Stochastic simulation of seawater intrusion in Longkou area, China

Yue Fan (fanyue678@163.com)
Changjiang River Scientific Research Institute https://orcid.org/0000-0002-3480-7166
Qinghua Wu
Changjiang River Scientific Research Institute
Haodong Cui
Changjiang River Scientific Research Institute
Wenxi Lu
Jilin University
Wanli Ren
Jilin University https://orcid.org/0000-0001-5213-9918

Research Article

Keywords: Seawater intrusion, Numerical simulation, Uncertainty analysis, Support vector regression, Surrogate model

Posted Date: May 9th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1571329/v1

Abstract

Seawater intrusion is a common groundwater pollution problem, which has a great impact on ecological environment and economic development. In this paper, a numerical simulation model of variable density groundwater was constructed to simulate and predict the future seawater intrusion in Longkou city, Shandong Province of China. The influence of the sensitive parameter uncertainty of the model on the simulation results was evaluated by using the Monte Carlo method. In order to reduce the computational load from repeatedly calling the simulation model, the surrogate model was established by using the Support Vector Regression (SVR) method. The research results indicate that the SVR surrogate model can fit the input-output relationship of the simulation model with high accuracy. The seawater intrusion in the Longkou area will gradually aggravate at a slow rate, and the area of seawater intrusion is estimated to be 67.45 to 77.07 km^2 after 30 years with 80% confidence.

1. Introduction

The coastal areas are economically developed and densely populated. About half of the world's population lives within 200 km of the coastline (Creel, 2003; Sreekanth and Datta, 2015). The numerous production activities and dense population create a large demand for groundwater in coastal areas, which leads to excessive exploitation of groundwater and causes seawater intrusion.

The numerical simulation is a good way to predict the development and variation of seawater intrusion. The basic theories of the seawater intrusion numerical simulation can be divided into two categories. One is the abrupt interface hypothesis and the other is the transition zone theory. The abrupt interface hypothesis is based on the static water balance between salt water and fresh water, assuming that seawater and freshwater are immiscible with each other and there is a mutation interface between them. This theory originated in the late 19th century (Ghyben, 1888; Herzberg, 1901), and has been widely adopted in the following century (Mercer et al. 1980; Essaid, 1990; Cheng et al. 2000; Shi et al. 2011).

The model based on the abrupt interface hypothesis ignores the hydrodynamic dispersion between seawater and freshwater, and the accuracy of the model is insufficient. In fact, seawater and freshwater can be miscible in any ratio, and the salt water and freshwater interface exists in the form of a transition zone. Due to the high salinity of seawater, the change of groundwater density caused by the change of solute concentration in the transition zone is not negligible. Therefore, since the 1990s, variable-density seawater intrusion numerical simulation based on the transition zone theory has gradually become the mainstream. Huyakom et al. (1987) proposed the flow equation and solute transport equation by considering the density variation factor based on the transition zone theory. Putti and Paniconi (1995) solved the variable-density water flow model and the solute transport model, and simulated seawater intrusion in southern Italy using three-dimensional finite element method. Langevin (2003) used SEAWAT code to estimate the exchange volume of submarine groundwater in the estuary of Biscayne Bay, Florida. Lin et al. (2009) established a numerical model to investigate the extension of seawater intrusion in the Gulf Coast of Alabama, USA. By coupling the simulation model and the optimization model, Dentoni et al. (2015) used simulation-optimization method to evaluate strategies for managing groundwater resources under natural and man-made pressure. Zhao et al. (2016) studied the variation of seawater intrusion in Dalian city of China by the three-dimensional variable density groundwater numerical simulation model. Fan et al. (2020a) proposed a multi-objective optimization model for groundwater exploitation in coastal areas based on the simulation model of seawater intrusion.

However, for a long period of time, the numerical simulation of seawater intrusion was performed based on certain conditions. When using the deterministic method to predict seawater intrusion, the parameter uncertainty was not taken into account, so the reliability of the prediction results need to be further evaluated. Therefore, it is necessary to study the influence of parameter uncertainty on the output results for the accurate prediction and early warning of seawater intrusion.

The Monte Carlo method is a common used method for analyzing model uncertainty (Jiang et al. 2017; Yan et al. 2019). The principle is to make thousands of calculations of simulation models under different parameter conditions and then perform a statistical analysis of the output results. Based on transition zone theory, the simulation of seawater intrusion with the three-

dimensional variable density groundwater numerical simulation model requires the coupling iterative calculation of flow model and solute transport model. Thus, the single operation time is longer than that of conventional groundwater simulation model. Once the Monte Carlo simulation is conducted, the large number of calls of simulation model will generate huge computational load and lengthy calculation time, which greatly limits the application of this method. Establishing the surrogate model of simulation model is an effective method to solve this problem. The support vector regression (SVR) is a neural network method with good performance. It has been widely used in many fields such as geology, biomedicine, mechanical manufacturing and so on (Wang et al. 2015; Kang et al. 2016; Wang et al. 2018). In this study, a surrogate model for the 3D variable density seawater intrusion simulation model was established by the SVR method, and the surrogate model was directly invoked in a large number of Monte Carlo simulations to shorten the computation time.

This paper took Longkou City, Shandong Province as an example, predicted the future seawater intrusion situation by establishing a three-dimensional variable density seawater intrusion numerical simulation model in the study area. The Monte Carlo method was used to consider the impact of the random change of sensitive parameters in the simulation model on the prediction results of seawater intrusion. In order to reduce the computational load caused by repeatedly invoking the simulation model, the SVR method was used to establish the surrogate model of the simulation model, and the surrogate model will be invoked directly in a large number of Monte Carlo simulation processes. Finally, the Monte Carlo simulation results were statistically analyzed to estimate the future seawater intrusion area under different confidence levels. This study provided more choice basis for the prediction of seawater intrusion.

2. Construction Of The Variable-density Seawater Intrusion Simulation Model

The sediments near the coastline and ocean surface sediments in the Longkou area of China have good permeability, so there is a good hydraulic connection between seawater and coastal aquifers, which provides favorable conditions for large-scale seawater intrusion. In the past 40 years, the rapid development of industrial and agricultural production in this region has caused a surge in water consumption, which has accelerated the overexploitation of groundwater resources and caused a drop in regional groundwater levels. So far, a large area of seawater intrusion has been formed. The seawater intrusion caused salty groundwater and aggravated the water supply crisis, which has become a bottleneck restricting industrial and agricultural production and urban development in this area.

The study area in this paper is located in the northwestern part of Longkou City (showed in Fig. 1), with a length of about 20.7 km from east to west, a width of about 18.8 km from north to south, and a total area of about 221 km^2. The area borders the Bohai Sea in the west and north, and the terrain is high in the southeast and low in the northwest.

2.1 Generalization of study area conditions

The target aquifer of this study is loose rock pore phreatic aquifer, which medium is mainly coarse sand and medium sand, with a small amount of gravel and pebbles, and a small amount of cohesive soil. Generally, it has good water permeability and water richness and belongs to medium -strong rich water layer. The lower is a thin layer of silt layer with weak permeability. The bottom is sandstone with extremely weak water permeability. In general, the aquifer was generalized as a heterogeneous anisotropic aquifer, and the water flow was generalized as a three-dimensional unsteady flow considering variable density. The schematic diagrams of the plane and section of the hydrogeological conceptual model were shown in Figs. 2 and 3.

In terms of boundary conditions, the northern and western boundary γ_1 of the study area is the Bohai Bay, which was generalized as the known head boundary. The eastern boundary γ_2 is the watershed between the Yellow River Basin and the Yongwen River Basin, which is generalized as the zero flux boundary. The southwestern boundary γ_4 is the watershed dividing the Balisha River basin from the Jiehe basin and is generalized to a zero flux boundary. The southern boundary γ_3 is the boundary between the plain area and the mountainous area, generalized as the lateral runoff recharge boundary. The top boundary of the aquifer is the phreatic surface, and the bottom boundary is the aquifer floor.

In terms of source and sink items, the main recharge items of groundwater include precipitation infiltration recharge, groundwater lateral runoff recharge, and river leakage recharge. The main discharge items include groundwater evaporation and groundwater extraction.

2.2 Construction of the numerical simulation model

For the mathematical expression of the numerical simulation model of variable density seawater intrusion, see Fan et al. (2020b). In this study, the simulation model was solved using the SEAWAT program written by the USGS. The study area was dissected horizontally on a 100m x 100m grid into 22,176 grids and vertically into three layers, for a total of 66,528 effective grids. The spatial discretization of the study area grid is shown in Fig. 4.

The model was calibrated and verificated using actual measured water level and water quality data from the study area. The location of the monitoring point was shown in Fig. 2. Using 2015.1 as the initial moment of the simulation, water level data were used from 2016.1 and 2017.1, and water quality data were used from 2018.8 and 2020.4. As can be seen from Fig. 5, the data fit well, indicating that the model constructed in this study can be used to reflect the actual groundwater movement patterns in the study area. The values of the parameters taken after calibration and verification are shown in Table 1.

Table 1
Summary of hydrogeological parameters in the simulation model

Zone	Description	Hydraulic conductivity K(m/d)	coefficient of precipitation recharge	Porosity	specific yield	storage coefficient (m^{-1})	Longitudinal dispersion(m)	Horizontal and vertical dispersion(m)
1	Coastal sediments in the west	15.50	0.08	0.30	0.12	0.00013	62.60	6.30
2	Coastal sediments in the north	24.80	0.12	0.30	0.15	0.00005	61.00	6.10
3	Middle reaches of the coastal basin	35.70	0.14	0.32	0.16	0.00006	68.00	6.80
4	Sediments in the central region	15.30	0.13	0.30	0.12	0.0002	62.60	6.30
5	Middle reaches of the Yongwen River basin	12.50	0.15	0.25	0.10	0.00014	56.00	5.60
6	Longkou city center	20.60	0.13	0.27	0.13	0.00018	58.60	5.90
7	Front hillside area of coastal small rivers basin	24.50	0.10	0.30	0.15	0.00005	61.00	6.10
8	Front hillside area of the Yongwen River basin	9.50	0.15	0.25	0.10	0.00015	54.50	5.50
C2	Silt or clay layer	0.25	—	0.40	0.007	0.01	60.00	6.00
C3	Sandstone layer	0.00086	—	0.09	0.02	3.3×10^{-6}	60.00	6.00

2.3 Model prediction

The model was used to make prediction of the future seawater intrusion. The forecast period of the model was set from January 2022 to January 2052. The precipitation in the future study area was based on the multi-year average value of 613.3mm/year. The groundwater extraction volume adopted the average value of the groundwater extraction volume in the region from 2016 to 2020.

Each of these parameters was input into the seawater intrusion simulation model and the model was run to make predictions. The seawater intrusion area was statistically analysed for the next 10 years (up to January 2032), 20 years (up to January 2042) and 30 years (up to January 2052). The results of the modelling projections are shown in Figs. 6–8 and the seawater intrusion area statistics are shown in Table 2.

Table 2
Statistical table of future changes in the area of seawater intrusion predicted by the model

Simulation time	Seawater intrusion area(km^2)	Increase of invasion area compared to January 2021(km^2)	Percentage of invasion area compared to 2021(%)
Present situation	71.78	0	0
2032.1	72.39	0.61	0.85
2042.1	72.98	1.20	1.67
2052.1	73.74	1.96	2.73

In this study, a concentration of chloride ions in groundwater greater than 250 mg/L was set as a marker for the occurrence of seawater intrusion. The spatial distribution of seawater intrusion shows that the area of seawater intrusion in the northern coastal area is gradually decreasing, while the line of seawater intrusion in the western coastal area is advancing inland year by year. The reason for this should be that the northern coastal area has less groundwater extraction and has the Yongwen River as a stable surface runoff recharge, so the area of seawater intrusion will gradually decrease over time. In contrast, groundwater extraction is more concentrated in the western coastal area, and there is a stable falling funnel in the middle reaches of the small coastal river basins, where the groundwater level is always below sea level, so seawater intrusion in the western coastal area will gradually increase. The overall situation of the whole study area shows that the area of seawater intrusion in the region will continue to increase at a slow rate in the future.

The numerical simulation model of seawater intrusion developed above is a deterministic model that does not contain any stochastic component and can only obtain unique prediction results. Under the influence of global climate change, the future sea level rise height and precipitation are highly stochastic, and it is difficult to assess the reliability of the prediction results if a deterministic approach is adopted. Therefore, there is a need to conduct research on the effect of uncertainty in the prediction results of seawater intrusion simulations due to stochastic changes in sensitive factors in simulation models.

3. Support Vector Regression Surrogate Model

The numerical simulation model of variable density groundwater takes a long time to solve. The large computational load when using Monte Carlo methods for uncertainty analysis of seawater intrusion predictions limits the efficiency of our research problem to some extent. The advent of the surrogate model has greatly alleviated these problems. The surrogate model is a data-driven model that can obtain input-output relationships similar to those of the simulation model with a smaller computational effort. In a large number of Monte Carlo experiments, the surrogate model can be invoked directly without the need to compute the simulation model extensively, which can greatly reduce the computational load and calculation time. (Hou and Lu, 2018).

In this study, the SVR method was used to develop the surrogate model for numerical simulation of variable density seawater intrusion. The core idea and theory of the method is to map the input data into a high-dimensional space through a non-linear mapping function and perform linear regression analysis in the high-dimensional space (Ouyang et al., 2017; Liu et al., 2019).

Assuming that the training sample set is, ,then the SVR equation can be expressed in the following form.

$$f(x) = \omega \cdot \varphi(x) + b$$

1

In the formula, $\varphi(x)$ refers to the nonlinear mapping function that maps the input variable to the high-dimensional feature space, ω is the weight vector, and b represents the constant term.

The $\varepsilon -$ insensitive function was used as the error function. The mathematical expression was shown below.

$$L_{\varepsilon}\left(y_i\right)=\begin{cases}0 & \left|y_i-(\omega\cdot\varphi(x)+b)\right|\leqslant\varepsilon\\ \left|y_i-(\omega\cdot\varphi(x)+b)\right|-\varepsilon & \left|y_i-(\omega\cdot\varphi(x)+b)\right|>\varepsilon\end{cases}$$

2

Set an insensitive zone with a width of 2ε in the error function, which is called \varepsilon zone. When the error is less than \varepsilon, the error can be ignored. But when the error is greater than \varepsilon, the value of the error function is the actual error minus \varepsilon.

Assuming that all training data falls within the \varepsilon band, the following formula holds.

\begin{gathered} \hbox{min} \ \frac{1}{2}\left\| {\omega ^2} \right\| \hfill \\ s.t.\left\{ {\begin{array}{*{20}{l}} {\left({\omega \cdot \varphi \left({{x_i}} \right)} \right)+b - {y_i} \leqslant \varepsilon } \\ {{y_i} - \left({\omega \cdot \varphi \left({{x_i}} \right)} \right) - b \leqslant \varepsilon } \end{array}} \right.{\text{ }}i=1,2, \cdots ,n \hfill \\ \end{gathered}
3

Since it is impossible for all the sample points to fall in the \varepsilon zone, two slack variables were introduced, and the formula (6) became the following form.

\begin{gathered} \hbox{min} \ \left\{ {\frac{1}{2}{{\left\| \omega \right\|}^2}+C\sum\limits_{{i=1}}^{n} {\left({{\xi _i}+\xi _{i}^{*}} \right)} } \right\} \hfill \\ s.t.\left\{ {\begin{array}{*{20}{l}} {\left({\omega \cdot \varphi \left({{x_i}} \right)} \right)+b - {y_i} \leqslant \varepsilon +{\xi _i}} \\ {{y_i} - \left({\omega \cdot \varphi \left({{x_i}} \right)} \right) - b \leqslant \varepsilon +\xi _{i}^{*}} \\ {{\xi _i},\xi _{i}^{*} \geqslant 0} \end{array}} \right. \hfill \\ \end{gathered}
4

In the formula, C>0 represents the degree of punishment for samples that exceed the error, and {\xi _i} and \xi _{i}^{*} are the upper and lower limits of the slack variable.

Construct a Lagrangian function.

\begin{gathered} L=\frac{1}{2}{\left\| \omega \right\|^2}+C\sum\limits_{{i=1}}^{n} {\left({{\xi _i}+\xi _{i}^{*}} \right) - \sum\limits_{{i=1}}^{n} {{\alpha _i}\left[{\varepsilon +{\xi _i}+{y_i} - \left({\omega \cdot \varphi \left({{x_i}} \right)} \right)+b} \right]\xi } } \hfill \\ {\text{ }}=\sum\limits_{{i=1}}^{n} {\alpha _{i}^{*}\left[{\varepsilon +\xi _{i}^{*} - {y_i}} \right]+} \left({\omega \cdot \varphi \left({{x_i}} \right)} \right)+b - \sum\limits_{{i=1}}^{n} {\left({{\eta _i}{\xi _i}+\eta _{i}^{*}\xi _{i}^{*}} \right)} \hfill \\ \end{gathered}
5

L is the Lagrangian operator, and {\eta _i},\eta _{i}^{*},{\alpha _i},\alpha _{i}^{*} are all the Lagrangian multiplier greater than or equal to 0. For the optimal solution, {\partial _b}L,{\partial _w}L,{\partial _{{\xi _i}}}L,{\partial _{\xi _{i}^{*}}}L are all 0.

{\partial _b}L=\sum\limits_{{i=1}}^{m} {\left({{\alpha _i} - \alpha _{i}^{*}} \right)} =0
6
{\partial _w}L=\omega - \sum\limits_{{i=1}}^{m} {\left({{\alpha _i} - \alpha _{i}^{*}} \right)} {x_i}{\text{=}}0
7
{\partial _{{\xi _i}}}L=C - {\alpha _i} - {\eta _i}=0
8
{\partial _{\xi _{i}^{*}}}L=C - \alpha _{i}^{*} - \eta _{i}^{*}=0
9

Substituting equations (6)–(9) into Eq. (5), the optimization problem of Eq. (4) can be rewritten as a dual form.

$$\begin{gathered} \hbox{min} \left\{ {\begin{array}{*{20}{l}} { - \frac{1}{2}\sum\limits_{{i,j=1}}^{m} {\left({{\alpha _i} - {\alpha _i}^{*}} \right)\left({{\alpha _j} - {\alpha _j}^{*}} \right)\left({\varphi \left({{x_i}} \right),\varphi \left({{x_j}} \right)} \right)} } \\ { - \varepsilon \sum\limits_{{i=1}}^{m} {{y_i}\left({{\alpha _i} - {\alpha _i}^{*}} \right)} } \end{array}} \right. \hfill \\ s.t\left\{ {\begin{array}{*{20}{l}} {\sum\limits_{{i=1}}^{m} {\left({{\alpha _i} - {\alpha _i}^{*}} \right)=0} } \\ {{\alpha _i},{\alpha _i}^{*} \in \left[{0,C} \right]} \end{array}} \right. \hfill \\ \end{gathered}$$
10

Combining Eq. (7) can get the regression function.

$$f\left(x \right)=\sum\limits_{{i=1}}^{n} {\left({{\alpha _i} - \alpha _{i}^{*}} \right)} \left({\varphi \left({{x_i}} \right)\cdot \varphi \left({{x_j}} \right)} \right)+b$$
11

Since the inner product is actually a Mercer core, the following formula holds.

$$K\left({{x_i},{x_j}} \right)=\left({\varphi \left({{x_i}} \right)\cdot \varphi \left({{x_j}} \right)} \right)$$
12

Substituting it into Eq. (11) can get the following equation.

$$f\left(x \right)=\sum\limits_{{i=1}}^{n} {\left({{\alpha _i} - \alpha _{i}^{*}} \right)K\left({x,{x_i}} \right)+b}$$
13

Among them, $K\left({x,{x_i}} \right)$ is the kernel function, and the Gaussian kernel function was used in this study.

$$K\left({x,{x_i}} \right){\text{=}}\exp \left({ - \frac{{{{\left\| {x - {x_i}} \right\|}^2}}}{{2{\sigma ^2}}}} \right)$$
14

Based on the above theory, the code of SVR surrogate model was written in MATLAB software.

4. Uncertainty Analysis Of Seawater Intrusion Simulation

4.1 Construction of the surrogate models

To consider the effect of uncertainty of sensitive factors in the model on the prediction results of seawater intrusion, the more sensitive factors in the model need to be selected first. According to the results of Fan et al. (2020b), the more sensitive parameters in the model are groundwater extraction and precipitation, followed by hydraulic conductivity, while all other parameters are less sensitive. Combined with the impact of future climate change on sea level rise height, this study will consider the impact of the uncertainty of stochastic changes in three factors, namely precipitation, groundwater extraction and sea level rise height, on the simulated prediction of seawater intrusion.

Table 3 showed the value ranges of precipitation, groundwater extraction, and sea level rise selected for this study. The data on the future sea level rise comes from “China Sea Level Bulletin (2020)”. Within the parameter value range, we input 240 sets of sampling data into the simulation model by using the Latin hypercube method, and output the chloride ion concentration in the three typical observation wells (showed in Fig. 9) and the seawater intrusion area, forming the input-output data set. Among them, 200 sets of data were selected for the training of the surrogate model, and 40 sets of data were selected for the test of the surrogate model accuracy.

Table 3
Sampling range of sensitive parameters in the model

Variable	Mean value	Sampling frame
Precipitation (mm/a)	613	(492, 734)
Groundwater exploration($\times 10^6$ m^3/a)	14.55	(11.64, 17.46)
Height of sea level rise (mm)	115	(51, 179)

In this study, two parameters, Correlation Coefficient (R^2) and Mean Relative Error (MRE), were used to characterise the extent to which the output of the surrogate model fitted the simulation model. Each parameter was calculated as follows.

{R^2}=1 - \frac{{\sum\limits_{{i=1}}^{n} {{{({y_i} - {{\hat {y}}_i})}^2}} }}{{\sum\limits_{{i=1}}^{n} {{{({y_i} - \bar {y})}^2}} }}
15
MRE=\frac{1}{n}\frac{{\sum\limits_{{i=1}}^{n} {\left| {{y_i} - {{\hat {y}}_i}} \right|} }}{{{y_i}}}
16
Where *n* is the number of samples, {y_i} represents the output value of the simulation model, {\hat {y}_i} is the output value of the alternative model, and \bar {y} refers to the average value of *n* output samples of the simulation model.

Using 40 sets of test data to verify the accuracy of the surrogate model, the output fit of the simulation model and surrogate model was shown in Fig. 10. The correlation coefficient R^2 reached 0.9957, the MRE was 0.20. This suggests that the data-trained SVR surrogate model can be used to replace the input-output relationship of the simulation model.

4.2 The result of uncertainty analysis

After that, 1000 groups were sampled within the range of values of sensitive factors in the model. The sampling results were fed into the trained surrogate model for a Monte Carlo test and the output of the individual observation wells for chloride concentration and seawater intrusion area was statistically analysed. The histograms of the distribution of chloride ion concentrations for each observation well after 30 years (June 2050) are shown in Figs. 11–13 and the histograms of the seawater intrusion area are shown in Fig. 14, with the statistical indicators shown in Table 4.

Comparing the above chart, it can be seen that with chloride ion concentration exceeding 250mg/L as a sign of seawater intrusion, the probability of seawater intrusion in Ob1 and Ob3 wells was relatively low, slightly greater than 20%, indicating that these two wells are not prone to seawater intrusion. The distribution of concentrations in the Ob1 well is more dispersed and the standard deviation of the 1000 Monte Carlo simulation output concentrations is larger. This indicates that the northern region of the study area is strongly influenced by uncertainties of sea level rise height, precipitation and groundwater extraction. Smaller fluctuations of the three model parameters described above can result in relatively large changes of concentrations in the Ob1 well. The standard deviation of the concentrations in the Ob3 well is smaller than in Ob1, being about one-third of that in the Ob1 well. It indicates that seawater intrusion in the western region of the study area is relatively little affected by the uncertainty of the three sensitive model parameters. In contrast, the minimum concentration in the 1000-group Monte Carlo test for the Ob2 well also exceeds 250 mg/L, which means that the probability of seawater intrusion at the Ob2 well location after 30 years is 100%. At the same time, the concentration distribution of the Ob2 well is more concentrated with a smaller standard deviation, indicating that the variation in concentration in this well is minimally affected by the uncertainty of the model parameters.

Combining the results of the uncertainty analysis of the three observation wells, it can be concluded that the implementation of seawater intrusion prevention and control measures in the northern of the study area will be the most effective and the seawater intrusion will be the least difficult to manage. The western coastal area is the next most difficult. In the central coastal area, a narrow peninsula exists in the northwest of Longkou City, which acts as a buffer against seawater intrusion

and retreat. Therefore, seawater intrusion in the central region will be more stable in the long term and the seawater intrusion will be the most difficult to manage.

Combined with Fig. 15 and Table 4, the area of seawater intrusion varies from 62.18km^2 to 83.25km^2, with a large variation of 21.07 km^2. It shows that the area of seawater intrusion across the region is strongly influenced by the uncertainty of the model's sensitive parameters. The average area of seawater intrusion after 30 years of 1000 sets of Monte Carlo simulations is 73.16 km^2, which is relatively close to the deterministic model predictions. It indicates that the predicted area of seawater intrusion obtained using the deterministic model in Section 2.3 has a high degree of confidence.

Table 4
The statistical indicator of chloride ion concentration in observation wells and the seawater intrusion area

Well ID Statistical indicators	ObW-1(mg/L)	ObW-2(mg/L)	ObW-3(mg/L)	Seawater intrusion area(km^2)
Maximum value	324.85	309.34	274.41	83.25
Minimum value	153.93	280.91	215.78	62.18
Mean value	218.80	293.00	240.19	73.16
Standard deviation	38.96	5.77	12.24	4.62
Probability of seawater intrusion	22.8%	100%	21.5%	—

The intervals that exist for each well concentration and seawater intrusion area at different probabilities based on Chebyshev's inequality were estimated in Table 5. Based on the table, the distribution of intervals for each observation well and seawater intrusion area at different confidence levels can be queried. At 80% confidence level, the seawater intrusion area in the study area for the next 30 years ranges from 67.45 to 77.07 km^2.

Table 5
Estimated intervals of seawater intrusion simulation results with different confidence levels

Well ID Confidence	ObW-1(mg/L)	ObW-2(mg/L)	ObW-3(mg/L)	Seawater intrusion area(km^2)
90%	199.06–238.53	285.40–300.60	229.13–251.25	65.46–79.06
80%	204.84–232.76	287.63–298.37	232.367–248.01	67.45–77.07
60%	208.93–228.67	289.20–296.80	234.66–245.72	68.86–75.66

4.3 Computational advantages of the surrogate model

This paper analyzed the calculation load of different methods. In the uncertainty analysis process of the numerical simulation of seawater intrusion, the computer used a PC equipped with an Intel i5 3.2 GHz processor and 8 GB RAM. It took an average of 3 minutes to run the SEAWAT program to solve the simulation model of seawater intrusion. If we directly use the simulation model for uncertainty analysis, it need to calculate the simulation model for 1000 times, which would take 50 hours in total. In this study, the surrogate model was used to replace the simulation model for uncertainty analysis. In the process of training and verifying the surrogate model, the simulation model need to run for 120 times to get the input-output data, and the total time was 6 hours. The calculation of the surrogate model took about 2-3s, which can be ignored. 1,000 sets of Monte Carlo experiments performed by using the surrogate model, saving 88% of the calculation time in total.

As increase of random trials number and the simulation model complexity, the application of surrogate models will save more time. The above research proves that when using the Monte Carlo method for uncertainty analysis, by adopting the surrogate model to replace the simulation model, the calculation time can be greatly reduced while ensuring the accuracy of the simulation.

5. Conclusion

The paper took Longkou city of China as an example study area, comprehensively applying multiple methods such as three-dimensional variable density seawater numerical simulation model, the surrogate model and the Monte Carlo method, to stochastically simulate the future seawater intrusion. The main results were concluded as follow.

1. In the forecast of future seawater intrusion, the area of seawater intrusion in the study area is expected to increase by 2.73% after 30 years, with an average increase of 0.065 km^2 per year. The seawater intrusion will gradually increase in the western coastal areas and decrease in the northern coastal areas.
2. The surrogate model of the 3D variable density seawater intrusion numerical simulation model by using the SVR method has a high accuracy. In stochastic simulations of seawater intrusion, the direct use of surrogate models for calculations can effectively reduce the computational load and improve computing efficiency.
3. Adopting the Monte Carlo method to consider the influence of random changes of sensitive factors on the simulation results, with 80% confidence, the seawater intrusion area of Longkou city was estimated to be 67.45 ~ 77.07 km^2 after 30 years.

Declarations

Acknowledgments

The authors acknowledge support provided by the Fundamental Research Funds for Central Public Welfare Research Institutes (Grant Nos. CKSF2021488/YT, CKSF2021485+YT) and the National Nature Science Foundation of China (Grant Nos.42072282, 41902260). Special gratitude is extended to the journal editors for their efforts in evaluating this study. The valuable comments provided by the anonymous reviewers are also gratefully acknowledged.

Authors contributions

Yue Fan: Conceptualization, Methodology, Software, Writing - Original Draft. Qinhua Wu: Simulation model building. Haodong Cui: Validation, Formal Analysis, Visualization. Wenxi Lu: Methodology. Wanli Ren: Conceptualization, Writing - Review & Editing, Supervision, Project Administration.

Funding

See the Acknowledgments section

Competing interest

None

Availability of data and materials

Available from the corresponding author upon reasonable request.

Code Availability

Available upon reasonable request.

Ethics Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent to Participate

Not applicable.

Consent for Publication

Not applicable

References

1. Cheng AD, Halhal D, Naji A, Ouazar D (2000) Pumping optimization in saltwater-intruded coastal aquifers.Water Resour Res 36(8):2155–2165. http://dx.doi.org/10.1029/2000WR900149
2. Creel L (2003) Ripple effects: population and coastal regions. Population reference bureau, Washington DC, pp 1–7
3. Dentoni M, Deidda R, Paniconi C, Qahman K, Lecca G (2015) A simulation-optimization study to assess seawater intrusion management strategies for the Gaza Strip coastal aquifer (Palestine). Hydrogeol J 23(2):249–264. http://dx.doi.org/10.1007/s10040-014-1214-1
4. Essaid HI (1990) A multilayered sharp interface model of coupled freshwater and saltwater flow in coastal systems: Model development and application. Water Resour Res 26(7):1431–1454. http://dx.doi.org/10.1029/WR026i007p01431
5. Fan Y, Lu W, Miao T, Li J, Lin J (2020a) Multiobjective optimization of the groundwater exploitation layout in coastal areas based on multiple surrogate models. Environ Sci Pollut Res 27(16):19561–19576. http://dx.doi.org/10.1007/s11356-020-08367-2
6. Fan Y, Lu W, Miao T, Li J, Lin J (2020b) Optimum design of a seawater intrusion monitoring scheme based on the image quality assessment method. Water Resour Manag 34(8):2485–2502. http://dx.doi.org/10.1007/s11269-020-02565-w
7. Ghyben BW (1888) Nota in verband met de voorgenomen putboring nabij, Amsterdam. The Hague, 21
8. Herzberg A (1901) Die wasserversorgung einiger Nordseebader. J Gasbeleucht Wasserversorg 44:842–844
9. Hou Z, Lu W (2018) Comparative study of surrogate models for groundwater contamination source identification at DNAPL-contaminated sites. Hydrogeol J 26(3):923–932. http://dx.doi.org/10.1007/s10040-017-1690-1
10. Huyakorn PS, Andersen PF, Mercer JW, White JHO (1987) Saltwater intrusion in aquifers: Development and testing of a three-dimensional finite element model. Water Resour Res 23(2):293–312. http://dx.doi.org/10.1029/WR023i002p00293
11. Jiang X, Na J, Lu W, Zhang Y (2017) Coupled Monte Carlo simulation and Copula theory for uncertainty analysis of multiphase flow simulation models. Environ Sci Pollut Res 24(31):24284–24296. http://dx.doi.org/10.1007/s11356-017-0030-2
12. Kang F, Xu Q, Li J (2016) Slope reliability analysis using surrogate models via new support vector machines with swarm intelligence. Appl Math Model 40(11–12):6105–6120. http://dx.doi.org/10.1016/j.apm.2016.01.050
13. Langevin CD (2003) Simulation of submarine ground water discharge to a marine estuary: Biscayne Bay, Florida. Groundwater 41(6):758–771. http://dx.doi.org/10.1111/j.1745-6584.2003.tb02417.x
14. Lin J, Snodsmith JB, Zheng C, Wu J (2009) A modeling study of seawater intrusion in Alabama Gulf Coast, USA. Environ Geol 57(1):119–130. http://dx.doi.org/10.1007/s00254-008-1288-y
15. Liu C, Hu Y, Yu T, Xu Q, Liu C, Li X, Shen C (2019) Optimizing the water treatment design and Management of the Artificial Lake with water quality modeling and surrogate-based approach. Water 11(2):391. http://dx.doi.org/10.3390/w11020391
16. Mercer JW, Larson SP, Faust CR (1980) Simulation of salt-water interface motion. Groundwater 18(4):374–385. http://dx.doi.org/10.1111/j.1745-6584.1980.tb03412.x

17. Ouyang Q, Lu W, Hou Z, Zhang Y, Li S, Luo J (2017) Chance-constrained multi-objective optimization of groundwater remediation design at DNAPLs-contaminated sites using a multi-algorithm genetically adaptive method. J Contam Hydrol 200:15–23. http://dx.doi.org/10.1016/j.jconhyd.2017.03.004
18. Putti M, Paniconi C (1995) Finite Element Modeling of Saltwater Intrusion Problems with an Application to an Italian Aquifer. Advanced Methods for Groundwater Pollution Control. Springer, Vienna, pp 65–84
19. Shi L, Cui L, Park N, Huyakorn PS (2011) Applicability of a sharp-interface model for estimating steady-state salinity at pumping wells-validation against sand tank experiments. J Contam Hydrol 124(1–4):35–42. http://dx.doi.org/10.1016/j.jconhyd.2011.01.005
20. Sreekanth J, Datta B (2015) Simulation-optimization models for the management and monitoring of coastal aquifers. Hydrogeol J 23(6):1155–1166. http://dx.doi.org/10.1007/s10040-015-1272-z
21. Wang C, Zhang J, Zhou J, Alting SA (2015) Prediction of film-cooling effectiveness based on support vector machine. Appl Therm Eng 84:82–93. http://dx.doi.org/10.1016/j.applthermaleng.2015.03.024
22. Wang Y, Meng X, Zhu L (2018) Cell group recognition method based on adaptive mutation PSO-SVM. Cells 7(9):135. http://dx.doi.org/10.3390/cells7090135
23. Yan X, Dong W, An Y, Lu W (2019) A Bayesian-based integrated approach for identifying groundwater contamination sources. J Hydrol 579:124160. http://dx.doi.org/10.1016/j.jhydrol.2019.124160
24. Zhao J, Lin J, Wu J, Yang Y, Wu J (2016) Numerical modeling of seawater intrusion in Zhoushuizi district of Dalian City in northern China. Environ Earth Sci 75(9):1–18. http://dx.doi.org/10.1007/s12665-016-5606-5

Figures

Figure 1

Administrative division of Longkou City and the study area location

Figure 2

Plan view of the conceptual hydrogeological model of the study area

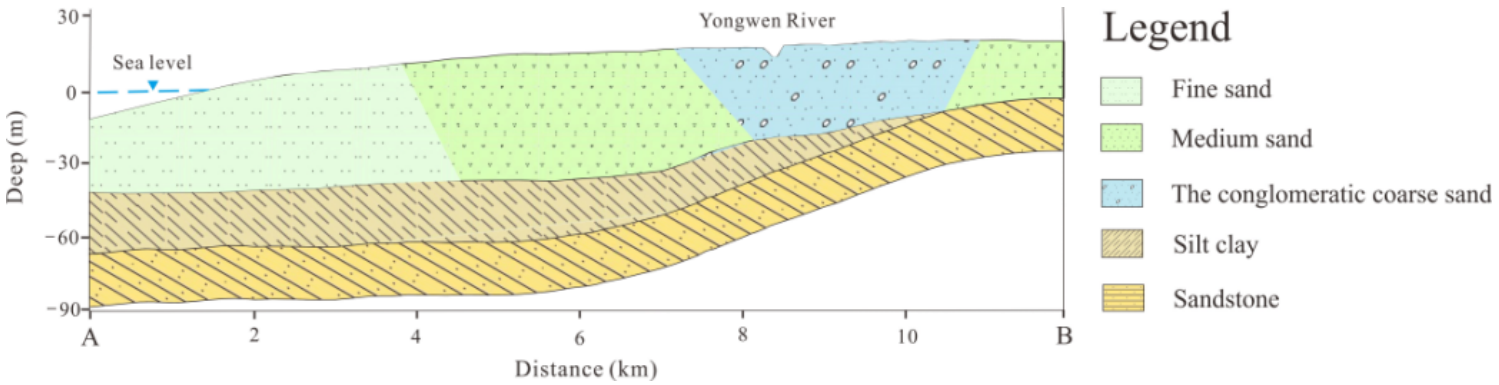

Figure 3

The section diagram of conceptual hydrogeological model of the study area

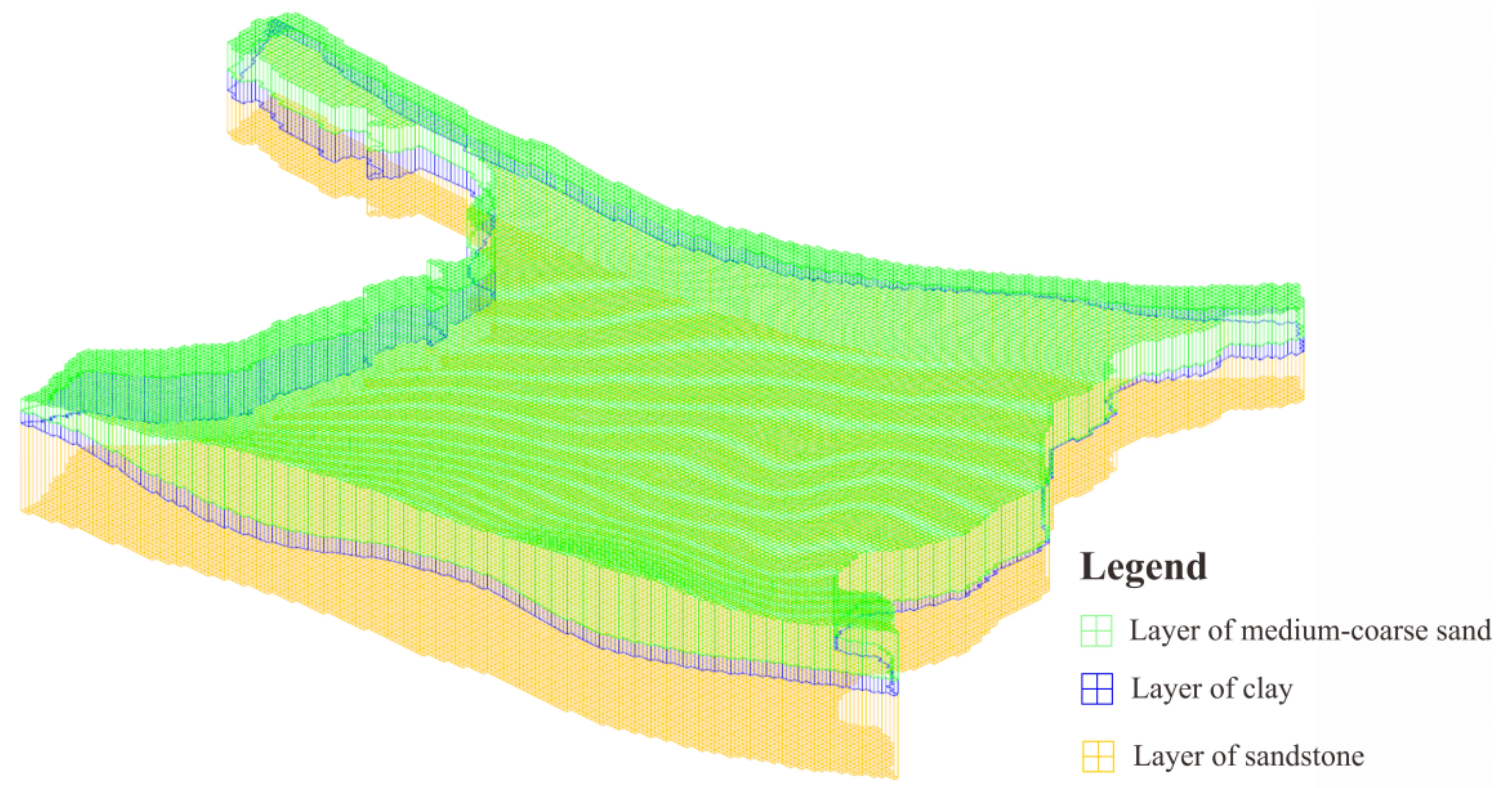

Figure 4

The space discrete diagram of the simulation model

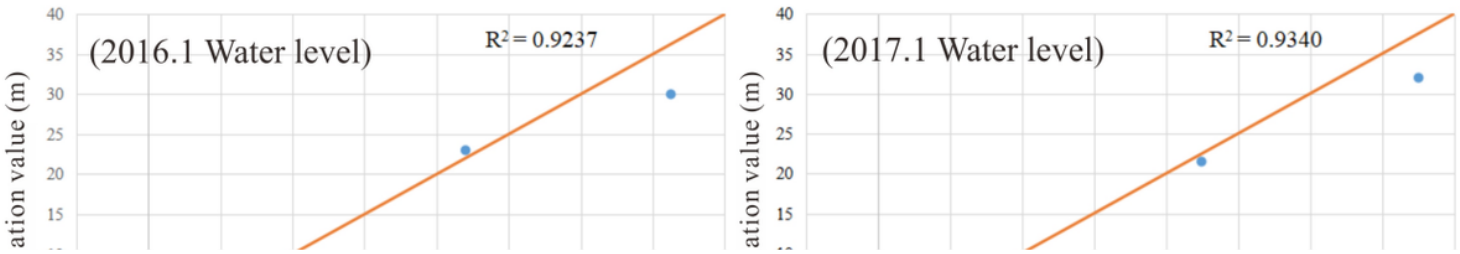

Figure 5

Fitting diagram between simulated data and measured data in calibration and verification period

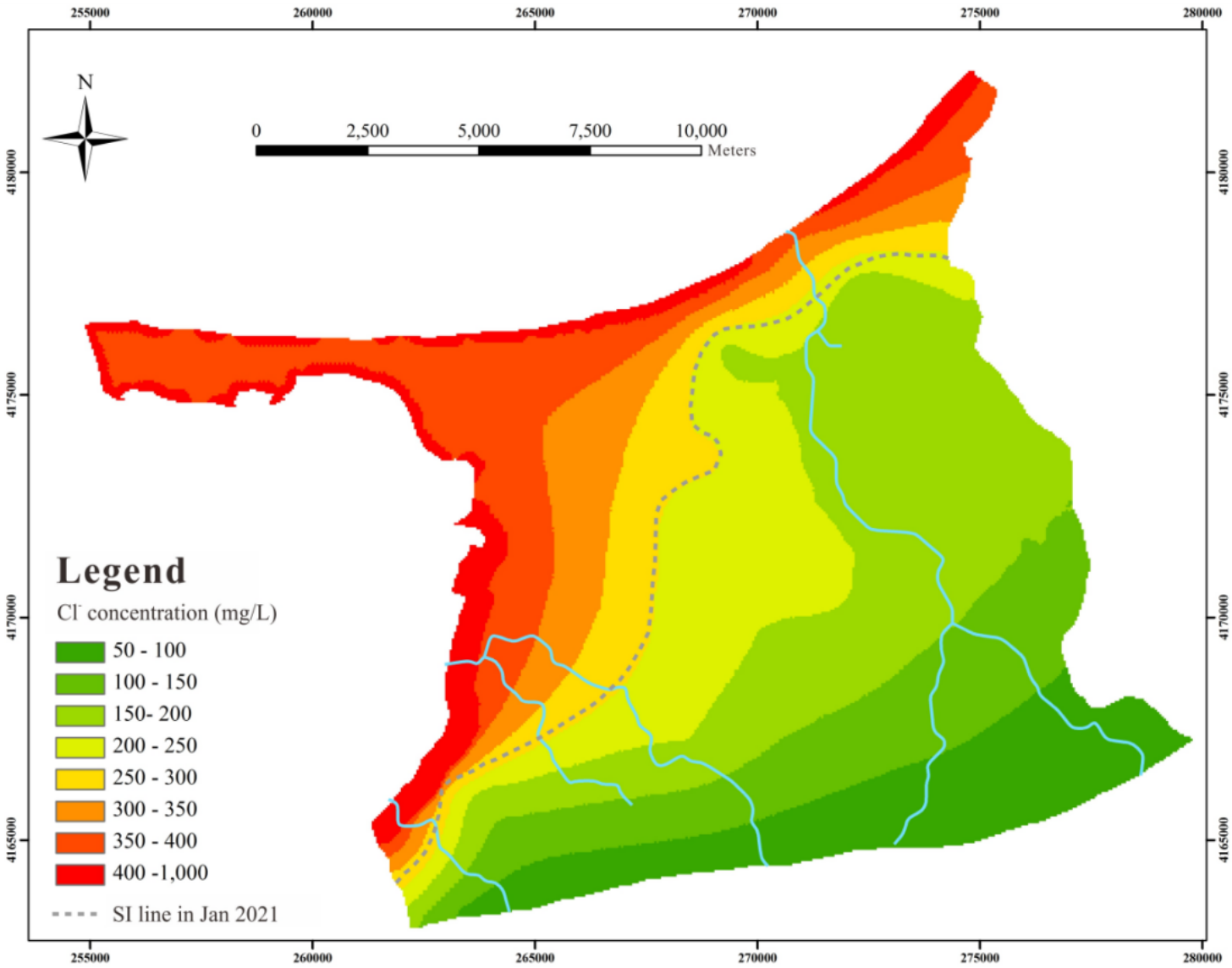

Figure 6

Model projections of seawater intrusion distribution in January 2032

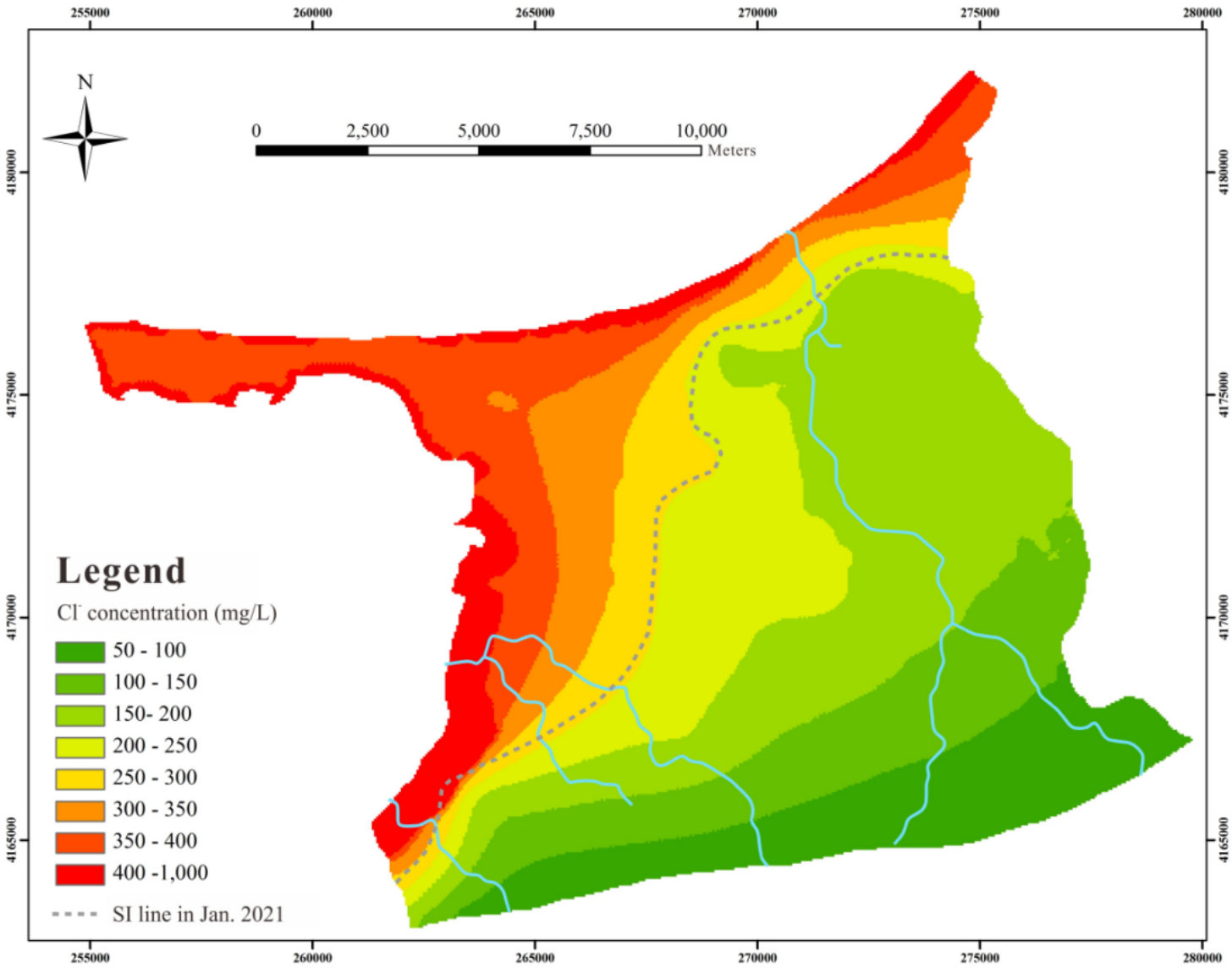

Figure 7

Model projections of seawater intrusion distribution in January 2042

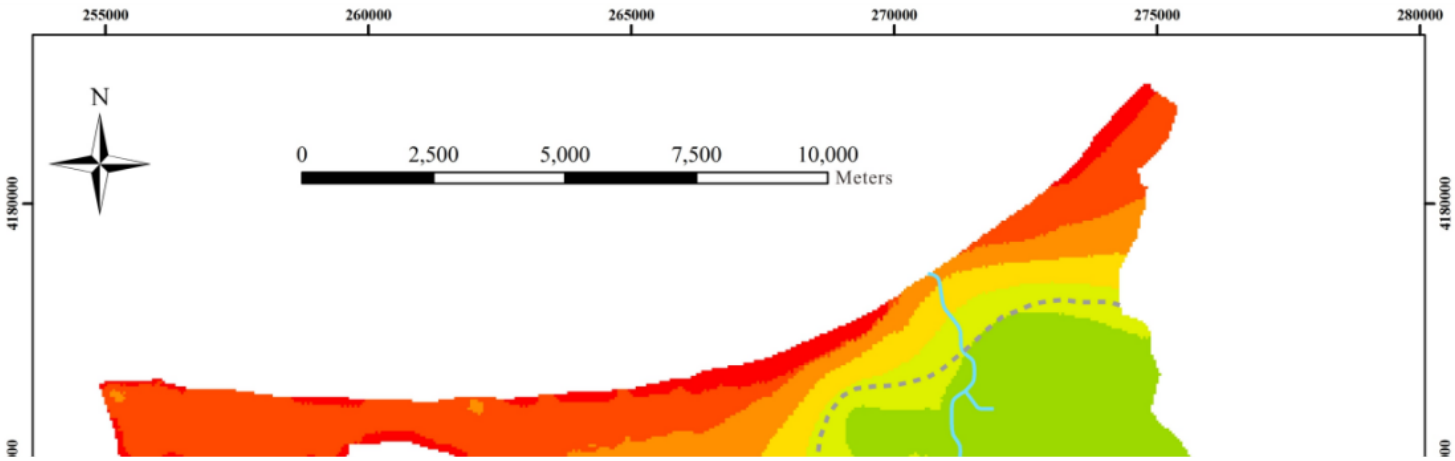

Figure 8

Model projections of seawater intrusion distribution in January 2052

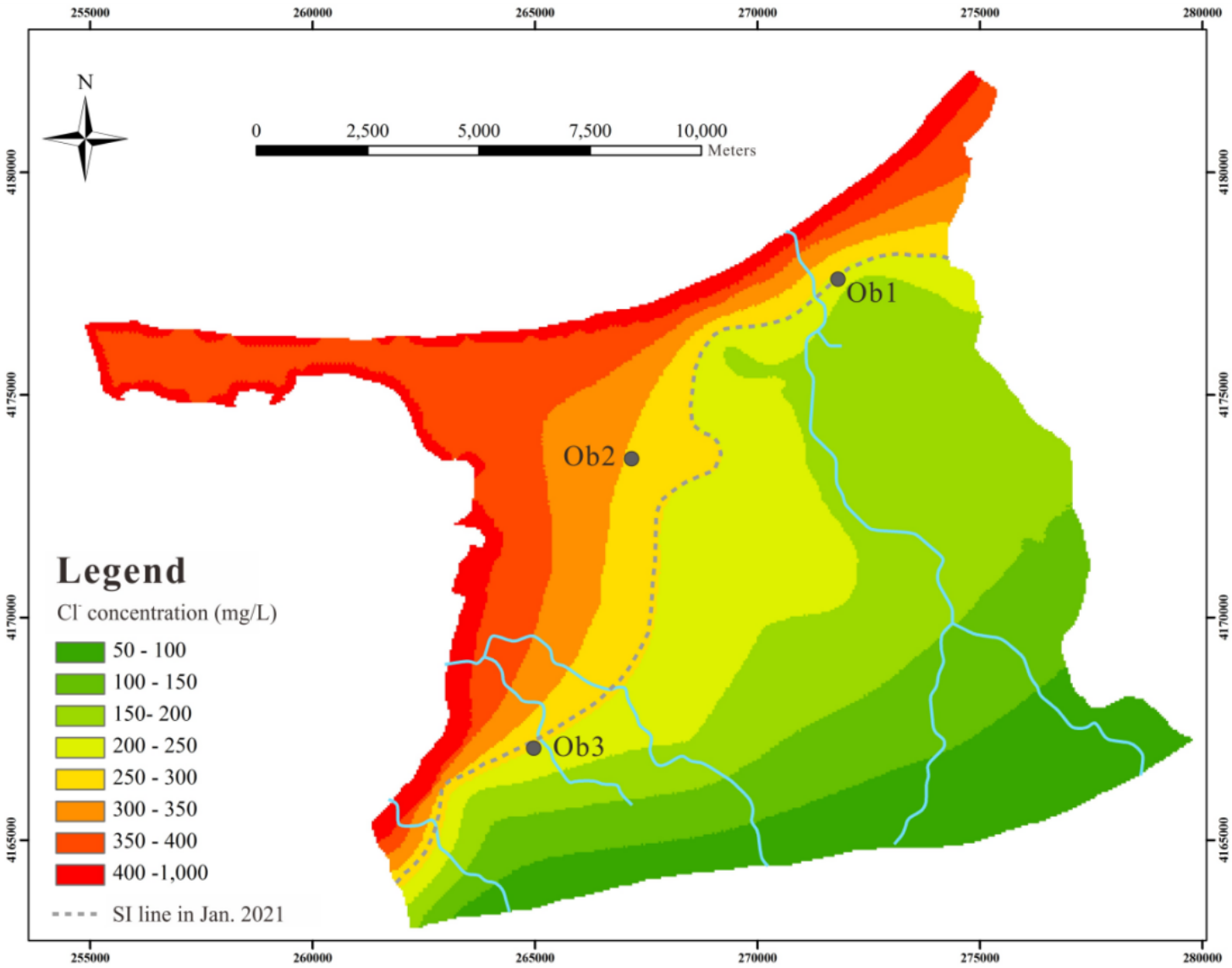

Figure 9

Location distribution of typical observation wells in the study area

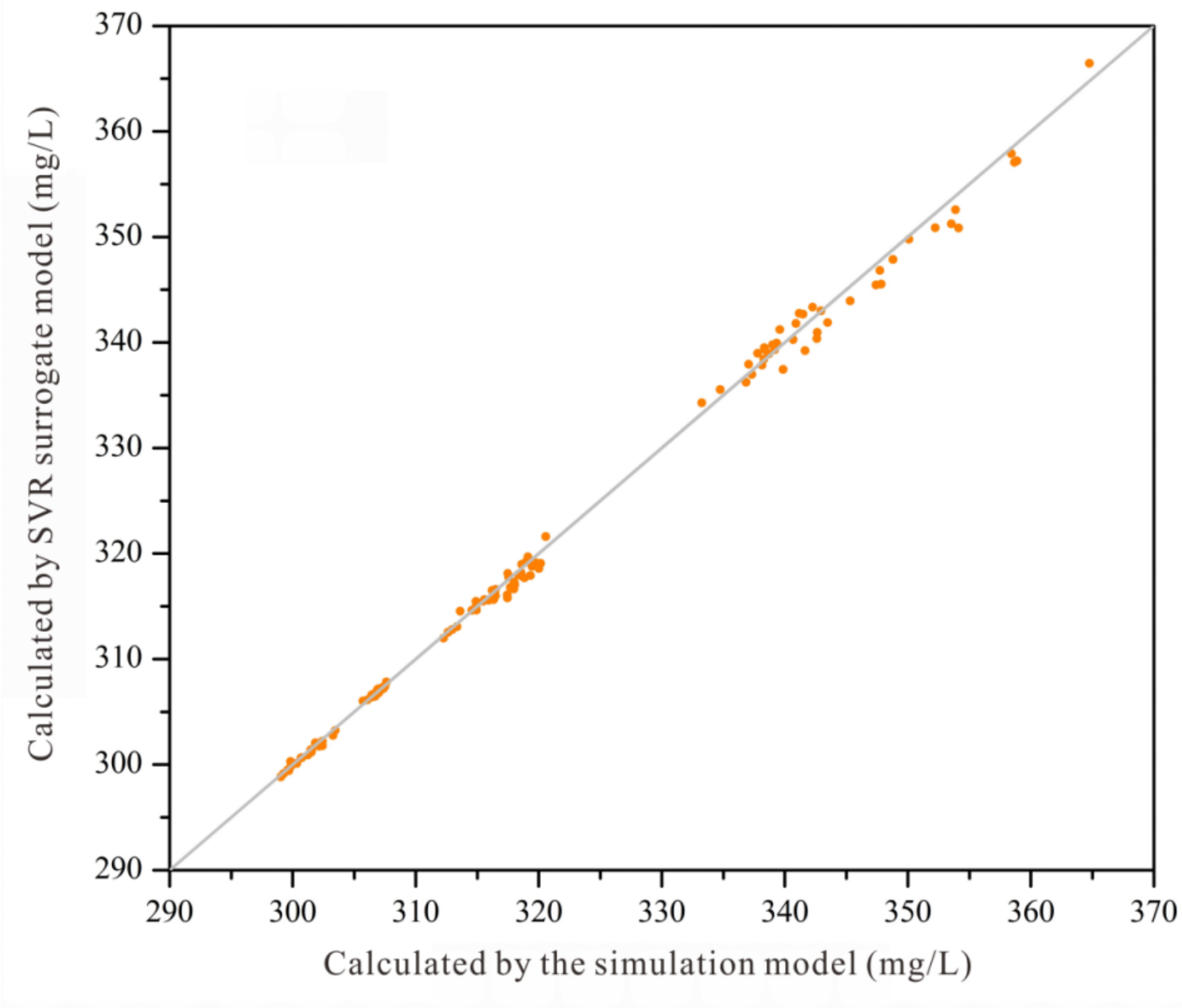

Figure 10

Fitting results of the SVR surrogate model

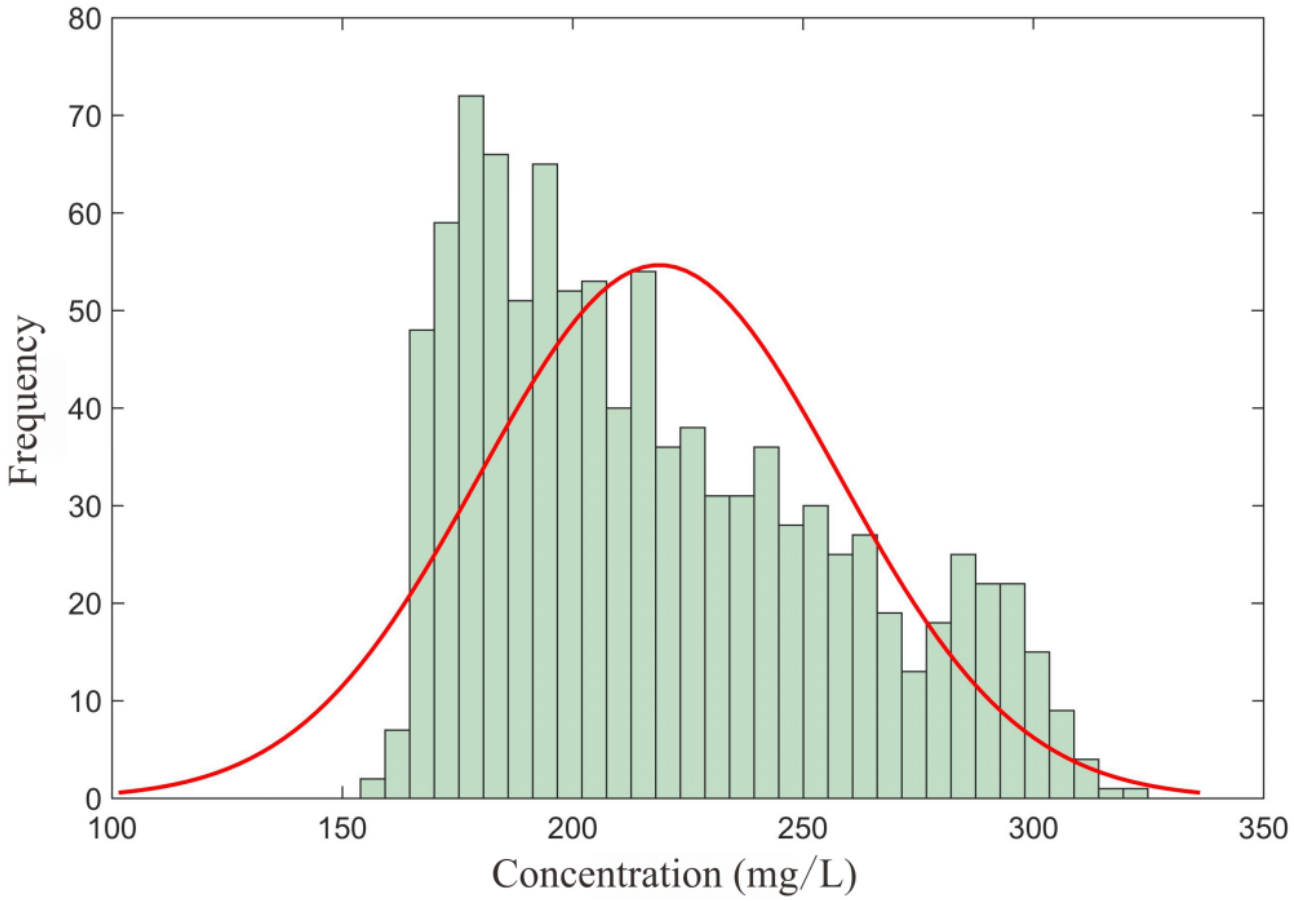

Figure 11

Histogram of chloride concentration distribution in Ob1 well in January 2052

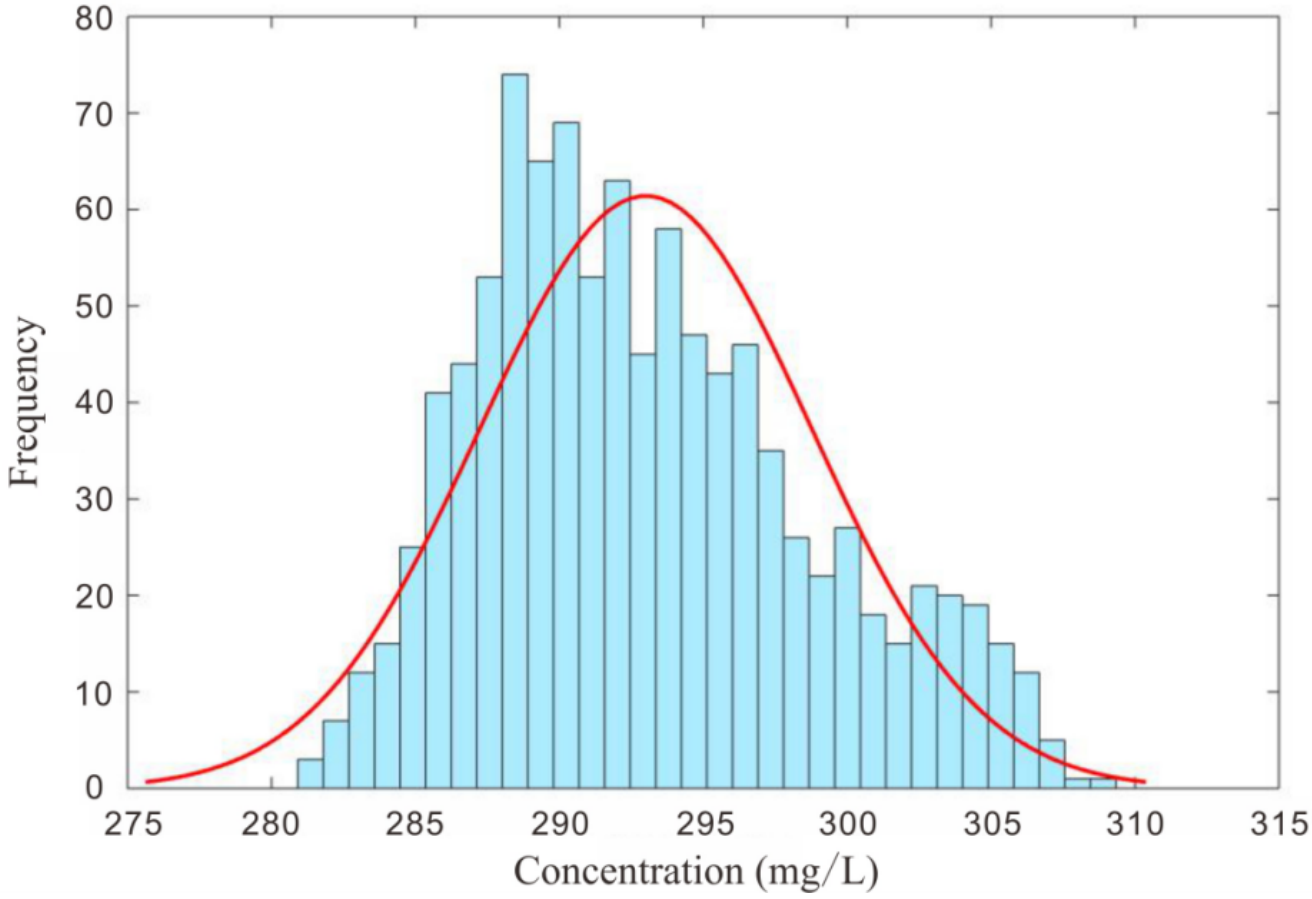

Figure 12

Histogram of chloride concentration distribution in Ob2 well in January 2052

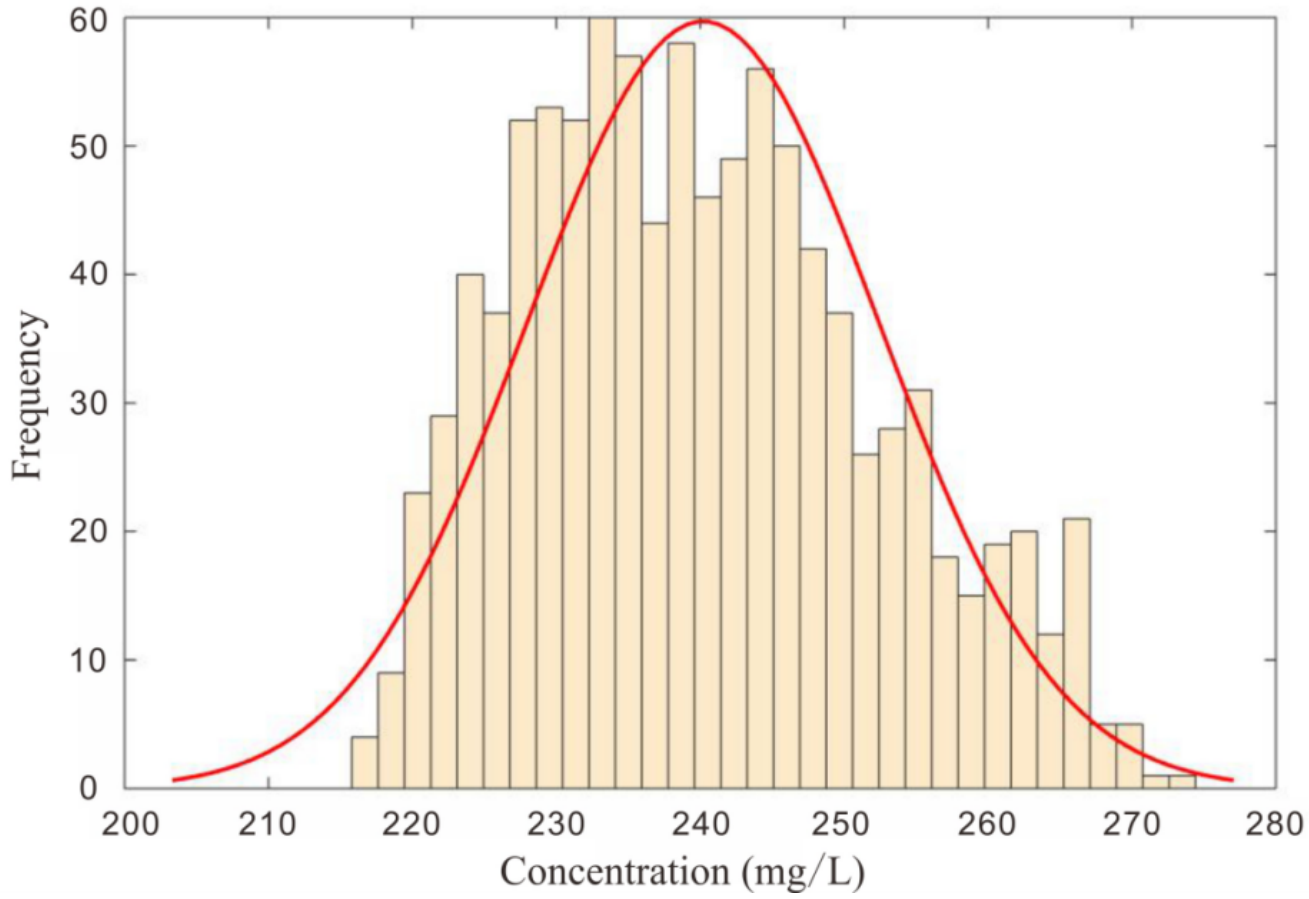

Figure 13

Histogram of chloride concentration distribution in Ob3 well in January 2052

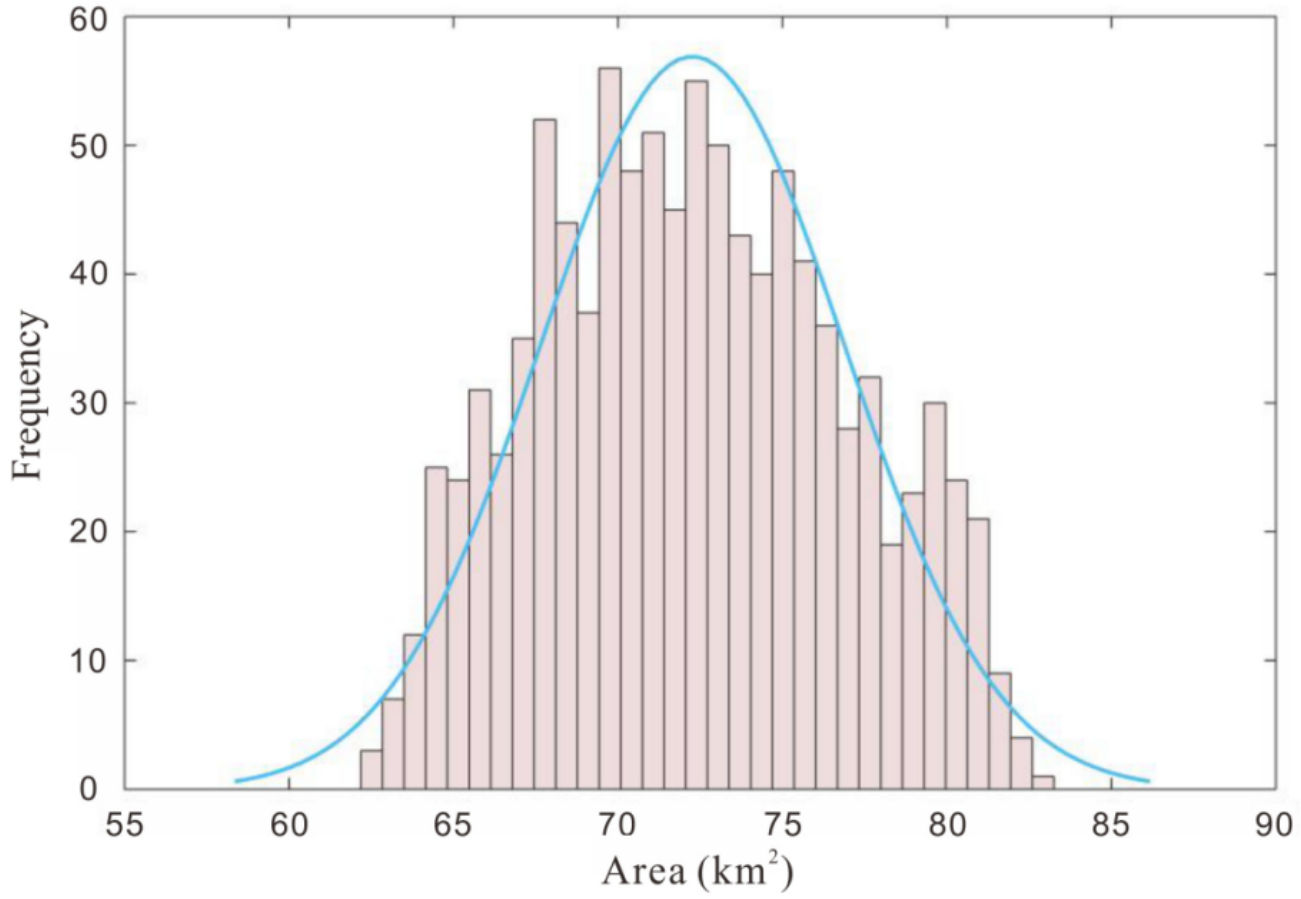

Figure 14

Distribution histogram of seawater intrusion area in January 2052