Journal of Petroleum Science and Engineering 123 (2014) 106–119

Contents lists available at ScienceDirect

Journal of Petroleum Science and Engineering

journal homepage: www.elsevier.com/locate/petrol

Data-driven forecasting of naturally fractured reservoirs based on nonlinear autoregressive neural networks with exogenous input

L. Sheremetov [a,*], A. Cosultchi [a], J. Martínez-Muñoz [a], A. Gonzalez-Sánchez [a], M.A. Jiménez-Aquino [b]

[a] Mexican Petroleum Institute, Eje Central Lázaro Cárdenas 152, San Bartolo Atepehuacan, Distrito Federal, Mexico
[b] PEMEX Exploración y Producción, Mexico

A R T I C L E I N F O

Article history:
Received 31 January 2014
Accepted 12 July 2014
Available online 30 July 2014

Keywords:
time series forecasting
oil production prediction
NARX neural networks
naturally fractured reservoirs

A B S T R A C T

In this paper we discuss the results of the modeling of naturally fractured reservoir based on the application of the nonlinear autoregressive neural network with exogenous inputs (NARX). We show that the NARX network can be efficiently applied to multivariate multi-step ahead prediction of reservoir dynamics. Predictability of the time series is studied using the Hurst exponent. We show that preliminary clustering of the time series can increase the precision of the forecasting. We evaluate the proposed approach using a real-world data set describing the dynamic behavior of a naturally fractured oilfield asset located in the coastal swamps of the Gulf of Mexico. This paper is not only intended for proposing a new model but to study carefully and thoroughly several aspects of the application of ANN models to reservoir modeling and to discuss conclusions that could be of the interest for petroleum engineers.

1. Introduction

The reservoir is described by a set of time series (TS) of fluids from petroleum wells, which are characterized by different starting points and mutual influence. Production performance is both controlled by the reservoir properties and is also affected by operational constraints and surrounding wells performance. The rock and fluid properties of the reservoirs are highly nonlinear and heterogeneous in nature. The situation is even worse for naturally fractured reservoirs (NFR), where natural fractures and faults (created over geologic time) are the primary channels both for hydrocarbon migration and for water breakthrough and gas coning. Thus, production TS comprise high-frequency multipolynomial components, represent a long memory process and are often discontinuous (or piecewise continuous) which make difficult to get the best model for such data.

Several important tasks of petroleum reservoir engineering are concerned with the forecasting of oil production. Usually, production prediction problem is considered within several different settings (He et al., 2001). The first is the prediction of existing wells which is based on that well's previous production data. The other one is the spatial prediction of a new infill drilling well which is based on nearby wells' history production data. Finally, the problem of backward prediction, known as “backcasting”, can also arise for some brown fields with no record of the measured wells' production. In this paper, due to the space limits we focus on the former case.

Traditional methods of forecasting in petroleum engineering include DCA, black oil model history matching, exploration analogies and exploration trend extrapolations (Weiss et al., 2002). These tools are based on subjective data interpretation: to pick the proper slope, to tune the parameters of the numerical simulation model in such a way that they keep the reasonable values to interpret reservoir geology.

TS forecasting, along with clustering and classification, is one of the traditional time series data mining tasks (Batyrshin and Sheremetov, 2008). Traditional prediction techniques based on TS analysis usually establish some requirements that should be fulfilled. For instance, the use of the ARMA method is limited to stationary TS (the ARIMA model assumes that the data become stationary after differencing), that implies that the mean, variance and autocorrelation structure do not change over time (Peña et al., 2001). Such assumptions do not fulfill for the TS describing the behavior of the reservoir.

For the past few decades, artificial neural networks (ANNs), among other artificial intelligence techniques, have been extensively applied in petroleum engineering due to their potential to handle nonlinearities and time-varying situations along with their ability to learn and adapt to new dynamic environments (Sheremetov et al., 2005;

* Corresponding author.
E-mail addresses: sher@imp.mx (L. Sheremetov), acosul@imp.mx (A. Cosultchi), jmmunoz@imp.mx (J. Martínez-Muñoz), agonzal@imp.mx (A. Gonzalez-Sánchez), marco.antonio.jimenez@pemex.com (M.A. Jiménez-Aquino).

http://dx.doi.org/10.1016/j.petrol.2014.07.013
0920-4105/

Mohaghegh, 2005; Bravo et al., 2013). Several ANN topologies have been studied in their application to short-term (1–2 years) univariate and multivariate prediction of oil production. Though, nonlinear autoregressive neural network with exogenous inputs (NARX) have been studied for univariate forecasting of TS (Menezes and Barreto, 2008), their application in multivariable settings in multi-step-ahead forecasting schemes has not been fully explored yet (Diaconescu, 2008).

In this paper we analyze the problem of flow rate forecasting of naturally flowing wells under a limited availability of operational data (irregularities in operational conditions, lack of production well tests, etc.) by using ANN models. Though we try to summarize the aspects that should be studied in order to thoroughly validate the application of the ANN for modeling of the oilfield behavior, in this paper we take as modest goals (i) analysis of the predictability of the production TS; (ii) applicability of the univariate and multivariate forecasting, (iii) analysis of different topologies of the NARX networks, and (iv) the application of clustering techniques to improve forecasting results. We will not attempt to study the feature selection process, the prediction capabilities of the NARX networks to forecast infill wells or thoroughly study long-term forecasting (for the periods up to 10 years). Obviously, to properly estimate the utility of the obtained conclusions we would require many tests.

The rest of the paper is organized as follows. In the next section we explain the motivation for research emphasizing both application-related and model-related aspects of production prediction. The basics of TS forecasting with NARX networks is considered in Section 3. Section 4 resumes the results of the provided experiments with real data from Jujo-Tecominoacán oilfield located in the coastal swamps of the Gulf of Mexico, used to validate the proposed approach. Section 5 provides a review of the related work followed by conclusions.

2. The NARX network in time series forecasting

We start with a short introduction to NARX networks just to make the motivation for this research more clear.

Recurrent neural network (RNN) is a class of neural network where connections between units form a directed cycle. This creates an internal state of the network which allows it to exhibit the dynamic temporal behavior. Unlike feedforward neural networks, RNNs can use their internal memory to process arbitrary sequences of inputs. RNNs cannot be easily trained for large numbers of neuron units nor for large numbers of input units. Successful training has been mostly in time series problems with several inputs. Such kind of architectures is usually trained by means of temporal gradient-based variants of the backpropagation algorithm. However, learning to perform tasks, in which the temporal dependencies present in the input/output signals span long time intervals, can be quite difficult using gradient-based learning algorithms (for networks like Time Delay Neural Network – TDNN). Learning of such long-term temporal dependencies is more effective with Nonlinear Autoregressive with eXogenous input (NARX) architectures because their input vector is built through two delay lines: sliding over the input and output signals (Menezes and Barreto, 2008).

The NARX is a recurrent neural network which has been demonstrated being well suited for modeling nonlinear systems and specially time series. Compared to classical prediction models of time series such as linear parametric autoregressive (AR), moving-average (MA) and autoregressive moving-average (ARMA) models (Box and Jenkins, 1970) recurrent NN (RNN) with a sufficiently large number of neurons is a realization of the nonlinear ARMA process (Haykin, 1999). Compared to feedforward ANN, they have the following advantages: (i) learning is more effective in NARX networks because the gradient descent is better and (ii) because of a feedback, these networks converge much faster and generalize better than other networks (Lin et al., 1996; Gao and Er, 2005). Embedded memory in recurrent NARX also helps reducing the effect of vanishing gradient (when the output of a system at time instant k depends on network inputs presented at times $r < < k$). In our previous paper, we show that NARX ANN outperformed considerably the traditional TDNN network for the problem at hand (Sheremetov et al., 2013). That is why NARX networks are used in this study.

In long-term prediction, the model's output is fed back to the input for a fixed number of time steps. This way, input components, previously composed of actual sample points, are gradually replaced by predicted values[1]. The output of the network is

$$\hat{y}(t+k) = f(u_i(t-1), \ldots, u_i(t-n), y(t-1), \ldots, y(t-n), \hat{y}(t-1), \ldots, \hat{y}(t-n))$$

where $i = 1,\ldots,m$.

The transfer function of the network f is the same as that of a one-output feedforward neural network; for more than one output f should have a subindex i:

$$f(\cdot) = g\left(\sum w_{h,j} h_q(\cdot)\right)$$

where g is the activation function of the output node, hyperbolic tangent, hyperbolic logarithm or linear, depending upon the sum of the activation functions of the nodes in the hidden layer. $w_{h,j}$ is the weight corresponding to the hidden node, h is the number of the hidden node. $h_q(\cdot)$ is the activation function of the hidden node. This activation function $h(\cdot)$is defined as

$$h(\cdot) = net\left(\sum w_{i,h} u_i\right),$$

where *net* is one of the hyperbolic tangent or hyperbolic logarithm functions:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad \mathrm{logh}(x) = \frac{1}{1+e^{-x}}$$

$w_{i,h}$ is the weight of input i that goes to the hidden node h. u_i is the i-th input to the network, whether any of the input variables or a feedback.

Linear activation functions are not used in the hidden layer, since they would not contribute to the nonlinearity of the transfer function of the network. However, the linear function can be used at the output node in those cases where the characteristic function does not contain a very high degree of nonlinearity. In this case the output layer absorbs the function linearly accumulating the contributions of the nodes of the hidden layers.

TS forecasting is usually performed in two different settings: (a) in a univariate setting, when the ANN is trained only with the time series data and (b) in a multivariate setting, when other variables (static and dynamic nature) are added as additional inputs. Those variables that change monthly enter the network just as TS. Variables regarded as static are converted to TS copying the same values month by month and changing them if any variation occurs. The process of the selection of the most appropriate set of input variables called "feature selection" is out of the scope of this paper.

In Fig. 1 a network topology is illustrated for the latter case. For each variable v_i there is a number of n delays and the output forecast k months ahead from the current month. First, based in former studies on forecasting oil production, it was decided to use a two-layer feedforward network (Schrader et al., 2005; Menezes

[1] If the prediction horizon tends to infinity, from some time in the future the input regressor is composed only of estimated values and the multi-step-ahead prediction task becomes a dynamic modeling task while the ANN model emulates the dynamic behavior of the system.