CHEMICAL ENGINEERING TRANSACTIONS

VOL. 33, 2013

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at: www.aidic.it/cet

Guest Editors: Enrico Zio, Piero Baraldi

ISBN 978-88-95608-24-2; **ISSN** 1974-9791

DOI: 10.3303/CET1333138

Optimization Models and Prediction of Drilling Rate (ROP) for the Brazilian Pre-Salt Layer

Carlos M. C. Jacinto[a], Paulo J. Freitas Filho*[,b], Sílvia M. Nassar[b] , Mauro Roisenberg[b] , Diego G. Rodrigues[b] , Mariana D. C. Lima[b]

[a]CENPES – PETROBRAS RESEARCH CENTER/PDGP/PCP - Rio de Janeiro, RJ, 21941-915, BRAZIL
[b]Dept. of Informatics and Statistics - PPGCC-Federal University of Santa Catarina
P.O.Box 476, INE – CTC – UFSC. Florianópolis, SC, 88040-900, BRAZIL
freitas@inf.ufsc.br

This article presents an on-going research that addresses the optimization of the cost of drilling wells in environments of high complexity and risk such as those related to the pre-salt region offshore Brazil. The minimization of these costs is directly related to the maximization of ROP (Rate of Penetration). The metric cost, i.e., the cost per meter of drilled not rely solely on the ROP. It directly involves minimization of two major components of that cost: the cost of drills and the operations cost. In short, the better combination number of bits used (generally smaller is better) versus meters drilled, increased ROP and reduced the cost per meter drilled. Finding the best combination is a difficult task. The ideal way to this nirvana is a good planning well and good control of the operation during the process. In such scenarios, it is essential to utilize software tools able to predict and improve the rate of penetration. Such tools must consider both extrinsic drilling skills training, as well as the effects of drilling parameters and drill wear. Such computer systems may, for instance, come to provide reliable alternatives to drilling plans and/or considering alternative plans presented, confirming them or rejecting them based on the information available. The process of creation and implementation of computational models capable of predicting and optimizing the rate of penetration (ROP) in the pre-salt wells is not trivial process. In this research the following techniques are investigated: a Bayesian inference approach for targeting the elicitation process and subsequent combination of models; and a Dynamic Evolving Neural-Fuzzy Inference System (DENFIS). We present the results of this investigation to date, the relevance of the proposed approach and the future prospects of their use for the delivery of viable solutions to the problem.

1. Introduction

In order to reduce costs it is necessary to accurately plan offshore oil drilling operations. The time required to successfully drill a well has to be estimated fairly precisely, since most of the costs associated are tied to the rental of equipment required for the operation as reported by Gandelman (2012); however, each operation has unique properties that make this task highly difficult. Many properties vary, such as rock type, rock porosity, gas presence, pressure, drill bit wear rate among others. All these properties affect the ROP, as well as many other parameters which are controlled by a drilling operator: weight on bit(WOB), revolutions per minute(RPM), bit type, bit diameter, bit wear rate, hydraulic horsepower per square inch(HSI).

Most of the work in the planning phase is restricted to adjusting bit type and diameter, RPM and WOB in order to achieve an acceptable ROP. To optimize this work many systems using artificial neural networks (ANN) were proposed in the past to predict the rate of penetration (ROP) for the project planning phase such as Bilgesu et al. (1997) and even choose automatically some parameters such as RPM and WOB in Fonseca et al. (2006). Unfortunately for the available data on the Brazilian pre-salt layer these systems did not achieve a reliable result due to the poor quality and scarcity of data. To overcome these problems we investigated two alternative approaches: a Bayesian Network (BN) inference approach for targeting the elicitation process and subsequent combination of models; and a Dynamic Evolving Neural-Fuzzy Inference System (DENFIS).

This work is structured into 8 sections. Besides this introduction, sections 2-5 are devoted to report the use and preliminary results for the Bayesian networks approach. Sections 6 and 7, report the use of neural networks for finding solutions to the problem of forecasting and optimization of ROP. In section 9 we present a resume and a discussion about the preliminaries results of this ongoing research.

2. The Bayesian Inference Approach

A Bayesian Network approach for such a problem is relatively recent and publications focus on how to determine a good topology for the network. Rajaieyamchee (2009) shows the use of *Influence diagrams* (ID), also known as *Bayesian Decision Networks*, to have a good quality when faced with real situations involving drilling in the North Sea. Giese (2011) uses ID and interviews with experts in the field to make a topology of a Bayesian network that aim assist in decision making for engineers when designing the treatment of drilling fluids in Saudi Arabia. Recently Al-Yami (2012) presents a topology to aid the drilling fluids practice in Saudi Arabia and also shows the Bayesian network as an efficient alternative of the flow charts, since it's not necessary to constantly update them.

3. Basic reference of Bayesian Networks

A Bayesian network (Pearl, 1988) is a model of representation and reasoning of uncertainty using the conditional probability between variables of a specific domain and expressed via directed acyclic graph (DAG). Its graphical structure can tackle the correlations between variables effectively, with appropriate language and efficient resources to represent the joint probability distribution over a set of random variables (Friedman, 1997).

Defining formally, a Bayesian network is a pair(G, P), where $G = (G, E)$ is a DAG where the nodes $V = \{V_1, \ldots, V_n\}$ represents the variables and edges $E = \{E_1, \ldots, E_n\}$ represents a direct correlation between each node in V. P is defined as a set of probabilistic parameters expressed through tables: given a particular variable a conditional probability distribution is made for each of their classes (states) joining each class of their parents. With that configuration, the network establishes that a variable is independent of all other variables except their descendants in the graph given the state of its parents.

Considering all the possible network topologies for a Bayesian network the well-known structure naïve Bayes is the simplest one. It assumes that all variables are mutually independent given the class context. Although this model does not reflect the reality in most real-world tasks it is very effective, because the parameters of each attribute can be learned separately, facilitating the learning process (McCallum, 1998). The naive Bayes topology is therefore a set of mutually independent variables that works as the input which collectively has a single parent (output node).

4. Methodology for the BN Approach

The naive Bayes structure was used for ROP's classification as a first exploration of the problem. To achieve this goal interviews were made with experts in the field and chosen four input nodes and the output node (ROP). Since the variables are numeric and continuous, it was required to turn them into classes for using the Bayesian approach. Table 1 shows the resulting configurations after this process of discretization.

Table 1: Nodes of the Bayesian network after discretization

Node's name	Type	Number of classes
Bit diameter	Input	6
UCS (unconfined compressive strength)	Input	6
WOB (weigh on bit)	Input	3
RPM (revolutions per minute)	Input	3
ROP (rate of penetration)	output	3

Although this approach shows the output as the result of a classification process, the ROP's prediction generally requires a numerical output. With the objective of attend this demand, a numerical range of ROP values were used to calculate the average and standard deviation for each classified register.

5. Results for the BN Approach

5.1 Dataset

To test the Bayesian network, we selected data from three wells in the pre-salt layer which in their totality have 4863 records. The data set was divided into two parts: 60 % for the inputs (training set) and the remaining 40 % to the output (test set). The final result can be seen in Table 2.

Table 2: Training and Test's set

Well	Type	Number of records	**Total**
#1	training	343	**2799**
#2	training	2456	
#3	test	2055	**2055**

5.2 Bayesian Classification

The naïve Bayes structure was used and its topology can be seen in Figure 1.

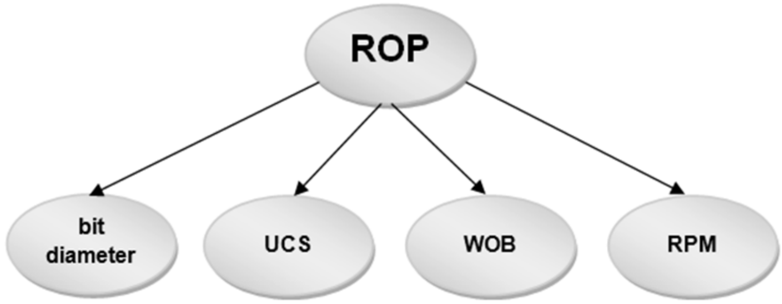

Figure 1: topology of the naive Bayes structure for ROP's classification

An important aspect to note is that some classes present in the test set were not observed in the training set. Looking at the data as a whole, we can see that the naive Bayes network is not able to learn and classify a new sample with good accuracy (18.59%). The classification matrix can be seen in Table 3 and shows that assertion, i.e. the percentage of correct answers was [(2 +249 +131) / 2055]. Note that the medium class was the lowest percentage of accuracy13.04% (249/1909).

Table 3: Classification matrix

Actual Class	Predict Class			**Total (n)**
	High	Medium	Low	
High	**2**	1	1	**3**
Medium	55	**249**	1605	**1909**
Low	5	7	**131**	**143**
Total (n)	**62**	**257**	**1736**	**2055**

The network was tested with a not training data assuming that all nodes are mutual conditional independent. To understand the behavior of the predictable data is important to look at the test set classification: it generally does not have high levels of probability to determine which class the record belongs. This observation shows a relatively indecision in the classification although the probabilistic rate in the wrong result is a little bigger; which can explain why most *Medium* records was incorrectly predicable as *Low*. The reasons for this result may be linked to data quality and lack of correlation between the wells in this field or even the structure of selected network. The naive Bayes assumes that all variables are conditional independent and therefore does not influence one another. We know that this model does not work in practice and therefore a hierarchical approach may be more fruitful.

6. Description of the DENFIS approach

Neuro-fuzzy inference systems consist of a set of rules and an inference method that can be combined with a connectionist model in order to incorporate the advantages of both approaches, achieving better performance, automatic learning and knowledge availability to the expert user. This knowledge can be represented by a set of extracted rules that can be analysed by either an expert or by an algorithm as reported in Kasabov and Song (2002).

The Dynamic Evolving Neural-Fuzzy Inference System (DENFIS) is a special application of Evolving connectionist systems (ECoS), both were introduced by Kasabov (1998) and Kasabov and Song (2002) respectively. ECoS are “evolving systems” in the sense of a system that changes as new data is presented. The advantages of an ECoS Artificial Neural Network (ANN) over a classic Multilayer Perceptron (MLP) ANN, besides having a Fuzzy Inference System (FIS) after the training stage, are: preventing over-training; outliers have less influence on the output of the network; avoiding problems with noisy data; training the network is generally faster; and it is far more resistant to the forgetting problem when new data is presented, since it evolves only with new data, without the need of presenting data that has already been learned. These advantages are discussed more in-depth in Watts (2009) with many examples and comparisons.

There are two types of DENFIS, one type works with online learning and the other works with off-line learning, both types use the Takagi-Sugeno type FIS as previously mentioned. Online learning is a good alternative for systems where data is being continuously generated and presented to the system. Off-line learning is generally faster and tends to have a smaller prediction error as reported in Kasabov and Song (2002). This article uses the off-line learning type as the focus is the planning phase and all the data that will be used is already available.

The basic mechanism that the DENFIS uses to generate the rules is a clustering algorithm called Evolving Clustering Method (ECM). The clusters are used to partition the input space, and then a MLP network with the backpropagation algorithm is used to find the optimal values for the consequent part of the rules. The antecedent part is determined by each cluster center. The rules created in the DENFIS are defined in Kasabov and Song (2002) as:

$$\begin{cases} if\ x_1\ is\ R_{11}\ and\ x_2\ is\ R_{12}\ and\ \cdots\ x_q\ is\ R_{1q},\ then\ y\ is\ f_1(x_1, x_2, \dots, x_q) \\ if\ x_1\ is\ R_{21}\ and\ x_2\ is\ R_{22}\ and\ \cdots\ x_q\ is\ R_{2q},\ then\ y\ is\ f_2(x_1, x_2, \dots, x_q) \\ \vdots \\ if\ x_1\ is\ R_{m1}\ and\ x_2\ is\ R_{m2}\ and\ \cdots\ x_q\ is\ R_{mq},\ then\ y\ is\ f_m(x_1, x_2, \dots, x_q) \end{cases}$$

Being *m* the number of cluster centres and *q* the size of the input vector. $x_1\ is\ R_{11}$ are the $m \times q$ antecedents for the *m* rules. x_j are the antecedent variables and R_{ij} are the fuzzy sets. The consequent parts of the rules, *y* is the consequent variable and functions f_i are employed. The membership functions employed are triangular functions that depend on three parameters: *a*; *b*; and *c*. This function (1) is defined as:

$$\mu(x) = \max(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0) \quad (1)$$

For an input vector $x^k = [x_1^k, x_2^k, \dots, x_q^k]$ the result of the inference y^k is defined by the equation (2) found in Kasabov and Song (2002):

$$y^k = \frac{\sum_{i=1}^{m} = \omega_i f_i(x_1^k, x_2^k, \dots, x_q^k)}{\sum_{i=1}^{m} = \omega_i}$$
$$where\ \omega_i = \prod_{j=1}^{q} \omega_{Rij}\ (x_j^k) \quad (2)$$

To calculate the function f_i MLP networks are used. So the learning is implemented as follows:

a) Use the ECM to cluster the data of the input vector
b) Optimize the clusters
c) Create the antecedent part for each fuzzy rule
d) Estimate the function f for each fuzzy rule by training a MLP network with the corresponding dataset.

This summarizes the learning algorithm of the offline type DENFIS.

The evolving clustering method(ECM) is a complete clustering method and can be used by itself. The DENFIS utilises it to support the inference of the fuzzy rules from the data. This inference is done in two distinct steps, one for the antecedents and another one for the consequent functions.

The ECM works as a one-pass, distance-based algorithm. Every point will have a maximum distance to its cluster center determined by a parameter called *Dthr*, which is a threshold value that has to be given to the ECM. This is the only parameter needed to run the algorithm and will affect directly the number of clusters that the ECM produces and, therefore, the number of rules the DENFIS generates. The distance used by

the ECM is a general Euclidian distance. The clustering process begins with the first example being presented, when this happens the first cluster is created and a cluster center(Cc) and a cluster radius(*Ru*) are associated to it and Ru is initially set to zero.
As new examples are presented to the ECM two things may happen: existing clusters may be updated by having their centers (*Cc*) and radiuses (*Ru*) modified; or new clusters may be created. By definition a cluster is updated until its radius reaches a value equals to *Dthr*, when it reaches this value it will stop being updated.

7. Results of the DENFIS approach

After training the DENFIS using the training dataset, with a *Dthr* value of 0.09, a total of 39 rules were created. The average error obtained for each data sample for the training set was 4.10%. This rulebase was used to predict the results for the test dataset and an average error of 2.89% for each data sample was obtained. As there was a log of a problematic or unusual bore hole drilling operation in the training set, this made the average error of the training set exceed the error of the test set. Since the value that is actually used for the planning phase is the average error, this is a good result. On the other hand, the rulebase created is complicated and an expert wouldn't be able to easily fine tune it. One example of the rules created is shown in (3).

$$\begin{cases} \begin{array}{c} Rule\ 1: if\ RPM\ is\ (0.02\ 0.08\ 0.13) \\ WOB\ is\ (0.10\ 0.16\ 0.22) \\ HSI\ is\ (0.02\ 0.08\ 0.14) \\ UCS\ is\ (0.00\ 0.06\ 0.11),\ then \\ y = 3.10 + 57.83 * RPM + 0.02 * WOB - 225.97 * HSI + 0.29 * UCS \\ \\ Rule\ 2: if\ RPM\ is\ (0.14\ 0.20\ 0.25) \\ WOB\ is\ (0.11\ 0.17\ 0.23) \\ HSI\ is\ (0.03\ 0.09\ 0.15) \\ UCS\ is\ (0.02\ 0.08\ 0.14),\ then \\ y = 0.55 + 0.32 * RPM + 0.05 * WOB - 4.62 * HSI + 0.18 * UCS \\ \\ Rule\ 3: if\ RPM\ is\ (0.49\ 0.55\ 0.61) \\ WOB\ is\ (0.05\ 0.11\ 0.17) \\ HSI\ is\ (0.02\ 0.08\ 0.14) \\ UCS\ is\ (0.01\ 0.07\ 0.13),\ then \\ y = -0.05 + 0.24 * RPM + 0.09 * WOB + 0.12 * HSI - 0.05 * UCS \\ \vdots \end{array} \end{cases} \quad (3)$$

Being all the functions in the antecedents, triangular membership functions expressed in the form *(a b c)* for each input variable.

8. Discussion

This article presented an on-going research that addresses the optimization of the cost of drilling wells in environments of high complexity and risk such as those related to the pre-salt region offshore Brazil. The process of creation and implementation of computational models capable of predicting and optimizing the rate of penetration (ROP) in the pre-salt wells is not trivial process. In this research the following techniques are investigated: a Bayesian inference approach for targeting the elicitation process and subsequent combination of models; and a Dynamic Evolving Neural-Fuzzy Inference System (DENFIS).
The use of a not hierarchical network (naive Bayes) to classify ROP was not enough for a good classification of input values. This kind of behavior shows the complexity of the domain: it is not possible to say that the variables are mutually independent. Other possible reasons for the low quality of classification are directly linked to the quality of the data and the division of classes for each variable that composes the input pattern's nodes. All these factors make this highly complex network architecture and its determination is not a simple task. As a sequel of this approach are the exploration of mention factors and the use of experts' knowledge with Bayesian learning algorithms such as K2 (Neapolitan, 2003), IC (Pearl, 1991), among others, in order to find efficient network topology.
The DENFIS was able to predict well the drilling operation, achieving a low error, but the rules created are extremely complicated to be used by an expert. This is a problem introduced by the type of FIS used (Takagi-Sugeno) and could be solved by using an approach that relies on the Mandami FIS.

References

Al-yami A. S., Schubert J., 2012, Using Bayesian Network to Model Drilling Fluids Practices in Saudi Arabia, in SPE International Production and Operations Conference & Exhibition, 23 pages, 14-16 May 2012, Doha, Qatar, DOI: 10.2118/152096-MS.

Fonseca, T. C., Mendes, J. R. P., Serapiao, A. B. S., Guilherme, I. R., 2006. A Genetic Neuro-Model Reference Adaptive Controller for Petroleum Wells Drilling Operations, in Computational Intelligence for Modelling, Control and Automation, 2006 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on, 3 pages, 28 Nov. - 1 Dec. 2006, Sydney, Australia, DOI: 10.1109/CIMCA.2006.8

Friedman N., Geiger D., Goldszmidt M., 1997, Bayesian network classifiers, Machine Learning - Special issue on learning with probabilistic representations, 29(2), 131-163, DOI: 10.1023/A:1007465528199

Gandelman, R. A. 2012. Predição da ROP e optimização em tempo real de parâmetros operacionais na perfuração de poços de petróleo offshore *(ROP prediction and real-time optimization of operational parameters on drilling offshore oil Wells)*, MS thesis, UFRJ, Brazil. (*Federal University of Rio de Janeiro*). (in portuguese)

Giese M., Bratvold R. B., 2010, Probabilistic Modeling for Decision Support in Integrated Operations, in SPE Intelligent Energy Conference and Exhibition, 23-25 March 2010, Utrecht, The Netherlands, 1, 173-185, DOI: 10.2118/127761-MS.

Kasabov, N. K., 1998, The ECOS Framework and the ECO Learning Method for Evolving Connectionist Systems, Journal of Advanced Computational Intelligence and Intelligent Informatics, 2(6), 195-202.

Kasabov, N. K., Song, Q., 2002, Denfis: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction, IEEE Transactions on Fuzzy Systems, 10(2), 144-154, DOI: 10.1109/91.995117.

McCallum A., Nigam K., 1998, A comparison of event models for naive bayes text classification, in AAAI-98 Workshop on Learning for Text Categorization, 26-27, Jul. 1998, Madison, Wisconsin, 41-48.

Neapolitan R. E., 2003, Learning Bayesian Networks, Pearson Prentice Hall, Upper Saddle River, NJ.

Pearl J., 1988, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, San Mateo, CA.

Pearl J., Verma T. S., 1991, A theory of inferred causation, in Second International Conference on the Principles of Knowledge Representation and Reasoning, 22-25, Apr., Cambridge, MA., 441-452.

Rajaieyamchee M. A., Bratvold R., 2009, Proceedings of SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers, 1, 1-17, DOI: 10.2118/124247-MS.

Watts, M. J. (2009), A decade of Kasabov's evolving connectionist systems: A review, IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, 39(3), 253-269, DOI: 10.1109/TSMCC.2008.2012254.