

Research Article

A Hybrid Model through the Fusion of Type-2 Fuzzy Logic Systems and Sensitivity-Based Linear Learning Method for Modeling PVT Properties of Crude Oil Systems

Ali Selamat,¹ Sunday Olusanya Olatunji,¹ and Abdul Azeez Abdul Raheem²

¹ Intelligent Software Engineering Laboratory, Faculty of Computer Science and Information Systems, University of Technology Malaysia, 81310 Skudai, Johor Bahru, Malaysia

² Centre for Petroleum and Minerals, The Research Institute, King Fahd University of Petroleum and Minerals (KFUPM), P.O. Box 1105, Dhahran 31261, Saudi Arabia

Correspondence should be addressed to Ali Selamat, aselamat@utm.my

Received 29 March 2012; Revised 8 September 2012; Accepted 10 September 2012

Academic Editor: Oscar Montiel Ross

Copyright © 2012 Ali Selamat et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sensitivity-based linear learning method (SBLLM) has recently been used as a predictive tool due to its unique characteristics and performance, particularly its high stability and consistency during predictions. However, the generalisation capability of SBLLM is sometimes limited depending on the nature of the dataset, particularly on whether uncertainty is present in the dataset or not. Since it made use of sensitivity analysis in relation to the data sets used, it is surely very prone to being affected by the nature of the dataset. In order to reduce the effects of uncertainties in SBLLM prediction and improve its generalisation ability, this paper proposes a hybrid system through the unique combination of type-2 fuzzy logic systems (type-2 FLSs) and SBLLM; thereafter the hybrid system was used to model PVT properties of crude oil systems. Type-2 FLS has been chosen in order to better handle uncertainties existing in datasets beyond the capability of type-1 fuzzy logic systems. In the proposed hybrid, the type-2 FLS is used to handle uncertainties in reservoir data so that the cleaned data from type-2 FLS is then passed to the SBLLM for training and then final prediction using testing dataset follows. Comparative studies have been carried out to compare the performance of the newly proposed T2-SBLLM hybrid system with each of the constituent type-2 FLS and SBLLM. Empirical results from simulation show that the proposed T2-SBLLM hybrid system has greatly improved upon the performance of SBLLM, while also maintaining a better performance above that of the type-2 FLS.

1. Introduction

Hybrid computational intelligence is any effective combination of intelligent techniques that performs superior or in a competitive way to simple standard intelligent techniques. The increased popularity of hybrid intelligent systems in recent times lies in the extensive success of these systems in many real-world complex problems [1]. Also, it is an established fact that every approach has its strengths and weaknesses; hence the need for hybrid models that is able to combine the strengths of the individual techniques while complementing the weaknesses of one method with the strength of the other. Therefore, this work seeks to take advantage of the unique capability of type-2 FLS, in

modelling uncertainties, to improve the performance of the sensitivity based linear learning method (SBLLM) in order to further boost the generalisation ability of SBLLM even in the face of uncertainties.

Type-2 fuzzy logic has been generally acknowledged as being better and ideal for uncertainty modeling [2–8]. Recently, type-2 FLS have been proposed as a novel framework for both classification and prediction in order to handle all forms of uncertainties [8, 9]. It is able to handle uncertainties that include those in measurements and data used to calibrate the parameters. It has been used in several fields and the results have been promising and very encouraging [10–12]. Therefore, there is a possibility that the type-2 FLS can handle uncertainty in reservoir data [13] as

the type-2 fuzzy logic has been specifically invented to deal with all forms of uncertainties [8] that are inherent in our day to day natural encounters and mode of reasoning.

Sensitivity based linear learning method (SBLLM) has been recently introduced as a learning technique for two-layer feedforward neural networks based on sensitivity analysis that uses a linear training algorithm for each of the two layers [14]. It was introduced in order to alleviate some of the limitations of the classical ANN. This algorithm tends to provide good generalization performance at extremely fast learning speed, while in addition, it gives the sensitivities of the sum of squared errors with respect to the input and output data without extra computational cost. It is very stable in performance as its learning curve stabilizes soon and behaves homogeneously not only if we consider just the end of the learning process, but also during the whole process, in such a way that very similar learning curves were obtained for all iterations of different experiments [14, 15]. Unfortunately, just like the classical ANN, SBLLM is unable to adequately model uncertainties in real life data. As a result of SBLLM inability to handle uncertainty, it will be a good contribution to seek to improve its performance through the use of type-2 FLS as its precursor, in a hybrid arrangement, for uncertainty handling, in order to take advantage of its unique capabilities. Therefore in this paper, we propose a hybrid approach that will combine the unique attributes of type-FLS with those of sensitivity based linear learning method (SBLLM) by way of improving SBLLM performance in order to achieve better generalisation ability in all situations including uncertainty oriented environment.

Characterization of reservoir fluids plays a very crucial role in developing a strategy on how to produce and operate a reservoir. Pressure-volume-temperature (PVT) properties are very crucial for geophysics and petroleum engineers, namely for the utilization in material balance calculations, inflow performance calculations, well log analysis, determining oil reserve estimations, and the amount of oil that can be recovered, the flow rate of oil or gas, and the simulation on reservoir outputs. The phase and volumetric behaviour of petroleum reservoir fluids is referred to as PVT [16, 17].

PVT properties include formation volume factor (FVF), solution gas-oil ratio (GOR), solution oil-gas ratio (OGR), liquid specific gravity, American petroleum institute (API) specific gravity, gas specific gravity, bubble-point pressure, saturation pressure, and so forth stated by [18]. Among those PVT properties, the bubble-point pressure (P_b) and the oil formation factor (B_{ob}) are the most important, because they are the most essential factors in reservoir and production computations [18]. The more the preciseness of estimating these properties, the better the calculations involved in reservoir simulation, production, and field development. Bubble-point pressure (P_b) is the pressure at which gas first begins to come out of the solution at constant temperature, while oil formation volume factor (B_{ob}) is defined as the volume of reservoir oil that would be occupied by one stock tank barrel oil plus any dissolved gas at the bubble point pressure and reservoir temperature as stated in [17, 19–21].

Ideally, these properties are determined from laboratory studies on samples collected from the bottom of the wellbore

or on the surface. However, such experimental data is very costly to obtain and the accuracy of the result is critical and not often known in advance. One of the solutions is to use the empirically derived correlations, which has been developed using equation of state (EOS), linear/non-linear statistical regression, or graphical techniques [17, 21]. Unfortunately, these correlations are constrained by several limitations, namely, the equation of state requires the extensive knowledge of the detailed compositions of the reservoir fluids and the determination of such quantities is expensive and time consuming; the accuracy of the EOS depends heavily on the nature of the fluid, on the type of equation selected, and on the operator-dependent tuning procedures. This method also involves several numerical computations. To overcome the shortcomings associated with the earlier correlation methods, researchers made use of artificial intelligence based methods foremost of which is the classical artificial neural network (ANN) and its variants. But still, the developed neural networks correlations often do not perform to expectations and are bedeviled with shortcomings that include, among others, instability, with its non homogeneous nature such that very different learning curves were obtained for different repeat of same experiment, and its characteristic low speed operation.

Researchers have done their best to address and overcome these problems of ANN. As a result, several variants of ANN and other methods like support vector machines (SVM) and functional networks (FN) have been proposed and used [21, 22], yet each has its limitations that still call for further research of this nature, particularly their inability to handle uncertainties and the need to ensure stability and consistency in predictions.

It is an established fact that geosciences disciplines are not clear-cut and, most of the time, are associated with uncertainties [13], hence the need for fuzzy logic based systems, particularly the newly introduced type-2 fuzzy logic system (type-2 FLS) that is able to adequately account for all forms of uncertainties [8]. For instance, prediction of core parameters from well log responses is difficult and is usually associated with uncertainties. Earlier methods try to minimize and ignore these uncertainties [13], while type-2 fuzzy logic derives useful information from the uncertainties and uses it as a good selection of parameters for increasing the accuracy of the predictions. SBLLM on its own is able to deploy its sensitivity analysis to ensure stable and consistent results always. Thus a combination of these unique methods in a hybrid arrangement will go a long way in improving the prediction accuracy while ensuring stability and consistency, which are requisite of good prediction system. This will definitely be in line with the often reported successes of hybrid systems as a result of the unique combination of methods that takes advantage of each constituent member while avoiding the shortcomings of each.

Therefore, this paper investigate the feasibility of using type-2 FLS as a pre-cursor to improve the generalisation ability of SBLLM in the face of uncertainty during prediction, in a hybrid framework setting; we develop a new hybrid model based on type-2 FLS and SBLLM and then use it for predicting PVT properties that include specifically bubble

point pressure (P_b) and oil formation volume factor (B_{ob}) using different standard databases of four input parameters, namely, solution gas-oil ratio, reservoir temperature, oil gravity, and gas relative density. We then investigate how individual constituent methods compare in their forecasting performance with the proposed hybrid model. Empirical results from simulations demonstrated that the proposed hybrid scheme produced better generalization performance, with high stability and consistency, which are requisite of good prediction models, compared to each of the constituent parts, particularly SBLLM.

The rest of this paper is organized as follows. Section 2 presents a review of related researches and Section 3 presents the proposed hybrid model and its constituent parts. Section 4 contains the empirical study and implementation process. Results and discussions are presented in section 5. The conclusion and future work recommendations are provided in Section 6.

2. Related Research

In the past few decades, engineers realized the importance of developing and using empirical correlations for PVT properties. The development of correlations for PVT calculations has been the subject of extensive research, resulting in a large volume of publications. In this section, we briefly review the most common empirical PVT correlations and the related prediction approach together with the measurement techniques that have been used in forecasting these PVT properties.

2.1. Common Empirical Models and Evaluation Studies. Standing [23] presented correlations for bubble point pressure and for oil formation volume factor. The correlations were based on laboratory experiments carried out on 105 samples from 22 different crude oils in California. Glaso [24] developed the Glasoempirical correlation for formation volume factor using 45 oil samples from North Sea hydrocarbon mixtures. Al-Marhoun [25] published his second correlation for oil formation volume factor. The correlation was developed with 11,728 experimentally obtained formation volume factors at, above, and below bubble point pressure. The data set represented samples from more than 700 reservoirs from all over the world, mostly from Middle East and North America. For more empirical correlation-related work, discussion, applications, and comparative studies; interested readers can see [26–42].

2.2. Predicting PVT Properties Based on Artificial Neural Networks. Artificial Intelligence schemes have been increasingly used in the field of PVT properties and other fields in oil and gas industry during the last few decades, the most popular of which are the neural networks. Artificial neural networks are parallel-distributed information processing models that can recognize highly complex patterns within available data. In recent years, neural network have gained popularity in petroleum applications. Many authors have discussed the applications of neural networks in petroleum engineering;

see [16, 19, 43–47] for details. Recently, it has been shown in both machine learning and data mining communities that artificial neural networks have the capacity to learn complex linear/nonlinear relationships amongst input and output data. The most common widely used neural network in literature is known as the feed forward neural networks with back propagation training algorithm [48]. This type of neural networks is an excellent computational intelligence modelling scheme in both prediction and classification tasks. Recently, feed forward neural networks were used to predict the PVT correlations, [16, 49–51].

The authors in [52] introduced a novel approach for predicting the complete PVT behavior of reservoir oils and gas condensates using a noniterative approach. The method uses key measurements that can be performed rapidly either in the lab or at the well site as input to a neural network, while in [51], two neural networks were trained separately to estimate the bubble point pressure (P_b) and oil formation volume factor (B_{ob}), respectively. The input data were solution gas-oil ratio, reservoir temperature, oil gravity, and gas relative density, while making use of two hidden layers (2HL) neural networks: the first neural network, (4-8-4-2) to predict the bubble point pressure and the second neural network, (4-6-6-2) to predict the oil formation volume factor. Both neural networks were built using a data set of size 520 observations from Middle East area. The input data set was divided into a training set of 498 observations and a testing set of 22 observations.

The authors in [19] used the feedforward learning scheme with log sigmoid transfer function in order to estimate the formation volume factor at the bubble point pressure, using data published in [50], while the authors in [53] developed two new models to predict the bubble point pressure, and the oil formation volume factor at the bubble-point pressure for Saudi crude oils. The models were based on artificial neural networks, and developed using 283 unpublished data sets collected from different Saudi fields. Recently, [46] made use of neural network to predict the PVT properties. As usual, they were confronted with the generic problems of the standard neural network like local minima convergence problem, trial and error syndrome, instability, and inconsistency.

For further works on the utilization of neural networks and its variants, such as radial basis function network, support vector machines, and abductive network, for predicting PVT properties, interested readers can refer to [16, 17, 21, 47, 49, 54–57].

We have noted that most of the reported cases of the use of fuzzy logic in modeling reservoir properties are restricted to the classical fuzzy logic (also known as type-1 fuzzy logic). However, type-1 fuzzy logic systems have recently been found inadequate for handling all forms of uncertainties [5, 8, 9, 58]. In response to this, type-2 fuzzy logic systems have been introduced as better computational intelligence approach for both prediction and classification to handle all forms of uncertainties [5]. The unique feature and advantage of type-2 fuzzy logic systems and those of sensitivity based linear learning methods (SBLLM) motivated the idea of this work.

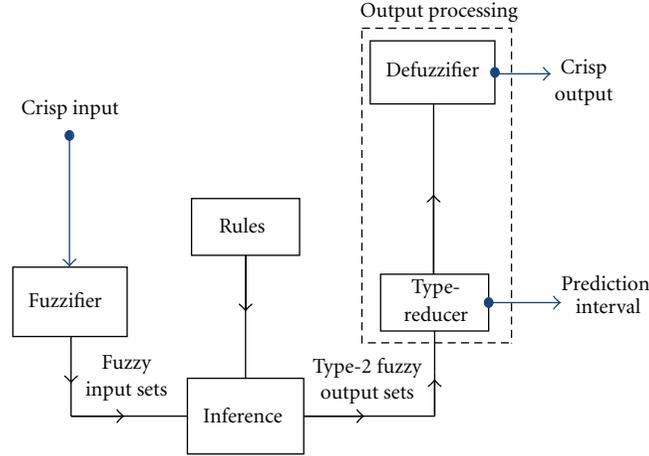


FIGURE 1: Schematic diagram of type-2 fuzzy logic systems (type-2 FLS).

In order to further boost the accuracy of predictions, particularly in this germane field of reservoir characterization where the need for accurate prediction is highly desirable, we have proposed a better and more reliable hybrid scheme that will be able to adequately model uncertainty in reservoir data and make accurate predictions while ensuring stable and consistent results.

In this regard, this work seeks to develop a new hybrid scheme based on type-2 fuzzy logic system and sensitivity based linear learning methods (SBLLM). These combinations have been chosen due to the fact that: type-2 fuzzy logic system is able to model all forms of uncertainties and sensitivity based linear learning methods (SBLLM) have a unique generalization ability coupled with higher stability and consistency. With these, the proposed hybrid model will be able to make use of these unique combinations for effective uncertainties handling while ensuring robust, consistent, and accurate performance.

3. The Proposed Hybrid Model and Its Constituent Frameworks

The proposed hybrid system is composed of type-2 fuzzy logic system and sensitivity based linear learning methods (SBLLM) learning schemes, uniquely combined together to form a better performing hybrid scheme.

3.1. Type-2 Fuzzy Logic System (Type-2 FLS). Type-2 adaptive fuzzy inference systems is an adaptive network that learns the membership functions and fuzzy rules, from data, in a fuzzy system based on type 2 fuzzy sets; see [8, 59] for details. “Type-2 fuzzy sets are fuzzy sets whose grades of membership are themselves fuzzy. They are intuitively appealing because grades of membership can never be obtained precisely in practical situations” [60]. Type-2 fuzzy sets can be used in situations where there is uncertainty about the membership grades themselves, for example, an uncertainty in the shape of the membership function or in some of its parameters. Consider the transition from ordinary sets to fuzzy sets; when

we cannot determine the membership of an element in a set as 0 or 1 we use fuzzy sets of type-1. Similarly, when the situation is very fuzzy that we have difficulty in determining the membership grade as a crisp number in $[0, 1]$, we use fuzzy sets of type-2. Thus, in general, “a fuzzy set is of type n , $n = 2, 3, \dots$ if its membership function ranges over fuzzy sets of type $n - 1$ ” [61].

Generally, a type-2 fuzzy logic system contains five components: fuzzifier, rules, inference engine, type-reducer, and defuzzifier that are interconnected as in Figure 1.

The fuzzifier takes the input parameters values as inputs. The output of the fuzzifier is the fuzzified measurements which will be the input to the inference engine. The resultant of the inference engine is type-2 fuzzy output sets which can be reduced to type-1 fuzzy set by the type reducer. The type reduced fuzzy set in this model is an interval set which gives the predicted external attribute measurement as a possible range of values. The defuzzifier calculates the average of this interval set to produce the predicted crisp external attribute measurement.

3.1.1. Inferencing in the Type-2 FLS. Fuzzy inference engine combines the fired fuzzy rules and map inputs into type-2 output fuzzy sets. Generally a type-2 FLS is a fuzzy logic system in which at least one of the fuzzy sets used in the antecedent and/or consequent parts and each rule inference output is a type-2 fuzzy set. Consider a type-2 Mamdani FLS [8] having n inputs $x_1 \in X_1, \dots, x_n \in X_n$ and one output $y \in Y$. The rule base contains L type-2 fuzzy rules expressed in the following form:

$$R^l : \text{IF } x_1 \text{ is } F_1^l \text{ and } x_2 \text{ is } F_2^l \text{ and } \dots \text{ and } x_n \text{ is } F_p^l \text{ THEN } y \text{ is } G^l, \quad (1)$$

where $l = 1, 2, \dots, L$, F_i^l and G^l are type-2 fuzzy set.

This rule represents a type-2 fuzzy relation between the input space $X \in X_1 \times X_2 \times \dots \times X_n$, and the output space

- (1) Arrange y_z in ascending order with $y_1 \leq y_2 \leq \dots \leq y_z$
- (2) Set $\theta_z = h_z$ for $z = 1, 2, \dots, Z$; and compute $y' = \sum_{z=1}^Z y_z \theta_z / \sum_{z=1}^Z \theta_z$
- (3) Find $e \in [1, Z - 1]$ such that $y_e \leq y' \leq y_{e+1}$
- (4) Set $\theta_z = h_z - \Delta_z$ for $z \leq e$ and $\theta_z = h_z + \Delta_z$ for $z > e$, and compute: $y'' = \sum_{z=1}^Z y_z \theta_z / \sum_{z=1}^Z \theta_z$
- (5) Stop if $y'' = y'$, otherwise set $y' = y''$ and return to step 3

ALGORITHM 1: Karnik-Mendel algorithm to compute y'_t .

Y of the system. We denote the membership function of this Type-2 relation as:

$$\mu_{F_1 \times \dots \times F_n} \longrightarrow G^l(x, y), \quad (2)$$

where $F_1^l \times \dots \times F_n^l$ denotes the Cartesian product of $F_1^l, F_2^l, \dots, F_n^l$ and $x = \{x_1, x_2, \dots, x_n\}$.

The antecedents in the fuzzy rules are connected by using the meet operation, the firing strength of the input fuzzy sets are combined with output fuzzy sets using the extended sup-star composition, and the multiple rules are combined using the join operation [8]. However, the computing load involved in deriving the system output from a general type-2 FLS model is high in practice, and the general practice is to use the interval type-2 FLS in which the fuzzy sets F_i^l and G^l are interval fuzzy sets through which the computing of type-2 FLS can be greatly simplified. The membership grades of interval fuzzy sets can be fully characterized by their lower and upper membership grades of the footprint of uncertainty (FOU) separately [8].

Without the loss of generality, let $\mu_{F_i^l}(x) = [\underline{\mu}_{F_i^l}(x), \bar{\mu}_{F_i^l}(x)]$ and $\mu_{G^l}(y) = [\underline{\mu}_{G^l}(y), \bar{\mu}_{G^l}(y)]$ for each sample (x, y) . The firing strength of interval type-2 FLS $\mu_{F^l}(x) = \bigcap_{i=1}^n \mu_{F_i^l}(x)$ is an interval [8], that is, $\mu_{F^l}(X) = [\underline{f}^i(X), \bar{f}^i(X)]$. In the proposed interval type-2 FLS, the meet operation under product t-norm is used, so that the firing strength is an interval type-1 set [8] as shown below:

$$\underline{f}^i(X) = \left[\underline{f}^i(X), \bar{f}^i(X) \right] = \left[\underline{f}^i, \bar{f}^i \right], \quad (3)$$

where $\underline{f}^i(X)$ and $\bar{f}^i(X)$ can be rewritten as follows with $*$ representing the t-norm product operation:

$$\begin{aligned} \underline{f}^i(X) &= \underline{\mu}_{F_1^l}(x_1) * \dots * \underline{\mu}_{F_p^l}(x_n), \\ \bar{f}^i(X) &= \bar{\mu}_{F_1^l}(x_1) * \dots * \bar{\mu}_{F_p^l}(x_n). \end{aligned} \quad (4)$$

3.1.2. Type Reduction. The results from the inference engine are type-2 fuzzy sets. There is then the need to reduce the type-2 fuzzy sets to type-1 fuzzy sets in order to give room for defuzzification so that the final crisp outputs can be generated. Centre-of-sets (COS) type-reducer algorithm developed by Mendel [8] and Karnik et al. [62] has been used in this study because it provides reasonable computational complexity compared to others like the expensive centroid type reducer, though other types can still be investigated

when the need arises as the research progresses. COS type reducer uses two steps in reducing the type-2 fuzzy sets as follows: (i) calculating the centroids of type-2 fuzzy rule consequences and (ii) calculating the reduced fuzzy sets. These stages are described in the following two subsections.

Computing the Centroids of Type-2 Fuzzy Rule Consequences. Suppose that the output of an interval type-2 FLS is represented by type-2 fuzzy sets G^t , where $t = 1, \dots, T$. T is the number of output fuzzy sets. In this first stage, the centroids of all the T output fuzzy sets are calculated and they will be used in calculating the reduced sets in the next stage. The centroid of the i th output fuzzy set y^t is a type-1 interval set which can be expressed in the following equation [8, 62]:

$$y^t = \left[y_l^t, y_r^t \right] = \int_{\theta_1 \in J_{y_l}} \int_{\theta_z \in J_{y_z}} \frac{1}{\sum_{z=1}^Z y_z \theta_z / \sum_{z=1}^Z \theta_z}, \quad (5)$$

where y_l^t and y_r^t are the leftmost and rightmost point of y^t , respectively.

Algorithm 1 is the iterative procedures developed by Mendel [8] Karnik et al. [62] to calculate the rightmost point y_r^t for each type-2 output fuzzy set, where Z is representing the number of discretised points for each output fuzzy set, $J_{y_z} = [L_z, R_z]$, $h_z = [h_z + R_z]/2$, $\Delta_z = [R_z - L_z]/2$, $z = 1, \dots, Z$. Figure 2 demonstrates how to calculate h_z , L_z , R_z , and Δ_z needed by the Algorithm 1. The leftmost point y_l^t can be calculated in the similar way except at step 4 of Algorithm 1 where we set $\theta_z = h_z + \Delta_z$ when $z \leq e$ and $\theta_z = h_z - \Delta_z$ when $z > e + 1$. This iterative procedure has been proven to converge in at most Z iterations to find y_l^t or y_r^t [8].

Computing the Reduced Fuzzy Sets. To calculate the type-reduced set, it is sufficient to compute its upper and lower bounds of the reduced set y_l and y_r , which can be expressed as follows:

$$y_l = \frac{\sum_{i=1}^M \underline{f}_i^i y_l^i}{\sum_{i=1}^M \underline{f}_i^i}, \quad y_r = \frac{\sum_{i=1}^M \bar{f}_i^i y_r^i}{\sum_{i=1}^M \bar{f}_i^i}, \quad (6)$$

where f_i^i and y_l^i are the firing strength and the centroid of the output fuzzy set of i th rule ($i = 1, \dots, M$) associated with y_l , respectively. Similarly, \bar{f}_i^i and y_r^i are the firing strength and the centroid of the output fuzzy set of i th rule ($i = 1, \dots, M$) associated with y_r , respectively.

Meanwhile, y_r can be calculated using the iterative Algorithm 2 as proposed in [5, 8]. Similarly, y_l can be calculated in the same way by setting $\bar{f}_i^i = \underline{f}_i^i$ for $i \leq R$

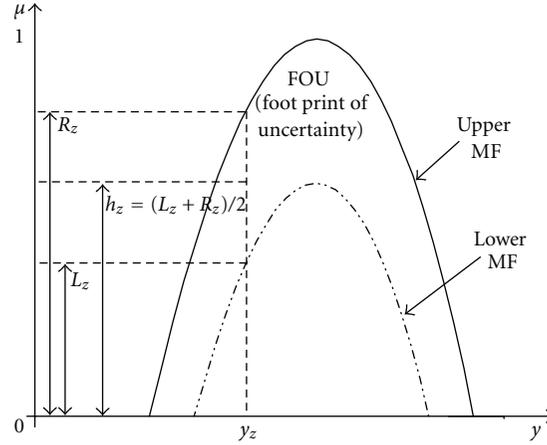


FIGURE 2: Diagrammatic representation of the computation of the parameters needed by each y_z in Algorithm 1 for computing y_l^i and y_r^i .

- (1) Arrange the precalculated y_r^i from Figure 2 in an ascending order, that is, $f_r^1 \leq f_r^2 \leq \dots \leq f_r^M$
- (2) Set $f_r^i = (f_r^i + \bar{f}_r^i)/2$ for $i = 1, 2, \dots, M$, and calculate y_r using (6)
- (3) Find $R \in [1, M - 1]$ such that $y_r^R \leq y_r' \leq f_r^{R+1}$
- (4) Set $f_r^i = \bar{f}_r^i$ for $i \leq R$ and $f_r^i = f_r^i + \bar{f}_r^i$ for $i > R$; and compute y_r' using (6)
- (5) Stop if $y_r = y_r'$, otherwise set $y_r' = y_r'$ and return to step 3

ALGORITHM 2: Iterative algorithm to compute y_r .

and $f_r^i = \bar{f}_r^i$ for $i > R$. The iterative procedure has also been proved to converge in no more than M iterations when computing either y_l or y_r [5].

3.1.3. Defuzzification. We defuzzify the type-reduced set to get a crisp output from the type-2 FLS. The final output of type-2 FLS is thus set to the average of y_l and y_r as shown below:

$$y(x) = \frac{y_l + y_r}{2}, \quad (7)$$

where $y(x)$ is the final crisp output.

3.1.4. The Steepest Descent Approach for Training FLS. The purpose of the training algorithm is to minimize the error function for E training epochs:

$$E(t) = \left(\frac{f(x^{(t)}) - y^{(t)}}{y^{(t)}} \right)^2, \quad t = 1, \dots, N. \quad (8)$$

Consider an FLS with Gaussian membership functions, centre of sums type-reducer, average defuzzification, max-product composition, and product implication; it could be expressed by the equation

$$y(x^{(i)}) = f_s(x^{(i)}) = \frac{\sum_{l=1}^M \bar{y}^l \prod_{k=1}^p \exp \left[-\frac{(x_k^{(i)} - m_{F_k^l})^2}{2\sigma_{F_k^l}^2} \right]}{\sum_{l=1}^M \prod_{k=1}^p \exp \left[-\frac{(x_k^{(i)} - m_{F_k^l})^2}{2\sigma_{F_k^l}^2} \right]}, \quad (9)$$

$$i = 1, \dots, N,$$

where: M is number of rules, p is number of antecedents, and N is number of data points; $m_{F_k^l}$ and $\sigma_{F_k^l}$ are the mean and standard deviation for the membership function, respectively.

Given an input-output training pair $(x^{(i)} : y^{(i)})$ also known as data point, we wish to design a fuzzy logic system (FLS) so that the error function is minimized. The steepest descent approach [8] can be applied to obtain the following recursions to update all the design parameters of this FLS in order to minimize the error function:

$$m_{F_k^l}(i+1) = m_{F_k^l}(i) - \alpha_m [f_s(x^{(i)}) - y^{(i)}] [\bar{y}^l(i) - f_s(x^{(i)})] \times \frac{[x_k^{(i)} - m_{F_k^l}(i)]}{\sigma_{F_k^l}^2(i)} \varphi_l(x^{(i)}), \quad (10)$$

(1) Initialize the parameters of all the membership functions for all the rules, $m_{F_k^l}(0)$, $\bar{y}^l(0)$, and $\sigma_{F_k^l}(0)$.

(2) Set an end criterion to achieve convergence.

(3) **Repeat**

(i) **for all** data points $(x^{(i)} : y^{(i)})$ only $i = 1, \dots, N$

(a) Propagate the next data point through the FLS.

(b) Compute error.

(c) Update the parameters of the membership functions using (10), (11), and (12).

(ii) **end for** (*end for each input-output pair*)

(iii) Compute the root mean square relative error (RMSRE) as in (13).

(iv) Test the end criterion. If satisfied break.

Until (*end for each epoch*)

ALGORITHM 3: Back propagation algorithm for FLS.

$$\bar{y}^l(i+1) = \bar{y}^l(i) - \alpha_{\bar{y}} [f_s(x^{(i)}) - y^{(i)}] \varphi_l(x^{(i)}), \quad (11)$$

$$\begin{aligned} \sigma_{F_k^l}(i+1) &= \sigma_{F_k^l}(i) - \alpha_{\sigma} [f_s(x^{(i)}) - y^{(i)}] [\bar{y}^l(i) - f_s(x^{(i)})] \\ &\quad \times \frac{[x_k^{(i)} - m_{F_k^l}(i)]^2}{\sigma_{F_k^l}^3(i)} \varphi_l(x^{(i)}). \end{aligned} \quad (12)$$

Now, the back propagation algorithm can be applied as in Algorithm 3 followed by RMSRE in (13):

$$\text{RMSRE} = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{[f_s(x^{(i)}) - y^{(i)}]^2}{y^{(i)}}}. \quad (13)$$

3.2. Sensitivity-Based Linear Learning Method (SBLLM). In [14], the authors proposed a new learning scheme in order to both speed up and avoid local minima convergence of the existing back propagation learning technique, while alleviating its other common weaknesses such as instability and inconsistency. This new learning strategy is called the sensitivity based linear learning method (SBLLM) scheme. It is a learning technique for two-layer feedforward neural networks based on sensitivity analysis, which uses a linear training algorithm for each of the two layers. First, random values are assigned to the outputs of the first layer; later, these initial values are updated based on sensitivity formulas, which use the weights in each of the layers and the process is repeated until convergence. Since these weights are learnt solving a linear system of equations, there is an important saving in computational time. The method also gives the local sensitivities of the least squared errors with respect to input and output data, with no extra computational cost, because the necessary information becomes available without extra calculations. This new scheme can also be used to provide an initial set of weights, which significantly improves the behaviour of other learning algorithms. The full theoretical basis for SBLLM and its performance has been demonstrated in [14], which contained its application to several learning problems examples in which it is compared with several learning algorithms and well known data sets.

The results have shown a learning speed generally faster than other existing methods.

Sensitivity analysis is a very useful technique for deriving how and how much the solution to a given problem depends on data; see [15, 63, 64] and the references therein for more details. However, in [14] it was shown that sensitivity formulas can also be used as a novel supervised learning algorithm for two-layer feedforward neural networks that presents a high convergence speed.

Generally, SBLLM process is based on the use of the sensitivities of each layer's parameters with respect to its inputs and outputs and also on the use of independent systems of linear equations for each layer to obtain the optimal values of its parameters. In addition, it gives the sensitivities of the sum of squared errors with respect to the input and output data.

3.2.1. The Learning Process for the Sensitivity Based Linear Learning Method. Consider the two-layer feedforward neural network in Figure 2, where I is the number of inputs, J is the number of outputs, K is the number of hidden units, $x_{0s} = 1$, $z_{0s} = 1$, S is the number of data samples, and the superscripts (1) and (2) are used to refer to the first and second layer, respectively.

This network can be considered to be composed of two one-layer neural networks as is shown in Figure 3.

According to [14], considering the one-layer network in Figure 4, the set of equations relating inputs and outputs is given by

$$y_{js} = f_j \left(\sum_{i=0}^I w_{ji} x_{is} \right); \quad j = 1, 2, \dots, J; \quad s = 1, 2, \dots, S. \quad (14)$$

where I is the number of inputs, J is the number of outputs, $x_{0s} = 1$, w_{ji} are the weights associated with neuron j , and S is the number of data points.

To learn the weight w_{ji} , the following sum of squared errors between the real and the desired output of the networks is usually minimized as:

$$P = \sum_{s=1}^S \sum_{j=1}^J \delta_{js}^2 = \sum_{s=1}^S \sum_{j=1}^J \left(y_{js} - f_j \left(\sum_{i=0}^I w_{ji} x_{is} \right) \right)^2. \quad (15)$$

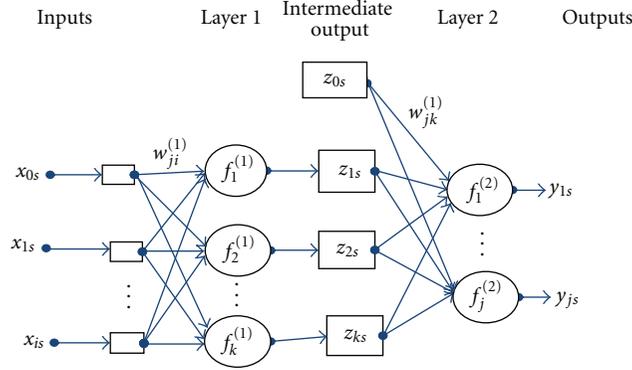


FIGURE 3: Two-layer feedforward neural network of the SBLLM.

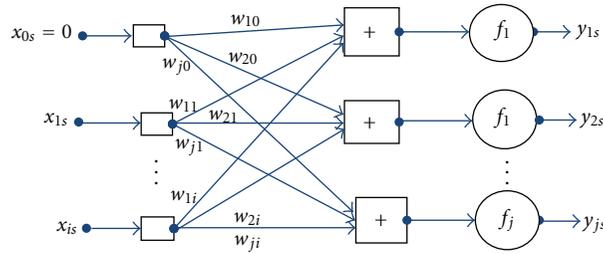


FIGURE 4: One-layer feedforward neural network combined to form the two-layer feedforward neural network of the SBLLM Scheme in Figure 2.

Assuming that the nonlinear activation functions, f_j , are invertible (which is the case for most of the commonly employed functions), otherwise, one can minimize the sum of squared errors before the nonlinear activation functions [14], that is,

$$Q = \sum_{s=1}^S \sum_{j=1}^J \varepsilon_{js}^2 = \sum_{s=1}^S \sum_{j=1}^J \left(\sum_{i=0}^I w_{ji} x_{is} - f_j^{-1}(y_{js}) \right)^2, \quad (16)$$

which leads to the system of equations:

$$\begin{aligned} \frac{\partial Q}{\partial w_{jp}} &= 2 \sum_{s=1}^S \left(\sum_{i=0}^I w_{ji} x_{is} - f_j^{-1}(y_{js}) \right) x_{ps} \\ &= 0; \quad p = 0, 1, \dots, I; \quad \forall j, \end{aligned} \quad (17)$$

that is,

$$\sum_{i=1}^I w_{ji} \sum_{s=1}^S x_{is} x_{ps} = \sum_{s=1}^S f_j^{-1}(y_{js}) x_{ps}; \quad p = 0, 1, \dots, I; \quad \forall j, \quad (18)$$

or

$$\sum_{i=0}^I A_{pi} w_{ji} = b_{pj}; \quad p = 0, 1, \dots, I; \quad \forall j, \quad (19)$$

where

$$A_{pi} = \sum_{s=1}^S x_{is} x_{ps}; \quad p = 0, 1, \dots, I; \quad \forall i, \quad (20)$$

$$b_{pj} = \sum_{s=1}^S f_j^{-1}(y_{js}) x_{ps}; \quad p = 0, 1, \dots, I; \quad \forall j.$$

Furthermore, for the one-layer neural network shown in Figure 4, according to [14, 15, 64], the sensitivities of the new cost function, Q , with respect to the input and output data can be obtained as:

$$\frac{\partial Q}{\partial y_{pq}} = - \frac{2 \left(\sum_{i=0}^I w_{pi} x_{iq} - f_p^{-1}(y_{pq}) \right)}{f_p'(y_{pq})}; \quad \forall p, q, \quad (21)$$

$$\frac{\partial Q}{\partial x_{pq}} = 2 \sum_{j=1}^J \left(\sum_{i=0}^I w_{ji} x_{iq} - f_j^{-1}(y_{jq}) \right) w_{jp}; \quad \forall p, q.$$

Based on [14], the learning method and the sensitivity formulas presented above can now be used to develop the SBLLM learning method for the two-layer feedforward networks shown in Figure 3. Now, consider the two-layer feedforward neural network in Figure 3, where I is the number of inputs, and include the four PVT properties independent variables namely solution gas-oil ratio, reservoir temperature, oil gravity, and gas relative density; J is the number of outputs, and it include the two target PVT properties (dependent variables) namely bubble point

pressure (P_b) and oil formation volume factor (B_{ob}); K is the number of hidden units, $x_{0s} = 1$, $z_{0s} = 1$; S is the number of data samples and the superscripts (1) and (2) are used to refer to the first and second layer, respectively.

Assuming that the intermediate layer outputs z are known, using (16), a new cost function for the two-layer feedforward neural network in Figure 3 is defined as:

$$Q(z) = Q^{(1)}(z) + Q^{(2)}(z)$$

$$= \sum_{s=1}^S \left[\sum_{k=1}^K \left(\sum_{i=0}^I w_{ki}^{(1)} x_{is} - f_k^{(1)-1}(z_{ks}) \right)^2 + \sum_{j=1}^J \left(\sum_{k=0}^K w_{jk}^{(2)} z_{ks} - f_j^{(2)-1}(y_{js}) \right)^2 \right]. \quad (22)$$

Thus, using the outputs z_{ks} we can learn, for each layer independently, the weights $w_{ki}^{(1)}$ and $w_{jk}^{(2)}$ by solving the corresponding linear system of (19). After this, the sensitivity (see (21)) with respect to z_{ks} is calculated thus:

$$\frac{\partial Q}{\partial z_{ks}} = \frac{\partial Q^{(1)}}{\partial z_{ks}} + \frac{\partial Q^{(2)}}{\partial z_{ks}}$$

$$= - \frac{2 \left(\sum_{i=0}^I w_{ki}^{(1)} x_{is} - f_k^{(1)-1}(z_{ks}) \right)}{f_k'^{(1)}(z_{ks})} + 2 \sum_{j=1}^J \left(\sum_{r=0}^K w_{jr}^{(2)} z_{rs} - f_j^{(2)-1}(y_{js}) \right) w_{jk}^{(2)}, \quad (23)$$

with $k = 1, \dots, K$, as $z_{0s} = 1$, for all s . After this, the values of the intermediate outputs z are modified using the Taylor series approximation:

$$Q(z + \Delta z) = Q(z) + \sum_{k=1}^K \sum_{s=1}^S \frac{\partial Q(z)}{\partial z_{ks}} \Delta z_{ks} \approx 0, \quad (24)$$

which leads to the following increments:

$$\Delta z = -\rho \frac{Q(z)}{\|\nabla Q\|^2} \nabla Q, \quad (25)$$

where ρ is a relaxation factor or step size.

3.2.2. Strengths and Weaknesses of SBLLM. (A) It has been established that the SBLLM offers an interesting combination of speed, reliability, and simplicity. In addition, based on the results obtained from the real-world experiments using the SBLLM learning algorithm, there are four main advantages of the SBLLM that can be summarized as follows, [14].

- (i) *High speed in reaching the minimum error:* it was demonstrated in [14] that in all cases the SBLLM obtains its minimum mean squared error (MSE_{\min}) just before the first four iterations and also sooner than the rest of the algorithms examined together. SBLLM gets its minimum error in an epoch for which the other algorithms are far from similar MSE values.

- (ii) *Good performance:* it can be deduced not only that SBLLM stabilizes soon, but also the minimum MSE that it reaches is quite good and comparable to that obtained by the second order methods. Other methods compared to it never succeeded in attaining this minimum MSE (MSE_{\min}) before the maximum number of epochs.

- (iii) *Homogeneous behavior:* the SBLLM learning curve stabilizes easily and within short time. The SBLLM behaves homogeneously not only if we consider just the end of the learning process, but also during the whole process, in such a way that very similar learning curves were obtained for all iterations of different experiments. This is indicative of ability to handle prediction with high stability and consistency which are requisite of good prediction system particularly in oil and gas reservoir modeling.

(B) Weaknesses of SBLLM. One of the major weaknesses of this technique is its inability to model uncertainties which is a very important capability being sought for in today's predictive solutions particularly in oil and gas modeling where uncertainties are very common. Also the prediction accuracy of SBLLM is usually dependent on the nature of the problem at hand, that is, nature of the data set. Thus there is need to compliment its usage with other better model like type-2 FLS to achieve better performance, particularly in the face of uncertainties.

3.3. The Proposed T2-SBLLM Hybrid Framework. In this proposed hybrid system, type-2 fuzzy logic system is first used for modeling and uncertainty handling using its procedural components of fuzzification, inferencing, type-reduction, and defuzzification processes to generate a clean output, while the Sensitivity based linear learning method (SBLLM) is then trained using the cleaned output from type-2 FLS and then used to make the final predictions. This is then applied to the problem of PVT property prediction in the field of reservoir engineering, but it is extendable for use in other prediction or classification oriented applications.

3.3.1. Conceptual Design of the Proposed T2-SBLLM Hybrid Models. Hybrid computational intelligence is defined as any effective combination of intelligent techniques that performs superior or in a competitive way to simple standard intelligent techniques. The increased popularity of hybrid intelligent systems in recent times lies in the extensive success of these systems in many real-world complex problems [1]. Also, it is an established fact that every approach has its strengths and weaknesses; hence the need for hybrid models that is able to combine the strengths of the individual techniques while complementing the weaknesses of one method with the strength of the other.

In this section, smart hybridization of type-2 FLS and sensitivity based linear learning method (SBLLM) is investigated, in order to take advantages of hybrid systems in general while using each component to complement the other in order to take advantages of each framework in particular,

where possible, while alleviating the weaknesses of component members in order to achieve effective, stable, consistent, and accurate predictive solution. The methodology in this work is based on the standard Computational Intelligence approach to hybridization of ion different techniques. The hybrids are designed in order to benefit immensely from the strength of the individual techniques while complementing the weaknesses of each technique with the advantages of the others through smart and optimum combination of the cooperative and competitive characteristics of the individual techniques.

In this hybrid system, type-2 fuzzy logic system is first used for modeling and uncertainty handling using its procedural components of fuzzification, inferencing, type-reduction, and defuzzification processes to generate a clean output, while the sensitivity based linear learning method (SBLLM) is then trained using the cleaned output from type-2 FLS and then used to make the final predictions. This is then applied to the problem of PVT properties and permeability prediction in the field of reservoir engineering, but it is extendable for use in other prediction or classification oriented applications.

In this case, a T2-SBLLM hybrid system that comprises of two building blocks made up of type-2 fuzzy logic systems (T2) and sensitivity based linear learning method (SBLLM), has been built. Having divided the input data into training and testing sets using the stratified sampling approach, the training and testing sets are passed to the type-2 FLS block where all possible forms of uncertainties are adequately handled. This capability of type-2 FLS has been shown in several works such as in [10, 62, 65] and this has been confirmed in this work also. It is able to adequately handle different forms of uncertainties using its extension to a third dimension. The type-2 FLS undergoes the necessary processes of fuzzification, inferencing, type-reduction, and final defuzzification to generate final crisp output. The final outputs from the type-2 FLS for both training and testing set are then passed to the SBLLM block for training and testing purposes, respectively. Thus, the output from the type-2 training process is then used to train the SBLLM in readiness for prediction using the unseen test dataset that happens to be the output from the type-2 FLS testing process. It must be noted that the target values are never manipulated by the type-2 FLS block in order to avoid creating bias as a result of such process. This means that the actual values (also known as target values) are just made to be read only for type-2 FLS. Finally, the new unseen test data, which is the type-2 FLS testing output, is then passed to the trained SBLLM model to perform the final prediction task.

The role performed by the type-2 FLS block in this model is to ensure that dataset containing uncertainties has been properly handled and cleaned to generate output that is later passed to the SBLLM block for training and then final prediction purposes. As a result of this, it ensures that, during the training process, clean and trusted data is allowed to enter the SBLLM block that is responsible for carrying out the prediction task after the training process, and this in turn facilitates better performance of the SBLLM model

produced. Figure 5 shows the conceptual design framework of the proposed T2-SBLLM hybrid system.

With the designed framework shown in Figure 5, the hybridization of type-2 FLS and sensitivity based linear learning method become clear to develop and implement. Meanwhile each of the designated steps in Figure 5 is briefly explained as follows.

Step 1. This is the step where the necessary real industrial dataset are made ready in preparation to be sent into the hybrid model.

Step 2. In this step, the available data is passed to type-2 fuzzy logic systems as the first component of the hybrid scheme. Several major subactivities take place here. These include fuzzification, type-2 FLS rule extraction, consequent matching, type reduction, and defuzzification processes. These processes ensure adequate uncertainty handling and modeling based on the entire processes of type-2 FLS. It must be noted that the dataset is already divided into training set and testing set and each is being processed separately, one after the other, at this stage. This means that, at the end of this step, two different outputs are generated, one from the training set and the other from the testing set. These two different outputs (which are the predicted target variable) are then presented to the next step of the hybrid scheme.

Step 3. In this step, the two different outputs of training and testing, coming from the type-2 FLS, are distinctly and, respectively, preserved in the form of training and testing sets to be passed to the next component of the hybrid model, which is SBLLM. Thus, this step separately passed the new training data set into the next step (Step 4) where training and calibration of the SBLLM take place, while the testing set is retained only to be passed to the Step 5 where testing of the trained SBLLM takes place.

Step 4. In this stage, SBLLM is being trained using the training output from type-2 FLS that has undergone adequate uncertainty handling procedures of type-2 FLS. Therefore, SBLLM is being trained using clean and better output from type-2 FLS. This ensures that SBLLM is being trained in such a way to prevent the effect of uncertainties on its performance, thereby facilitating the generation of better and improved model with the ultimate goal of achieving better performance accuracy.

Step 5. In this step, the trained SBLLM in previous step is being tested using the testing dataset retained in Step 3 for onward transfer to this present step. Thus, the SBLLM is being tested using the output from the type-2 FLS testing section. This ensures that clean and better data is passed to the SBLLM for final testing and the ultimate prediction needed. The testing at this step is the final testing for the entire hybrid model presented.

Step 6. This is the stage where final predicted output from the hybrid model is collected and necessary performance measure analysis and computations are carried out to

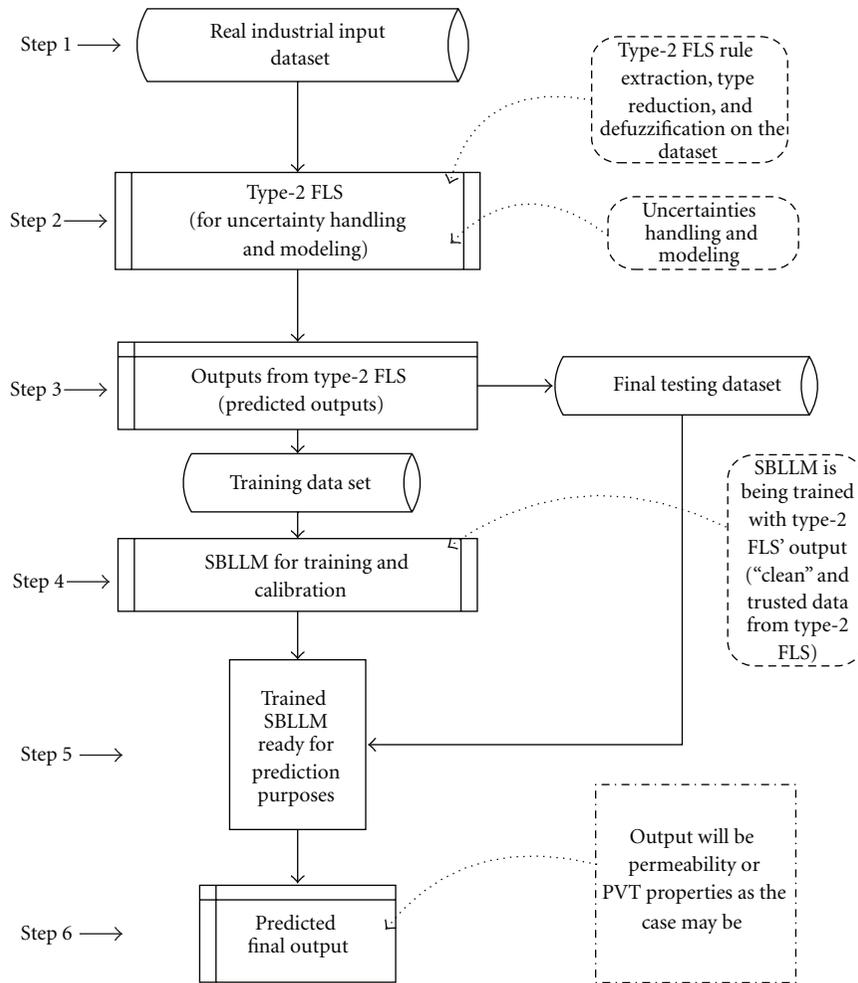


FIGURE 5: Schematic design framework for the proposed T2-SBLLM hybrid model for predicting PVT properties.

identify the performance accuracy gains of the proposed hybrid model over the individual constituting models. The most common statistical quality measures that are utilized in both data mining and petroleum engineering journals were employed in this stage, and their description is presented in Section 4.2.

To further make the flow of data through the entire processes clearer, a data flow diagram for the proposed T2-SBLLM hybrid model is presented as in Figure 6.

It must be stated here that the proposed model followed strictly the standard training and testing procedures in which the testing set is kept away as unseen data before it is sent to the model for testing. From the data flow diagram depicted in Figure 6, it is made clear how the dataset is first partitioned into training set and testing set. The standard procedure for the sequential implementation of hybrid model is followed in this work, where the preceding method (T2FLS) always acted on the dataset to generate an output that is then fed to the next model in the hybrid setup (i.e., SBLLM in this present work.). The well log attributes are first sent to the T2FLS for training. Based on the training set, T2FLS is made

to generate an output representing its estimated permeability values. These estimated permeability values produced from the training section of T2FLS are then passed into the SBLLM as input to be used for training. Having trained the SBLLM model with the estimated permeability from T2FLS the system proceeded to the testing process as follows.

As shown in Figure 6, the testing set is separated into the attributes and the target (actual) permeability. The testing set attributes are first sent into T2FLS for testing and the T2FLS generates its final estimates of the target output (PVT properties). The estimated PVT property values produced from the testing section of T2FLS are then passed into the trained SBLLM as input to be used in carrying out its testing procedure. At the end of this SBLLM testing, final predicted values for the PVT properties are then generated as the final output from the hybrid scheme. These predicted outputs are then compared to the preserved target (actual) PVT properties in order to determine the accuracy of the final predicted output.

Depicted in Figure 7 is the network diagram demonstrating how the estimated values of PVT properties from T2FLS are made to enter the SBLLM network as input to the system.

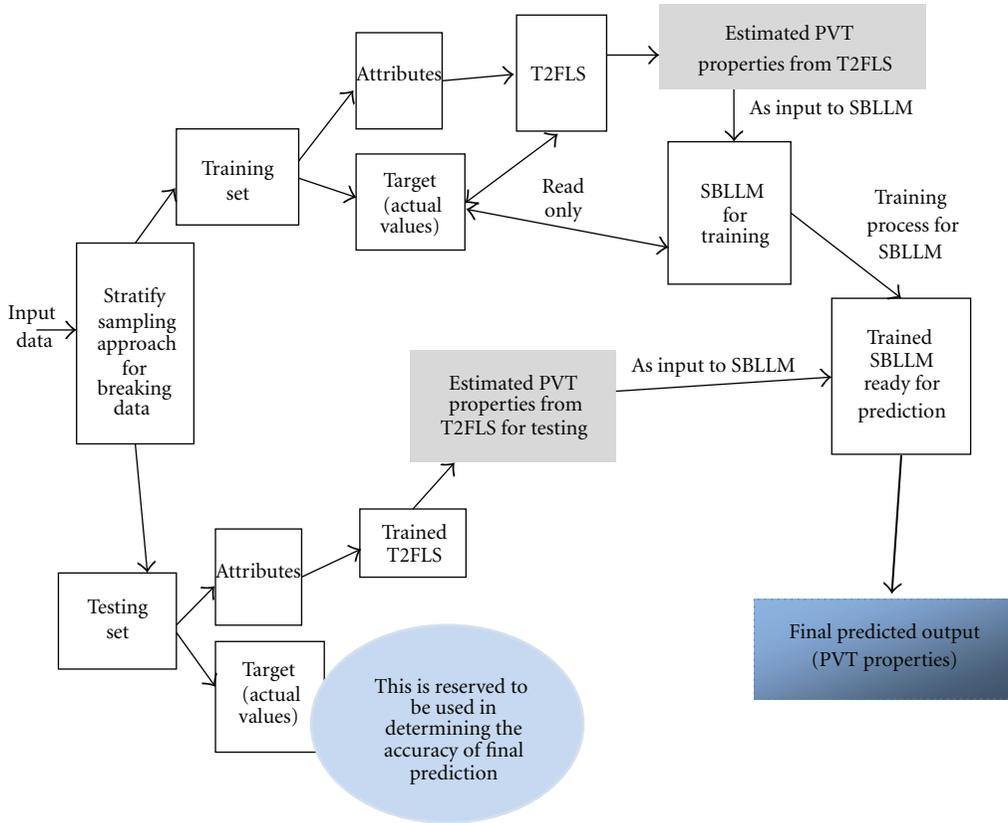


FIGURE 6: Data flow diagram of the proposed T2-SBLLM hybrid model.

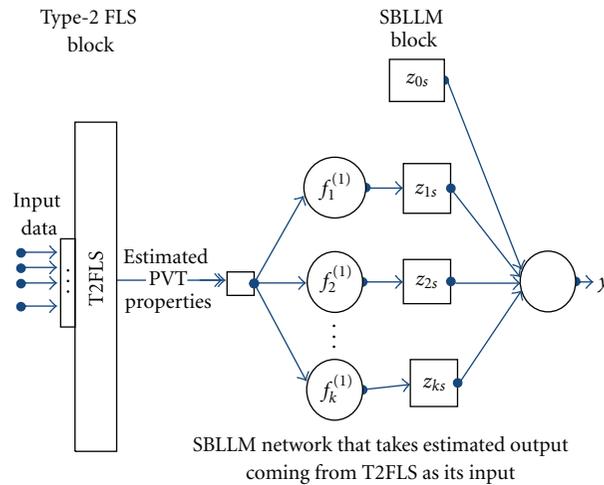


FIGURE 7: Network diagram of the proposed hybrid showing the flow of data from T2FLS into SBLLM for both training and testing phases (y stands for the final PVT properties output from the hybrid model).

It must be noted that the network diagram depicted in Figure 7 holds for both training and testing phases. For the testing phase, the training data set is fed into the T2FLS and the flow continues as depicted in the figure. While in case of testing phase, the T2FLS is fed with testing data set and the estimated permeability output from T2FLS then goes into the SBLLM as its input as shown in the figures.

It must however be noted that the training set is different from the testing set that is set aside for ascertaining the

predictive capability of the proposed models. Also, it must be noted that the T2FLS is first used on both training and testing set one after the other before the estimated output from each case is then passed into the SBLLM block for training and testing, respectively, as detailed in Figures 6 and 7.

3.3.2. *Optimal Parameters Search Procedure for SBLLM.* The parameters associated with the SBLLM were optimized through a test-set-cross-validation on the available data set.

The details of the test-set-cross-validation for optimizing the SBLLM parameters goes thus: For each run of generated training and testing set, the values of RMSE and correlation coefficient were monitored for a group of parameters that includes a number of hidden neurons (N) and the activation functions (AF). Searching through all possible values of the parameters in a given range will identify the best performance measures and the corresponding values of the parameters for the fixed set of features. In our experiment, this process was repeated for every SBLLM activation function available, each time with an incremental step of parameters. The optimal values of the parameters and the kernel option associated with the best performance measure were identified. A summary of the procedure is as follows.

Step 1. Choose the initial “activation function” option from the list of available options.

Step 2. Identify the best values of the number hidden neuron through a test-set-cross-validation and store the corresponding performance measures.

Step 3. If there is no activation function option left, then go to Step 4. Otherwise, add the next activation function option and go to Step 2.

Step 4. Identify the best performance measure and its associated parameters values.

Step 5. Use the optimized activation function option and the parameters values to train the final SBLLM

Step 6. Calculate the performance measures for both the training and testing sets using the system obtained in the previous Step 5.

This is presented in mathematical form as follows.

Let the set A contain all the possible activation function options, the element of A is of the shape $A_i(j)$, where i is the activation function number, j is the selected number of hidden neurons, nf is the total number of activation functions available, and nh is the maximum number of hidden neuron assumed. Also pm represents performance measure taken, ix represents index for best activation function, and jx represents index for best number of layer.

The algorithm then goes thus:

initialization; $jx = 0, vx = 0, ix = 0,$

for $i = 1 \rightarrow nf,$

for $j = 1 \rightarrow nh,$

$pm = f(A_i(j))$ {performance measure for the present parameters combination},

if pm is better than vx then $vx = pm,$

$ix = i; jx = j$ {storing the index of the better parameter},

end,

end.

4. Empirical Study, Results, and Discussion

In order to carry out an empirical study, three distinct databases were acquired. To evaluate the performance of each modelling scheme, the entire database was divided, using the stratified sampling approach, into training set and testing set. The training set (70% of the entire dataset) was used for training and building the proposed implemented models (internal validation) while the testing set (remaining 30%) was used for testing and validating the models. For testing and evaluation of the newly proposed hybrid framework and to carry out effective comparative studies, the most common statistical quality measures that are utilized in both data mining and petroleum engineering journals were employed in this study, and their brief descriptions are given shortly.

4.1. Acquired Datasets. For this study, three distinct datasets have been acquired. The complete databases were earlier utilized in distinct published research articles. They include: (a) 160 observations-database; (b) 283 observations-database; and (c) 782 observations-database. Details of each are as follows.

- (a) 160-Dataset This first database was drawn from the article [66] containing published correlations for estimating bubble point pressure and oil formation volume factor for Middle Eastern oils. This database contains 160 observation data drawn from Middle Eastern reservoirs.
- (b) 283-Dataset This second database was contained in the works of [19, 53]. This database has 283 data points collected from different Saudi fields to predict the bubble point pressure and the oil formation volume factor at the bubble-point pressure for Saudi crude oils.
- (c) 782-Dataset This third database was obtained from the works of [20, 50], This database contains 782 observations after removing the redundant 21 observations from the actual 803 data points. This data set was gathered from Malaysia, the Middle East, the Gulf of Mexico, and Colombia.

One of the unique attributes of the three databases is that they all share the same input attributes (independent variables) and these include gas-oil ratio, API oil gravity, relative gas density, and reservoir temperature.

4.1.1. Uncertainties and Complexity Involved in Oil and Gas Reservoir Data. The complexity involved in reservoir exploration, which invariably leads to different forms of uncertainties, randomness, or irregularity in data are hereby discussed, based on literatures and opinion of experts in the field of reservoir engineering. It is an established fact that reservoir characterization involves handling uncertainties [13]. For instance, there is no causal, mathematically describable relationship between the porosity and permeability of sedimentary rocks. While, at least theoretically, porosity is independent of grain size, permeability is strongly dependent on it, through the specific surface factor figuring in the

Kozeny-Carman equation [67]. Because of the lack of a well-defined porosity-permeability correlation, in permeability prediction from well logs other rock-physical properties are utilized, such as electrical, radioactive, and sonic properties of the rock, obtained from well logs. As the underlying physical principles connecting these very different, and generally indirectly measured quantities, and relating them to permeability, are not known as yet, the only way to proceed seems to be to rely on probabilistic techniques in one form or other and apply multivariable regression analysis, fuzzy algorithms, or artificial neural-network techniques (Professor Gabor Korvin, personal communication, October 17, 2010).

The problem of permeability prediction is especially complicated in carbonate rocks whose depositional and diagenetic history can be very complex, so that their permeability cannot be causally upscaled from core scale to reservoir scale, or even up to a few feet scale seen by the well logs. Larger than core-sized, vuggy, or fractured intervals in carbonates can result in permeability which at the scale of a few feet are significantly higher than the matrix permeability measured in core plugs. Swarms of fractures, if connected, yield very large flow rates, if disconnected, very low flow rates—and this capricious variability is not recorded (or deeply hidden) by the core permeability data (Professor Gabor Korvin, personal communication, October 17, 2010).

A further problem arising when matching core data and well-log data is that the depth value indicated on the well log is never more precise than one part in a thousand, that is an exact depth correspondence between core and well-log value is impossible [13, 68].

The rock properties, k , otherwise called permeability can be determined by time-based recording of only one variable, the pressure change in each reservoir. Experimental error during data acquisition propagates through the data reduction process leading to uncertainty in experimental results. In addition, unlike steady-state systems, the pressure-time curves are influenced by the compressive storage of the reservoirs and both the dimensions and properties of the sample [69]. Thus, uncertainty in permeability and PVT properties may arise from errors in measurement of sample dimension, fluid pressure, or reservoir storages. Since reservoirs are typically small for tight rock samples and irregular in shape due to the combination of tubing, fittings, valves, and pressure transducers, thus the uncertainty is expected to be higher (Prof. Gabor Korvin, personal communication, August 23, 2009).

4.2. Criteria for Performance Evaluation. Different quality measures can be used to judge the performance and accuracy of the models. This is done by carrying out statistical error analysis. To evaluate and compare the performance and accuracy of the proposed SBLLM models with the earlier mentioned standard neural networks and the three common empirical correlations in literatures, the most common statistical quality measures that are utilized in both petroleum engineering and data mining journals, namely, average absolute percent relative error (E_a), standard deviation (SD)

and correlation coefficient (R^2) have been employed; see [48, 50] for details regarding their mathematical formulae. However, their brief descriptions are given below.

(i) *Correlation Coefficient*. The correlation coefficient measures the statistical correlation between the predicted and actual values. This method is unique in model evaluations. A higher number means a better model, with a “1” meaning perfect statistical correlation and a “0” meaning there is no correlation, indicating a failed performance. The formula is

$$\frac{\sum (y_a - y'_a)(y_p - y'_p)}{\sqrt{\sum (y_a - y'_a)^2 \sum (y_p - y'_p)^2}}, \quad (26)$$

where y_a and y_p are the actual and predicted values while y'_a and y'_p are the mean of the actual and predicted values.

(ii) *Average Absolute Percent Relative Error (E_a)*. The relative error is the absolute error divided by the magnitude of the exact value. The percent error is the relative error expressed in terms of per 100. And thus, the average absolute percent relative error (E_a) is the average of the absolute percent relative error for all the cases. Mathematically put as:

$$E_a = \sum_{i=1}^n \left| \frac{y_a^i - y_p^i}{y_a^i} \times 100 \right|, \quad (27)$$

where n is the total number of samples, y_a is the original (actual) values known, and y_p is the predicted values

(iii) *Standard Deviation (SD)*. Standard deviation is a measure of the average distance between individual data points and their mean. It is a measure of how stable a model result is when repeated over several runs. It provided a measure of confidence; the higher the standard deviation (“sigma”), the lower the prediction reliability. This is very useful for understanding the risk of data extrapolation. A model will be adjudged to be consistent and stable if the standard deviation is very low.

The mathematical formula for determining standard deviation is:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}. \quad (28)$$

An equivalent version of this formula is:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}}. \quad (29)$$

This can be interpreted as the square root of a summation divided by n , where n is the total number of samples, X_i is ith sample, and \bar{X} is the mean.

Percentage of Improvement. This is a quality measure devised to make clear and understandable, the amount of improvement a particular model had over another model. It can be

a percentage increase as in the case of correlation coefficient (R^2) or percentage decrease for the case of standard deviation (SD). For instance, for a model to be adjudged as better than another in terms of R^2 , it must achieve a higher R^2 but in terms of SD, it has to achieve a lower value. Based on these two possible cases, to which any of the quality measures must belong to either of the two but not both, the formulae for calculating the percentage increase and decrease, respectively, are formulated as follows:

$$\begin{aligned} \text{percentage increase} &= \frac{\text{difference between the two values}}{\text{the lower of the two values}} \\ &\quad \times 100, \\ \text{percentage decrease} &= \frac{\text{difference between the two values}}{\text{the higher of the two values}} \\ &\quad \times 100. \end{aligned} \tag{30}$$

4.3. Experimental Environments and Settings. To evaluate performance of the proposed T2-SBLLM hybrid modeling scheme, the acquired database is divided, using the stratified sampling approach, into 80% training set and 20% testing set for estimating how the investigated model performed on new unseen data. For testing and evaluation of the newly developed framework, and to carry out effective comparative studies viz-a-viz other earlier methods, the most common statistical quality measures that are utilized in both petroleum engineering and data mining journals were employed in this study, and they were already discussed in the preceding section. The training set is first pass through type-2 FLS block for proper uncertainty handling and then the output coming from the type-2 FLS is used to train the SBLLM model. Thereafter, testing data is then used to evaluate the predictive capability of the trained SBLLM model. We repeat both internal and external validation processes for each of the considered models. The obtained results are presented in tables that follow shortly.

The training sets were used to build the models while the testing sets were utilized in evaluating the predictive capability of the models. As for the implementation, we did not use any ready-made software, the entire coding has been done using MATLAB though some MATLAB inbuilt functions, most especially in the case of SBLLM, and few others made available online, have been called and used in some cases. Also part of the type-2 fuzzy logic functions made available in [8] were also made use of.

In the case of the type-2 FLS based model, the implementation process proceeded by supplying the system with the available input data sets, one sample at a time, and the rules and membership functions are automatically learned from the available input data. Gaussian membership function has been used based on two different learning criteria that include least squares and back-propagation. The same combination was utilized in training FLS membership function parameters. Further details on initializing, training, and validating type type-2 FLS have been presented in Section 3, and additional details could be found in [8, 70, 71].

TABLE 1: Testing results for 160-dataset when predicting B_{ob} and P_b . R^2 : correlation coefficient, SD: standard deviation, E_a : average absolute percent relative error (AAPRE), P_b : bubble-point pressure, B_{ob} : oil formation volume factor.

Prediction methods	B_{ob}			P_b		
	R^2	SD	E_a	R^2	SD	E_a
Type-2 fuzzy model	0.994	0.09166	1.493	0.931	2.461	20.65
SBLLM	0.995	0.09154	1.200	0.961	4.060	35.54
T2-SBLLM hybrid	0.996	0.01427	1.013	0.986	2.134	17.65

As for the sensitivity based linear learning method (SBLLM) implementation, the number of hidden neuron was set to be 1000 while the activation function used was sigmoidal (sig) activation function. All these were arrived at using the optimisation procedure described earlier.

5. Results and Discussions

The results of comparisons using external validation checks (testing on unseen data) have been summarized in Tables 1, 2, and 3. The performance results from hybridized type-2-SBLLM model outperformed each of the constituting individual models, which is in line with the general establish fact, to date, that a hybrid scheme often performs better than the individual constituent parts. The proposed hybrid model showed high accuracy in predicting both PVT property values with a stable and accurate performance and achieved the lowest standard deviation in all cases, lowest absolute percent relative error, and the highest correlation coefficient in most cases in comparison to the individual constituent model. Detailed discussion of the results for each model follows shortly.

Judging from the results summarized in Tables 1 through 3, it is clear that the proposed hybrid scheme is better than individual methods because a good forecasting scheme should have the highest correlation coefficient (R^2), the lowest standard deviation (SD), and the lowest absolute percent relative error (E_a).

From the tables presented, it could be easily observed that the proposed hybrid system performs better than the two individual constituent models that include type-2 FLS and SBLLM. The hybrid has proven to be a better way to boost the performance of SBLLM as the results indicated that it has greatly improved upon the performance of SBLLM, up to 96.9% improvement in terms of standard deviation (SD), 8.6% improvement in terms of correlation coefficient (R^2), and finally up to 95% improvement in terms of average absolute percent relative error (E_a). The performance of the newly proposed hybrid greatly outperforms the standard SBLLM model thereby serving as a better improved form of SBLLM, while it has also performed better than the type-2 FLS. These results are in line with the established fact that hybrid models usually performed better than any of their individual constituting models. Further result analyses are presented as follows.

It could be easily observed, for instance, in estimating bubble-point pressure (P_b) based on the 782 dataset, that

TABLE 2: Testing results for 283-dataset when predicting B_{ob} and P_b . R^2 : correlation coefficient, SD: standard deviation, E_a : average absolute percent relative error (AAPRE), P_b : bubble-point pressure, B_{ob} : oil formation volume factor.

Prediction methods	B_{ob}			P_b		
	R^2	SD	E_a	R^2	SD	E_a
Type-2 fuzzy model	0.9948	0.2662	1.069	0.9978	0.8873	1.034721
SBLLM	0.9959	0.0372	0.852	0.9597	0.1427	30.205
T2-SBLLM hybrid	0.9978	0.00815	0.510	0.9968	0.0943	2.1034

TABLE 3: Testing results for 782-dataset when predicting B_{ob} and P_b . R^2 : correlation coefficient, SD: standard deviation, E_a : average absolute percent relative error (AAPRE), P_b : bubble-point pressure, B_{ob} : oil formation volume factor.

Prediction methods	B_{ob}			P_b		
	R^2	SD	E_a	R^2	SD	E_a
Type-2 fuzzy model	0.9998	0.1625	0.1	0.9894	2.347	0.3432
SBLLM	0.9826	0.4069	2.742	0.9138	1.3328	38.6214
T2-SBLLM hybrid	0.9975	0.0421	1.9866	0.9927	1.031	3.6273

T2-SBLLM hybrid system had 8.6% improvement over that of sensitivity based linear learning method (SBLLM) and 0.33% improvement over that of type-2 FLS in terms of correlation coefficient (R^2). In terms of standard deviation (SD), T2-SBLLM hybrid model had 22.7% improvement over SBLLM and 56.1% over type-2 FLS, while in terms of absolute percent relative error (E_a), T2-SBLLM hybrid model had 90.6% improvement over SBLLM and 90.5% over type-2 FLS. Similarly, for the case involving estimating oil formation volume factor (B_{ob}) using the 783-dataset, T2-SBLLM hybrid system had 1.5% improvement over that of sensitivity based linear learning method (SBLLM) while type-2 FLS had 0.2% improvement over it, in terms of correlation coefficient (R^2). In terms of standard deviation (SD), T2-SBLLM hybrid model had 89.7% improvement over SBLLM and 74.1% over type-2 FLS, while in terms of absolute percent relative error (E_a), T2-SBLLM hybrid model had 27.6% improvement over SBLLM. Other reported results also follow similar trends with T2-SBLLM hybrid model taking the lead always.

Moreover, for the case involving estimating oil formation volume factor (B_{ob}) based on the 283-dataset, T2-SBLLM hybrid model had 78.1% improvement over that of sensitivity based linear learning method (SBLLM) and 96.9% improvement over that of type-2 FLS, in terms of standard deviation (SD). In terms of absolute percent relative error (E_a), T2-SBLLM hybrid model had 40.1% improvement over SBLLM and 52.3% over type-2 FLS, while in terms of correlation coefficient (R^2), T2-SBLLM hybrid model had 0.2% improvement over SBLLM and 0.3% over type-2 FLS. Meanwhile, for the case involving estimating bubble-point pressure (P_b) using the 283-dataset, T2-SBLLM hybrid system had 3.9% improvement over that of sensitivity based linear learning method (SBLLM) in terms of correlation coefficient (R^2), 33.9% in terms of standard deviation (SD) and 90% in term of absolute percent relative error (E_a).

As for the case involving estimating bubble-point pressure (P_b) based on the 160-dataset, T2-SBLLM hybrid model had 47.4% improvement over that of sensitivity based linear learning method (SBLLM) and 13.3% improvement over

that of type-2 FLS in terms of standard deviation (SD). In terms of absolute percent relative error (E_a), T2-SBLLM hybrid model had 50.3% improvement over SBLLM and 14.5% over type-2 FLS, while in terms of correlation coefficient (R^2), T2-SBLLM hybrid model had 2.6% improvement over SBLLM and 5.9% over type-2 FLS. Meanwhile, for the case involving estimating oil formation volume factor (B_{ob}) using the 160-dataset, T2-SBLLM hybrid system had 84.4% improvement over that of sensitivity based linear learning method (SBLLM) and 84.43% improvement over that of type-2 FLS in terms of standard deviation (SD). In terms of absolute percent relative error (E_a), T2-SBLLM hybrid model had 15.6% improvement over SBLLM and 32.2% over type-2 FLS, while in terms of correlation coefficient (R^2), T2-SBLLM hybrid model had 0.1% improvement over SBLLM and 0.2% over type-2 FLS in terms of correlation coefficient (R^2).

From the overall reported experimental results, it could be easily noted that T2-SBLLM hybrid model performed better in all fronts. This is evident as its quality measure values are consistently better than others in all fronts. Even though the performance of type-2 FLS and SBLLM might be closer on very few cases, there was no single case where they are closer in performance in terms of standard deviation (SD), which is a measure of the stability of the predictive systems. As for the SBLLM model, the newly proposed T2-SBLLM hybrid model outperformed it throughout the reported experimental results. This is an indication that the proposed approach has greatly improved the capability of the classical SBLLM through the incorporation of type-2 FLS as preprocessor to the actual SBLLM, in the form of what is popularly regarded as hybrid system. The overall results also indicate that the newly proposed T2-SBLLM hybrid model is able to consistently deal with the nature of reservoir data due to its ability to cater for all forms of uncertainties, using its type-2 FLS component while ensuring a better generalization and higher stability and consistency using its SBLLM component.

6. Conclusion and Recommendations

A new hybrid model combining type-2 fuzzy logic system (T2) and sensitivity based linear learning method (SBLLM) has been proposed and implemented. The proposed hybrid serves as a better improvement over the classical SBLLM by way of using type-2 FLS as precursor to the SBLLM model. The proposed T2-SBLLM hybrid has also been used to model the PVT properties of crude oil systems using three distinct published databases. This is to investigate the performance and accuracy of the proposed T2-SBLLM hybrid system, while at the same time solving the challenging prediction problems in oil and gas industry. Further conclusion emanating from this research and future work recommendation are presented as follows.

A new hybrid modelling scheme, through the appropriate combination of type-2 FLS and SBLLM, has been investigated, developed, and implemented, as predictive solution that takes care of all forms of uncertainties, while ensuring stable and consistent predictions. It has been shown, through adequate simulation works, that the newly proposed hybrid system is able to provide a better prediction for PVT properties of crude oil systems. Validation of the framework has been carried out using published databases. In-depth comparative studies have been carried out between this new framework and each of the constituent models that include type-2 FLS and SBLLM. Also the proposed method has been specially compared to the classical sensitivity based linear learning machines in order to show the improvement provided by the new hybrid model over SBLLM. The empirical results have confirmed the superiority of proposed T2-SBLLM hybrid over SBLLM in all fronts. Similar improvement patterns were also recorded against type-2 FLS though at a lesser degree of improvement. Thus, the overall empirical results from the experimental works and simulations show that the proposed model outperformed all the other individual constituting models in all fronts. Given any new data, the proposed T2-SBLLM will be able to handle uncertainties that might be present and perform the required prediction effectively with stable and consistent results through unique combination of uncertainty handling capability of type-2 FLS and the unique generalization capability, consistency and stability of sensitivity based linear learning methods.

The proposed system will also be useful for other classification problems, bearing in mind that regression model can easily be modified to take care of classification problems unlike the classification model that cannot easily be made to carry out regression. Thus this work should be seen as a cutting edge solution in the field of pattern recognition, in general, as a tool for both regression and classifications.

As an intake from the promising results from this work, it is suggested, as a form of future work, that this newly proposed system should be considered as viable tools for other reservoir engineering problems like that of PVT properties, porosity, history matching, lithofacies, and other reservoir engineering properties, while also looking into its usefulness in other germane fields such as time series

forecasting, bioinformatics, intrusion detection systems, and others.

Acknowledgments

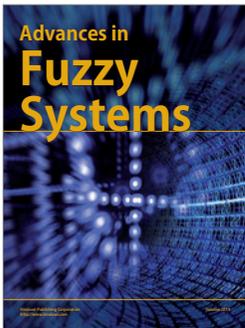
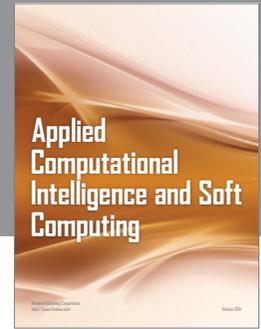
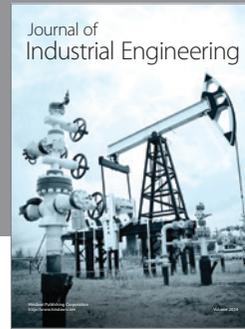
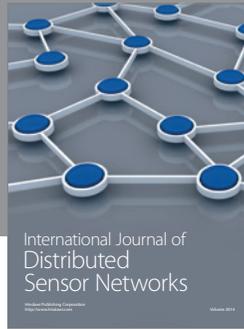
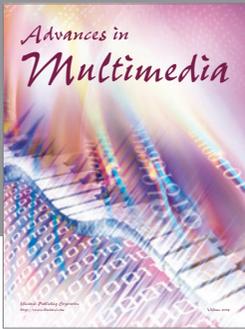
The Zamalah/Institutional Ph.D. scholarship provided by UTM and The Ministry of Higher Education of Malaysia is hereby acknowledged. King Fahd University of Petroleum and Minerals, Saudi Arabia (KFUPM) is hereby acknowledged for the use of some facilities in the course of this research work.

References

- [1] A. Giovanni and L. Vincenzo, "Using FML and fuzzy technology in adaptive ambient intelligence environments," *International Journal of Computational Intelligence Research*, vol. 1, no. 2, pp. 171–182, 2005.
- [2] N. N. Karnik and J. M. Mendel, "Type-2 fuzzy logic systems: type-reduction," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2046–2051, IEEE, October 1998.
- [3] Y. Li, X. Sun, J. Hua, and C. Gong, "Modelling redundant structure in ecosystem by type-2 fuzzy logic system," *Ecological Modelling*, vol. 211, no. 1-2, pp. 113–120, 2008.
- [4] Q. Liang and N. N. Karnik, "Connection admission control in ATM networks using survey-based type-2 fuzzy logic systems," *IEEE Transactions on Systems, Man and Cybernetics Part C*, vol. 30, no. 3, pp. 329–339, 2000.
- [5] Q. Liang and J. M. Mendel, "Interval type-2 fuzzy logic systems: theory and design," *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 5, pp. 535–550, 2000.
- [6] Q. Liang and J. M. Mendel, "Overcoming time-varying co-channel interference using type-2 fuzzy adaptive filters," *IEEE Transactions on Circuits and Systems II*, vol. 47, no. 12, pp. 1419–1428, 2000.
- [7] M. H. Fazel Zarandi, B. Rezaee, I. B. Turksen, and E. Neshat, "A type-2 fuzzy rule-based expert system model for stock price analysis," *Expert Systems with Applications*, vol. 36, no. 1, pp. 139–154, 2009.
- [8] J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice-Hall, 1st edition, 2001.
- [9] J. M. Mendel, "Fuzzy sets for words: a new beginning," in *Proceedings of the 12th IEEE International conference on Fuzzy Systems*, pp. 37–42, Los Angeles, Calif, USA, May 2003.
- [10] D. Wu and J. M. Mendel, "A vector similarity measure for linguistic approximation: interval type-2 and type-1 fuzzy sets," *Information Sciences*, vol. 178, no. 2, pp. 381–402, 2008.
- [11] M. R. Mashinchi and A. Selamat, "Constructing a customer's satisfactory evaluator system using GA-based fuzzy artificial neural networks," *World Applied Sciences Journal*, vol. 5, no. 4, pp. 432–440, 2008.
- [12] M. Reza Mashinchi and A. Selamat, "An improvement on genetic-based learning method for fuzzy artificial neural networks," *Applied Soft Computing Journal*, vol. 9, no. 4, pp. 1208–1216, 2009.
- [13] K. I. Ali, R. Mohammadreza, and A. M. Seyed, "A fuzzy logic approach for estimation of permeability and rock type from conventional well log data: an example from the Kangan reservoir in the Iran Offshore Gas Field," *Journal of Geophysics Engineering*, vol. 3, no. 4, pp. 356–369, 2006.

- [14] E. Castillo, B. Guijarro-Berdiñas, O. Fontenla-Romero, and A. Alonso-Betanzos, "A very fast learning method for neural networks based on sensitivity analysis," *Journal of Machine Learning Research*, vol. 7, pp. 1159–1182, 2006.
- [15] E. Castillo, A. Conejo, A. S. Hadi, and A. Fernández-Canteli, "A general method for local sensitivity analysis with application to regression models and other optimization problems," *Technometrics*, vol. 46, no. 4, pp. 430–444, 2004.
- [16] O. Omole, O. A. Falode, and A. D. Deng, "Prediction of nigerian crude oil viscosity using artificial neural network," *International Journal for Petroleum Processing*, vol. 51, no. 3, pp. 181–188, 2009.
- [17] E. A. Osman, M. A. Al-Marhoun, and K. Fand, "Artificial neural networks models for predicting PVT properties of oil field brines," in *Proceedings of the 14th SPE Middle East Oil and Gas Show and Conference (MEOS '05)*, pp. 1501–1516, Bahrain, March 2005.
- [18] M.B. Standing, "Oil-system correlation," in *Petroleum Production Handbook*, T. C. Frick, Ed., McGraw-Hill, New York, NY, USA, 2nd edition, 1962.
- [19] A. O. Kumoluyi and T. S. Daltaban, "Higher-order neural networks in petroleum engineering," in *Proceedings of the 64th Annual Western Regional Meeting*, pp. 555–570, Longbeach, Calif, USA, March 1994.
- [20] H. M. Goda, "Prediction of the PVT data using neural network computing theory," in *Proceedings of the 27th Annual SPE International Technical Conference and Exhibition in Abuja, Abuja, Nigeria*, 2003.
- [21] E. A. El-Sebakhy, "Forecasting PVT properties of crude oil systems based on support vector machines modeling scheme," *Journal of Petroleum Science and Engineering*, vol. 64, no. 1–4, pp. 25–34, 2009.
- [22] A.E. El-Sebakhy, A. Abdulraheem, M. Ahmed et al., "Functional network as a novel approach for prediction of permeability in a carbonate reservoir," in *Proceedings of the SPE Conference*, 2007.
- [23] M. B. Standing, "A pressure-volume-temperature correlation for mixtures of california oils and gases," *Drilling and Production Practice*, pp. 275–287, 1947.
- [24] O. Glaso, "Generalized pressure-volume temperature correlations," *Journal of Petroleum Technology*, vol. 32, pp. 785–795, 1980.
- [25] M.A. Al-Marhoun, "New correlation for formation volume factor of oil and gas mixtures," *Journal of Canadian Petroleum Technology*, vol. 31, no. 3, pp. 2–22, 1992.
- [26] R. Labedi, "Use of production data to estimate volume factor, density and compressibility of reservoir fluids," *Journal of Petroleum Science and Engineering*, vol. 4, no. 4, pp. 375–390, 1990.
- [27] M. E. Dokla and M. E. Osman, "Correlation of PVT properties for UAE crudes," *SPE Formation Evaluation*, vol. 7, no. 1, pp. 41–46, 1992.
- [28] H. Y. Al-Yousef and M. A. Al-Marhoun, "Discussion of correlation of PVT properties for United Arab Emirates crudes," in *Proceedings of the Society of Petroleum Engineers Conference*, UAE, 1993.
- [29] S. M. Macary and M. H. El-Batanony, "Derivation of PVT correlations for the gulf of Suez crude oils," in *Proceedings of the EGPC 11th Petroleum Exploration and Production Conference*, Cairo, Egypt, 1992.
- [30] A. M. Saleh, I. S. Maggoub, and Y. Asaad, "Evaluation of empirically derived PVT properties for Egyptian oils," in *Proceedings of the Middle East Oil Show and Conferenc*, SPE, Bahrain, 1987.
- [31] M. I. Omar and A. C. Todd, "Development of new modified black oil correlations for Malaysian crudes," in *Proceedings of the Asia Pacific Oil and Gas Conference*, pp. 211–219, Singapore, February 1993.
- [32] G. E. Petrosky and F. F. Farshad, "Pressure-volume-temperature correlations for Gulf of Mexico crude oils," in *Proceedings of the SPE Annual Technical Conference and Exhibition*, pp. 395–406, Houston, Tex, USA, October 1993.
- [33] T. Kartoatmodjo and Z. Schmidt, "Large data bank improves crude physical property correlations," *Oil and Gas Journal*, vol. 92, no. 27, pp. 51–55, 1994.
- [34] R. A. Almehaideb, "Improved PVT correlations for UAE crude oils," in *Proceedings of the 10th Middle East Oil Show & Conference*, pp. 109–120, Bahrain, March 1997.
- [35] W. D. McCain, "Reservoir-fluid property correlations—state of the art," *SPE Reservoir Engineering*, vol. 6, no. 2, pp. 266–272, 1991.
- [36] A. M. Elsharkawy, A. A. Elgibaly, and A. A. Alikhan, "Assessment of the PVT correlations for predicting the properties of Kuwaiti crude oils," *Journal of Petroleum Science and Engineering*, vol. 13, no. 3-4, pp. 219–232, 1995.
- [37] M. A. Mahmood and M. A. Al-Marhoun, "Evaluation of empirically derived PVT properties for Pakistani crude oils," *Journal of Petroleum Science and Engineering*, vol. 16, no. 4, pp. 275–290, 1996.
- [38] H. H. Hanafy, S. M. Macary, Y. M. ElNady, A. A. Bayomi, and M. H. El Batanony, "Empirical PVT correlations applied to Egyptian crude oils exemplify significance of using regional correlations," in *Proceedings of the SPE International Symposium on Oilfield Chemistry*, pp. 733–737, Houston, Tex, USA, February 1997.
- [39] A. A. Al-Shammasi, "Bubble point pressure and oil formation volume factor correlations," in *Proceedings of the 11th SPE Middle East Oil Show & Conference*, pp. 241–256, Bahrain, February 1999.
- [40] A. A. Al-Shammasi, "A review of bubblepoint pressure and oil formation volume factor correlations," *SPE Reservoir Evaluation and Engineering*, vol. 4, no. 2, pp. 146–160, 2001.
- [41] O. O. Bello, K. M. Reinicke, and P. A. Patil, "Comparison of the performance of empirical models used for the prediction of the PVT properties of crude oils of the niger delta," *Petroleum Science and Technology*, vol. 26, no. 5, pp. 593–609, 2008.
- [42] S. S. Ikiensikimama and O. Ogboja, "New bubblepoint pressure empirical PVT correlation," in *Proceedings of the 33rd Annual SPE International Technical Conference and Exhibition*, Abuja, Nigeria, 2009.
- [43] S. Mohaghegh, "Virtual-intelligence applications in petroleum engineering—part I—artificial neural networks," *Journal of Petroleum Technology*, vol. 52, no. 9, pp. 64–73, 2000.
- [44] S. Mohaghegh and S. Ameri, "An artificial neural network as a valuable tool for petroleum engineers," in *Proceedings of the Society of Petroleum Engineers*, 1994.
- [45] J. K. Ali, "Neural networks: a new tool for the petroleum industry," in *Proceedings of the European Petroleum Computer Conference*, Aberdeen, UK, 1994.
- [46] A. R. Moghadassi, F. Parvizian, S. M. Hosseini, and A. R. Fazlali, "A new approach for estimation of PVT properties of pure gases based on artificial neural network model," *Brazilian Journal of Chemical Engineering*, vol. 26, no. 1, pp. 199–206, 2009.
- [47] H. Hassanzadeh, M. Pooladi-Darvish, A. M. Elsharkawy, D. W. Keith, and Y. Leonenko, "Predicting PVT data for CO₂—brine mixtures for black-oil simulation of CO₂ geological storage,"

- International Journal of Greenhouse Gas Control*, vol. 2, no. 1, pp. 65–77, 2008.
- [48] R. O. Duda, P. E. Hart, and D. G. Stock, *Pattern Classification*, John Wiley and Sons, New York, NY, USA, 2001.
- [49] A. Sharifi, A. Moghadassi, F. Parvizian, and S. M. Hosseini, “Prediction of Pvt properties of ammonia by using artificial neural network and equations of state,” *ARPN Journal of Engineering and Applied Sciences*, vol. 3, no. 6, pp. 18–27, 2008.
- [50] E. A. Osman, O. A. Abdel-Wahhab, and M. A. Al-Marhoun, “Prediction of oil PVT properties using neural networks,” in *Proceedings of the SPE Middle East Oil Show*, pp. 893–906, Bahrain, March 2001.
- [51] R. B. Gharbi and A. M. Elsharkawy, “Neural network model for estimating the PVT properties of Middle East crude oils,” in *Proceedings of the 10th Middle East Oil Show & Conference*, pp. 151–166, Bahrain, March 1997.
- [52] N. Varotsis, V. Gaganis, J. Nighswander, and P. Guieze, “Novel non-iterative method for the prediction of the PVT behavior of reservoir fluids,” in *Proceedings of the SPE Annual Technical Conference and Exhibition*, Houston, Tex, USA, October 1999.
- [53] M. A. Al-Marhoun and E. A. Osman, “Using artificial neural networks to develop new PVT correlations for Saudi crude oils,” in *Proceedings of the 10th Abu Dhabi International Petroleum Exhibition and Conference (ADIPEC '02)*, Abu Dhabi, UAE, 2002.
- [54] A. Moghadassi, F. Parvizian, and S. Hosseini, “A new approach based on artificial neural networks for prediction of high pressure vapor-liquid equilibrium,” *Australian Journal of Basic and Applied Sciences*, vol. 3, no. 3, pp. 1851–1862, 2009.
- [55] A. M. Elsharkawy, “Modeling the properties of crude oil and gas systems using RBF network,” in *Proceedings of the Asia Pacific Oil & Gas Conference*, pp. 35–46, Perth, Australia, October 1998.
- [56] E. A. Osman and R. E. Abdel-Aal, “Abductive networks: a new modeling tool for the oil and gas industry,” in *Proceedings SPE Asia Pacific Oil and Gas Conference and Exhibition*, pp. 487–493, Melbourne, Australia, October 2002.
- [57] A. E. El-Sebakhy, A. Abdurraheem, M. Ahmed et al., “Functional network as a novel approach for prediction of permeability in a carbonate reservoir,” in *SPE Conference*, 2007.
- [58] J. M. Mendel and R. I. B. John, “Type-2 fuzzy sets made simple,” *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 2, pp. 117–127, 2002.
- [59] E. Cox, “Adaptive fuzzy systems,” *IEEE Spectrum*, vol. 30, no. 2, pp. 27–31, 1993.
- [60] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York, NY, USA, 1982.
- [61] L. A. Zadeh, “The concept of a linguistic variable and its application to approximate reasoning-I,” *Information Sciences*, vol. 8, no. 3, pp. 199–249, 1975.
- [62] N. N. Karnik, J. M. Mendel, and Q. Liang, “Type-2 fuzzy logic systems,” *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 6, pp. 643–658, 1999.
- [63] E. Castillo, A. Cobo, J. Manuel Gutiérrez, and E. Pruneda, “Working with differential, functional and difference equations using functional networks,” *Applied Mathematical Modelling*, vol. 23, no. 2, pp. 89–107, 1999.
- [64] E. Castillo, J. M. Gutiérrez, and A. S. Hadi, “Sensitivity analysis in discrete Bayesian networks,” *IEEE Transactions on Systems, Man, and Cybernetics Part A*, vol. 27, no. 4, pp. 412–423, 1997.
- [65] M. H. Fazel Zarandi, B. Rezaee, I. B. Turksen, and E. Neshat, “A type-2 fuzzy rule-based expert system model for stock price analysis,” *Expert Systems with Applications*, vol. 36, no. 1, pp. 139–154, 2009.
- [66] M. A. Al-Marhoun, “PVT correlations for Middle East crude oils,” *Journal of Petroleum Technology*, vol. 40, no. 5, pp. 650–666, 1988.
- [67] S. Mohaghegh, B. Balan, and S. Ameri, “State-of-the-art in permeability determination from well log data—part 2—verifiable, accurate permeability predictions, the touch-stone of all models,” in *Proceedings of the 1995 Eastern Regional Conference*, pp. 43–47, Morgantown, WV, USA, September 1995.
- [68] N. Hurtado, M. Aldana, and J. Torres, “Comparison between neuro-fuzzy and fractal models for permeability prediction,” *Computational Geosciences*, vol. 13, no. 2, pp. 181–186, 2009.
- [69] I. Song, A. P. Rathbun, and D. M. Saffer, “Uncertainty of permeability and specific storage due to experimental error during data acquisition for pulse-transient technique,” in *Proceedings of the American Geophysical Union Fall Meeting*, 2011.
- [70] J. M. Mendel and R. I. B. John, “Type-2 fuzzy sets made simple,” *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 2, pp. 117–127, 2002.
- [71] N. N. Karnik, J. M. Mendel, and Q. Liang, “Type-2 fuzzy logic systems,” *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 6, pp. 643–658, 1999.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

