



An automatic model selection-based machine learning approach to predict seawater intrusion into coastal aquifers

Dilip Kumar Roy¹ · Chitra Rani Paul¹ · Tasnia Hossain Munmun¹ · Bithin Datta²

Received: 7 November 2023 / Accepted: 5 April 2024 / Published online: 2 May 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

Concerns about seawater intrusion resulting from unplanned mining of groundwater from coastal aquifers have become a global issue. To address this, it is crucial to implement a well-structured plan for groundwater withdrawal that can effectively manage and restrict salinity levels within the aquifer to acceptable levels. An integrated simulation–optimization (S–O) system has been a suitable tool for developing an optimum groundwater withdrawal scheme for managing seawater intrusion. The success of an S–O methodology largely depends on the use of computationally competent surrogates for the intricate simulation model. Although several surrogate models have been recently proposed to create management models for seawater intrusion using integrated S–O approach, the majority of these surrogates have been developed based on the subjective judgement. To fill this research gap, this study proposes an automatic model selection (AutoML)-based machine learning (ML) approach to predict the mechanisms of seawater intrusion in coastal aquifers. The AutoML was performed by optimizing the hyperparameters of a number of ML algorithms and selecting the best performing algorithm utilizing the asynchronous successive halving algorithm (ASHA). The best performing models were developed at 16 monitoring locations (MLs) using the predictor–response training data originated from a simulation model. Results revealed the capability of the ASHA optimization-based AutoML scheme to optimally select the best performing models with adequate prediction accuracies for the particular MLs. The selected best models at various monitoring locations exhibited high performance with accuracy (= 1), R (~0.99), NS (~0.99), IOA (~0.99), and KGE (~0.99), which are close to 1, indicating excellent model accuracy. Furthermore, the models demonstrated an RMSE value range of 0.0003–1.4987 mg/L, a relatively small range that is considered favorable for any predictive modeling approach. This study reveals the suitability and efficacy of the automatically selected surrogate models to develop an S–O-based management model to address real-world coastal aquifer management challenges related to seawater intrusion.

Keywords Coastal aquifer · Seawater intrusion · Automatic model selection · Machine learning algorithm · Integrated simulation–optimization

Introduction

Coastal groundwater plays a vital role in maintaining freshwater reserves essential for fulfilling the agricultural, domestic, and industrial water requirements of populations residing near coastal regions. With over 50% of the world's

population residing in these regions and the potential for this number to increase to 75% in the next century (Neumann et al. 2015), the demand for freshwater resources is increasing. This growing trend in coastal settlements inevitably leads to the excessive utilization of valuable groundwater reserves. Overexploiting coastal groundwater can lead to saline water intrusion, which, in turn, can have severe consequences on natural equilibrium. Addressing the complex and interconnected threats of salinization is crucial for the well-being of populations living in coastal regions (Feist et al. 2023). To prevent seawater intrusion, it is imperative to adopt a prudent and sustainable groundwater pumping plan in coastal aquifers. Effective management of groundwater resources serves as a pivotal factor in fostering sustainable

✉ Dilip Kumar Roy
dilip.roy@my.jcu.edu.au

¹ Irrigation and Water Management Division, Bangladesh Agricultural Research Institute, Gazipur 1701, Bangladesh

² Discipline of Civil Engineering, College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia

development within any given region, notably contributing to agricultural, residential, industrial, and economic stability (Masoud 2020). A solution to sustainable groundwater resources management lies in the implementation of multiple objective management models that can prescribe an optimum groundwater pumping plan whilst keeping the aquifer's saltwater levels within allowable bounds. These models are developed using the concept of a combined simulation–optimization (S–O) methodology, wherein the computationally efficient surrogate models replace the simulation model (Dhar and Datta 2009; Sreekanth and Datta 2010). To support the development of an integrated S–O-based seawater intrusion management model, this study proposes an automatic model selection (AutoML)-based machine learning (ML) algorithms to develop surrogate models for predicting salinity intrusion in a coastal aquifer located in Bangladesh.

Recently, there has been an increasing attention among researchers in developing computationally efficient surrogate models to predict seawater intrusion mechanisms in coastal aquifers. These surrogates have been employed as computationally feasible alternatives to the intricate numerical models in the combined S–O-based seawater intrusion management models (Dhar and Datta 2009; Roy and Datta 2018a). Prediction accuracies of the surrogate models are of utmost importance as they largely determine the overall precision and dependability of the developed management model. In the last few years, different ML-based algorithms have been employed to construct the surrogate models for seawater intrusion problems with different performance abilities. The ML-based methodologies involve the utilization of various cutting-edge ML algorithms, which have been utilized to construct models for managing saltwater intrusion. A synopsis of recent investigations concerning the prediction of saltwater intrusion through ML algorithms for the development of management models in coastal aquifers is outlined in Table 1. Overall, the prediction capabilities of these prediction models have achieved satisfactory results. However, selecting the most appropriate model from a multitude of algorithms designed for estimating seawater intrusion, each with varying performance levels, can be a formidable task. Hence, there is a need to seek an approach that simplifies the model selection task by automating the process. This automated model selection process has not been previously applied to choose the most suitable saltwater intrusion prediction models. To address this research gap, this study introduces an AutoML approach to select the best saltwater intrusion prediction models from a number of different alternative ML-based algorithms. These prediction models are known as the surrogate or meta-models in the context of a combined S–O approach to construct a management model to control seawater intrusion.

Implementing new techniques of seawater intrusion prediction is crucial to lowering modelling errors and model

parameter uncertainties (Moravej et al. 2020). However, it is still a complex scientific problem for seawater intrusion prediction to reach the requirement of higher prediction accuracy with the desired model. Previous studies of seawater intrusion prediction compared a few ML-based modelling approaches and proposed the best predictive model based on the comparison results. This approach is constrained by the selection of appropriate candidate models for comparison and finding the top performing model. In other words, traditional ML-based approaches to seawater intrusion prediction involve manual model selection, which can be time-consuming, and subjective. To address these limitations, it is often beneficial to compare multiple approaches through optimizing their optimizable hyperparameters using optimization algorithms to arrive at the most appropriate prediction model for a given data set. This AutoML refers to the process of automatically choosing the most suitable regression model for a given data set. The objective is to determine the model that offers the best fit to the data and provides the most accurate predictions. To this end, the present study proposes an AutoML technique to select the top-performing model for enhancing prediction accuracy and streamlining the modelling process. The asynchronous successive halving algorithm (ASHA) (Li et al. 2020) was used in the AutoML technique to search for the top-performing model through hyperparameters tuning.

Selection of models for a particular problem is quite difficult given that a vast majority of ML-based models are available. Training numerous models and optimizing their hyperparameters can frequently span days or even weeks. Consequently, employing an automated model optimization technique for the development and comparison of various models can significantly expedite the process. For model hyperparameters tuning, Bayesian and ASHA optimization can be used to speed up the process. This research paper focuses on the application of the ASHA, for AutoML in seawater intrusion prediction tasks. The ASHA algorithm is a state-of-the-art model selection method that balances exploration and exploitation through adaptive resource allocation. Rather than individually training each model with distinct hyperparameter configurations, a selection of different models can be chosen, and their default hyperparameters can be fine-tuned using the ASHA optimization method. Optimization algorithms aim to discover the ideal combination of hyperparameters for a specific model by minimizing the model's objective function. These optimization algorithms select new hyperparameters strategically in each iteration and generally reach the optimal hyperparameter set more efficiently than a straightforward grid search. Consequently, this study aims to employ ASHA optimization to train several regression models on a given training data set and determine which model performs most effectively on a test data set. Our goal is to look into the effectiveness

Table 1 Summary of recent studies on predicting saltwater intrusion using ML algorithms for developing saltwater intrusion management models in coastal aquifers

Year	Study area	ML algorithms	Input variables	Input–output training patterns	References
2009	Illustrative study area (2.52 km ²)	Artificial Neural Network (ANN)	33	3600	Dhar and Datta (2009)
2007	Illustrative study area (2.59 km ²)	ANN	18	636	Bhattacharjya et al. (2007)
2005	Illustrative study area (2.52 km ²)	ANN	24	2400	Bhattacharjya and Datta (2005)
2009	Real aquifer	ANN	11	273	Kourakos and Mantoglou (2009)
2010, 2011	Illustrative study area (2.52 km ²)	Genetic Programming (GP) and ANN	33	230	Sreekanth and Datta (2010, 2011a; b)
2018	Illustrative study area (4.35 km ²)	Adaptive Neuro Fuzzy Inference System (ANFIS), Fuzzy Inference System (FIS), Genetic Algorithm tuned FIS, Gaussian Process Regression (GPR), Multivariate Adaptive Regression Spline (MARS), Probabilistic Linear Regression (PLR), Relevance Vector Regression, Support Vector Regression (SVR), Regression Tree	80	1500	Roy and Datta (2017a, 2018b, 2020a; b, c, d, e, b, c)
2017	Real aquifer (90.5 km ²)	Gaussian Process Regression (GPR)	11	5000	Rajabi and Ketabchi (2017)
2019	Illustrative study area (21.0 km ²)	GPR	40	4000	Kopsiaftis et al. (2019)
2022	Illustrative study area (12.5 km ²)	GPR	18	350	Saad et al. (2022)
2021	Illustrative study area (2.5 km ²)	Group Method of Data Handling	52	800	Lal and Datta (2021)
2023	Illustrative study area (21.0 km ²)	Random Forest	11	5000	Kopsiaftis et al. (2023)
2018	Real aquifer (421.0 km ²)	M5 Model Tree	66 (divided into 5 clusters)	–	Ranjbar and Mahjouri (2018)

of the AutoML-based surrogate model selection using the ASHA optimization algorithm to predict seawater intrusion processes within coastal aquifers. By automating the model selection process, this research intends to overcome the limitations of manual model selection process, which often involves subjectivity and leads to suboptimal choices.

Our proposed approach automates and excludes the manual phases needed to transition from a data set to a predictive model. We present a novel approach that automates the model selection process for seawater intrusion prediction, providing a more objective and efficient alternative to manual selection. Therefore, the contribution of this study includes (a) building multiple regression models for a given training data set of salinity intrusion through optimizing

their hyperparameters using the ASHA optimization algorithm, and (b) selecting the top-performing models that can be used as surrogate models in a combined S–O-based management model for controlling seawater intrusion.

Methodology

The study seeks to introduce an AutoML-based approach for model selection in predicting seawater intrusion processes within coastal aquifer systems subject to spatiotemporal groundwater withdrawal. This AutoML approach aims to replace manual surrogate model selection for a combined S–O framework to construct seawater intrusion management

models. The methodology involves several key steps: simulating groundwater flow and solute transport phenomena through a 3D simulation model, generating spatiotemporal groundwater extraction values within practical limits through a sampling strategy, creating predictor–response data set for training with the assistance of a calibrated and validated simulation model, and developing surrogate models using the AutoML approach. This research is based on data produced using a 3D seawater intrusion simulation model that has been applied to an actual coastal aquifer in southern Bangladesh, as initially proposed in Roy and Datta (2020b). To enhance readability, the methodologies are summarized briefly in the subsequent paragraphs.

Simulation of aquifer processes

Aquifer dynamics were modeled in the context of transient flow and salt transport using a 3D density-driven combined flow and solute conveyance numerical code known as FEMWATER (Lin et al. 1997). FEMWATER is a robust and widely adopted finite-element-based modelling approach that has demonstrated its effectiveness in numerous prior investigations, specifically in addressing density-driven combined flow and salinity transport phenomena in coastal aquifers (Sreekanth and Datta 2011b; Roy and Datta 2017a). The controlling equation for 3-D flow within FEMWATER is expressed through the modified Richards equation (Lin et al. 1997):

$$\frac{\rho}{\rho_0} F \frac{\delta h}{\delta t} = \nabla \cdot \left[K \cdot \left(\nabla h + \frac{\rho}{\rho_0} \nabla z \right) \right] + \frac{\rho}{\rho^*} q, \quad (1)$$

where F denotes storage coefficient, h represents pressure head, K refers to hydraulic conductivity tensor, z is the potential head, q represents a source or a sink, ρ denotes water density at chemical concentration C , ρ_0 symbolizes referenced water density at zero chemical concentration, ρ^* is the density of injection fluid or that of the withdrawn water.

The hydraulic conductivity, denoted as K , can be represented as

$$K = \frac{\rho g}{\mu} k = \frac{(\rho/\rho_0)}{\mu/\mu_0} \frac{\rho_0 g}{\mu_0} k_s k_r = \frac{\rho/\rho_0}{\mu/\mu_0} k_{so} k_r, \quad (2)$$

where μ represents the dynamic viscosity of water at a given chemical concentration denoted as C , μ_0 denotes reference dynamic viscosity at zero chemical concentration, k_s is the saturated permeability tensor, k_r is the relative permeability or relative hydraulic conductivity, k_{so} refers to referenced saturated conductivity tensor.

In the context of the seawater intrusion problem, the connection between concentrations of fluid density and saltwater is described by the following equation:

$$\frac{\rho}{\rho_0} = 1 + \varepsilon \frac{C}{C_{\max}}, \quad (3)$$

where ε represents the density reference ratio (dimensionless), C denotes the material concentration in the aqueous phase, and C_{\max} is the highest material concentration.

The 3D transport equation is expressed through the following equation:

$$\begin{aligned} \theta \frac{\delta C}{\delta t} + \rho_b \frac{\delta S}{\delta t} + V \cdot \nabla C - \nabla \cdot (\theta D \cdot \nabla C) \\ = - \left(\alpha' \frac{\delta h}{\delta t} + \lambda \right) (\theta C + \rho_b S) - (\theta K_w C + \rho_b K_s S) \\ + m - \frac{\rho^*}{\rho} q C + \left(F \frac{\delta h}{\delta t} + \frac{\rho_0}{\rho} V \cdot \nabla \left(\frac{\rho}{\rho_0} \right) - \frac{\delta C}{\delta t} \right) C, \end{aligned} \quad (4)$$

where ρ_b is the bulk density of the medium, C refers to material concentration in aqueous phase, S is the material concentration in adsorbed phase, t denotes time, V represents discharge, ∇ is the del operator, D denotes dispersion coefficient tensor, λ represents decay constant, $M = qC_m$ = artificial mass rate, q is the source rate of water, C_m is the material concentration in the source, K_w represents the first order biodegradation rate constant through dissolved phase, K_s is the first order biodegradation rate through adsorbed phase, K_d refers to the distribution coefficient, θ refers to moisture content, α' is the modified compressibility of water, β' is the modified compressibility of the medium, n refers to porosity, S is the saturation. Equations (1) and (4) describe, respectively, the flow and salinity transport phenomena. The salinity intrusion phenomenon becomes highly nonlinear due to the interconnection between the density coupling coefficient and Darcy velocities in these two equations. Consequently, the equations for flow and salinity transport processes were concurrently resolved using a numerical simulation model grounded on finite elements.

Multiple simulation to generate predictor–response data set for training

The calibrated and validated seawater intrusion model was employed to produce sets of predictor–response arrays to train AutoML-based surrogate models, using randomized inputs representing spatiotemporal pumping conditions. The spatiotemporal groundwater abstraction stress imposed on the aquifer was linked to the extraction of water from a series of pumping and barrier abstraction wells situated at distinct positions and time intervals.

These randomized inputs were derived from a uniform sampling distribution, specifically employing Halton Sequences (HA) (Halton 1960). This selection of sampling method was adopted in the present study due to its demonstrated advantages over more frequently employed methods (Loyola et al. 2016). These sampled inputs were then fed into the simulation model to compute the resultant salinity values at designated monitoring locations (MLs). Each run of the simulation model, with inputs and outputs, constitutes a single input–output pattern. Executing the simulation model multiple times produces a collection of such input–output patterns. The aquifer properties, along with the simulation conditions, stayed consistent across various simulation runs. The sole variable that changed in successive simulations was the transient groundwater abstraction values. This variation allowed for the acquisition of distinct estimations of salinity concentrations, which were exclusively influenced by the applied groundwater abstraction stress on the aquifer system.

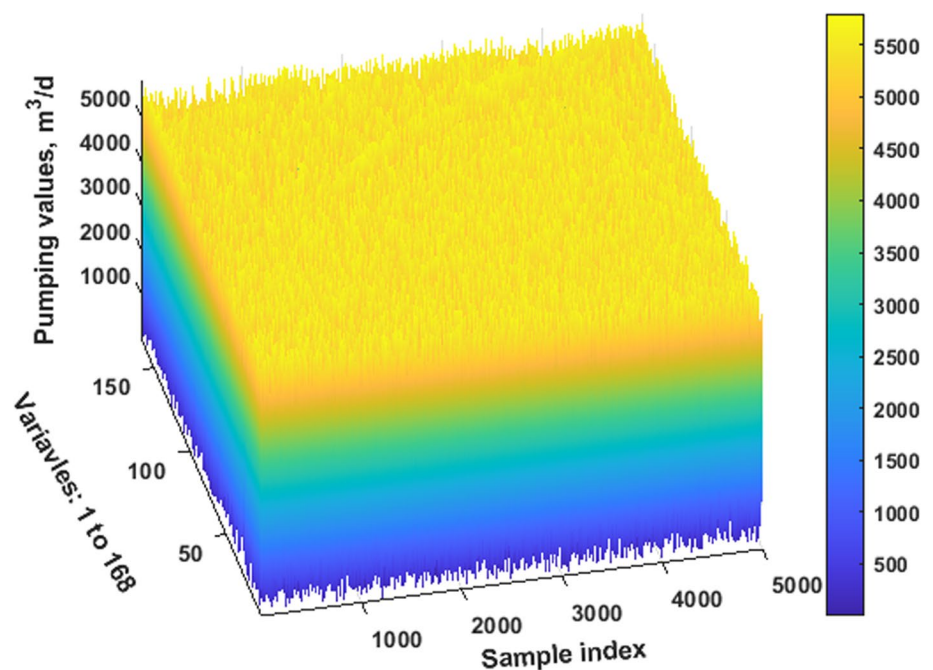
The simulation was conducted over a 3-year management period utilizing 43 production and 13 barrier wells. As a result, the study accounted for 168 decision variables $[(43 + 13) \times 3 = 168]$, representing spatiotemporal groundwater extraction values. To adequately capture the input–output training patterns, which encompassed numerous input variables (168 in this study), an adequately sufficient number of input–output training patterns were generated. The input pumping patterns were generated using Halton sequences, with specified upper and lower bounds for pumping values. These bounds were set at 0 and 6000 m^3/day , respectively.

The selection of the upper bounds of pumping was based on actual groundwater withdrawal values recorded in the study area in April 2017 (Roy and Datta 2020b). A three-dimensional visualization of the spatiotemporal groundwater extraction values has been presented in Fig. 1 depicting the groundwater abstraction patterns over time.

The necessary input–output training patterns for developing a coupled S–O-based saltwater intrusion management model vary based on system complexity and decision variable count. Complex problems may demand numerous patterns for a reliable surrogate model. Typically, a large data set is generated, and training starts with a subset, incrementally adding patterns until no significant improvement is observed. Despite its time demand, this method yields surrogate models capturing true input–output trends across the decision space. Coupled with optimization algorithms, these models provide optimal groundwater extraction values. However, generating training patterns through multiple calibrated model simulations is time-intensive. An alternative is adaptive training, which uses fewer patterns to develop reasonably accurate emulators of saltwater intrusion processes (Sreekanth and Datta 2010; Christelis and Mantoglou 2016; Christelis et al. 2018).

To develop adaptive surrogate models, Sreekanth and Datta (2010) employed an expanding set method, progressively adding training patterns generated by an optimization strategy based on the search direction. Initially, limited training patterns were used with surrogate models alongside an optimization algorithm to find near-optimal solutions. Once an approximate optimal solution was found, additional training data near this solution were generated to improve

Fig. 1 Three-dimensional view of the transient groundwater pumping values



accuracy. The surrogate model was then retrained, and the optimization routine rerun to achieve a more accurate optimal solution. This process iterated until the desired level of surrogate model accuracy was attained.

Christelis and Mantoglou (2016) employed an adaptive approach to train surrogate models for optimizing pumping in coastal aquifers, dealing with ten wells and ten corresponding constraints per well. They devised an online Radial Basis Function (RBF) surrogate model training scheme integrated with an optimization algorithm. Their method involved adding infill points to the initial sampling plan using the best solutions found by the RBF model during optimization, prioritizing local exploitation over global improvement. While providing accurate predictions in limited regions of the decision space, this approach may overlook the global optimum (Forrester et al. 2008). It also integrates feedback from numerical simulations to assess current best solutions. Christelis et al. (2018) extended this method to larger problems within computational constraints, evaluating its performance.

In an alternative approach, Song et al. (2018) adaptively trained a surrogate model during evolutionary search to meet a desired fidelity level, preventing error accumulation in forecasting and ensuring convergence to the true Pareto-optimal front. Initially, a set number of training samples spanning the feasible variable space established the preliminary surrogate model. Offspring individuals were then evaluated using this model, integrated with the current population, and subjected to fast non-domination sorting. About 20% of the population size was selected based on convergence and diversity criteria as promising individuals at intervals of several generations. In addition, a modified local search operator enhanced local optimality. Selected local optima were reevaluated using a numerical simulation model and merged with approximate Pareto optimal solutions. Solutions evaluated by the simulation model were archived to improve surrogate model accuracy. This approach focuses on local optimal solutions, offering more accurate predictions within limited regions of the input variable decision space.

However, in its pursuit of computational efficiency in generating necessary input–output patterns for surrogate

model training, this approach overlooks the accuracy and uncertainty of surrogate model predictions. Furthermore, these adaptive methods necessitate evaluating the generated training patterns (obtained through optimization algorithms) using numerical simulation models, inevitably demanding additional computational time and effort compared to the initial fewer input–output training patterns. Relying on optimal solutions from limited decision space regions and local optimality, the adaptive approach may miss the global optimum while generating additional training patterns from existing ones. Considering these limitations, our study aims to prioritize the development of highly accurate surrogate models using the requisite number of input–output training patterns.

Surrogate model development

Surrogate models play a pivotal role in a combined S–O method, which is the central component for developing a management model for seawater intrusion control that prescribes ideal water abstraction rates from coastal aquifers. To aid in constructing a combined S–O-based management model for seawater intrusion, this study proposes automatic selection of ML-based surrogate models instead of manually selecting a few of them and determining the top-performing one. The AutoML process was built upon the data set used for model development. The spatiotemporal pumping rates, originating from 43 production wells and 13 barrier wells, spanning a 3-year timeframe, served as input data for the surrogate models. Consequently, the surrogate model's input comprised a total of 168 (56 wells \times 3 time steps). On the other hand, the surrogate model's output represents the concentrations recorded at designated MLs throughout all time steps. Salinity concentrations were monitored at 16 locations. Consequently, 16 models were developed representing 1 model at each monitoring location. This input–output configuration for the surrogate models can be expressed as follows:

Salinity concentration at a certain

$$\text{monitoring location} = f(\text{Spatiotemporal groundwater extraction from the production and barrier wells}). \quad (5)$$

Equation 5 can be mathematically represented by

$$\left\{ C_i^j, i = ML1, \dots, ML16, j = 1, \dots, 3 \right\} = f \left\{ \left(PW_k^j, k = 1, \dots, 43, j = 1, \dots, 3 \right), \left(BW_l^j, l = 1, \dots, 13, j = 1, \dots, 3 \right) \right\}, \quad (6)$$

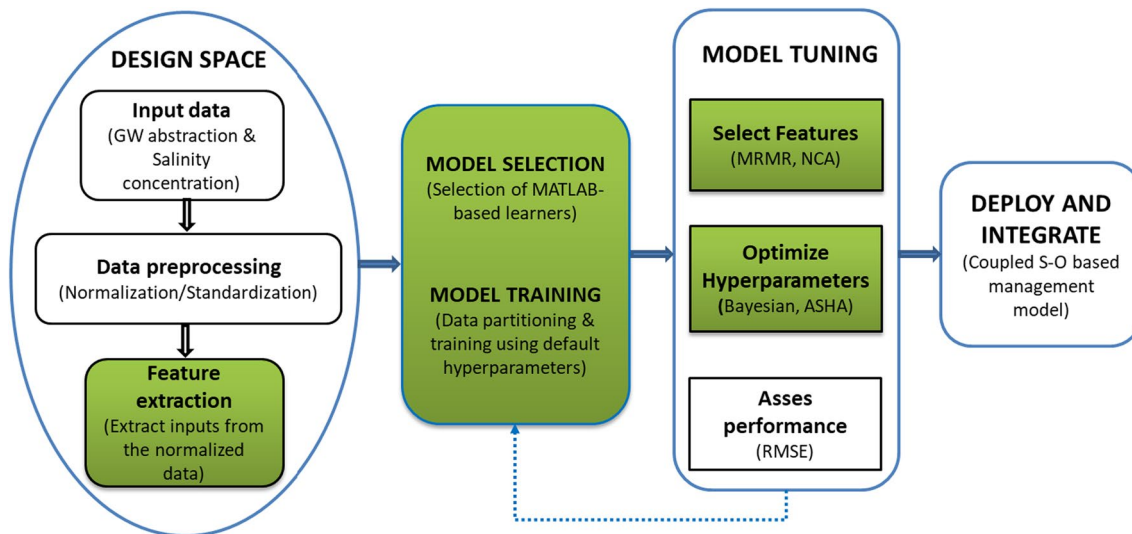


Fig. 2 Machine learning-based modelling workflows with automated machine learning (AutoML) approach. The steps where AutoML is applicable are displayed in a light green shade

where i and j represent indexes for monitoring places and management time periods, accordingly; k and l indicate indices for the production and barrier well locations, respectively; C_i^j denotes saltwater concentration levels at the i th monitoring places for the j th management time steps; PW_k^j symbolizes water abstraction from the k th production well places for the j th management time steps; and BW_l^j represents water abstraction from the l th barrier well places for the j th management time steps.

The proposed surrogate model selection using AutoML approach is presented in Fig. 2.

Selection of candidate machine learning algorithms

The precision of a surrogate model primarily hinges on the judicious choice of models and their prediction accuracy.

To predict salinity intrusion in coastal aquifer systems, surrogate models were created through the training and testing of ML algorithms. Nonetheless, training multiple ML algorithms and identifying their optimal parameter configurations presents a time-intensive endeavor. To streamline and simplify this procedure, a set of commonly employed ML algorithms with distinct hyperparameters was chosen as the candidate algorithms. Table 2 presents these candidate ML algorithms alongside their adjustable hyperparameters.

In the process of model development at a specific monitoring location, the hyperparameters of the seven ML algorithms listed in Table 1 were fine-tuned utilizing the ASHA (Li et al. 2020) optimization algorithm. The ASHA optimization served a dual purpose: not only did it fine-tune the parameters of the seven ML algorithms, but it also

Table 2 Machine learning algorithms and their tunable hyperparameters

Machine learning algorithms	Hyperparameters
Ensemble Regression (ER) Model	Method (Least-squares boosting, Bootstrap aggregation), Number of ensemble learning cycles, Learning rate for shrinkage, Minimum number of leaf node observations, Maximal number of decision splits, Number of predictors to select at random for each split
Gaussian Process Regression (GPR) Model	Sigma, Basis function, Kernel function, Kernel scale, Kernel parameters, Standardize
Kernel Regression (KR) Model	Epsilon, Kernel scale, Lambda, Learner, Number of dimensions of expanded space
Linear Regression (LR) Model for high-dimensional data	Lambda, Learner, Regularization
Artificial Neural Network (ANN) Model	Activations, Lambda, Layer sizes, Standardize, Layer biases initializer, Layer weights initializer
Support Vector Machine (SVM) Regression Model	Box constraint, Epsilon, Kernel scale, Kernel function, Polynomial order, Standardize
Binary Decision Regression Tree (BDRT)	Minimum number of leaf node observations, Maximal number of decision splits

determined the best-performing ML algorithm among the seven options.

Selection of the best models through hyperparameters optimization

It is of crucial importance to discern which model, from the multitude of options available, excels in achieving the desired task. Subsequently, fine-tuning its hyperparameters becomes paramount for optimal performance. AutoML streamlines this by optimizing both the model and its related hyperparameters in one operation. The majority of ML algorithms require careful selection of hyperparameters (Probst et al. 2019). The choice of hyperparameters significantly influences the performance of ML models (Zhang et al. 2016), and haphazard selection can yield subpar results. Many studies utilized trial-and-error hyperparameter selection (Zhang et al. 2016; Jeong and Park 2019), grid and/or random search (Bergstra and Bengio 2012) to develop ML-based models. Some even employ heuristic optimization methods like genetic algorithm and particle swarm optimization (Bergstra and Bengio 2012). Therefore, a precise and dependable automated hyperparameter optimization approach is of great desirability (Zhang et al. 2016) and proves valuable for ensuring an impartial comparative evaluation among various ML alternatives. Moreover, when evaluating various ML models, fair assessments can only be made if all models are similarly optimized (or receive equivalent consideration) for the specific problem under consideration.

ASHA (Li et al. 2020) optimization algorithm is an upgraded variant of Successive Halving algorithm (SHA) (Karnin et al. 2013; Jamieson and Talwalkar 2015). It is an easy-to-use hyperparameter optimization technique that takes advantage of aggressive early halting and is appropriate for large-scale hyperparameter optimization problems (Li et al. 2020). As demonstrated on a workload with 500 workers, ASHA outperforms current advanced hyperparameter tuning techniques, scales in a linearly manner with the number of workers in distributed environments, and is appropriate for huge parallelism (Li et al. 2020). The user does not need to indicate in advance how many configurations to evaluate, because it is asynchronous, but it still needs the same inputs as SHA. A thorough description of ASHA can be found in Li et al. (2020), and is not duplicated in this work. As far as the authors are aware, this is the first effort that ASHA optimization algorithm is used to tune the hyperparameters of several ML algorithms to automatically select the best proxy model to predict salinity intrusion in coastal aquifer systems.

In this research, surrogate models were developed by tuning hyperparameters of seven commonly used ML

algorithms and finding the best model for the seawater intrusion data set at 16 MLs. In this approach, rather than utilizing distinct combinations of hyperparameters to train every model, seven ML algorithms were selected and their default hyperparameters were optimized using ASHA optimization algorithm. This optimization algorithm searches for an optimal set of hyperparameters for a particular model through optimizing the cost function of the model (minimization of the Mean Squared Error (MSE)). The optimization algorithm deliberately selects new hyperparameters in each iteration and produces an optimal set of hyperparameters for a given training data set and identifies the model with the highest performance on a test data set. The ASHA optimization algorithm used seven ML algorithms with various hyperparameters values and trained them using a tiny portion of the learning data set. If the $\log(1 + \text{valLoss})$ value for a particular model is promising, where valLoss is the cross-validation MSE, the model is retained for the next phase and retrained using more training data. Repetitively, this process utilizes increasingly greater volumes of data for training efficient models. The selected best model was one that produced the lowest training and test errors. Table 3 presents the selected best models and their optimized hyperparameters at different MLs.

Evaluation of surrogate model performance

The effectiveness of the proposed AutoML-based surrogate models was assessed using a range of statistical evaluation metrics, including Accuracy (A), Correlation Coefficient (R), Nash–Sutcliffe Efficiency Coefficient (NSE), Willmott's Index of Agreement (IOA), Kling–Gupta Efficiency (KGE), Root Mean Squared Error (RMSE), Normalized RMSE (NRMSE), Mean Absolute Error (MAE), Median Absolute Deviation (MAD), Mean Bias Error (MBE), and Percentage Bias (PBIAS). The mathematical expressions defining these performance indices are provided below:

$$A = 1 - \text{abs}\left(\text{mean}\frac{C_{i,P} - C_{i,S}}{C_{i,S}}\right), \quad (7)$$

$$R = \frac{\sum_{i=1}^n (C_{i,S} - \bar{C}_S)(C_{i,S} - \bar{C}_P)}{\sqrt{\sum_{i=1}^n (C_{i,S} - \bar{C}_S)^2} \sqrt{\sum_{i=1}^n (C_{i,P} - \bar{C}_P)^2}}, \quad (8)$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^n (C_{i,S} - C_{i,P})^2}{\sum_{i=1}^n (C_{i,S} - \bar{C}_S)^2}, \quad (9)$$

Table 3 Selected best models and their optimal parameters at different monitoring locations

Monitoring locations	Selected model	Optimal parameters
ML1	GPR	Sigma: 0.0014, Basis Function: Constant, Kernel Function: Rational Quadratic, Kernel Scale: 514.21, Standardize=True
ML2	GPR	Sigma: 0.1747, Basis Function: Linear, Kernel Function: Matern 52, Kernel Scale: 36.087, Standardize=True
ML3	ANN	Activations: None, Standardize: True, Lambda: 3.4381×10^{-05} , Layer Weights Initializer: ‘glorot’, Layer Biases Initializer: ones, Layer Size: [1 8]
ML4	ANN	Activations: None, Standardize: True, Lambda: 0.3064, Layer Weights Initializer: ‘he’, Layer Biases Initializer: ones, Layer Size: [7 1 49]
ML5	GPR	Sigma: 0.00022487, Basis Function: None, Kernel Function: Rational Quadratic, Kernel Scale: 739.48, Standardize=True
ML6	GPR	Sigma: 0.0046, Basis Function: Linear, Kernel Function: Matern 52, Kernel Scale: 6.1465, Standardize=False
ML7	GPR	Sigma: 22.363, Basis Function: Linear, Kernel Function: Matern 52, Kernel Scale: 4476.5, Standardize=False
ML8	GPR	Sigma: 0.0768, Basis Function: Linear, Kernel Function: Matern 52, Kernel Scale: 7.7765, Standardize=True
ML9	GPR	Sigma: 67.488, Basis Function: Linear, Kernel Function: Rational Quadratic, Kernel Scale: 403.3, Standardize=True
ML10	ANN	Activations: None, Standardize: True, Lambda: 2.0371×10^{-06} , Layer Weights Initializer: ‘he’, Layer Biases Initializer: zeros, Layer Size: [41 55]
ML11	GPR	Sigma: 0.0396, Basis Function: Linear, Kernel Function: Matern 52, Kernel Scale: 62.349, Standardize=False
ML12	GPR	Sigma: 0.00013, Basis Function: Linear, Kernel Function: Matern 52, Kernel Scale: 48.677, Standardize=False
ML13	ANN	Activations: None, Standardize: True, Lambda: 7.8003×10^{-05} , Layer Weights Initializer: ‘he’, Layer Biases Initializer: ones, Layer Size: 296
ML14	GPR	Sigma: 0.0966, Basis Function: Linear, Kernel Function: Exponential, Kernel Scale: 51.134, Standardize=True
ML15	ANN	Activations: Sigmoid, Standardize: True, Lambda: 8.6446×10^{-06} , Layer Weights Initializer: ‘glorot’, Layer Biases Initializer: zeros, Layer Size: [2 13]
ML16	ANN	Activations: None, Standardize: True, Lambda: 1.8712×10^{-06} , Layer Weights Initializer: ‘he’, Layer Biases Initializer: ones, Layer Size: [20 1]

GPR Gaussian Process Regression, ANN Artificial Neural Network

$$IOA = 1 - \frac{\sum_{i=1}^n (C_{i,S} - C_{i,P})^2}{\sum_{i=1}^n (|C_{i,P} - \bar{C}_S| + |C_{i,S} - \bar{C}_S|)^2}, \tag{10}$$

$$NRMSE = \frac{RMSE}{C_S} \times 100, \tag{15}$$

$$KGE = 1 - ED = 1 - \sqrt{(R - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}, \tag{11}$$

$$MAE = \text{mean}[|C_{i,S} - C_{i,P}|], \tag{16}$$

$$\alpha = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (C_{i,P} - \bar{C}_P)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (C_{i,S} - \bar{C}_S)^2}}, \tag{12}$$

$$MAD(C_S, C_P) = \text{median} (|C_{S,i=1} - C_{P,i=1}|, |C_{S,i=2} - C_{P,i=2}|, \dots, |C_{S,i=n} - C_{P,i=n}|) \tag{17}$$

for $i = 1, 2, \dots, n,$

$$\beta = \frac{\frac{1}{n} \sum_{i=1}^n C_{i,P}}{\frac{1}{n} \sum_{i=1}^n C_{i,S}}, \tag{13}$$

$$MBE = \frac{1}{n} \sum_{i=1}^n (C_{i,P} - C_{i,S}), \tag{18}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (C_{i,S} - C_{i,P})^2}, \tag{14}$$

$$PBIAS = \frac{\sum_{i=1}^n (C_{i,P} - C_{i,S})}{\sum_{i=1}^n C_{i,S}} \times 100, \tag{19}$$

where $C_{i,S}$ represents the simulated salinity concentrations derived from the simulation model and is measured in milligrams per liter (mg/L), $C_{i,P}$ represents the AutoML-based surrogate model predicted salinity concentrations (mg/L),

\overline{C}_S denotes the average or mean simulated salinity concentrations as calculated by the simulation model (mg/L), \overline{C}_P symbolizes the mean AutoML-based surrogate model predicted salinity concentrations (mg/L), n represents the total number of input–output data sets, ED indicates Euclidian distance from the ideal data points, α signifies the corresponding differences between the simulated and model predicted salinity concentrations, β represents the ratio of the average predicted and average simulated salt concentration levels, reflecting the bias in the predictions.

Application of the proposed methodology

The proposed methodology was assessed using the simulation model and data presented in Roy and Datta (2020b). For the ease of readership, brief descriptions of the study area, data, and model calibration process are presented in the following sections.

Study area and data

The research domain is situated within the Barguna district, which is part of the coastal regions of Bangladesh. It is positioned in the southern region of the country, spanning the geographical coordinates of 21° 48' to 22° 29' north latitudes and 89° 52' to 90° 22' east longitudes. In this study area, the coastal aquifer is experiencing contamination due to salinity intrusion caused by extensive groundwater extraction. The primary source of water for agricultural purposes in this region is groundwater, used for cultivating major crops, including rice, wheat, and potatoes. The study area encompasses two administrative units, known as upazillas, within the Barguna district: Patharghata covering an area of 258.63 square kilometers and Barguna Sadar spanning 339.54 square kilometers. To the south, this region is bordered by the Bay of Bengal. Figure 3 presents a visual representation of the research domain.

Aquifer material layers were selected based on lithological data from the study area. Up to a depth of 300 m, the area consists mainly of alluvium soil types, including clay, silt, sand, and occasional gravel (Faneca Sánchez et al. 2015). Given that most physical processes occur within the first few meters of the aquifer, a thickness of 150 m was chosen. This thickness was divided into four layers: sandy silt (40 m), sandy loam (50 m), sand (40 m), and sandy clay (20 m). Each layer was assumed to have homogeneous aquifer material, with only vertical heterogeneity in terms of hydraulic conductivity considered. These values align with previous studies in the Bengal Delta (Faneca Sánchez et al. 2015; Michael and Voss 2009). An anisotropy ratio (k_x/k_y)

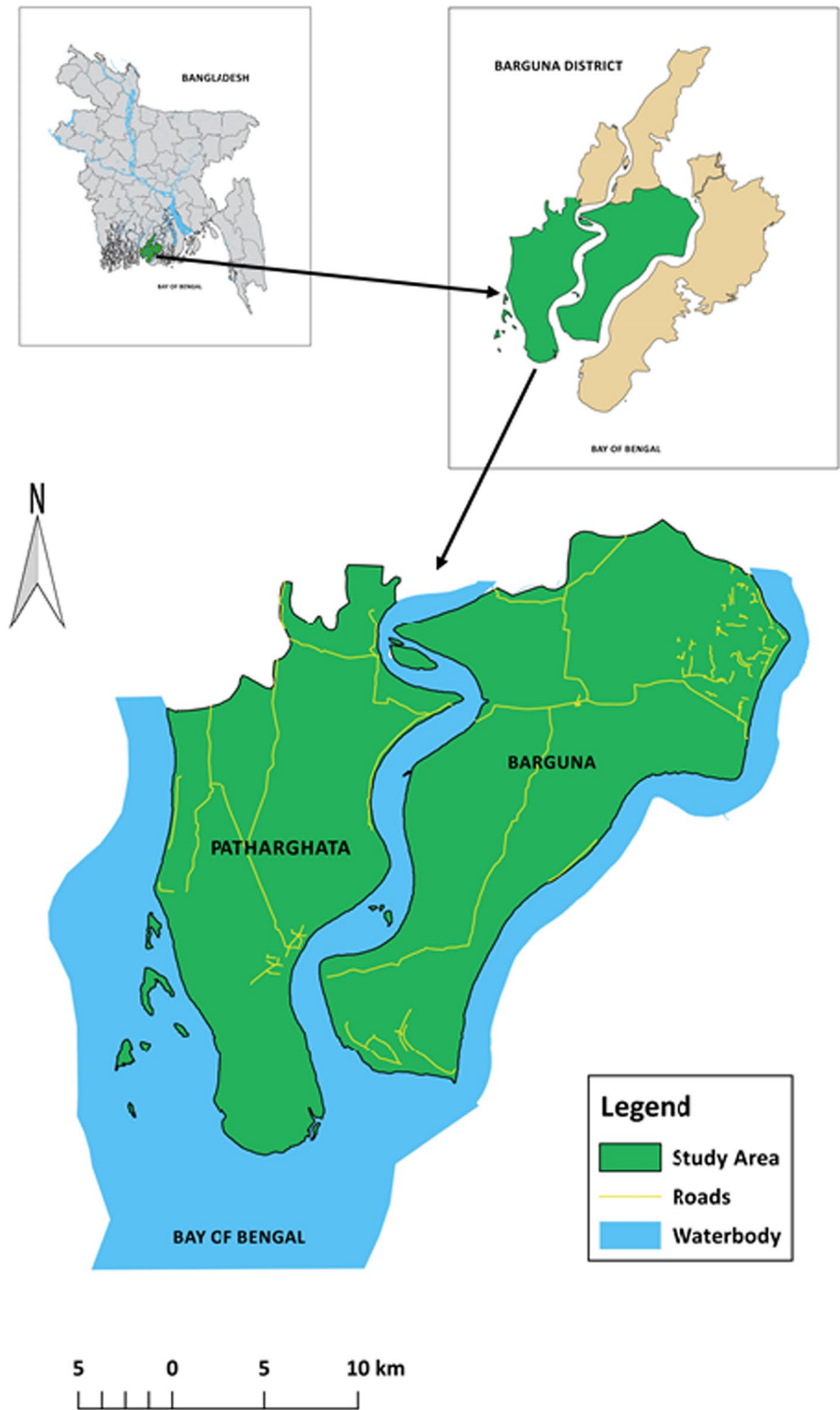
of 2.0 was applied, with k_x representing horizontal hydraulic conductivity in the X-direction, and k_y representing horizontal hydraulic conductivity in the Y-direction. The vertical hydraulic conductivity (k_z) was set as one-tenth of the values in the X-direction. Figure 4 illustrates a 3-D view of the model domain with finite-element meshes.

The Northern boundary, defined as the administrative boundary, was designated as a no-flow boundary. While conclusively determining it as such is challenging, this condition was assumed due to the negligible hydraulic gradient (approximately 1: 20,000) near this boundary in the study area. Consequently, lateral groundwater movement across this boundary (illustrated in Fig. 3) was considered insignificant (Faneca Sánchez et al. 2015), leading to its designation as a no-flow boundary for modeling purposes. The sea face boundary was set as both a constant head and constant concentration boundary, with values of zero meters above sea level (MSL) and 35,000 mg/l, respectively. Upstream ends of rivers were assigned specified head values that linearly varied along the streams, reaching zero meters at the sea face boundary. Specific head values of 0.8 m, 0.86 m, and 0.70 m were designated at the upstream ends of the Haringhata, Bishkhali, and Burishwar Rivers, respectively. Tidal river salinity concentrations remained constant at 10,000 mg/l throughout the simulation period.

Figure 5 depicts a plan view of the study area, delineating boundaries and wells. In this figure, production wells, barrier wells, and monitoring locations are denoted by P1–P43, B1–B13, and M1–M16, respectively. During calibration and validation, pumping from barrier wells was disregarded, and hydraulic heads were monitored at M1 and M2. Following successful calibration and validation, barrier extraction wells were implemented as hydraulic control measures for saltwater intrusion processes. In addition, 14 additional monitoring locations were employed to track salinity concentrations for developing the saltwater intrusion management model in the study area.

The calibration process began by establishing steady-state hydraulic head conditions across the finite-element nodes of the model domain. This involved running the transient simulation model over an 80-year period (from April 1930 to April 2009) in stages of 10 years each, using average pumping rates. The outputs at the end of each 10-year interval served as initial conditions for the subsequent stages until a stable hydraulic head condition was reached by April 2009. These stable hydraulic head values across different nodes were then utilized as initial conditions for the calibration process. The calibrated model was subsequently validated over a 3-year period from April 2015 to April 2017, using outputs from April 2014 as initial conditions for the validation period.

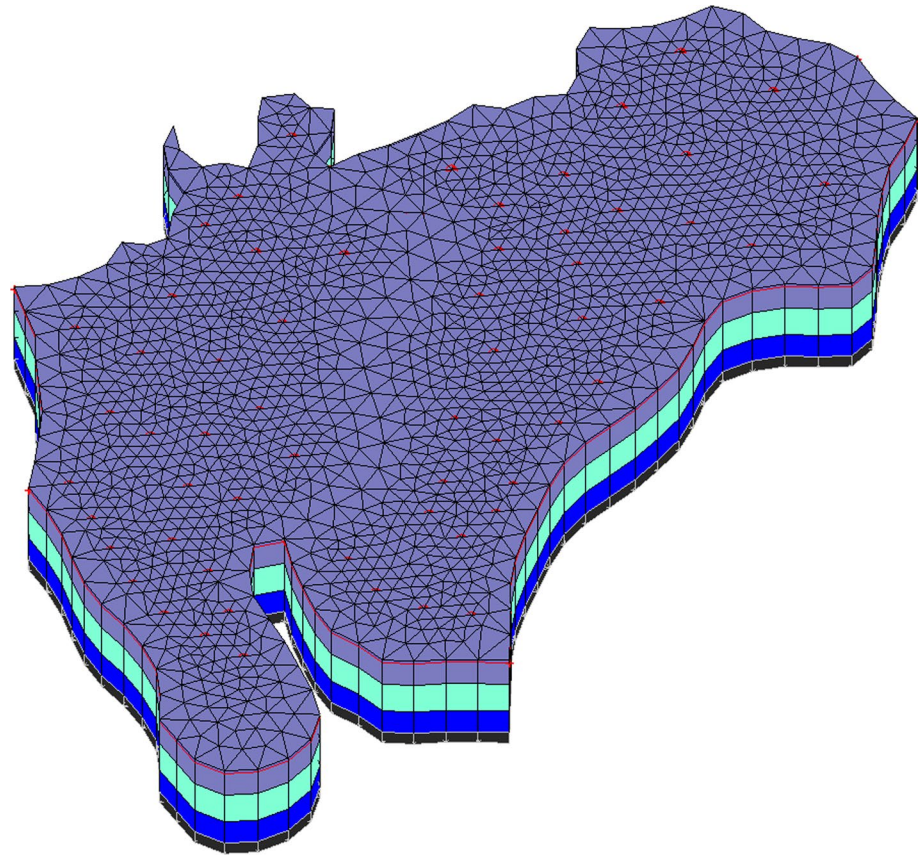
Fig. 3 Location and aerial map of the study area



The process of model calibration spanned five years, covering the period from April 2010 to April 2014. During this time frame, hydraulic head measurements were consistently recorded at the specified MLs in April of each year (2010, 2011, 2012, 2013, and 2014). This calibration effort involved

the refinement of hydraulic conductivity and recharge values to bring the simulated hydraulic heads into closer alignment with the observed hydraulic head values. Following the calibration phase, the model was authenticated throughout the subsequent three years, spanning from 2015 to 2017. The

Fig. 4 Three-dimensional view of the study area (after Roy and Datta 2020b)



hydraulic head values recorded in 2014 were employed as the starting settings for the simulation during the validation phase. During this validation stage, the boundary conditions of the model were kept consistent with those established during the calibration phase.

The calibration and validation processes utilized a uniform time step of 5 days to analyze hydraulic heads. Computational efficiency is crucial when conducting numerous simulations for training a meta-model in an integrated S–O approach. Hence, a sensitivity analysis was conducted to assess the impact of simulation time steps on computed hydraulic heads. Time steps of 1 day, 5 days, 10 days, and 73 days were employed during the calibration periods from April 2010 to April 2014. Results revealed negligible differences in hydraulic head estimates across varying time steps during calibration. However, significant computational efficiency was observed with a 73-day time step, reducing simulation times considerably. Specifically, simulation times for time steps of 1 day, 5 days, 10 days, and 73 days were 14.61 min, 4.82 min, 3.17 min, and 0.95 min, respectively. Consequently, a time step of 73 days was adopted for multiple simulations with different transient pumping values, facilitating the generation of input–output training patterns for surrogate model training.

The calibrated and validated model was subjected to multiple simulations with varying spatiotemporal groundwater pumping values. This was done to capture a diverse range of predictor–response data sets needed to train the surrogate models. The stress imposed on the coastal aquifer due to spatiotemporal groundwater withdrawal resulted through the usage of both pumping and barrier wells situated at specified places and time intervals. The spatiotemporal water pumping amounts served as inputs to the numerical model to predict salinity concentrations at indicated MLs. This process generated a series of paired data sets, consisting of the spatiotemporal groundwater pumping values as predictors and the corresponding salinity concentration amounts as responses. Salinity measurements were collected from 16 MLs, and multiple such data sets were created by running the aquifer simulations numerous times, each time using diverse groups of spatiotemporal water pumping values derived from both pumping and barrier wells. Throughout these simulations, the aquifer properties, along with the other simulation conditions, stayed consistent. The only variable factor was the varying transient groundwater extraction values. This approach allowed us to capture various scenarios representing the obtained salinity concentrations exclusively owing to the applied spatiotemporal groundwater abstraction

Fig. 5 Plan view of the study area showing the boundaries and wells (after Roy and Datta 2020b)

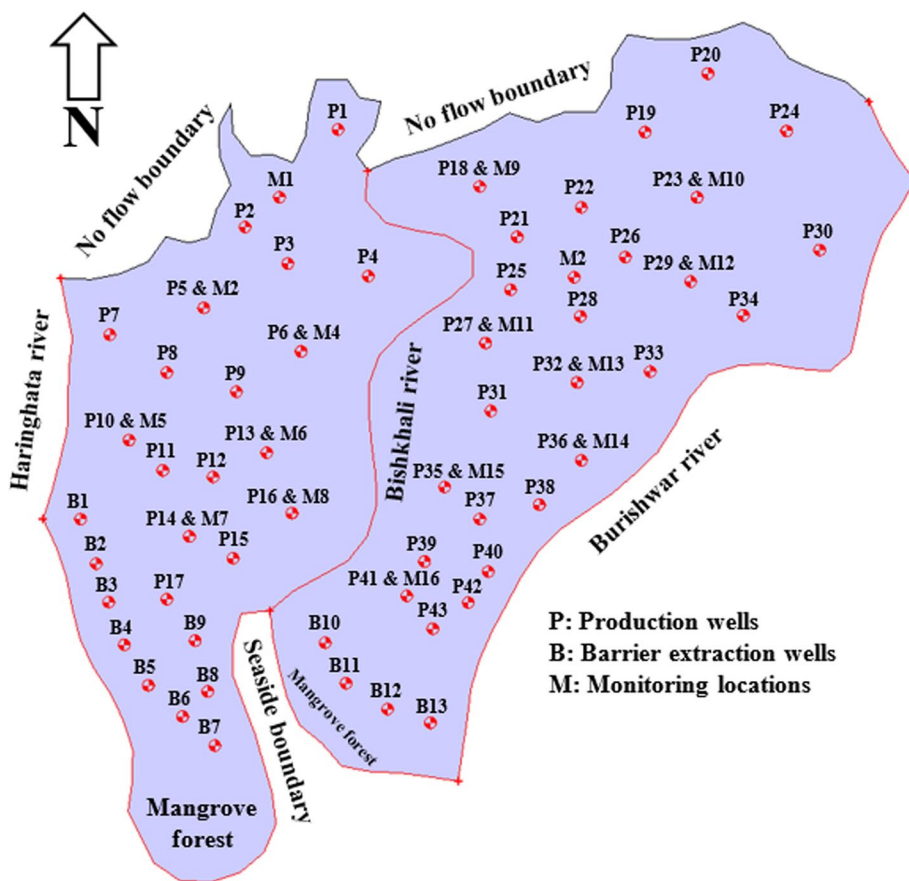


Table 4 Statistics of the obtained salinity concentrations at the monitoring locations

Monitoring locations	Descriptive statistics, mg/L			
	Mean	Standard deviation	Skewness	Kurtosis
ML1	2584.81	0.06	-0.01	-0.73
ML2	2346.92	0.06	0.00	-0.38
ML3	3044.19	10.68	-0.06	-0.36
ML4	3239.45	11.53	-0.06	-0.37
ML5	4979.61	21.71	-0.06	-0.48
ML6	3428.25	12.22	-0.09	-0.37
ML7	3821.92	13.79	-0.07	-0.37
ML8	3638.37	12.80	-0.05	-0.36
ML9	2487.49	8.60	-0.03	-0.34
ML10	2403.23	8.32	-0.03	-0.34
ML11	2733.83	9.50	-0.05	-0.35
ML12	2358.32	8.31	-0.06	-0.36
ML13	2819.77	9.85	0.00	-0.35
ML14	3228.97	11.24	-0.03	-0.33
ML15	3228.36	11.70	-0.06	-0.33
ML16	4194.71	15.94	-0.05	-0.39

stress on the aquifer. Table 4 displays the descriptive statistics of the salinity concentrations produced at the 16 MLs in response to the spatiotemporal groundwater abstraction stress imposed to the aquifer.

Evaluation of the developed surrogate models

To assess the effectiveness of surrogate models in mimicking the aquifer processes within the study area, various statistical metrics were employed. A data set comprising 5000 predictor–response arrays was utilized for both training and validation purposes. Of these predictor–response arrays, 80% were allocated to train the surrogate models, while the other 20% were set aside to assess the effectiveness of the models that were generated. The RMSE and MAE criteria during the training and testing phase were employed to assess the efficiencies of the selected best surrogate models at different MLs. The 12 statistical performance indicators were computed using the test data set, which included 1000 predictor–response arrays. These indices were categorized into two groups: benefit indices (where higher values are favorable) and cost indices (where lower values are favorable). The

Table 5 Training and test errors during training of the selected models

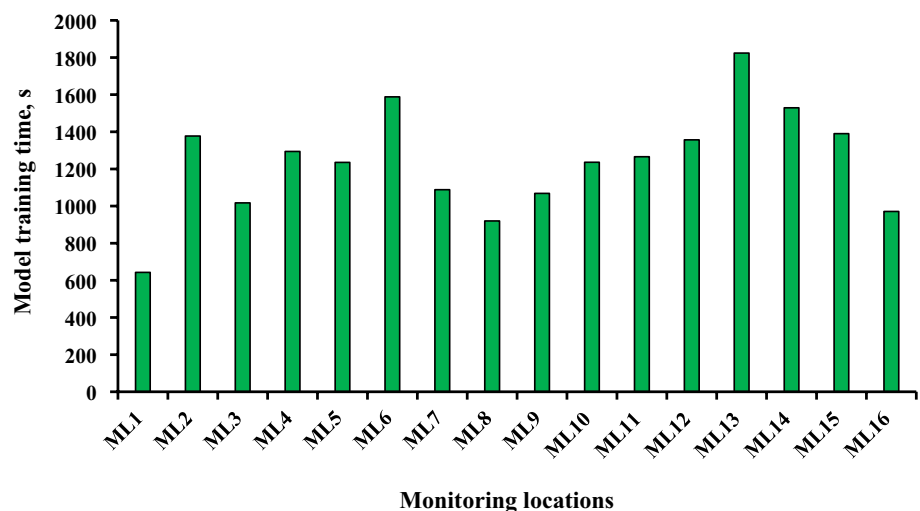
Monitoring locations	RMSE, mg/L		MAE, mg/L	
	Training phase	Testing phase	Training phase	Testing phase
ML1	0.0003	0.0003	0.0002	0.0002
ML2	0.0005	0.0005	0.0004	0.0004
ML3	0.4647	0.4751	0.3765	0.3840
ML4	0.4559	0.4700	0.3677	0.3807
ML5	0.0362	1.4987	0.0293	1.2317
ML6	0.0274	0.6359	0.0219	0.5112
ML7	0.6320	0.6553	0.5033	0.5286
ML8	0.0317	0.5642	0.0257	0.4566
ML9	0.3535	0.3527	0.2842	0.2857
ML10	0.3484	0.2825	0.3611	0.2915
ML11	0.0213	0.4769	0.0172	0.3885
ML12	0.0194	0.3955	0.0157	0.3198
ML13	0.4605	0.4817	0.3746	0.3908
ML14	0.0323	0.4182	0.0261	0.3311
ML15	0.5562	0.6333	0.4311	0.4963
ML16	0.6664	0.7077	0.5345	0.5754

benefit performance indices include accuracy, R, NS, IOA, and KGE, whereas the cost indices used for the evaluation were the RMSE, NRMSE, MAPRE, MAE, MAD, MBE, and PBIAS. In general, the surrogate models yielded greater values for the benefit indices and smaller values for the cost indices at the designated 16 MPs. These outcomes indicate the effectiveness of the proposed surrogate models in capturing the predictor–response relationships between spatiotemporal water abstractions and the resulting salinity concentration values at the designated MPs.

Training performance of the selected best models

The errors during the training and testing phases were computed and compared to assess the training performance and guarantee that the models developed at various MPs were neither over fit nor under fit. Table 5 provides training and test errors noticed during the training of selected models at specific MPs, as demonstrated by the RMSE and MAE values. The training phase error reflects the degree to which the models match or represent the training data, while the testing phase error provides insights into their generalization to new, unseen data. Smaller disparities between the errors in the training and testing phases indicate better model performance. Table 5 shows that all models exhibit very low errors during the phases of testing and training, suggesting excellent predictive performance. Moreover, the disparities between the errors in the training and testing phases are minimal, indicating no overfitting during model development. This holds true for models developed at all MLs, except for ML5, ML6, ML8, ML11, ML12, and ML14, where the differences between the training and testing phase errors were slightly higher. However, it is worth noting that the testing errors at these MLs were relatively smaller, indicating the models' good generalization capability. Overall, Table 5 serves as a valuable reference for assessing the efficacy of the selected models in terms of training performance, focusing on error metrics for both training and testing phases.

The training time requirements for developing ML-based surrogate models are of significant importance when considering applicability and scalability. Scalability becomes more accessible with models that train quickly, enabling them to handle large data sets and real-time demands efficiently. In this study, we developed seawater intrusion prediction models to support the creation of a

Fig. 6 Training time requirements for models developed at different monitoring locations

combined S–O-based management model for controlling seawater intrusion in coastal aquifers. These prediction models serve as effective surrogates for the complicated seawater intrusion simulation model. Therefore, it is imperative that the prediction models remain as simple as possible, with minimal training time requirements. Figure 6 offers a visual depiction of the training time requirements for models developed at various MLs. This figure offers valuable insights into the computational resources and time investments needed for effective model training. It serves as a visual aid for understanding how training time varies across different locations and can inform decision-making regarding model development strategies. Observations from Fig. 6 reveal that training times vary for models developed at different MLs. The time required for training these models ranged from as low as 643 s at ML1 to as high as 1824s at ML13, depending on the numerical values of the training data set at these MLs. Considering a substantial training data set of 5000 input–output patterns with 168 input variables, these training times remained relatively short when using a moderately configured computer (Intel(R) Core (TM) i7-8700 CPU @ 3.20 GHz 3.19 GHz & 16.0 GB RAM). Following successful training, the trained models demonstrated the ability to produce predictions for a new, unseen data set within a fraction of a second. This capability of these developed models enables their utilization as effective substitutes for the intricate simulation model inside the framework of a combined S–O-assisted management model for controlling seawater intrusion.

Performance assessment of the selected best models on an unidentified data

After appropriately training the selected best models at various MLs, these models underwent testing with a new, unseen test data set, employing 12 statistical performance evaluation indices categorized into benefit and cost indices. The outcomes are presented in Table 5 (benefit indices) and Table 6 (cost indices). Table 6 provides a detailed assessment of the selected best models' performance on an unseen test data set across a range of MLs with respect to accuracy, R, NS, IOA, and KGE criteria. The automatically selected models developed at the designated MLs exhibited high performances with respect to the computed benefit indices. The benefit indices attained similar performances while different benefit indices varied in magnitude among different MLs. All models exhibited superior performances based on accuracy (= 1), R (~ 0.99), NS (~ 0.99), IOA (~ 0.99) and KGE (~ 0.99) criteria. These findings align well with those presented in Roy and Datta (2020b), who reported a similar performance using a MARS-based seawater intrusion prediction model. An

ideal accuracy value is 1.0, and any value close to 1.0 indicates excellent model performance (Elbeltagi et al. 2020). The R and IOA values exceeding 0.8 suggest very good performance of the model (Willmott 1981; Kirch 2008), while NS values exceeding 0.8 also suggest excellent model performance (Nash and Sutcliffe 1970). The model performances were deemed excellent with according to the KGE criterion with KGE values very close to 1 (Gupta et al. 2009; Kling et al. 2012). It can be argued that the selected models can effectively represent the numerical simulation model as the benefit indices showed indications of high model performance. Consequently, these models can function effectively as suitable surrogates for the intricate simulation model within the context of the combined S–O-based management model for controlling seawater intrusion.

Table 7 presents a comprehensive evaluation of the effectiveness of the top-selected surrogates on an unobserved data set, encompassing various MLs, in relation to the computed cost evaluation indices. Models developed at different MLs exhibited equivalent performance when evaluated using the specified cost evaluation indices. In general, the numerical values of all cost indices were remarkably small, indicating the exceptional performance of the surrogate models (Heinemann et al. 2012; Hyndman and Koehler 2006; Legates and McCabe 1999; Li et al. 2013; Pal 2017; Pham-Gia and Hung 2001; Yapo et al. 1996). The computed cost evaluation indices, including RMSE, NRMSE, MAPRE, MAE, MAD, MBE, and

Table 6 Performance assessment of the selected best models on an unused test data set based on benefit indices

Monitoring locations	Performance indices (benefit)				
	Accuracy	R	NS	IOA	KGE
ML1	1.0000	1.0000	1.0000	1.0000	0.9998
ML2	1.0000	1.0000	0.9999	1.0000	0.9997
ML3	1.0000	0.9990	0.9980	0.9995	0.9987
ML4	1.0000	0.9992	0.9983	0.9996	0.9991
ML5	1.0000	0.9977	0.9954	0.9988	0.9968
ML6	1.0000	0.9987	0.9973	0.9993	0.9979
ML7	1.0000	0.9989	0.9978	0.9994	0.9973
ML8	1.0000	0.9990	0.9981	0.9995	0.9957
ML9	1.0000	0.9992	0.9983	0.9996	0.9973
ML10	1.0000	0.9991	0.9982	0.9995	0.9983
ML11	1.0000	0.9988	0.9976	0.9994	0.9965
ML12	1.0000	0.9988	0.9977	0.9994	0.9981
ML13	1.0000	0.9988	0.9976	0.9994	0.9974
ML14	1.0000	0.9993	0.9986	0.9996	0.9952
ML15	1.0000	0.9986	0.9972	0.9993	0.9952
ML16	1.0000	0.9991	0.9981	0.9995	0.9990

Table 7 Performance assessment of the selected best models on an unused test data set based on cost indices

Monitoring locations	Performance indices (cost)						
	RMSE, mg/L	NRMSE, %	MAPRE, %	MAE, mg/L	MAD, mg/L	MBE, mg/L	PBIAS, %
ML1	0.0003	0.0000	0.0000	0.0002	0.0001	0.0001	0.0000
ML2	0.0005	0.0000	0.0000	0.0004	0.0002	0.0001	0.0000
ML3	0.4751	0.0156	0.0126	0.3840	0.1917	-0.0022	-0.0001
ML4	0.4700	0.0145	0.0118	0.3807	0.1920	-0.0055	-0.0002
ML5	1.4987	0.0301	0.0247	1.2317	0.5739	-0.0744	-0.0015
ML6	0.6359	0.0185	0.0149	0.5112	0.2606	-0.0013	0.0000
ML7	0.6553	0.0171	0.0138	0.5286	0.2672	0.0067	0.0002
ML8	0.5642	0.0155	0.0126	0.4566	0.2383	-0.0067	-0.0002
ML9	0.3527	0.0142	0.0115	0.2857	0.1373	-0.0101	-0.0004
ML10	0.3611	0.0150	0.0121	0.2915	0.1567	-0.0009	0.0000
ML11	0.4769	0.0174	0.0142	0.3885	0.1921	-0.0048	-0.0002
ML12	0.3955	0.0168	0.0136	0.3198	0.1560	-0.0099	-0.0004
ML13	0.4817	0.0171	0.0139	0.3908	0.2004	0.0125	0.0004
ML14	0.4182	0.0130	0.0103	0.3311	0.1699	-0.0097	-0.0003
ML15	0.6333	0.0196	0.0154	0.4963	0.2537	0.0479	0.0015
ML16	0.7077	0.0169	0.0137	0.5754	0.2633	0.0137	0.2633

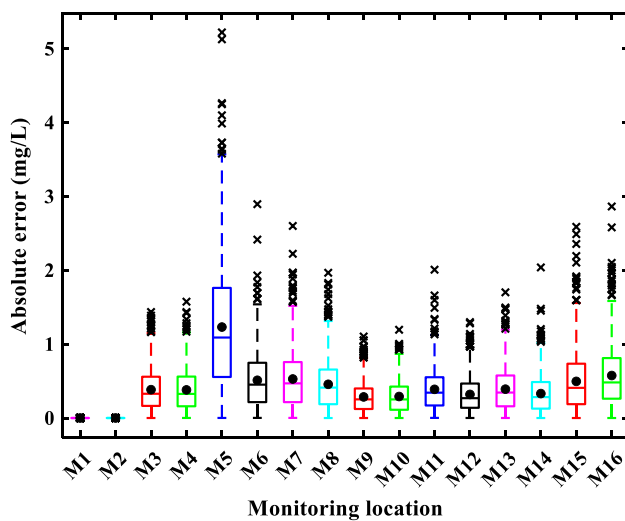


Fig. 7 Absolute error boxplots of the models selected at different monitoring locations. M1, M2, M3, ..., M16 represent the top-performing models selected for the monitoring locations ML1, ML2, ML3, ..., ML16. A horizontal line inside each box indicates the median absolute error between the actual and predicted salinity concentrations for the respective prediction model. The black circle denotes the mean absolute error, while the 'x' symbols represent the outliers

PBIAS, strongly suggest that the developed surrogate models authentically represent the numerical simulation model from which they were derived. Consequently, these models emerge as the prime candidates for serving

as computationally efficient alternatives to the simulation model in the context of the framework of a combined S–O approach for the creation of a management model for controlling seawater intrusion

The prediction accuracies of the selected top-performing models at different MLs, as presented in the absolute error boxplots (Fig. 7), reveal that all models generated lower absolute errors between the simulation model outputs and surrogate predicted saltwater concentrations. However, it is noteworthy that models M1, M2, M3, M4, M9, M10, M12, and M14 exhibited greater efficiency in comparison with other surrogate models. Furthermore, Fig. 7 illustrates that the prediction accuracies of models M1 and M2 surpassed those of the other models, while the performance of model M5 was the poorest among all the models. These performance variations may be attributed to the dependence of model performance on the numerical values and data properties used for training. In addition, optimal parameter sets for the selected best-performing models, determined by the ASHA optimization algorithm, can impact performance. However, it is important to emphasize that each iteration of the ASHA optimization used similar optimization parameters and the same random seeds in the implementation engine (MATLAB) to minimize model selection errors. Therefore, it becomes apparent that the training data (input–output training patterns) at various MLs considerably affect the effectiveness of the selected best-performing models.

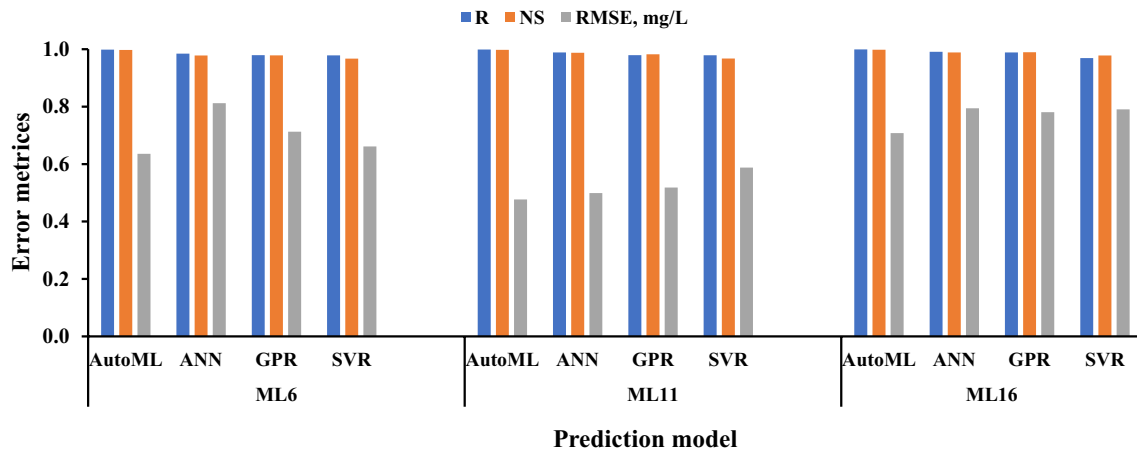


Fig. 8 Performance comparison of the proposed AutoML approach with ANN, GPR, and SVR models

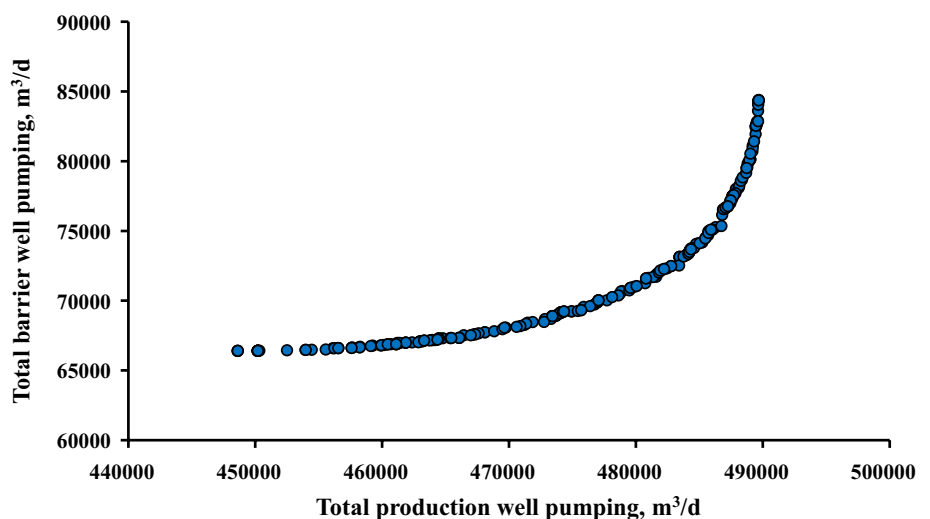
Performance comparison of the proposed AutoML approach with ANN, GPR, and SVR models

Figure 8 presents a comparison of the performance of the best-selected models against the subjectively chosen ANN, GPR, and SVR models using the same unseen test data set across three monitoring locations (ML6, ML11, and ML16) with respect to R, NS, and RMSE criteria. The results reveal that the proposed AutoML approach outperforms the subjectively selected ANN, GPR, and SVR models based on the computed performance evaluation criteria. This performance-centric analysis highlights the superiority or parity of the proposed AutoML approach over subjectively chosen models, offering valuable insights into the evolving landscape of machine learning model selection and optimization. These findings have the potential to redefine best practices in predictive modeling.

Practical implications

The practical importance of this research resides in its capability to demonstrate the capability of an AutoML-based ML approach to predict seawater intrusion in coastal aquifer systems. The proposed AutoML-based prediction models (surrogates) can be used as computationally viable proxies of the complex simulation models in a combined S–O method to create management plans for regional-scale coastal aquifers within intricate aquifer systems. To demonstrate this, a management model aimed at controlling seawater intrusion was developed by integrating the selected best performing models at the 16 MLs with a controlled elitist multiple objective genetic algorithm (CEMOGA) (Deb and Goel 2001). The suggested management model takes into account 168 decision variables that encompass spatiotemporal water extraction rates. Specifically, variables X-1 to

Fig. 9 Pareto optimal front



X-129 denote groundwater extraction from pumping wells, while variables X-130 to X-168 pertain to barrier extraction bores. For instance, variables X-1 to X-3 indicate groundwater extraction from the pumping well P-1 during the first, second, and third time periods, respectively. Similarly, variables X-4 to X-6 represent groundwater extraction from the pumping well P-2, and so forth, spanning three designated time periods. On the other hand, variables X-130 to X-132, X-133 to X-135, X-136 to X-138, and so forth, up to X-166 to X-168, correspond to groundwater withdrawal from barrier extraction wells B-1, B-2, B-3, and so on, throughout the initial, second, and third time periods, in that order.

The 16 selected models, which predict salinity at specified MLs, were integrated into CEMOGA to forecast salinity within the combined S–O framework for the development of the seawater intrusion management model. In addition, these externally linked surrogate models were employed to verify the constraint satisfactions concerning the highest salinity levels that are permitted at particular MLs. With CEMOGA, the highest generation size was 16,000, the crossover percentage was 0.92, the population size was 2000, and the Pareto front population fraction was 0.70. The value of function tolerance was configured at 1×10^{-5} , while the constraint tolerance configured at 1×10^{-4} . The best parameters for CEMOGA were selected after several trials, even though an exhaustive sensitivity assessment was not performed. To find the optimum solution, the optimization procedure examined 116,001 functions across 116 generations. It took 142 min for the optimization routine to reach to the optimal solutions. The management model offered optimal solutions for the seawater intrusion management issue using a Pareto optimum front comprising various feasible options that demonstrate the trade-off between opposing goals. Figure 9's Pareto optimum front illustrates that increasing water extraction from barrier wells can lead to higher water withdrawal from pumping wells.

The results obtained from our solution indicate that this approach is not only feasible but also reasonably accurate. Importantly, this methodology can potentially be adapted for use in other aquifer systems along the coast. However, it is essential to note that for every new application, the seawater intrusion models must undergo calibration and validation specific to the research site. In summary, even though initially constructed for a particular aquifer system, our suggested approach holds promise for broader application across different coastal aquifer contexts.

The findings presented in the manuscript hold broader implications that extend beyond the immediate scope of saltwater intrusion prediction. The AutoML methodology developed in this study demonstrates its adaptability and potential applicability to various fields within ML-based

and environmental modeling. With the proposed modeling approach's enhanced predictive accuracy and efficiency, it could be employed in other environmental studies, hydrogeological assessments, and water resource management. The insights gained from this research contribute to the advancement of ML-based techniques, fostering their integration into diverse scientific domains for more effective and informed decision-making processes. Its adaptability allows for application in various environmental prediction scenarios, such as groundwater contamination in different geological settings or the forecasting of other complex hydrological phenomena. This cross-disciplinary applicability underscores the broader impact and significance of the research findings, opening avenues for advancements in predictive modeling and decision support systems across diverse scientific and environmental disciplines.

In addition, the proposed computationally effective proxies (AutoML-based prediction models) of the numerical simulation code, FEMWATER can be further employed to develop management strategies in other related fields of hydrology and water resources management. Furthermore, prediction models developed using the proposed AutoML technique can be integrated with various multiple objective optimization algorithms to obtain Pareto optimal solutions. These solutions will aid decision-makers in selecting the best alternatives from a range of feasible options. The proposed methodology can be extended to other conditions by employing the trained and validated prediction models on new, unseen data sets representing those conditions. This process will also confirm the generalization capability of the developed AutoML-based saltwater prediction models. This has been performed by testing the model's generalization capability on an unseen test data set for this case study.

Validation of the surrogate-based coupled S–O approach

The accuracy of the surrogate-based saltwater intrusion management model was evaluated by monitoring the adherence to defined constraints. It was observed that the saltwater concentrations derived from the optimization model's solution, facilitated by surrogate models within the optimization framework, consistently remained below the predetermined maximum allowable levels at all monitoring sites. This indicates the fulfillment of imposed constraints, with no violations detected throughout the optimization process. Furthermore, the obtained saltwater concentrations closely aligned with the specified values, suggesting convergence of the optimization model towards the upper bounds of the imposed constraints.

Table 8 Surrogate model predicted and FEMWATER simulated salinity concentrations calculated from optimal solutions of groundwater abstraction

Obs	Saltwater concentration, mg/l									
	M1		M2		M3		M4		M5	
	Surrogates	SM	Surrogates	SM	Surrogates	SM	Surrogates	SM	Surrogates	SM
1	2584.752	2584.809	2346.861	2346.853	3037.917	3039.298	3234.157	3232.618	4919.685	4920.600
2	2584.750	2584.808	2346.862	2346.851	3037.807	3038.998	3233.802	3231.984	4914.364	4920.597
3	2584.771	2584.815	2346.871	2346.894	3043.730	3042.171	3238.282	3238.060	4929.148	4920.617
4	2584.766	2584.811	2346.869	2346.875	3040.954	3041.465	3237.700	3236.718	4928.352	4920.613
5	2584.769	2584.815	2346.869	2346.890	3041.405	3041.737	3237.482	3237.880	4928.702	4920.611

M monitoring locations, *SM* simulation model

In the second stage, the efficacy of the optimal groundwater abstraction strategies derived from the optimization model was validated through comparison with results obtained from FEMWATER. Five solutions were randomly selected from various regions of the Pareto optimal front for this analysis. Table 8 presents a comparison between the saltwater concentration values predicted by the surrogate model and those simulated by FEMWATER, utilizing the optimal groundwater abstraction strategies prescribed by the saltwater intrusion management model. The results for 5 monitoring locations are displayed in Table 8, with similar trends observed at other monitoring locations. It is evident from Table 8 that the solutions yielded by the FEMWATER model closely aligned with the predictions of the surrogate model.

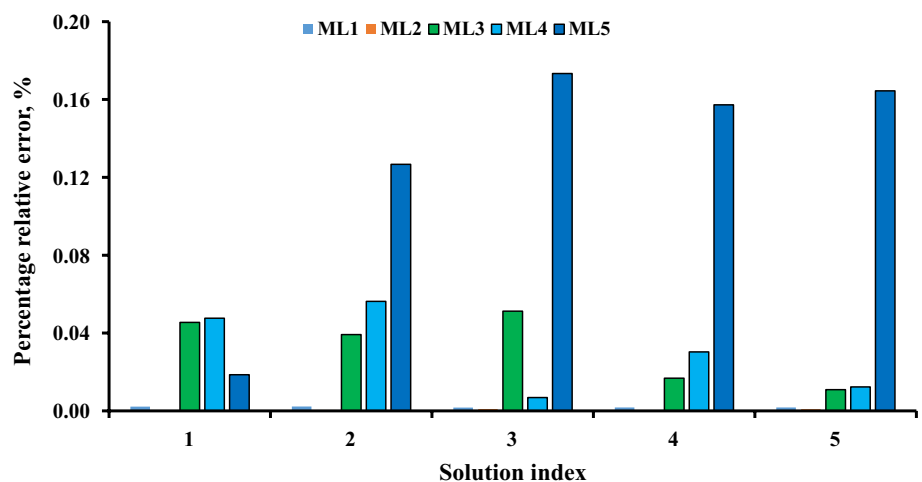
The percentage relative errors between the surrogate predicted and FEMWATER simulated salinity concentrations using five randomly selected optimal groundwater abstractions is presented in Fig. 10. The relative percentage errors, consistently below 3% in all instances, indicate a highly satisfactory constraint satisfaction outcome and a strong correlation between the simulated and computed

groundwater head values. This finding aligns well with the results reported by Roy and Datta (2017b), who suggested that a relative percentage error of less than 5% is deemed suitable and efficient for achieving an optimal management strategy using an integrated S–O approach.

Conclusions

Utilizing a combined S–O approach to devise an optimum groundwater withdrawal strategy has proven effective in mitigating seawater intrusion within coastal aquifers. While various ML-based surrogate models have recently emerged as computationally efficient alternatives to simulation models for addressing seawater intrusion issues, there has been a notable absence of studies focusing on the automatic selection of these models through optimization algorithms. This study fills this gap by demonstrating the applicability of the ASHA optimization algorithm to automatically select the surrogate models that can aid in developing surrogate assisted combined S–O method to solve complex

Fig. 10 Percentage relative errors between the surrogate predicted and FEMWATER simulated salinity concentrations using optimal groundwater abstractions



coastal aquifer management issues. In our investigation, we employed FEMWATER, a numerical code that relies on finite-element methods and is designed for three-dimensional simulations of integrated flow and salt transport. This code was employed to model and analyze the mechanisms behind seawater intrusion within an aquifer system situated in the coastal areas of Bangladesh. To support our analysis, a diverse set of input data was gathered from various sources, encompassing a study area of approximately 598 square kilometers.

To address the challenges of selecting the most appropriate prediction models, this study proposes an innovative approach for selecting the best surrogate models using ASHA hyperparameter tuning method to perform automated model selection. Notably, this study is the first to utilize the ASHA algorithm for automating the model selection process to provide accurate predictions of salinity intrusion. Top-performing models at the 16 MLs were identified. The best models chosen for different monitoring locations demonstrated exceptional performance, demonstrating high accuracy with metrics, such as accuracy (= 1), R (~0.99), NS (~0.99), IOA (~0.99), and KGE (~0.99), all close to 1 and indicative of excellent model accuracy. In addition, the models exhibited a narrow range of RMSE values, ranging from 0.0003 to 1.4987 mg/L. This limited range is considered favorable for any predictive modeling approach. Moreover, the identified top-performing models at diverse monitoring locations can be utilized to formulate effective management strategies for determining optimal groundwater pumping rates to mitigate saltwater intrusion. In addition, the prediction models identified through the proposed AutoML technique can be seamlessly integrated with various multiple objective optimization algorithms, resulting in the derivation of Pareto optimal solutions. These solutions, characterized by a balance of trade-offs, will assist decision-makers in the selection of optimal alternatives from a spectrum of feasible options. The evaluation findings show that the ASHA optimization algorithm can be effectively employed to create accurate and robust seawater intrusion prediction models, making them suitable for developing a surrogate-assisted seawater intrusion management plan applicable to complex coastal aquifer study areas.

This study makes a significant contribution to the field of hydrogeology and environmental modeling. Unlike existing research, this study introduces a novel AutoML methodology within a machine learning framework for predicting seawater intrusion into coastal aquifers. The approach not only enhances predictive accuracy and efficiency but also demonstrates adaptability for broader applications in machine learning and environmental modeling. Compared to similar research endeavors, this study distinguishes itself by its focus on an AutoML strategy, which streamlines the modeling process and ensures the identification of the

most suitable model for the given data set. This innovative methodology contributes to the efficiency and reliability of the predictive models, setting it apart from traditional approaches. Furthermore, the study provides insights into the potential applicability of the developed AutoML technique in diverse fields, fostering advancements in decision-making processes for water resource management.

Author contributions All authors contributed to the study's conception and design. Dilip Kumar Roy: conceptualization, methodology, formal analysis, software, validation, visualization, writing—original draft; Chitra Rani Paul: assisted in analyzing the results, writing—reviewing and editing; Tasnia Hossain Munmun: assisted in writing and editing the manuscript; Bithin Datta: overall supervision, writing—review and editing, provided suggestions for improving the manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Availability of data and materials Data sets and other materials are available with the authors, and may be accessible at any time upon request.

Code availability MATLAB codes are available with the first author and will be provided upon request.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Compliance with ethical standards There are no potential conflicts of interest. The research does not include human participants and/or animals.

Consent to participants The research does not include human participants and/or animals.

Consent to publish The authors give their consent to publish in the Earth Systems and Environment journal if accepted for publication.

References

- Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimization. *J Mach Learn Res* 13:281–305
- Bhattacharjya RK, Datta B (2005) Optimal management of coastal aquifers using linked simulation optimization approach. *Water Resour Manag* 19:295–320. <https://doi.org/10.1007/s11269-005-3180-9>
- Bhattacharjya RK, Datta B, Satish MG (2007) Artificial neural networks approximation of density dependent saltwater intrusion process in coastal aquifers. *J Hydrol Eng* 12:273–282. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2007\)12:3\(273\)](https://doi.org/10.1061/(ASCE)1084-0699(2007)12:3(273))
- Christelis V, Mantoglou A (2016) Pumping optimization of coastal aquifers assisted by adaptive metamodelling methods and radial basis functions. *Water Resour Manag* 30:5845–5859. <https://doi.org/10.1007/s11269-016-1337-3>
- Christelis V, Regis RG, Mantoglou A (2018) Surrogate-based pumping optimization of coastal aquifers under limited computational

- budgets. *J Hydroinform* 20:164–176. <https://doi.org/10.2166/hydro.2017.063>
- Deb K, Goel T (2001) Controlled elitist non-dominated sorting genetic algorithms for better convergence BT—evolutionary multi-criterion optimization. In: Zitzler E, Thiele L, Deb K et al (eds) Springer. Berlin, Heidelberg, pp 67–81
- Dhar A, Datta B (2009) Saltwater intrusion management of coastal aquifers. I: linked simulation-optimization. *J Hydrol Eng* 14:1263–1272. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000097](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000097)
- Elbeltagi A, Deng J, Wang K et al (2020) Modeling long-term dynamics of crop evapotranspiration using deep learning in a semi-arid environment. *Agric Water Manag* 241:106334. <https://doi.org/10.1016/j.agwat.2020.106334>
- Faneca Sánchez M, Bashar K, Janssen G, Vogels M, Snel J, Zhou Y, Stuurman RJ, Oude Essink G (2015) SWIBANGLA: managing saltwater intrusion impacts in coastal groundwater systems of Bangladesh Final Report. <https://doi.org/10.13140/2.1.2447.0721>. Deltares report number: 1207671–000-BGS-0016
- Feist SE, Hoque MA, Ahmed KM (2023) Coastal salinity and water management practices in the bengal delta: a critical analysis to inform salinisation risk management strategies in asian deltas. *Earth Syst Environ* 7:171–187. <https://doi.org/10.1007/s41748-022-00335-9>
- Forrester AIJ, Sóbester A, Keane AJ (2008) Engineering design via surrogate modelling: a practical guide. John Wiley & Sons Ltd, Oxford
- Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009) Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *J Hydrol* 377:80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Halton JH (1960) On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numer Math* 2:84–90. <https://doi.org/10.1007/BF01386213>
- Heinemann AB, VanOort PAJ, Fernandes DS, Maia ADHN (2012) Sensitivity of APSIM/ORYZA model due to estimation errors in solar radiation. *Bragantia* 71:572–582. <https://doi.org/10.1590/S0006-87052012000400016>
- Hyndman RJ, Koehler AB (2006) Another look at measures of forecast accuracy. *Int J Forecast* 22:679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Jamieson K, Talwalkar A (2015) Non-stochastic best arm identification and hyperparameter optimization
- Jeong J, Park E (2019) Comparative applications of data-driven models representing water table fluctuations. *J Hydrol* 572:261–273. <https://doi.org/10.1016/j.jhydrol.2019.02.051>
- Karnin Z, Koren T, Somekh O (2013) Almost optimal exploration in multi-armed bandits. 30th Int Conf Mach Learn ICML 2013 2275–2283
- Kirch W (2008) Pearson's correlation coefficient. In: Kirch W (ed) Encyclopedia of public health. Springer, Netherlands, pp 1090–1091
- Kling H, Fuchs M, Paulin M (2012) Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J Hydrol* 424–425:264–277. <https://doi.org/10.1016/j.jhydrol.2012.01.011>
- Kopsiaftis G, Protopapadakis E, Voulodimos A et al (2019) Gaussian process regression tuned by bayesian optimization for seawater intrusion prediction. *Comput Intell Neurosci* 2019:2859429. <https://doi.org/10.1155/2019/2859429>
- Kopsiaftis G, Kaselimi M, Protopapadakis E et al (2023) Performance comparison of physics-based and machine learning assisted multi-fidelity methods for the management of coastal aquifer systems. *Front. Water* 5:1195029. <https://doi.org/10.3389/frwa.2023.1195029>
- Kourakos G, Mantoglou A (2009) Pumping optimization of coastal aquifers based on evolutionary algorithms and surrogate modular neural network models. *Adv Water Resour* 32:507–521. <https://doi.org/10.1016/j.advwatres.2009.01.001>
- Lal A, Datta B (2021) Application of the group method of data handling and variable importance analysis for prediction and modelling of saltwater intrusion processes in coastal aquifers. *Neural Comput Appl* 33:4179–4190. <https://doi.org/10.1007/s00521-020-05232-8>
- Legates DR, McCabe GJ Jr (1999) Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour Res* 35:233–241. <https://doi.org/10.1029/1998WR900018>
- Li M-F, Tang X-P, Wu W, Liu H-B (2013) General models for estimating daily global solar radiation for different solar radiation zones in mainland China. *Energy Convers Manag* 70:139–148. <https://doi.org/10.1016/j.enconman.2013.03.004>
- Li L, Jamieson K, Rostamizadeh A et al (2020) A system for massively parallel hyperparameter tuning. In: Proceedings of the 3rd MLSys Conference. Austin, TX, USA
- Lin H-CJ, Richards DR, Talbot CA et al (1997) FEMWATER: a three-dimensional finite element computer model for simulating density-dependent flow and transport in variable saturated media. Technical Rep No CHL-97–12. US Army Engineer Water
- Loyola RDG, Pedergrana M, Gimeno-García S (2016) Smart sampling and incremental function learning for very large high dimensional data. *Neural Netw* 78:75–87. <https://doi.org/10.1016/j.neunet.2015.09.001>
- Masoud M (2020) Groundwater resources management of the shallow groundwater aquifer in the desert fringes of El Beheira Governorate. *Egypt Earth Syst Environ* 4:147–165. <https://doi.org/10.1007/s41748-020-00148-8>
- Michael HA, Voss CI (2009) Controls on groundwater flow in the Bengal Basin of India and Bangladesh: regional modeling analysis. *Hydrogeol J* 17(7):1561. <https://doi.org/10.1007/s10040-008-0429-4>
- Moravej M, Amani P, Hosseini-Moghari S-M (2020) Groundwater level simulation and forecasting using interior search algorithm-least square support vector regression (ISA-LSSVR). *Groundw Sustain Dev* 11:100447. <https://doi.org/10.1016/j.gsd.2020.100447>
- Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—a discussion of principles. *J Hydrol* 10:282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Neumann B, Vafeidis AT, Zimmermann J, Nicholls RJ (2015) Future coastal population growth and exposure to sea-level rise and coastal flooding—a global assessment. *PLoS One* 10:e0118571
- Pal R (2017) Validation methodologies. pp 83–107
- Pham-Gia T, Hung TL (2001) The mean and median absolute deviations. *Math Comput Model* 34:921–936. [https://doi.org/10.1016/S0895-7177\(01\)00109-1](https://doi.org/10.1016/S0895-7177(01)00109-1)
- Probst P, Wright MN, Boulesteix A-L (2019) Hyperparameters and tuning strategies for random forest. *Wires Data Min Knowl Discov* 9:e1301–e1301. <https://doi.org/10.1002/widm.1301>
- Rajabi MM, Ketabchi H (2017) Uncertainty-based simulation-optimization using Gaussian process emulation: application to coastal groundwater management. *J Hydrol* 555:518–534. <https://doi.org/10.1016/j.jhydrol.2017.10.041>
- Ranjbar A, Mahjouri N (2018) Development of an efficient surrogate model based on aquifer dimensions to prevent seawater intrusion in anisotropic coastal aquifers, case study: the Qom aquifer in Iran. *Environ Earth Sci* 77:418. <https://doi.org/10.1007/s12665-018-7592-2>
- Roy DK, Datta B (2018b) Selection of meta-models to predict saltwater intrusion in coastal aquifers using entropy weight based decision

- theory. In: 2018 IEEE Conference on Technologies for Sustainability (SusTech 2018). Long Beach, California, USA
- Roy DK, Datta B (2017a) Fuzzy C-mean clustering based inference system for saltwater intrusion processes prediction in coastal aquifers. *Water Resour Manag*. <https://doi.org/10.1007/s11269-016-1531-3>
- Roy DK, Datta B (2017b) Genetic algorithm tuned fuzzy inference system to evolve optimal groundwater extraction strategies to control saltwater intrusion in multi-layered coastal aquifers under parameter uncertainty. *Model Earth Syst Environ* 3:1707–1725. <https://doi.org/10.1007/s40808-017-0398-5>
- Roy DK, Datta B (2017c) A surrogate based multi-objective management model to control saltwater intrusion in multi-layered coastal aquifer systems. *Civ Eng Environ Syst* 34:238–263. <https://doi.org/10.1080/10286608.2018.1431777>
- Roy DK, Datta B (2018a) A review of surrogate models and their ensembles to develop saltwater intrusion management strategies in coastal aquifers. *Earth Syst Environ* 2:193–211. <https://doi.org/10.1007/s41748-018-0069-3>
- Roy DK, Datta B (2018c) Optimal management of groundwater extraction to control saltwater intrusion in multi-layered coastal aquifers using ensembles of adaptive neuro-fuzzy inference system. *World Environ Water Resour Congr 2017*:139–150
- Roy DK, Datta B (2018d) Trained meta-models and evolutionary algorithm based multi-objective management of coastal aquifers under parameter uncertainty. *J Hydroinformatics* 20:1247–1267. <https://doi.org/10.2166/hydro.2018.087>
- Roy DK, Datta B (2018e) An ensemble meta-modelling approach using the Dempster-Shafer theory of evidence for developing saltwater intrusion management strategies in coastal aquifers. *Water Resour Manag*. <https://doi.org/10.1007/s11269-018-2142-y>
- Roy DK, Datta B (2020a) Saltwater intrusion prediction in coastal aquifers utilizing a weighted-average heterogeneous ensemble of prediction models based on Dempster-Shafer theory of evidence. *Hydrol Sci J* 65:1555–1567. <https://doi.org/10.1080/02626667.2020.1749764>
- Roy DK, Datta B (2020b) Modelling and management of saltwater intrusion in a coastal aquifer system: a regional-scale study. *Groundw Sustain Dev* 11:100479. <https://doi.org/10.1016/j.gsd.2020.100479>
- Saad S, Javadi AA, Chugh T, Farmani R (2022) Optimal management of mixed hydraulic barriers in coastal aquifers using multi-objective Bayesian optimization. *J Hydrol* 612:128021. <https://doi.org/10.1016/j.jhydrol.2022.128021>
- Song J, Yang Y, Wu J et al (2018) Adaptive surrogate model based multiobjective optimization for coastal aquifer management. *J Hydrol* 561:98–111. <https://doi.org/10.1016/j.jhydrol.2018.03.063>
- Sreekanth J, Datta B (2010) Multi-objective management of saltwater intrusion in coastal aquifers using genetic programming and modular neural network based surrogate models. *J Hydrol* 393:245–256. <https://doi.org/10.1016/j.jhydrol.2010.08.023>
- Sreekanth J, Datta B (2011a) Comparative evaluation of genetic programming and neural network as potential surrogate models for coastal aquifer management. *Water Resour Manag* 25:3201–3218. <https://doi.org/10.1007/s11269-011-9852-8>
- Sreekanth J, Datta B (2011b) Coupled simulation-optimization model for coastal aquifer management using genetic programming-based ensemble surrogate models and multiple-realization optimization. *Water Resour Res*. <https://doi.org/10.1029/2010WR009683>
- Willmott CJ (1981) On the validation of models. *Phys Geogr* 2:184–194. <https://doi.org/10.1080/02723646.1981.10642213>
- Yapo PO, Gupta HV, Sorooshian S (1996) Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. *J Hydrol* 181:23–48. [https://doi.org/10.1016/0022-1694\(95\)02918-4](https://doi.org/10.1016/0022-1694(95)02918-4)
- Zhang F, Deb C, Lee SE et al (2016) Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique. *Energy Build* 126:94–103. <https://doi.org/10.1016/j.enbuild.2016.05.028>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.