices for the process being studied.

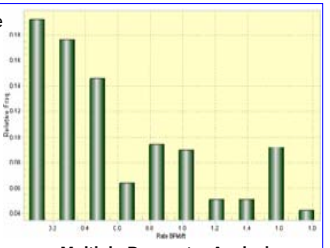


Single-Parameter Analysis

This figure shows the result of multiple parameter analysis, and the matrix below shows the example of summarizing the results of both single and multiple parameters analyses.

Parameter	Single Parameter Analysis: Percent Population	Dominant Trend	Change in Value	Combinatorial Analysis: Distribution	Dominant Trend	Recommendations
Water	Majority	Decreasing	Low	Skewed	Low Values	Do Not Recommended
Oil	All	Increasing	Moderate	Skewed	High Values	Use Recommended
Acid	All	Increasing	Moderate	Skewed	Low Values	INCONCLUSIVE

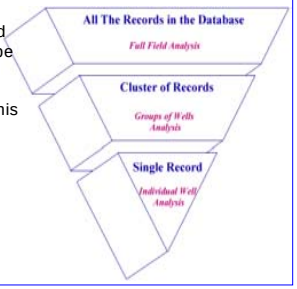
The Recommendation Matrix



Multiple-Parameter Analysis

Upon completion of the analysis for the entire database the records in the database are classified and the process is repeated for each class (category or cluster) of the records. The classification can be based on information already provided in the database (native) or based on k-mean clustering or fuzzy c-mean clustering performed on the database. Once the classification of the records is completed, the best practices analysis is performed on each group of the records and the result are tabulated in recommendation matrices as shown before. The last part of the analysis would be performed on individual records.

Single and combinatorial analysis is performed for one record at a time. Based on the nature of the database and the type of process it describes this step can be important or redundant. The relationship between these analyses is shown here. This figure shows an inverted pyramid that increases in precision and decreases in averaging as one moves from the top of the pyramid (analyzing all the records in the database or full field analysis in the case oil industry) to cluster of records (groups of wells) and finally to the single record (individual wells).



CONCLUSIONS

A new methodology has been developed and introduced for identification of best practices in any process that can be described by a dataset. The new methodology is based on soft computing techniques such as artificial neural networks, genetic algorithms and fuzzy logic and has been named Intelligent Best Practices Analysis. Intelligent Best Practices Analysis (IBPA) is a two-step process that includes a descriptive and a predictive analysis. During the descriptive analysis fuzzy-averaging of parameters is performed in order to identify the trends that are present in the database. These trends and patterns usually provide a strong foundation for the best practices that are ultimately identified.

Predictive best practices analysis is a drill-down process that starts with all the records in the database and ends with individual records. During the complete database analysis the best practices are identified based on all the records in the database. The database is then divided into groups, based on different criteria, and each group is analyzed separately in order to either verify, refine or dispute the best practices that were identified during complete database analysis.

The last step of the predictive analysis is working with individual records and using the results of past two steps as a guide to enhance our understanding of each individual record. It has been observed that although the best practices that were identified during the complete database analysis and groups of records analysis would work for most of the records, there will be records in the database that would not necessarily follow the identified trends as expected. This makes analysis of individual records a highly recommended exercise. This process was successfully applied to several oil filed related databases.