10.2478/msr-2014-0030

Soft Measurement of Water Content in Oil-Water Two-Phase Flow Based on RS-SVM Classifier and GA-NN Predictor

Dongzhi Zhang, Bokai Xia

College of Information and Control Engineering, China University of Petroleum (East China), Qingdao 266580, China, email: dzzhang@upc.edu.cn

Measurement of water content in oil-water mixing flow was restricted by special problems such as narrow measuring range and low accuracy. A simulated multi-sensor measurement system in the laboratory was established, and the influence of multi-factor such as temperature, and salinity content on the measurement was investigated by numerical simulation combined with experimental test. A soft measurement model based on rough set-support vector machine (RS-SVM) classifier and genetic algorithm-neural network (GA-NN) predictors was reported in this paper. Investigation results indicate that RS-SVM classifier effectively realized the pattern identification for water holdup states via fuzzy reasoning and self-learning, and GA-NN predictors are capable of subsection forecasting water content in the different water holdup patterns, as well as adjusting the model parameters adaptively in terms of online measuring range. Compared with the actual laboratory analyzed results, the soft model proposed can be effectively used for estimating the water content in oil-water mixture in all-round measuring range.

Keywords: Oil-water mixture, water content measurement, soft model, multi-sensor, model prediction.

1. INTRODUCTION

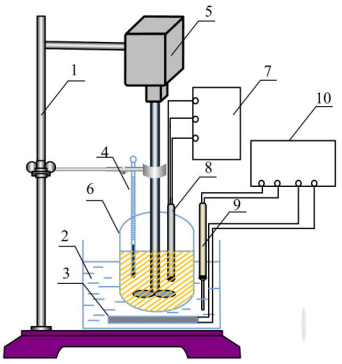
THE MEASUREMENT of water content has attracted a widespread attention in oil-water two-phase flow because it is crucially important in aspects of petroleum logging, refining, production and export [1-2]. So far, many efforts have been made to develop several typical instrumental techniques, such as capacitance method [3-4], microwave technique [5-6] and near-infrared spectroscopy [7-8]. The capacitance sensor depends on the significant difference in the dielectric permittivity of water and oil, changing the capacitance proportionally with the varying permittivity of the oil/water mixture. The microwave resonance technique also relies on measuring the dielectric permittivity of an oil/water mixture. Water content measurement is achieved by sending microwave signals through the oil/water mixture flow, and the microwave signal changes with the water content. Near-infrared spectroscopy is employed for detection of water content via measuring the absorption, reflection and scatter of the infrared beam caused by the dielectric difference of the fluid composition. The present instrumental techniques for water content in oil/water mixture mostly focus on the difference in dielectric property for water and oil, and suffer from the influence of multi-factors on the measurement precision and the measuring range. For instance, the temperature results in the change of dielectric permittivity for oil/water mixture, and the salinity content leads to the change of density, conductance and viscosity for the mixture. These issues make the accuracy measurement level in the industrial applications currently about 10 %, which is unsatisfactory to meet the higher-accuracy requirements for digital oilfield development in petroleum industry [9-11].

Up to now, there are some difficulties in the postprocessing for the commercial instrument to reduce biased data and results affected by the complicated nonlinear factor of oil-water two-phase flow, as well as the cross-sensitivity of multi-factors therein. Therefore, model-based soft measurement method becomes a significant way to solve this problem, which can improve the defect of hardware instrument through software compensation. Zhao et al. established a prediction model of artificial neural network (ANN) for the measurement of oil-water two-phase flow, and obtained good results for the water content measurement in the range of 51 % to 91 % [12]. However, ANN is based on the traditional empirical risk minimization and suffers from several drawbacks, such as the difficulty in obtaining a stable solution, over-fitting, and thus it cannot estimate one global optimal value due to the random selection of initial weights. Xia et al. and Zhang et al. developed soft sensors using genetic algorithm (GA) optimized radical basis function (RBF) neural network and wavelet neural network models for water-content-in-oil, respectively, and certified their effectiveness for the soft model [10, 13]. The support vector machine (SVM) as a newly developed soft measurement method based on the statistical learning, also has been gradually applied to the estimation of phase fraction in multiphase flow, providing a promising perspective for the explanation and prediction of water content in crude oil [14].

Although some investigations have reported on the transduction and software techniques in the present instruments, special problems such as narrow measuring range, and low accuracy restrict their applications. In this paper, a simulated multi-sensor measurement system in the laboratory was established, and the influence of multi-factor such as temperature, and salinity content on the measurement was investigated by numerical simulation combined with experimental test. Subsequently, a novel compound model based on rough set-support vector machine (RS-SVM) classifier and genetic algorithm-neural network (GA-NN) predictors was presented to achieve the high accuracy measurement of water content in all-round measuring range.

2. EXPERIMENTAL SYSTEM

The multi-sensor measuring system employed in this paper is shown in Fig.1., which is composed of water bath boiler, temperature controller, blender, thermometer, dielectric sensor, tri-neck boiling container, temperature sensor, etc. The testing sample is oil-water two-phase flow, which has water content varying from 0 to 100 %, temperature varying from 20 to 90°C, and salinity content varying from 0 to 50 g/L. A blender is employed to obtain a uniform status of oil-water mixture.



1. hob, 2. water bath boiler, 3. resistive heater, 4. thermometer, 5. blender, 6. tri-neck boiling container, 7. measuring circuit, 8. dielectric sensor, 9. temperature sensor, 10. temperature controller

Fig.1. Illustration of experimental equipment.

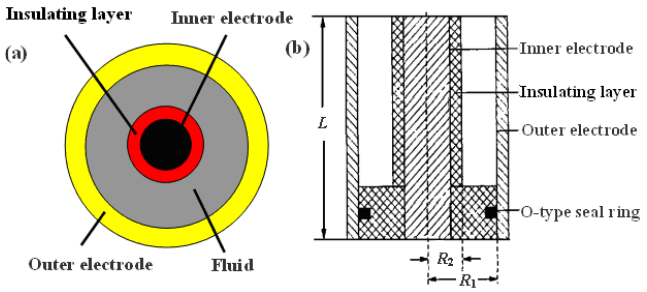


Fig.2. Illustration of sensor structure.

The dielectric sensor is designed to be a coaxial-torus condenser, and its structure is shown in Fig.2. The fluid is filled in the toroidal cavity between the inner and outer electrodes, the capacitance of the sensor is a function of liquid permittivity, which is given by

$$C_x = \frac{2\pi\varepsilon_0\varepsilon L}{\ln(R_1/R_2)} \tag{1}$$

where ε_0 and ε are the relative dielectric permittivity of vacuum and the considered fluid, respectively. L is the length of sensor probe, R_1 is the inner radius of outer electrode, and R_2 the outer radius of inner electrode. The liquid permittivity is measured by using the electromagnetic-wave harmonic technique. The harmonic circuit is motivated by a 40 MHz quartz oscillator, and the sensor used as the tuning capacitor C_x. Fig.3. shows the schematic diagram of the sensor working principle. The sensor capacitance changes along with the water content ratio in the oil-water mixture, and is converted to voltage signal by a microprocessor and associated components. During the experiment, the first step is manual matching of the mixture sample of water and oil with different water content ratio, next is heating the mixture sample to an expected temperature, and adding salinity content to the sample for forming mineralized fluid, and finally collecting the measuring data using a logger.

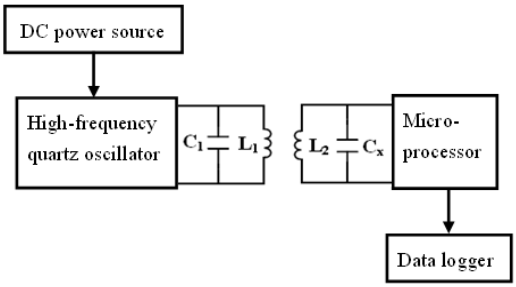


Fig.3. Schematic diagram of sensor working principle.

3. MULTI-FACTOR INFLUENCE

For oil-water two-phase flow, its effective permittivity is related to water content and mixture state of oil and water, but not equal to the arithmetic mean of dielectric coefficient of oil and water by reason of different polarization for polar water molecules and non-polar oil molecules under alternating electric field. The relative permittivity of oil-water mixture is depicted by

$$\varepsilon_{ow} = k[\varepsilon_o(1-\lambda)+\varepsilon_w\lambda]+(1-k)\frac{\varepsilon_w\varepsilon_o}{\lambda\varepsilon_o+(1-\lambda)\varepsilon_w} \tag{2}$$

where λ is the water content ratio, ε_o, ε_w, ε_{ow} are the relative permittivity of oil, water and oil-water mixture, respectively. k is a coefficient related to water content and distribution of water phase, and its empirical formula is given as

$$k = \frac{2\lambda^{0.25}}{5-3\lambda} \tag{3}$$

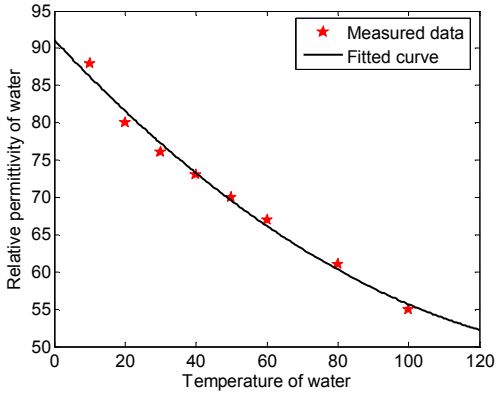


Fig.4. Relative permittivity of water vs. its temperature.

The relative permittivity of oil ε_o is affected slightly by its temperature, it commonly changes in 2.2~2.5 and 2.3 is widely used. However, the relative permittivity of water ε_w is affected greatly by the temperature, and the relationship between them is shown in Fig.4. The fitting equation for the relative permittivity of water ε_w as a function of temperature t can be depicted as ε_w =91.1238-0.5077t+0.0015t^2.

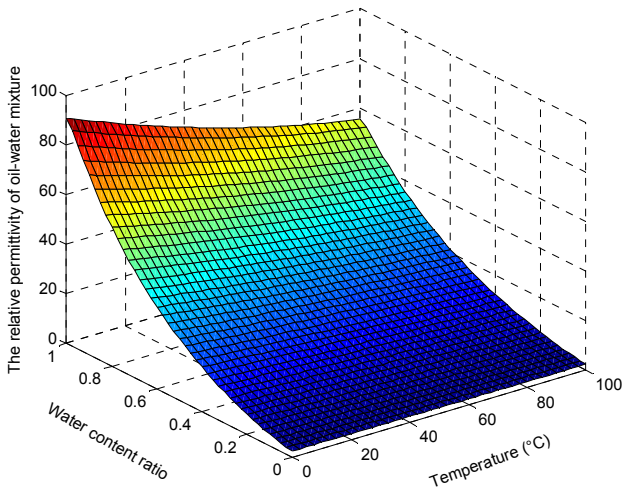


Fig.5. Numerical simulation for the influence of temperature and water content ratio on the relative permittivity of oil-water mixture.

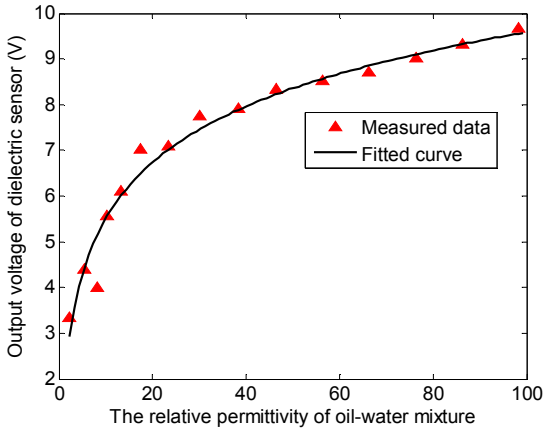


Fig.6. Output voltage vs. relative permittivity of oil-water mixture.

The numerical simulation for the interaction influence of temperature and water content ratio on the relative permittivity of oil-water mixture is shown in Fig.5. The relative permittivity of oil-water mixture ε_{ow} decreases with rising temperature if under the same water content ratio, but increases with rising water content ratio if under the same temperature. The characteristics of the dielectric sensor under different permittivity of oil-water mixture are investigated. At 60°C, the output voltage measured and relative permittivity of oil-water mixture is shown in Fig.6. The output voltage u and the relative permittivity of oil-water mixture ε_{ow} can be fitted by:

$$u = a + b\ ln\ (\varepsilon_{ow}) \tag{4}$$

where a=1.4217, b=1.7687 are fitted coefficients. From Fig.6. we can see that the sensitivity of the dielectric sensor is higher at low ε_{ow} than that at high ε_{ow}. This characteristic allows the dielectric sensor to achieve high-precision at low water content and worse at higher water content.

For the oil-water two-phase flow with varying water content ratio, the influence of different salinity content on the output voltage u of dielectric sensor is investigated at 25°C. The result is plotted in Fig.7. The salinity content is imposed to the oil-water mixture as 400, 800, 2400, 5000, 25000 mg/L, respectively. The influence of salinity content on the output voltage u is relatively small when $\lambda \leq 0.3$, while the output voltage becomes salient with visible inflexion when $0.3<\lambda<0.7$, and then decreases when $\lambda \geq 0.7$. This phenomenon is related to different characteristic in flow state of water-in-oil, phase-transformation and oil-in-water. The water-in-oil type occurs if the water content is lower than 30 %, the oil-in-water type occurs if the water content is higher than 70 %, and the phase transition occurs if the water content is in the range of 30 %-70 %.

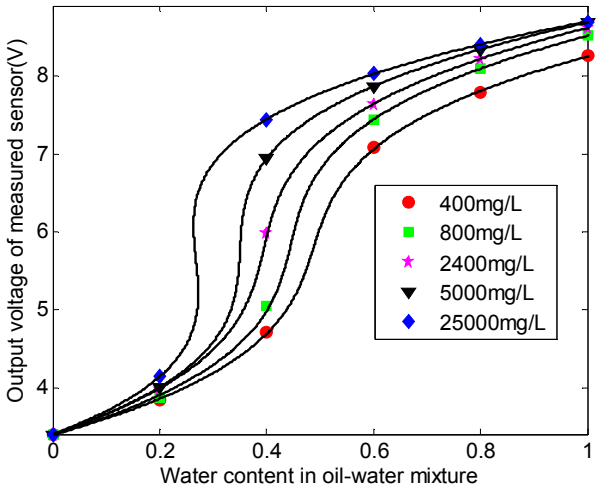


Fig.7. Relationship of sensor voltage vs. water content under different salinity content.

The output voltage of dielectric sensor for the oil-water mixture with varying salinity content is plotted in Fig.8. The measurement is implemented at 25°C for the oil-water mixture with water content ratio λ=0.4, 0.6, 0.8 and 1. The sensor voltage and salinity content s present a linear characteristic when $s<3$ *g/L*, and show a transition when 3 *g/L*$<s<$10 *g/L*, while becoming constant and saturated when $s>10$ *g/L*. See from the experiment results, the online measurement of water content in water-oil mixture based on the widely used method of dielectric permittivity is affected by some factors including temperature, salinity content, and oil-water mixture state.

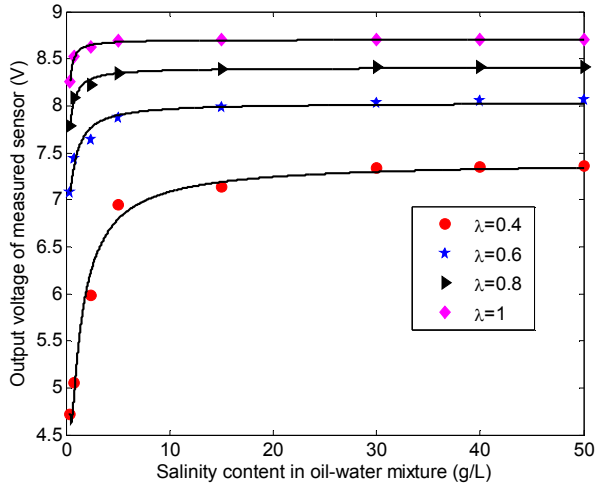


Fig.8. Relationship of voltage signal vs. salinity content at different water content.

4. COMPOUND MODEL OF RS-SVM CLASSIFIER AND GA-NN PREDICTOR

The present instruments for water content measurement in oil-water two-phase flow have their specific measuring range and are hardly for all-range. Therefore, the automatic adjustment of the soft model parameters in terms of the water holdup pattern of oil-water mixture is still a focus and difficulty for the high-precision online measurement of water content. The internal phase state of oil-water two-phase flow has different characteristics of unsteady and time-variations, commonly the oil-water emulsion can be divided into patterns of water-in-oil ($\lambda \leq 0.3$), phase-transformation ($0.3<\lambda<0.7$) and oil-in-water ($\lambda \geq 0.7$) [10]. In order to improve the on-line measuring accuracy and widen the measuring range, a novel compound model is employed for measuring water content by combining RS preprocessor, SVM classifier and GA-NN predictors. The structural diagram for the compound model is shown in Fig.9. The multi-variable array concluding temperature, salinity content, and dielectric voltage is divided into two parts: train-set and test-set. The influencing variables are employed as model inputs and the corresponding water content is used as output of the compound model. The RS-SVM classifier is presented for realizing correct recognition of low, middle, high water holdup patterns via fuzzy reasoning and self-learning. The three GA-NN predictors are employed to provide subsection forecast of water content for the different water holdup patterns, and the model parameters can be adjusted adaptively in terms of online measuring range automatic switching. The principle of this compound model is, firstly, recognizing the water-content-pattern for the current oil-water sample, and then automatically switching to the corresponding pattern range when the corrected value of water content is out of current measuring scope. The combination of RS-SVM classifier and GA-NN predictor can adjust and correct every detected water content value in terms of measured multi-variable data. So, the measurement errors can be reduced to the greatest extent.

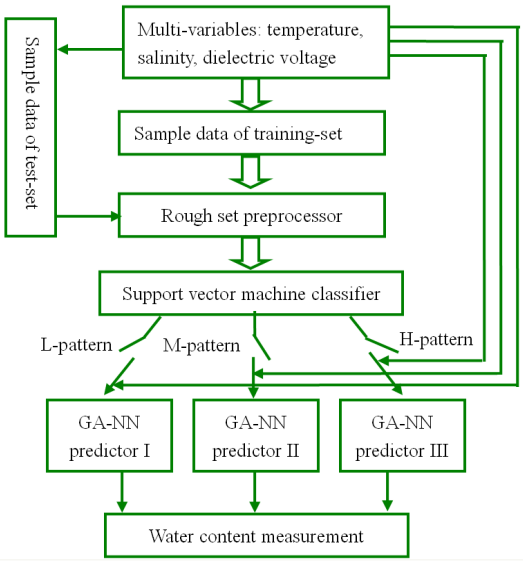


Fig.9. Structural diagram of compound model with RS-SVM classifier and GA-NN predictors.

4.1. Rough set preprocessor.

Rough set theory proposed firstly by Pawlak has attracted much attention for dealing with vague and uncertain problems, and is used in many fields such as data mining, machine learning, pattern identifying and decision-making. It is based on the indiscernibility relation that describes indistinguishable objects, and concepts are represented by lower and upper approximations. The recognition of oil-water patterns depends on useful knowledge or rules extracted from the experimental measured multi-variable array. The kernel algorithm in the RS preprocessor adopts a clustering method of self-organizing map neural network (SOM-NN) algorithm based on unsupervised learning [15].

The decision system for oil-water patterns can be described as $S=(U,C\bigcup\{d\},V,f)$, where $U=\{u_1, u_2,\ldots,u_n\}$ is the universe of measured data entity, $C=\{c_1, c_2, c_3\}$ is the condition attribute, c_1, c_2, c_3 are the continuous attribute values for temperature, salinity content and voltage of dielectric sensor, respectively, d is the decision attribute of oil-water patterns, $V=\{V_{c1}, V_{c2}, V_{c3}, V_d\}$, $V_{c1}, V_{c2}, V_{c3}, V_d$ are the corresponding numerical range of c_1, c_2, c_3, d, respectively, f is the mapping function $U\times C\rightarrow V$.

Here, the numerical range V_{c1} for temperature is divided equally into k_{c1} intervals ($k_{c1}=3$ is given this paper), which is expressed as follows:

$$\begin{cases} h=[\max(V_{c1})-\min(V_{c1})]/k_{c1} \\ p^{i}_{c1}=\left[\min(V_{c1})+i\times h,\min(V_{c1})+(i+1)\times h\right) \\ p_{c1}=\{p^{i}_{c1}, i=1,2,\cdots,k_{c1}\} \end{cases} \quad (5)$$

here $\max(V_{c1})$ and $\min(V_{c1})$ are the upper and lower bounds of V_{c1}, h is the interval stepsize, p^{i}_{c1} the divided sub-interval set with corresponding quantified value $i=1, 2, \ldots, k_{c1}$.

The numerical range V_{c2} for salinity is divided into three intervals according to salinity holdup as light saltwater (≤3 *g/L*), saltwater (3~10 *g/L*) and heavy saltwater (≥10 *g/L*), the corresponding discrete values are 1, 2, 3 for each kind. V_d is divided into low (0~0.3), middle (0.3~0.7) and high (0.7~1) three intervals, and the corresponding quantified values are 1, 2, 3. The discretization of V_{c3} uses the novel clustering method based on self-organizing map artificial neural network (SOM-NN), which is described by:

(1) Initialize the clustering number (neural nodes of competition layer) M=1 to V_{c3}.

(2) Provide random number in [0, 1] to weight vector of neural network W_i=(w_{i1}, w_{i2}, …, w_{in}). Give the initialization $\eta(0)\in(0, 1)$ for learning rate $\eta(k)$, $N_g(0)$for neighborhood field $N_g(k)$, and final learning epoch T.

(3) Provide sample pattern X=(x_1, x_2, …, x_n) to the input layer of neural network, and then it is normalized by

$$\overline{X}=\frac{X}{\|X\|}=\frac{(x_1,x_2,\cdots,x_n)}{(\sum_{i=1}^{n}x_i^{\,2})^{\frac{1}{2}}} \tag{6}$$

(4) Normalize the connecting weights vector W_i by

$$\overline{W_i}=\frac{W_i}{\|W_i\|}=\frac{(w_{i1},w_{i2},\cdots,w_{in})}{(\sum_{j=1}^{n}w_{ij}^{\,2})^{\frac{1}{2}}} \tag{7}$$

and calculate the Euclidean distance between $\overline{W_i}$ and $\overline{X}$ by

$$d_i=\left[\sum_{j=1}^{n}(\overline{x_j}-\overline{w_{ij}})^2\right]^{\frac{1}{2}},\ \text{here } i=1, 2, \ldots, M \tag{8}$$

(5) Find the minimum distance d_s and the winning neuron g: d_s=min[d_i], here i=1, 2, …, M.

(6) Modify the connecting weights of neurons both in neighbor's field $N_g(k)$ of competition layer and input layer, and the modified principle is given by

$$\overline{W_i}(k+1)=\overline{W_i}(k)+\eta(k)[\overline{X}-\overline{W_i}(k)] \tag{9}$$

here i=1,2, ⋯, M, $\eta(k)\in(0, 1)$.

(7) Update the learning rate $\eta(k)$ and neighbor's field $N_g(k)$ in terms of the rules given by

$$\eta\ (k)=\eta(0)(1-\frac{k}{T}) \tag{10}$$

$$N_g(k)=INT[N_g(0)(1-\frac{k}{T})] \tag{11}$$

(8) Adjust k by adding 1, then return to step (3) until k=T. Divide V_{c3} into M sub-intervals by using bound points and thresholds, and the corresponding quantified values are 1, 2, …, M.

(9) Check the consistency of the decision table constituted by the discrete values of c_1, c_2, c_3 and d. If it is achieved, then the discretization of V_{c3} is successful; if not, adjust M by adding 1, return to step (1) and train again.

(10) Reduce the redundant information and create the minimum rule set with consistency check.

The SOM-NN provides an effective method for data discretization and classification. Finally, M =5 is obtained in this paper, and a consistent decision table is achieved. And next, the decision table characterized by small samples is self-learned by a support vector machine classifier.

4.2. Support vector machine classifier.

Support vector machine (SVM) is employed to deal with small sample and incomplete decision table in this paper. The presence of well-trained SVM classifier can not only avoid complex rule reasoning, but also establish the nonlinear mapping for condition attributes and decision attributes [16-18]. Given linearly separable training samples {(x_i, y_i)| i=1, 2, …, n }, where $x_i\in R^n$ is a vector of input samples and $y_i\in${-1, 1} indicates the class to which the point x_i belongs. The linear discriminant function for SVM classifier is given as

$$f(x, w)=\text{sign}(w\cdot x+b) \tag{12}$$

where w is the weight vector, and b is the bias.

For solving the problem of nonlinear classification, a non-negative slack variable ξ_i ($\xi_i\geq0$) is introduced to make the hyperplane $w^Tx+b=0$ subject to

$$y_i(w^Tx_i+b)\geq1-\xi_i \tag{13}$$

Pattern sample points x_i can be correctly separated when $0<\xi_i<1$, and improperly separated when $\xi_i\geq1$. This problem can now be realized through solving the following optimization problem

$$\psi(w,\xi)=\frac{1}{2}\|w\|^2+\gamma\sum_{i=1}^{n}\xi_i \tag{14}$$

here γ is a penalty parameter which obviously balances the empirical risk and the capacity of the machine. This serves to avoid an under-fitting and also over-fitting of the training data. To solve this optimization problem, Lagrange function is constructed through a quadratic programming as

$$L(w,b,\xi,\alpha)=\frac{1}{2}\|w\|^2+\gamma\sum_{i=1}^{n}\xi_i^2-\alpha_i\left[y_i(w^Tx_i+b)+\xi_i-1\right] \tag{15}$$

where α_i are Lagrange multipliers, ξ_i is the error on the ith training data point. The solution of (15) can be obtained by partial differentiation with respect to w, b, ξ and α, setting

the derivatives equal to zero. Elimination of w and ξ_i through substitution results in the following set of linear equations

$$\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & K(x_1,x_1)+\gamma^{-1} & \cdots & K(x_1,x_n) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & K(x_n,x_1) & \cdots & K(x_n,x_n)+\gamma^{-1} \end{bmatrix} \cdot \begin{bmatrix} b \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix} \tag{16}$$

where $K(x_i, x)=\Phi(x_i)^T\Phi(x)$ is a kernel function satisfying the Mercer theory and is employed to map the input data into a high dimensional feature space. Therefore, the non-linear discriminant model for sample data can be obtained as

$$f(x) = sign(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b) \tag{17}$$

4.3. GA-NN predictors.

Genetic algorithm-neural network (GA-NN) predictors in this paper are designed to provide one-to-one prediction for the water content in the low, middle and high water holdup patterns. Generally, multi-layer BP algorithm based on the gradient descent of computation error function is traditionally used for neural network (NN) learning, but still has some defects such as large computation scale, slow learning speed, over-fitting and blind initialization of structure parameters. Genetic algorithm (GA) is an adaptive searching optimization method with characteristics of global convergence. The combination of NN and GA is effective to obtain higher generalization ability, better global convergence and swifter learning ability than that of single NN model [10, 13]. The learning algorithm of GA-NN predictors is established by:

(1) Set the NN structure as m_1-m_2-m_3 for the number of neurons in the input, hidden and output layer, respectively, and tansig and purelinear function are used as transfer functions for the hidden and the output layer, respectively.

(2) Create original population with real number coding, and use it as weights and thresholds of NN. The population is built by N chromosome strings in length of $S=m_1m_2+m_2m_3+m_2+m_3$ for each. The weights and thresholds of NN are coded by

$$w_i^{(k)} = (w_{i1}^{(k)},\ w_{i2}^{(k)}, \cdots,\ w_{iS}^{(k)}) \tag{18}$$

here i=1, 2, …, N, k=0, 1, …, k_{max}, and thus the coding matrix for the chromosome population is indicated as

$$w(k)=[w_1^{(k)}\ \ w_2^{(k)}\ \ \cdots\ \ w_N^{(k)}]^T=\begin{bmatrix} w_{11}^{(k)} & w_{12}^{(k)} & \cdots & w_{1S}^{(k)} \\ w_{21}^{(k)} & w_{22}^{(k)} & \cdots & w_{2S}^{(k)} \\ \vdots & \vdots & \vdots & \vdots \\ w_{N1}^{(k)} & w_{N2}^{(k)} & \cdots & w_{NS}^{(k)} \end{bmatrix} \tag{19}$$

here k is the evolving generation number, $k_{\max}$ is the final generation number, N is population scale.

(3) Allocate the genes of chromosome as the parameters of NN, and run the NN to calculate the sum of square error (SSE) between the output $\bar{y}_l$ and desired output y_l, and the fitness function for GA is given by

$$f(w_i^{(k)}) = \frac{1}{E(w_i^{(k)})} = \frac{1}{\sum_{l=1}^{m}\left(\bar{y}_l^{(k)} - y_l^{(k)}\right)^2} \tag{20}$$

here m is the number of training samples.

(4) Sort the individual chromosome in terms of fitness calculated, and give the selection probability of individual chromosome using the method of roulette wheel by

$$q_i^{(k)} = f(w_i^{(k)}) \Big/ \sum_{i=1}^{N} f\left(w_i^{(k)}\right) \tag{21}$$

Crossover operation of individual chromosome employs the non-uniform across operator, here let $w_i^{(k)}$ and $w_j^{(k)}$ be two parent individuals in generation k, the two new offspring individuals $x_i^{(k+1)}$ and $x_j^{(k+1)}$ are produced by

$$\begin{cases} w_i^{(k+1)} = \alpha \times w_i^{(k)} + (1-\alpha) \times w_j^{(k)} \\ w_j^{(k+1)} = (1-\alpha) \times w_i^{(k)} + \alpha \times w_j^{(k)} \end{cases} \tag{22}$$

here $\alpha \in (0, 1)$ is a uniform distributed random variable.

The mutation operation adopts the method of non-uniform mutation defined by

$$w_i^{'} = \begin{cases} w_i + \Delta\ (k,\ k_{\max},\ U_{\max}^i - v_i),\ \text{rand}(0,1)=0 \\ w_i + \Delta\ (k,\ k_{\max},\ v_i - U_{\min}^i),\ \text{rand}(0,1)=1 \end{cases} \tag{23}$$

here $w_i^{'}$ is the new individual of w_i after mutation, $[U_{\min}^i\ \ U_{\max}^i]$ is the value range of w_i, $\Delta(k, k_{\max}, d)$ generate a non-uniform distributed random number from $[0, d]$.

(5) The optimal individual used as original weights for NN, and the best weights obtained through further optimization by using Levenberg-Marquardt (LM) `algorithm`. The weights and thresholds of NN are adjusted by

$$w^{(k+1)} = w^{(k)} - [J^T(w^{(k)})J(w^{(k)}) + \mu I]^{-1} J^T(w^{(k)})e^{(k)} \tag{24}$$

where μ is the correction factor, e is the learning error of NN, and $J(w) = \partial E / \partial W$ is the Jocabian matrix.

5. SIMULATION RESULTS

To investigate the measurement of water-content-in-oil effectively, an experimental measurement system in the laboratory was established. Some samples of oil-water mixture with different water content, temperature and salinity content were tested, and the measured data were collected for establishing the compound model. In order to make the prediction result more accurate and reliable, training and predicting data both cover the full range of the water holdup. Measured data of 46 groups from low-water-content-pattern, 54 groups from middle-water-content-pattern and 39 groups from high-water-content-pattern were employed to construct the compound model. Here we take penalty gene γ=1000 and the quadratic polynomial as kernel function.

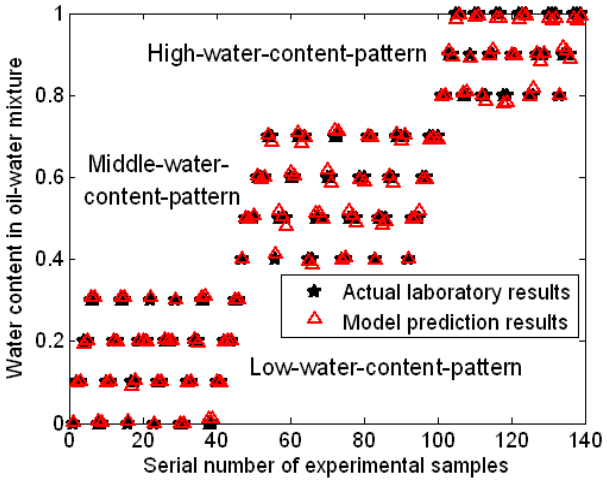


Fig.10. Model prediction results in low, middle and high-water-content-pattern.

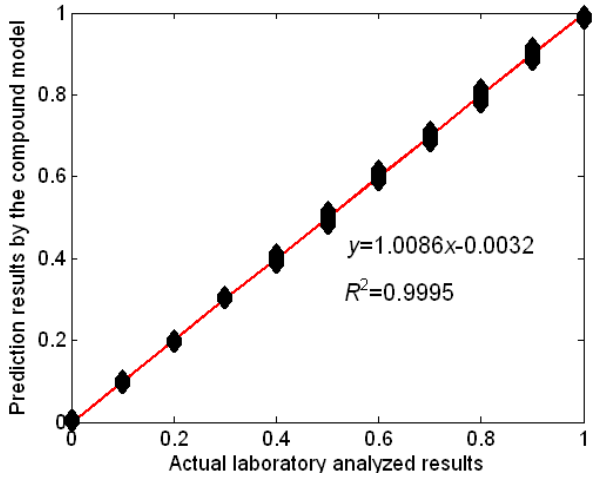


Fig.11. Results comparison between prediction results and actual laboratory analyzed results.

The effective measurement results are verified through the combination of RS-SVM classifier and GA-NN predictor. Tri-layer network structure including input layer, hidden layer and output layer was adopted for GA-NN predictors in this paper. The learning rate was set to η=0.3, momentum factor to μ=0.15, training error to E_{max}=0.0025, maximum training epochs to t_{max}=20000. Compared with the actual laboratory analyzed results, the model prediction results in low, middle and high-water-content-pattern are shown in Fig.10., indicating that this model is effective, and the precision might meet the demands for oil-water two-phase flow measurements in a wide measuring range. The result comparison between prediction results and actual laboratory analyzed results is illustrated in Fig.11., the regression equation can be represented as y=1.0086x−0.0032, and the regression coefficient, R^2, is 0.9995, indicating a highly predicting accuracy. The model error between the prediction results and actual laboratory analyzed results is plotted in Fig.12. The compound model gets mean absolute error (MAE) 0.76 % and maximum absolute error (AE_{max}) 2.3 % in the measuring range 0~100 %, which is much better than that of back-propagation neural network, it overcomes the defects of the present measuring instruments and methods such as low accuracy, and narrow measurement range.

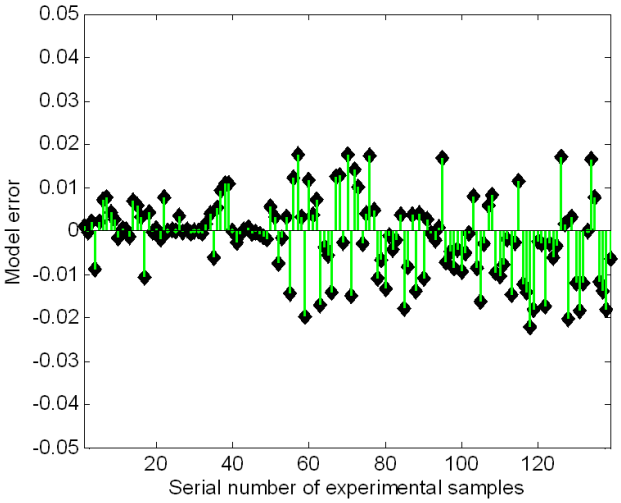


Fig.12. Model error between the prediction results and actual laboratory analyzed results.

6. CONCLUSIONS

The measurement of water-content-in-oil based on the method of dielectric coefficient suffers under the influence of temperature, salinity content and non-linear characteristics of sensor. To improve the measuring accuracy in a wide measuring range, an experimental system in the laboratory is established, the relations and influencing rules between multi-factors are investigated by numerical simulation combining with experimental test. Subsequently, a compound model based on RS-SVM classifier and GA-NN predictors is presented to improve the prediction precision in all-round measuring range of water content. Experimental results show that RS-SVM classifier can effectively realize the pattern identification for water holdup states via fuzzy reasoning and self-learning, and GA-NN predictors are capable of subsection forecasting water content in the different water holdup patterns, as well as adjusting the model parameters adaptively in terms of online

measuring range. Compared with the actual laboratory analyzed results, the experiment results show that this method is effective to eliminate the cross-influence of the multi-factors, and the precision might meet the demands for water content measurements in a wide measuring range.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant No. 51407200), the Promotive Research Foundation for the Excellent Middle-Aged and Youth Scientists of Shandong Province of China (Grant No. BS2012DX044), the Fundamental Research Funds for the Central Universities of China (Grant No. 12CX04065A), and the Science and Technology Development Plan Project of Qingdao (Grant No. 13-1-4-179-jch).

REFERENCES

[1] Thorn, R., Johansen, G.A., Hjertaker, B.T. (2013). Three-phase flow measurement in the petroleum industry. *Measurement Science and Technology*, 24, 012003.

[2] Tan, C., Dong, F. (2010). Modification to mass flow rate correlation in oil-water two-phase flow by a V-cone flow meter in consideration of the oil-water viscosity ratio. *Measurement Science and Technology*, 21, 045403.

[3] Li, Y., Yang, W., Xie, C., Huang, S., Wu, Z., Tsamakis, D., Lenn, C. (2013). Gas/oil/water flow measurement by electrical capacitance tomography. *Measurement Science and Technology*, 24, 074001.

[4] Assaad, M., Aslam, M.Z. (2012). An interface circuit design based on differential capacitive sensors for accurate measurement of water contents in crude oil. In *IEEE International Conference on Circuits and Systems (ICCAS)*. IEEE, 263-266.

[5] Fortuny, M., Oliveira, C.B.Z., Melo, R.L.F.V., Nele, M., Coutinho, R.C.C., Santos, A.F. (2007). Effect of salinity, temperature, water content, and pH on the microwave demulsification of crude oil emulsions. *Energy Fuels*, 21, 1358-1364.

[6] Makeev, Y.V., Lifanov, A.P., Sovlukov, A.S. (2013). Microwave measurement of water content in flowing crude oil. *Automation and Remote Control*, 74, 157-169.

[7] Araujo, A.M., Santos, L.M., Fortuny, M., Melo, R.L.F.V., Coutinho, R.C.C., Santos, A.F. (2008). Evaluation of water content and average droplet size in water-in-crude oil emulsions by means of near-infrared spectroscopy. *Energy Fuels*, 22, 3450-3458.

[8] Tripathi, M.M., Hassan, E.B.M., Yueh, F.Y., Singh, J.P., Steele, P.H., Ingram, L.L. (2009). Reflection-absorption-based near infrared spectroscopy for predicting water content in bio-oil. *Sensors & Actuators*, *B*: *Chemical*, 136 (1), 20-25.

[9] Mohamed, A.M.O., Gamal, M.E., Zekri, A.Y. (2003). Effect of salinity and temperature on water cut determination in oil reservoirs. *Journal of Petroleum Science* and *Engineering*, 40, 177-188.

[10] Xia, B., Bo, Y. (2010). A soft sensor for water-content-in-oil based on GA-RBF neural network. *Measurement and Control*, 43, 179-182.

[11] Chen, X., Zhu, W., Zhao, J., Ren, L. (2010). Method of predicting water content in crude oil based on measuring range automatic switching. *Journal* of *Beijing Institute of Technology (English Edition)*, 19, 87-91.

[12] Zhao, X., Jin, N., Li, W. (2005). Soft measurement method of phase volume fraction for oil/water two-phase flow. *Journal of Chemical Industry and Engineering* (*China*), 56, 1875-1879.

[13] Zhang, D., Hu, G. (2011). Measurement of oil-water flow based on inverse model of wavelet neural network with genetic optimization. *Optics and Precision Engineering*, 19, 1588-1595.

[14] Li, N., Liu, C. (2011). Application of SVM to the prediction of water content in crude oil. In *International Conference on Control, Automation and Systems Engineering (CASE)*. IEEE, 5997528.

[15] Li, W., Zhang, S., He, G. (2013). Semisupervised distance-preserving self-organizing map for machine-defect detection and classification. *IEEE Transactions on Instrumentation and Measurement*, 62, 869-879.

[16] Campbell, C., Ying, Y. (2011). *Learning with Support Vector Machines*. Morgan and Claypool Publishers.

[17] Steinwart, I., Christmann, A. (2008). *Support Vector Machines*. Springer.

[18] Zhang, C., Zhang, T., Yuan, C. (2011). Oil holdup prediction of oil-water two-phase flow using thermal method based on multiwavelet transform and least squares support vector machine. *Expert Systems with Applications*, 38, 1602-1610.

Received October 03, 2013.
Accepted August 14, 2014.