

Spatial-Temporal Separation Based on the Dynamic Recurrent Wavelet Neural Network Modelling for ASP Flooding

Shurong Li^{1,*}, Yulei Ge²

¹Automation School, Beijing University of Posts and Telecommunications, Beijing, China

²College of Information and Control Engineering, China University of Petroleum (East China), Qingdao, China

Email address:

lishurong@bupt.edu.cn (Shurong Li)

*Corresponding author

To cite this article:

Shurong Li, Yulei Ge. Spatial-Temporal Separation Based on the Dynamic Recurrent Wavelet Neural Network Modelling for ASP Flooding. *American Journal of Applied Mathematics*. Vol. 5, No. 6, 2017, pp. 154-167. doi: 10.11648/j.ajam.20170506.12

Received: December 13, 2017; **Accepted:** December 27, 2017; **Published:** January 10, 2018

Abstract: In this paper, a three-dimensional spatial-temporal decomposition modelling method is proposed to build the alkali-surfactant-polymer (ASP) flooding model, in which a new dynamic recurrent wavelet neural network (DRWNN) is presented to identify the temporal coefficients. At first, the detailed mathematical model of ASP flooding is described which is a complex distributed parameter system. Then a three-dimensional spatial-temporal modelling method is inferred based on Karhunen-Loeve (K-L) decomposition to decompose the water saturation of reservoir into a series of spatial basis functions and corresponding temporal coefficients. Furthermore, the recurrent wavelet neural network is used to acquire the identification model, in which the injection concentrations of ASP flooding and temporal coefficients are taken as the input and output information. In order to improve the capability of dynamic modelling, DRWNN is proposed through adding feedback layers and setting the different weights with time to achieve dynamic memory of the past information. Considering the gradient descent method for the neural networks training easily leads to local minimum and slow convergence speed, the spectral conjugate gradient method is introduced to optimize the weights of DRWNN. At last, DRWNN is used to build the relation between the moisture content of production wells and the water saturation of the corresponding grids. Thus, the final approximate model of ASP flooding is finished. The accuracy is proved by model with four injection wells and nine production wells through data from the mechanism model.

Keywords: ASP Flooding, Karhunen-Loeve Decomposition, Dynamic Recurrent Wavelet Neural Network, Spectral Conjugate Gradient Method

1. Introduction

With the old oil fields entering the later period of development, moisture content of reservoir is increasing, and the oil production is reducing [1]. How to update the technical means to ensure oil recovery is one of the most important measures to stabilize oil production. ASP flooding is an important tertiary oil recovery technique which is widely studied [2]. It can enhance the oil production evidently by use of the interaction among alkali, surfactant and polymer to improve the physicochemical property of reservoir. However, there is lack of a uniform mathematical model description because of the uncertainty of alkali reaction. It is important to

build an accurate and easily to be applied model in the research of ASP flooding.

In addition, since the mechanism model of ASP flooding is a complexed distributed parameter system, it is hard to carry out optimal control strategies for the oil production, because of the features of infinite dimensions, spatial-temporal coupling and complex nonlinear behavior. The inputs of ASP flooding are the injection concentrations of alkali, surfactant and polymer; the output is the water cut of production wells. The state variables of ASP flooding contain water saturation, pressure and grid concentrations. In applications, different methods are used to search the injection concentrations of ASP flooding to get the best performance index. One of

common ways is to solve the mechanism equations of ASP flooding, such as fluid equations and seepage equations, etc. [3] However, the method based on mechanism model usually involves a lot of math operations, and the process of mathematic treatment always needs a fair amount of calculation. Though many people have done researches on modelling of ASP flooding, nearly all the works are about improving or enriching the primary model. The main work they have done is to consider more influencing factors or inductive researches; the model is still complex and difficult to be applied. So it is important to consider using model identification method to approximate the ASP flooding system so as to simplify the mathematical operation process, and it will be helpful to improve the computational efficiency and reduce the computational complexity of the algorithm [4].

Traditional system identification methods do not consider spatial information [5]. To model the spatial-temporal dynamics, spatial information is a very important part of system. Spatial-temporal decomposition technology [6, 7], which comes from Fourier series expansion, is a very useful modelling method for distributed parameter systems. It can reflect the information in time domain and space domain of the system by decomposition. A spatiotemporal variable can be expanded into an infinite number of spatial basis functions and corresponding temporal coefficients. Generally speaking, the first few primary spatial basis functions can reflect the maximal inner information of the system, which provides a good approximation because of their separation properties [8]. In this way, the spatial-temporal method can obtain a finite-dimensional model.

An important condition of modelling is to guarantee the accuracy of model which is highly dependent on the choice of spatial basis functions. In particular, Karhunen-Loeve (K-L) decomposition, which is called proper orthogonal decomposition and principal component analysis [9, 10], is a popular spatial-temporal decomposition approach to find principal spatial structures and reduce the dimension of the data. Among all linear expansions, K-L expansion is the most efficient in the sense that, for a given approximation error, the number of K-L bases required is minimal. As a result, K-L decomposition can help to reduce the model dimension and the number of estimated parameters.

Once the spatial basis functions are designed properly, the corresponding states can be determined by projecting the spatiotemporal data onto these spatial basis functions. To model the input-state dynamics, the relationship between inputs and temporal coefficients should be identified. In the process of modelling, the dynamic recurrent neural network is a good way for the system with unknown nonlinearity. The recurrent wavelet neural network (RWNN) [11, 22] combining wavelet transform with dynamic neural network can get better modelling capability. It is widely used in nonlinear dynamical modelling problems because of the advantages of neural networks such as reliable theory basis, explicit practical sense, simple algorithm realization and strong adaptability, etc.

In order to improve the dynamic modelling capability of

RWNN, a new dynamic recurrent wavelet neural network (DRWNN) is proposed, in which the feedback layers are added for the network and different weights are set for feedback layers which is decreasing with time to make the network have the dynamic memory function. However, the gradient descent method is commonly adopted to train the parameters of network, which may cause the local minimum solutions and slow convergence speed. Especially in large search space for multimodal function, the function usually can't find out the global optimal value [12].

Many researchers have studied the optimization methods, which can be divided into five classes in general: gradient descent methods, Newton methods, conjugate gradient methods, heuristic optimization methods and Lagrange multiplier methods [13]. But the gradient descent method has slow converge speed, the Newton method needs to storage and compute the Hessian matrix which has huge computing complexity, the results of heuristic optimization method has a certain uncertainty, and the Lagrange multiplier method is often used to deal with multi-objective optimization problems. While the conjugate gradient method, which balances the convergence speed and computing complexity, has good stability and with no need of any extraneous parameter. It is widely used in nonlinear optimization. PRP method is viewed as one of the most effective conjugate gradient (CG) methods, but the convergence property is not so good [14]. Reference [15] gave an improved PRP method which can also be named as NPRP method. By modifying the step-length parameter, the sufficient descent property and global convergence are guaranteed on the condition of strong wolf line search. To improve the performance of numerical methods further, on the basis of [15] a new spectral NPRP (SNPRP) conjugate gradient method was presented through adding the spectral parameter and revising the step-length [16]. The author also gave the proof about global convergence and fast descent. On account of these advantages, the SNPRP conjugate gradient method is adopted to train the weights of DRWNN.

In this paper, a new model of ASP flooding is obtained in terms of spatial-temporal decomposition. K-L decomposition is used to model the state parameters of reservoir (water saturation) into a series of temporal coefficients and spatial basis functions. To get the relation between the injection concentrations of ASP flooding and temporal coefficients, the DRWNN is proposed in which different weights decreasing with time are added for feedback layers to improve the dynamic modelling capability and the weights are trained by the SNPRP conjugate gradient method. Furthermore, build the model between the grid water saturation and the moisture content of production wells with DRWNN. Then the final model of ASP flooding is finished with the injection concentrations as the inputs and the moisture content as the outputs.

This article is organized as follows: The mechanism model description of ASP flooding is described in section 2. In section 3, the Karhunen-Loeve decomposition of three-dimensional state of ASP flooding is introduced. In section 4, a new dynamic recurrent wavelet neural network,

named DRWNN, is proposed to identify the relationships in the modelling. Besides that, in order to improve the ability of DRWNN, the SNPRP conjugate gradient method is adopted to train the weights. Section 5 gives an illustrative simulation example. Finally, a few conclusions are presented in section 6.

2. Mechanism Model Description of ASP Flooding

Suppose that the region of the reservoir is $(x, y, z) \in \Omega$, $\Omega \in R^3$, the main five components of ASP flooding are oil, water, polymer, surfactant, alkali. The mathematical model of ASP flooding can be described as: [19, 20]

$$\begin{aligned} & \nabla \cdot \left(\frac{Kk_{rw}c_p}{B_w R_k \mu_w} \nabla (p_w - \rho_w g h) \right) + \nabla \cdot \left[(D_w + D_{wp}) \frac{\phi_p S_w}{B_w} \nabla c_p \right] \\ & + q_c = \frac{\partial}{\partial t} \left(\rho_r (1 - \phi) C_{rp} + \frac{\phi_p S_w c_p}{B_w} \right). \end{aligned} \quad (3)$$

The flow equation for alkali phase:

$$\begin{aligned} & \nabla \cdot \left(D_{OH} \frac{\phi S_w}{B_w} \nabla c_{OH} \right) + \nabla \cdot \left(\frac{Kk_{rw}c_{OH}}{B_w R_k \mu_w} \nabla (p_w - \rho_w g h) \right) \\ & + R_{OH} + q_e = \frac{\partial}{\partial t} \left[\frac{K_a}{(1 + K_b c_{OH})^2} (1 - \phi) S_w c_{OH} + \frac{\phi S_w c_{OH}}{B_w} \right]. \end{aligned} \quad (4)$$

The flow equation for surfactant phase:

$$\begin{aligned} & \nabla \cdot \left(\frac{Kk_{rw}c_{ws}}{B_w R_k \mu_w} \nabla (p_w - \rho_w g h) \right) + \nabla \cdot \left[(D_o + D_{os}) \frac{\phi_s S_o}{B_o} \nabla c_{os} \right] \\ & + \nabla \cdot \left[(D_w + D_{ws}) \frac{\phi_s S_w}{B_w} \nabla c_{ws} \right] + \nabla \cdot \left(\frac{Kk_{ro}c_{os}}{B_o R_k \mu_o} \nabla (p_o - \rho_o g h) \right) \\ & + q_d = \frac{\partial}{\partial t} \left(\frac{\phi_s S_o c_{os}}{B_o} + \frac{\phi_s S_w c_{ws}}{B_w} + \rho_r (1 - \phi) C_{rs} \right). \end{aligned} \quad (5)$$

The initial conditions:

$$\begin{aligned} p(x, y, z, t)|_{t=0} &= p^0, S_w(x, y, z, t)|_{t=0} = S_w^0, \\ c_{\Theta}(x, y, z, t)|_{t=0} &= c_{\Theta}^0, (x, y, z) \in \Omega. \end{aligned} \quad (6)$$

The boundary conditions:

$$\left. \frac{\partial p}{\partial n} \right|_{\partial \Omega} = 0, \quad \left. \frac{\partial S_w}{\partial n} \right|_{\partial \Omega} = 0, \quad \left. \frac{\partial c_{\Theta}}{\partial n} \right|_{\partial \Omega} = 0. \quad (7)$$

Polymer can improve the viscosity of pore water, so it can be represented as:

$$\mu_w = \mu_{w0} \left[1 + (ap_1 C_p + ap_2 C_p^2 + ap_3 C_p^3) C_{sep}^{sp} \right], \quad (8)$$

where μ_{w0} is the viscosity of pure water; ap_1, ap_2, ap_3, sp

The flow equation for oil phase:

$$\nabla \cdot \left[\frac{Kk_{ro}}{B_o \mu_o} \nabla (p_o - \rho_o g h) \right] + q_o = \frac{\partial}{\partial t} \left[\frac{\phi(1 - S_w)}{B_o} \right]. \quad (1)$$

The flow equation for water phase:

$$\nabla \cdot \left(\frac{Kk_{rw}}{B_w R_k \mu_w} \nabla (p_w - \rho_w g h) \right) + q_w = \frac{\partial}{\partial t} \left(\frac{\phi S_w}{B_w} \right). \quad (2)$$

The flow equation for polymer phase:

are empirical constants; C_{sep} is the salinity of reservoir.

Alkali can react with substances in the reservoir to produce new substances and cause alkali consumption. This process also influences oil-water interfacial tension, and the concrete mathematical description is

$$\sigma_{wo} = \sigma_{wo}(C_s) A_{st}(C_a), \quad (9)$$

where $\sigma_{wo}(C_s)$ is oil-water interfacial tension when oil displacement agent is only surfactant.

The different oil displacement agents can affect the strength of the rock adsorption and can be described as:

$$\Gamma_{rl} = \frac{a_l C_l}{1 + b_l C_l}, (l = a, s, p), \quad (10)$$

where a_{ll}, b are adsorption parameters.

Because the alkali will lower adsorption of polymer and surfactant on the surface of the pores of rock, the adsorption should be recalculated when alkali is injected to reservoir, and the formula for calculation can be described as:

$$\Gamma_{rl} = C_l A_{ad} (C_a), (l = s, p), \tag{11}$$

where $A_{ad} (C_a)$ is adsorption coefficient.

The other parameters and model explanations can be found in the appendix.

3. The Three-Dimensional Spatial-Temporal Decomposition Modelling Method for ASP Flooding

ASP flooding system is a complex distributed parameter system which is very difficult to be modeled in general method. In this paper, a new three-dimensional spatial-temporal model method is proposed based on Karhunen-Loeve decomposition method [8]. During the modelling process for ASP flooding, the main problem is to identify an appropriate spatial-temporal model according to the input, output and system state data. The whole model for space/time separation of ASP flooding can be divided into three stages:

- (1) Decompose the states of ASP flooding into spatial basis functions and temporal coefficients;
- (2) Identify the relationship between the injection concentrations and the temporal coefficients;
- (3) Model the moisture content of production wells and corresponding grid states.

3.1. Three-Dimensional Spatial-Temporal Method Based on KL Decomposition

For the system of ASP flooding, suppose that the output of system is $\{Y(x_i, y_j, z_k, t)\}_{i=1, j=1, k=1, t=1}^{N_x, N_y, N_z, L}$ which is sampled in time and space, where $t = 1, 2, \dots, L$ denotes L sample points of output in time domain, $i = 1, 2, \dots, N_x$, $j = 1, 2, \dots, N_y$, $k = 1, 2, \dots, N_z$ is $N_x \times N_y \times N_z$ sample points in space domain.

According to the theory of Fourier series, the spatial-temporal output $Y(x, y, z, t)$ can be decomposed into a series of orthonormal spatial basis functions and temporal coefficients. So the decomposition form can be expressed as:

$$Y(x, y, z, t) = \sum_{i=1}^{\infty} \varphi_i(x, y, z) T_i(t), \tag{12}$$

where $\{\varphi_i(x, y, z)\}_{i=1}^{\infty}$ represents the orthonormal spatial basis functions and $\{T_i(t)\}_{i=1}^{\infty}$ represents the temporal coefficients.

In engineering applications, ensuring the accuracy of the approximation, the model output is usually truncated to the finite form:

$$Y_n(x, y, z, t) = \sum_{i=1}^n \varphi_i(x, y, z) T_i(t), \tag{13}$$

where n denotes the order of model, it mainly depends on the system spatial basis functions which reflect the dynamic characteristic of system.

As the description above, it is clear that the main problem of modelling is to get the dominant spatial basis functions $\{\varphi_i(x, y, z)\}_{i=1}^n$ among the system output $\{Y(x_i, y_j, z_k, t)\}_{i=1, j=1, k=1, t=1}^{N_x, N_y, N_z, L}$. The problem can be converted into a minimization problem as the following objective function:

$$\begin{cases} \min_{\varphi(x,y)} \left\langle \left\| Y(x, y, z, t) - \sum_{i=1}^n \varphi_i(x, y, z) T_i(t) \right\|^2 \right\rangle, \\ \text{s.t. } (\varphi_i, \varphi_i) = 1, \varphi_i \in L^2(\Omega), i = 1, 2, \dots, n. \end{cases} \tag{14}$$

where $\|f(x, y, z)\| = (f(x, y, z), f(x, y, z))^{1/2}$ denotes norm,

$\langle f(x, y, z, t) \rangle = \frac{1}{L} \sum_{t=1}^L f(x, y, z, t)$ denotes the mean value of a set.

Because the spatial basis functions are unit orthogonal, they are submitted to the below equation:

$$\begin{aligned} &(\varphi_i(x, y, z), \varphi_j(x, y, z)) \\ &= \iiint_{\Omega} \varphi_i(x, y, z) \varphi_j(x, y, z) dx dy dz = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases} \end{aligned} \tag{15}$$

Thus, the temporal coefficients can be described as:

$$T_i(t) = (\varphi_i(x, y, z), Y(x, y, z, t)), i = 1, 2, \dots, n. \tag{16}$$

Considering the orthonormal constraint $(\varphi_i, \varphi_i) = 1$, through Lagrangian function construction method, the constrained optimization problem in equation (14) can be structured in the following form:

$$\begin{aligned} J = &\left\langle \left\| Y(x, y, z, t) - \sum_{i=1}^n \varphi_i(x, y, z) T_i(t) \right\|^2 \right\rangle \\ &+ \sum_{i=1}^n \lambda_i ((\varphi_i(x, y, z), \varphi_i(x, y, z)) - 1), \end{aligned} \tag{17}$$

Combine the known conditions and do the transformation to equation (17), then take the variation with respect to

$\varphi_i(x, y, z)$,

$$\delta J = \frac{1}{L} \sum_{t=1}^L \left[-2Y(x, y, z, t) \sum_{i=1}^n T_i(t) \right] \delta \varphi_i(x, y, z) + \sum_{i=1}^n 2\lambda_i \iiint_{\Omega} \varphi_i(x, y, z) \delta \varphi_i(x, y, z) dx dy dz. \quad (18)$$

The necessary condition of extreme value for this functional problem is $\delta J = 0$. Since $\varphi_i(x, y, z)$ can be an arbitrary function, the below equation can be developed.

$$\frac{1}{L} \sum_{t=1}^L \left[Y(x, y, z, t) \sum_{i=1}^n T_i(t) \right] = \sum_{i=1}^n \lambda_i \varphi_i(x, y, z). \quad (19)$$

Substitute $T_i(t)$ into (19), then

$$\sum_{i=1}^n \iiint_{\Omega} R((x, y, z), (\xi, \zeta, \varsigma)) \varphi_i(\xi, \zeta, \varsigma) d\xi d\zeta d\varsigma = \sum_{i=1}^n \lambda_i \varphi_i(x, y, z), \quad (20)$$

where $R((x, y, z), (\xi, \zeta, \varsigma)) = \langle Y(x, y, z, t) Y(\xi, \zeta, \varsigma, t) \rangle$ is the correlation function between two points in the space.

Considering the mutual independence of value for $i = 1, \dots, n$, to make sure that (20) is always equal to zero, if and only if (21) is true. This is the necessary condition of extremum for (14).

$$\begin{cases} \iiint_{\Omega} R((x, y, z), (\xi, \zeta, \varsigma)) \varphi_i(\xi, \zeta, \varsigma) d\xi d\zeta d\varsigma = \lambda_i \varphi_i(x, y, z), \\ (\varphi_i, \varphi_i) = 1, i = 1, 2, \dots, n. \end{cases} \quad (21)$$

3.2. Snapshots Method

In this way, the original optimization problem is transformed to the solution of integral equation (21). However, in normal conditions, this equation is hard to solve. When the number of time samples is less than space samples, the snapshots method [17] is a good way to solve (21). This method can significantly reduce calculation quantities. The snapshot is obtained from sampling for each position point in discrete time. For an example, $Y(x, y, z, 3)$, $x = 1, 2, \dots, N_x$, $y = 1, 2, \dots, N_y$, $z = 1, 2, \dots, N_z$ and $N = N_x \cdot N_y \cdot N_z$ denotes one snapshot at $t = 3$.

When time nodes are less than space nodes, suppose the spatial basis function $\varphi_i(x, y, z)$ can be expressed as a linear combination of snapshots as follows:

$$\varphi_i(x, y, z) = \sum_{t=1}^L \gamma_{it} Y(x, y, z, t). \quad (22)$$

For eigenvalue problem of equation (21), below equation

can be got.

$$\begin{aligned} \iiint_{\Omega} R((x, y, z), (\xi, \zeta, \varsigma)) \varphi_i(\xi, \zeta, \varsigma) d\xi d\zeta d\varsigma &= \lambda_i \varphi_i(x, y, z) \\ &= \iiint_{\Omega} \frac{1}{L} (Y(x, y, z, 1), Y(x, y, z, 2), \dots, Y(x, y, z, L)) \\ &\quad \times \begin{pmatrix} Y(\xi, \zeta, \varsigma, 1) \\ Y(\xi, \zeta, \varsigma, 2) \\ \dots \\ Y(\xi, \zeta, \varsigma, L) \end{pmatrix} \varphi_i(\xi, \zeta, \varsigma) d\xi d\zeta d\varsigma \\ &= \lambda_i \varphi_i(x, y, z). \end{aligned} \quad (23)$$

Substitute (22) into (23), at the same time, and let $\bar{Y}(x, y, z, t)^T = (Y(x, y, z, 1), Y(x, y, z, 2), \dots, Y(x, y, z, L))$, $\bar{Y}((\xi, \zeta, \varsigma, t)) = (Y(\xi, \zeta, \varsigma, 1), Y(\xi, \zeta, \varsigma, 2), \dots, Y(\xi, \zeta, \varsigma, L))^T$, $\gamma_i = (\gamma_{i1}, \gamma_{i2}, \dots, \gamma_{iL})^T$, we can get:

$$\begin{aligned} \iiint_{\Omega} \frac{1}{L} \bar{Y}(x, y, z, t)^T \bar{Y}(\xi, \zeta, \varsigma, t) \begin{pmatrix} Y(\xi, \zeta, \varsigma, 1) \\ Y(\xi, \zeta, \varsigma, 2) \\ \dots \\ Y(\xi, \zeta, \varsigma, L) \end{pmatrix} \begin{pmatrix} \gamma_{i1} \\ \gamma_{i2} \\ \dots \\ \gamma_{iL} \end{pmatrix} d\xi d\zeta d\varsigma \\ = \iiint_{\Omega} \frac{1}{L} \bar{Y}(x, y, z, t)^T \bar{Y}(\xi, \zeta, \varsigma, t) \bar{Y}(\xi, \zeta, \varsigma, k)^T \gamma_i d\xi d\zeta d\varsigma \\ = \lambda_i \bar{Y}(x, y, z, t)^T \gamma_i. \end{aligned} \quad (24)$$

Define $C = \frac{1}{L} \iiint_{\Omega} \bar{Y}(\xi, \zeta, \varsigma, t) \bar{Y}(\xi, \zeta, \varsigma, k)^T d\xi d\zeta d\varsigma$, so

$$\bar{Y}(x, y, z, t)^T (C \gamma_i - \lambda_i \gamma_i) = 0 \quad (25)$$

In this way, $N \times N$ eigenvalue problem in equation (21) is converted to the $L \times L$ problem:

$$C \gamma_i = \lambda_i \gamma_i, \quad (26)$$

where γ_i is the i th eigenvector, and λ_i is the corresponding eigenvalue. The matrix C can be obtained from the output samples, so the spatial basis functions can be computed through equation (26). Because C is symmetric and positive semidefinite, the computed eigenfunctions are orthogonal.

The dimension of the model largely determines the modelling accuracy and complexity. When deciding the number of the dimension n , the eigenvalues are arranged in descending order. The eigenfunctions corresponding to the first several eigenvalues are the most advanced representative in the system output. So the proper number of spatial basis functions can be selected. Generally speaking, the larger the number of dimension n is, the more accurate the accuracy of modelling will be. However, the modelling complexity will increase with the increasing model order. As long as the proper order of model is chose, the modelling complexity and

accuracy can be guaranteed, and the ideal mathematical model would be achieved.

Define that the total energy of system is equal to the sum of all eigenvalues. Then the bigger the value is, the more energy the basis function reflects. For every eigenvalue λ_i , the proportion that the energy it reflects account for the total energy is

$$\hat{E}_i = \frac{\lambda_i}{\sum_{j=1}^K \lambda_j} \tag{27}$$

Generally speaking, when the proportion is more than 99% , we think these bases can reflect most energy, the corresponding n is the needed dimension. Then the spatial basis functions and temporal coefficients are obtained.

4. Dynamic Recurrent Wavelet Neural Network Modelling Based on SNPRP Conjugate Gradient Method

In order to establish the relationship between the temporal coefficients $T(t)$ and the input $u(t)$, suppose that the dynamic between $T(t)$ and $u(t)$ can be expressed by a nonlinear autoregressive with exogenous input (NARX) model [18]:

$$\hat{T}_i(t) = \hat{F}_i[\hat{T}_i(t-1), \dots, \hat{T}_i(t-n_T), u(t-1), \dots, u(t-n_u)] \tag{28}$$

where n_u and n_T are the maximum input and output lags, respectively. The modelling process of NARX model which is black-box model doesn't need to know its internal mechanism, and the mathematical model can be established based on its input and output data. Because the neural network has good nonlinear approximation capability, powerful operation ability, strong fault tolerance and robustness, so it is widely used in the field of nonlinear system identification.

Dynamic neural networks can directly reflect the system dynamic characteristic, and is widely used in the dynamic system modelling. To improve the capability of dynamic modelling, in this paper, the feedback layers for dynamic recurrent wavelet neural network is added and different weights for them in chronological order are set. A new algorithm which called dynamic memory feedback wavelet neural network (DRWNN) based on the above-mentioned idea is proposed. Considering the weakness of the traditional gradient descent method for the neural networks training, the SNPRP conjugate gradient method is used to optimize the weights of DRWNN.

4.1. The Network Structure

Considering the conventional dynamic neural networks can only memory the output of last hidden layers, by adding the multi-layer of feedback to DRWNN, it can memory several dynamic feedback of hidden layers. In the structure of DRWNN, the number of feedback layers reflects the memory of the historical data, as shown in Figure 1.

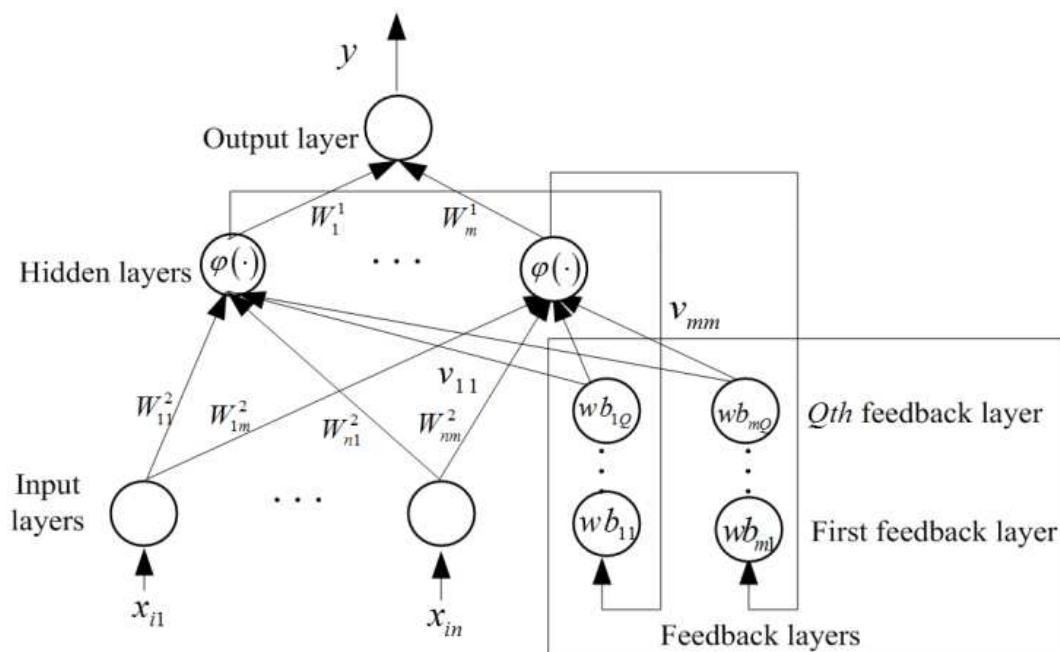


Figure 1. The structure chart of DRWNN.

The structure of DRWNN consists of four layers: the input layer, the hidden layer, the output layer and the feedback layer. The network contains n input neurons, an output neurons,

m hidden layer neurons, Q feedback layers. $x_i(t) \in R^n$ is the input with n dimensions for the time t ; $y(t) \in R$ is the

output; $H_i(t) \in R^m$ is the output of hidden layers; $x_c(t) \in R^n$ is the output of feedback layers.

Through adding Q feedback layers to the hidden layers to dynamically memory the output of hidden layers. Suppose the data of the q th feedback layer is $H(k-q)$, and its weight is $Wb_q = [wb_{1q}, wb_{2q}, \dots, wb_{mq}]$, $wb \in (0,1)$. In order to make the network achieve dynamic oblivion function, wb_{iq} is set as wb_i^q . Because of the weights ranging from 0 to 1, wb_i^q will be smaller with time which can gradually forget further information. Suppose α is the feedback gain, the output of feedback layer can be computed by the following formula:

$$x_{ci}(t) = \alpha \sum_{q=1}^Q (wb_{iq} \cdot H_i(t-q)). \quad (29)$$

The dynamic equations of DRWNN can be expressed as:

$$\begin{cases} y(t) = \sum_{i=1}^m W_i^1 H_i(t), \\ H_i(t) = \varphi\left(\frac{h_i(t) - b_i(t-1)}{a_i(t-1)}\right), \\ h_i(t) = \sum_{j=1}^n W_{ij}^2 x_j(t-1) + \alpha \sum_{k=1}^m \sum_{q=1}^Q v_{ik} wb_{iq} H_k(t-q), \end{cases} \quad (30)$$

where W_i^1 is the linked weight between the hidden layer of neuron i and output layer of neuron; W_{ij}^2 is the linked weight between the input layer of neuron j and the hidden layer of neuron i ; v_{ik} is the linked weight between the feedback layer k and the hidden layer of neuron i ; $H_i(t)$ is the output of the hidden layer of neuron i . $\varphi(\cdot)$ is Morlet wavelet function which can be described as:

$$\varphi(t') = \cos(1.75t') e^{-t'^2/2}, \quad (31)$$

where $t' = \frac{h_i(t) - b_i(t-1)}{a_i(t-1)}$, a_i is dilatation coefficient and

b_i is translation coefficient.

4.2. The SNPRP Conjugate Gradient Method

The SNPRP conjugate gradient method is a new conjugate gradient method which is presented in [16]. It has the sufficient descent property and global convergence property under the condition of strong line search. Furthermore, it needs less stored information which ensures the high computational efficiency. Then a brief introduction to this method is given as follows.

Considering a general unconstrained optimization problem

$\min\{f(\mathbf{x}) | \mathbf{x} \in R^n\}$, where $f: R^n \rightarrow R$ is a one order continuously differentiable nonlinear function, and its gradient vector is $\mathbf{g}(\mathbf{x}) \triangleq \nabla f(\mathbf{x})$, $\mathbf{g}_k = \mathbf{g}(\mathbf{x}_k)$. For a general spectral conjugate gradient method [23], below equation is satisfied.

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \eta_k \mathbf{d}_k, \quad (32)$$

$$\mathbf{d}_k = \begin{cases} -\mathbf{g}_k, & k=1, \\ -\frac{1}{\delta_k} \mathbf{g}_k + \beta_k \mathbf{d}_{k-1}, & k \geq 2. \end{cases} \quad (33)$$

where \mathbf{d}_k is the search direction, β_k is the parameter, and η_k is the step-length factor. The spectral coefficient is defined as

$$\delta_k = \frac{s_{k-1}^T \mathbf{y}_{k-1}}{\|s_{k-1}\|^2}, \mathbf{y}_{k-1} = \mathbf{g}_k - \mathbf{g}_{k-1}. \quad (34)$$

In SNPRP conjugate gradient method, the parameter β_k is defined as

$$\beta_k^{\text{SNPRP}} = \frac{\delta_{k-1} \left(\|\mathbf{g}_k\|^2 - \frac{\|\mathbf{g}_k\| \|\mathbf{g}_{k-1}\| \mathbf{g}_k^T \mathbf{g}_{k-1}}{\|\mathbf{g}_{k-1}\|^2} \right)}{\delta_k \|\mathbf{g}_{k-1}\|^2}. \quad (35)$$

The corresponding convergence and stability has been proved in [16].

The algorithm process is as follows.

Give the $\mathbf{x}_1 \in R^n$, $0 \leq \varepsilon$, $\mathbf{d}_1 = -\mathbf{g}_1$, $k=1$;

If $\|\mathbf{g}_k\| \leq \varepsilon$, then end the whole calculation;

Compute the step-length factor η_k ;

Execute computation as (32), and let $\mathbf{g}_{k+1} = \mathbf{g}(\mathbf{x}_{k+1})$;

Compute the spectral parameter, β_k^{SNPRP} and \mathbf{d}_{k+1} as equations (34), (35) and (33) separately. Let $k = k+1$, then go to step 2).

Then the SNPRP conjugate gradient method is used to train the weights of DRWNN.

4.3. The Training Algorithm of DRWNN

Suppose the $y(t)$ and $y_e(t)$ denote the real output and the expected output at time step t , the error is

$$e(t) = y_e(t) - y(t). \quad (36)$$

The cost function is

$$E(t) = \frac{1}{2} (y_e(t) - y(t))^2 = \frac{1}{2} (e(t))^2, \quad (37)$$

The total error function from the time step 1 to N is

$$E = \sum_{t=1}^N \frac{1}{2} (y_e(t) - y(t))^2 = \sum_{t=1}^N \frac{1}{2} (e(t))^2. \quad (38)$$

The training goal of DRWNN is to adjust the parameters to decrease E gradually. From the chain rule of neural network on the basis of SNPRP conjugate method, the weights can be inferred as follows,

$$\begin{aligned} W_i^1(t) &= W_i^1(t-1) + \eta_{t-1}^1 \left\{ -\frac{1}{\delta_{t-1}^1} \frac{\partial E(t)}{\partial W_i^1(t-1)} + \beta_{t-1}^1 d_{t-2}^1 \right\} \\ &= W_i^1(t-1) + \eta_{t-1}^1 \left\{ -\frac{1}{\delta_{t-1}^1} e(t) H_i(t) + \beta_{t-1}^1 d_{t-2}^1 \right\}. \end{aligned} \quad (39)$$

Define $\Theta_i(t) = \{W_{ij}^2(t), v_{ik}(t), a_i(t), b_i(t)\}$, then

$$\begin{aligned} \Theta_i(t) &= \Theta_i(t-1) + \eta_{t-1}^s \left\{ -\frac{1}{\delta_{t-1}^s} \frac{\partial E(t)}{\partial \Theta_i(t-1)} + \beta_{t-1}^s d_{t-2}^s \right\} \\ &= \eta_{t-1}^s \left\{ -\frac{1}{\delta_{t-1}^s} e(t) W_i^1(t-1) \frac{\partial H_i(t)}{\partial \Theta_i(t-1)} + \beta_{t-1}^s d_{t-2}^s \right\} \\ &\quad + \Theta_i(t-1), \end{aligned} \quad (40)$$

where $i, k = 1, 2, \dots, m$, $j = 0, 1, \dots, n$, $s = 2, 3, 4, 5$ corresponds to the variables $W_{ij}^2(t), v_{ik}(t), a_i(t), b_i(t)$, η_{t-1}^s is the learning coefficients of W^1, W^2, v, a, b .

$$d_{t-1} = \begin{cases} -\frac{\partial E(t)}{\partial \Psi_i(t-1)}, & t = 1, 2, \\ -\frac{1}{\delta_{t-1}^s} \frac{\partial E(t)}{\partial \Psi_i(t-1)} + \beta_{t-1}^s d_{t-2}^s, & t \geq 3, \end{cases} \quad (41)$$

where $\Psi_i(t) = \{W_i^1(t), \Theta_i(t)\}$, $s = 1$ corresponds to the variable $W_i^1(t)$.

For the equations above,

$$\begin{aligned} \frac{\partial H_i(t)}{\partial W_{ij}^2(t-1)} &= \frac{\partial H_i(t)}{\partial h_i(t)} \\ &\cdot \left[x_j(t-1) + \alpha \sum_{k=1}^m \sum_{q=1}^Q v_{ik}(t-1) w_{b_{iq}} \frac{\partial H_k(t-q)}{\partial w_{ij}^2(t-q-1)} \right]. \end{aligned} \quad (42)$$

$$\begin{aligned} \frac{\partial H_i(t)}{\partial v_{ik}(t-1)} &= \frac{\partial H_i(t)}{\partial h_i(t)} \\ &\cdot \left[\alpha H_k(t-1) + \alpha \sum_{k=1}^m \sum_{q=1}^Q v_{ik}(t-1) w_{b_{iq}} \frac{\partial H_k(t-q)}{\partial v_{ik}(t-q-1)} \right]. \end{aligned} \quad (43)$$

$$\begin{aligned} \frac{\partial H_i(t)}{\partial a_i(t-1)} &= \varphi'(t') \\ &\cdot \left[\frac{\partial(t')}{\partial a_i(t-1)} + \alpha \sum_{k=1}^m \sum_{q=1}^Q v_{ik}(t-1) w_{b_{iq}} \frac{\partial H_k(t-q)}{\partial a_i(t-q-1)} \right]. \end{aligned} \quad (44)$$

$$\begin{aligned} \frac{\partial H_i(t)}{\partial b_i(t-1)} &= \varphi'(t') \\ &\cdot \left[\frac{\partial(t')}{\partial b_i(t-1)} + \alpha \sum_{k=1}^m \sum_{q=1}^Q v_{ik}(t-1) w_{b_{iq}} \frac{\partial H_k(t-q)}{\partial b_i(t-q-1)} \right]. \end{aligned} \quad (45)$$

where $\varphi'(t') = -1.75 \sin(1.75t') e^{-t'^2/2} - t' \cos(1.75t') e^{-t'^2/2}$,

$$\begin{aligned} \frac{\partial(t')}{\partial a_i(t-1)} \quad \text{and} \quad \frac{\partial(t')}{\partial b_i(t-1)} &\text{ can be calculated by} \\ t' &= \frac{h_i(t) - b_i(t-1)}{a_i(t-1)}. \end{aligned}$$

The initial training value is $x_{c,i}(t) = 0$, $\frac{\partial H_i(t)}{\partial W_{ij}^2(t-1)} = 0$,

$$\frac{\partial H_i(t)}{\partial v_{ik}(t-1)} = 0, \quad \frac{\partial H_i(t)}{\partial a_i(t-1)} = 0, \quad \frac{\partial H_i(t)}{\partial b_i(t-1)} = 0, \quad \forall i, j, k, t = 1.$$

With the method above, the final DRWNN is determined.

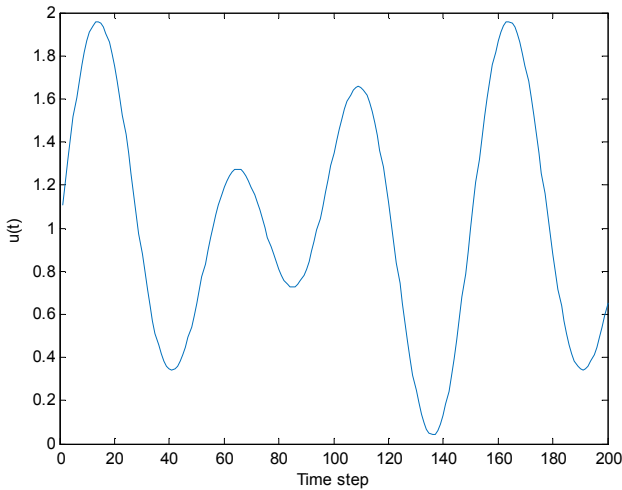
4.4. Experimental Results and Analysis

In order to verify the modelling ability of the proposed DRWNN, A single input single output dynamical nonlinear system is used for the testing. The dynamic equation is described as:

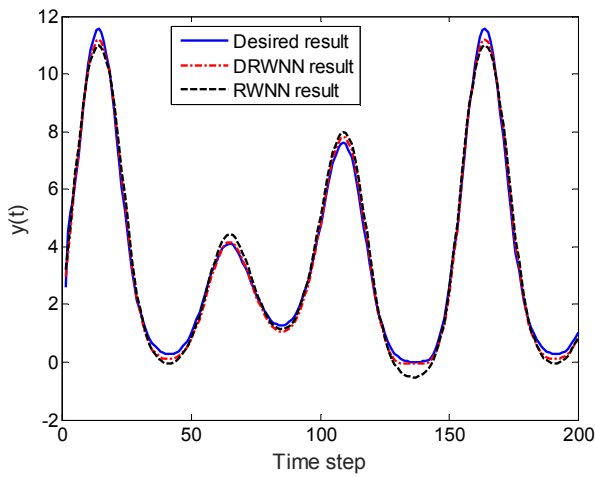
$$\begin{cases} y(t+1) = f(t) + g(t) + \omega_y(t+1), \\ f(t) = 3y(t) / (6.2 + 2y(t-1) + y^2(t-2)), \\ g(t) = u^2(t) + u^3(t), \\ u(t) = 1.0 + 0.6 \sin(2k\pi / 50) + 0.4 \sin(2k\pi / 75). \end{cases} \quad (46)$$

where $u(t)$ is the input, $y(t)$ is the output, $\omega_y(t)$ is a zero mean white noise.

In this section, the neural network used for testing contains 6 hidden layers, 1 input layer, 1 output layer, and the number of feedback layers of DRWNN is 3. $u(t)$ and $y(t)$ of the first 200 time steps are taken as the training samples. The training signal is shown in Figure 2. In order to compare the performance, the RWNN is introduced to model with the same training way [11]. The training number of both RWNN and DRWNN is 200 times.



Input signal



Output signal

Figure 2. Training signals.

The simulation results are shown in Figure 2. To explain the results better, the maximal absolute error E_m is defined, the mean error \bar{E} and the root-mean-square error $RMSE_1$ as follows,

$$E_m = \max |y_i^* - y_i|, i = 1, 2, \dots, N, \quad (47)$$

$$\bar{E} = \frac{\sum_{i=1}^N |y_i^* - y_i|}{N}, \quad (48)$$

$$RMSE_1 = \sqrt{\frac{\sum_{i=1}^N (y_i^* - y_i)^2}{N}}, \quad (49)$$

where y_i^* is the real value, y_i is the calculated value, and N is the number of samples. The training results are shown as Table 1.

Table 1. The error of training samples.

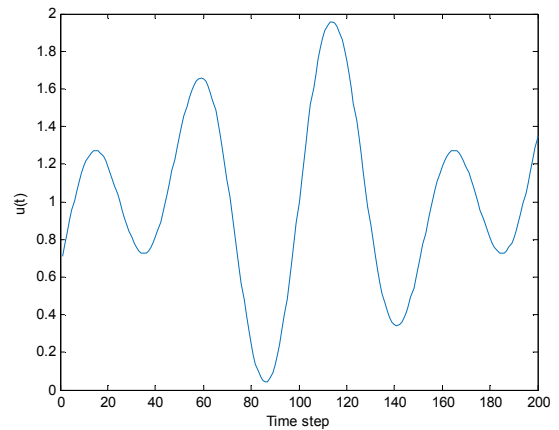
—	$RMSE_1$	E_m	E
RWNN	0.2549	3.1343	0.3729
DRWNN	0.1924	2.9806	0.3249

It can be shown that no matter for the maximal absolute error, the mean error or the root-mean-square error, the results of DRWNN is always less than that of RWNN, and the model output curve of DRWNN is closer than that of RWNN. So DRWNN has better modelling ability.

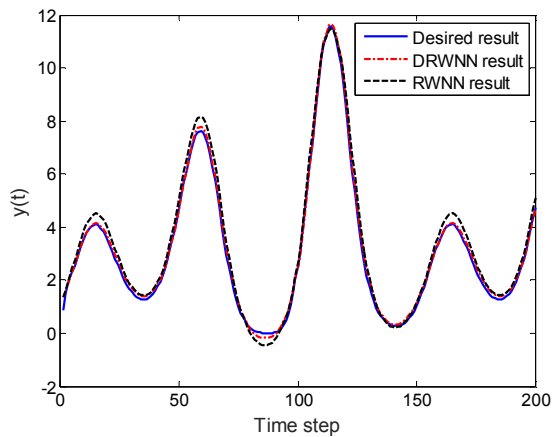
Table 2. The error of testing samples.

—	$RMSE_1$	E_m	E
RWNN	0.1700	1.2023	0.2599
DRWNN	0.1075	1.0082	0.3167

In order to verify the generalization ability of DRWNN, the input signal $u(t)$ from the time step 201 to 400 is served as the testing signal $\tilde{u}(t)$, and the output $\tilde{y}(t)$ is obtained. The training signal is shown in Figure 3 and the corresponding errors are shown in Table 2. These results demonstrate that the output of DRWNN is much closer to the real output, which verifies the DRWNN has better model generalization ability. Using the SNPRP conjugate gradient method to optimize the weights of DRWNN is effective.



Input signal



Output signal

Figure 3. Testing signals.

5. Modelling for ASP Flooding with Proposed Spatial-Temporal Separation Method

5.1. Reservoir Description

Suppose that the reservoir of ASP flooding consists of four injections and nine production wells. All wells distribute uniformly; there is one injection well at the center of every four production wells. The distribution of wells is shown in Figure 4.

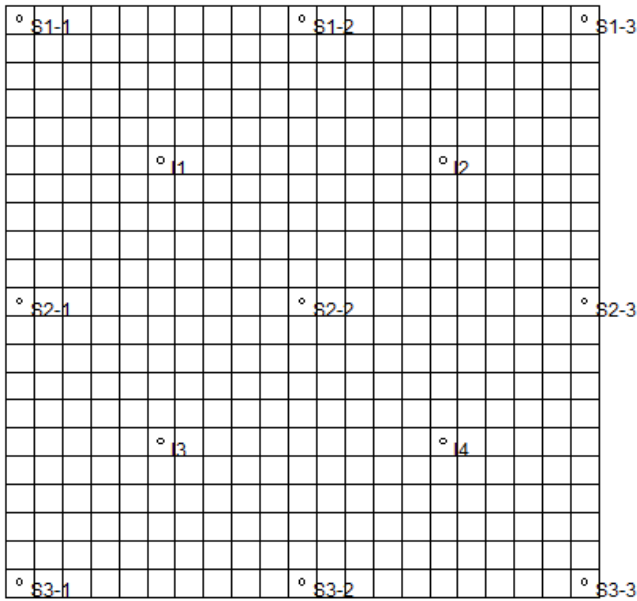


Figure 4. The distribution diagram of well position.

As for this reservoir, the length is 630m; the width is 630m, and the thickness is 19.990m. There are 7 layers in all; the thickness of each layer is 2.857m; and the net thickness is 1.4286m; the depth of upper surface is 2420m; the porosity of every layer is 0.3, and the pore volume is $1.1097 \times 10^6 \text{m}^3$. The initial grid concentration of ASP is 0 g/L. The initial water saturation is shown in Figure 5. The grids of reservoir are divided in three directions x, y, z . The grids in x and y are divided into 21 units, respectively, and the grids in z are divided into 7 units. The total number is $21 \times 21 \times 7$.

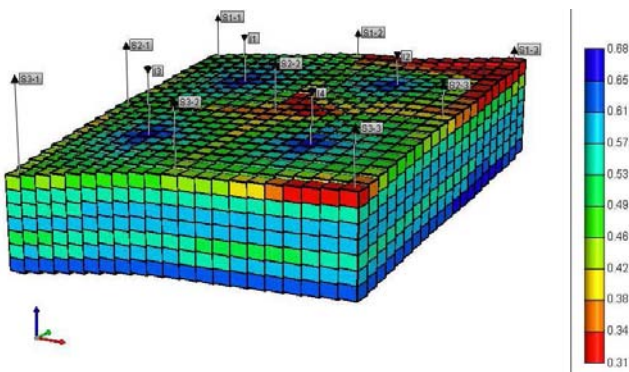


Figure 5. The distribution diagram of initial water saturation.

The water injection rate of each injection well is $83 \text{ m}^3/\text{d}$. The rate of production wells is as defined in Table 3. Using reservoir simulation software CMG, the whole oil production process can be obtained by simulation. The whole production time lasts for 96 months, and sampled output can be got by the snapshots method. The injecting time of ASP is served as the initial time, so the total time nodes are $L=97$ and the total space nodes are $N = 21 \times 21 \times 7$.

Table 3. The liquid volume quantity of production wells.

wells	S1-1	S1-2	S1-3	S2-1	S2-2	S2-3	S3-1	S3-2	S3-3
m^3/day	20.75	41.5	20.75	83	41.5	41.5	20.75	41.5	20.75

5.2. Modelling and Verification for ASP Flooding

Given the ASP flooding above, the approximate model is built with the method proposed in this paper. In order to sufficiently motivate system, the simulation is run on the software CMG for 50 times. The software randomly generates different injection strategies for every month, and the process of injection totally lasts for 48 months. Then the three-dimensional spatial-temporal modelling method is used to get the spatial basis functions and the temporal coefficients of the 50 sets of grid saturation. So as to get the best structure to represent the ASP flooding system, use one group of the spatial basis functions to reconstitute the system with other 49 groups of temporal coefficients and calculate the average mean square errors of all sampling points for these 49 times. After executing this process for 50 times, a group of spatial basis functions with the smallest mean square error is served as the spatial basis functions for modelling. The relationship between injection concentrations of ASP and temporal coefficients is get by DRWNN, then combine the spatial basis functions to build the model between grid water saturation and injection concentrations of ASP. At the same time, build the model between the grid water saturation and moisture content of production wells with DRWNN. The mathematical form can be described as follows,

$$f_w(t) = \hat{F}_2 \left\{ \hat{f}_w(t-1), \dots, \hat{f}_w(t-n_w), \hat{S}_w(t-1), \dots, \hat{S}_w(t-n_s) \right\}, \quad (50)$$

where n_w is the lag of moisture content, n_s is the lag of grid water saturation.

Then the model between the injection concentration of ASP and moisture content of production wells can be built after integrating the models above, which can be expressed as equations (13), (28) and (50).

In order to compare the influence of the number of the spatial basis functions, the performance indicator is defined as follows:

$$RMSE = \sqrt{\frac{\int \sum e(x, y, z, t)^2 dx dy dz}{\int dx dy dz \sum \Delta t}}, \quad (51)$$

where $e(x, y, z, t) = \hat{S}_w(x, y, z, t) - S_w(x, y, z, t)$.

Generally speaking, the more the number of the spatial basis function is, the more information of system the model reflects, and the higher accuracy the model has. But at the same time, the model will be sensitive to the external disturbance and the generalization ability will decrease. Besides, the dimension of the model increases. On the other hand, the less quantity the spatial basis function is; the model will reflect less information and the error can be very big. It is important to choose the proper amounts of the spatial basis functions. In order to test the influence of different numbers spatial basis function on the accuracy of modelling, different numbers of spatial basis functions is chose to reconstruct the system with the temporal coefficients and calculate the average RMSE which is shown in Table 4.

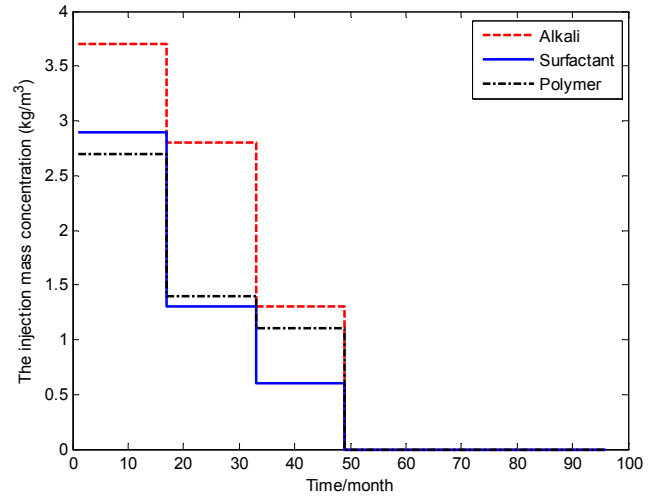
Table 4. RMSE for different number of spatial basis functions.

n	1	2	3	4	5
RMSE	0.0751	0.0502	0.0296	0.0209	0.0197

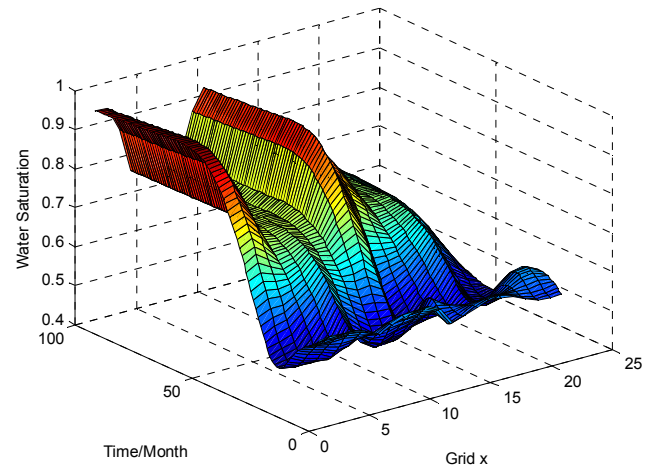
From Table 4, the accuracy of model increase with the number of spatial basis functions. However, when the number is over 4, the RMSE of model will not increase which indicate that the spatial basis functions can reflect the whole reservoir, so the number of spatial basis functions adopted in this paper is 4. Besides that, the number of feedback layers of the DRWNN networks for injection concentrations and temporal coefficients is 3; the corresponding hidden layers is 6; the lags of temporal coefficients is $n_T = 4$; the lags of injection concentrations is $n_u = 3$. The number of feedback layers of DRWNN networks for moisture content of production wells and the corresponding grid water saturation is 2; the hidden layers is 8; the lags of moisture content is $n_w = 3$; the lags of grid water saturation is $n_s = 2$.

Model for the ASP flooding with the spatial-temporal separation method based on DRWNN which is proposed in this paper. The errors for this model are that the mean RMSE for grid water saturation is 0.0285, and the mean RMSE for moisture content of production wells is 1.1378%. This demonstrates the good modelling ability.

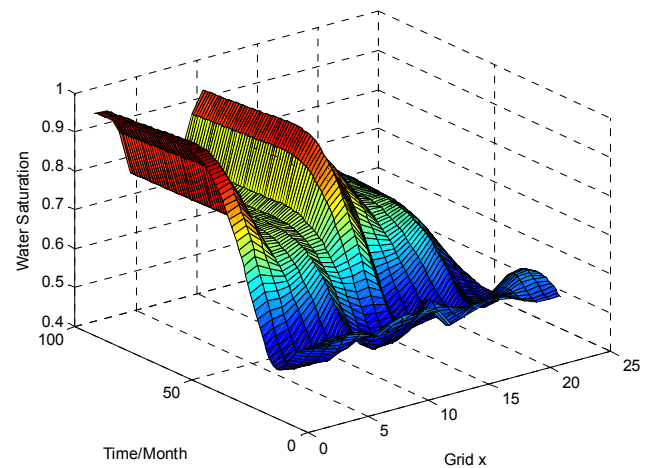
In order to verify the generalization ability, the injection concentrations of ASP flooding are given randomly: $u_P = (2.7, 1.4, 1.1)$, kg/m $u_A = (3.7, 2.8, 1.3)$ kg/m³, $u_S = (2.9, 1.3, 0.6)$ kg/m³. The displacing agent injection totally lasts for 48 months which are divided into 3 slugs uniformly, and the rest time is water flooding. The specific injection can be found in Figure 6-a. Because the whole system is five-dimensional considering the time and value of water saturation, it can't be plotted in the picture directly. In order to illustrate clearly, the 21 grids ranging from (1,11,3) to (21,11,3) at the third layer is chose. The result and error of modelling are shown in Figures 6 and 7.



The injection mass concentration



The output of real system



The output of modelling

Figure 6. The comparison diagram of water saturation.

To analyze the generalization ability further, we compute the error with equation (48). The mean absolute errors for the moisture content of production wells and for the grid water saturation are 1.2970% and 0.0318%, respectively. Figure 8 compares the moisture content of modelling and that of the

simulation software. It can be known that the error is very small, which verifies that the whole model has better generalization ability.

From the above, the three-dimensional spatial-temporal separation modelling method based on DRWNN can model for ASP flooding well. It has good modelling accuracy and generalization ability.

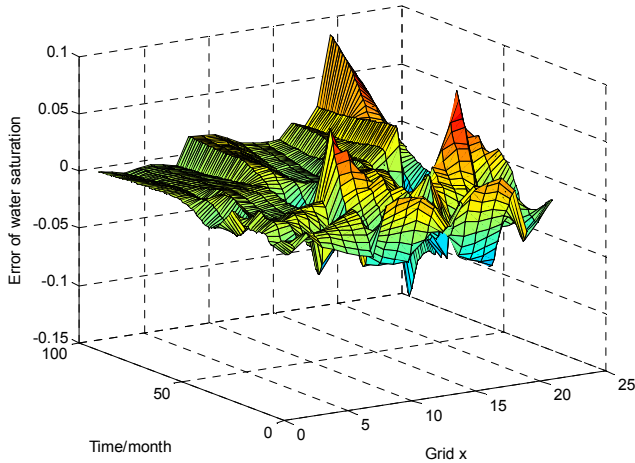


Figure 7. The error diagram of water saturation.

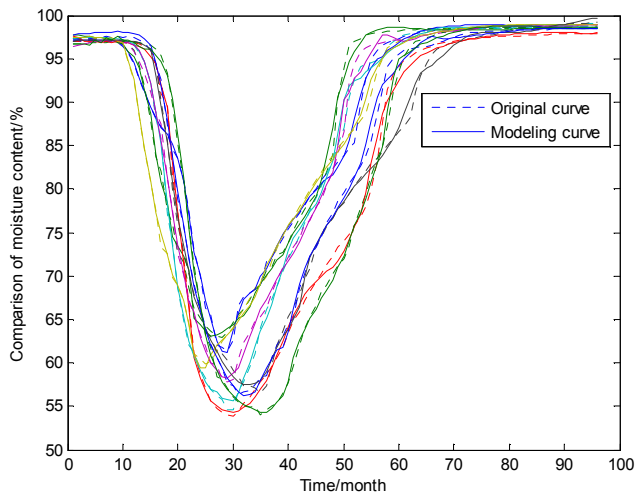


Figure 8. The comparison diagram of moisture content of production wells.

6. Conclusions

In this paper, a three-dimensional spatial-temporal separation modelling approach is proposed to establish the identification model for ASP flooding. At first, spatial state (the water saturation) is expanded onto a series of dominant spatial basis functions and temporal coefficients. Then a new dynamic recurrent wavelet neural network, in which feedback layers are added and different weights are set with time to achieve dynamic memory of the past information, is proposed to build the relationship between the temporal coefficients and the injection concentrations of ASP flooding. To avoid the local minimum and low convergence, the SNPRP conjugate gradient method is adopted to train the weights of DRWNN. Besides that, DRWNN is also used to

establish the model of the moisture content of production wells and the corresponding grid water saturation. The simulation of ASP flooding is carried out to show the effectiveness of this spatial-temporal modelling method. The result shows that the method proposed in this paper has good accuracy and generalization ability. It is suitable to be used for modelling for complex distributed parameter systems like ASP flooding.

Acknowledgements

Authors would like to thank the associate editor and the anonymous referees for their valuable comments and suggestions. This work is supported by National Natural Science Foundation under Grant No. 60974039, National Natural Science Foundation under Grant No. 61573378, Natural Science Foundation of Shandong province under Grant No. ZR2011FM002, the Fundamental Research Funds for the Central Universities under Grant No. 15CX06064A.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Appendix

For the mechanism model of ASP flooding, the specific meaning of parameters is as follows below [19, 20]:

K is the absolute permeability of rock; k_{ro} , k_{rw} are the relative permeability of water and oil; p_o , p_w are the pressure of water and oil; S_o , S_w are the oil saturation and water saturation; B_o , B_w denote the volume factors of oil and water; $p_{cow}(x, y, z, t)$ is the capillary force; c_Θ , $\Theta = \{p, s, OH\}$ are the concentrations of polymer, surfactant and alkali; c_{ij} is the mass concentration of component i in solution j ; μ_o , μ_w are the viscosity of water and oil; ϕ , ϕ_p , ϕ_s are the rock porosity, reachable porosity of polymer and reachable porosity of surfactant; $\phi_{p,s} = \phi f_{a,s}$, f_a , f_s are the reachable porosity factors; R_k is descending coefficient of relative permeability; K_a is the speed factor of the ion exchange and adsorption capacity; K_b is the adsorption constant of surfactant; v_w is the seepage velocity; R_{OH} is the alkali consumption; C_{rp} , C_{rs} are the adsorption quality of unit mass of rock of polymer, surfactant; r_o , r_w are the flow coefficients of oil and water; q_o , q_w are the flow rate of oil and water in the standard state; q_c , q_d , q_e are the transport velocity of shaft flooding agents; D_i , $i \in \{w, o, OH\}$ is the diffusion coefficient; D_{ij} , $i \in \{w, o\}$, $j \in \{s, p\}$ is the diffusion coefficient of component j in solution i .

The flow terms are defined as follows:

$$q_o = \begin{cases} -(1-f_w)q_{out}, & (x, y, z) \in \Psi_p, \\ 0, & (x, y, z) \notin \Psi_p, \end{cases}$$

$$q_w = \begin{cases} -f_w q_{out}, & (x, y, z) \in \Psi_p, \\ q_{in}, & (x, y, z) \in \Psi_w, \\ 0, & (x, y, z) \notin \Psi_p \cup \Psi_w, \end{cases}$$

$$q_c = \begin{cases} q_w c_p, & (x, y, z) \in \Psi_p, \\ q_w c_{pin}, & (x, y, z) \in \Psi_w, \\ 0, & (x, y, z) \notin \Psi_p \cup \Psi_w, \end{cases}$$

$$q_d = \begin{cases} q_w c_{ws} + q_o c_{os}, & (x, y, z) \in \Psi_p, \\ q_w c_{sin}, & (x, y, z) \in \Psi_w, \\ 0, & (x, y, z) \notin \Psi_p \cup \Psi_w, \end{cases}$$

$$q_e = \begin{cases} q_w c_{OH}, & (x, y, z) \in \Psi_p, \\ q_w c_{OHin}, & (x, y, z) \in \Psi_w, \\ 0, & (x, y, z) \notin \Psi_p \cup \Psi_w. \end{cases}$$

The water cut is $f_w = \frac{r_w}{r_w + r_o}$.

The concentration of surfactant is defined as

$$c_s = \frac{q_w c_{ws} + q_o c_{os}}{q_w + q_o}.$$

The alkali consumption is

$$R_a = -\phi S_w \frac{\partial}{\partial t} (r_1 + r_2 + \dots),$$

where r_i denotes the alkali consumption per unit volume of reaction i .

The oil and water relative permeability k_{ro} , k_{rw} can be described as

$$K_{rw,ro} = A \cdot (1 - S_w)^B \cdot (S_w - C)^D.$$

where A,B,C,D are the identification coefficients, the specific method is shown in [21].

The descending of permeability is caused by the adsorption of polymer, it can be described as

$$R_k = 1 + \frac{(R_k^{\max} - 1) \cdot q_p}{q_p^{\max}},$$

where q_p denotes the adsorption quantity of polymer.

References

- [1] E. Verheyen. Oil extraction imperils Africa's Great Lakes. *Science*. 354. 6312 (2016) 561-562.
- [2] Y. Zhu, Q. Hou, W. Liu, et al. Recent Progress and Effects Analysis of ASP Flooding Field Tests. *Canadian Psychiatric Association Journal*, 10 (5) (2012) 387-392.
- [3] F. Abadli. Simulation Study of Enhanced Oil Recovery by ASP (Alkaline, Surfactant and Polymer) Flooding for Norne Field C-segment. Department of Petroleum Engineering & Applied Geophysics, (2012).
- [4] Y. L. Ge, S. R. Li, S. L. Lu, et al. Spatial-Temporal ARX Modelling and Optimization for Polymer Flooding. *Mathematical Problems in Engineering*, (2014).
- [5] G. Prando, A. Chiuso, G. Pilonetto. Maximum Entropy vector kernels for MIMO system identification. *Automatica* 79 (2017) 326-339.
- [6] D. Coca, S. A. Billings. Identification of finite dimensional models of infinite dimensional dynamical systems. *Automatica*, 38 (11) (2002) 1851-1865.
- [7] H. Deng, H. X. Li, G. Chen. Spectral-approximation-based intelligent modelling for distributed thermal processes. *IEEE Transactions on Control Systems Technology*, 13 (5) (2005) 686-700.
- [8] C. K. Qi, H. X. Li. A time/space separation-based Hammerstein modelling approach for nonlinear distributed parameter processes. *Computers & Chemical Engineering*, 33 (7) (2009) 1247-1260.
- [9] C. Wan, M. Pan, et al. Performance improvement of magnetic anomaly detector using Karhunen-Loeve expansion. *Iet Science Measurement & Technology*. 11 (5) (2017) 600-606.
- [10] N. Hamilton, M. Tutkun, R. B. Cal. Low-order representations of the canonical wind turbine array boundary layer via double proper orthogonal decomposition. *Physics of Fluids*. 28 (2) (2016) 683-739.
- [11] F. Y. Zhao, Z. Y. Ma. Nonlinear dynamical system simulation based on recurrent wavelet neural network. *Journal of System Simulation*, 19 (7) (2007) 1453-1539.
- [12] N. Yu. Universal gradient methods for convex optimization problems. *Mathematical Programming*, 152 (1-2) (2015) 381-404.
- [13] F. Rothlauf. *Optimization Methods. Design of Modern Heuristics*. Springer Berlin Heidelberg, 2011, 45-102.
- [14] Y. H. Dai, Y. X. Yuan. *The Nonlinear Conjugate Gradient Method*. Shanghai: Shanghai Science and Technology Press, 2001.
- [15] L. Zhang. An improved Wei-Yao-Liu nonlinear conjugate gradient method for optimization computation. *Applied Mathematics & Computation*, 215 (6) (2009) 2269-2274.
- [16] Z. Q. Li, P. R. Lin, Z. X. Wei. A New Spectral Conjugate Gradient Method for Solving Unconstrained Minimization Problem. *Journal of Southwest University (Natural Science Edition)*, 38 (7) (2016) 115-120.

- [17] C. Hua, N. Li, S. Y. Li. Time-space ARX modelling and predictive control for distributed parameter system. *Control Theory & Applications*, 28 (12) (2011) 1711-1716.
- [18] C. K. Qi, H. X. Li. Nonlinear dimension reduction based neural modelling for distributed parameter processes. *Chemical Engineering Science*, 64 (19) (2009) 4164-4170.
- [19] S. R. Li, X. D. Zhang. Optimal control of polymer flooding for enhanced oil recovery, Dongying: China university of petroleum press, 2013, 19-23.
- [20] C. Z. Yang, et al. Enhanced oil recovery for chemical flooding, Beijing: Petroleum Industry Press, 2007.10, 328-344.
- [21] Y. L. Ge, S. R. Li, K. X. Qu. A Novel Empirical Equation for Relative Permeability in Low Permeability Reservoirs. *Chinese Journal of Chemical Engineering*, 22 (11) (2014) 1274-1278.
- [22] F. M. F. El-Sousy, K. A. Abuhasel. Adaptive Nonlinear Disturbance Observer Using Double Loop Self-Organizing Recurrent Wavelet-Neural-Network for Two-Axis Motion Control System. *IEEE Transactions on Industry Applications* 99 (2017).
- [23] Z. Wan, C. Hu, Z. Yang. A spectral PRP conjugate gradient methods for nonconvex optimization problem based on modified line search. *Discrete and Continuous Dynamical Systems - Series B (DCDS-B)*. 16 (4) (2017) 1157-1169.