

Original Research

Application of Dual-Response Surface Methodology and Radial Basis Function Artificial Neural Network on Surrogate Model of the Groundwater Flow Numerical Simulation

Yanping Yi^{1,2*}, Wenxi Lu¹, Defa Hong³, Hongchao Liu², Lei Zhang²

¹College of Environment and Resources, Jilin University, Changchun 130021, China

²Songliao Institute of Water Environment Science, Songliao River Basin Water Resources Protection Bureau, Changchun 130021, China

³Changchun Institute of Urban Planning and Design, Changchun 130033, China

Received: 12 December 2016

Accepted: 7 February 2017

Abstract

The surrogate model is an effective way to connect the simulation and optimization models in groundwater flow numerical modeling; it could overcome the limitations of embedding and calling simulation models in the optimization model by conventional methods, which greatly reduces the computational load caused by directly calling the simulation model in the solving process of the optimization model. In this paper, the dual-response surface method and radial basis function artificial neural network method were applied to establish the surrogate model of groundwater flow numerical simulation in Jinqian Industrial Park, Inner Mongolia, China. The Latin hypercube sampling method was used to determine random pumping load of the five pumping wells, which were taken as the input data groundwater flow numerical simulation model for calculating 10 observation wells drawdown data sets (output data sets). Based on the input and output data sets, the dual-response surface method and radial basis function artificial neural network method were used to establish the surrogate model of groundwater simulation model, and the validity of surrogate models were comparatively tested. The results showed that both the results of two surrogate models fit well with the results of the simulation model, which indicates that two surrogate models were capable of approaching the groundwater flow numerical simulation model; compared with the dual response surface model, the RBF neural network model had more advantages in terms of sample size requirements, fitting the accuracy of simulation results.

Keywords: surrogate model, dual-response surface methodology, radial basis function artificial neural network, groundwater numerical simulation model

Introduction

The efficient integration of the numerical simulation and optimization management models in groundwater management has been a great concern to water managers. The numerical simulation model focuses on solving the problem of forecasting, while the optimization management model focuses on solving problems of optimization of a decision-making program in groundwater management [1], while the notification and efficient integration of two mathematical models are important issues that need to be resolved in the management of water resources.

The current methods used to solve the efficient integration of optimization management model and numerical simulation model of groundwater management are mainly the embedding method, response matrix method, state equation method, and so on [2]. However, these methods have their own scope of application, and were often limited in practical use [3]. For example, the dimension of embedding the increasing model led to increasing the amount of calculation, for multi-period unsteady flow problems, it would result in the “curse of dimensionality”; state variables of response matrix method in the optimization model did not appear in the form of decision variables, it was only applied to model coupling of the linear system [4].

To address the limitations of conventional methods, we try to build an approximate surrogate model which is functionally close to the groundwater flow simulation model, and the surrogate model could be directly called in the process of the optimization model iterative solution (to solve the surrogate model) to overcome the limitations of embedding and calling simulation models in the optimization model by conventional methods, greatly reducing the calculation load by calling the simulation model in the calculation process of the optimal model [5]. The surrogate model theory had been widely used in petroleum, mechanical design, and other fields, but it has not been reported a lot in the groundwater resource management. The common surrogate model included the dual response surface model based on regression analysis, the artificial neural network model, and the support vector machine model and so on, where the dual response surface method and radial basis function neural network method were relatively well developed research methods [6-8]. Response surface methodology was first proposed by Box and Wilson in 1951 [9], and was applied in the chemical industry, food industry, and biology. Until the late of 20th century, Kim and Lin found that the variables in the traditional response surface method often generate errors in many cases [10]. Meanwhile, Vining and Myers proposed a “dual response surface method” to surrogate the traditional response surface method, and this method had been widely used and constantly improved in recent years [11-13].

Artificial neural network research was explored in the 1940s. In 1985 the radial basis function method of multivariate interpolation was proposed by Powell [14]. And in 1988 the radial basis function network was used

in artificial neural network design for the first time by Broomhead and Moody and Darken. It was successfully applied to pattern recognition, which constituted a radial basis function (RBF) neural network [15]. RBF has the advantages of fast convergence and high precision approximation and small network size, which overcomes the problems of local minimum and slow convergence of the BP network [16-17]. It has attracted increasing attention of researchers in the field of pattern recognition, chemical information processing, and signal control theory and so on, and in recent years it has been widely used in the fields of meteorology, hydrology, and hydrogeology, etc.

This paper attempts to establish a surrogate model for a groundwater numerical simulation model with dual response surface method and radial basis function neural network. Jinqian Industrial Park water sources in Inner Mongolia, China, to discuss in detail the main process of an surrogate model with two methods, the comparison between two surrogate models was performed to demonstrate the reliability and advantages and disadvantages of two models, which may provide theoretical support for the integration of the groundwater optimization management model and groundwater numerical simulation model.

Methodology

Dual-Response Surface Methodology

The dual-response surface method is an objective optimization method that establishes two response surfaces to realize objective optimization, where one surface fits the mean of response, μ , and the other surface fits its standard deviation, σ^2 , which can ensure the solution minimizing the design objective volatility in the interference of controlled or uncontrolled factors, but also ensures that the target tended to be optimal [18-19].

In the study of the relationship of pumping and drawdown, the pumping dataset was first input into the groundwater simulation model to calculate the drawdown data set, and then two response surface-mean and standard deviations were established by polynomial regression equations with the “output data set,” where the response relationship between groundwater pumping intensity and drawdown was close to the groundwater flow numerical simulation model in function.

It is assumed that there were n pumping wells, m observation wells, and the number of pumping trials was N .

- Establish regression equation of the observation wells drawdown mean.

From the impact of the different pumping test program on the drawdown of groundwater, the statistical relationship between drawdown mean of observation well and pumping intensity of each pumping well could be analyzed.

The average drawdown of observation wells in a pumping test is expressed as:

$$\bar{y} = \frac{1}{m} \sum_{i=1}^m y_i \tag{1}$$

...where y_i was the drawdown value of the i^{th} observation well and $i = 1,2,3,L m$ was the total number of observation wells.

Set pumping load x_1, x_2, L, x_n of each pumping well as an independent variables, and the drawdown mean of each observation well as the dependent variable, the polynomial regression equation could be established as:

$$\bar{y} = b_0 + \sum_{j=1}^n b_j x_j + \sum_{j=1}^n b_{jj} x_j^2 + \sum_{p<j}^n b_{pj} x_p x_j + e_m \tag{2}$$

...where β was the regression coefficient of the surface; e_μ was the fitting error; $j = 1,2,3,L n$; and n was the total number of pumping wells.

According to the pumping test data of each group:

$$\begin{matrix} y^{(1)}, x_1^{(1)}, x_2^{(1)}, L, x_n^{(1)} \\ y^{(2)}, x_1^{(2)}, x_2^{(2)}, L, x_n^{(2)} \\ M \quad M \quad M \quad M \\ y^{(N)}, x_1^{(N)}, x_2^{(N)}, L, x_n^{(N)} \end{matrix}$$

Regression equations could be established respectively, regression coefficient β could be determined by the least squares method, $x_n^{(N)}$ was the pumping intensity of the n^{th} pumping wells in the N^{th} pumping test, and N was the total number of pumping tests.

Establishing the regression equation of residual standard deviation of drawdown of the observation wells:

– According to the relationship between drawdown residual standard deviation of observation wells and pumping intensity of each pumping well, the differential degree of each observation wells drawdown in a pumping test under certain pumping intensity distribution schemes could be expressed and measured.

Residual sum of squares Q of observation wells drawdown in a pumping test could be calculated with equation (3):

$$Q = \sum_{i=1}^m (\bar{y} - y_i)^2 \tag{3}$$

...where m is the total number of observation wells; \bar{y} is the drawdown mean of all m observation wells in a pumping test; y_i is drawdown value of the i^{th} observation wells; $i = 1,2,3,L m$; and m is the total number of observation wells.

The residual standard deviation S of observation wells drawdown in a pumping test could be calculated as:

$$S = \sqrt{\frac{Q}{m - n - 1}} \tag{4}$$

...where m is the total number of observation wells and n is the total number of pumping wells.

Set pumping intensity x_1, x_2, L, x_n of each pumping well in the pumping test as independent variables, and the residual standard S of each observation well drawdown as the dependent variables, the polynomial regression equation could be established as:

$$S = g_0 + \sum_{j=1}^n g_j x_j + \sum_{j=1}^n g_{jj} x_j^2 + \sum_{p<j}^n g_{pj} x_p x_j + e_s \tag{5}$$

...where γ was the regression coefficient of the surface; e_σ was the fitting error; $j = 1,2,3,L n$; and n was the total number of pumping wells.

According to the pumping test data of each group:

$$\begin{matrix} S^{(1)}, x_1^{(1)}, x_2^{(1)}, L, x_n^{(1)} \\ S^{(2)}, x_1^{(2)}, x_2^{(2)}, L, x_n^{(2)} \\ M \quad M \quad M \quad M \\ S^{(N)}, x_1^{(N)}, x_2^{(N)}, L, x_n^{(N)} \end{matrix}$$

Regression equations could be established respectively, γ could be obtained by the least squares method, N was the total number of pumping tests, and $x_n^{(N)}$ was the pumping intensity of the n^{th} pumping wells in the N^{th} pumping test.

The statistical relationship between residual standard deviation of observation wells drawdown and pumping intensity was calculated, and for each pumping test the residual standard deviation of observation wells drawdown could be calculated. Then this indicator could be used to describe and measure the inconsistency (difference) degree of each observation well drawdown.

By establishing two response surfaces, polynomial on the pumping intensity and mean of groundwater drawdown and the residual standard deviation could be listed. About solving polynomial regression coefficients, in general, the polynomial regression was transformed into multiple linear regressions, so the fitting error could be zero, that is, for any polynomial comprised of many variables:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1^2 + b_4 x_1 x_2 + b_5 x_2^2 + \dots$$

We could assume to make $z_1 = x_1, z_2 = x_2, z_3 = x_3, z_4 = x_1 x_2, z_5 = x_2^2, \dots$, the above formula could be transformed into a multiple linear regression problem, then the above equation could be:

$$y = b_0 + b_1 z_1 + b_2 z_2 + b_3 z_3 + b_4 z_4 + b_5 z_5 + \dots$$

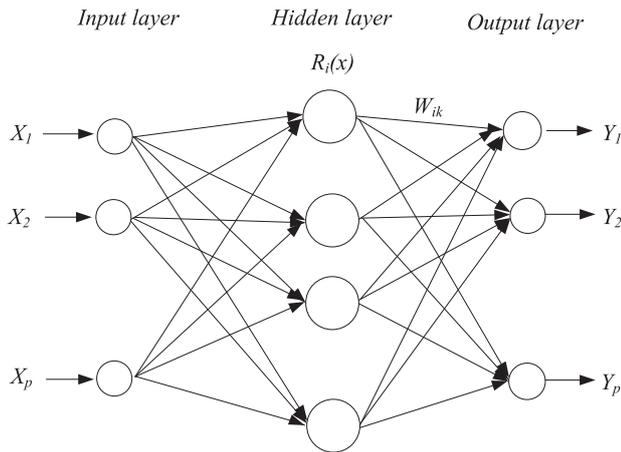


Fig. 1. Topological structure of RBF neural network.

Polynomial regression coefficients b_1, b_2, \dots, b_n and constants term b_0 could be obtained by least squares method.

Radial Basis Function Artificial Neural Network

The radial basis function (RBF) neural network used herein was a feed-forward neural network of local approximation; its topological structure is shown in Fig. 1 and Fig. 2 [20]. The mapping from the input layer to the radial base was nonlinear, and from the radial base layer to the output layer implementation was a linear mapping [21].

Choosing a suitable radial basis function was the key to achieving the network. Radial basis function was a local distribution, central radially symmetric, non-negative attenuation non-linear function; its two parameters (base center and the base width) would decide the range, which produced a significant response to input. Gaussian function has many advantages as the basis function, such as simple forms of expression, and would not increase too much complexity even for multi-variable input, radial symmetry, good smoothness, and presented any order derivatives and so on [22]. Therefore, it was often taken as a radial basis function RBF network, which was expressed as:

$$R_i(x) = \exp\left(\frac{-P X - c_i P^2}{2s_i^2}\right) \quad (6)$$

... where $R_i(x)$ is the output of the i^{th} nodes in hidden layer, $i = 1, 2, 3, \dots, m$; and σ_i is the variance of the i^{th} nodes in hidden layer, which determined the width of ranging to center point of the base functions.

The product of the vector distance and deviation between the weight vector w_j of each neuron j ($j = 1, 2, 3, \dots, L, R$) in radial base layer and the p^{th} input vector x_p is taken as input value in radial basis function to enter the expression:

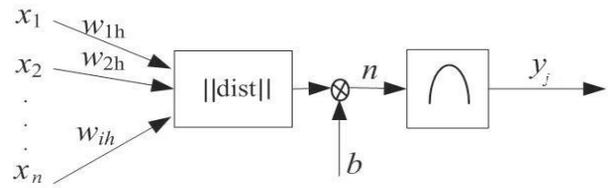


Fig. 2. RBF neural network hidden layer neurons of the input and output.

$$dist = x_p - w_j$$

$$z_j = P \cdot dist \cdot b_j = \sqrt{\prod_{i=1}^m (w_{ij} - x_{pi})^2} \cdot b_j, j = 1, 2, \dots, L, R \quad (7)$$

Thus the output of radial base neuron j was:

$$y_j = e^{-z_j^2} = e^{-(x_p - w_j \cdot P \cdot b_j)^2}, j = 1, 2, \dots, L, R \quad (8)$$

Results and Discussion

Overview of Study Area

The selected study area lies between latitude 105-110° North and longitude 39°69'-40°30' East (Fig. 3), the Piedmont alluvial plain area, the aquifer type was unconsolidated rock pore water, and its aquifer thickness was 42.72-68.01 m. The aquifer lithology was mainly grayish breccia, gravel sand gray-blue, gray-yellow sand, local clay, and clay containing interlayer, and the thickness was in uneven distribution and was not continuous distribution. Buried in shallow groundwater, its quantity was abundant and quality was good. The groundwater of Piedmont alluvial plains was mainly from runoff recharge of mountain groundwater, surface water of mountain (mainly in the flood), and seepage recharge from atmospheric precipitation. The groundwater flowed from

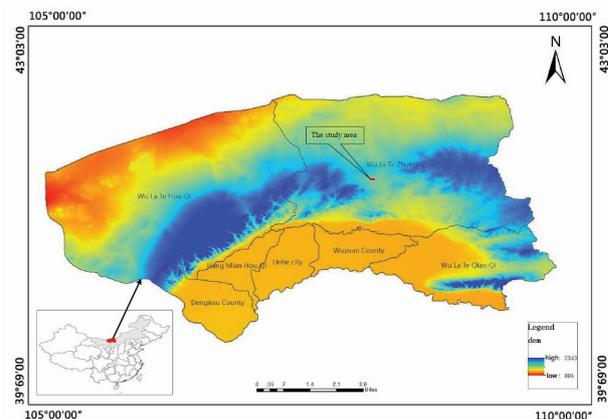


Fig. 3. The study area.

the Piedmont to alluvial plain runoff, at the same time, impacted by a large number of groundwater exploitation, the study area is recharged from lateral runoff of groundwater on both sides of the plain area.

Groundwater Flow Numerical Simulation Model

The simulation calculation area in this research was mainly in the Piedmont alluvial plain area. The target layer of simulation calculation was Quaternary Holocene - Pleistocene unconfined aquifer. The upper boundary of the simulation region was diving free surface, through which the diving water exchange water vertically with the out system; the lower part was the Pleistocene silty clay layer with extremely poor permeability, generalized as aquiclude; the inner zone to the border was generalized as flux boundary.

The mathematical model could be described by following partial differential equations and its definite condition [23]:

$$\begin{cases} \frac{\partial}{\partial x} \left(k(H - Z_b) \frac{\partial H}{\partial x} \right) + \frac{\partial}{\partial y} \left(k(H - Z_b) \frac{\partial H}{\partial y} \right) + W - E - P = \\ H(x, y, t)|_{t=0} = H_0(x, y) \\ k(H - Z_b) \frac{\partial H}{\partial n} \Big|_{\Gamma_2} = q(x, y, t) \\ \mu \frac{\partial H}{\partial t} \quad (x, y) \in D, t \geq 0 \\ (x, y) \in D, t = 0 \\ (x, y) \in \Gamma_2, t > 0 \end{cases} \quad (9)$$

...where H is groundwater head (m), $H_0(x, y)$ is

groundwater initial head (m), Z_b is aquifer floor elevation (m), k is aquifer permeability coefficient (m/d), μ is aquifer-specific yield (dimensionless), W is aquifer vertical recharge intensity (including rainfall, flood, and irrigation water leakage recharge) (m/d), E is groundwater evaporation excretion intensity (m/d), P is aquifer mining intensity (m/d), G_2 is known flow boundary, $q(x, y, t)$ is aquifer lateral displacement per unit width (m^3/d – positive when inflow, negation when outflow, 0 when impermeable boundary), n is outward normal direction on the boundary, and D is range of calculation area.

The simulation period in the research was identified as 7 September 2008 to 18 March 2009. The Groundwater flow field of 7 September was taken as the initial flow field of identification period, and the groundwater flow field of 18 March was taken as the final flow field of identification period, with a total of 183 days. The results of the model test (Fig. 4) declared that the groundwater simulation model could objectively reflect the motion characteristics of groundwater flow.

Surrogate Model

Data Sources and Processing

Five pumping wells were selected in the key simulation calculation area of the research area using the Latin hypercube sampling method to the extent practicable, and the representative pumping load was selected as an input data set and the 30 group different pumping tests program was developed, which provided data preparation for establishing a surrogate model of the groundwater flow numerical simulation model.

The pumping load data sampling results by Latin hypercube sampling methods for the selected five pumping wells are shown in Table 1 and Fig. 5 [24-25].

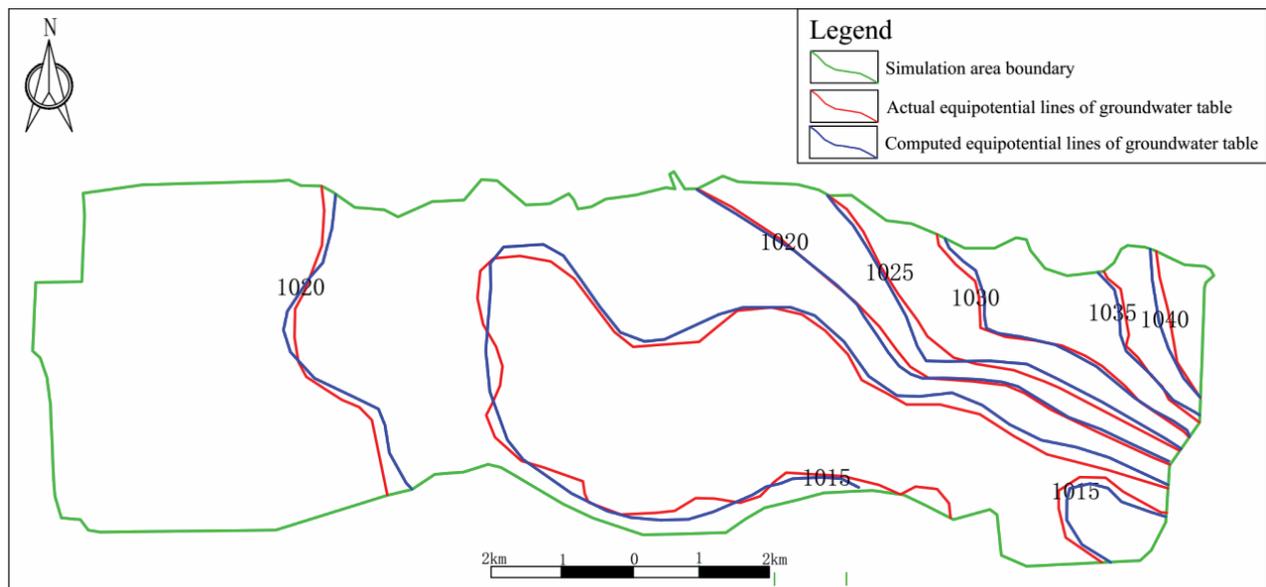


Fig. 4. Flow-fitting chart of the test period at the end of time.

Table 1. LHS sample results.

Well Number	Q ₁	Q ₂	Q ₃	Q ₄	Q ₅
1	45	30	304	165	625
2	150	130	1,089	19	400
3	100	195	119	110	985
4	430	260	28	213	23
5	370	300	1,219	390	1,100
6	265	393	492	278	108
7	525	461	1,437	336	1254
8	385	530	185	441	170
9	600	595	1,219	1,106	1,432
10	735	657	365	500	215
11	1,335	726	981	997	1,319
12	800	790	1,674	554	278
13	665	842	1,325	629	1,608
14	1,407	918	743	1,288	336
15	866	990	1,561	1,505	1,384
16	1,069	1,055	230	599	456
17	800	1,120	1,620	1,546	1,537
18	938	1,188	425	1,210	1,653
19	1,742	1,255	1,507	720	518
20	1,000	1,319	552	1,455	1,198
21	1,809	1,386	620	778	579
22	1,135	1,452	1,365	1,391	1,496
23	1,900	1,500	1,277	884	684
24	1,671	1,569	1,035	835	729
25	1,268	1,649	928	946	1,132
26	1,742	1,715	679	1,669	806
27	2,000	1,778	1,175	1,037	1,021
28	1,532	1,846	800	1,153	869
29	1,875	1,905	1,158	1,325	934
30	1,600	1,979	865	1,615	916

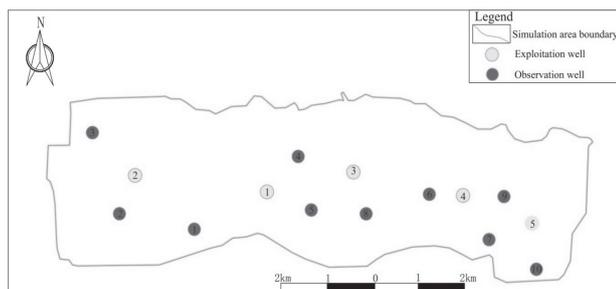


Fig. 5. The distribution of exploitation and observation wells at the simulation areas.

Dual Response Surface Model

1. Establishing regression equation of the observation wells drawdown mean

According to the theory of the dual-response surface method, the regression equation of the observation wells drawdown mean was established by existing data:

$$\bar{y} = 0.472678207 + 1.5311 \cdot 10^{-5} x_1 - 9.28894 \cdot 10^{-6} x_2 + 1.08451 \cdot 10^{-5} x_3 + 9.97093 \cdot 10^{-5} x_4 + 4.84779 \cdot 10^{-6} x_5 - 3.06998 \cdot 10^{-8} x_1^2 - 1.35851 \cdot 10^{-8} x_2^2 + 3.54148 \cdot 10^{-9} x_3^2 - 1.3116 \cdot 10^{-8} x_4^2 + 5.33059 \cdot 10^{-9} x_5^2 + 1.47388 \cdot 10^{-8} x_1 x_2 - 3.98616 \cdot 10^{-9} x_1 x_3 - 2.89901 \cdot 10^{-5} x_1 x_4 + 1.58054 \cdot 10^{-8} x_1 x_5 + 1.01887 \cdot 10^{-8} x_2 x_3 + 3.0194 \cdot 10^{-8} x_2 x_4 - 1.05236 \cdot 10^{-8} x_2 x_5 + 3.75841 \cdot 10^{-8} x_3 x_4 - 1.87759 \cdot 10^{-9} x_3 x_5 - 2.30902 \cdot 10^{-5} x_4 x_5$$

Then each group pumping test program data was substituted into the regression equation to obtain the regression values of drawdown mean under different programs, as shown in Table 2.

Table 2 shows that the regression values of groundwater drawdown mean under the 1, 2, 3, and 4 pumping test programs were relatively small. The establishment of observation well drawdown means regression equation could know the impact of different pumping test programs on groundwater drawdown. If the regression values of groundwater drawdown mean were relatively small, the pumping of groundwater would have little effect on the water level.

2. Establishing the regression equation of residual standard deviation on the observation wells drawdown

Table 2. The mean water level drawdown of different pumping schemes.

Scheme	Mean value of groundwater level (m)	Scheme	Mean value of groundwater level (m)
1	0.496968	16	0.536002
2	0.491058	17	0.619618
3	0.490845	18	0.572725
4	0.498168	19	0.584738
5	0.537995	20	0.590627
6	0.51251	21	0.566305
7	0.543607	22	0.610618
8	0.51537	23	0.590645
9	0.580342	24	0.579652
10	0.528012	25	0.5785
11	0.577758	26	0.616283
12	0.566497	27	0.59905
13	0.56024	28	0.590392
14	0.589012	29	0.613212
15	0.615923	30	0.618205

Another response surface: the regression equation of residual standard deviation on the observation wells drawdown was established by using input and output data for the residual standard deviation of observation wells drawdown that was obtained by using equations (3) and (4):

$$S = 0.433750445 + 2.96329 \cdot 10^{-7} x_1 - 2.75595 \cdot 10^{-5} x_2 + 5.71047 \cdot 10^{-6} x_3 + 3.32465 \cdot 10^{-5} x_4 + 8.36091 \cdot 10^{-6} x_5 - 4.74304 \cdot 10^{-9} x_1^2 - 2.42234 \cdot 10^{-8} x_2^2 + 6.49954 \cdot 10^{-9} x_3^2 - 4.03544 \cdot 10^{-9} x_4^2 + 9.81431 \cdot 10^{-9} x_5^2 + 2.61134 \cdot 10^{-8} x_1 x_2 - 7.50606 \cdot 10^{-9} x_1 x_3 - 5.36625 \cdot 10^{-8} x_1 x_4 + 2.84232 \cdot 10^{-8} x_1 x_5 + 1.9281 \cdot 10^{-8} x_2 x_3 + 5.26149 \cdot 10^{-8} x_2 x_4 - 1.97629 \cdot 10^{-8} x_2 x_5 + 2.81508 \cdot 10^{-9} x_3 x_4 - 3.20401 \cdot 10^{-9} x_3 x_5 - 4.08703 \cdot 10^{-8} x_4 x_5$$

Then each group pumping test program data was substituted into the regression equation to obtain the regression values of drawdown residual standard deviation under different programs, as shown in Table 3.

Through the establishment of the regression equation of residual standard deviation on the observation wells drawdown, in selected pumping test programs, the drawdown residual standard deviation could be used as an indicator to describe and measure the degree of difference of each observation well groundwater drawdown.

Thus far, two response surface regression equations were obtained, meaning that the establishment of dual response surface model was completed.

To verify whether the dual response surface model could surrogate the groundwater flow numerical model, any five different pumping test programs in the pumping load allowable range of selected five pumping wells were developed as the verification data. Then the authentication data were input into groundwater flow numerical simulation model and dual response surface model to

Table 3. Drawdown residual standard deviation of the return value under different pumping test programs.

Scheme	Value	Scheme	Value
1	0.445394	16	0.422995
2	0.450187	17	0.450238
3	0.445217	18	0.414549
4	0.432405	19	0.45171
5	0.462338	20	0.418785
6	0.439356	21	0.422094
7	0.470694	22	0.439964
8	0.435627	23	0.440896
9	0.445685	24	0.434966
10	0.434011	25	0.42757
11	0.433921	26	0.411204
12	0.470098	27	0.433976
13	0.460647	28	0.421552
14	0.422883	29	0.429834
15	0.447714	30	0.423135

Table 4. Results of the mean and the residual standard deviation of running groundwater numerical simulation model and dual response surface model (m).

Scheme	Groundwater numerical simulation model	Dual response surface model	Average relative error
Mean value of water level	0.512664	0.574099	0.056
	0.593256	0.627166	
	0.512534	0.499825	
	0.532882	0.566621	
	0.563761	0.55627	
Water level drawdown residual standard deviation	0.46123	0.440306	0.045
	0.53165	0.495028	
	0.39012	0.412614	
	0.40175	0.418193	
	0.42518	0.432202	

calculate the mean and the residual standard deviation of observation wells drawdown, the model results are shown in Table 4.

The average relative error of water level drawdown mean by dual response surface model and the fitting results by simulation model was 0.056; the average relative error of drawdown residual standard deviation and fitting results by the simulation model was 0.045.

The fitting errors of both water level drawdown mean and residual standard deviation by dual response surface model and the results of groundwater numerical simulation model were small, indicating that dual response surface approximated the simulation model in function, which could effectively surrogate the groundwater flow numerical simulation model.

RBF Neural Network Model

In order to simulate the effect of simultaneous extraction of five pumping wells on groundwater drawdown, the RBF neural network model was used to build a nonlinear relationship model between pumping intensity and observation wells drawdown. Thirty different sets of pumping test program by Latin hypercube sampling were used as training samples and the verified data in Table 5 were used as the test samples, which meant 1 to 30 groups were selected as training samples; 31 to 35 groups were selected as test samples.

With the input (pumping) and output (drawdown) datasets, two RBF neural network models were established: pumping well pumping load-observation wells drawdown mean and pumping well pumping load-observation wells drawdown residual standard deviation. Matlab 6.5 is taken as a platform to call the function newrb () to create an RBF neural network model, and the calling format is [26-27]:

Table 5. RBF neural network training data samples and test samples.

Sample number	Pumping capacity (m ³ /d)					Mean value of water level drawdown (m)	Residual standard deviation of water level in observation wells (m)	Remarks
	Q ₁	Q ₂	Q ₃	Q ₄	Q ₅	\bar{y}	S	
1	45	30	304	165	625	0.496968	0.447849	Training sample
2	150	130	1,089	19	400	0.491058	0.443612	
3	100	195	119	110	985	0.490845	0.443836	
4	430	260	28	213	23	0.498168	0.434339	
5	370	300	1,219	390	1,100	0.537995	0.465451	
6	265	393	492	278	108	0.51251	0.447243	
7	525	461	1,437	336	1,254	0.543607	0.473492	
8	385	530	185	441	170	0.51537	0.430239	
9	600	595	1,219	1,106	1,432	0.580342	0.44418	
10	735	657	365	500	215	0.528012	0.431454	
11	1,335	726	981	997	1,319	0.577758	0.433291	
12	800	790	1674	554	278	0.566497	0.470527	
13	665	842	1,325	629	1,608	0.56024	0.457654	
14	1,407	918	743	1,288	336	0.589012	0.421565	
15	866	990	1,561	1,505	1,384	0.615923	0.448406	
16	1,069	1,055	230	599	456	0.536002	0.421353	
17	800	1,120	1,620	1,546	1,537	0.619618	0.450367	
18	938	1,188	425	1,210	1,653	0.572725	0.415609	
19	1742	1255	1507	720	518	0.584738	0.453291	
20	1,000	1,319	552	1,455	1,198	0.590627	0.417635	
21	1,809	1,386	620	778	579	0.566305	0.423437	
22	1,135	1,452	1,365	1,391	1,496	0.610618	0.439895	
23	1,900	1,500	1,277	884	684	0.590645	0.440086	
24	1,671	1,569	1035	835	729	0.579652	0.435076	
25	1,268	1,670	928	946	1,132	0.5785	0.43143	
26	1,742	1,715	679	1,669	806	0.616283	0.415575	
27	2,000	1,778	1,175	1,037	1,021	0.59905	0.432311	
28	1,532	1,846	800	1,153	869	0.590392	0.421759	
29	1,875	1,905	1,158	1,325	934	0.613212	0.427985	
30	1,600	1,979	865	1,615	916	0.618205	0.420699	
31	720	356	983	1,067	562	0.512664	0.46123	Test sample
32	348	1,385	862	1,468	289	0.593256	0.53165	
33	185	942	257	359	827	0.512534	0.39012	
34	529	487	246	1,322	1,479	0.532882	0.40175	
35	765	654	234	925	1,632	0.563761	0.42518	

net = newrb (P , T , goal , spread , mn , df)

...where P is the input vector; T was the target vector; goal is the mean square error with the default of 0; spread is the distribution density of radial basis functions – the greater of the spread value, the smoother the function (the default value of spread was 1); mn is the maximum number of neurons; and df is the added number of neurons between two displays.

Then simulation training of the network was taken to verify identification and prediction performance. The calling format is:

y = sim (net, X_test)

... where X_test was the testing sample of the model.

First we established the RBF neural network model on pumping wells pumping load-observation wells drawdown mean. The input layer represented the pumping wells pumping load, the number of neurons was consistent with the number of pumping wells, as the number of 5; output layer represented the observation wells drawdown mean, and the number of neurons was set as 1. Setting each parameter value in newrb () function: goal = 0.001, spread = 50,000, mn = 20, df = 1.

By programming in matlab codes, the operating results obtained by RBF neural network model were as follows:

$\bar{y} = 0.5628 \quad 0.5926 \quad 0.5184 \quad 0.5581 \quad 0.5446$
 $err = 0.0501 \quad -0.0007 \quad 0.0059 \quad 0.0253 \quad -0.0191$

Then we established the RBF neural network model of pumping well pumping load-observation wells drawdown residual standard deviation. The input layer still represented pumping wells pumping load, the number of neurons was consistent with the number of pumping wells, as the number of 5; output layer represented the observation wells drawdown residual standard deviation, and the number of neurons was set as 1. Setting each

parameter values in newrb () function: goal = 0.000001, spread = 3,273, mn = 50, df = 1.

By matlab programming, the operating results obtained by RBF neural network model were:

$S = 0.4481 \quad 0.4904 \quad 0.4217 \quad 0.4074 \quad 0.4198$
 $err = -0.0131 \quad -0.0412 \quad 0.0316 \quad 0.0057 \quad -0.0053$

We found that the average relative error of water level drawdown mean obtained by RBF neural network model and the calculated results by simulation model was 0.038; the average relative error of drawdown residual standard deviation and the simulation model results was 0.042.

Comparative Analysis of Results

Both the dual response surface model and RBF neural network model could approximate groundwater flow numerical simulation model in function. A comparative analysis of the two models are performed to see which model is more robust, and it was found that the approximate surrogate model is more suitable for the groundwater flow numerical simulation model.

In the use of the same test data to verify the validity of the surrogate models, it was found that the accuracy of two surrogate models existed some difference, and the average relative error of water level drawdown means by dual response surface model and the fitting results by simulation model was 0.056; the average relative error of drawdown residual standard deviation and fitting results by the simulation model was 0.045, while the average relative error of water level drawdown mean obtained by RBF neural network model and the calculated results by simulation model was 0.038; the average relative error of drawdown residual standard deviation and the simulation model results was 0.042. We calculated that the fitting error of drawdown mean obtained by dual response surface model was 32% higher than the fitting error of drawdown mean obtained by RBF neural network model, the fitting

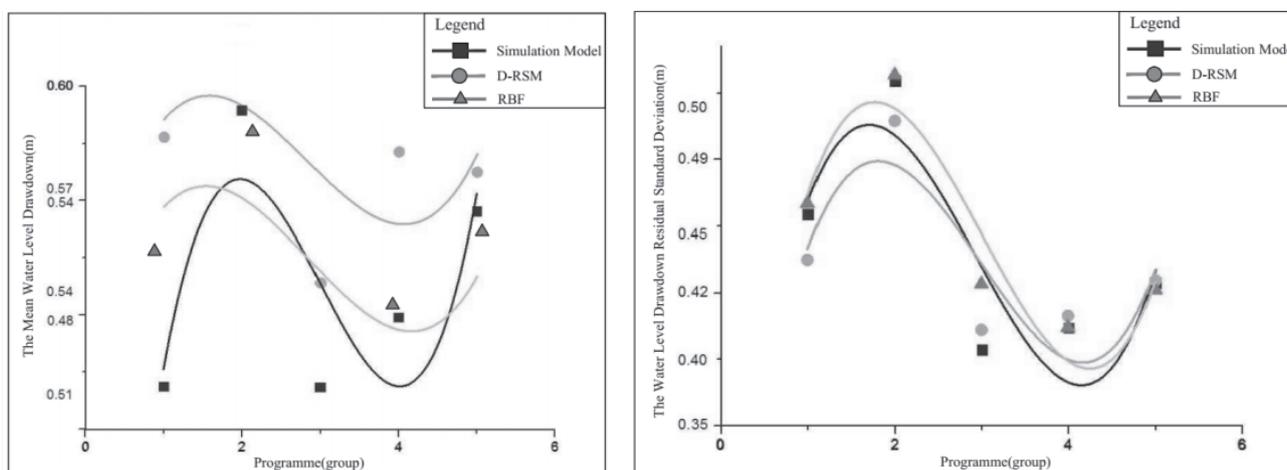


Fig. 6. Fitting of surrogate model and simulation model.

error of drawdown the residual standard deviation obtained by dual response surface model was 6% higher than the fitting error of drawdown the residual standard deviation obtained by RBF neural network model. This showed that the results accuracy of RBF neural network model was slightly higher than the dual response surface model, which was more approximate to the groundwater flow numerical simulation model. Fitting of the surrogate model and simulation model is shown in Fig. 6.

Therefore, compared with the dual response surface model, the calculation process of the RBF neural network model was simple and easy for operation, with a low requirement for sample number and the fitting error of running results with calculated results obtained by the simulation model was small, more direct expression, and more suitable as approximate surrogate model of groundwater flow numerical simulation model.

Conclusions

In this paper we used Latin hypercube sampling to obtain the input (pumping load) and output (drawdown) data sets in groundwater flow numerical simulation, and these data sets were applied to establish dual response surface and RBF neural network models as surrogate models of groundwater flow numerical simulation. In the same validation data, the results of two surrogate models and the calculated results of the simulation model fit well, and both of the two surrogate models could approximate groundwater flow numerical simulation model in function. By comparative analysis of calculated results of two surrogate models that we focused on in this paper, we found that the RBF neural network model was more competitive in terms of sample size requirement, and results fitting accuracy of the simulation model, etc., which was more conducive to practical use.

The surrogate model of groundwater simulation model established in this paper overcame the limitations of conventional coupling methods, it could be directly called in the solving process of the iteration optimization model, and could greatly reduce computational load caused by direct calling of the simulation model in the solving process of the iteration optimization model. The surrogate model method provides important theoretical and practical significance for solving nesting and fusion of the groundwater optimization management model, and numerical simulation models a series of problems. Besides, it had potential and space on development and management research work of groundwater resources.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (No. 41672232) and the National Water Pollution Control and Management of Science and Technology Major Projects (No. 2012ZX07601002-002). The authors wish to thank the

editor and three anonymous referees whose comments and suggestions improved the quality of this manuscript.

Author Contributions

Yanping Yi performed the experimental work, data collection, and model establishment, and prepared the manuscript. Wenxi Lu had the original idea and supervised the research, while Hongchao Liu and Lei Zhang performed data analysis and manuscript revision.

References

- BAIRD K.J., III T.M. Simulating riparian evapotranspiration: a new methodology and application for groundwater models. *Journal of Hydrology*, **312** (312), 176, **2005**.
- WILLIS R. A Unified Approach to Regional Groundwater Management. *Groundwater Hydraulics*, **9**, 392, **2013**.
- VERMEULEN P.T.M., HEEMINK A.W., STROET C.B.M.T. Reduced models for linear groundwater flow models using empirical orthogonal functions. *Advances in Water Resources*, **27** (1), 57, **2004**.
- SINGH A. Simulation – optimization modeling for conjunctive water use management. *Agricultural Water Management*, **141** (141), 23, **2014**.
- AN Y., LU W., CHENG W. Surrogate Model Application to the Identification of Optimal Groundwater Exploitation Scheme Based on Regression Kriging Method – A Case Study of Western Jilin Province. *International Journal of Environmental Research & Public Health*, **12** (8), 8897, **2014**.
- FAN C., HUANG Y., WANG Q. Sparsity-promoting polynomial response surface: A new surrogate model for response prediction. *Advances in Engineering Software*, **77**, 48, **2014**.
- BAO P.F., MIAO W.D., XIE R., SHI Y.L. A BP-ANN Based Surrogate Modeling for Predicting Engineering Analysis of Forging Press. *Advanced Materials Research*, **915-916** (20), 987, **2014**.
- CICCAZZO A., PILLO G.D., LATORRE V. Support vector machines for surrogate modeling of electronic circuits. *Neural Computing and Applications*, **24** (1), 69, **2014**.
- ŠUMIĆ Z., VAKULA A., TEPIĆ A., et al. Modeling and optimization of red currants vacuum drying process by response surface methodology (RSM). *Food Chemistry*, **203**, 465, **2016**.
- LIN D.K.J., TU W. Dual Response Surface Optimization. *Journal of Quality Technology*, **27** (1), 34, **1995**.
- VINING G.G., MYERS R.H. Combining Taguchi and Response Surface Philosophies: A Dual Response Approach. *Journal of Quality Technology*, **22** (1), 38, **1990**.
- YENIAY O., UNAL R., LEPSCH R.A. Using dual response surfaces to reduce variability in launch vehicle design: A case study. *Reliability Engineering System Safety*, **91** (4), 407, **2006**.
- BEHBAHANI M., SEO Y. Investigation on haloacetic acid (HAA) degradation by iron powder: Application of response surface methodology. *Process Safety & Environmental Protection*, **103**, 153, **2016**.
- FANG Y., FEI J., MA K. Model reference adaptive sliding mode control using RBF neural network for active power

- filter. *International Journal of Electrical Power & Energy Systems*, **73**, 249, **2015**.
15. TIEN C.M.T., NGO-CONG D., MAI-DUY N., TRAN C.D., TRAN-CONG T. High-order fluid solver based on a combined compact integrated RBF approximation and its fluid structure interaction applications. *Computers & Fluids*, **131**, 151, **2016**.
 16. LIU Y., ZHANG W., JIANG Y., et al. A high-order finite volume method on unstructured grids using RBF reconstruction. *Computers & Mathematics with Applications*, **72** (4), 1096, **2016**.
 17. KIM J.S., JUNG S. Implementation of the RBF neural chip with the back-propagation algorithm for on-line learning. *Applied Soft Computing*, **29** (C), 233, **2015**.
 18. CHAMOLI S. ANN and RSM approach for modeling and optimization of designing parameters for a V down perforated baffle roughened rectangular channel. *AEJ - Alexandria Engineering Journal*, **30** (3), 429, **2015**.
 19. WANG Z., KLEBANOV N., GEORGAKIS C. DRSM Model for the Optimization and Control of Batch Processes. *IFAC-PapersOnLine*, **49** (7), 55, **2016**.
 20. FUSELIER E.J., SHANKAR V., WRIGHT G.B. A high-order radial basis function (RBF) Leray projection method for the solution of the incompressible unsteady Stokes equations. *Computers & Fluids*, **128**, 41, **2016**.
 21. BENKHALDOUN F., HALASSI A., OUAZAR D., et al. Slope limiters for radial basis functions applied to conservation laws with discontinuous flux function. *Engineering Analysis with Boundary Elements*, **66**, 49, **2016**.
 22. BAGHERI M., MIRBAGHERI S.A., EHTESHAMI M., et al. Modeling of a sequencing batch reactor treating municipal wastewater using multi-layer perceptron and radial basis function artificial neural networks. *Process Safety & Environmental Protection*, **93**, 111, **2014**.
 23. FRANCÉS A.P., LUBCZYNSKI M.W., ROY J., SANTOS F.A.M., ARDEKANI M.R.M. Hydrogeophysics and remote sensing for the design of hydrogeological conceptual models in hard rocks – Sardón catchment (Spain). *Journal of Applied Geophysics*, **110** (110), 63, **2014**.
 24. CAMP M.V., WALRAEVENS K. Pumping test interpretation by combination of Latin hypercube parameter sampling and analytical models. *Computers & Geosciences*, **35** (10), 2065, **2009**.
 25. SALLABERRY C.J., HELTON J.C., HORA S.C. Extension of Latin hypercube samples with correlated variables. *Reliability Engineering System Safety*, **93** (7), 1047, **2008**.
 26. WU Q., WANG X., SHEN Q. Research on dynamic modeling and simulation of axial-flow pumping system based on RBF neural network. *Neurocomputing*, **186**, 200, **2016**.
 27. FANG Y., FEI J., MA K. Model reference adaptive sliding mode control using RBF neural network for active power filter. *International Journal of Electrical Power & Energy Systems*, **73**, 249, **2015**.