# Development of a sustainable groundwater management strategy and sequential compliance monitoring to control saltwater intrusion in coastal aquifers

**Thesis submitted by**

**Dilip Kumar Roy**
**B. Sc. (Agricultural Engineering, BAU, Bangladesh)**
**M. Sc. (Irrigation and Water Management, BAU, Bangladesh)**
**M. Sc. (Environmental Technology and Engineering, Ghent University, Belgium)**

**July 2018**

**for the degree of Doctor of Philosophy (PhD)**
**in the College of Science and Engineering**
**James Cook University**

JAMES COOK UNIVERSITY
AUSTRALIA

## Dedication

This thesis and years spent completing the research works to complete this thesis is dedicated to

*The memory of my father "Kshitish Chandra Roy" (1951 – 2018). His inspiration and encouragement continued until 01 February 2018. His words of inspiration in pursuit of excellence still remain.*

## Acknowledgements

Principally, I am profoundly thankful to almighty God for giving me the patience and strength to complete this research work and writing the thesis.

I would like to express heartfelt respect for and sincere appreciation to my primary supervisor Dr Bithin Datta for his unreserved support, invaluable suggestions, constructive criticism, and fruitful guidance throughout the duration of this research. I am grateful to him for his immense help and continuous encouragement in the timely completion of this thesis.

Thanks are also extended to my co-supervisor Assoc. Prof Dr Siva Sivakugan, and other faculty and staff members of the College of Science and Engineering, James Cook University. In addition, I would like to thank my fellow researchers in engineering, and the friends who made my stay at Townsville most enjoyable.

My sincere thanks go to the Graduate Research School, James Cook University, for providing me with the financial support for the duration of my study. I am also thankful to the College of Science and Engineering, James Cook University, Australia for providing financial assistance to allow me to attend two international conferences in the USA and Japan.

My thanks are also extended to Bangladesh Agricultural Research Institute and Ministry of Agriculture, Bangladesh for providing leave (deputation) to pursue this PhD research in Australia.

Finally, I owe much to all of my family members for their unconditional encouragement and support throughout my absence from them during my academic endeavours.

Dilip Kumar Roy

## Statement of the contribution of others

The concepts and the results reported in this thesis are produced under the supervision of Dr Bithin Datta, who assisted in the conceptualization of the problems and the various approaches to solve the problems. Dr Datta reviewed the manuscripts of the published articles in journals, conference proceedings, and the text of this thesis.

Financial assistance for this PhD project was received from JCUPRS scholarship, which includes fortnightly living expenses and tuition fee waiver. Apart from this scholarship, conference attendance costs were provided by the College of Science and Engineering.

Bangladesh Agricultural Research Institute provided 42 months leave for conducting this PhD study.

## Abstract

The coastal areas of the world are characterized by high population densities, an abundance of food, and increased economic activities. These increasing human settlements, subsequent increases in agricultural developments and economic activities demand an increasing amount quantity of freshwater supplies to different sectors. Groundwater in coastal aquifers is one of the most important sources of freshwater supplies. Over exploitation of this coastal groundwater resource results in seawater intrusion and subsequent deterioration of groundwater quality in coastal aquifers. In addition, climate change induced sea level rise, in combination with the effect of excessive groundwater extraction, can accelerate the seawater intrusion. Adequate supply of good quality water to different sectors in coastal areas can be ensured by adoption of a proper management strategy for groundwater extraction. Optimal use of the coastal groundwater resource is one of the best management options, which can be achieved by employing a properly developed optimal groundwater extraction strategy. Coupled simulation-optimization (S-O) approaches are essential tools to obtain the optimal groundwater extraction patterns. This study proposes approaches for developing multiple objective management of coastal aquifers with the aid of barrier extraction wells as hydraulic control measure of saltwater intrusion in multilayered coastal aquifer systems. Therefore, two conflicting objectives of management policy are considered in this research, i.e. maximizing total groundwater extraction for advantageous purposes, and minimizing the total amount of water abstraction from barrier extraction wells. The study also proposes an adaptive management strategy for coastal aquifers by developing a three-dimensional (3-D) monitoring network design. The performance of the proposed methodologies is evaluated by using both an illustrative multilayered coastal aquifer system and a real life coastal aquifer study area.

Coupled S-O approach is used as the basic tool to develop a saltwater intrusion management model to obtain the optimal groundwater extraction rates from a combination of feasible solutions on the Pareto optimal front. Simulation of saltwater intrusion processes requires solution of density dependent coupled flow and solute transport numerical simulation models that are computationally intensive. Therefore, computational efficiency in the coupled S-O approach is achieved by using an approximate emulator of the accompanying physical processes of coastal aquifers. These emulators, often known as surrogate models or meta-models, can replace the computationally intensive numerical simulation model in a coupled S-O approach for achieving computational efficiency. A number of meta-models have been developed and compared in this study for integration with the optimization algorithm in order to

develop saltwater intrusion management model. Fuzzy Inference System (FIS), Adaptive Neuro Fuzzy Inference System (ANFIS), Multivariate Adaptive Regression Spline (MARS), and Gaussian Process Regression (GPR) based meta-models are developed in the present study for approximating coastal aquifer responses to groundwater extraction. Properly trained and tested meta-models are integrated with a Controlled Elitist Multiple Objective Genetic Algorithm (CEMOGA) within a coupled S-O approach. In each iteration of the optimization algorithm, the meta-models are used to compute the corresponding salinity concentrations for a set of candidate pumping patterns generated by the optimization algorithm. Upon convergence, the non-dominated global optimal solutions are obtained as the Pareto optimal front, which represents a trade-off between the two conflicting objectives of the pumping management problem. It is observed from the solutions of the meta-model based coupled S-O approach that the considered meta-models are capable of producing a Pareto optimal set of solutions quite accurately. However, each meta-modelling approach has distinct advantages over the others when utilized within the integrated S-O approach.

Uncertainties in estimating complex flow and solute transport processes in coastal aquifers demand incorporation of the uncertainties related to some of the model parameters. Multidimensional heterogeneity of aquifer properties such as hydraulic conductivity, compressibility, and bulk density are considered as major sources of uncertainty in groundwater modelling system. Other sources of uncertainty are associated with spatial and temporal variability of hydrologic as well as human interventions, e.g. aquifer recharge and transient groundwater extraction patterns. Different realizations of these uncertain model parameters are obtained from different statistical distributions. FIS based meta-models are advanced to a Genetic Algorithm (GA) tuned hybrid FIS model (GA-FIS), to emulate physical processes of coastal aquifers and to evaluate responses of the coastal aquifers to groundwater extraction under groundwater parameter uncertainty. GA is used to tune the FIS parameters in order to obtain the optimal FIS structure. The GA-FIS models thus obtained are linked externally to the CEMOGA in order to derive an optimal pumping management strategy using the coupled S-O approach. The evaluation results show that the proposed saltwater intrusion management model is able to derive reliable optimal groundwater extraction strategies to control saltwater intrusion for the illustrative multilayered coastal aquifer system. The optimal management strategies obtained as solutions of GA-FIS based management models are shown to be reliable and accurate within the specified ranges of values for different realizations of uncertain groundwater parameters.

One of the major concerns of the meta-model based integrated S-O approach is the uncertainty associated with the meta-model predictions. These prediction

uncertainties, if not addressed properly, may propagate to the optimization procedures, and may deteriorate the optimality of the solutions. A standalone meta-model, when used within an optimal management model, may result in the optimization routine producing actually suboptimal solutions that may undermine the optimality of the groundwater extraction strategies. Therefore, this study proposes an ensemble approach to address the prediction uncertainties of meta-models. Ensemble is an approach to assimilate multiple similar or different algorithms or base learners (emulators). The basic idea of ensemble lies in developing a more reliable and robust prediction tool that incorporates each individual emulator's unique characteristic in order to predict future scenarios. Each individual member of the ensemble contains different input-output mapping functions. Based on their own mapping functions, these individual emulators provide varied predictions on the response variable. Therefore, the combined prediction of the ensemble is likely to be less biased and more robust, reliable, and accurate than that of any of the individual members of the ensemble. Performance of the ensemble meta-models is evaluated using an illustrative coastal aquifer study area. The results indicate that the meta-model based ensemble modelling approach is able to provide reliable solutions for a multilayered coastal aquifer management problem.

Relative sea level rise, providing an additional saline water head at the seaside, has a significant impact on an increase in the salinization process of the coastal aquifers. Although excessive groundwater withdrawal is considered as the major cause of saltwater intrusion, relative sea level rise, in combination with the effect of excessive groundwater pumping, can exacerbate the already vulnerable coastal aquifers. This study incorporates the effects of relative sea level rise on the optimized groundwater extraction values for the specified management period. Variation of water concentrations in the tidal river and seasonal fluctuation of river water stage are also incorporated. Three meta-models are developed from the solution results of the numerical simulation model that simulates the coupled flow and solute transport processes in a coastal aquifer system. The results reveal that the proposed meta-models are capable of predicting density dependent coupled flow and solute transport patterns quite accurately. Based on the comparison results, the best meta-model is selected as a computationally cheap substitute of the simulation model in the coupled S-O based saltwater intrusion management model. The performance of the proposed methodology is evaluated for an illustrative multilayered coastal aquifer system in which the effect of climate change induced sea level rise is incorporated for the specified management period. The results show that the proposed saltwater intrusion management model provides acceptable, accurate, and reliable solutions while significantly improving computational efficiency in the coupled S-O methodology.

The success of the developed management strategy largely depends on how accurately the prescribed management policy is implemented in real life situations. The actual implementation of a prescribed management strategy often differs from the prescribed planned strategy due to various uncertainties in predicting the consequences, as well as practical constraints, including noncompliance with the prescribed strategy. This results in actual consequences of a management strategy differing from the intended results. To bring the management consequences closer to the intended results, adaptive management strategies can be sequentially modified at different stages of the management horizon using feedback measurements from a deigned monitoring network. This feedback information can be the actual spatial and temporal concentrations resulting from the implementation of actual management strategy. Therefore, field-scale compliance of the developed coastal aquifer management strategy is a crucial aspect of an optimally designed groundwater extraction policy. A 3-D compliance monitoring network design methodology is proposed in this study in order to develop an adaptive and sequentially modified management policy, which aims to improve optimal and justifiable use of groundwater resources in coastal aquifers. In the first step, an ensemble meta-model based multiple objective prescriptive model is developed using a coupled S-O approach in order to derive a set of Pareto optimal groundwater extraction strategies. Prediction uncertainty of meta-models is addressed by utilizing a weighted average ensemble using Set Pair Analysis. In the second step, a monitoring network is designed for evaluating the compliance of the implemented strategies with the prescribed management goals due to possible uncertainties associated with field-scale application of the proposed management policy. Optimal monitoring locations are obtained by maximizing Shannon's entropy between the saltwater concentrations at the selected potential locations. Performance of the proposed 3-D sequential compliance monitoring network design is assessed for an illustrative multilayered coastal aquifer study area. The performance evaluations show that sequential improvements of optimal management strategy are possible by utilizing saltwater concentrations measurements at the proposed optimal compliance monitoring locations.

The integrated S-O approach is used to develop a saltwater intrusion management model for a real world coastal aquifer system in the Barguna district of southern Bangladesh. The aquifer processes are simulated by using a 3-D finite element based combined flow and solute transport numerical code. The modelling and management of seawater intrusion processes are performed based on very limited hydrogeological data. The model is calibrated with respect to hydraulic heads for a period of five years from April 2010 to April 2014. The calibrated model is validated for the next three-year period from April 2015 to April 2017. The calibrated and partially

validated model is then used within the integrated S-O approach to develop optimal groundwater abstraction patterns to control saltwater intrusion in the study area. Computational efficiency of the management model is achieved by using a MARS based meta-model approximately emulating the combined flow and solute transport processes of the study area. This limited evaluation demonstrates that a planned transient groundwater abstraction strategy, acquired as solution results of a meta-model based integrated S-O approach, is a useful management strategy for optimized water abstraction and saltwater intrusion control. This study shows the capability of the MARS meta-model based integrated S-O approach to solve real-life complex management problems in an efficient manner.

## List of publications published during PhD candidature

## Peer reviewed journal articles

1. Roy, D. K., & Datta, B. (2017a). Fuzzy c-mean clustering based inference system for saltwater intrusion processes prediction in coastal aquifers. *Water Resources Management, 31*(1), 355-376. doi:10.1007/s11269-016-1531-3

2. Roy, D. K., & Datta, B. (2017b). Multivariate adaptive regression spline ensembles for management of multilayered coastal aquifers. *Journal of Hydrologic Engineering, 22*(9), 04017031.

3. Roy, D. K., & Datta, B. (2017c). Genetic algorithm tuned fuzzy inference system to evolve optimal groundwater extraction strategies to control saltwater intrusion in multi-layered coastal aquifers under parameter uncertainty. *Modeling Earth Systems and Environment, 3*(4), 1707-1725. doi:10.1007/s40808-017-0398-5

4. Roy, D. K., & Datta, B. (2017d). Saltwater intrusion processes in coastal aquifers – modelling and management: a review. *Desalination and Water Treatment, 78*, 57-89.

5. Roy, D. K., & Datta, B. (2018a). A surrogate based multi-objective management model to control saltwater intrusion in multi-layered coastal aquifer systems. *Civil Engineering and Environmental Systems*, 1-26. doi:10.1080/10286608.2018.1431777

6. Roy, D. K., & Datta, B. (2018b). Comparative efficiency of different artificial intelligence based models for predicting density dependent saltwater intrusion processes in coastal aquifers and saltwater intrusion management utilizing the best performing model. *Desalination and Water Treatment, 105*, 160–180.

7. Roy, D. K., & Datta, B. (2018c). Influence of sea level rise on multi-objective management of saltwater intrusion in coastal aquifers. *Journal of Hydrologic Engineering, 23* (8), 04018035.

8. Roy, D. K., & Datta, B. (2018). Trained meta-models and evolutionary algorithm based multi-objective management of coastal aquifers under parameter uncertainty. *Journal of Hydroinformatics,* (Accepted).

9. Roy, D. K., & Datta, B. (2018). A review of surrogate models and their ensembles to develop saltwater intrusion management strategies in coastal aquifers. *Earth Systems and Environment,* (Accepted).

10. Roy, D. K., & Datta, B. Adaptive management of coastal aquifers using entropy-set pair analysis based three-dimensional sequential monitoring network design. *Journal of Hydrologic Engineering,* (Submitted).

11. Roy, D. K., & Datta, B. An ensemble meta-modelling approach using Dempster-Shafer theory of evidence for developing saltwater intrusion management strategies in coastal aquifers. *Water Resources Management,* (Submitted).

## Conference proceedings

1. Roy, D. K., & Datta, B. (2017a). Optimal management of groundwater extraction to control saltwater intrusion in multi-layered coastal aquifers using ensembles of adaptive neuro-fuzzy inference system. *World Environmental and Water Resources Congress 2017, American society of civil engineers, May 21–25, 2017, Sacramento, California, USA*, 139-150.

2. Roy, D. K., & Datta, B. (2017b). Prediction of saltwater intrusion processes using a WAVELET-FCM-ANFIS hybrid model. *Seventh International Conference on Geotechnique, Construction Materials and Environment, Mie, Japan*.

# Table of Contents

# List of Tables

# List of Figures

## Notations used in this thesis

### List of symbols

| Symbols | Meaning |
|---|---|
| $\overline{C_o}$ | Mean values of the observed salinity concentrations |
| $\overline{C_P}$ | Mean values of the predicted salinity concentration values |
| $BF_i$ | $i^{th}$ Basis functions |
| $B_i^j$ | Abstraction from the $i^{th}$ barrier extraction wells at $j^{th}$ time steps |
| $C_{i,o}$ | Observed salinity concentration values in the training dataset |
| $C_{i,p}$ | Predicted salinity concentrations in the training dataset |
| $C_i$ | Saltwater concentrations at $i^{th}$ monitoring locations at the closure of the management period |
| $C_i^j$ | Salinity concentrations at the $i^{th}$ monitoring locations at $j^{th}$ time steps |
| $C_n^O$ | Observed values of seawater concentrations |
| $C_n^P$ | Predicted values of seawater concentrations |
| $H_{ij}$ | Salinity concentration based entropy at $j^{th}$ potential location for $i^{th}$ layer |
| $K_d$ | Distribution coefficient |
| $K_s$ | First order biodegradation rate through adsorbed phase |
| $K_w$ | First order biodegradation rate constant through dissolved phase |
| $K_x$ | Horizontal hydraulic conductivity values in the X-direction |
| $K_y$ | Horizontal hydraulic conductivity values in the Y-direction |
| $K_z$ | Vertical hydraulic conductivity in the Z-direction |
| $N_{MF(input)}$ | Number of input Membership Functions |
| $N_{MF(output)}$ | Number of output Membership Functions |
| $N_X$ | Number of predictors |
| $N_Y$ | Number of responses |

| | |
|---|---|
| $\boldsymbol{N_a}$ | Number of data points whose absolute percentage relative error value is less than a specified threshold value $a\%$. |
| $\boldsymbol{N_{clusters}}$ | Number of clusters |
| $\boldsymbol{OutMARS_i}$ | Output from the $i^{th}$ MARS meta-models |
| $\boldsymbol{Output_{m_i}}$ | Output from the $i^{th}$ MARS meta-models |
| $\boldsymbol{Output_{WAE}}$ | Weighted average prediction from all meta-models |
| $\boldsymbol{P_i^j}$ | Abstraction from the $i^{th}$ pumping wells at $j^{th}$ time steps |
| $\boldsymbol{Q^{BW}(selected)}$ | Total amount of water extraction from the barrier wells for the selected pumping Management strategy. |
| $\boldsymbol{\|V\|}$ | Magnitude of $V$ |
| $\boldsymbol{Weight_i}$ | Weight assigned to $i^{th}$ meta-model |
| $\boldsymbol{W_m}$ | Set pair weight of the $m^{th}$ meta-model |
| $\boldsymbol{a_L}$ | Longitudinal dispersivity |
| $\boldsymbol{a_T}$ | Lateral dispersivity |
| $\boldsymbol{a_{ij}}$ | Binary decision variable taking the value of either 1 or 0 |
| $\boldsymbol{a_m}$ | Molecular diffusion coefficient |
| $\boldsymbol{c_j}$ | Centre of the $j^{th}$ cluster |
| $\boldsymbol{f_{MSE}}$ | Cost function to be minimized |
| $\boldsymbol{k_r}$ | Relative permeability or relative hydraulic conductivity |
| $\boldsymbol{k_s}$ | Saturated permeability tensor |
| $\boldsymbol{k_{so}}$ | Referenced saturated conductivity tensor |
| $\boldsymbol{{}_k^tQ_{BW}}$ | Water extraction from $k^{th}$ barrier extraction well throughout $t^{th}$ time phase |
| $\boldsymbol{l_i}$ | Maximum number of monitoring locations to be selected in layer $i$. |
| $\boldsymbol{n_i}$ | Total number of potential monitoring locations in hydrogeologic layer $i$ |
| $\boldsymbol{n_j}$ | Total number of hydrogeologic layers considered |
| $\boldsymbol{{}_r^tQ_{PW}}$ | Water extraction from the $r^{th}$ pumping well throughout $t^{th}$ time phase |
| $\boldsymbol{x_i}$ | $i^{th}$ data points |

| | |
|---|---|
| $\boldsymbol{\alpha'}$ | Modified compressibility of water |
| $\boldsymbol{\beta'}$ | Modified compressibility of the medium |
| $\boldsymbol{\gamma_k}$ | Coefficients of $BF_i(X)$ |
| $\boldsymbol{\mu_0}$ | Reference dynamic viscosity at zero chemical concentration |
| $\boldsymbol{\mu_{ij}}$ | Degree of membership of $x_i$ in the $j^{th}$ cluster |
| $\boldsymbol{\rho^*}$ | Density of injection fluid or that of the withdrawn water |
| $\boldsymbol{\rho_0}$ | Referenced water density at zero chemical concentration |
| $\boldsymbol{\rho_b}$ | Bulk density of the medium |
| $\boldsymbol{\nabla}$ | del operator |
| $\boldsymbol{h}$ | Pressure head |
| $\boldsymbol{C}$ | Chemical concentration |
| $\boldsymbol{D}$ | Dispersion coefficient tensor |
| $\boldsymbol{F}$ | Storage coefficient |
| $\boldsymbol{I}$ | Numbers of identical characteristics |
| $\boldsymbol{K}$ | Hydraulic conductivity tensor |
| $\boldsymbol{N}$ | Number of data points |
| $\boldsymbol{OutEnMARS}$ | Combined output from the ensemble of MARS meta-models |
| $\boldsymbol{S}$ | Saturation |
| $\boldsymbol{V}$ | Discharge |
| $\boldsymbol{b}$ | Maximum permissible total number of monitoring wells summed over all layers |
| $\boldsymbol{m}$ | Fuzzy partition matrix exponent |
| $\boldsymbol{n}$ | Porosity |
| $\boldsymbol{q}$ | A source or a sink |
| $\boldsymbol{t}$ | Time |
| $\boldsymbol{z}$ | Potential head |
| $\boldsymbol{\delta}$ | Kronecker delta tensor |
| $\boldsymbol{\theta}$ | Moisture content |
| $\boldsymbol{\lambda}$ | Decay constant |
| $\boldsymbol{\mu}$ | Dynamic viscosity of water at chemical concentration $C$ |

| | |
|---|---|
| $\boldsymbol{\rho}$ | Water density at chemical concentration $C$ |
| $\boldsymbol{\tau}$ | Tortuosity |

## List of abbreviations

| Abbreviations | Meaning |
|---|---|
| **2-D** | Two-dimensional |
| **3-D** | Three-dimensional |
| **ANFIS** | Adaptive Neuro Fuzzy Inference System |
| **ANN** | Artificial Neural Network |
| **BBS** | Bangladesh Bureau of Statistics |
| **BWs** | Barrier extraction wells |
| **CEMOGA** | Controlled Elitist Multiple Objective Genetic Algorithm |
| **CV** | Coefficient of Variation |
| **ED** | Euclidian distance |
| **En-MARS** | Ensemble of MARS |
| **EPR** | Evolutionary Polynomial Regression |
| **FCM** | Fuzzy C-Mean Clustering |
| **FIS** | Fuzzy Inference System |
| **GA** | Genetic Algorithm |
| **GA-FIS** | Genetic Algorithm tuned hybrid FIS |
| **GD** | Gradient Descent |
| **GP** | Genetic Programming |
| **GPR** | Gaussian Process Regression |
| **GRA** | Gray Relational Analysis |
| **GRC** | Gray Relational Coefficient |
| **HA** | Halton sequences |
| **IOA** | Index of Agreement |
| **KGE** | Kling–Gupta Efficiency |
| **LHS** | Latin Hypercube Sampling |
| **LSE** | Least Squares Estimate |
| **MAD** | Median Absolute Deviation |
| **MAPRD** | Mean Absolute Percentage Relative Difference |

| | |
|---|---|
| **MAPRE** | Mean Absolute Percentage Relative Error |
| **MARS** | Multivariate Adaptive Regression Spline |
| **MF** | Membership Function |
| **MLs** | Monitoring locations |
| **MSE** | Mean Squared Error |
| **NS** | Nash–Sutcliffe Efficiency Coefficient |
| **PARE** | Percentage Absolute Relative Errors |
| **PSO** | Particle Swarm Optimization PSO |
| **PWs** | Production wells |
| **R** | Correlation Coefficient |
| **RBF** | Radial Basis Function |
| **RE** | Relative Error |
| **RMSE** | Root Mean Squared Error |
| **RRMSE** | Relative Root Mean Squared Error |
| **RWs** | Freshwater recharge wells |
| **SA** | Simulated Annealing |
| **SD** | Standard Deviation |
| **S-O** | Simulation-optimization |
| **SPA** | Set Pair Analysis |
| **TS** | Threshold Statistics |
| **WAE-MARS** | Weighted Average Ensemble of Multivariate Adaptive Regression Spline |
| **X1-X80** | Decision variables for the management model |

# Chapter 1: Introduction

This chapter briefly describes saltwater intrusion management issues in coastal aquifers. The first part of this chapter outlines approaches to controlling saltwater intrusion through adopting optimal groundwater extraction strategies. The second part deals with this study's research objectives, and the third part summarizes the organization of the thesis.

## 1.1 Outline

The coastal areas of the world are characterized by high population densities, with about 50% of the world's population living within 60 km of the shoreline (Oude Essink, 2001b). Coastal areas attract human settlement by providing abundant food (fisheries and agriculture) and economic activities (trade, harbors, ports and infrastructure). Therefore, a shortage of fresh groundwater supply in the domestic, agricultural, and industrial sectors in these areas is inevitable due to the pressure of increased human settlements, agricultural developments, and economic activities. Seawater intrusion and the subsequent deterioration of groundwater quality in coastal aquifers usually results from over extraction of groundwater resources to meet irrigation, domestic and industrial needs (Nocchi and Salleolini, 2013). In addition, tidal fluctuations affect the dynamics of groundwater flow patterns in coastal aquifers, such as fluctuations in the groundwater table, and in groundwater discharge to the sea (Liu et al., 2012). In an unconfined coastal aquifer system, tidal oscillations significantly reduce the overall extent of the seawater intrusion due to an increase in the total discharge of water towards the sea from the aquifer (Kuan et al., 2012). Chen and Hsu (2004) demonstrated that the saltwater intrusion profile varies only in vertical direction in an unconfined aquifer due to the influence of tidal fluctuations. Narayan et al. (2007) reported that the effect of tidal fluctuations on groundwater levels is limited to the areas close to the coast, and therefore, tidal influence on saltwater intrusion can be neglected when compared with the effects due to groundwater pumping. In addition, sea level rise can accelerate saltwater intrusion processes in aquifer systems, and several centuries would be required to equilibrate this sea level rise induced saltwater intrusion, even if the sea level returns to its original position (Webb and Howard, 2011). However, the effect of this sea level rise induced increase in hydraulic heads of groundwater system is generally confined to within few kilometres of the coastline and main rivers (Oude Essink et al., 2010). Although excessive groundwater withdrawal is considered as the major cause of saltwater intrusion (Narayan et al., 2007), relative sea level rise in combination with the effect of excessive groundwater pumping can exacerbate the already vulnerable coastal aquifers (Langevin and Zygnerski, 2013).

The beneficial use of coastal groundwater resources requires the development and implementation of sustainable management strategies. This thesis details methodologies for

efficient and computationally feasible regional scale management strategies for the sustainable use of coastal aquifers. The methodologies include recommendations for meta-model based coupled simulation-optimization approaches to develop a saltwater intrusion management model; an ensemble meta-modelling approach in the management model to address the prediction uncertainty of meta-models; a coastal aquifer management model under groundwater parameter uncertainty; a management model incorporating climate change induced sea level rise; and uncertainty based 3-D monitoring network design. The performance of the developed methodologies is evaluated by using a multilayered coastal aquifer system. A meta-model based coupled simulation-optimization approach is also applied to develop a regional scale saltwater intrusion management model for a real life coastal aquifer system.

Over extraction of groundwater resources is one of the main reasons for salinity intrusion in coastal aquifers. Hence, the judicial use of coastal groundwater resources is the best management option, which can be achieved by employing properly developed optimal groundwater extraction strategies. Sustainable beneficial water extraction from the coastal aquifers can be maximized by creating a reverse gradient from the seaside along the coast to control saltwater intrusion hydraulically. Therefore, the present study has incorporated the possible use of barrier wells along the coastal boundary as one of the options in the management strategy. Coupled simulation-optimization (S-O) approaches are essential tools to obtain the optimal groundwater extraction rates from a combination of feasible solutions in the Pareto optimal front. In a coupled S-O approach, the simulation part is linked to an optimization algorithm in order to derive global optimal solutions. The simulation of saltwater intrusion processes requires solution of density dependent coupled flow and salt transport numerical simulation models that are computationally intensive. Therefore, the linking of these numerical simulation models to a linked S-O methodology for the management of coastal aquifers requires enormous computational time. For instance, a linked S-O methodology for a small 3-D coastal aquifer may require as large as a 30-day computer run time (Dhar and Datta, 2009a). One way of achieving computational efficiency in such a linked S-O approach is the use of reasonably accurate approximate meta-models, which are very useful tools to approximately simulate the coupled flow and salt transport processes in coastal aquifers (Banerjee et al., 2011; Bhattacharjya et al., 2007; Kourakos and Mantoglou, 2009; Kourakos and Mantoglou, 2013). These approximate meta-models can replace the computationally intensive numerical simulation model to achieve computational efficiency in a linked S-O approach. Repeated use of these meta-models by the optimization algorithm does not require significant computational time because these meta-models are simplified representations of the complex simulation models.

The Artificial Neural Network (ANN) (Banerjee et al., 2011; Bhattacharjya and Datta, 2009), Genetic Programming (GP) (Sreekanth and Datta, 2010), Cubic Radial Basis Function (RBF) (Christelis and Mantoglou, 2016), and Evolutionary Polynomial Regression (EPR) models

(Hussain et al., 2015) are some of the commonly used meta-models in substitution for complex numerical simulation models for saltwater intrusion management problems in coastal aquifers. In all cases, the objective is to find appropriate meta-models that are rather simple but quite reliable and accurate in their prediction capability. Such meta-models can be good choices for a linked S-O methodology to solve large-scale coastal aquifer management problems. However, the commonly used meta-models in saltwater intrusion management problem have certain limitations. The drawbacks of ANN models include a tendency to premature convergence in local minima, the "Black-Box" nature of the models, higher computational burden, and susceptibility to model overfitting etc. (Holman et al., 2014). In addition, ANN models are not stable when the number of training datasets is insufficient (Hsieh and Tang, 1998). GP, an explicit mathematical formulation (Shiri and Kişi, 2011), produces simple regression models (Sreekanth and Datta, 2011a) that can easily be linked within an optimization algorithm to achieve computational efficiency in linked S-O methodology. However, GP requires extensive training time in evaluating many model structures before the optimal structure can be identified (Sreekanth and Datta, 2011a). In addition, GP suffers from being trapped in local minima (Pillay, 2004). RBF is quite simple in formulation (Sóbester et al., 2014), and easy to implement in any number of dimensions, with reasonable accuracy for certain types of radial functions (Piret, 2007). However, the RBF method has stability issues that can be a serious concern. Computational cost is another significant drawback of using the RBF meta-model. Polynomial Regression has stability problems when the polynomial order is high for polynomial fits. In addition, individual observations of the training datasets can have an unexpected influence on remote parts of the curve in Polynomial Regression (Green and Silverman, 1993).

To overcome some of the limitations of these commonly used meta-models, a number of meta-models have been proposed in this study for linking with the optimization algorithm to develop a saltwater intrusion management model. Fuzzy Inference System (FIS), Adaptive Neuro Fuzzy Inference System (ANFIS), Multivariate Adaptive Regression Spline (MARS), and Gaussian Process Regression (GPR) based meta-models are developed for approximating coastal aquifer responses to extraction of groundwater resources, and linked externally with a Controlled Elitist Multiple Objective Genetic Algorithm (CEMOGA) based optimization algorithm to obtain optimal groundwater extraction strategies for saltwater intrusion management. Although meta-models provide considerable computational efficiency in a coupled S-O based saltwater management model, further computational efficiencies can be achieved by implementing parallel computing facilities. Parallel computation is performed by distributing the objective functions and constraints to all physical processors of the PC rather than running them serially with parallel gradient estimation. Therefore, the parameters are automatically passed to worker machines during the execution of parallel computations. This work aims at using a parallel pool of worker

machines by utilizing the physical cores of a standard computer [Intel® Core ™ i7-2600 CPU @3.40 GHz, 8GB RAM] by utilizing parallel computing toolbox of MATLAB.

The use of meta-models in approximating density dependent coupled flow and solute transport processes in coastal aquafers imposes a certain degree of uncertainty in predictions. This prediction uncertainty can be minimized by using an ensemble of meta-models, which protects against single but wrong meta-models (Goel et al., 2007). A methodology based on utilizing ensemble meta-models in a simple and weighted averaging approach is employed to obtain optimal groundwater extraction strategies for controlling saltwater intrusion in coastal aquifers. A new approach is also proposed in this study, of incorporating groundwater parameter uncertainties in meta-models by training different meta-models from different realizations of uncertain model parameters and combining them in an ensemble framework.

The reliability and accuracy of the optimal management strategy based on a coupled S-O approach depends on how accurately the simulation model captures and simulates the physical processes. Uncertainties associated with some of the groundwater parameters results in inaccurate characterizations of these physical processes in the coastal groundwater system. Therefore, a coastal aquifer management model should incorporate the uncertainty associated with groundwater parameters. The multidimensional heterogeneity of aquifer properties such as hydraulic conductivity, compressibility, and bulk density are considered to be major sources of uncertainty in groundwater modelling system (Ababou and Al-Bitar, 2004). Other sources of uncertainty are associated with the spatial and temporal variability of hydrologic and human interventions, e.g. aquifer recharge and transient groundwater extraction patterns. In developing saltwater intrusion management models, the meta-models and ensemble meta-model based methodologies are further extended to account for uncertainties related to groundwater parameters.

The success of any coastal aquifer management strategy largely depends on how accurately the prescribed management policy is implemented in field situations. Total compliance with prescribed strategies is rarely satisfied in the field, which triggers the need for developing adaptive management strategies that can be sequentially modified at different stages of the management horizon, based on feedback information from the actually implemented strategies. This feedback information can be obtained from an optimally designed compliance-monitoring network. Saltwater intrusion in coastal aquifers is three-dimensional. Therefore, multilevel monitoring is required for the efficient management of coastal aquifers (Dhar and Datta, 2009b). To address this issue, this study aims to propose a 3-D sequential compliance monitoring network design, in order to develop an adaptive management strategy in a multilayered coastal aquifer system.

## 1.2 Research objectives

The research aim of this study is the formulation of sustainable groundwater management strategies to control saltwater intrusion in coastal groundwater aquifers. The tools used for modelling include a finite element based density dependent coupled flow and solute transport simulation model, meta-model based coupled S-O approaches, a nonlinear optimization algorithm, and a parallel processing framework to speed up the optimization procedure. Different uncertainties associated with physical parameter estimations, anthropogenic activities, and uncertainties related to meta-model predictions are addressed in order to develop optimal groundwater extraction strategies. The adaptive management of coastal aquifers is also proposed by using a 3-D network of compliance monitoring wells. Meta-model based coupled S-O approaches have been extended for application to a real-world coastal aquifer system. The specific research objectives are to:

1. Develop multiple objective saltwater intrusion management models for multilayered coastal aquifer systems by using coupled S-O approach. In order to achieve this objective, a number of computationally efficient meta-models are proposed and the relative advantages of these meta-models over previously used meta-models in saltwater intrusion management problems are illustrated.

2. Incorporate prediction uncertainty of meta-models and groundwater parameter uncertainties in developing the optimal groundwater extraction strategies. This is achieved by proposing both simple and weighted average ensembles of meta-models.

3. Incorporate the influences of sea level rise, seasonal variation of river water level, and river water concentration in the saltwater intrusion management model.

4. Introduce a parallel processing framework to speed up meta-model assisted coupled S-O approaches in multiple objective saltwater intrusion management models.

5. Develop an adaptive saltwater intrusion management model by using uncertainty based 3-D sequential compliance monitoring network design.

6. Extend the proposed meta-model based coupled S-O approach for application in a real life coastal aquifer system in Bangladesh.

## 1.3 Organization of the thesis

This thesis consists of eleven chapters, together with this introductory chapter, which presents different aspects of saltwater intrusion management problem in coastal aquifers and ways of addressing this problem. The overall research aim and explicit research objectives of the thesis are also presented in this introductory chapter.

Chapter 2 details the recent literature on saltwater intrusion problems in coastal aquifers and describes the state-of-the-art of various methodologies on the management and monitoring of coastal groundwater resources.

Chapter 3 illustrates a comparative evaluation of barrier extraction and recharge wells as hydraulic control measures to control saltwater intrusion in coastal aquifers.

Chapter 4 presents the development and implementation of artificial intelligence based meta-models for approximating coupled flow and solute transport processes in coastal aquifers.

Chapter 5 presents coastal aquifer management strategies in terms of the optimal use of groundwater resources to control saltwater intrusion. Several computationally efficient meta-models are proposed to replace numerical simulation model in the coupled S-O approach. The advantages of these proposed meta-models over the previously used meta-models in saltwater intrusion management problems are also presented.

Chapter 6 deals with coupled S-O approaches to develop an uncertainty based saltwater intrusion management model that incorporates uncertainties in groundwater parameters.

Chapter 7 presents an ensemble of meta-models to address the prediction uncertainty of meta-models when linked to the nonlinear optimization algorithm in developing coastal aquifer management models.

Chapter 8 describes the influence of sea level rise on multiple objective management of coastal aquifers.

Chapter 9 deals with developing a 3-D sequential compliance monitoring network design.

Chapter 10 deals with applying the meta-model based coupled S-O approach in a real world coastal aquifer system in the Barguna district of the southern Bangladesh.

Chapter 11 presents a summary of the proposed approaches, conclusions, and direction for future research.

# Chapter 2: Literature review

Partial contents of this chapter have been published and copyrighted, as outlined below:

Roy, D. K., & Datta, B. (2017d). Saltwater intrusion processes in coastal aquifers – modelling and management: a review. *Desalination and Water Treatment, 78*, 57-89.

## 2.1 Summary

This chapter highlights some of the factors influencing saltwater intrusion processes in coastal aquifers and discusses methodologies for the development of management strategies by utilizing solutions of mathematical models. In recent years, significant advancements have been made in the development of mathematical tools for driving optimal management strategies aimed at controlling saltwater intrusion in coastal aquifers. This chapter also reviews the different modelling techniques utilized, and offers a brief account of the management strategies prescribed by several researchers. Two types of modelling approaches exist in the literature: descriptive models aiming at simulating the physical processes analytically or numerically for evaluating the impact of a chosen management strategy, and prescriptive models for choosing optimal aquifer management strategies. Prescriptive models are developed through a linked simulation-optimization (S-O) approach, in which the simulation model is linked to an optimization algorithm to evolve Pareto optimal groundwater extraction strategies. This linking is computationally prohibitive because of the multiple calls of the simulation models by the optimization algorithm within the optimization framework. Therefore, the recent trend is to develop artificial intelligence based meta-models trained by using random inputs and corresponding numerical outputs in order to develop coastal aquifer management strategy. Different meta- modelling approaches aimed at reducing the computational burden in a linked S-O methodology are also highlighted in this chapter. In addition, this chapter provides a brief outline of the development of methodologies for optimal monitoring network design and sequential compliance monitoring of the real life consequences of implementing prescribed management strategies for coastal aquifers.

## 2.2 Causes of saltwater intrusion in coastal aquifers

The previous literature pointed out several causes of saltwater intrusion in coastal aquifers and highlighted some of the methodologies useful in the development of management strategies. The most common causes are saltwater intrusion due to over extraction for

agricultural or other water supply, saltwater intrusion due to tidal effects or inundations caused by storm surges, and saltwater intrusion due to climate change induced sea level rise. The following sub-sections provide an overview of some of the efforts in modelling the saltwater intrusion scenario caused by these drivers in coastal aquifers.

### 2.2.1 Saltwater intrusion due to over extraction of groundwater

Due to their close proximity to saltwater, coastal groundwater supplies are particularly vulnerable to chloride contamination. The situation becomes worse when groundwater resources in coastal aquifers are overexploited by intensive pumping, leading to the inland movement of the freshwater-saltwater interface, resulting in a high salt concentration in the aquifer water. Groundwater pumping has enormous influence on the lateral or horizontal as well as vertical saltwater intrusion processes in coastal aquifers. Qi and Qiu (2011) considered saltwater intrusion as a natural hazard and indicated that irresponsible human activity such as overexploitation of groundwater resources in coastal aquifers would aggravate the extent of this natural hazard. While many studies pointed out that groundwater abstraction is more significant cause of saltwater intrusion than the rising sea levels (Loáiciga et al., 2012; Sherif and Singh, 1999; Yechieli et al., 2010), others (Carneiro et al., 2010; Sanford and Pope, 2010) argued that groundwater extraction alone might not be the major cause of seawater intrusion.

Datta et al. (2009) utilized a finite element based flow and salt transport numerical simulation model under limited data availability to predict future seawater intrusion scenario due to uncontrolled use of groundwater in a real life coastal aquifer system in Andhra Pradesh, India. The effectiveness of the pumping strategies was evaluated using a calibrated and partially validated model. The modelling results demonstrated that the present rate of pumping would have a detrimental effect on the aquifer, and that a carefully planned pumping strategy was needed to control the spatial and temporal saltwater intrusion processes. Hugman et al. (2015) attempted to evaluate the impact of pumping in coastal aquifers on saltwater intrusion utilizing numerical modelling approaches to simulate different past and present scenarios of groundwater pumping. Four different groundwater abstraction scenarios, including natural state/no abstraction, abstraction for public supply, abstraction for public and private supply, and abstraction for private supply, were evaluated. Seawater intrusion as a result of overexploitation of groundwater and insufficient recharge was also reported as the predominant factor controlling groundwater salinization in the coastal region of the South China Sea near the Leizhou Peninsula (Zhang et al., 2015).

Sherif et al. (2014) evaluated three different pumping scenarios: maintaining the current pumping rates, reducing the current pumping rates, and increasing the current

pumping rate, to quantify saltwater intrusion and salinity distribution in the coastal aquifer of Wadi Ham, United Arab Emirates. Their results indicated that the seawater intrusion process is accelerated by any increase in groundwater pumping, and an increase in the pumping rate by 50% would cause a significant landward movement of the transition zone. On the other hand, reducing the pumping rate by 50% would have a positive result, and would result in the mitigation of seawater intrusion and the improvement of groundwater quality after 10 years of reduced pumping. Reduced pumping from the wells and placing the wells further away from the coast with a better distribution of water withdrawal was proposed by Nocchi and Salleolini (2013).

Pham and Lee (2015) assessed the impact of sea level rise and groundwater extraction on seawater intrusion in the coastal aquifers for a period of 90 years, assuming no change of hydraulic parameters and future hydrologic conditions over the study period, and ignoring the tidal effect and change of shoreline. They considered three different scenarios of sea level rise and groundwater extraction. They concluded that sea level rise due to climate change has very little effect on the magnitude of seawater intrusion, and also that groundwater extraction is the main cause of seawater intrusion. Varying scenarios of groundwater pumping and tidal effects were considered in developing a 3-D density dependent numerical model to simulate the seawater intrusion in the coastal aquifer of Shenzhen city, China. The model results demonstrated that, while tide induced saltwater intrusion is significant near the estuarine Dasha River, groundwater exploitation up to a certain limit could also decrease the tendency of seawater intrusion into the aquifer (Lu et al., 2013). Intense groundwater pumping for irrigation since 1980 had created a saltwater intrusion problem in the central part of the Korba aquifer at the east of the Cap Bon peninsula in Tunisia (Kerrou et al., 2010). A 3-D transient density dependent groundwater model for the area demonstrated that the aquifer had been overexploited, and if the current rate of exploitation continues for the next 50 years, it would take at least 150 years to return to its natural condition, if no exploitation was permitted for this period.

Unsal et al. (2014) assessed the effects of climate change (sea level rise and decreased recharge) and over extraction for a hypothetical circular aquifer system, considering both quantity (groundwater reserve) and quality (freshwater-saltwater interface movement in the lateral and vertical directions) of groundwater resources for evaluating long-term sustainability of aquifer systems. A comparison of the effects of groundwater abstraction and 21st century sea level rise to seawater intrusion in a coastal area near the city of Monterey, California, USA, also revealed that groundwater abstraction is the main reason of seawater intrusion (Loáiciga et al., 2012).

### 2.2.2 Saltwater intrusion due to sea level rise

According to the present best estimate of the global mean sea level, a rise of 50 cm will be observed in the coming century (Oude Essink, 2001a). This will have a serious effect on the today's vulnerable coastal aquifers, with an increase in the salinization process of coastal aquifers. The most significant impacts of climate change on the coastal groundwater resources are sea level rise, decreased recharge, and more storm surge events. Climate change is likely to have an effect on sea level rise, and it is anticipated that a rise of 50 cm by the year 2100 will lead to an increased seawater intrusion in the coastal aquifers of many parts of the world (Shrivastava, 1998). Climate change induced sea level rise imposes additional saline water heads at the seaside, leading to expectation of more seawater intrusion in coastal aquifers (Yang et al., 2015). Low-lying sandy coastal aquifers, especially, are more susceptible to the effect of sea level rise. It is presumed that an increase in sea-level rise due to climate change is also a cause of saltwater intrusion in coastal aquifers, although anthropogenic activities such as over extraction, and excess paving in urbanized areas are the major causes of saltwater intrusion (Unsal et al., 2014).

A numerical modelling study in the low-lying Dutch Delta in the Netherlands indicated that sea level rise would indeed increase the hydraulic heads in groundwater system, and the impact of sea level rise would be confined within 10 km of the coastline and main rivers (Oude Essink et al., 2010). On the other hand, Chang et al. (2011) demonstrated that sea level rise would have no long-term effect on the steady state movement of the salt wedge. In addition, their transient confined flow simulation results revealed that the formations affected by saltwater intrusion due to sea level rise have a self-setback mechanism, and the intruded salts could be driven back to their original position. In contrast, Webb and Howard (2011) showed that sea level rise induced saltwater intrusion is more problematic for aquifer systems, and several centuries would be needed to equilibrate after the sea level rise has ceased. Sea level rise results in a high ratio of hydraulic conductivity to recharge and high effective porosity in the aquifer systems. This high ratio of hydraulic conductivity to recharge and high effective porosity are responsible for developing a large degree of disequilibrium in the aquifer systems. In contrast, aquifer systems with a low ratio of hydraulic conductivity to recharge and low effective porosity required only decades to equilibrate after the termination of sea level rise (Colombani et al., 2015).

In a confined alluvial aquifer with fine-medium sands and with a normal recharge rate, for example, the shallow coastal Golden Grove aquifer of Jamaica, the effect of sea level rise and storm surge on saltwater intrusion is not significant (Shrivastava, 1998). In contrast, Chang et al. (2011) demonstrated that the influence of sea level rise on saltwater intrusion is less in unconfined aquifers compared with that in confined aquifers. The presence of highly

permeable layers in aquifer systems can control the seawater intrusion processes due to sea level rise. For instance, Kim et al. (2013) evaluated homogenous and highly permeable aquifer systems, and showed that an aquifer with a highly permeable layer is less sensitive to sea level rise compared to a homogenous aquifer. They investigated three scenarios of sea level rise (0, 0.5, and 1 m) for both aquifer systems in Jeju Island, Korea. Their results showed that a sea level rise of 1 cm contributed to an inland seawater encroachment of 11 m for homogenous, and 6-8 m for highly permeable aquifer systems.

Recent studies on the impact of sea level rise due to climate change mainly focused on various hydrogeological settings and sea level rise scenarios. However, land surface inundation due to sea level rise has a more extensive impact on saltwater intrusion than the effects of pressure changes at the shoreline in unconfined coastal aquifers with realistic parameters (Ataie-Ashtiani et al., 2013). In addition, the relative importance of sea level rise, groundwater extraction and groundwater recharge largely depends on the specific location considered (Green and MacQuarrie, 2014). The effect of sea level rise on saltwater intrusion or movement of the freshwater-seawater interface in an inland direction in coastal aquifers largely depends on the coastal topography next to the shoreline (Yechieli et al., 2010). In the Mediterranean and Dead Sea coastal aquifers, Yechieli et al. (2010) showed that a horizontal slope of 2.5‰ would cause 400 m inland shift of the freshwater-saltwater interface with a sea level rise of 1 m in around 100 years. They also found that a vertical slope would yield no shift within the same period, and reduced recharge or over extraction of groundwater would enhance the inland shift of the interface. The following paragraphs will outline some of the recent works conducted on saltwater intrusion in the coastal aquifers due to sea level rise.

A study utilizing three different scenarios of future climate change in the Nile Delta aquifer in Egypt and the Madras aquifer in India, revealed that a 50 cm rise in water level in the Mediterranean Sea and in the Bay of Bengal will cause an additional intrusion of 9.0 km in the Nile Delta aquifer, but only 0.4 km in the Madras aquifer (Sherif and Singh, 1999). Oude Essink (2001b) employed three sea level rise scenarios: no rise, a rise of 0.5 m per century, and a fall of 0.5 m per century, in a Noord-Holland province. Their results showed that a relative sea level rise of 0.5 m per century would increase the saltwater intrusion in all low-lying areas of the northern part of the Noord-Holland province. In the well fields of the coastal aquifer in Broward County, Florida, USA, a sea level rise in excess of 48 cm over the next 100 years would result in chloride contamination (Dausman and Langevin, 2005). Giambastiani et al. (2007) demonstrated that a sea level rise of 0.475 m per century in an unconfined coastal aquifer in Comacchio, Italy, would result in an increase of 40% salt load and 800 m inland shifting of the mixing zone between fresh and saline groundwater.

Langevin and Zygnerski (2013) demonstrated that seawater intrusion near the well field in a shallow coastal aquifer system in south-eastern Florida, USA was dominated by abstraction of water from the well field, and that historical sea level rise could exacerbate the extent of saltwater intrusion. They reported a saltwater intrusion of about 1 km for about 25 cm sea level rise during the simulation period (105 years, from 1900 to 2005). For a sandstone coastal aquifer in the Richibucto region of New Brunswick in Atlantic Canada, Green and MacQuarrie (2014) simulated various combinations of different scenarios: two groundwater recharge scenarios: 40 and 85 mm/year; and two sea-level rise scenarios of 0.93 and 1.86 m; pumping: increased by a factor of 2.3 for the period from 2011 to 2100. They concluded that sea level rise has the least significant effect on saltwater intrusion in the sandstone aquifers in Atlantic Canada. Yang et al. (2015) investigated two scenarios in northern Germany: the influences of mean sea level rise of 1 m in the next 100 years, and a storm surge event on groundwater quality for a two-dimensional (2-D) cross-sectional coastal aquifer near Bremerhaven, considering the aquifer heterogeneity. Their results indicated that 1 m sea level rise in the next 100 years would move the seawater-freshwater interface 1,250 m further inland. They also predicted a 2,050 m landward expansion of the salinized area. However, they found the salinization effect in the deeper and more highly conductive layers of the aquifer. Modelling results of the impact of a storm surge under a dyke failure situation demonstrated that the overtopping seawater infiltrated into the aquifer and only the top part of the aquitard would be salinized, while the groundwater in the deeper aquifer would remain unaffected.

The saltwater intrusion problem can also be represented as the loss of freshwater reserve in coastal aquifers. In a coastal aquifer in Israel, changes in the groundwater head for an assumed sea level rise of 0.5 m accounted for about 23% of the freshwater loss in the aquifer (Melloul and Collin, 2006). The relationship between sea level rise and the global groundwater depletion rate were demonstrated by Wada et al. (2010), who stated that global reserve of groundwater is depleted at a rate of 0.8 mm/year (one quarter of sea level rise) against a sea level rise of 3.1 mm/year. In contrast, Rozell and Wong (2010) found an increase of 1% of freshwater lens volume in a Shelter Island, New York State, USA, even in an unfavourable climatic condition of sea level rise (0.61 m) and decreased groundwater recharge (2% precipitation decrease). The scenario of the most favourable condition (15% precipitation increase and only 0.18 m sea level rise), resulted in only a 3% increase in the freshwater lens volume. The aquifer is underlain by a clay layer restricting the maximum depth of the aquifer, and this allows for this unexpected groundwater volume increase under unfavourable climate change conditions. Land surface inundation due to sea level rise has a considerable impact on the fresh groundwater lenses in small oceanic circular islands. However, aquifer recharge, the thickness of the aquifer and hydraulic conductivity have more

profound influence on the volume of fresh groundwater lenses compared with sea level rises (Ketabchi et al., 2014).

Saltwater intrusion also depends on the nature of the inland boundaries. Saltwater intrusion due to sea level rise is more pronounced in aquifers with head controlled inland boundaries compared with an aquifer system with flux controlled inland boundaries, because head-controlled systems are associated with maximum seawater intrusion due to sea level rise (Werner and Simmons, 2009; Werner et al., 2012). In addition, Werner et al. (2012) confirmed that in constant discharge confined aquifers, seawater intrusion is not induced by sea level rise. Mazi et al. (2013) used a generalized analytical solution in modelling an unconfined coastal aquifer to investigate the effects of climate driven scenarios of sea level rise on saltwater intrusion. They considered both flux and head control conditions, and concluded that the flux control case is far more resilient than the head control case, in terms of initial seawater intrusion. They introduced three important thresholds or tipping points: spatial, temporal, and managerial tipping points, and found high nonlinearity between the sea level rise and seawater intrusion. The aquifer responses to sea level rise would shift abruptly from a stable state of mild change responses, to a new stable state of large responses, to even small changes that lead rapidly to complete (deep) seawater intrusion if these tipping points were passed. Lu et al. (2015) derived analytical solutions to predict the distance of interface toe movement in response to 1 m sea level rise in confined and unconfined coastal aquifers, respectively, with a general head boundary condition. They also compared the results with those obtained from using constant head (upper bound) and constant flux (lower bound) inland boundaries with the same initial system condition. Depending on the values of two general head boundary parameters (hydraulic conductance and reference head), the predicted values of the interface toe movement lies between those using a constant head (upper bound) and constant flux (lower bound) inland boundaries.

Watson et al. (2010) investigated the transience of saltwater intrusion with regard to sea level rise in an idealized unconfined coastal aquifer system subjected to an instantaneous sea level rise. A simplified conceptual framework comprising homogeneous aquifer conditions and constant aquifer stresses was considered, in order to extend the findings of recent studies concerning the saltwater intrusion response to sea level changes. It was observed that the simplified steady state sharp interface solution overestimated the 100-year landward toe movement for most of the cases studied. In addition, the steady state sharp interface estimates of toe shift span 40%-250% of the toe shifts obtained from the 100-year dispersive interface simulations. This large range indicates that steady state sharp interface estimates are at best only crude initial estimates of 100-year “planning time frame”.

### 2.2.3 Saltwater intrusion due to other reasons

In addition to the factors discussed in sections 2.2.1 and 2.2.2, there are other factors reported to be responsible for the salinization of the aquifer. Aquifer salinization can occur from the karst structures developed as a result of bacterial reduction of the dissolved sulphate. A good example of this type of aquifer salinization was reported in the Guanahacabibes Peninsula of Pinar del Rio Province, western Cuba (Boschetti et al., 2015). The orientation of fractures in a krastified coastal aquifer and fluid density were also reported to be the drivers of seawater intrusion in a krastified coastal aquifer on the island of Crete (Dokou and Karatzas, 2012). The other reported causes of seawater intrusion include extensive agricultural land drainage systems in the Pleistocene Crag aquifer in the Thurne catchment in northeast Norfolk, UK (Simpson et al., 2010), leakage during drilling works (Zhang et al., 2015), and groundwater fluxes due to changes in rainfall patterns (Chang and Clement, 2012). In some aquifers, several processes occur simultaneously and some or all of the processes are responsible for aquifer salinization. For instance, aquifer salinization in the irrigated coastal shallow aquifers of Aousja-Ghar El Melh and Kalâat el Andalous, in north-eastern Tunisia were reported to be caused by the combined effect of water-rock interaction, evapotranspiration, irrigation return flow, sea aerosol spray, and agricultural fertilizers (Bouzourra et al., 2015). They also noted that carbonate and sulphate precipitation occurred because of increasing concentrations of solutes in groundwater due to evaporation.

Vegetation dynamics is also an important parameter to be considered for groundwater study, and neglecting to consider the uptake of groundwater by vegetation might result in an overestimation of both the amount and quality of available freshwater resources, and might lead to an unrealistic and unsustainable groundwater management strategy (Comte et al., 2014). They simulated the evolution of groundwater salinity in Grande Glorieuse, a low-lying coral island in the Western Indian Ocean. Their results revealed that the combined effect of sea level rise and climate change (rainfall and evapotranspiration) would result in an average increase in salinity of 140% (+8 $kg/m^3$), whereas the increase in salinity could reach up to 300% (+10 $kg/m^3$) in low-lying area with high vegetation density.

## 2.3 Approaches for prediction and management of saltwater intrusion in coastal aquifers

Saltwater intrusion modelling approaches can be classified into two broad categories: sharp interface and diffuse interface models. Sharp interface models comprise analytical solutions and numerical models, while diffuse interface models are usually based on numerical models. Figure 2.1 illustrates the flow chart of the common approaches utilized for prediction and management of saltwater intrusion processes in coastal aquifers. In the sharp interface

modelling approach, a sharp interface is assumed between the fresh and saline groundwater. However, in reality a diffuse interface exists between them. In this diffuse interface, the density of water gradually decreases from the seawater side to the freshwater side.

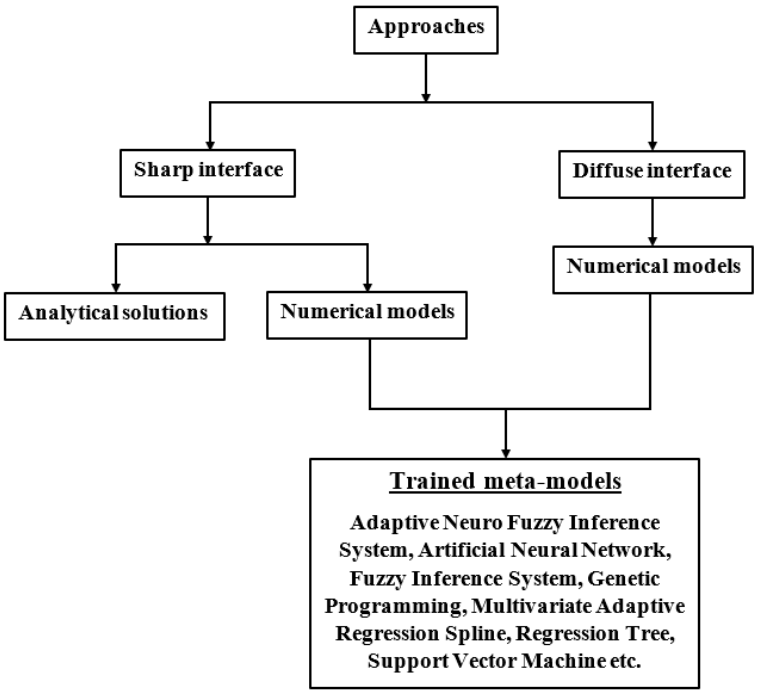


**Figure 2.1 Modelling approaches used to predict and manage saltwater intrusion**

### 2.3.1 Analytical solutions

Analytical solutions provide a relatively simple and clear concept of a problem. They are often used as benchmarks for numerical solutions. Analytical solutions are a physically based approach and computationally less intensive, compared with a non-physically based approach (Werner et al., 2012).

The first analytical solution of saltwater intrusion problem was given by the Ghyben-Herzberg relationship (Cheng et al., 2000). It is based on the assumption of hydrostatic conditions in a homogeneous unconfined coastal aquifer. The modified Ghyben-Herzberg relation uses observed piezometric head in the saltwater zone. The Ghyben-Herzberg relation can be used primarily to locate the saltwater-freshwater interface. Strack (1976) developed a 2-D analytical solution to locate the position of the saltwater-freshwater interface in coastal aquifers. Strack's solution assumed a homogeneous and isotropic aquifer with sharp interface between freshwater and saltwater. It also assumed pumping wells to be fully penetrating into the aquifer and intersecting the lower boundary of the aquifer. This assumption is typically wrong in a real situation and, therefore, Beebe et al. (2011) recommended that Strack's solution was not suitable for predicting saltwater intrusion in pumping wells. Bower et al.

(1999) developed an analytical model for predicting the extent of saltwater upconing in a leaky confined aquifer using a sharp interface assumption.

Benchmarking the performance of a numerical code against standard analytical solutions is necessarily the first step in testing the accuracy of the numerical approximations (Simpson and Clement, 2003). Henry and Elder problems, commonly used for the benchmark test, are associated with variable density flow in a variably saturated porous medium.

### 2.3.2 Numerical modelling

Several saltwater intrusion models were developed based on sharp interface assumptions. Sbai et al. (1998) used a sharp interface assumption to simulate saltwater intrusion under steady state and transient conditions using the finite element method. Marin et al. (2001) assumed sharp interface to develop a quasi-3-D finite difference model to simulate groundwater flow in Karstic aquifer of north-western Yucatan, Mexico. Based on a sharp interface assumption, Fleet and Baird (2001) developed a 2-D model to predict saltwater intrusion due to over extraction from the coastal aquifer in Tripoli, Libya.

To reduce the computational burden in saltwater intrusion problems, Llopis- Albert and Pulido- Velazquez (2015) proposed a transient sharp interface model that does not require a calibration process and the implementation of a density dependent model. Their results revealed that the sharp interface approach provides a good agreement in terms of piezometric heads regarding the density dependent transport model. This was confirmed by conducting a numerical experiment in a 3-D unconfined synthetic aquifer with spatial and temporal distribution of recharge and pumping wells. The overall conclusion was that, when applied to transient problems, the sharp interface approach provided a considerably accurate result, as well as improving the steady state results. Therefore, they concluded that this approach could replace the application of computationally intensive density dependent models, reducing the computational cost.

Freshwater and saltwater mixes because of hydrodynamic dispersion and diffusion, and consequently develops a transition zone in coastal aquifers. The thickness of the transition zone depends on structure of the aquifer, extraction from the aquifer, variability of recharge, tides, and climate change, and could range from a few meters to several kilometres in over pumped aquifers (Todd, 1974). For the simulation of the physical processes in coastal aquifers, an accurate and realistic modelling approach is the use of density dependent miscible flow and transport simulation modelling. For saturated and unsaturated soils, Cheng et al. (1998) developed a 2-D density dependent flow and solute transport model. Sakr (1999) investigated the limitation of the sharp interface approach by presenting a finite element density dependent solute transport model. Advective and dispersive transport of saltwater

below a partially penetrating pumping well was presented by means of a numerical model (Shalabey et al., 2006). A 3-D variable density flow model considering the development of a transition zone was developed and applied to a heterogeneous coastal aquifer to investigate the pumped water quality degradation due to pumping rate, the salinity of freshwater inflow and the thickness of the aquifer (Doulgeris and Zissis, 2014). The developed model was verified against the Henry seawater intrusion problem, the Elder salt convection problem, and experimental results of the 3-D salt pool problem. Results demonstrated that the concentration of salts in the pumped water increased with the increase in aquifer thickness regardless of the aquifer transmissivity.

Das and Datta (2000) developed a large-scale nonlinear optimization based methodology to obtain solution of the discretised density dependent, 3-D transient saltwater intrusion process in coastal aquifers. The discretised governing equations were specified as binding constraints in the nonlinear optimization model with the objective of satisfying these equations. The solution results were validated using benchmark saltwater intrusion problems.

## 2.4 Management models in saltwater intrusion problems

The management models for managing the saltwater intrusion processes in coastal aquifers require development of regional scale management strategies. These management models for prescribing spatially and temporally varying large-scale management strategies are generally based on different optimization algorithms, and often integrate an optimization model with simulation models to accurately simulate the flow and transport processes in an aquifer.

Willis and Finney (1988) were the first to develop a planning model for the control of seawater intrusion in regional groundwater systems. Later, Finney et al. (1992) proposed another optimization model for the control of saltwater intrusion. The optimization technique was further implemented in the development of a coastal aquifer management strategy based on optimized pumping (Bhattacharjya and Datta, 2009; Das and Datta, 1999b; Dhar and Datta, 2009a; Sreekanth and Datta, 2011b). The accuracy of the management strategy depends on the level of approximation and the way in which the strategy is incorporated. To develop saltwater intrusion management models, the simulation models are combined with the management models, either by using the governing equations as binding constraints in the optimization model or by using a response matrix or external simulation model. The simulation model component of the management model is generally based on the finite difference or finite element approximations of the partial differential equations of groundwater flow and solute transport (Das and Datta, 1999b).

The coupled S-O approach (linking groundwater simulation models with optimization techniques in a single framework) is the more efficient approach for

determining optimum management strategies of coastal aquifers (Ataie-Ashtiani et al., 2014). This coupling of simulation and optimization models can be developed by utilizing an embedding technique (binding the discretised governing equations into the constraints of the optimization model) (Das and Datta, 1999b), a response matrix (Gorelick, 1983; Yang et al., 2001), or by employing external linking (Das and Datta, 2001). The commonly used techniques to incorporate the simulation model within the management model are the embedding technique and the response matrix approach (Gorelick, 1983).

Das and Datta (1999b) demonstrated the potential feasibility of the embedding technique for the development of multiple objective coastal aquifer management models. In this procedure, the finite difference form of density dependent coupled flow and transport equations are embedded as constraints, and the constraint method is used to obtain the Pareto-optimal or non-inferior set of solutions for the multiple objective models. Later, Das and Datta (2000) proposed a nonlinear optimization method to solve the embedded governing equations for the simulation of seawater intrusion in coastal aquifers. They used a gradient search process to solve the discretised flow and transport equations simultaneously. They compared the results of the proposed methodology from those of the other researchers and found substantial agreement of the results with the known solutions. However, the embedding technique is computationally inefficient for a large-scale heterogeneous aquifer system and has several limitations when used for a saltwater intrusion management model in a large-scale aquifer system (Das and Datta, 1999a; Willis and Finney, 1988). Moreover, in the embedding approach, a large number of decision variables need to be employed in the optimization model (Bhattacharjya and Datta, 2009). On the other hand, the response matrix approach, based on the principle of superposition and linearity, is unsatisfactory for nonlinear systems (Rosenwald and Green, 1974). The response matrix approach is applicable when the system is linear or approximately linear and the boundary conditions are homogeneous. Any change in boundary condition, location of the source/sink, and observation wells requires several simulations to generate the responses and requires recalculation of the response matrix (Das and Datta, 2000).

However, for highly complex and nonlinear 3-D numerical models, a management tool based on external linking of the simulation model with an optimization algorithm is more feasible (Ataie-Ashtiani et al., 2014; Bhattacharjya and Datta, 2005). For the saltwater intrusion management models, a linked S-O approach was proposed as an alternative to embedding technique and response matrix approach (Emch and Yeh, 1998; Gorelick, 1983). Several authors incorporated an external linking technique to develop saltwater intrusion management models by externally linking coastal aquifer simulation models with an optimization model (Bhattacharjya and Datta, 2005; Sreekanth and Datta, 2011b; Sreekanth and Datta, 2014b). In a linked S-O technique, the simulation model provides necessary

information to the optimization model at every stage of iteration in order to reach an optimal solution (Sreekanth and Datta, 2011a).

Bhattacharjya and Datta (2005) developed a single-objective management model to develop a saltwater intrusion management model. This was operated by externally linking a properly trained and tested ANN meta-model as an approximate simulator, with the GA based optimization model in a linked S-O framework. The developed methodology was evaluated using an illustrative study area, and the evaluation result suggested that GA-ANN based linked S-O approach can be applied for the optimal management of coastal aquifers. However, the management of groundwater resources in coastal regions usually involves several contradictory objectives which affect each other adversely. Some examples of these objectives include maximizing pumping from the production wells for the beneficial use while maintaining the salinity level in the pumped water within a permissible limit, minimizing pumping from the barrier wells, maximizing profit from the pumped water, minimizing the cost of pumping, etc.

Multiple objective optimization of coupled simulation and optimization models can address these trade-offs between the contradictory objectives (Ataie-Ashtiani et al., 2014; Das and Datta, 1999b; Grundmann et al., 2012; Sreekanth and Datta, 2010). A multiple objective optimization technique for a coastal aquifer management problem generates a large number of so-called Pareto optimal solutions, posing a significant challenge in interpreting and communicating those solutions to decision makers (Subagadis et al., 2014). Therefore, post Pareto optimality analysis is necessary to communicate the Pareto optimal solutions from the multiple objective decision making problems (Reddy and Nagesh, 2007; Subagadis et al., 2014).

To solve multiple objective saltwater intrusion management problems, Bhattacharjya and Datta (2009) utilized an ANN based meta-model externally linked with an optimization model. The results obtained from this linked S-O framework were compared with those obtained from the embedded technique and the classical nonlinear optimization technique. They concluded that the optimization model needs to handle only few variables and explicit constraints in a linked S-O technique. Therefore, this technique performed better than the embedding technique. Dhar and Datta (2009a) evaluated the computational efficiency of the developed multiple objective management model by using three different modelling scenarios: direct linking of the numerical model with the optimization model, ANN based meta-model as a replacement of the numerical simulation model, and use of a partially trained meta-model. The performance evaluation of the developed methodologies was carried out in an illustrative coastal aquifer system, indicating that the developed methodologies were applicable for multiple objective management of saltwater intrusion in coastal aquifers. Sreekanth and Datta (2011a) evaluated the performance of an ANN and a GP based saltwater

intrusion management model by using an illustrative coastal aquifer system to determine the optimal groundwater extraction strategies from coastal aquifers. Sreekanth and Datta (2011c) proposed another optimal pumping management strategy, developed by externally linking a properly trained and tested ANN model with a multiple objective optimization algorithm. Recently, Dentoni et al. (2015) developed a saltwater intrusion management model for the Gaza strip coastal aquifer in Palestine, using a linked S-O technique. They showed that the developed optimal scheme significantly lowered the salt concentration of the extracted water (by 23%), while keeping the overall extraction close to the user defined total amount. Moreover, they reported an average increase of groundwater heads by 4.5% and a decrease in the seawater intrusion affected area by 5% with the optimized solution.

To develop saltwater intrusion management models, Abd-Elhamid and Javadi (2011) proposed a methodology combining abstraction of water from the aquifer and recharge to the aquifer using desalinated water. Later, Javadi et al. (2015) modified this methodology by using treated wastewater as a source of water for aquifer recharge. Javadi et al. (2015) considered different management scenarios to find the best management practices to control seawater intrusion in coastal aquifers, with a view to minimizing the cost of management process while minimizing the total salinity in the aquifer. A multiple objective optimization formulation was utilized to find optimal solutions for each scenario of the saltwater intrusion management problem. Results indicated that the proposed methodology maximizes the impedance of freshwater-saline water interface and induces the seaward movement of the seawater body while providing the least cost and least salt concentration in the aquifer. They also claimed that their methodology required a lower volume of water for aquifer recharge to control seawater intrusion. They concluded that treated wastewater is the most cost effective option of recharging the aquifer and the use of desalinated water should be restricted to domestic use only. Different combinations of recharge and abstraction were also considered in the development of saltwater management models by Hussain et al. (2015). Minimizing the cost of management approaches and the salinity level in the aquifer was chosen as the objective function to evaluate the efficiency of each management scenario. Among the management alternatives, the application of treated wastewater and/or storm water, coupled with continuous abstraction of brackish water and its desalination and use, were found to be the most cost effective alternatives to control seawater intrusion.

A management plan for simultaneous minimization of economic and environmental costs was proposed for the island of Santorini, utilizing the formulation of a multi-objective optimization framework (Kourakos and Mantoglou, 2011). Subagadis et al. (2014) presented an integrated approach to support management decisions of coupled groundwater-agricultural hydro systems under uncertainty. They also demonstrated the functionality of the proposed methodology in finding the appropriate management interventions for the saltwater

intruded Al-Batinah coastal agricultural region in Oman. This methodology addressed the two contradictory management objectives: sustainable aquifer management and profitable agricultural production. Chen et al. (2014) developed a management model for the large-scale conjunctive use of surface water and groundwater resources. This model aimed to maximize the supply for irrigation and for public supplies by imposing constraints on groundwater level drawdown for the Chou-Shui alluvial fan system in central Taiwan. They showed that if the drawdown is limited to 1 m/year, the deficiency could be increased by an average of 25.2 million $m^3$/year for the Chou-Shui alluvial fan. A regional scale model utilizing a scenario simulation employing the numerical model OpenGeoSys was developed for the southern Al-Batinah region on the northern coast of Oman to assess the development of the groundwater levels and salinity intrusion in the past and to project a future scenario (Walther et al., 2014).

## 2.5 Meta modelling

Sanford and Pope (2010) described the challenges of utilizing models to forecast seawater intrusion, and suggested that accurate simulation of regional scale (many km) saltwater intrusion and accurate forecasting, even for a single well, is computationally prohibitive, even with a massively powerful computer. For a simulation, sensitivities are approximated by calculating the ratios of the resulting marginal (incremental) changes in the output variables to the changes in the input variables that caused them (Blanning, 1975). Therefore, the simulation part of the S-O framework has a relatively high computational cost. Moreover, to get an optimum solution in the irregularly shaped feasible regions of the problems, the simulations must be performed many hundreds or thousands of times in a typical optimization problem. Runtime of the optimization can be reduced by introducing parallel or distributed algorithms on the simulation level such as domain decomposition (Farhat and Roux, 1991) or on the optimization level such as parallel direct search (Torczon and Trosset, 1998). Another way of reducing the computational effort required is to apply approximation techniques, in order to reduce the number of simulations (Nakayama et al., 2001). Sufficiently accurate and computationally less intensive approximate solution of the flow and transport processes in coastal aquifer may also be useful (Bhattacharjya and Datta, 2005).

To obtain an optimal management strategy, density dependent flow and transport simulation models need to be coupled with a suitable optimization algorithm. External linking of the simulation model with the optimization based management model is an appropriate solution of coupling, which is very complex and difficult. Moreover, this technique demands a considerable computational time, due to the large number of iterations between the optimization and simulation model required to obtain an optimal management

strategy (Sreekanth and Datta, 2011a). This modelling bottleneck can be solved either by using parallel algorithms and better computer configurations (Dougherty, 1991), or by using an approximate simulator, called a meta-model or surrogate model.

To achieve computational efficiency, simulation models can be replaced with a properly trained and tested surrogate model or meta-model in the S-O framework. The use of the meta-modelling approach in optimization processes to reduce the computational time was proposed by Blanning (1975). A meta-model can be used as a surrogate to calculate fitness values, which are normally based on time consuming simulations. Such a meta-model can be effectively integrated into the search process to gradually substitute for the large portion of simulation (Behzadian et al., 2009). Hemker et al. (2008) showed that the surrogate approaches can locate design points better than other approaches, while significantly improving the computational efficiency. However, meta-models should preserve the size and scope of the problem domain and should reflect the empirical relationships between decision variables and selected model outcomes (Virginia and Leah, 2000). The use of meta-models in an S-O framework requires an additional calibration effort, in order for the meta-models to represent the system and produce good predictions. Meta-models require relatively small execution times to produce model predictions, in spite of the extra computational time required for the meta-model calibration (Kourakos and Mantoglou, 2013).

Meta-models should be trained with the representative input-output patterns from the entire decision space (Sreekanth and Datta, 2011a). Dividing the dataset into two mutually exclusive subsets (the training set and the validation set), called the holdout method, is commonly used during the model validation and selection process. A part of the data is usually used to train the meta-model, while the rest of the data is used to validate the meta-model (Namura et al., 2011). This results in the overfitting of the training data and the under fitting of the validation data (Luo and Lu, 2014). The holdout method can be improved by a method called cross-validation, in which both the training and validation data are used. However, the cross-validation method is rarely used in the field of groundwater optimization field for the estimation of the meta-model accuracy (Razavi et al., 2012). A comprehensive review of meta-model-based optimization approaches in general can be found in Jin (2005), and in Jin and Branke (2005). Polynomial Regression (PR), ANN, kriging, Support Vector Machine, and MARS are common methods to build meta-models that can be used to substitute the complex numerical models (Luo and Lu, 2014).

However, there are some disadvantages in utilizing these existing meta-models for the prediction of saltwater intrusion in coastal aquifers. The structure of ANN models (e.g. number of hidden layers and the number of neurons in the hidden layer) needs to be fixed a priori (Sreekanth and Datta, 2011a), which requires a compromise between model complexity and accuracy. Other disadvantages of ANN are its “Black-Box” nature, its higher

computational burden, and its susceptibility to overfitting (Tu, 1996). On the other hand, GP models are based on explicit mathematical formulations (Shiri and Kisi 2011) that are intended to develop surrogate models, which are simple regression models (Sreekanth and Datta, 2011a). Besides, GP needs to evaluate millions of model structures before finding the optimal model structure (Sreekanth and Datta, 2011a), which requires extensive training time. Another limitation of GP is its propensity to converge prematurely to local optima, and its consequent inability to find solutions in some instances (Pillay, 2004) . In situations where multimodal fitness functions (e.g. containing many peaks or depressions) are used, GP systems are guaranteed to converge prematurely (Pillay, 2004).

### 2.5.1 Fuzzy Inference System as meta-model

Fuzzy Inference System (FIS), which is based on fuzzy set theory, has recently received considerable attention. FIS is recognized as a successful computing framework due to its applicability in a multiplicity of fields (Jang et al., 1997). FIS is capable of capturing nonlinear relationships between input and response variables and is an effective tool for modelling nonlinear processes (Sugeno and Yasukawa, 1993; Takagi and Sugeno, 1985). The most widely used FISs in different applications are Mamdani FIS, Sugeno FIS, and Tsukamoto FIS. They differ in the way they use consequent parts of their fuzzy rules, and the way in which they accomplish the aggregation and defuzzification steps. The Sugeno fuzzy model (Sugeno, 1985), also known as Takagi-Sugeno-Kang model, is especially suited for modelling nonlinear systems by interpolating between multiple linear models. Despite the potential capability of emulating complex and nonlinear systems, the application of FIS to approximate physical processes of coastal aquifer systems is quite limited. The present study intends to utilize an FIS model as a computationally efficient substitute for the complex numerical simulation model within a linked S-O methodology to develop saltwater intrusion management model.

Recently, a data driven method based on fuzzy logic theory has been used as an efficient computational tool in different water management problems (Allen et al., 2007; Ayvaz et al., 2007; Kord and Asghari Moghaddam, 2014). In addition, a limited number of research works examined the use of fuzzy logic in groundwater management problems (Altunkaynak, 2010; Muhammetoglu and Yardimci, 2006). However, thus far fuzzy logic has not been used as a meta-model to predict saltwater concentrations at different monitoring locations in a coastal aquifer under the influence of transient pumping stresses. In the proposed study, an FCM clustering algorithm (Bezdek et al., 1984) is used to develop a Sugeno type FIS. The FIS meta-model has the advantage of predicting salinity concentrations

at multiple monitoring locations. Therefore, it overcomes the need for multiple meta-models for multiple output problems.

However, one of the major challenges in using a fuzzy logic based prediction modelling approach is to manage large dimensional input datasets, in particular to handle a large number of spatially and temporally varying groundwater extraction patterns for saltwater intrusion management in coastal aquifers. In such situations, reducing the dimensionality of the input space by using the Fuzzy C-Mean Clustering (FCM) algorithm (Bezdek et al., 1984) provides a reasonably practical solution. FCM is used to compress the entire input space into a number of identical clusters. This clustering technique significantly reduces the number of fuzzy IF-THEN rules and the number of modifiable parameters (linear and nonlinear) of the generated FIS. The present study utilizes FCM technique to reduce the dimensionality of the input space in developing the FIS based meta-model to approximate saltwater intrusion processes in coastal aquifers.

### 2.5.2 Adaptive Neuro Fuzzy Inference System as meta-model

The Adaptive Neuro Fuzzy Inference System (ANFIS) proposed by Jang (1993), is a multilayer adaptive network based FIS that integrates the benefits of both the neural networks and fuzzy logic approaches. It incorporates the basic advantages of ANN, such as massive parallelism, robustness, and learning in data rich environments (Sudheer and Mathur, 2010). Many studies utilized ANFIS in groundwater modelling applications, for example, in predicting groundwater level (Emamgholizadeh et al., 2014), characterizing groundwater quality parameters (Khaki et al., 2015), forecasting river flow (He et al., 2014), assessing groundwater quality (Khaki et al., 2015), spatial distribution of groundwater quality (Khashei-Siuki and Sarbazi, 2015), predicting water quality index (Sahu et al., 2011), estimating groundwater level (Sreekanth et al., 2011), predicting daily discharge responses of a large karstic aquifer (Kurtulus and Razack, 2010) and predicting electrical conductivity of groundwater (Tutmez et al., 2006). However, the ANFIS based meta-models have not been utilized to approximate coupled flow and salt transport processes in a multilayered coastal aquifer system with transient pumping stress applied to the aquifer.

### 2.5.3 Multivariate Adaptive Regression Spline as meta-model

Multivariate Adaptive Regression Spline (MARS) introduced by Friedman (1991) is a rapid, flexible, and accurate artificial intelligence technique for predicting both continuous and binary output variables (Salford-Systems, 2016). MARS is a nonparametric modelling technique in which no prior assumption is made regarding the functional relationship between input and output variables (Friedman, 1991). MARS builds the functional

relationship between input and output variables from the regression data using a set of coefficients and Basis functions (Friedman, 1991). The number of Basis functions and parameters associated with these basis functions are automatically determined by the training data. The major benefits of MARS lie in its capability to capture the predictor-response relationship of high-dimensional data patterns while producing simple and easy-to-interpret surrogate models (Zhang and Goh, 2016). Another advantage of MARS is its ability to build models based on the relative importance of predictor variables in determining the response.

Several previous studies verified the use of a MARS based meta-model in different research areas, for example, in runoff prediction for watersheds (Adamowski et al., 2012; Sharda et al., 2006; Sharda et al., 2008), in scour depth estimation below free overfall spillways (Samadi et al., 2015), in streamflow forecasting (Coulibaly and Baldwin, 2005), in rainfall prediction (Beuchat et al., 2011), and in groundwater potential mapping (Zabihi et al., 2016) etc. These studies have demonstrated the potential applicability of the MARS approach with acceptable prediction accuracy in different problem domains. However, the application of the MARS approach in saltwater intrusion management problems has been quite limited.

### 2.5.4 Gaussian Process Regression as meta-model

Gaussian Process Regression (GPR) is a stochastic process (Jacobs and Koziel, 2015) that performs modelling within a Bayesian framework in which the model variables are assumed to follow a Gaussian prior distribution. GPR, as a popular artificial intelligence tool, has been successfully applied in many engineering applications (Forrester et al., 2008). GPR models are nonlinear approximation models that are able to provide probabilistic information on prediction. In other words, they provide confidence in the prediction. According to the GPR approach, the way a machine learns is formulated in terms of a Bayesian estimation problem, in which the parameters of the machine are assumed as random variables drawn from a Gaussian distribution (Bazi et al., 2012). Given a set of training data, GPR provides a flexible Bayesian framework for identifying nonlinear relationships between predictors and the response variable (Sun et al., 2014). GPR is a nonparametric modelling approach, i.e. no assumption is made about the shape of the function to estimate. GPR provides a "principled, practical, and probabilistic approach to learning in kernel machines" (Rasmussen and Williams, 2005). Recently, Rajabi and Ketabchi (2017) demonstrated the potential applicability of GPR based emulators for single objective problem setting in a S-O based methodology in a coastal groundwater management problem. The capability of the GPR based meta-models needs to be evaluated for a multiple objective saltwater intrusion management models.

## 2.6 Overview of the optimization techniques in saltwater intrusion modelling

The efficient search for global optimal solutions is the most important part of the S-O process, and different optimization processes have been proposed by researchers to achieve this goal (Rao, 1996). The commonly used metaheuristic search algorithms are Simulated Annealing (SA) (Kirkpatrick et al., 1983), Tabu search (Fred, 1989), and Genetic Algorithms (GA) (Goldberg, 1989; Holland, 1992). GA (Goldberg, 1989; Holland, 1992) and SA (Kirkpatrick et al., 1983) are the two most commonly used non-gradient based heuristic search techniques utilized in the S-O framework. GA (Goldberg, 1989; Holland, 1992), Particle Swarm Optimization (Kennedy and Eberhart, 1995), and SA (Kirkpatrick et al., 1983) optimization methods are computationally expensive, in spite of the capability of finding global optima (Goel and Stander, 2009). To improve the convergence rate of the search process, Ingber (1989) proposed an adaptive SA by modifying the existing conventional SA algorithm. However, there is no guarantee that a global optimal solution be found using these two search techniques, although they can achieve nearly optimal solutions at a reasonable computational cost (Reeves, 1993). Heuristic techniques have the ability to locate solutions with greater efficiency in combinatorial optimization problems. On the other hand, branch and bound techniques are implicit enumeration techniques that have lower efficiency compared with heuristic techniques (Wagner, 1995). They have also advantages over the gradient based techniques because heuristic techniques can handle the discontinuities and nonlinearities of the real world problems (McKinney and Lin, 1994; Ritzel et al., 1994). However, the number of times the objective function must be calculated is still very high in heuristic search techniques. For example, to complete a single search, GA required 1,250 evaluations of the objective function (Wagner, 1995) whereas 2,344 evaluations of the objective functions were required by SA algorithm (Marryott et al., 1993).

Selection of an appropriate optimization algorithm is a crucial first step towards the development of a linked S-O methodology for solving coastal aquifer management problems. In multiple objective problem setting, GA (Deb et al., 2000) has received more attention, and is employed in several recent studies (Bhattacharjya and Datta, 2009; Hussain et al., 2015; Sreekanth and Datta, 2010) to solve saltwater intrusion management problems. As an improvement on the original Elitist Genetic Algorithm, Deb and Goel (2001) proposed a Controlled Elitist Multiple Objective Genetic Algorithm (CEMOGA), which demonstrates better convergence for a number of difficult test problems. To ensure convergence to an optimal Pareto front, CEMOGA maintains the diversity of population by controlling the elite members of the population as the algorithm progresses. The present study proposes the use

of the CEMOGA optimization algorithm integrated with a meta-model based simulation modelling approach in order to control saltwater intrusion in a multilayered coastal aquifer system.

## 2.7 Ensemble of meta-models to address prediction uncertainty

Replacing numerical simulation models by meta-models in a linked S-O framework adds a certain amount of uncertainty in the predicted variable (Sreekanth and Datta, 2011b). This predictive uncertainty resulting from the residuals affects the optimality and feasibility of the Pareto optimal solution of a coupled S-O model. Meta-models are associated with certain amount of uncertainty in prediction that results from the residuals. Using an ensemble of individual meta-models is one of the most effective ways of reducing the predictive uncertainty of the meta-modelling approach, in which outputs obtained from a group of separately trained meta-models are integrated to acquire one unified output (Zhou et al., 2002). An ensemble of meta-models is able to capture the accurate trend of data by combining the outputs from a certain number of individual models, thus providing better prediction capability than individual models within the ensemble (Goel et al., 2007; Jovanović et al., 2015). An ensemble of surrogate models recompenses prediction errors by integrating the outputs from all individual models of the ensemble, and is able to achieve better prediction than individual models (Jovanović et al., 2015; Sreekanth and Datta, 2011b). This approach is also advantageous because an individual meta-model often fails to map the true pattern of data from the entire parameter space. Nevertheless, individual meta-models within the ensemble should be adequately accurate and sufficiently diverse. Accuracy can be achieved by the proper choice of surrogate models and by adjusting the optimal combination of model parameters. Diversity in individual models can be maintained through utilization of several approaches including optimizing model structures and parameters, varying the training algorithm, and the use of different realizations of the training dataset (Sharkey, 1999; Zhang and Berardi, 2001). The random sampling without replacement (Hastie et al., 2008) technique can be utilized to obtain different realizations of the training dataset in order to form each individual ensemble member.

Various methods have been proposed in the literature to generate ensemble members: using different model architectures, adopting different training algorithms, or manipulating the training data set (Sharkey, 1999; Shu and Ouarda, 2008). The basic idea behind the development of an ensemble of meta-models for future prediction is to use each model's unique feature for capturing different patterns within the data set (Khashei and Bijari, 2011). Among other approaches to generate ensemble members, altering the training data using resampling techniques often provides better results (Shu and Burn, 2004; Zaier et al., 2010).

For combining the outputs of ensemble members, the most commonly used techniques are the average and stacked generalization (Shu and Ouarda, 2007; Wolpert, 1992).

## 2.8 Numerical modelling under groundwater parameter uncertainty

The extent of future saltwater intrusion scenarios in coastal aquifers can be obtained by simulating physical processes and spatial and temporal groundwater extraction patterns. The reliability and accuracy of the optimal management strategy based on a coupled S-O approach depends on how accurately the simulation model captures and simulates the accompanying physical processes. However, prediction of saltwater intrusion processes in coastal aquifers is a challenging task, due to the inherent uncertainties associated with model structure and uncertainties related to accompanying model parameters (Sreekanth and Datta, 2011c; Sreekanth et al., 2012). Another source of uncertainty arises from inappropriate and inadequate characterization of the accompanying physical processes in the subsurface system. Multidimensional heterogeneity of aquifer properties such as hydraulic conductivity, compressibility, and bulk density are considered to be major sources of uncertainty in the groundwater modelling system (Ababou and Al-Bitar, 2004). Other sources of uncertainty are associated with spatial and temporal variability of hydrologic as well as human interventions, e.g. aquifer recharge and transient groundwater extraction patterns.

Highly nonlinear and dynamic groundwater flow and solute transport processes are associated with a certain amount of uncertainty related to groundwater flow parameters. Therefore, the developed meta-models should also incorporate this parameter uncertainty to a certain extent. This can be done by training a number of meta-models by using different realizations of uncertain model parameters and combining them in the framework of ensemble meta-modelling approach.

## 2.9 Adaptive management of coastal aquifers using sequential compliance monitoring network design

Despite the usefulness of a management strategy in prescribing the optimal pumping management plan for a specified time horizon, field level implementation may deviate from the optimal pumping prescribed for implementation. This operational uncertainty, along with uncertainties arising from changing groundwater parameters, demands the adaptive management of saltwater intrusion processes using an optimally designed sequential compliance monitoring network, in order to assess the compliance of a proposed management strategy after implementation. Identifying a set of optimal monitoring locations is an important part of the adaptive saltwater intrusion management strategy. These optimal

monitoring wells are placed at locations where there is high uncertainty in terms of deviation in salinity concentrations resulting from a set of uncertain inputs to the system.

The field response to developed optimal management strategies might deviate from the prescribed values by linked S-O based groundwater management solutions, due to the uncertainties resulting from the poor characterization of the groundwater system and field scale implementation deviation. Therefore, for any groundwater management strategy to be effective, a properly designed groundwater quality monitoring network is essential, in order to evaluate the field level compliance of the prescribed groundwater management strategy. Sreekanth and Datta (2014a) proposed an adaptive management approach in which sequential revision of optimal management strategies was performed by incorporating field level compliance monitoring information from a designed monitoring network. The S-O technique, design of optimal monitoring network, and incorporating feedback information to the original simulation model are the three major components of their developed methodology. Sequential modification of the strategies was found to achieve better compliance, while optimally satisfying the salinity requirements. In order to minimize the redundancy in monitoring data, monitoring sites were situated at locations where uncertainty in the salinity concentration value is highest and the correlation between the concentrations of the monitored locations is lowest. Monitoring information from a large number of points ensures better characterization of the system, and thereby generates more accurate solutions for groundwater management by reducing the uncertainty. However, it is not practical to install a large number of monitoring wells due to budgetary constraints. Compliance monitoring at the field level is associated with the collection of relevant data that is time consuming and costly. Therefore, field scale implementation of the proper monitoring strategy is always constrained by the availability of funds (Dhar and Patil, 2011).

Previous literature on monitoring network design was based on different issues such as addressing multiple time steps (Chadalavada and Datta, 2008; Dhar and Datta, 2007; Mugunthan and Shoemaker, 2004), minimization of redundancy (Dhar and Datta, 2010; Nunes et al., 2004), multi-objective formulations (Kollat and Reed, 2007; Reed and Minsker, 2004; Sreekanth and Datta, 2014a), uncertainty (Meyer and Brill Jr., 1988), optimal location of monitoring wells (Bashi-Azghadi and Kerachian, 2010), vulnerability mapping and geo statistics (Baalousha, 2010). Earlier studies on monitoring network design in groundwater contamination emphasized minimum well density (Grabow et al., 1993), initial detection of contamination (Meyer et al., 1994), and cost-effective groundwater monitoring design (Reed et al., 2000). However, in problems related to contamination plume movement, most of the optimization techniques for monitoring network design were based on an implicit objective of minimizing the monitoring cost (Prakash and Datta, 2014).

## 2.10 Research motivation

A review of the existing literature on saltwater intrusion management in coastal aquifers revealed the possibility of developing computationally efficient coupled S-O based methodologies. More accurate emulators of the coupled flow and solute transport processes of coastal aquifers can be introduced in the coupled S-O approach for developing multiple objective saltwater intrusion management models. It appears from the literature that the existing meta-models in saltwater intrusion management model have certain limitations. There is scope for improving their prediction capability by proposing more accurate alternatives to the existing meta-models.

Groundwater flow and transport processes in coastal aquifers are associated with uncertainties in different model parameters. Previous studies of saltwater intrusion management problems considered uncertainties in hydraulic conductivity and aquifer recharge. However, other model parameters such as bulk density and compressibility of aquifer material are also considered to be sources of uncertainty that need to be addressed in saltwater intrusion modelling. Moreover, there is scope for extension of the meta-model based couple S-O approach under groundwater parameter uncertainties in which the parameter uncertainties are explicitly incorporated in the meta-models. There is also possibility of extending the ensemble meta-modelling approach to the uncertainty based weighted average ensemble approach.

Saltwater intrusion in coastal aquifers is 3-D and multilevel monitoring is expected for efficient management of coastal aquifers (Dhar and Datta, 2009b). Therefore, monitoring network design is a 3-D problem and dimensionality reduction would result in a considerable amount of information loss. The management strategy proposed by Dhar and Datta (2009b) was not adaptive in nature, i.e. feedback information was not used to modify the optimal pumping strategies sequentially for the future time horizons. Later, Sreekanth and Datta (2014a) proposed a 2-D compliance monitoring network design with feedback information included to manage saltwater intrusion in coastal aquifers. The present study proposes a 3-D sequential compliance monitoring network design for the adaptive management of saltwater intrusion in coastal aquifers.

The current research aims to explore these concepts, which have been applied to some other domains of water resources problems, but have not yet been tested in the field for saltwater intrusion management problems in coastal aquifers. Finally, the meta-model based coupled S-O approach needs to be applied in a real life regional scale coastal aquifer system for prescribing optimal groundwater extraction patterns in order to control saltwater intrusion.

The next chapter discusses the plausible benefits of utilizing barrier wells and fresh water recharge wells as part of alternative coastal aquifer management strategy for control of saltwater intrusion.

# Chapter 3: Comparative evaluation of barrier extraction and freshwater recharge wells to control saltwater intrusion in coastal aquifers

This chapter discusses the utility and benefits of using a set of barrier extraction well along the sea face boundary and incorporating fresh water recharge at selected locations in the aquifer.

## 3.1 Summary

Pumping induced saltwater intrusion in coastal aquifers is a challenging problem due to the increased abstraction of groundwater resources to meet the growing demand for freshwater supplies. Sustainable beneficial water abstraction from coastal aquifers can be ensured by controlling saltwater intrusion through reversing the hydraulic gradient along the coast by using a set of barrier extraction wells. Another plausible way of controlling saltwater intrusion is to create a freshwater lens near the shoreline by utilizing a set of freshwater recharge wells. In this chapter, the comparative efficiencies of a set of barrier extraction and freshwater recharge wells are presented as a measure of salinity control in coastal aquifers. The simulation of density reliant coupled flow and solute transport processes was performed for an illustrative 3-D multilayered coastal aquifer system. The results obtained demonstrated the superiority of barrier extraction wells over recharge wells in a shorter simulation time for controlling saltwater intrusion of the illustrative coastal aquifer system.

## 3.2 Density dependent coupled flow and solute transport numerical simulation model

A 3-D density reliant coupled transient flow and solute transport model, FEMWATER (Lin et al., 1997) was used to simulate the aquifer processes under transient flow and salt transport conditions. The 3-D flow and solute transport processes were solved by using the following equations (Lin et al., 1997)

$$\frac{\rho}{\rho_0} F \frac{\partial h}{\partial t} = \nabla . \left[ K . \left( \nabla h + \frac{\rho}{\rho_0} \nabla z \right) \right] + \frac{\rho}{\rho^*} q \tag{3.1}$$

$$F = \alpha' \frac{\theta}{n} + \beta' \theta + n \frac{dS}{dh} \tag{3.2}$$

where, $F$ = storage coefficient, $h$ = pressure head, $K$ = hydraulic conductivity tensor, $z$ = potential head, $q$ = a source or a sink, $\rho$ = water density at chemical

concentration $C$, $\rho_0$ = referenced water density at zero chemical concentration, $\rho^*$= density of injection fluid or that of the withdrawn water, $\theta$ = moisture content, $\alpha'$ = modified compressibility of water, $\beta'$ = modified compressibility of the medium, $n$ = porosity, $S$ = saturation.

The hydraulic conductivity tensor, $K$ is represented by

$$K = \frac{\rho g}{\mu} k = \frac{(\rho/\rho_0)}{\mu/\mu_0} \frac{\rho_0 g}{\mu_0} k_s k_r = \frac{\rho/\rho_0}{\mu/\mu_0} k_{so} k_r \tag{3.3}$$

where, $\mu$ = dynamic viscosity of water at chemical concentration $C$, $\mu_0$ = reference dynamic viscosity at zero chemical concentration, $k_s$ = saturated permeability tensor, $k_r$ = relative permeability or relative hydraulic conductivity, $k_{so}$ = referenced saturated conductivity tensor.

Both the dynamic viscosity and density of water vary with the chemical concentration. They take the form of the following two equations

$$\frac{\rho}{\rho_0} = a_1 + a_2 C + a_3 C^2 + a_4 C^3 \tag{3.4}$$

$$\frac{\mu}{\mu_0} = a_5 + a_6 C + a_7 C^2 + a_8 C^3 \tag{3.5}$$

where, $a_1, a_{2,\dots,} a_8$ indicates the coefficients that defines the dependence of density and viscosity of water on chemical concentration; and $C$ represents the chemical concentration.

The Darcy velocity term $V$ is given by the following equation

$$V = -K.\left(\frac{\rho_0}{\rho}\nabla h + \nabla z\right) \tag{3.6}$$

The 3-D solute transport equation is expressed as

$$\theta \frac{\partial C}{\partial t} + \rho_b \frac{\partial S}{\partial t} + V.\nabla \mathrm{C} - \nabla.(\theta D.\nabla \mathrm{C})$$

$$= -\left(\alpha' \frac{\partial h}{\partial t} + \lambda\right)(\theta C + \rho_b S) - (\theta K_w C + \rho_b K_s S) \tag{3.7}$$

$$+m - \frac{\rho^*}{\rho} qC + \left(F \frac{\partial h}{\partial t} + \frac{\rho_0}{\rho} V.\nabla\left(\frac{\rho}{\rho_0}\right) - \frac{\partial C}{\partial t}\right) C$$

where, $\rho_b$ = bulk density of the medium, $C$ = material concentration in aqueous phase, $S$ = material concentration in adsorbed phase, $t$ = time, $V$ = discharge, $\nabla$ = del operator,

$D$ = Dispersion coefficient tensor, $\lambda$ = decay constant, $M = qC_m$ = artificial mass rate, $q$ = source rate of water, $C_m$ = material concentration in the source, $K_w$ = first order biodegradation rate constant through dissolved phase, $K_s$ = first order biodegradation rate through adsorbed phase, $K_d$ = distribution coefficient.

The dispersion coefficient tensor $D$ in Equation (3.7) is expressed as

$$\theta D = a_T|V|\delta + (a_L - a_T)\frac{VV}{|V|} + a_m\theta\tau\delta \tag{3.8}$$

where, $|V|$ = magnitude of $V$, $\delta$ = Kronecker delta tensor, $a_T$ = lateral dispersivity, $a_L$ = longitudinal dispersivity, $a_m$ = molecular diffusion coefficient, and $\tau$ = tortuosity.

## 3.3 Application of the methodology in an illustrative coastal aquifer study area

### 3.3.1 Description of the study area

A multilayered unconfined coastal aquifer system similar to one developed in Roy and Datta (2017a), as shown in Figure 3.1, was used to evaluate the performance of the utilization of barrier extraction and freshwater recharge wells. The areal extent and the total thickness of the illustrative study area were 4.35 km$^2$ and 80 m, respectively. The total thickness of the aquifer was divided into four distinct layers of materials. Each of the material layers was 20 m thick, and the aquifer material within each layer was assumed homogeneous. The ratio of horizontal hydraulic conductivity values in the X- and Y-direction $(K_x/K_y)$ was taken as 2.0. The value of vertical hydraulic conductivity in the Z-direction $(K_z)$ was taken as one tenth of the values of horizontal hydraulic conductivity in the X-direction. The southern and eastern sides of the study area were surrounded by the coastal boundary and a river, respectively. An initial head of 0 m was assigned to both ends of the seaside boundary, whereas the upstream end of the rived had an initial head of 1 m, which was allowed to vary along the stream. The western side of the aquifer was considered as no-flow boundary. The coastal boundary was assigned with a constant concentration of 35 kg/m$^3$. The top of the aquifer was subjected to a natural recharge of 0.000137 m/d, distributed uniformly over the entire top phreatic surface of the aquifer.

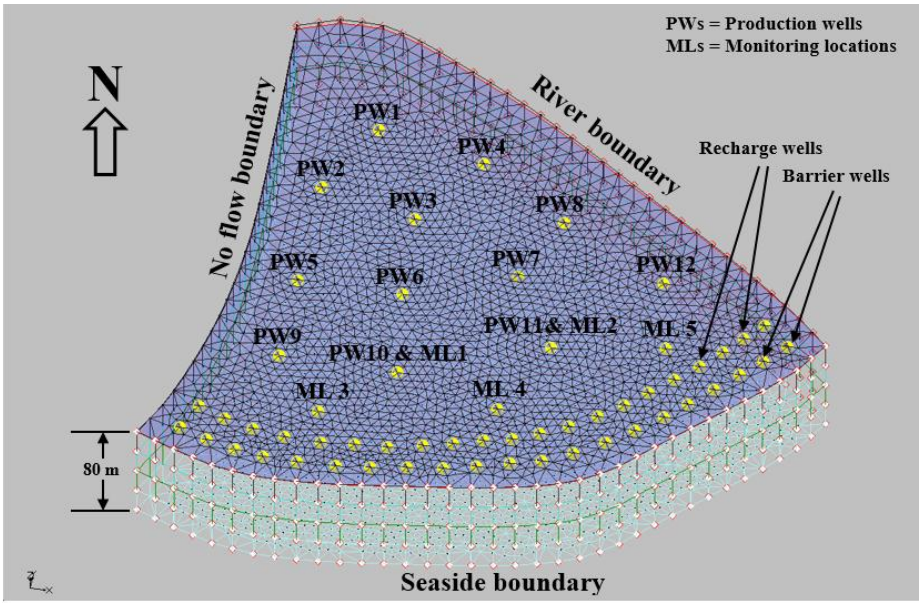


**Figure 3.1 Three dimensional view of the illustrative study area**

Figure 3.1 shows the 3-D view of the study area with finite element meshes and the locations of the wells. The study considered 12 production wells (PW1-PW12) to abstract water for beneficial purposes, 21 recharge wells (RW1-RW21) and 21 barrier extraction wells (BW1-BW21), placed near the coast to create a hydraulic barrier for controlling saltwater intrusion. The approximate distance between two adjacent wells was kept to 150 m, for both recharge and barrier extraction wells, to ensure a complete hydraulic barrier along the coast. Well screens were located at the second and third vertical layers of the aquifer. The entire study area was discretised into prism-shaped finite elements of 90 m. A finer mesh of 45 m was used near the wells. The total simulation period of three years (1095 days) was divided into 219 uniform time steps of five days each. Salinity concentrations were monitored at five monitoring locations (ML) at the end of the simulation period of three years. The recharge water for artificial recharge was assumed to have a salinity concentration of 0.1 kg/m$^3$. Simulation parameters are shown in Table 3.1.

**Table 3.1 Aquifer parameters**

| Parameters | Units | Material layers | | | |
|---|---|---|---|---|---|
| | | Layer 1 | Layer 2 | Layer 3 | Layer 4 |
| Hydraulic conductivity in x-direction | m/d | 5 | 10 | 15 | 3 |
| Hydraulic conductivity in y-direction | m/d | 2.5 | 5 | 7.5 | 1.5 |
| Hydraulic conductivity in z-direction | m/d | 0.5 | 1.0 | 1.5 | 0.3 |
| Compressibility | $md^2/kg$ | $1.33 \times 10^{-15}$ | $1.33 \times 10^{-17}$ | $1.33 \times 10^{-17}$ | $1.33 \times 10^{-16}$ |
| Bulk density | $kg/m^3$ | 1650 | 1600 | 1550 | 1700 |
| Porosity | - | 0.46 | 0.41 | 0.43 | 0.38 |
| Longitudinal dispersivity | m | 80 | 80 | 80 | 80 |
| Lateral dispersivity | m | 35 | 35 | 35 | 35 |
| Molecular diffusion coefficient | $m^2/d$ | 0.69 | 0.69 | 0.69 | 0.69 |
| Density reference ratio | - | $7.14 \times 10^{-7}$ | $7.14 \times 10^{-7}$ | $7.14 \times 10^{-7}$ | $7.14 \times 10^{-7}$ |
| Recharge | m/d | 0.000137 | - | - | - |

*Recharge is distributed uniformly over the first layer of the study area

### 3.3.2 Results and Discussion

A sensitivity analysis was conducted to evaluate the effects of barrier extraction wells, artificial recharge via injection wells, and a combination of the both on saltwater intrusion processes in coastal aquifers. This sensitivity analysis was performed to identify which options had the most influential effect on saltwater intrusion processes. These were quantified by observing changes in heads and saltwater concentrations over time at specified MLs.

The development of pressure heads with respect to utilization of production wells (PWs), barrier extraction wells (BWs), and freshwater recharge wells (RWs) and some possibilities of their combinations is presented in Figure 3.2. It is observed from Figure 3.2 that the initial pressure head was declined due to water extracted from PWs, which was declined further when BWs were combined with PWs. Recharging freshwater through RWs combined with water extraction from PWs built up the pressure head again. The pressure head decreased again when BWs were added to PWs + RWs. This fluctuation of the pressure head confirms the effects of both of these wells, reflecting a balance between water coming in and water going out of the system. It is obvious from Figure 3.2 that recharging freshwater

through a set of RWs was associated with building up the head, however, the developed head might not be sufficient to reduce salinity concentrations at the specified MLs for the specified simulation period of three years.

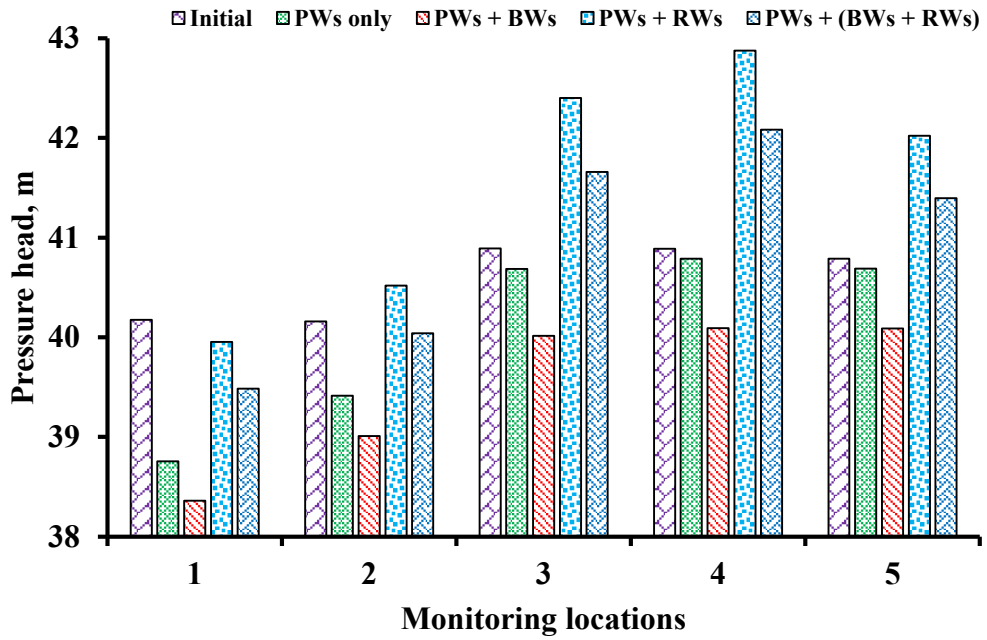


**Figure 3.2 Comparison of heads at different monitoring locations PWs = Production wells, BWs = Barrier extraction wells, RWs = Freshwater recharge wells**

Salinity concentrations at five MLs at the end of the simulation period for a particular pumping pattern are presented in Figure 3.3. It is noted that MLs were placed at different locations to ensure the variation of salinity concentrations at these locations. Figure 3.3 shows that groundwater extraction had effects on increased salinity concentrations in the aquifer when compared to initial states of the aquifer before pumping started. An increase in salinity concentrations resulting from water extraction from PWs could be reduced by allowing water extraction from a set of BWs. In contrast, injecting water through a set of RWs did not help reduce salinity concentrations at specified MLs, rather the salinity concentrations increased from a combined operation of PW + RW when compared to PW pumping alone. A plausible explanation of this may be that during the three years of the simulation period an adequate freshwater head did not develop to prevent migration of saltwater plume. However, a longer simulation period was not considered for this study because the primary objective was to develop a saltwater intrusion management model. A longer management period may be practically meaningless due to the changing physical processes of the aquifer system. Combined utilization of BWs and RWs is also presented in this study for comparison purposes. It is observed from Figure 3.3 that this combined implementation partly compensated for the ill effects of using RWs alone. Therefore, BWs

were proved to be more effective saltwater intrusion control measures than RWs and BWs +RWs, at least for this example problem. The use of BWs over RWs is also well justified when considering the shortage of global freshwater supplies.

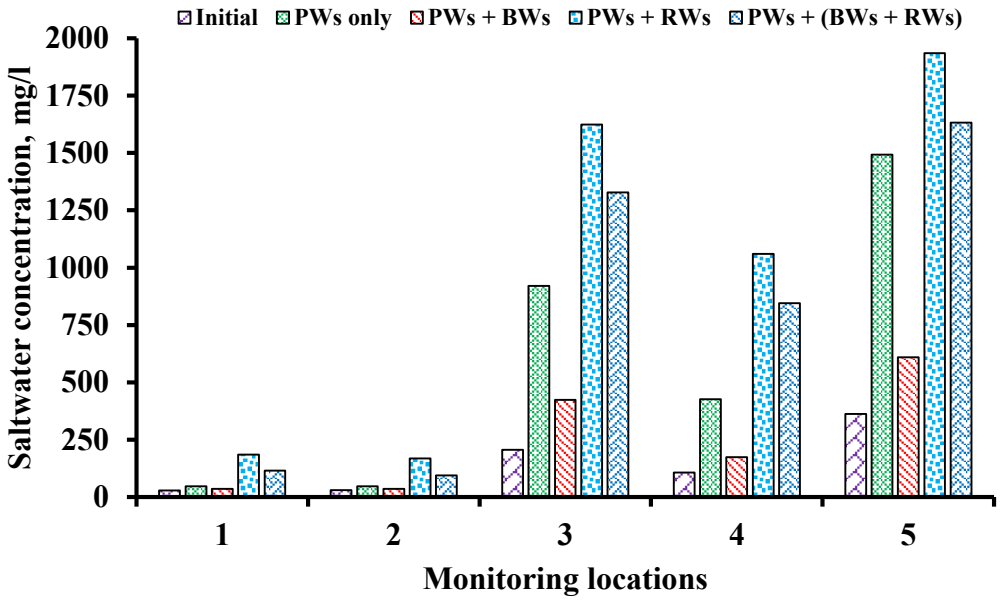


**Figure 3.3 Comparison of salinity concentrations at different monitoring locations PWs = Production wells, BWs = Barrier extraction wells, RWs = Freshwater recharge wells**

## 3.4 Conclusions

This chapter presented the relative effectiveness of a set of barrier extraction wells, freshwater recharge wells, and a combination of these wells in reducing the extent of saltwater intrusion processes in coastal aquifers. Results indicated that barrier extraction wells were more effective in controlling salinity intrusions in the illustrative coastal aquifer system for the specified simulation period of three years. While increase in the pressure head was observed in the specified MLs when freshwater was recharged through recharge wells, the developed head might not be sufficient to create a freshwater lens to prevent saltwater intrusion, especially within this simulation period. The physical processes of a dynamic coastal aquifer system might change within a very short period, and as the primary aim of simulating the aquifer processes was to develop management policies for optimized pumping from the aquifer, a simulation period of three years was considered optimal for illustrative purposes. The development of various trained meta-models as approximate simulators of the numerical flow and transport simulation model is discussed in the next chapter.

# Chapter 4: Development of artificial intelligence based meta-models for saltwater intrusion processes prediction in coastal aquifers

Partial contents of this chapter have been published and copyrighted, as outlined below:

Roy, D. K., & Datta, B. (2017a). Fuzzy c-mean clustering based inference system for saltwater intrusion processes prediction in coastal aquifers. *Water Resources Management, 31*(1), 355-376.

Roy, D. K., & Datta, B. (2018a). A surrogate based multi-objective management model to control saltwater intrusion in multi-layered coastal aquifer systems. *Civil Engineering and Environmental Systems*, 1-26.

This chapter discusses the development of various trained meta-models as approximate simulators of the numerical flow and transport simulation model.

## 4.1 Summary

This chapter presents development of meta-models to approximate density reliant coupled flow and solute transport processes in multilayered coastal aquifer systems. Meta-models are used to represent the nonlinear mapping of the input-output relationships of complex physical processes. The reliability of meta-models depends on how accurately they capture the complex functional relationships between the inputs and outputs of a physical system. Simulation of complicated physical processes of a coastal aquifer to an acceptable degree of accuracy using a suitable meta-model can be useful to achieve computational efficiency in modelling. Two meta-models: Sugeno type Fuzzy Inference System (FIS) and Multivariate Adaptive Regression Spline (MARS), were developed to predict salinity concentrations at specified MLs with respect to groundwater extraction. FIS incorporates human reasoning in analysing different combinations of antecedent transient pumping values and returns concentration values at different MLs. On the other hand, MARS adaptively selects the most influential input variables in determining the outputs. In this chapter, these two meta-models are proposed in order to address some of the limitations of the existing meta-models for saltwater intrusion processes in coastal aquifers.

An illustrative multilayered coastal aquifer system was used to evaluate the performance of the developed meta-models. As introduced in chapter 3, barrier extraction wells have some advantages over freshwater recharge wells as salinity control measures in coastal aquifers. Barrier wells create a hydraulic barrier along the coast to control saltwater intrusion. A 3-D density reliant coupled flow and solute transport numerical simulation

model, FEMWATER, was used to simulate the aquifer processes for generating required input-output training patterns of the meta-models. Trained and validated meta-models were utilized as approximate simulators of the coupled flow and salt transport processes. Performance evaluation results indicated that the developed FIS and MARS based meta-models can be applied to approximate complex physical processes of multilayered coastal aquifers and can be suitable for incorporation in a coupled S-O technique in order to develop optimum pumping strategy. Both FIS and MARS meta-models have unique advantages. For example, FISs are suitable for multiple output problems and MARS models are adaptive in nature that have very high computational efficiency.

## 4.2 Development of the FIS meta-model

### 4.2.1 Generation of input-output training patterns

The simulation model was used with randomized inputs (transient pumping) to generate sets of input-output arrays for training a FIS-based meta-model. The random transient pumping set, obtained from a uniform sampling distribution using Latin Hypercube Sampling (LHS) (Pebesma and Heuvelink, 1999) was used as input to the simulation model to obtain the corresponding salinity concentrations at specified MLs. Inputs to the simulation model and outputs from the simulation model in a single run produce one input-output pattern. A number of such input-output patterns were generated by multiple runs of the simulation model.

### 4.2.2 Developed FIS meta-model

The proposed FIS meta-model is able to predict salinity concentrations at multiple locations for multiple output problems. Although ANN has this capability, the superiority of fuzzy logic based approaches such as ANFIS over ANN is well documented in the literature (Emamgholizadeh et al., 2014; Gong et al., 2016). However, ANFIS is unable to predict multiple outputs, and separate ANFIS structures need to be constructed for multiple output problems. To overcome this limitation, a FIS based meta-model was proposed to predict multiple outputs in a single global structure. In addition, the capability of a fuzzy logic approach in order to incorporate vague, imprecise, and fuzzy data eliminates the need to normalize the data before being presented to the FIS.

FIS meta-model was trained from a dataset with 80 input and 25 output variables. The number of input and output variables was very large in order to develop a FIS using conventional approaches e.g. grid partitioning approach. Therefore, we adopted the FCM clustering algorithm to divide the input space into identical clusters that significantly reduced the number of rules in FIS generation. Clustering was performed from a known number of

clusters by minimizing the fuzzy overlap between clusters. FCM minimized the following objective function during clustering to avoid fuzzy overlap

$$J_m = \sum_{i=1}^{D}\sum_{j=1}^{N} \mu_{ij}^{m} \left\| x_i - c_j \right\|^2 \tag{4.1}$$

where, $D$ = number of data points, $N$ = number of clusters, $m$ = fuzzy partition matrix exponent, $x_i = i^{th}$ data points, $c_j$ = centre of the $j^{th}$ cluster, $\mu_{ij}$ = degree of membership of $x_i$ in the $j^{th}$ cluster.

FIS is very effective tool for modelling nonlinear systems (Sugeno and Yasukawa, 1993; Takagi and Sugeno, 1985). The FIS structure developed and utilized was a Sugeno type FIS with 80 input variables (transient pumping values), and 25 output variables (salinity concentration values). In the Sugeno system, each rule linearly depends on the input variables. Therefore, the Sugeno method is ideal for acting as an interpolating supervisor of multiple linear controllers. This capability makes the developed FIS model ideal for highly nonlinear systems that are dynamic in nature (MATLAB, 2017d). To evaluate the utility of the FIS to predict salinity concentrations at different MLs, 15 rules representing 15 clusters were chosen as sufficient. The number of clusters was selected based on trial and error. The optimal number of clusters was chosen based on Root Mean Square Error (RMSE) values between the actual and predicted saltwater concentrations in the validation dataset. Number of clusters that minimized the RMSE value was used to develop the FIS model. Table 4.1 summarizes the components of the developed FIS. Figure 4.1 shows the design architecture of the developed FIS. In Figure 4.1, the inputs (1-80) were crisp (non-fuzzy) numbers with a range between 0 and 1300 $m^3$/day. The second pane is the heart of the FIS, in which crisp inputs were fuzzified, number of rules was specified based on the number of clusters, all rules were evaluated, and finally the combined rules were defuzzified. The last pane represents the corresponding crisp output values (1-25) generated by the FIS.

In the developed FIS, the inputs to the FIS model were equivalent to 80 (16 wells × 5 time steps). The FIS model output corresponds to the concentrations at the selected MLs for all time steps. Hence, there were 25 (5 monitoring locations × 5 time steps) outputs from the FIS. The inputs and outputs to the FIS can be written in the form

$$\begin{aligned} &Concentration\ at\ monitoring\ locations \\ &= f(Water\ extraction\ from\ production\ and\ barrier\ extraction\ wells) \end{aligned} \tag{4.2}$$

**Table 4.1 Properties of the developed FIS**

| Parameters | Specifications | Parameters | Specifications |
|---|---|---|---|
| Name of FIS | Sugeno8025 | And method | 'prod' |
| Type of FIS | 'Sugeno' | Or Method | 'probor' |
| Inputs/Outputs | [80 25] | Implementation Method | 'prod' |
| No. of Input MFs | 80 | Aggregation Method | 'sum' |
| No. of Output MFs | 25 | Defuzzification Method | 'wtaver' |
| Number of Clusters | 15 | Input MF type | Gaussian |
| Number of rules | 15 | Output MF type | Linear |
| Rule Weight | 1 | | |
| Rule Connection | 1 | | |

The expression in Equation 4.2 can be mathematically expressed as (after Bhattacharjya and Datta (2005))

$$\left\{C_i^j, i = M_1, \dots, M_5, j = 1, \dots, 5\right\} = f\begin{Bmatrix} P_i^j, i = 1, \dots, 11, j = 1, \dots, 5, \\ B_i^j, i = 1, \dots, 5, j = 1, \dots, 5 \end{Bmatrix} \quad (4.3)$$

where, $i$ and $j$ are the indices for locations and time steps, respectively; $C_i^j$ represents the salinity concentrations at the $i^{th}$ monitoring locations at $j^{th}$ time steps; $P_i^j$ is the abstraction from the $i^{th}$ pumping wells at $j^{th}$ time steps; $B_i^j$ is the abstraction from the $i^{th}$ barrier extraction wells at $j^{th}$ time steps. Equation 4.3 represents the approximate simulation model based on the FIS meta-model.

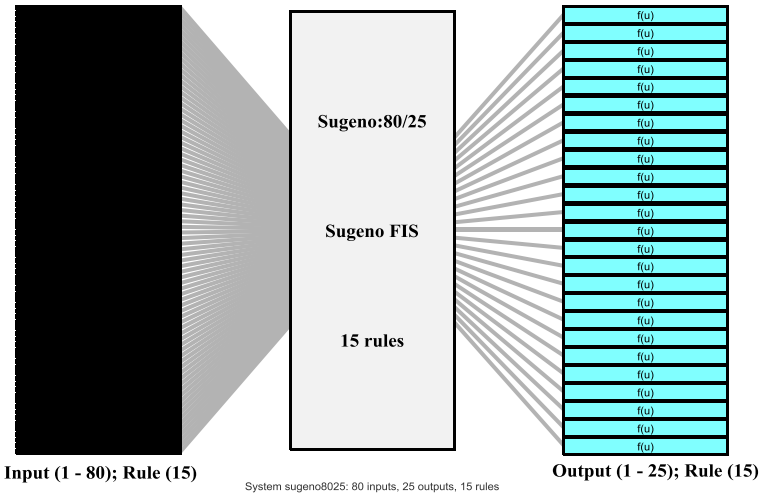


**Figure 4.1 Architecture of the developed FIS**

### 4.2.3 Training procedure

In order to train the FIS meta-model, sufficient datasets were generated using the numerical simulation model for different randomized pumping scenarios to make sure the data contained the necessary representative range and features. A total of 1500 input-output patterns were generated. The entire dataset was divided randomly into training, validation, and prediction sets, using random sampling without replacement (Hastie et al., 2008). Seventy percent (70%) of the dataset was used for training, 15% of the dataset was used to validate the generated FIS structure, while the rest (15%) was used for evaluating the prediction capability of the developed FIS. The validation and prediction datasets were sufficiently distinct from the training data set. However, training, validation, and prediction input dataset were all within the range of the specified pumping values (0 to 1300 $m^3$/day). Clustering termination condition was set by specifying the difference in objective function values (Equation 4.1) between two successive iterations to $1\times10^{-5}$. To improve clustering results and to reduce fuzzy overlap during clustering, the fuzzy partition matrix exponent was set to 1.1. As per the imposed stopping criterion, termination of the clustering process occurred when the difference between the objective function (Equation 4.1) values in two successive iterations was smaller than the pre-set minimum value.

### 4.2.4 Performance measures

The performance of the trained FIS meta-model was evaluated using the following indices

Mean Squared Error (MSE)

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(C_n^O - C_n^P)^2 \tag{4.4}$$

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(C_n^O - C_n^P)^2} \tag{4.5}$$

Mean Absolute Percentage Relative Error (MAPRE)

$$MAPRE = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{C_n^O - C_n^P}{C_n^O}\right| \times 100 \tag{4.6}$$

Nash–Sutcliffe Efficiency Coefficient (NS)

$$NS = 1 - \frac{\sum_{n=1}^{N}(C_n^O - C_n^P)^2}{\sum_{n=1}^{N}(C_n^O - C^O)^2} \tag{4.7}$$

Correlation Coefficient (R)

$$R = \frac{\sum_{n=1}^{N}(C_n^O - C^O) \times (C_n^P - C^P)}{\sqrt{\sum_{n=1}^{N}(C_n^O - C^O)^2} \times \sqrt{\sum_{n=1}^{N}(C_n^P - C^P)^2}} \tag{4.8}$$

Threshold Statistics (TS)

$$TS = \frac{N_a}{N} \times 100 \tag{4.9}$$

where, $C_n^O$ and $C_n^P$ are the observed and predicted values of seawater concentrations, respectively; $C^O$ and $C^P$ are the mean values of the observed and predicted saltwater concentrations, respectively, and $N$ denotes number of data points; $N_a$ is the number of data points whose absolute percentage RE value is less than a specified threshold value $a$%.

### 4.2.5 Performance evaluation using an illustrative multilayered coastal aquifer study area

The developed FIS model's performance was evaluated using a small illustrative coastal aquifer system as shown in Figure 4.2. The study area for illustrative purposes was a hypothetical coastal aquifer of 4.35 km$^2$, and was heterogeneous in nature. Aquifer heterogeneity was considered by incorporating layered zones of different hydraulic conductivities. Four distinct layers of hydraulic conductivities were adopted in the present study. Aquifer parameters and the hydraulic conductivity values used for each distinct layer are presented in Table 3.1 of chapter 3. For simplicity, stratified (layered) heterogeneity along the vertical direction was considered, and material within the same layer was considered homogeneous. The aquifer considered had a layered stratification in which heterogeneity in hydraulic conductivity was assumed significant only in the vertical direction.

The aquifer was surrounded by the sea and a river system on the southern and eastern sides, respectively. The length of the coastline was 3.2 km, whereas the river had a length of 2.9 km. The total thickness of the aquifer was 80 m, consisting of four distinct layers of materials, each having 20 m thickness. A constant head value of 0 m was assigned at both ends of the seaside boundary, while the seaside concentration was kept constant at 35 kg/m$^3$. The upstream end of the river had a specified head of 1 m that varied linearly along the stream from the specified value at the upstream end to zero at the seaside. The compressibility and dynamic viscosity of water were taken as 6.69796 × 10$^{-20}$ md$^2$/kg and 131.328 kg/md,

respectively. The western side and bottom of the model domain were considered as no flow boundaries. A groundwater recharge rate of 0.05 m/year was assumed to be distributed uniformly over the surface area of the entire study area. Groundwater recharge in the form of base flow from the riverbed was assumed to be negligible. The entire model domain was discretised into triangular finite elements of 190 m size. Element size near the production and barrier wells was set to 95 m. The well screens were located in the second and third layers of the aquifer. Groundwater was extracted from 11 potential pumping locations represented by PW1 to PW11 in Figure 4.2. To create a hydraulic barrier along the coastline to prevent saltwater intrusion, 5 barrier wells, named BW1 to BW5 in Figure 4.2, were also considered. When pumping, a drop in the hydraulic head occurred near these barrier wells. To maintain equilibrium, water from the aquifer moved towards the sea, leading to a decreased saltwater intrusion. The salinity concentrations after each time step were monitored at 5 MLs, represented by MP1 to MP5. The pumping locations PW6, PW9, and PW10 were also used as MLs. The two other MLs, denoted as MP4 and MP5, were located close to the seaside boundary.

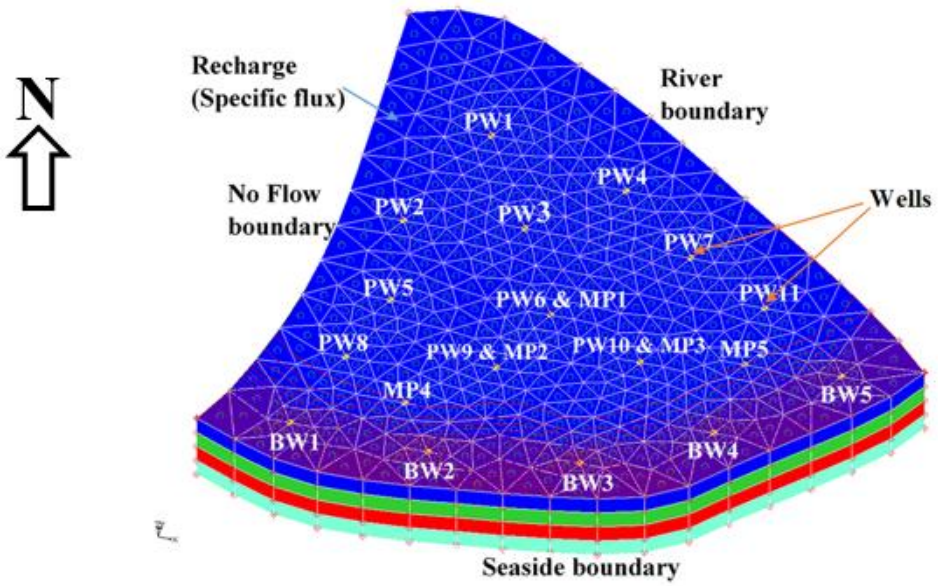


**Figure 4.2 Illustrative study area (three-dimensional mesh and model boundaries)**

The 3-D transient simulation was initiated from an initially prevailing steady state condition of the aquifer. To obtain the initial steady state, a constant pumping of 200 $m^3$/day from each of the 11 production wells for a period of 5 years was specified. The barrier well pumping was not used for this purpose. After 5 years of pumping, the observed heads at different nodes of the model domain became constant. These constant heads and resulting concentration values at different nodes of the model domain were set as initial conditions for simulating the aquifer processes. Figure 4.3 illustrates the initial concentration values, and concentration contours at the end of the simulation period (5th time step). It can be seen from

the concentration contours in Figure 4.3 that the saltwater front moves further inland compared to the initial condition, due to the pumping stress applied at different time steps. However, the movement of the saltwater front was restricted by introducing the barrier well pumping near the coastline. Various random combinations of production and barrier well pumping values were assigned to the well fields of the aquifer. Therefore, the developed FIS-based meta-model could be easily utilized for deriving optimal spatial and temporal pumping strategies, considering the combined effect of beneficial withdrawal for water supply and barrier well pumping.

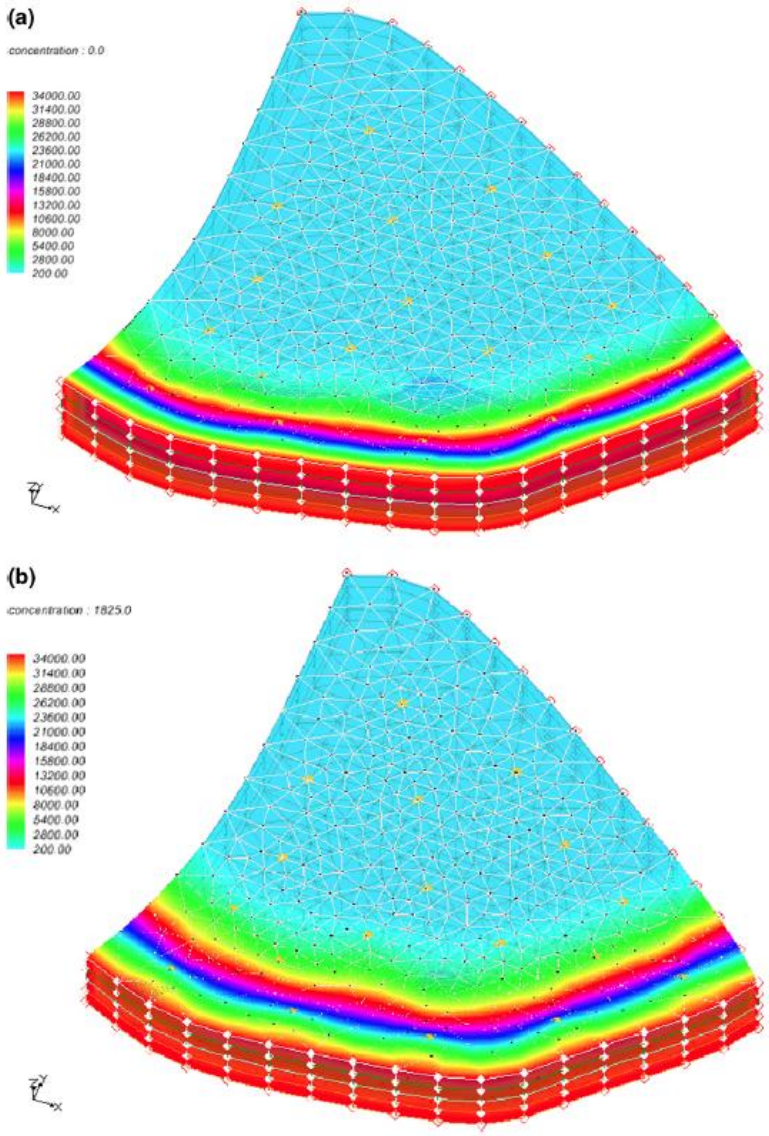


**Figure 4.3 Concentration contours; (a) initial concentration, (b) concentration at the end of simulation period (time step 5)**

4.2.6 Performance of the developed FIS meta-model based on statistical indices

The capability of the developed FIS to predict saltwater concentrations at different MLs of a heterogeneous coastal aquifer was evaluated based on its performance with training,

validation, and prediction datasets. It was observed that the Standard Deviation (SD) and Coefficient of Variation (CV) of the observed and predicted concentration values in the prediction dataset did not vary substantially. The maximum absolute difference in SD between the actual and prediction dataset was 3.804, whereas the maximum absolute difference in CV was found to be 0.446 %. The model's performance with regard to RMSE, MAPRE, NS, and R values on training, validation, and prediction patterns is summarized in Table 4.2. It is apparent that all the performance measures were similar for training and validation datasets. The prediction capability of the developed FIS was good, based on MAPRE, NS, and R values. The highest MAPRE value obtained in the prediction set was 1.65 %, and the lowest obtained value was 0.04 %, while the average value of MAPRE was 0.76 %. Likewise, the highest, the lowest, and the average value of NS in the prediction dataset were 1, 0.95, and 0.98, respectively. Model performance was also satisfactory, based on R values. The R values obtained were very close to 1 with a maximum, minimum and average value of 99.98, 97.74, and 99.20, respectively. It was found that the statistical inference index values did not differ noticeably between the training and validation datasets (Table 4.2). The developed FIS provided smaller RMSE and MAPRE values, as well as higher R and NS values for both training and validation datasets. The highest difference in RMSE between the training and validation datasets was 1.78, and this value was observed at monitoring location ML2 for time step 5. The maximum values of the difference in MAPRE, NS, and R were 0.26, 0.02, and 0.60 respectively. These differences in maximum values were observed at monitoring location ML2 at time step 1.

In the present study, the original values of the concentration without normalization were used for the purpose of training and validating the developed FIS. As normalized data were not used, the value of MAPRE provided better evaluation of the model performance. MAPRE provides information on the distribution of errors and is a measure for testing robustness of the developed model (He et al., 2014). The MAPRE index also provides an indication about whether a model tends to overestimate or underestimate (He et al., 2014). The value of MAPRE for both training, validation, and prediction datasets for all MLs at different time steps were very small (Table 4.2). The analysis based on the MAPRE index indicated that the errors were more symmetric around zero. The developed model produced relative error values of less than 2% for all estimates. This indicates that the developed FIS model seems to perform well from the relative error viewpoint.

**Table 4.2 Performance of the FIS on training, validation, and prediction set of data**

| Time steps | MPs | Training | | | | Validation | | | | Prediction | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | RMSE | MAPRE (%) | NS | R (%) | RMSE | MAPRE (%) | NS | R (%) | RMSE | MAPRE (%) | NS | R (%) |
| | 1 | 0.03 | 0.08 | 0.95 | 97.70 | 0.03 | 0.09 | 0.95 | 97.62 | 0.03 | 0.08 | 0.96 | 97.74 |
| | 2 | 4.17 | 1.38 | 0.97 | 98.24 | 4.87 | 1.64 | 0.95 | 97.64 | 5.02 | 1.65 | 0.95 | 98.04 |
| 1 | 3 | 2.68 | 1.25 | 0.98 | 98.90 | 2.93 | 1.41 | 0.98 | 98.78 | 2.81 | 1.30 | 0.97 | 98.84 |
| | 4 | 4.00 | 0.13 | 1.00 | 99.89 | 4.29 | 0.14 | 1.00 | 99.87 | 4.08 | 0.14 | 1.00 | 99.89 |
| | 5 | 1.11 | 0.03 | 1.00 | 99.98 | 1.15 | 0.04 | 1.00 | 99.98 | 1.09 | 0.04 | 1.00 | 99.98 |
| | 1 | 0.08 | 0.22 | 0.97 | 98.50 | 0.09 | 0.25 | 0.96 | 98.10 | 0.10 | 0.26 | 0.96 | 98.38 |
| | 2 | 7.48 | 1.43 | 0.98 | 98.75 | 8.35 | 1.63 | 0.97 | 98.37 | 8.51 | 1.62 | 0.97 | 98.67 |
| 2 | 3 | 5.26 | 1.40 | 0.98 | 99.18 | 5.79 | 1.55 | 0.98 | 98.98 | 5.43 | 1.38 | 0.98 | 99.14 |
| | 4 | 7.58 | 0.18 | 1.00 | 99.86 | 7.77 | 0.18 | 1.00 | 99.87 | 7.72 | 0.18 | 1.00 | 99.86 |
| | 5 | 1.79 | 0.04 | 1.00 | 99.98 | 1.80 | 0.04 | 1.00 | 99.99 | 1.75 | 0.05 | 1.00 | 99.98 |
| | 1 | 0.19 | 0.49 | 0.97 | 98.58 | 0.19 | 0.49 | 0.97 | 98.38 | 0.24 | 0.55 | 0.97 | 98.50 |
| | 2 | 11.79 | 1.44 | 0.98 | 98.87 | 13.01 | 1.56 | 0.97 | 98.66 | 12.75 | 1.56 | 0.98 | 98.83 |
| 3 | 3 | 9.13 | 1.47 | 0.98 | 99.15 | 8.60 | 1.40 | 0.98 | 99.20 | 9.37 | 1.49 | 0.98 | 99.16 |
| | 4 | 10.59 | 0.19 | 1.00 | 99.86 | 11.23 | 0.19 | 1.00 | 99.87 | 10.76 | 0.19 | 1.00 | 99.86 |
| | 5 | 2.43 | 0.05 | 1.00 | 99.98 | 2.77 | 0.05 | 1.00 | 99.98 | 3.04 | 0.05 | 1.00 | 99.98 |
| 4 | 1 | 0.39 | 0.84 | 0.97 | 98.74 | 0.42 | 0.97 | 0.97 | 98.48 | 0.38 | 0.80 | 0.98 | 98.71 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 17.11 | 1.38 | 0.98 | 98.90 | 18.05 | 1.50 | 0.97 | 98.72 | 18.48 | 1.53 | 0.98 | 98.86 |
| | 3 | 13.44 | 1.43 | 0.98 | 99.17 | 14.17 | 1.55 | 0.98 | 99.09 | 14.73 | 1.51 | 0.98 | 99.13 |
| | 4 | 13.13 | 0.18 | 1.00 | 99.87 | 14.21 | 0.20 | 1.00 | 99.86 | 13.90 | 0.20 | 1.00 | 99.87 |
| | 5 | 3.35 | 0.05 | 1.00 | 99.98 | 3.73 | 0.06 | 1.00 | 99.98 | 3.88 | 0.06 | 1.00 | 99.98 |
| | 1 | 0.72 | 1.28 | 0.98 | 98.75 | 0.70 | 1.33 | 0.97 | 98.76 | 0.73 | 1.27 | 0.97 | 98.74 |
| | 2 | 23.08 | 1.33 | 0.98 | 98.92 | 24.86 | 1.49 | 0.97 | 98.67 | 23.96 | 1.37 | 0.98 | 98.88 |
| 5 | 3 | 18.34 | 1.36 | 0.98 | 99.18 | 19.58 | 1.44 | 0.98 | 99.10 | 20.45 | 1.51 | 0.98 | 99.14 |
| | 4 | 14.87 | 0.17 | 1.00 | 99.89 | 16.34 | 0.18 | 1.00 | 99.87 | 15.79 | 0.19 | 1.00 | 99.88 |
| | 5 | 4.06 | 0.05 | 1.00 | 99.98 | 4.51 | 0.06 | 1.00 | 99.98 | 4.58 | 0.06 | 1.00 | 99.98 |
| **Avg.** | | **7.07** | **0.72** | **0.98** | **99.23** | **7.58** | **0.76** | **0.98** | **99.11** | **7.58** | **0.76** | **0.98** | **99.20** |

*RMSE = Root Mean Squared Error, MAPRE = Mean Absolute Percentage Relative Error, R = Coefficient of Correlation

Another statistical parameter that evaluates the capability of a hydrological model to predict any system behaviour is the NS. A model's prediction capability is 100% perfect when the NS value is equal to 1. The prediction of any model is deemed acceptable when NS criterion has a value greater than or equal to 0.8 (Shu and Ouarda, 2008). In the present study, the NS values for both training, validation, and prediction datasets were greater than 0.8, which indicate that the developed FIS predictions were within the acceptable range. It was assumed that the FEMWATER simulation results were as accurate as could be obtained in real situations. The predicted concentration values were those obtained from the developed FIS meta-model. It was observed that the FIS estimates were very close to the corresponding values of actual salinity concentrations at all MLs for all time steps. The $R^2$ values for all the estimates were also very close to unity with the lowest value obtained was 0.9473 at monitoring location MP2 at time step 1. The value of $R^2$ varied between 0.9473 and 0.9997. The model showed good prediction accuracy for all MLs and for all time steps.

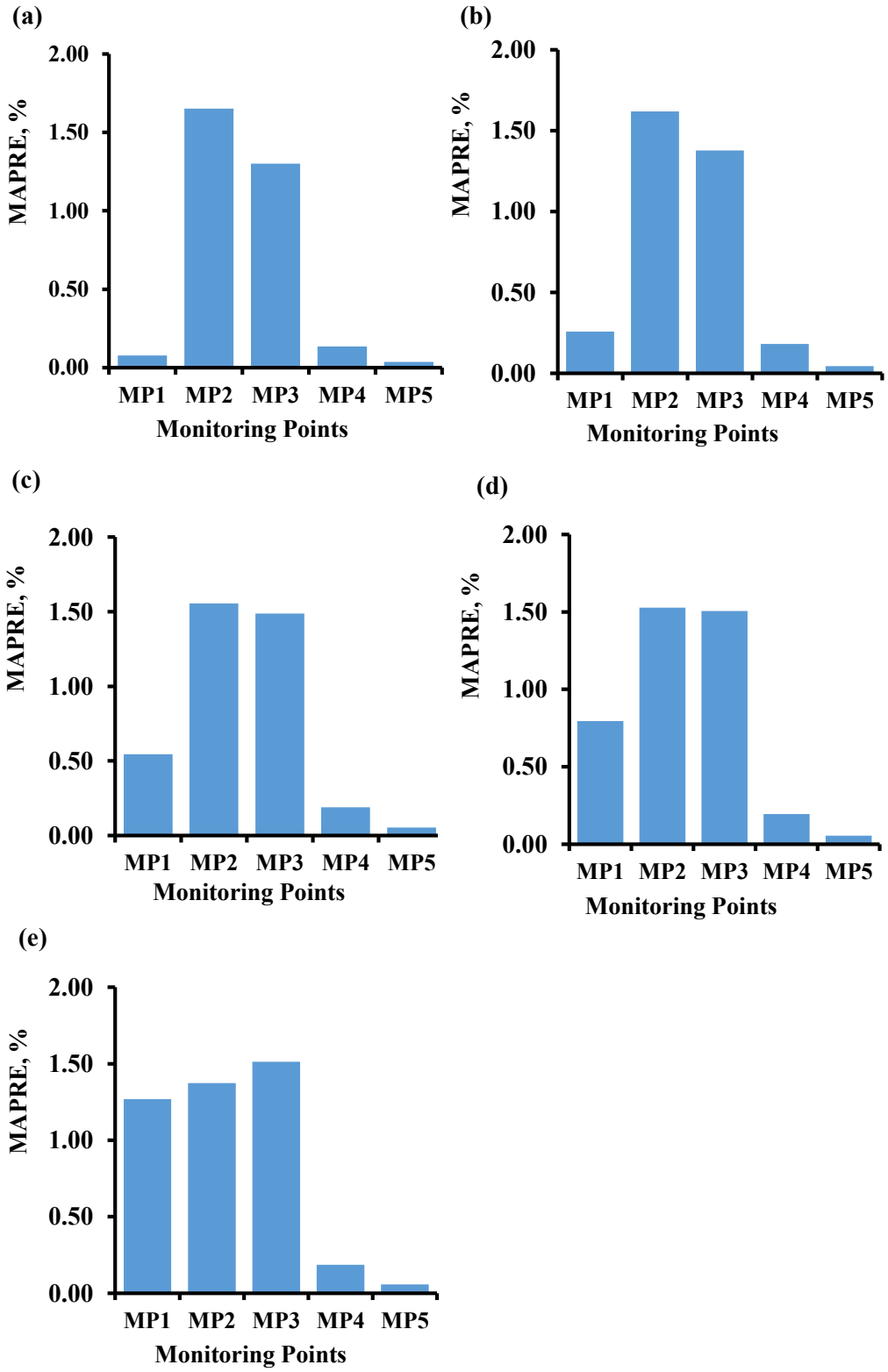


**Figure 4.4 MAPRE values between actual and predicted concentrations in the prediction dataset at different time steps for Monitoring Locations 1 – 5; (a) time step 1, (b) time step 2, (c) time step 3, (d) time step 4, (e) time step 5**

The prediction capability was slightly lower at monitoring location MP2. However, the prediction at this location was also very reliable and accurate in terms of obtained $R^2$ values. The model performance was also acceptable in terms of RMSE values. As normalized data were not used in the present study, the values of RMSE looked a bit higher. However, considering the higher concentration values at some MLs, the higher values of RMSE is justifiable. The highest value of RMSE obtained in the prediction dataset was 23.96 at monitoring location MP2 at time step 5 (Table 4.2) whereas the lowest value obtained was 0.03 at the monitoring location ML1 at time step 1 with an average value of 7.58 (Table 4.2).

Figure 4.4 shows the MAPRE values at different MLs for various time steps. The values of MAPRE were very small for all MLs. From the results of performance evaluation, it is logical to reach the conclusion that the developed FIS is able to simulate aquifer processes for the demonstrative study area representing a typical heterogeneous coastal aquifer system to a reasonable level of precision.

### 4.2.7 Generalization capability of the FIS for new datasets

The generalization capability of the developed FIS meta-model for a new set of data was evaluated by presenting a completely new set of data to the FIS model. These results should enhance the evaluation of the FIS by incorporating an entirely new set of inputs and outputs. This pumping dataset was completely different from that of the training, validation, and prediction dataset used earlier. This dataset was generated from a new LHS within the pumping range of 0-1300 $m^3$/day. A total number of 15 data points with 80 different pumping values representing 80 pumping variables were generated. The corresponding concentrations at 5 different MLs at 5 time periods were determined by utilizing the same simulation model with same initial and boundary conditions. The pumping values were presented to the FIS model to generate the outputs in terms of salinity concentrations at different MLs. Table 4.3 summarizes the statistical indices of the actual and predicted values of saltwater concentrations at different MLs and for various time steps using a completely new data set. For all MLs, larger values of R and NS as well as smaller values of RMSE and MAPRE were found. The values of R and NS for all the cases were very close to 1. The values of MAPRE for all the locations were less than 3%. This indicates the ability of the developed FIS to mimic the coupled flow as well as salt transport processes for a hypothetical illustrative coastal area with good accuracy. However, this range of possible inputs needs to be representative of the possible range of the field values for accurate prediction.

**Table 4.3 Statistical indices of the observed and predicted concentration values from a new dataset**

| Time steps | MPs | R | RMSE | MAPRE (%) | NS |
|---|---|---|---|---|---|
| | 1 | 0.987 | 0.014 | 0.033 | 0.967 |
| | 2 | 0.972 | 5.626 | 2.059 | 0.878 |
| 1 | 3 | 0.994 | 2.968 | 1.401 | 0.872 |
| | 4 | 1.000 | 1.957 | 0.069 | 0.999 |
| | 5 | 1.000 | 0.625 | 0.025 | 0.999 |
| | 1 | 0.994 | 0.042 | 0.119 | 0.987 |
| | 2 | 0.997 | 7.114 | 1.423 | 0.961 |
| 2 | 3 | 0.990 | 4.299 | 1.170 | 0.961 |
| | 4 | 1.000 | 3.733 | 0.089 | 0.999 |
| | 5 | 1.000 | 1.283 | 0.025 | 0.999 |
| | 1 | 0.998 | 0.078 | 0.236 | 0.985 |
| | 2 | 1.000 | 11.965 | 1.340 | 0.971 |
| 3 | 3 | 0.984 | 11.475 | 1.918 | 0.930 |
| | 4 | 1.000 | 5.675 | 0.102 | 0.999 |
| | 5 | 1.000 | 3.050 | 0.066 | 0.998 |
| | 1 | 0.999 | 0.117 | 0.209 | 0.994 |
| | 2 | 0.996 | 29.207 | 2.232 | 0.906 |
| 4 | 3 | 0.995 | 17.142 | 1.664 | 0.956 |
| | 4 | 1.000 | 8.456 | 0.124 | 0.999 |
| | 5 | 0.999 | 7.949 | 0.138 | 0.996 |
| | 1 | 0.990 | 0.451 | 0.650 | 0.976 |
| | 2 | 0.983 | 36.030 | 1.961 | 0.922 |
| 5 | 3 | 0.997 | 30.819 | 2.298 | 0.960 |
| | 4 | 1.000 | 16.090 | 0.200 | 0.998 |
| | 5 | 0.999 | 11.053 | 0.154 | 0.994 |

## 4.3 Development of MARS meta-model

MARS, a non-parametric adaptive regression technique, is capable of building flexible regression models (Friedman, 1991). It divides the entire solution space into various intervals of input variables, and builds regression models by fitting individual Splines or Basis functions to each interval (Bera et al., 2006). MARS based meta-models were developed to

simulate saltwater intrusion processes in coastal aquifers, and to predict salinity concentrations at different MLs. The MARS based meta-models predict salinity concentrations through input-output mapping by integrating both a forward and a backward stepwise procedure. To avoid the development of unnecessarily complex models, and to prevent model overfitting, MARS incorporates a backward stepwise procedure that eliminates irrelevant input variables in determining the output variable (Salford-Systems, 2016). Input-output mapping of the developed MARS model can be expressed as

$$BF_i(X) = \begin{matrix} max(0, X_j - \propto) \\ OR \\ max(0, \propto - X_j) \end{matrix} \quad (4.10)$$

$$Y = f(X) = \beta \pm \gamma_k * BF_i(X) \quad (4.11)$$

where, $i$ and $j$ are the indices for Basis functions and input variables (groundwater extractions), respectively; $BF_i$ represents $i^{th}$ Basis functions, the value of which is determined by taking the maximum value between 0 and the difference in values between $X_j$ and $\propto$; $X_j$ represents $j^{th}$ input variables (groundwater extractions); $\propto$ is a constant referred to as knot; $\beta$ indicates a constant value; $\gamma_k$ is the corresponding coefficients of $BF_i(X)$; $' \pm '$ indicates that the next entity may be added to or subtracted from the previous entity. $' \pm '$ depends on the sign of the coefficient of variable. The product of coefficient of variable and the $i^{th}$ $BF$ is subtracted when it has a negative sign, and vice versa; and $Y$ denotes predicted saltwater concentration value at a specified ML. The right hand side of Equation 4.10 is called hinge functions. A hinge function is defined by a variable and a knot.

MARS builds regression based meta-models by adding Basis functions, a mechanism that defines variable intervals. MARS constructs relatively flexible models by fitting piecewise linear regressions. The nonlinearity of the system to be modelled is approximated by using separate regression slopes in distinct intervals of the input variable space. The variables to use and the end points of the intervals for each variable are obtained through an intensive search technique (Salford-Systems, 2016).

Individual MARS meta-models were developed for each ML. The maximum number of Basis functions was selected based on the number of input variables considered to develop the MARS meta-models. These input variables were the transient groundwater extraction values from a combination of 11 production and 5 barrier wells for a management period of 5 years. The study considered 80 (16 wells × 5 years) input groundwater extraction values distributed in space and time. Therefore, 200 Basis functions were chosen based on the 'rule of thumb' (more than two times the number of variables) (Salford-Systems, 2016). As a result, 100 forward steps were allowed to build the MARS meta-models in this study. By

default, MARS generates one knot for every observed data value. A knot at a value of x = 20 within a mirrored pair of hinge functions is presented in Figure 4.5. This default value allows MARS regression to change slope or direction anywhere and as often as the data dictate. MARS meta-models can be made less locally adaptive by increasing the number of data points required between knots. However, it is always preferable to use optimum number observations between the knots in constructing MARS meta-models.

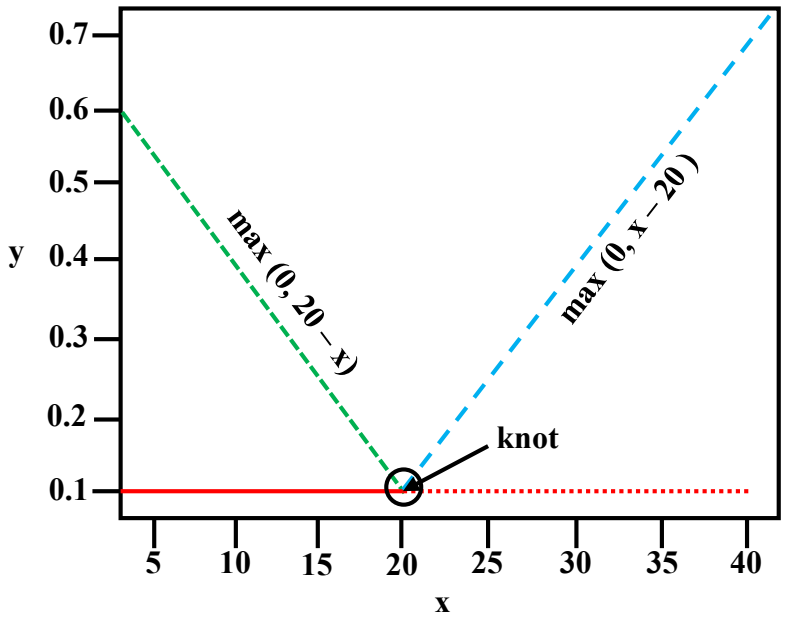


**Figure 4.5 A mirrored pair of hinge functions with a knot at x = 20**

The minimum number of observations between knots was selected by conducting numerical experiments by changing this parameter to a reasonable number of times. No penalty was added to the variables, enabling MARS to give equal priority to all input variables in the forward stepping process of model development. However, in the backward stepping process, MARS sparingly chose the most relevant input variables required to predict the output variables. This backward step kept the developed model as simple as possible, with less possibility of model overfitting. A commercial software package, Salford Predictive Modeller® (SPM, 2016) was used to build the MARS models.

### 4.3.1 Number of input-output training patterns

The present study considered eighty input pumping variables having a domain size of 0-1300 $m^3$/day, making the search space an 80-sided hypercube (Sreekanth and Datta, 2010). Thus, a meta-model incorporating these variables needs a large number of input-output datasets. However, in adaptive meta-models like MARS utilizing adaptive search space methodology in which forward and backward steps are used to modify the search space in relation to variable importance greatly reduces the number of input-output data pairs. The optimum

number of input-output training patterns was selected by observing the significant improvements of MARS models' prediction accuracy. After a certain threshold value, an increase in the number of input-output training patterns did not provide any significant improvement in prediction capability. This particular threshold value was used as the optimum number of training patterns for MARS based meta-models. In the present study, 1500 sets of input-output patterns were found appropriate for the purpose of training and validating the MARS meta-models. These 1500 input-output training patterns were selected based on a compromise between accuracy and computational feasibility, as well as computational time requirements. After using an initial 1000 patterns, an additional 100 patterns were added sequentially to the existing patterns, and the MARS models' prediction accuracy was evaluated by observing the training and testing RMSE values. After 1500 patterns were evaluated, the addition of further patterns did not improve the prediction accuracy significantly. Therefore, 1500 patterns were chosen as an appropriate number for developing MARS based meta-models.

### 4.3.2 Relative importance of input variables

Figure 4.6 illustrates the relative importance of input variables in predicting the saltwater concentrations at ML5 as an example. It is represented as impact factor expressed in a 0-1 scale. Impact factor refers to the contribution of a variable in developing MARS meta-models. The variable that has the highest influence on the model development has an impact factor of 1 whereas the variable that is not contributing to the model development gets an impact factor of 0. Twenty-four, 15, 15, 12, and 7 numbers of variables out of 80 input variables were assigned to have 0 impact factor by MARS in obtaining the salinity concentrations at monitoring locations ML1, ML2, ML3, ML4, and ML5, respectively. For instance, at ML1, 24 input pumping variables were ignored by the MARS meta-model, implying that determination of salinity concentrations at ML1 was not influenced by these ignored input variables. Likewise, the numbers of input variables ignored by MARS in developing the final models at ML2, ML3, ML4, and ML5 are 15, 15, 12, and 7, respectively.

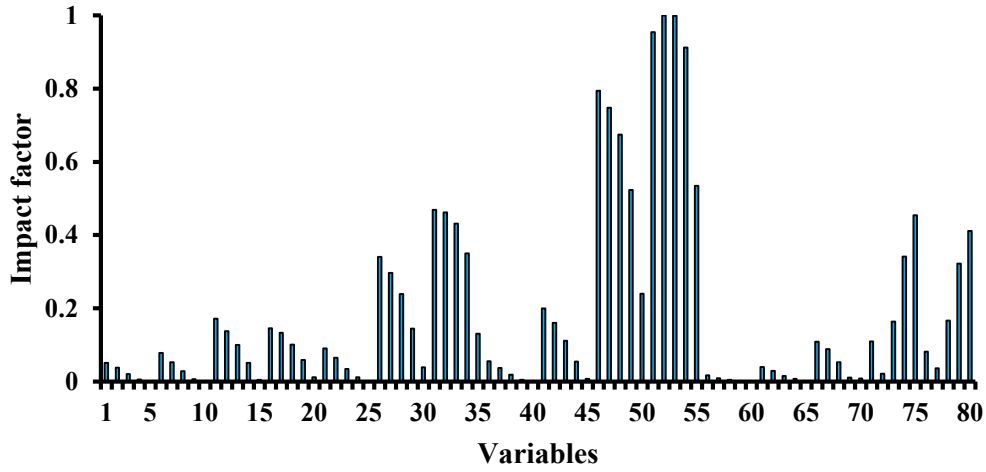


**Figure 4.6 Impact factors of variables predicting salinity concentrations at ML5**

Pumping at a given location may not substantially affect the salinity concentrations at specified MLs. Therefore, these choices of relevant input variables, i.e. the extraction at different locations, were actually made by the MARS meta-model. It is difficult to list each of these relevant extraction variables for each of the MLs. For example, to predict salt concentrations at monitoring location MP5, the relevant extraction values not chosen by the MARS meta-model were: water extractions from production well 1 at time step 5 (variable X5), from production well 2 at time step 5 (variable X10), from production well 5 at time step 5 (variable X25), from production well 8 at time step 5 (variable 40), from barrier well 1 at time step 4 (variable X59), from barrier well 1 at time step 5 (variable X60), and from barrier well 2 at time step 5 (variable X65).

### 4.3.3 Minimum observations between knots

An important parameter in developing MARS based meta-models is the minimum observations between the knots. This parameter was set to its default value in the process of datasets selection. After the correct proportions of datasets were selected, a further trial was conducted to determine the optimum number of observations between the knots (Figure 4.7). RMSE criterion between the training and testing datasets was used to determine the minimum number of observations between knots. The minimum numbers of observations between knots during the MARS meta-model development phase were 4, 15, 2, 1, and 20 at monitoring locations ML1, ML2, ML3, ML4, and ML5 respectively. However, this sequential approach of selecting the minimum number of observations between knots does not guarantee that the optimal value found is 'optimal'. The MARS meta-models thus developed were tested for their performance on a completely new set of validation data. For consistency, the 300 validation datasets were kept the same for testing the performance of all five MARS meta-models developed at the five MLs.

### 4.3.4 Performance of the MARS based meta-models

Each developed MARS meta-model was utilized to predict salinity concentrations at specified MLs, with respect to transient pumping stress applied to the aquifer. The accuracy of the MARS meta-model was evaluated based on its performance with a completely new test dataset, which was different from training and testing datasets during the training phase of individual model development, and is presented in Table 4.4. It is apparent from Table 4.4 that model performance was satisfactory based on the selected statistical indices. MARS produced smaller values of RMSE and Mean Absolute Percentage Relative Error (MAPRE) as well as higher values of R and NS for the new testing datasets.

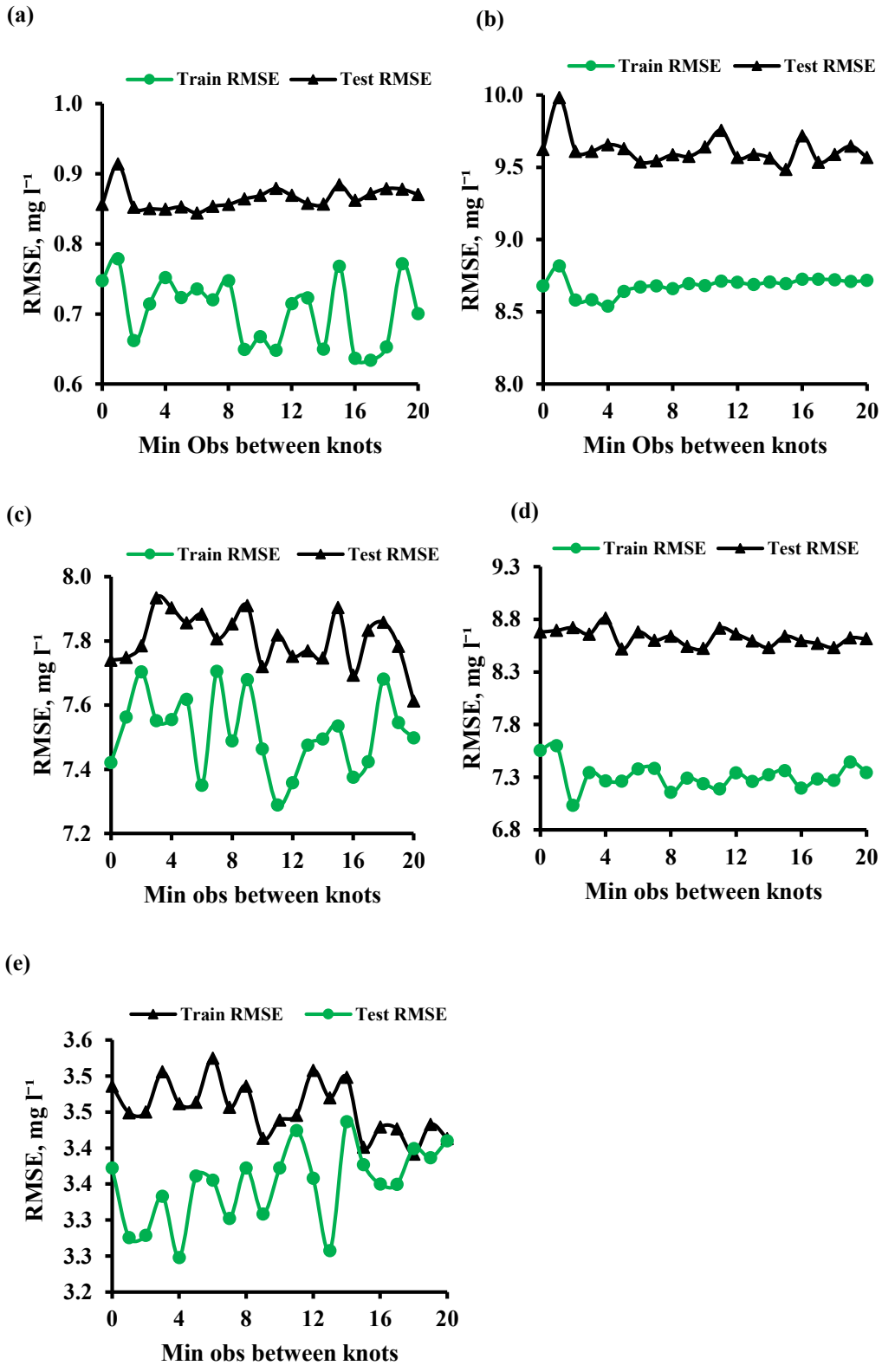


**Figure 4.7 Selection of minimum observation between knots during the training phase; (a) monitoring location ML1, (b) monitoring location ML2, (c) monitoring location ML3, (d) monitoring location ML4, and (e) monitoring location ML5**

RMSE provides a good measure of prediction capability because it incorporates both the variance and bias of the prediction error. Therefore, RMSE criterion provides an indication of how well the model fits the test data. RMSE values calculated on the training and testing data set, as well as on a completely new realization of the test dataset for MARS models developed at ML1-ML5, are presented in Table 4.4. It is observed from Table 4.4 that the training and testing RMSE values did not differ significantly. This indicates that MARS model did not overfit during the training phase of model development. The performances of MARS models on a new test dataset were also satisfactory, based on RMSE criterion. However, the downside of using RMSE is that it gives more weight to the outlying observations. Therefore, MAPRE and NS criteria were also used to evaluate prediction performance of the developed MARS meta-models. MAPRE provides better evaluation of the model performance by providing information on the distribution of errors. The developed models had a relatively small MAPRE values (average MAPRE values were less than 5% for all estimates) at all MLs. This indicates that MARS is able to perform well from the relative error viewpoint. On the other hand, NS criteria also provides the evaluation of performance of a model in terms of prediction capability. The prediction capability of any model is deemed acceptable when NS values are greater than or equal to 0.8 (Shu and Ouarda, 2008). MARS meta-models for all MLs provided NS values greater than 0.8, indicating that the MARS meta-model's prediction errors were within acceptable limits. The values of R for all estimates were very close to 1, indicating that the model performance was also satisfactory on correlation coefficient viewpoint.

**Table 4.4 Statistical indices for performance evaluation of MARS meta-model**

| MLs | MAPRE,% | | | NS | | | RMSE, mg/l | | | R | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Train | Test | New Test | Train | Test | New Test | Train | Test | New Test | Train | Test | New Test |
| ML1 | 1.26 | 1.49 | 1.33 | 0.97 | 0.70 | 0.79 | 0.75 | 0.988 | 0.984 | 0.985 | 1.26 | 1.49 |
| ML2 | 0.48 | 0.51 | 0.43 | 0.99 | 8.69 | 9.48 | 8.21 | 0.998 | 0.998 | 0.998 | 0.48 | 0.51 |
| ML3 | 0.53 | 0.56 | 0.54 | 0.99 | 7.70 | 7.78 | 7.85 | 0.998 | 0.998 | 0.998 | 0.53 | 0.56 |
| ML4 | 0.08 | 0.09 | 0.08 | 0.99 | 7.55 | 8.67 | 7.39 | 0.999 | 0.999 | 0.999 | 0.08 | 0.09 |
| ML5 | 0.04 | 0.04 | 0.05 | 0.99 | 3.41 | 3.41 | 3.40 | 0.999 | 0.999 | 0.999 | 0.04 | 0.04 |

*MAPRE = Mean Absolute Percentage Relative Error, NS = Nash-Sutcliffe Efficiency Coefficient, RMSE = Root Mean Squared Error, R = Correlation Coefficient

The values of Threshold Statistics (TS) (Raghavendra and Deka, 2015) were calculated to estimate the percentage of data points whose relative error values were less than the specified threshold values. Four threshold values were selected to ensure that all relative

error values fall within the highest assigned threshold value. Performance of the MARS meta-models in terms of TS criterion indicates that at ML1, MARS had 51.33% estimates lower than 1% relative error. More than 97% estimates were lower than 5% relative error, whereas 99% estimates were lower than 8% relative error at ML1. All the estimates for other MLs follow similar trend and is presented in Table 4.5.

**Table 4.5 Threshold statistics based on percentage relative error criterion**

| MLs | Threshold Statistics (for new test data), % | | | |
|---|---|---|---|---|
| | "<1%" | "<2%" | "<5%" | "<8%" |
| ML1 | 51.33 | 80.67 | 97.67 | 98.67 |
| ML2 | 92.00 | 99.67 | 100.0 | 100.0 |
| ML3 | 87.67 | 98.33 | 100.0 | 100.0 |
| ML4 | 100.0 | 100.0 | 100.0 | 100.0 |
| ML5 | 100.0 | 100.0 | 100.0 | 100.0 |

MARS provided a transparent relationship between the transient groundwater extraction values and the resulting saltwater concentration values at specified MLs. Prediction of salinity concentrations was straightforward from the regression equations developed by MARS meta-models. During the development of this regression equation, MARS learned and extracted knowledge from the dataset, and provided an explicit relationship between the input and output variables (Yang et al., 2003). In addition, MARS parsimoniously selected input variables that were most relevant in obtaining the output variables for developing the model. In fact, the number of MARS meta-models depends on the number of output variables, permitting this parsimonious selection process (Sreekanth and Datta, 2011a). In a physical system, certain sets of stresses influence the inputs more than others do. In this study, input stresses were the groundwater extractions, and the outputs were the saltwater concentrations. It is obvious that extractions at certain locations might not have significant impact on the salt concentrations at certain specified MLs. The inbuilt capability of the MARS model was utilized to select these more relevant input variables (groundwater extractions) for predicting salt concentrations at different MLs. Therefore, in developing model at any particular ML, MARS chose only those input variables that were the most influential in determining saltwater concentration at that location. Thus, these most relevant groundwater extraction values might vary for different MLs.

## 4.4 Summary and conclusions

A FCM clustering based FIS and an adaptive MARS based meta-models were developed to predict saltwater concentrations at different MLs in response to groundwater abstraction from an illustrative multilayered coastal aquifer system. The trained and validated meta-models were utilized as approximate simulators (emulators) of the complex density dependent coupled flow and salt transport processes in coastal aquifers. The illustrative coastal aquifer study area was stratified, and it incorporated vertical heterogeneity in hydraulic conductivity. The saltwater intrusion process was induced by the combination of pumping from production wells for beneficial use and barrier well pumping. Training and validation datasets, in terms of transient groundwater extraction and the resulting salinity concentrations at different MLs, were obtained from a finite-element based 3-D combined flow and salt transport numerical simulation model. Performance evaluations conducted in this study required randomized generation of input stresses and corresponding response patterns for the chosen study area. A typical coastal study area was chosen for synthetically generating field data, using random inputs in a numerical simulation model. Results obtained from the developed meta-models and the numerical model simulated values were judged and evaluated by comparing prediction performance. The errors in prediction for the illustrative study area were quite small, and therefore it was not felt necessary to compare the evaluation results with those obtained utilizing meta-models based on other algorithms, e.g., an ANN and GP.

The developed FIS is capable of implementing crisp nonlinear mapping specified by fuzzy IF-THEN rules that encode expert or common sense knowledge of the problem. A Sugeno FIS is accurate, computationally efficient, works well with adaptive techniques, suitable for mathematical analysis, and has guaranteed continuity of the output surface. It has the ability to represent concentration field at multiple locations with a single training. In that sense, it overcomes the need for multiple meta-models for multiple output problems. Therefore, this new meta-model is potentially applicable as a new and more efficient meta-model for application in coastal aquifers. The time required to train and validate the FIS was almost negligible compared to other existing meta-models for salinity predictions. Moreover, once trained, the FIS can provide the prediction results very quickly.

This study also illustrates the use of MARS as a useful soft-computing tool in predicting salinity concentrations in a linked S-O methodology for optimal groundwater extraction strategies. Separate MARS based meta-models were developed at individual MLs. MARS provides simple regression equation by utilizing a set of Basis functions, making it a good candidate meta-model in the S-O approach. Moreover, the adaptive nature of the MARS algorithm, resulting from its inherent characteristics of selecting the most influential input variables, reduces the chanced of model overfitting and increases the robustness of MARS

based meta-models. MARS maps the nonlinear relationship between transient groundwater extraction values and salinity concentrations at different MLs. From the performance evaluation viewpoint, the proposed MARS meta-model can be successfully applied to predict salinity concentrations in coastal aquifers.

However, the practical utility of the developed meta-models lies in their use when repeated simulations of the flow and transport processes are necessary. The enormous reduction in computational time for predicting the impact of any pumping strategy on the resulting salinity scenario makes these meta-models ideal for linkage within an S-O model for developing regional scale coastal aquifer management strategies. No doubt further evaluations using field data from coastal aquifers are necessary to establish the applicability of such meta-models within the linked S-O methodology. The next chapter discusses the development and evaluation of two objective saltwater intrusion management models by utilizing trained FIS and MARS based meta-models as an approximate of the numerical simulation models in the coupled S-O model.

# Chapter 5: Management of coastal aquifers using the meta-model based coupled simulation-optimization approach

Partial contents of this chapter have been published and copyrighted, as outlined below:

Roy, D. K., & Datta, B. (2018b). Comparative efficiency of different artificial intelligence based models for predicting density dependent saltwater intrusion processes in coastal aquifers and saltwater intrusion management utilizing the best performing model. *Desalination and Water Treatment, 105*, 160–180.

Roy, D. K., & Datta, B. (2018a). A surrogate based multi-objective management model to control saltwater intrusion in multi-layered coastal aquifer systems. *Civil Engineering and Environmental Systems*, 1-26.

This chapter discusses the development and evaluation of two objective saltwater intrusion management models by utilizing trained FIS and MARS based meta-models as an approximate of the numerical simulation models in the S-O model.

## 5.1 Summary

Simulation of saltwater intrusion processes requires the application of density dependent coupled flow and salt transport numerical simulation models that are computationally intensive. Therefore, linking these numerical simulation models to a linked S-O methodology for management of coastal aquifers results in enormous computational time for solution. For instance, linked S-O methodology for a small 3-D coastal aquifer may require as much as a 30-day computer run time (Dhar and Datta, 2009a). One way of achieving computational efficiency in such a linked S-O approach is the use of reasonably accurate approximate meta-models, which are very useful tools to approximately simulate the coupled flow and salt transport processes in coastal aquifers (Banerjee et al., 2011; Bhattacharjya et al., 2007; Kourakos and Mantoglou, 2009; Kourakos and Mantoglou, 2013; Sreekanth and Datta, 2011a). These approximate meta-models can replace the computationally intensive numerical simulation models to achieve computational efficiency in a linked S-O approach. Repeated use of these meta-models by the optimization algorithm does not pose significant computational time because these meta-models are simplified representation of the complex simulation models. One of the prime objectives of the current study is to develop computationally efficient meta-models to replace complex numerical simulation models within a coupled S-O approach in developing saltwater intrusion management models in coastal aquifers. This chapter presents the development of multiple objective saltwater intrusion management models by utilizing trained FIS and MARS based meta-models as approximate simulators of the numerical simulation models. These meta-models were then

externally linked to a Controlled Elitist Multiple Objective Genetic Algorithm (CEMOGA) within a computationally feasible coupled S-O approach in order to develop optimal groundwater extraction strategies for a coastal aquifer system.

Two management models prescribing optimal groundwater extraction strategies in coastal aquifers were developed by using FIS and MARS based meta-models, respectively. CEMOGA (Deb and Goel, 2001) was used to search for optimal groundwater extraction strategies, while maintaining the maximum allowable saltwater concentration limits at the specified MLs. Two conflicting objectives of groundwater extraction strategy were considered in this study. The first objective ensures the maximum withdrawal of groundwater for beneficial purposes. The second objective minimizes the water extraction from barrier pumping wells in order to control saltwater intrusion by establishing a hydraulic head barrier near the coastal boundary. The multiple objective management model provides a trade-off between these conflicting objectives in terms of a Pareto optimal front, which consists of several non-dominated feasible alternative groundwater extraction strategies that meet the pre-specified allowable saltwater concentration limits at specified MLs. Using any one of the Pareto optimal solutions, an optimal pumping strategy can be specified for the management time frame to control saltwater intrusion within the permissible limits. However, selection of a single Pareto optimal solution as the prescribed preferred optimal strategy will require further preference ordering of the solutions (Datta and Peralta, 1986).

Attaining computational feasibility and efficiency is the main advantage of meta-model based linked S-O models for the regional scale management of coastal aquifers. Computational efficiency can be significantly improved by implementing parallel computation capabilities in which objective function (s) and all other constraints of the optimization problem are evaluated in parallel by distributing them into physical cores of a PC. In this study, four physical cores of a standard seven core PC were utilized to evaluate the objective functions and constraints of the multiple objective saltwater intrusion management problem. Integration of parallel processing capabilities within the optimization model was shown to improve both computational efficiency and the feasibility of solving such large-scale multiple objective problems.

The performance of the proposed FIS and MARS meta-model based coupled S-O approach was evaluated for illustrative coastal aquifer systems. The performance evaluation of the saltwater intrusion management model using FIS based meta-models was executed for a coastal aquifer study area proposed in Sreekanth and Datta (2011a). The MARS based management model was developed for an illustrative study area described in Subsection 4.2.5 of chapter 4. The results of the management models indicated that both FIS and MARS based meta-models provided acceptable, accurate, and reliable groundwater extraction patterns to limit the saltwater concentrations within the pre-specified maximum allowable limits.

Therefore, the methodology developed utilizing the FIS and MARS based meta-models is potentially applicable for developing optimal groundwater extraction strategies for sustainable regional scale management of coastal aquifers.

## 5.2 Saltwater intrusion management model

The proposed saltwater intrusion management model was multi-objective in nature in that two conflicting objectives of groundwater extraction strategy were considered. The first objective ensured the maximum withdrawal of groundwater for beneficial purposes. The second objective minimized the water extraction from barrier pumping wells to control saltwater intrusion by establishing a hydraulic head barrier near the coastal boundary.

### 5.2.1 Multiple objective optimization formulation

The mathematical formulation for the proposed saltwater intrusion management methodology is expressed by the following equations

$$Maximize: f_1(Q_{PW}) = \sum_{r=1}^{R}\sum_{t=1}^{T} {}_r^tQ_{PW} \tag{5.1}$$

$$Minimize: f_2(Q_{BW}) = \sum_{k=1}^{K}\sum_{t=1}^{T} {}_k^tQ_{BW} \tag{5.2}$$

Subject to

$$C_i = f(Q_{PW}, Q_{BW}) \tag{5.3}$$

$$C_i \leq C_{max}\ \forall_i \tag{5.4}$$

$$Q_{PW}(min) \leq {}_r^tQ_{PW} \leq Q_{PW}(max) \tag{5.5}$$

$$Q_{BW}(min) \leq {}_k^tQ_{BW} \leq Q_{BW}(max) \tag{5.6}$$

where, ${}_r^tQ_{PW}$ represents water extraction from the $r^{th}$ pumping well throughout $t^{th}$ time phase; ${}_k^tQ_{BW}$ stands for water extraction from $k^{th}$ barrier extraction well throughout $t^{th}$ time phase; $C_i$ symbolizes saltwater concentrations at $i^{th}$ monitoring locations at the closure of the management period. Equation 5.3 indicates salinity concentration is a function of both production and barrier extraction wells; Equation 5.4 specifies the maximum allowable salt concentration at specified monitoring locations; Equation 5.5 and Equation 5.6 provide the lower and upper limits on the water extraction rate from the pumping wells and barrier

extraction wells, respectively. Subscripts $PW$ and $BW$ stand for production bores and barrier extraction wells, respectively. $R$, $K$, and $T$ stand for the entire pumping wells, barrier extraction wells, and time periods, respectively. The first objective of maximization of groundwater extraction from the pumping wells for beneficial use is represented by Equation 5.1, and the second objective of minimizing the water extraction from barrier pumping wells is given by Equation 5.2.

### 5.2.2 Optimization algorithm-CEMOGA

A population based search algorithm, CEMOGA (Deb and Goel, 2001) was utilized for optimization. This algorithm demonstrated better convergence for a number of complex optimization problems (Deb and Goel, 2001) . The key feature of CEMOGA lies in its ability to prefer individuals, who despite having low fitness values, help in increasing the diversity of the population. The diversity of the population is preserved by regulating the population's elite members during the progress of the algorithm, making the new population more diverse. More specifically, this regulated elitist tactic allows a particular fraction of the population (dominated populations) to be part of the current preeminent non-dominated solutions. This inclusion of a particular portion of the dominated solutions in the non-dominated solutions greatly reduces the effect of elitism. 'Pareto Fraction' and 'Distance Function' are the two parameters that control the extent of elitism. The first parameter restricts the number of individuals (elite members) on the Pareto front, whereas the second one is intended to preserve the diversity on the Pareto front by giving preference to individuals who are reasonably far-off on the front (Deb and Goel, 2001).

### 5.2.3 Coupled S-O model

In the coupled S-O approach, the simulation part was replaced by properly trained and validated meta-models. The meta-models were then linked to the CEMOGA to develop global Pareto optimal groundwater extraction strategies. The optimization algorithm, CEMOGA generates an initial set of groundwater extraction values, and the resulting salinity concentrations were evaluated by the meta-models. Constraints in terms of maximum permissible salinity concentrations at specified MLs were checked. The process was continued until all imposed constraints were satisfied. Schematic representation of the saltwater intrusion management model developed in the present study is illustrated in Figure 5.1. Two coupled S-O based saltwater intrusion management models, viz., FIS-CEMOGA and MARS-CEMOGA were developed in this research.

### 5.2.4 Parallel computing

Parallel computing was performed in the MATLAB (MATLAB, 2017c) environment by distributing the objective functions and constraints to all physical processors of the PC rather than running them serially with parallel gradient estimation. Therefore, parameters were automatically passed to worker machines during the execution of parallel computations. A parallel pool of four workers was implemented by utilizing the four physical cores of a standard computer [Intel® Core™ i7-2600 CPU @3.40 GHz, 8GB RAM] by utilizing parallel computing toolbox of MATLAB.

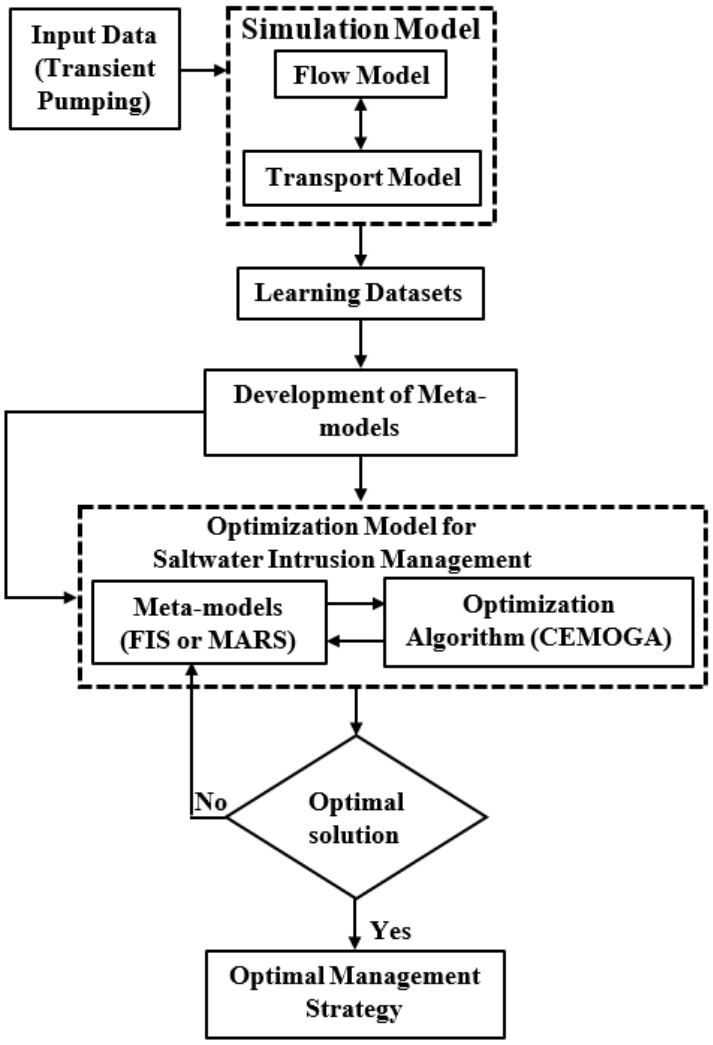


**Figure 5.1 Flow diagram of the management model**

## 5.3 FIS-CEMOGA based saltwater intrusion management model

The performance of the FIS-CEMOGA based management model was evaluated using a small illustrative coastal aquifer system of 2.52 $km^2$ aerial extent, and an aquifer thickness

of 60 m (Sreekanth and Datta, 2011a). The aquifer system with finite element meshes, and location of different production and barrier pumping wells with MLs is illustrated in Figure5.2. The seaside boundary (2.04 km long) was assumed to have a constant head of 0.0 m and a constant concentration of 35 kg/m$^3$. The other two sides of the model domain (total length = 4.85 km) and the bottom of the aquifer were considered as no flow boundaries. The average thickness of the aquifer was 60 m, which was vertically discretised into 3 layers of 20 m each. However, the aquifer properties were the same at all of these 3 layers. The sizes of the triangular finite elements were decided based on the mesh dependency test. Element sizes smaller than 150 m induced an additional computational burden but did not achieve any further improvement in terms of computational efficiency and accuracy. Therefore, the entire model domain was divided into triangular finite elements of 150 m size. A finer element size of 60 m was used near the production and barrier extraction wells. Abstraction of water was carried out from 8 production wells placed inside the model domain for beneficial purposes, and 3 barrier extraction wells placed near the coastline to hydraulically control saltwater intrusion.

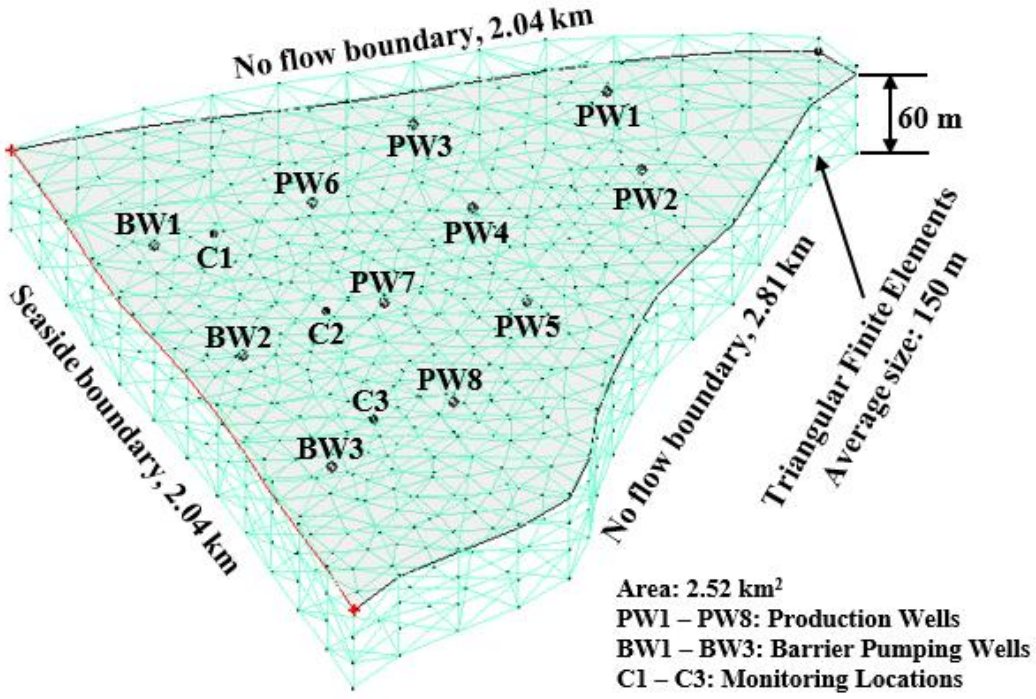


**Figure 5.2 Three-dimensional representation of the study area with finite elements and location of different production and barrier wells**

Barrier extraction wells were placed near the coastline to create a hydraulic barrier along the coastline. When water was pumped out of the barrier extraction wells, a gradient was created that was filled in by the adjacent freshwater due to the natural gradation of freshwater towards the ocean, thereby preventing saltwater from entering into the aquifer. Aquifer recharge of 0.2 m/year was assumed to be uniformly distributed over the entire model

domain. Therefore, it was unlikely that the aquifer was filled with saltwater within a simulation period more than 3 years. Moreover, the illustrative aquifer including the initial and boundary conditions were chosen carefully to avoid trivial scenarios. Aquifer parameters used in the simulation is provided in Table 5.1. (Sreekanth and Datta, 2011a).

**Table 5.1 Aquifer properties**

| Parameters | Assigned values | Unit |
|---|---|---|
| Hydraulic conductivity in x-direction | 25 | m/day |
| Hydraulic conductivity in y-direction | 25 | m/day |
| Hydraulic conductivity in z-direction | 0.25 | m/day |
| Molecular diffusion coefficient | 0.69 | $m^2$/day |
| Lateral dispersivity | 35 | m |
| Longitudinal dispersivity | 80 | m |
| Density reference ratio | 0.025 | - |
| Soil porosity | 0.2 | - |
| Aquifer recharge | 0.2 | m/year |

Total numbers of pumping wells considered in this study were 11 (8 production wells + 3 barrier extraction wells). The simulation was performed over a period of 3 years. The simulation was performed in the transient mode in which total simulation period of 3 years was divided into 219 uniform time steps of 5 days each. Trials were conducted using four time steps of 1 day, 3 days, 5 days, and 15 days. The simulation time step of 5 days was selected by considering a trade-off between the computational time requirement and the accuracy of the simulation. The management period of 3 years was divided into 3 uniform time steps of one year each. Abstraction of water in this 1-year time step was assumed to be constant. Therefore, the study considered 33 pumping variables (11 pumping locations × 3 time steps) that were fed to the models as inputs. The lower and upper bounds of pumping values for these variables were set as 0 and 1300 $m^3$/day, respectively. Salinity concentrations were monitored at the end of the simulation period at three potential MLs.

### 5.3.1 Management model performance

Prediction accuracy and feasibility of incorporation of any meta- model within a linked S-O methodology determines the robustness of the approach. As mentioned in chapter 4, FCM based FIS prediction models are good candidates for replacing and approximating numerical simulation models within a linked S-O approach to determine Pareto optimal groundwater

extraction policies. The FIS model was externally linked as binding constraint within the optimization framework in order to ensure that the assigned maximum allowable salt concentration limits for the specified MLs were not exceeded. The optimal parameter values for the optimization algorithm were chosen by conducting several trials using different combinations of these parameters. Based on these trials, a population size of 2000, crossover fraction of 0.9, mutation probability of 0.1, and migration fraction of 0.2 with an interval of 20 in the forward direction were selected. The function and constraint tolerances were set to $1\times10^{-5}$ and $1\times10^{-3}$, respectively. A Pareto front population fraction of 0.7 was chosen to allow 1400 (2000 × 0.7) non-dominated feasible optimal solutions in the Pareto front. The optimization algorithm needed to evaluate 7,896,000 (3948 generations × 2000 populations) groundwater extraction patterns to arrive at an optimal solution.

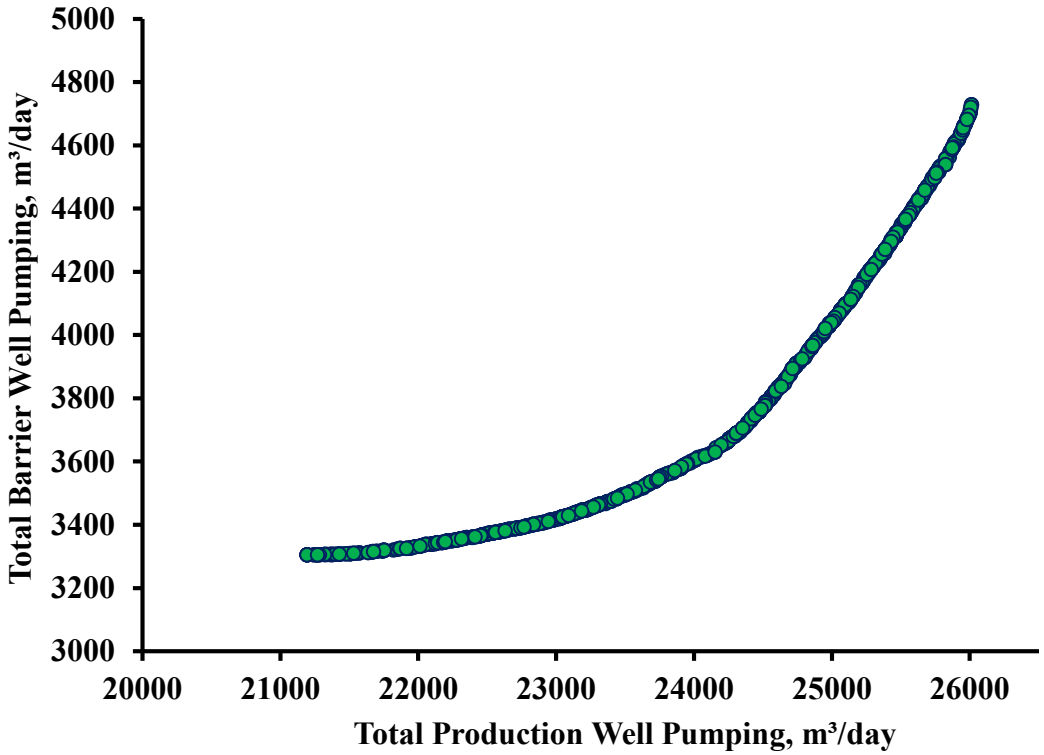


**Figure 5.3 Pareto optimal front of the developed saltwater intrusion management model**

The saltwater intrusion management model in the present study took into consideration two conflicting objectives of the groundwater extraction strategy. The optimization model considered 33 input variables relating to pumping of water from 11 locations for 3 time steps, and 3 output variables that correspond to saltwater concentration at specified MLs at the completion of the management time horizon of 3 years. The optimization model provided an optimal solution in the form of a Pareto optimal front that represents the non-dominated trade-off between these two conflicting objectives. The Pareto optimal front presented in Figure 5.3 provides several sets of optimal groundwater extraction

values obtained while ensuring that the maximum permissible saltwater concentrations at specified MLs were not exceeded. It is observed from Figure 5.3 that increasing the rate of extraction from production wells requires an additional amount of water to be withdrawn from the barrier extraction wells. Water extracted from barrier extraction wells generally cannot be used for beneficial purposes due to high salinity contents. Therefore, managers can choose the rate of barrier well pumping based on the demand for beneficial water use, while keeping the pre-set maximum allowable saltwater concentrations at certain MLs in mind. These Pareto optimal solutions show the conflicting nature of the two objectives, a necessary condition for a multiple objective optimal management model.

The validity of the proposed coastal aquifer management model was evaluated by observing the actual violation of any one of the constraints within the optimization model. It is noted that the saltwater concentrations obtained from the optimization model solution (determined by the global FIS model within the optimization framework) were smaller than the pre-specified maximum allowable saltwater concentrations at all MLs. This implies that all the imposed constraints were satisfied, and no constraint violation occurred during the search process. Moreover, obtained saltwater concentrations were very close to the prescribed values, which indicate that the optimization model converged to the upper limit of the saltwater concentration constraints. Ten solutions were selected randomly from different regions of the Pareto front to check the constraint satisfaction and constraint violations. This is shown in Table 5.2.

**Table 5.2 Constraint satisfaction within the management model in reaching the global Pareto optimal solution**

| Solution Number | C1 ≤ 500 mg/l | | | C2 ≤ 600 mg/l | | | C3 ≤ 600 mg/l | | |
|---|---|---|---|---|---|---|---|---|---|
| | FIS | Diff. | % Diff. | FIS | Diff. | % Diff. | FIS | Diff. | % Diff. |
| 1 | 499.81 | 0.19 | 0.04 | 587.09 | 12.91 | 2.15 | 599.96 | 0.04 | 0.01 |
| 2 | 499.81 | 0.19 | 0.04 | 587.19 | 12.81 | 2.14 | 599.84 | 0.16 | 0.03 |
| 3 | 499.89 | 0.11 | 0.02 | 585.50 | 14.50 | 2.42 | 599.86 | 0.14 | 0.02 |
| 4 | 499.91 | 0.09 | 0.02 | 585.06 | 14.94 | 2.49 | 599.85 | 0.15 | 0.03 |
| 5 | 499.85 | 0.15 | 0.03 | 587.38 | 12.62 | 2.10 | 599.96 | 0.04 | 0.01 |
| 6 | 499.94 | 0.06 | 0.01 | 587.46 | 12.54 | 2.09 | 599.98 | 0.02 | 0.00 |
| 7 | 499.86 | 0.14 | 0.03 | 586.21 | 13.79 | 2.30 | 599.95 | 0.05 | 0.01 |
| 8 | 499.92 | 0.08 | 0.02 | 587.59 | 12.41 | 2.07 | 599.92 | 0.08 | 0.01 |
| 9 | 499.84 | 0.16 | 0.03 | 587.36 | 12.64 | 2.11 | 599.97 | 0.03 | 0.01 |
| 10 | 499.89 | 0.11 | 0.02 | 586.80 | 13.20 | 2.20 | 599.97 | 0.03 | 0.00 |

*FIS = FIS output within the optimization model, Diff. = Difference between imposed constraint and the FIS output

It is observed in Table 5.2 that FIS predicted saltwater concentration values within the optimization model were very close to the pre-assigned saltwater concentration limits at different MLs. Therefore, the limited performance evaluation of the proposed FIS-CEMOGA based saltwater intrusion management methodology demonstrated its potential applicability to prescribe accurate Pareto optimal solutions for optimal groundwater extraction from a set of beneficial pumping wells and barrier extraction wells in a coastal aquifer.

## 5.4 MARS-CEMOGA based saltwater intrusion management model

The performance of the MARS-CEMOGA based saltwater intrusion management model was demonstrated using the multilayered coastal aquifer system presented in Subsection 4.2.5 of chapter 4. The coupled S-O based multiple objective optimization formulation was similar to the one described in Subsection 5.2.1. However, the proposed management model considered 80 decision variables in solving the optimization formulation. These variables correspond to water extraction from16 potential pumping locations (eleven production wells plus five barrier wells) for five time steps. Variables X1–X5, X6–X10, X11–X15, X16–X20, X21–X25, X26–X30, X31–X35, X36–X40, X41–X45, X46–X50, and X51–X55 represent water extraction from the production wells PW1–PW11 during the 1st, 2nd, 3rd, 4th, and 5th year of management period. Variables X56–X60, X61–X65, X66–X70, X71–X75, and X76–X80 symbolize extraction from barrier wells BW1–BW5 for the five years of management period. The minimum and maximum bounds of these decision variables were set as 0 and 1300 $m^3$/day, respectively. MLs were located at three different salinity zones of the aquifer: ML1 was located at region where salinity concentration was very low, ML2 and ML3 were located at moderately saline zones of the aquifer, and ML4 and ML5 were situated slightly close to the coastline where the salinity concentrations were relatively high. The MARS meta-model had eighty predictors (transient pumping values) corresponding to sixteen pumping locations for five time steps. For each ML, a separate MARS meta-model was developed, each forecasting salinity concentration for single ML.

In the management model, the permissible limits of salt concentrations were set with a view to permitting the use of extracted water for different purposes. For ML1, the maximum permissible salt concentration was set as 40 mg/l. Therefore, the extracted water from pumping wells located north to ML1 in the model domain are of very high quality, and could be used for a wide range of purposes. The maximum allowable salt concentrations at ML2 and ML3 were set as 1000 and 800 mg/l, respectively. So, the water extracted from this zone of the aquifer could be used for irrigating salt tolerant crops. The highest acceptable limit of salt concentrations at ML4 and ML5 were respectively set as 6400 and 5500 mg/l. The water extracted from this region could be used for aquaculture. The maximum allowable salt

concentration limits set here were intended to demonstrate the performance of the proposed MARS based saltwater intrusion management model in an illustrative coastal aquifer system. However, in a real world application, the exact demand and type of water use should be taken into account to set the maximum allowable saltwater concentration for different regions within the study area.

The multiple objective global optimization toolbox of MATLAB (MATLAB, 2017b) was used to run the optimization problem. The optimization algorithm used a population size of 1600, a crossover fraction of 0.9, and a migration fraction of 0.2. A Pareto front population fraction of 0.7 was used. Considering the number of decision variables, a large number of population was intentionally chosen in order to allow the genetic algorithm to use a more diverse search space. Other parameters of the genetic algorithm were selected based on numerical experiments using several combinations of different parameters. The value of the function tolerance and constraint tolerance were set as $1\times10^{-5}$ and $1\times10^{-4}$, respectively. A formal sensitivity analysis using different Genetic Algorithm (GA) parameters may provide detailed information of the optimization procedure. However, in this limited evaluation, we did not perform such a sensitivity analysis of the GA parameters.

### 5.4.1 Management model performance

Developed MARS based meta-models are computationally efficient meta-models that can be implemented in a linked S-O problem setting quite effectively. The MARS-CEMOGA saltwater intrusion management model took 22.84 minutes to evaluate 2867201 functions (Generation: 1791 × Number of population: 1600) in a parallel computing platform of MATLAB environment consisting of 4 workers corresponding to 4 physical cores of a standard PC [Intel® Core™ i7-2600 CPU @ 3.40 GHz, 8.00 GB RAM]. To compare performance efficiency of the parallel computing platform, the same optimization problem was run in serial using the same PC, and with the same random state of the previous run with parallel computing platform. The same problem took around 60 minutes when the objective functions and constraints were evaluated in series. Ketabchi and Ataie-Ashtiani (2015) also reported computational time saving in a management model by utilizing parallel processing to run the optimization formulation. However, the result of the present study cannot be compared with those obtained by Ketabchi and Ataie-Ashtiani (2015) because they differ in study conditions and computer configurations.

Five MARS meta-models predicting saltwater concentrations at five different MLs were linked externally to the optimization algorithm within the linked S-O methodology. MARS meta-models were developed based on the most influential input transient groundwater extraction values. Therefore, different combinations of relevant input variables

were evaluated simultaneously to determine the optimal groundwater extraction values from a set of production and barrier extraction wells. These optimal extraction values obtained by the linked S-O model ensured that saltwater concentrations at any one of the MLs did not exceed the maximum allowable saltwater concentrations. The S-O model provided several optimal solutions of optimal groundwater pumping from each of the production and barrier extraction wells for the management period of five years. For each alternate optimal solution, total extraction from the spatially and temporally distributed production wells versus total extraction from the spatially and temporally distributed barrier wells were calculated and presented in the form of a Pareto optimal front. The Pareto optimal front obtained by MARS-CEMOGA methodology for this management model is shown in Figure 5.4. The Pareto front contains 1120 solutions, each of which relates to the groundwater extraction from a combination of eleven production wells and five barrier extraction wells. Therefore, this Pareto front provides several alternate solutions that maximize water extraction from production wells and minimize barrier well extraction, while maintaining the salt concentrations at specified MLs within the specified acceptable limits. The Pareto optimal front illustrates the trade-off between two conflicting objectives, and the front is extensive, indicating the trade-offs between the two objectives are well captured. It is observed from Figure 5.4 that an approximate water volume of 23,000 $m^3$/day can be extracted from the production wells with a very small amount of water (280 $m^3$/day) extracted from the barrier wells. However, as the amount of water extracted from production wells needs to be increased, an increased amount of barrier well extraction is also required to maintain the water quality as specified in terms of salinity.

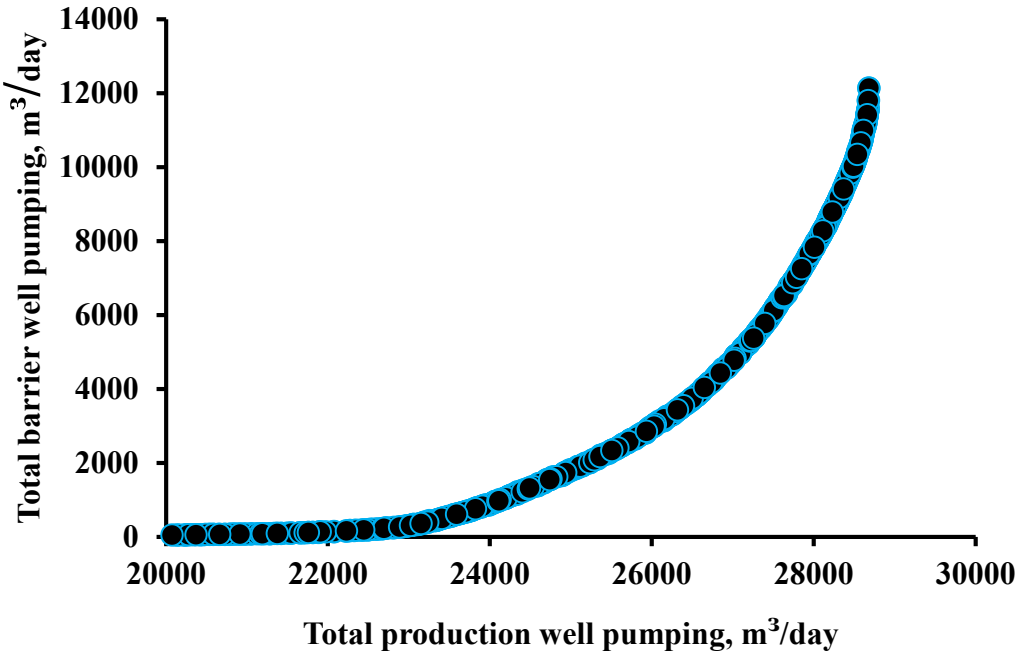


**Figure 5.4 Pareto optimal front of the management model**

5.4.2 Verification of the management model

Salinity concentrations obtained from optimal solutions were very close to the pre-specified salinity concentrations, which imply that optimal solutions converged to the upper limit of constraints. Optimal solutions obtained as solution of the management model were then validated by using the prescribed optimal groundwater extraction values to numerically simulate the coupled flow and salt transport processes. For this, 20 solutions were randomly selected from different regions of the Pareto front. Groundwater extraction rates of these selected solutions were used as inputs to the numerical simulation model. The resulting salinity concentrations at different MLs obtained from the simulation model were compared with MARS predicted saltwater concentration values.

**Table 5.3 Percentage relative error between MARS predicted and FEMWATER simulated saltwater concentrations using optimal groundwater extraction values given by the optimization model**

| Solution | ML1: C1 ≤ 40 mg/l | ML2: C2 ≤ 1000 mg/l | ML3: C3 ≤ 800 mg/l | ML4: C2 ≤ 6400 mg/l | ML5: C2 ≤ 5500 mg/l |
|---|---|---|---|---|---|
| 1 | 2.379 | 2.973 | 1.672 | 0.159 | 0.065 |
| 2 | 2.977 | 2.915 | 1.274 | 0.173 | 0.040 |
| 3 | 2.517 | 2.941 | 1.608 | 0.160 | 0.058 |
| 4 | 2.732 | 2.947 | 1.453 | 0.178 | 0.044 |
| 5 | 3.129 | 2.907 | 1.166 | 0.180 | 0.034 |
| 6 | 3.121 | 2.914 | 1.297 | 0.176 | 0.037 |
| 7 | 2.593 | 2.954 | 1.659 | 0.160 | 0.061 |
| 8 | 2.923 | 2.922 | 1.282 | 0.180 | 0.034 |
| 9 | 2.844 | 2.933 | 1.347 | 0.173 | 0.043 |
| 10 | 2.645 | 2.954 | 1.502 | 0.176 | 0.047 |
| 11 | 2.460 | 2.971 | 1.669 | 0.165 | 0.064 |
| 12 | 2.628 | 2.956 | 1.544 | 0.171 | 0.048 |
| 13 | 2.593 | 2.954 | 1.659 | 0.160 | 0.061 |
| 14 | 2.305 | 3.016 | 1.878 | 0.162 | 0.087 |
| 15 | 2.868 | 2.934 | 1.336 | 0.176 | 0.039 |
| 16 | 2.858 | 2.939 | 1.355 | 0.182 | 0.036 |
| 17 | 2.755 | 2.958 | 1.401 | 0.176 | 0.040 |
| 18 | 2.949 | 2.918 | 1.289 | 0.176 | 0.036 |
| 19 | 3.074 | 2.903 | 1.214 | 0.177 | 0.031 |
| 20 | 3.060 | 2.894 | 1.225 | 0.175 | 0.030 |
| **MAX** | **3.129** | **3.016** | **1.878** | **0.182** | **0.087** |
| **MIN** | **2.305** | **2.894** | **1.166** | **0.159** | **0.030** |

*C1, C2, C3, C4, and C5 are the maximum permissible salt concentrations at monitoring location ML1, ML2, ML3, ML4, and ML5 respectively

It was found that saltwater concentrations predicted by MARS meta-models within the optimization model were very close to simulation model results at all MLs. The accuracy of the developed management model was justified on the basis of percentage relative error between the MARS predicted and simulation model results as depicted in Table 5.3. The relative error was less than 5% for all the estimates. This error reflects the combined error of the S-O process resulting from the error due to training of the meta-model and the error in the optimization process itself, with an assumption that the simulation model is as accurate

as can be obtained in field conditions. Therefore, the developed saltwater intrusion management model is suitable and efficient in obtaining the optimal management strategy for combined operation of production and barrier extraction wells from a multilayered coastal aquifer system.

## 5.5 Conclusions and recommendations

This chapter presents the development and application of a methodology for multiple objective groundwater extraction management of coastal aquifer systems using a linked S-O approach. To achieve computational efficiency in the linked S-O framework, the coupled density dependent flow and salt transport numerical simulation model was replaced by properly trained and tested meta-models. For this purpose, this study illustrates the use of FIS and MARS based meta-models as useful soft computing tools in predicting salinity concentrations in a linked S-O methodology for optimal groundwater extraction strategy. These meta-model based coupled S-O approaches were solved to demonstrate the feasibility of incorporating the proposed meta-models to a CEMOGA based optimization algorithm within a linked S-O methodology, in order to determine Pareto optimal strategies for groundwater abstraction. The performance of the developed saltwater intrusion management models was evaluated using illustrative coastal aquifer study areas. Evaluation results indicated that extraction of water according to the prescribed management strategy successfully limited the salt concentrations at MLs to pre-specified limits. The results demonstrated the potential applicability of the FIS and MARS based meta-models as computationally efficient substitutes of numerical simulation model within a coupled S-O approach for obtaining saltwater intrusion management strategy for coastal aquifer systems.

The multiple objective management model provides a Pareto optimal front from which several alternative non-dominated solutions for extraction strategies can be obtained. Using any one of the Pareto optimal solutions, an optimal pumping strategy can be specified for the management period, in order to control saltwater intrusion within the permissible limits. However, selection of a single Pareto optimal solution as the prescribed preferred optimal strategy will require further preference ordering of the solutions (Datta and Peralta 1986). The optimal solution obtained by utilizing a surrogate model was verified using the actual simulation model, in which optimal pumping values derived from an optimal solution served as inputs to the simulation model. It is demonstrated for an illustrative coastal aquifer study area that extraction of water according to the prescribed management strategy can limit the salt concentrations at MLs to pre-specified limits. The results of the present study demonstrate the potential applicability of using FIS and MARS based meta-models to replace a numerical simulation model in a linked S-O methodology for developing coastal

groundwater management strategy. It was also demonstrated that a considerable amount of reduction in computation time could be achieved by implementing parallel computing platform for the solution of linked S-O methodology. The proposed methodology has the potential to improve the computational efficiency of aquifer management models, making it feasible to develop long-term sustainable optimal management strategies on a regional scale.

The present study focuses on an illustrative coastal aquifer system in which aquifer material within each layer was assumed to be homogeneous. Future research may be directed towards considering aquifer heterogeneity by using random a hydraulic conductivity field. In addition, the present study considered a parallel pool of worker machines within the local clusters of a PC for executing the FIS-CEMOGA and MARS-CEMOGA optimization routines. For more complex optimization settings, a parallel pool consisting of more PCs or any other high performance-computing platform can be used. Further evaluation using the data from a real world coastal aquifer system with saltwater intrusion problems is required to establish the wide applicability of the developed FIS and MARS based saltwater intrusion management models. Although the results presented here are limited in scope, these evaluation results may provide more generalized guidance for the optimal extraction of groundwater resources from coastal aquifers. The next chapter covers the development of management strategies incorporating uncertainties in the estimated flow and transport parameters.

# Chapter 6: Coastal aquifer management under groundwater parameter uncertainty

Partial contents of this chapter have been published and copyrighted, as outlined below:

Roy, D. K., & Datta, B. (2017c). Genetic algorithm tuned fuzzy inference system to evolve optimal groundwater extraction strategies to control saltwater intrusion in multi-layered coastal aquifers under parameter uncertainty. *Modeling Earth Systems and Environment, 3*(4), 1707-1725.

This chapter covers the development of management strategies incorporating uncertainties in the estimated flow and transport parameters.

## 6.1 Summary

The methodologies developed and applied thus far are based on deterministic values of groundwater parameters. However, the stochastic nature of the complex flow and solute transport processes in coastal aquifers demands incorporation of stochasticity of groundwater parameters in terms of uncertainties related to some of the model parameters. Multidimensional heterogeneity of aquifer properties such as hydraulic conductivity, compressibility, and bulk density are considered as major sources of uncertainty in a groundwater modelling system. Other sources of uncertainty are associated with spatial and temporal variability of hydrologic as well as human interventions, e.g. aquifer recharge and transient groundwater extraction patterns. This chapter addresses uncertainties arising from estimating hydrogeological model parameters (hydraulic conductivity, compressibility, bulk density, and aquifer recharge), and in accurately estimating spatial and temporal variation of groundwater extraction patterns. In addition, FIS based meta-models are advanced to the Genetic Algorithm (GA) tuned hybrid FIS model (GA-FIS) to emulate physical processes of coastal aquifers and to evaluate responses of the coastal aquifers to groundwater extraction under groundwater parameter uncertainty. GA was used to tune the FIS parameters in order to obtain the optimal FIS structure. The GA-FIS models thus obtained were linked externally to the Controlled Elitist Multiple Objective Genetic Algorithm (CEMOGA) in order to derive optimal pumping management strategies using a coupled S-O approach. The performance of the hybrid GA-FIS-CEMOGA based saltwater intrusion management model was compared with that of a basic Adaptive Neuro Fuzzy Inference System (ANFIS) based management model (ANFIS-CEMOGA). The parameters of the ANFIS model were tuned using a hybrid algorithm. To achieve computational efficiency, the proposed optimization routine was run on a parallel processing platform. An illustrative multilayered coastal aquifer system was used to evaluate the performances of both management models. The illustrative aquifer

system considered uncertainties associated with the hydrogeological parameters e.g. hydraulic conductivity, compressibility, bulk density, and aquifer recharge. The evaluation results show that the proposed saltwater intrusion management models are able to obtain reliable optimal groundwater extraction strategies to control saltwater intrusion for the illustrative multilayered coastal aquifer system. However, a closer look at the performance evaluation results demonstrates the superiority of the GA-FIS-CEMOGA based management model over the ANFIS-CEMOGA based saltwater intrusion management model.

## 6.2 Development of GA-FIS and ANFIS based meta-models

ANFIS meta-models were developed from an initial FIS structure whose parameters were tuned using a hybrid algorithm. This study developed GA tuned coupled FIS models and compared these FIS models with hybrid algorithm tuned ANFIS models.

### 6.2.1 ANFIS architecture

FISs are suitable for nonlinear mapping of predictor-response relationships. A Sugeno type FIS, also known as Takagi-Sugeno-Kang FIS, introduced in 1985 (Sugeno, 1985), is ideal for this nonlinear mapping. FISs utilize a set of fuzzy IF-THEN rules in establishing this predictor-response relationship. For a first order Sugeno type FIS with two inputs ($\propto$ and $\beta$), one output ($\gamma$), and two fuzzy rules, the simplest form of fuzzy IF-THEN rules are expressed as:

$$Rule\ 1: If\ \propto is\ P_1\ and\ \beta\ is\ Q_1\ then\ f_1 = p_1\alpha + q_1\beta + r_1 \quad (6.1)$$

$$Rule\ 2: If\ \propto is\ P_2\ and\ \beta\ is\ Q_2\ then\ f_2 = p_2\alpha + q_2\beta + r_2 \quad (6.2)$$

This is illustrated in Figure 6.1. Figure 6.1 shows the equivalent type-3 ANFIS architecture originated from type-3 fuzzy reasoning represented in Equations (6.1) and (6.2). The resulting ANFIS structure has five layers, namely a fuzzy layer, a product layer, a normalized layer, a defuzzification layer, and a total output layer. A detailed description of these layers can be found in Jang (1993) and is not repeated in this document.

An ANFIS allows advantages from both fuzzy logic theory and artificial neural networks. The simple structure of a Sugeno type ANFIS has a good learning capability compared to other types of ANFIS architectures (Jang et al., 1997). Desired ANFIS structures are obtained from an initial FIS structure whose parameters are optimized by using either a hybrid algorithm (Jang, 1993) or a population based optimization algorithm, e.g. GA

(Goldberg, 1989). In situations where the considered problem has a high dimensional large number of input variables, compressing the dataspace by using a suitable clustering algorithm may be of great value. Fuzzy C-Mean Clustering (FCM) (Bezdek et al., 1984) is a useful tool for compressing the dataset by dividing it into a group of identical clusters. This technique greatly reduces the number of modifiable parameters (linear and nonlinear) and fuzzy IF-THEN rules of a FIS.

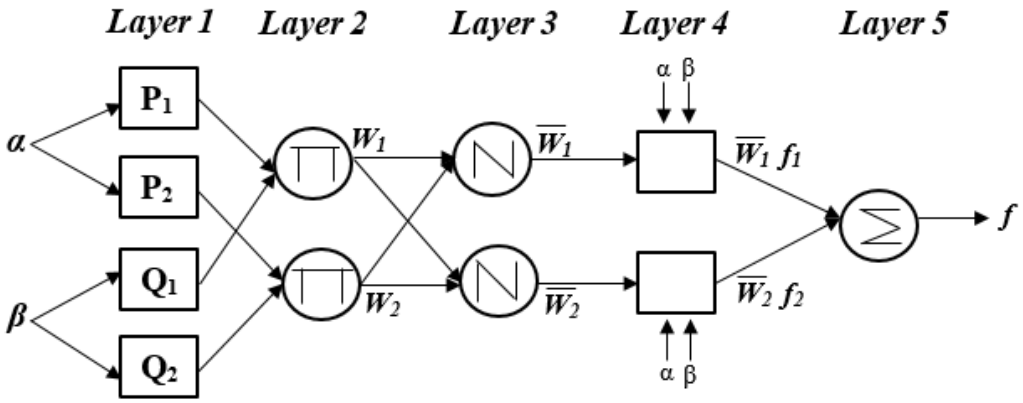


**Figure 6.1 ANFIS architecture based on a two-input first-order Sugeno FIS**

6.2.2 Training rule and algorithm for adaptive networks

The Gradient Descent (GD) and Chain Rule, proposed in early 1970s by Werbos (1974) were the basic learning rules of adaptive networks. However, this gradient based method is slow in convergence and it has the tendency to become trapped in local minima. Therefore, to speed up the learning process and to avoid premature convergence in local optima Jang (1993) proposed a hybrid learning rule (Jang, 1991), which integrates the GD approach and the Least Squares Estimate (LSE) to identify optimal parameters of adaptive networks. Each iteration of this hybrid learning approach is associated with both a forward and a backward pass. Suppose that the parameter set $S$ can be decomposed into two subsets such that $S = S1 \oplus S2$, where $\oplus$ indicates the direct summation. Then, for a given constant parameter values in $S_1$, the obtained parameter values in $S_2$ are guaranteed to be the global optimal values in the parameter space of $S_2$ (Jang, 1993). This hybrid learning algorithm not only reduces the gradient method's search space dimension, but also achieves a substantial amount of computational efficiency (Jang, 1993).

6.2.3 Optimum number of clusters

Selecting the optimum number of clusters is an important pre-processing step of the FIS model development using the FCM algorithm. The optimum right number of clusters is

decided based on the type of problem and dimension of the input space. A model with a simple architecture is preferable in most cases (Mohammadi et al., 2016). In this research, the optimum number of clusters was selected by conducting several trials using different number of clusters and observing the resulting Root Mean Squared Error (RMSE) between the actual salinity concentration values and predicted responses obtained using selected FIS models. The number of clusters that produced a minimum RMSE value as well as least variance in RMSE values between learning and testing sets of data were chosen as adequate. The lowest variance in RMSE values between training and test datasets was checked in order to protect against model overfitting. The optimum number of clusters was selected for both the Hybrid Algorithm (HA) trained ANFIS and GA trained GA-FIS models. The trial was conducted on salinity concentration values obtained at five MLs. It is noted that model architecture with two clusters provided a reasonably accurate prediction model at almost all MLs, except at Ml4, in which GA-FIS used three clusters to provide an accurate prediction model. The number of input and output Membership Functions (MF) at each ML can be expressed as:

$$N_{MF(input)} = N_X \times N_{clusters} \tag{6.3}$$

$$N_{MF(output)} = N_Y \times N_{clusters} \tag{6.4}$$

where, $N_{MF(input)}$ is the number of input Membership Functions, $N_{MF(output)}$ is the number of output Membership Functions, $N_X$ is the number of predictors, $N_Y$ is the number of responses, $N_{clusters}$ is the number of clusters. All GA-FIS and ANFIS models were developed using commands and functions of Fuzzy Logic Toolbox of MATLAB (MATLAB, 2017a).

### 6.2.4 Genetic algorithm

The working principles of GA are based on the theory of natural genetics and natural selection. The fundamental ideas of GA differ from classical optimization methods in that GAs utilize a coding of variables rather than the variables themselves (Deb, 1999). GAs use a probability based search technique to provide a population of solutions instead of a single solution. Unlike traditional optimization approaches that are based on fixed transition rules to swap between solutions, GA's search procedure begins with an initial random population. Therefore, GA's search procedure may progress in any direction and is not associated with any crucial decision at the beginning of the search process. The working principle of GA is summarized as follows:

1. Generate an initial random set of solutions.
2. Evaluate each candidate solution in relation to the underlying problem.
3. Check a termination criterion.
4. Stop search process and provide the optimum solution if termination criterion is satisfied.
5. Modify population of solutions using three main operators if termination criterion is not satisfied.

The modified population of solutions is expected to be better than the previous population of solutions. After each iteration, the generation counter keeps note of every completed generation of the GA search. Figure 6.2 shows a flowchart of the fundamental working principle of GA. As shown in the flowchart, the working principle is simple and straightforward. However, despite the fact that operations are simple, GAs are nonlinear, multifaceted, complex, and stochastic in nature (Deb, 1999).

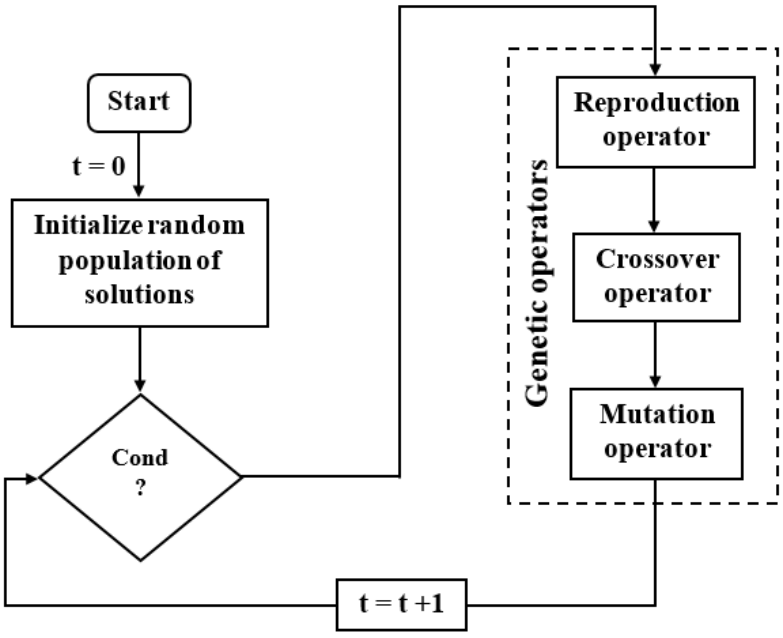


**Figure 6.2 Flowchart of the operational principles of GA**

Three basic genetic operators, e.g. reproduction, crossover, and mutation operators form the major part of the GA search mechanism. The reproduction operator identifies better (above-average) solutions in a population, creates multiple copies of these solutions, and replaces these multiple copies of solutions by eliminating worse solutions in the population. The reproduction operator does not generate any new solutions in the population. It can only create more copies of good solutions at the cost of "not-so-good" solutions. Crossover and

mutation operators perform the task of creating new solutions. In a typical crossover operator, two strings are randomly taken from the mating pool and some parts of the strings are swapped between the strings. Crossover operators are principally accountable for the search mechanism of GAs. Mutation operators are also responsible for the search aspect sparingly. The role of mutation operators in the GA search process is to maintain diversity in the population. The mutation operator executes swapping between a 1 and 0 with a small mutation probability $P_m$. The role of these three genetic operators can be summarized as follows: (a) the reproduction operator chooses good strings, (b) the crossover operator recombines good substrings from two good strings to create a relatively better string, and (c) the mutation operator swaps a string locally to form a comparatively better string (Deb, 1999). If bad strings are generated in any generation, these strings will be eliminated by the reproduction operator in subsequent future generations (Deb, 1999).

### 6.2.5 Proposed GA-FIS coupled model

One major drawback of applying fuzzy logic based modelling approaches in a high dimensional dataset is the selection of suitable rules to ensure the best performance of the model. A first order Sugeno type FIS solves this issue through modifying the rules, and learns adaptively to provide optimal sets of parameters for the FIS architecture. Tuning the parameters of antecedent and consequent parts of the rule base is usually performed using the gradient method. However, this gradient method is generally slow in convergence and is likely to become trapped in local minima. To address this issue, Jang (1993) proposed a hybrid learning rule which integrates GD and LSE to identify optimal parameters. In this hybrid algorithm, the nonlinear premise parameters are estimated by GD through error backpropagation, whereas the linear consequent parameters are estimated by recursive LSE. To achieve further improvement in the tuning process of FIS models, population based search algorithms replace the traditional search techniques. Commonly used optimization algorithms to optimize FIS parameters are GA (Ishigami et al., 1995; Zanaganeh et al., 2009), Particle Swarm Optimization (PSO) (Oliveira and Schirru, 2009; Rini et al., 2016), and Cuckoo search (Araghi et al., 2015). These populations based hybrid FIS learning techniques have been applied in different research domains, e.g. in traffic signal control (Araghi et al., 2015), in sensor monitoring (Oliveira and Schirru, 2009), in determining optimum parameters of a protective spur dike (Basser et al., 2015) and in determining spatiotemporal groundwater quality parameters (Jalalkamali, 2015). In this study, for the first time, this hybrid learning technique executed by using GA is proposed to develop FIS models in approximating saltwater intrusion processes in a multilayered coastal aquifer system under the influence of spatiotemporal groundwater extraction values. In this work, GA-FIS hybrid

models were developed to predict salinity concentrations at specified ML situated in the aquifer. The GA-FIS models had 80 inputs and 1 output.

Spatiotemporal groundwater extraction values from a set of production bores and barrier extraction wells were used as the inputs to the GA-FIS models. The outputs were the resultant salinity concentrations observed at specified MLs at the end of the simulation period of 5 years. Both antecedent and consequent parameters were tuned using GA to obtain the optimal FIS model architecture. During the training phase, GA evaluated different parameters until it obtained the optimal parameters for the FIS models. The cost function of the GA optimization approach was the Mean Squared Error (MSE) values that reflect the training error. Therefore, the objective was to minimize the MSE values between the targets (actual) and FIS outputs (predicted) on the training dataset. The cost function is represented as:

$$Minimize, f_{MSE} = \frac{\sum_{i=1}^{n}\left(C_{i,o} - C_{i,p}\right)^2}{n} \tag{6.5}$$

where, $f_{MSE}$ is the cost function to be minimized, $i = 1, 2, \ldots, n$ is the number of training data, $C_{i,o}$ are the observed salinity concentration values in the training dataset, and $C_{i,p}$ is the predicted salinity concentrations in the training dataset.

Optimal GA parameters were determined by conducting a set of trials using different combinations of these parameters. Based on these trials, the optimal combination of GA parameter sets were estimated as: population size 200, crossover fraction 0.90, mutation fraction 0.85, mutation rate 0.005, and selection pressure 8. These parameter values (GA parameters) were obtained through a trial and error procedure, and were regarded as optimum values at least for this example problem.

Once the best GA-FIS model structures with optimized parameters were obtained, the developed models were presented with a new set of test data, and testing errors (RMSE values) were computed. Five GA-FIS models were developed at five MLs. The performances of the proposed GA-FIS hybrid models were compared with those of the basic ANFIS models. In ANFIS models, the antecedent parameters were optimized using GD through error backpropagation and the consequent parameters were tuned using a recursive LSE approach.

## 6.3 Application of the proposed methodology

An illustrative multilayered coastal aquifer system with a set of uncertain model parameters was used to evaluate the performance of the proposed methodology. Hydraulic conductivity, compressibility, and bulk density were assumed to be homogeneous but uncertain within each vertical layer of materials. Different realizations of these uncertain model parameters

were used in each vertical material layer. Different realizations of compressibility and aquifer recharge were also used. Aquifer recharge was uniformly spread over the top layer of the aquifer. These uncertain model parameters were randomly paired with each other and combined with the transient groundwater extraction values obtained from a set of production bores and barrier extraction wells. These input values, along with other initial and boundary conditions, were used as inputs to the numerical simulation model in order to obtain salinity concentrations at specified MLs at the end of the management period. The multilayered coastal aquifer system considered was similar to one presented in Subsection 4.2.5 of chapter 4.

Different realizations of these uncertain model parameters were obtained from different statistical distributions. A representative set of hydraulic conductivity realizations were obtained from a lognormal distribution with a specific mean and standard deviation of the associated normal distribution (Table 6.1).

**Table 6.1 Parameter distributions with mean and standard deviation values used in simulation**

| Parameters | Unit | Material layer | Distribution | Mean | Standard deviation |
|---|---|---|---|---|---|
| Hydraulic conductivity in x-direction | m/d | 1 | lognormal | 5.02 | 0.30 |
| Hydraulic conductivity in y-direction | m/d | 1 | lognormal | 2.52 | 0.15 |
| Hydraulic conductivity in z-direction | m/d | 1 | lognormal | 0.50 | 0.03 |
| Compressibility | $md^2/kg$ | 1 | uniform (LHS) | $7.37\times10^{-16}$ | $3.50\times10^{-16}$ |
| Bulk density | $kg/m^3$ | 1 | normal (LHS) | 1650 | 5 |
| Recharge | m/d | 1 | uniform (LHS) | 0.00019 | $3.48\times10^{-5}$ |
| Hydraulic conductivity in x-direction | m/d | 2 | lognormal | 9.97 | 0.28 |
| Hydraulic conductivity in y-direction | m/d | 2 | lognormal | 4.98 | 0.14 |
| Hydraulic conductivity in z-direction | m/d | 2 | lognormal | 1.00 | 0.03 |
| Compressibility | $md^2/kg$ | 2 | uniform (LHS) | $7.35\times10^{-18}$ | $3.49\times10^{-18}$ |
| Bulk density | $kg/m^3$ | 2 | normal (LHS) | 1600 | 5 |
| Hydraulic conductivity in x-direction | m/d | 3 | lognormal | 14.95 | 0.29 |
| Hydraulic conductivity in y-direction | m/d | 3 | lognormal | 7.47 | 0.15 |
| Hydraulic conductivity in z-direction | m/d | 3 | lognormal | 1.49 | 0.03 |
| Compressibility | $md^2/kg$ | 3 | uniform (LHS) | $7.35\times10^{-18}$ | $3.49\times10^{-18}$ |
| Bulk density | $kg/m^3$ | 3 | normal (LHS) | 1550 | 5 |
| Hydraulic conductivity in x-direction | m/d | 4 | lognormal | 3.05 | 0.30 |
| Hydraulic conductivity in y-direction | m/d | 4 | lognormal | 1.53 | 0.15 |
| Hydraulic conductivity in z-direction | m/d | 4 | lognormal | 0.31 | 0.03 |
| Compressibility | $md^2/kg$ | 4 | uniform (LHS) | $7.35\times10^{-17}$ | $3.49\times10^{-17}$ |
| Bulk density | $kg/m^3$ | 4 | normal (LHS) | 1700 | 5 |

*Recharge is distributed uniformly over the first layer of the study area

Lognormal distribution is a probability distribution in which the uncertain model parameter hydraulic conductivity is divided into $N$ equally probable intervals from which a single value is chosen randomly. Aquifer recharge and compressibility realizations were generated from LHS uniform distributions for specific lower and upper bounds within the parameter space. Realizations of bulk density were obtained from LHS technique from a $p$-dimensional multivariate normal distribution with specific mean and covariance. The values of hydraulic conductivity, recharge, compressibility, and bulk density were then shuffled randomly and combined to obtain multivariate random realizations of uncertain model parameters. One hundred realizations of each parameter were generated for each material layer. A total number of 3000 uniformly distributed groundwater extraction values were generated from the variable space, with a range of 0-1300 $m^3$/day. For each randomized uncertain model parameter set, 30 sets of transient pumping values at the well locations were assigned. The 3000 randomized combined realizations of uncertain model parameters and transient groundwater extraction values obtained were then used along with other initial and boundary conditions as inputs to the simulation model in order to obtain the corresponding salinity concentrations at specified MLs.

### 6.3.1 Performances of the GA‐FIS and ANFIS models

Performances of the GA-FIS models to approximate density dependent coupled flow and salt transport processes were evaluated based on a set of different statistical indices. Each developed GA-FIS model was utilized to predict salinity concentrations at a specified ML. Performances of GA-FIS models were compared with those of basic ANFIS models. Results are summarized in Table 6.2, Table 6.3, Table 6.4, Table 6.5, Figure 6.3, and Figure 6.4.

**Table 6.2 Performance of the proposed GA-FIS and ANFIS models on training and testing phase**

| MLs | Training and Test RMSE, mg/l | | | | | |
|---|---|---|---|---|---|---|
| | ANFIS | | | GA-FIS | | |
| | Train RMSE | Test RMSE | Difference | Train RMSE | Test RMSE | Difference |
| ML1 | 0.81 | 1.12 | 0.30 | 0.93 | 1.06 | 0.13 |
| ML2 | 48.55 | 64.47 | 15.92 | 53.28 | 56.59 | 3.31 |
| ML3 | 63.10 | 85.39 | 22.28 | 70.09 | 76.42 | 6.33 |
| ML4 | 221.40 | 300.66 | 79.26 | 245.09 | 269.39 | 24.30 |
| ML5 | 187.86 | 245.02 | 57.16 | 207.03 | 223.70 | 16.67 |

Table 6.2 presents the training and testing RMSE values as well as the differences between these two RMSE values at all MLs. In all cases, ANFIS models produced smaller

training errors than GA-FIS models. However, ANFIS models produced higher testing errors than GA-FIS at all MLs. As such, the difference in errors between training and testing datasets was smaller in GA-FIS than in ANFIS models. Therefore, GA-FIS models seem to be more robust and are likely to provide smaller errors when presented with a completely new set of unseen data.

**Table 6.3 Performance of the proposed ANFIS models on test dataset**

| MLs | RMSE, mg/l | | MAPRE, % | | IOA (6.6) | | R | |
|---|---|---|---|---|---|---|---|---|
| | ANFIS | GA-FIS | ANFIS | GA-FIS | ANFIS | GA-FIS | ANFIS | GA-FIS |
| ML1 | 1.12 | 1.06 | 2.69 | 2.58 | 0.92 | 0.92 | 0.85 | 0.87 |
| ML2 | 64.47 | 56.59 | 6.19 | 5.51 | 0.89 | 0.91 | 0.80 | 0.84 |
| ML3 | 85.39 | 76.42 | 7.03 | 6.40 | 0.91 | 0.93 | 0.84 | 0.87 |
| ML4 | 300.66 | 269.39 | 4.96 | 4.50 | 0.79 | 0.82 | 0.65 | 0.71 |
| ML5 | 245.02 | 223.70 | 3.66 | 3.38 | 0.81 | 0.84 | 0.67 | 0.73 |

*MAPRE = Mean Absolute Percentage Relative Error, IOA = Index of Agreement

Performances of both GA-FIS and ANFIS models based on RMSE, MAPRE, IOA and R criteria are presented in Table 6.3. It is observed form Table 6.3 that both models produced relatively lower values of RMSE and MAPRE, and higher values of IOA and R. Although both GA-FIS and ANFIS models were sufficiently accurate in predicting the responses, GA-FIS exhibited relatively better performance than ANFIS models at all MLs. Therefore, it is concluded that the GA-FIS models' prediction accuracy in terms of capturing the trends of responses at different regions is quite satisfactory based on RMSE criteria. However, the drawback of RMSE criterion is that it provides more weights to the outlying observations. Therefore, MAPRE criteria were used to provide comparatively better information on the distribution of errors. MAPRE criteria also demonstrated the better performance of GA-FIS compared to ANFIS models. Developed models were also evaluated from the correlation coefficient viewpoint. All GA-FIS models produced higher values of R compared to their ANFIS counterparts.

However, R values at MLs 4 and 5 were comparatively lower (close to 70%). Therefore, IOA (Willmott, 1984) was proposed to ascertain the proposed model's prediction capability as well as to overcome the insensitivity of R to differences in the actual and predicted means and variances (Legates and McCabe, 1999). Both models produced reasonable and acceptable values of IOA at all MLs. Performances of GA-FIS models were also better based on the IOA criterion. IOA criterion is calculated as

$$d = 1 - \frac{\sum_{i=1}^{n}\left(C_{i,o} - C_{i,p}\right)^2}{\sum_{i=1}^{n}\left(\left|C_{i,p} - \bar{C}_o\right| + \left|C_{i,o} - \bar{C}_o\right|\right)^2} \quad (6.6)$$

where, $C_{i,o}$ = observed salinity concentration values, $C_{i,p}$ = predicted salinity concentration values, and $\bar{C}_o$ = mean values of the observed salinity concentrations.

Figure 6.3 illustrates actual versus predicted salinity concentrations at ML1 as an example. For brevity of presentation, only the first 30 samples out of a total number of 600 test datasets are presented in Figure 6.3. For this performance evaluation purpose, the actual concentrations were those synthetically obtained as solution of the numerical simulation model in response to water abstraction from the aquifer. The predicted concentrations denote the concentrations predicted by the meta-models. It is observed from Figure 6.3 that both GA-FIS and ANFIS models predictions were very similar to those of the actual salinity concentrations. Both models were able to capture the trends of data quite accurately. However, GA-FIS models provided relatively better predictions than ANFIS models.

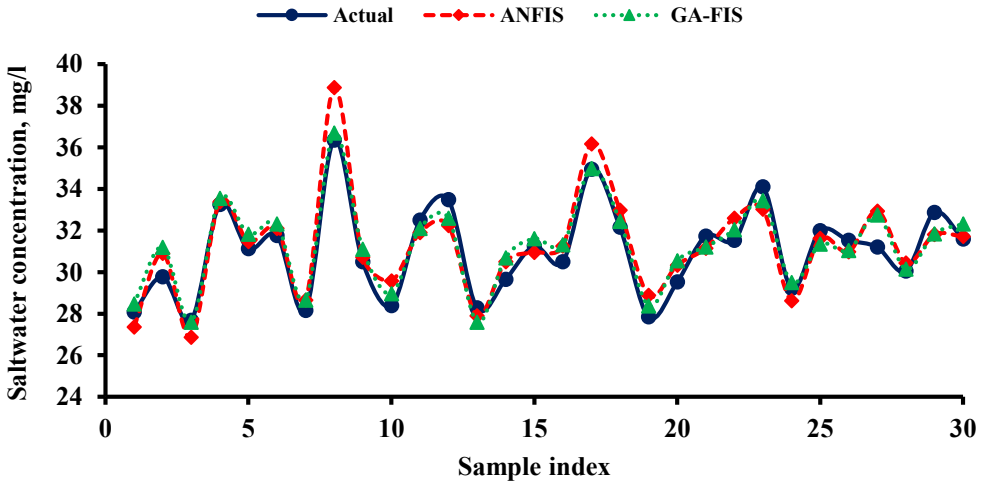


**Figure 6.3 Actual and predicted salinity concentrations at ML1**

Figure 6.4 illustrates boxplots of absolute errors between the actual and predicted salinity concentration values obtained by GA-FIS and ANFIS models at different MLs. In Figure 6.4, the red and blue horizontal lines indicate medians of the absolute errors produced by ANFIS and GA-FIS models, respectively. The mean of absolute errors by both models is represented by small black circles. Figure 6.4 also demonstrates the superiority of GA-FIS models over ANFIS models at all MLs based on absolute error viewpoint.

Performances of the proposed models were also evaluated based on the TS criterion, which provides distribution of errors. TS criterion is an indication of the percentage of sample indices whose Relative Error (RE) values are smaller than the pre-defined threshold values.

Four threshold values (<5%, <10%, <15%, and <20%) were used in the present study so that RE values obtained fell within the selected threshold values. It is apparent from Table 6.4 that GA-FIS models outperformed ANFIS models in terms of TS criterion.

**Table 6.4 Threshold statistics between the actual and predicted saltwater concentration values on the test dataset**

| MLs | Threshold statistics, % | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | <5% | | <10% | | <15% | | <20% | |
| | ANFIS | GA-FIS | ANFIS | GA-FIS | ANFIS | GA-FIS | ANFIS | GA-FIS |
| ML1 | 89.50 | 91.00 | 98.50 | 99.17 | 100.0 | 100.0 | 100.0 | 100.0 |
| ML2 | 46.83 | 46.67 | 81.33 | 87.33 | 95.83 | 99.17 | 99.33 | 100.0 |
| ML3 | 39.50 | 42.83 | 73.50 | 79.00 | 92.17 | 97.17 | 98.83 | 99.83 |
| ML4 | 54.33 | 57.00 | 92.00 | 96.50 | 99.17 | 99.67 | 99.83 | 100.0 |
| ML5 | 71.83 | 76.50 | 98.00 | 99.67 | 100.0 | 100.0 | 100.0 | 100.0 |

Another important variable that should be considered while evaluating the performance of any prediction model is the computational time required to train the model. Table 6.5 presents training time requirement in the development of both GA-FIS and ANFIS models. It is noted that the difference in training time requirement between different MLs by both GA-FIS and ANFIS models was not very substantial. However, at all MLs, GA-FIS required more time to train. This is because GA performs a thorough search in order to provide an optimal set of parameter values. However, once trained, these models are able to provide prediction results very quickly (fraction of a minute). Therefore, it is concluded that GA-FIS provides relatively better approximations of the coupled flow and salt transport processes than ANFIS models, at least based on the limited evaluations presented here.

**Table 6.5 Training time requirement (min)**

| Models | ML1 | ML2 | ML3 | ML4 | ML5 |
|---|---|---|---|---|---|
| ANFIS | 2.77 | 2.74 | 2.70 | 2.76 | 2.72 |
| GA-FIS | 12.20 | 11.28 | 10.76 | 11.60 | 16.98 |

*MLs = Monitoring Locations

### 6.3.2 Performance of the management model

Saltwater intrusion management models were developed by integrating GA-FIS and ANFIS models separately with a population based multiple objective optimization algorithm,

CEMOGA. The proposed management models provide an optimal solution of groundwater abstraction in the form of Pareto optimal fronts that show the trade-offs between the two conflicting objectives of groundwater abstraction. The optimization routine was run on a parallel computing platform by distributing the objective functions and the constraints among four physical cores of a PC. The performance of the GA-FIS-CEMOGA based management model was compared with that of the ANFIS-CEMOGA based management model. The GA-FIS-CEMOGA model evaluates 1,804,600 functions (1289 generations × 1400 populations) to decide on the global optimal solution.

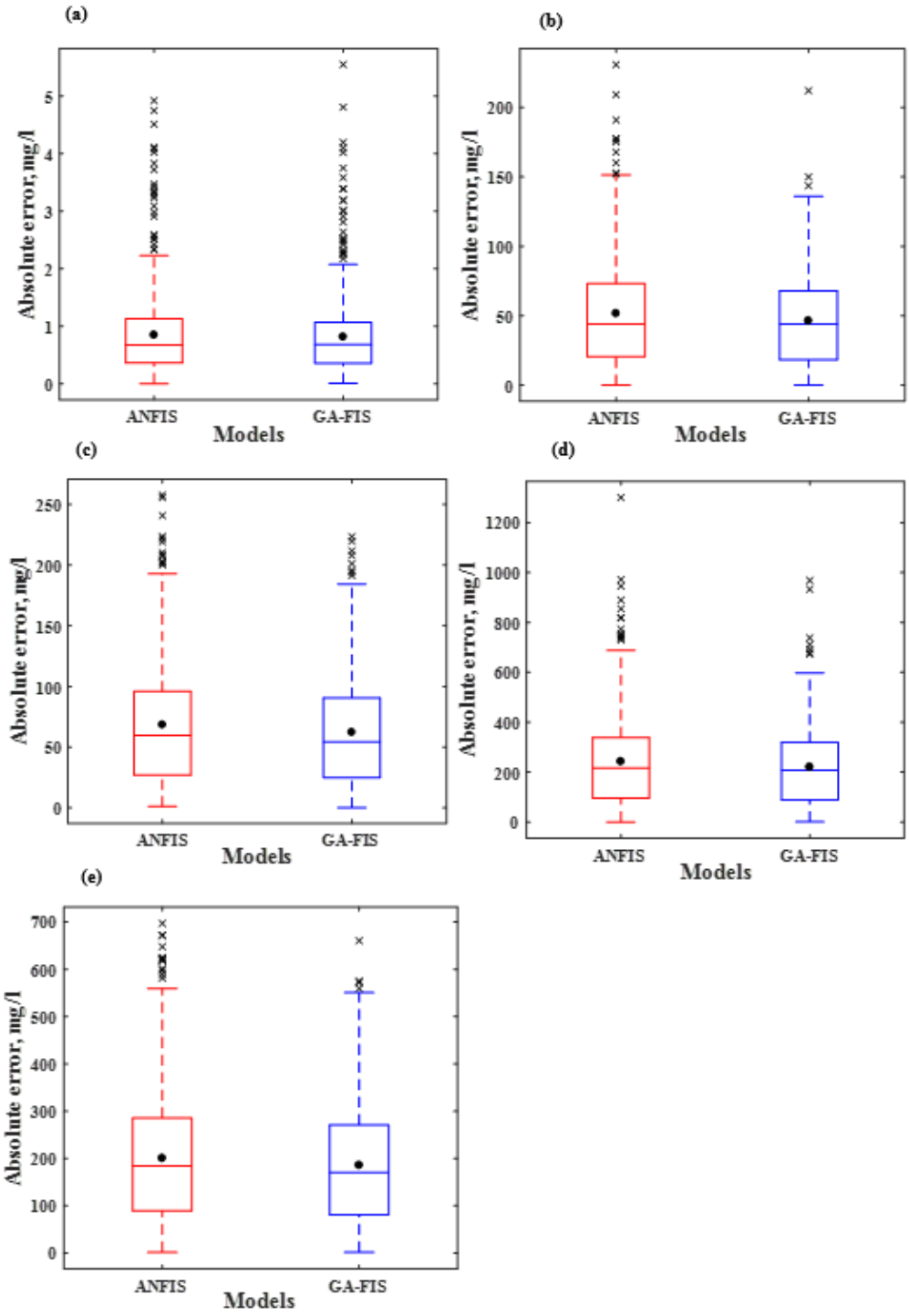


**Figure 6.4 Box plots of absolute errors between actual and predicted saltwater concentration values at (a) ML1, (b) ML2, (c) ML3, (d) ML4, and (e) ML5**

On the other hand, ANFIS-CEMOGA model performed 2,018,800 (1442 generations × 1400 populations) function evaluations before reaching the global Pareto optimal solution. CEMOGA parameters were selected by conducting a set of numerical experiments using a variation of several combinations of different parameters. Based on these numerical trials, the CEMOGA used a population size of 1400, crossover rate of 0.95, and

Pareto front population fraction of 0.7. The function and constraint tolerances were set as $1\times10^{-5}$ and $1\times10^{-3}$, respectively. The optimal groundwater extraction strategy, in the form of a Pareto optimal front, is presented in Figure 6.5. The Pareto front provides 980 non-dominated solutions, from which the managers can choose the right combination of production and barrier well pumping. These solutions are based on limiting the salinity concentrations at specified MLs to the pre-defined maximum allowable limits.

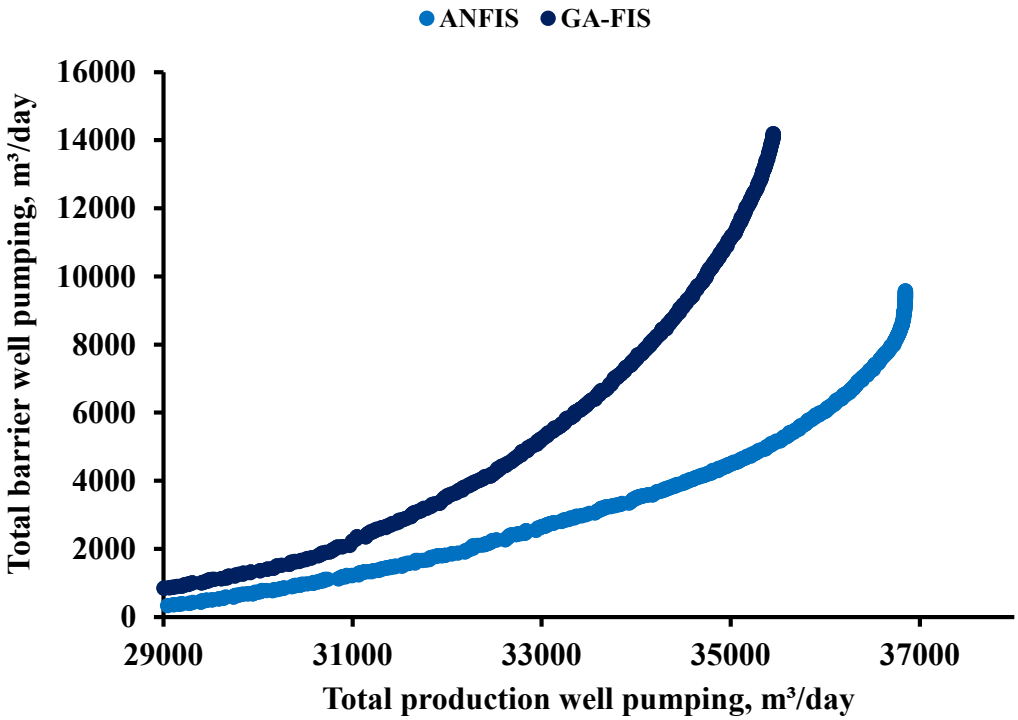


**Figure 6.5 Pareto optimal front of the developed management model**

The total amount of water abstraction from both the production bores and barrier extraction wells are presented in Table 6.6. Ten solutions were selected randomly from different regions of the Pareto optimal front. It is observed from Table 6.6 that ANFIS models seem to be more efficient in terms of total amount of beneficial production well pumping with respect to barrier extraction well pumping. However, in order to reach a conclusion, verification of the performance of both of these models is required. This verification was conducted by running the numerical simulation model using these pumping values, and by checking the constraint violations as well as the errors in prediction.

**Table 6.6 Optimal solutions obtained using GA-FIS and ANFIS models**

| Solutions | Total pumping × $10^3$ $m^3$/day (GA-FIS) | | Total pumping × $10^3$ $m^3$/day (ANFIS) | |
|---|---|---|---|---|
| | Production well | Barrier well | Production well | Barrier well |
| 1 | 35.43 | 13.86 | 34.88 | 4.37 |
| 2 | 35.44 | 14.05 | 33.63 | 3.18 |
| 3 | 35.45 | 14.19 | 34.02 | 3.51 |
| 4 | 35.37 | 13.39 | 33.97 | 3.45 |
| 5 | 33.59 | 6.50 | 34.04 | 3.53 |
| 6 | 35.34 | 13.10 | 35.02 | 4.52 |
| 7 | 35.43 | 13.83 | 33.67 | 3.21 |
| 8 | 34.27 | 8.45 | 31.82 | 1.74 |
| 9 | 32.95 | 5.16 | 34.00 | 3.48 |
| 10 | 35.44 | 13.96 | 34.33 | 3.78 |

### 6.3.3 Verification of the management model

The performance of the developed saltwater intrusion management models was verified by comparing the solution results obtained from the optimization routine with those obtained from the numerical simulation model. To do this, ten solutions of optimal groundwater extraction values (which were same as the solutions used in Table 6.6) were randomly selected from different regions of the Pareto optimal front. These solutions were used as inputs to the simulation model, developed by using the average values of uncertain model parameters in order to obtain the corresponding saltwater concentration values at each ML.

**Table 6.7 Percentage absolute relative error between optimal solutions and simulated salinity concentrations**

| Solutions | PARE, % (GA-FIS) | | | | | PARE, % (ANFIS) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ML1 | ML2 | ML3 | ML4 | ML5 | ML1 | ML2 | ML3 | ML4 | ML5 |
| 1 | 0.04 | 2.16 | 4.01 | 0.11 | 0.49 | 0.90 | 0.65 | 12.36 | 2.35 | 6.39 |
| 2 | 0.05 | 2.13 | 3.97 | 0.07 | 0.46 | 0.28 | 0.73 | 10.04 | 2.11 | 5.16 |
| 3 | 0.06 | 2.09 | 3.96 | 0.03 | 0.45 | 0.36 | 0.10 | 11.00 | 2.27 | 5.61 |
| 4 | 0.01 | 2.24 | 4.08 | 0.20 | 0.54 | 0.32 | 0.31 | 10.51 | 2.19 | 5.42 |
| 5 | 0.60 | 2.25 | 3.56 | 0.87 | 0.84 | 0.29 | 0.28 | 11.04 | 2.27 | 5.55 |
| 6 | 0.01 | 2.31 | 4.15 | 0.28 | 0.59 | 0.91 | 0.72 | 12.43 | 2.54 | 6.45 |
| 7 | 0.04 | 2.17 | 4.02 | 0.12 | 0.50 | 0.25 | 0.51 | 10.52 | 2.11 | 5.30 |
| 8 | 0.35 | 2.44 | 3.85 | 0.75 | 0.85 | 0.77 | 1.56 | 7.84 | 2.50 | 3.54 |
| 9 | 0.80 | 2.12 | 3.20 | 1.02 | 0.87 | 0.27 | 0.18 | 11.13 | 2.17 | 5.62 |
| 10 | 0.04 | 2.14 | 3.99 | 0.09 | 0.48 | 0.46 | 0.04 | 11.42 | 2.32 | 5.73 |

*PARE = Percentage Absolute Relative Error

Table 6.7 represents the Percentage Absolute Relative Errors (PARE) between the management models' predicted saltwater concentration values and simulation models' output. It is observed from Table 6.7 that PARE values were less than 5% for all estimates for the GA-FIS based management model. On the other hand, the ANFIS based management model produced relatively higher values of PARE, especially at MLs 3 and 5. It is noted that the ANFIS based management model produced lower PARE values at ML2 compared to the GA-FIS based management model. Therefore, this example problem demonstrates the superiority of the GA-FIS based saltwater intrusion management model, based on the PARE viewpoint.

**Table 6.8 Absolute values of constraint violations at selected monitoring locations**

| Solution | ML1 (< 30 mg/l) | | ML2 (< 700 mg/l) | | ML3 (< 700 mg/l) | | ML4 (< 4500 mg/l) | | ML5 (< 5000 mg/l) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | GA-FIS | ANFIS | GA-FIS | ANFIS | GA-FIS | ANFIS | GA-FIS | ANFIS | GA-FIS | ANFIS |
| 1 | 0.90 | 0.77 | 0.00 | 0.87 | 0.67 | 0.51 | 0.76 | 0.22 | 0.01 | 14.52 |
| 2 | 0.89 | 0.79 | 0.27 | 0.20 | 0.15 | 0.04 | 1.50 | 6.35 | 0.21 | 5.80 |
| 3 | 0.87 | 0.79 | 0.47 | 1.64 | 0.00 | 0.74 | 1.22 | 3.07 | 0.51 | 12.65 |
| 4 | 0.97 | 0.85 | 0.06 | 0.60 | 0.81 | 0.24 | 5.68 | 1.78 | 0.95 | 8.59 |
| 5 | 1.73 | 0.82 | 0.03 | 0.48 | 1.65 | 1.39 | 0.25 | 3.27 | 0.60 | 7.07 |
| 6 | 1.02 | 0.81 | 0.48 | 0.45 | 0.15 | 0.17 | 12.17 | 3.58 | 1.25 | 12.91 |
| 7 | 0.91 | 0.78 | 0.04 | 1.61 | 0.48 | 3.29 | 3.82 | 4.60 | 0.11 | 10.91 |
| 8 | 1.60 | 0.88 | 0.07 | 0.58 | 0.88 | 12.91 | 6.62 | 0.50 | 2.74 | 5.28 |
| 9 | 1.82 | 0.84 | 1.74 | 0.58 | 2.01 | 0.92 | 2.37 | 2.64 | 0.35 | 11.49 |
| 10 | 0.90 | 0.83 | 0.05 | 0.03 | 0.22 | 0.43 | 1.91 | 1.56 | 0.27 | 5.33 |

Performances of the management models were also verified by comparing the violation of the constraints at each ML. This is presented in Table 6.8. Constraint violations by the ANFIS models were higher for most of the estimates. At ML5, ANFIS models provided a considerable amount of constraint violations.

Based on the above discussions, it can be concluded that the proposed saltwater intrusion management models can be applied to obtain optimal groundwater abstraction values in a multilayered coastal aquifer system under parameter uncertainty. However, GA-FIS based meta-models are more efficient in developing the saltwater intrusion management model, at least for this illustrative example problem.

## 6.4 Summary and conclusions

This chapter presents the use of a coupled GA-FIS-CEMOGA based management model to predict saltwater intrusion processes and to develop a regional scale optimal management strategy in a multilayered coastal aquifer system under hydrogeological parameter uncertainty. The performance of the GA-FIS-CEMOGA model was compared with the performance of an ANFIS-CEMOGA based management model. In the first step, the GA-FIS and ANFIS based models were developed to approximate density dependent coupled flow and salt transport processes. GA was used to tune the antecedent and consequent parameters for obtaining the best FIS model structures. The parameters of ANFIS based

models were tuned using the hybrid algorithm. Both models were trained using datasets consisting of predictor-response arrays of groundwater withdrawal and resultant saltwater concentrations, obtained as solutions of a density dependent 3-D coupled flow and salt transport numerical simulation model. In the second step, two management models were developed by external linking of these meta-models to the optimization algorithm, CEMOGA. It was demonstrated that both GA-FIS and ANFIS models captured the nonlinear relationship between the spatiotemporal groundwater extraction values and the resulting salinity concentrations quite accurately. Comparison results indicated that the GA-FIS models provided better prediction capabilities compared to the ANFIS models. This study applied GA for extracting fuzzy IF-THEN rules in the FCM method in order to develop the hybrid GA-FIS model.

The management model performance evaluation result also indicated the superiority of the GA-FIS-CEMOGA based model. It is noted that comparison of Pareto fronts of both the GA-FIS-CEMOGA and the ANFIS-CEMOGA based management models appear to indicate that the ANFIS-CEMOGA model allows increased amount of beneficial pumping associated with the same amount of barrier well extraction. However, this apparently permissible increased amount of pumping corresponding to a specific amount of barrier well extraction appears to be the consequence of relatively inaccurate prediction of salinity by the ANFIS-CEMOGA, and not due to better optimal solutions. Based on the PARE criterion, it can be concluded that GA-FIS-CEMOGA based management model provided more accurate estimates of the optimal groundwater extraction strategies and therefore provided more accurate representation of the optimal permissible pumping. This is a very important consideration in all real life applications.

The proposed GA-FIS models did not include parameter uncertainty directly. However, these models indirectly incorporated uncertainty because they were developed from the solution results of a numerical model that addressed parameter uncertainty. The present study considered a geologically multilayered coastal aquifer system in which four vertical material layers varied in randomized realizations of uncertain model parameters. However, the aquifer materials within each layer are considered homogeneous but uncertain. Future research might profitably be directed towards studying the applicability of the GA-FIS based approximation of the complex saltwater intrusion process for developing optimal management strategies in heterogeneous coastal aquifer systems. The next chapter presents the ensemble meta-modelling approach for addressing the prediction uncertainties of meta-models.

# Chapter 7: Management of coastal aquifers using ensemble-modelling approach to address prediction uncertainty of meta-models

Partial contents of this chapter have been published and copyrighted, as outlined below:

Roy, D. K., & Datta, B. (2017b). Multivariate adaptive regression spline ensembles for management of multilayered coastal aquifers. *Journal of Hydrologic Engineering, ASCE 22*(9), 04017031.

This chapter covers management of coastal aquifers using ensemble meta-modelling approach to address prediction uncertainty of meta-models.

## 7.1 Summary

The approaches developed and presented in the previous chapter focused on developing saltwater intrusion management model by using artificial intelligence based meta-models linked to an optimization algorithm. These computationally efficient replacements for the numerical simulation model add a certain amount of uncertainty to the predicted variable. This predictive uncertainty results from the residuals and affects the optimality and feasibility of the Pareto optimal solution of a coupled S-O model. One way to minimize this prediction uncertainty is to use an ensemble of several models in which the outputs from a group of separately trained models are combined to make one integrated prediction. Ensemble of meta-models recompenses prediction errors by integrating the outputs from all individual models of the ensemble, and is able to achieve better prediction than individual models. This approach is also advantageous because an individual meta-model often fails to map the true pattern of data from the entire parameter space. However, individual models of an ensemble should have sufficient accuracy in terms of future prediction capability and adequate diversity in terms of representing different regions of the input space. Accuracy can be achieved by the proper choice of meta-models, and by adjusting the optimal combination of model parameters. Diversity of the individual models can be achieved by several approaches, such as optimizing the model parameters and structures, varying the training algorithm, and using different realizations of the training dataset. In this study, individual ensemble members were built by using different realizations of the training dataset obtained by random sampling without replacement approach.

This chapter presents an ensemble of MARS (En-MARS) based meta-models within a coupled S-O methodology to develop multiple objective optimal groundwater extraction strategies in a multilayered coastal aquifer system. A 3-D density dependent coupled flow

and salt transport numerical simulation code FEMWATER was used to generate the training patterns of groundwater extraction strategies and the resulting saltwater concentrations. The prediction capability of En-MARS was compared with that of the best MARS model in the ensemble. Following the development of the En-MARS meta-models, they were linked to an optimization algorithm within a regional scale saltwater intrusion management model which was employed to address two conflicting objectives of groundwater extraction. The optimal solutions obtained from the En-MARS models were verified by running the numerical simulation model. The results indicated that MARS based ensemble modelling approach is able to provide reliable solutions for a multilayered coastal aquifer management problem. The adaptive nature of MARS models, the use of ensembles, and the utilization of parallel processing resulted in a computationally efficient, accurate and reliable methodology for coastal aquifer management that also incorporates uncertainties in modelling.

## 7.2 Methodology

The proposed saltwater intrusion management model integrates two main components: ensemble based meta-models to simulate the physical processes of a multilayered coastal aquifer system and an optimization model to obtain optimal groundwater extraction strategies constrained by pre-specified saltwater concentration limits. The MARS based meta-models were trained from the input-output datasets generated by a 3-D density dependent coupled flow and salt transport numerical simulation model FEMWATER (Lin et al., 1997). Different realizations of the input-output datasets obtained using a random sampling without replacement approach (Hastie et al., 2008) forming the entire datasets were used to train different MARS meta-models that formed the ensemble. The En-MARS models were then linked externally to the CEMOGA for deriving optimal groundwater extraction strategies.

### 7.2.1 Ensemble meta-models

Various methods have been proposed in the literature to generate ensemble members, for instance, by using different model architectures, by adopting different training algorithms, or by manipulating the training data set (Sharkey, 1999; Shu and Ouarda, 2008). The basic idea behind the development of an ensemble of meta-models for future prediction is to use each model's unique feature for capturing different patterns within the data set (Khashei and Bijari, 2011). Among other approaches to generate ensemble members, altering the training data using resampling techniques often provides better results (Shu and Burn, 2004; Zaier et al., 2010). The present study adopts different realizations of training data sets to build individual MARS meta-models. Figure 7.1 provides the flow diagram of the ensemble generation process using MARS meta-models.

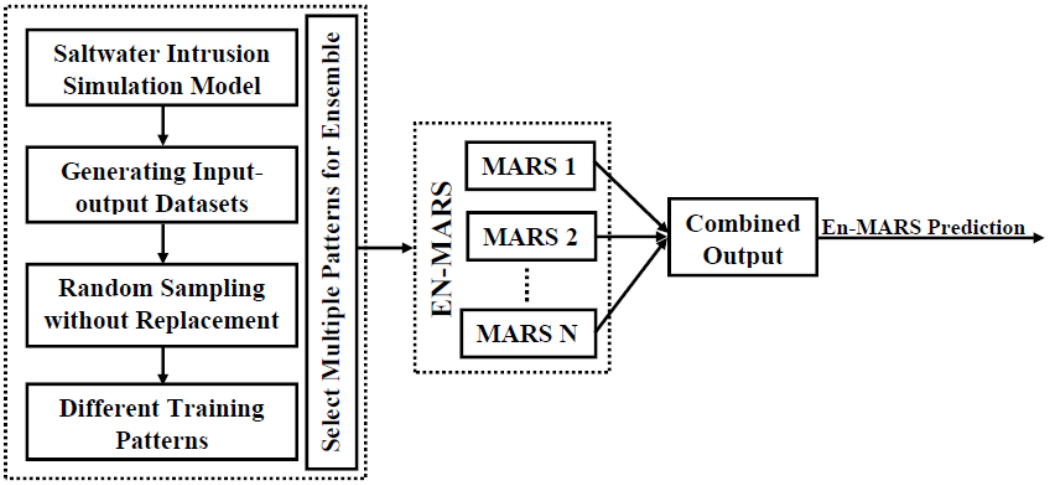


**Figure 7.1 Flow diagram of the ensemble based meta-model**

Several previous studies have used different numbers of individual meta-models to build ensembles. For instance, a good generalization capability was achieved by using 10-20 individual models in the ensemble (Shu and Burn, 2004; Zaier et al., 2010). However, Sreekanth and Datta (2011b) used 30 individual models to form an ensemble of GP models by sequentially adding models to the ensemble and checking the resulting root mean square error. Khalil et al. (2015) evaluated 2-20 ANN models to predict groundwater levels under conditions of mine-tailings. In the present study, sufficient numbers of ensemble members for each ML were selected initially. Then individual models were added sequentially to the ensemble, and the optimum ensemble size was selected by observing the resulting RMSE of the ensemble.

To combine outputs from the individual members of the ensemble, the most commonly used techniques are the average and stacked generalization (Shu and Ouarda, 2007; Wolpert, 1992). In this study, a simple averaging technique was used to combine the output from the ensemble members of individual MARS meta-models. To obtain the average of the MARS ensemble output, the output of the individual models was added, and the result was divided by the number ($n$) of MARS models. Uniform weighting was applied to calculate the output from the En-MARS models, which is calculated using the following equation:

$$OutEnMARS = \sum_{i=1}^{n} \frac{OutMARS_i}{n} \tag{7.1}$$

where, $OutEnMARS$ is the combined output from the ensemble of MARS meta-models and $OutMARS_i$ is the output from the $i^{th}$ MARS meta-models.

### 7.2.2 Random sampling without replacement

Different realizations of the actual input-output data sets were generated using the random sampling without replacement technique (Hastie et al., 2008). Each realization of these data sets was then used to train individual MARS meta-models to be used later for building En-MARS models. Because the training data sets were sufficiently different from each other, the resulting trained MARS meta-models were distinctly different from each other. As the final values of model parameters are dependent on the training data set, their values have the potential to differ slightly among MARS meta-models trained on different training data sets.

An input-output pattern $T$ of size $N$ was obtained from LHS and the numerical simulation model. For generating different realizations of these data sets, $N$ was split up into training (80%) and testing (20%) data sets. Random sampling without replacement $R$ was then obtained from $K$ permutations of $N$ where $K$ is the desired number of samples. Thus, each random sample had different input-output patterns with differential weighting. Therefore, each model was sufficiently distinct with respect to its prediction capability because these models represented different regions of the decision space (all possible combinations of data shuffling) (Sreekanth and Datta, 2011b).

### 7.2.3 Coupled simulation-optimization using En-MARS

The present study adopts an externally linked S-O methodology to obtain optimal groundwater extraction strategies for multilayered coastal aquifer system. En-MARS meta-models were used to simulate the aquifer processes, in order to obtain saltwater concentrations at specified MLs. Each ensemble member was separately linked to the optimization algorithm. This procedure provides optimal values of groundwater extraction strategies based on the pre-specified constraints on maximum allowable salt concentrations at different MLs. The proposed methodology used a multiple objective problem setting in which the optimal groundwater extraction values were presented in a Pareto optimal front that demonstrated the trade-offs between the conflicting objectives. A schematic representation of the ensemble-based proposed coastal aquifer management model is illustrated in Figure 7.2.

## 7.3 Results and discussion

An illustrative multilayered coastal aquifer system of 4:35 km$^2$ areal extent, introduced in Subsection 4.2.5 in chapter 4, was used to evaluate the performance of the proposed En-MARS based saltwater-intrusion management model. The solution results of the En-MARS

based saltwater intrusion management model using linked S-O technique are presented in the following subsections.

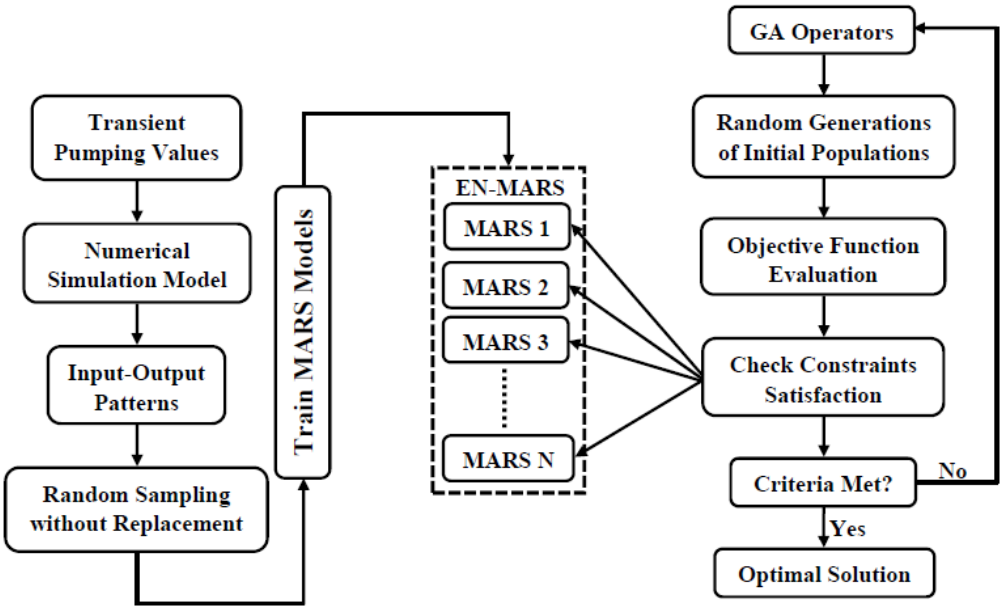


**Figure 7.2 Flow chart of the ensemble based linked simulation-optimization methodology**

7.3.1 Performance evaluation of the proposed En-MARS models

To reduce prediction uncertainty in modelling, an ensemble of MARS based meta-models was developed and evaluated in the present study. The En-MARS model was constructed by combining multiple individual MARS meta-models to learn the same data set, with each being presented with different realizations generated using random sampling without replacement from the entire data set. To maintain the same standard, each individual meta-model within En-MARS was developed using similar model parameters at all MLs. However, each individual model was trained and validated by using different sets of training and testing patterns for a particular monitoring location. Fifty replicates of training and testing data sets were used to train individual MARS meta-models in the ensemble for each ML. Test data sets were used to check for model overfitting during the learning process of all models. Once the learning was accomplished, each individual MARS meta-model was used to predict salinity concentrations on a new test sample. To maintain consistency, the same test dataset was used for prediction purposes of all meta-models in the ensemble at each individual ML. Because the final values of model parameters are dependent on the training data set, their values may differ slightly between MARS meta-models trained using different training data sets.

The RMSE criteria were used to evaluate the performance of the En-MARS models. RMSE values were computed over a new test dataset, which was sufficiently different from the test data sets used to test for model overfitting of the individual MARS meta-models. Individual MARS meta-models produced different RMSE values against this new test dataset because these models were trained from different realizations of the data set. Therefore, these models had different predictive capabilities, based on the training data set used for training. Combining them in an ensemble reduced the uncertainty associated with meta-model predictions. Presenting a different realization of test data set to the trained meta-models during ensemble formation was intentionally done, with a view to checking whether the ensemble still performed consistently with the new test dataset. Individual MARS meta-models were added sequentially to the ensemble, and the resulting RMSE was evaluated. The RMSEs of En-MARS models predicting saltwater concentrations at monitoring locations ML1-ML5 are plotted against the number of MARS models within the ensemble, as shown in Figure 7.3, Figure 7.4, Figure 7.5, Figure 7.6, and in Figure 7.7. In contrast to Sreekanth and Datta (2011b), adding an individual MARS meta-model to the ensemble started from the second model, and continued up to model 50. The addition of models into the ensemble was limited to 50 because the lowest RMSE value was attained within 50 individual MARS models in most cases. Because training data sets were chosen randomly without replacement from the total data set, these MARS models have different prediction capabilities. Therefore, it is impossible to know in advance which model will perform better with the new test data realization. Adding a worse model to the existing ensemble may increase the ensemble's RMSE. Based on the RMSE values, the number of MARS meta-models in the ensemble was chosen as 32, 23, 26, 50, and 8, respectively, for monitoring locations ML1, ML2, ML3, ML4, and ML5.

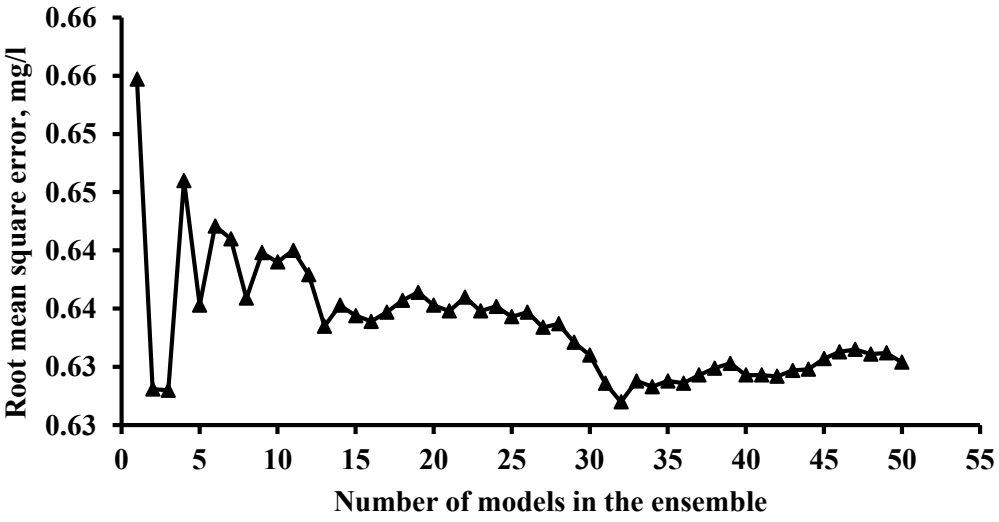


**Figure 7.3 Root-mean square errors for sequential addition of models in the ensemble at ML1**

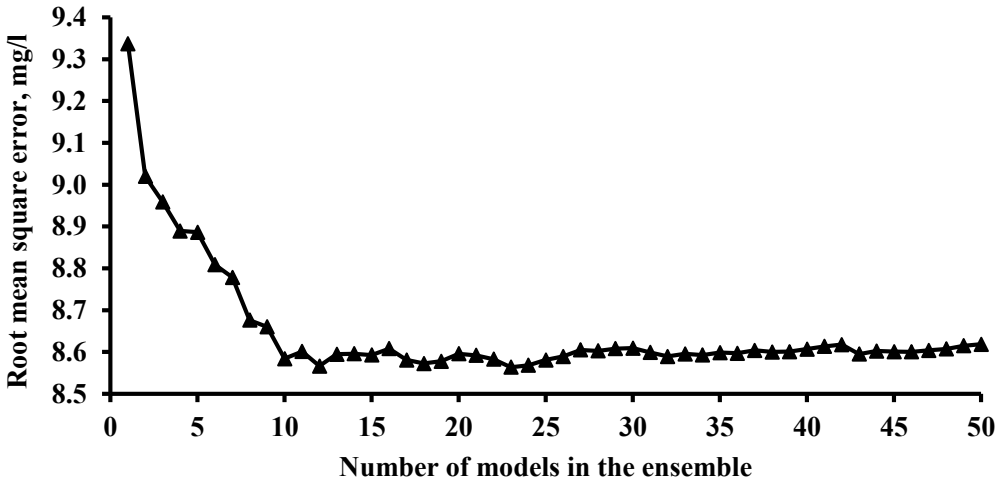


**Figure 7.4 Root-mean square errors for sequential addition of models in the ensemble at ML2**

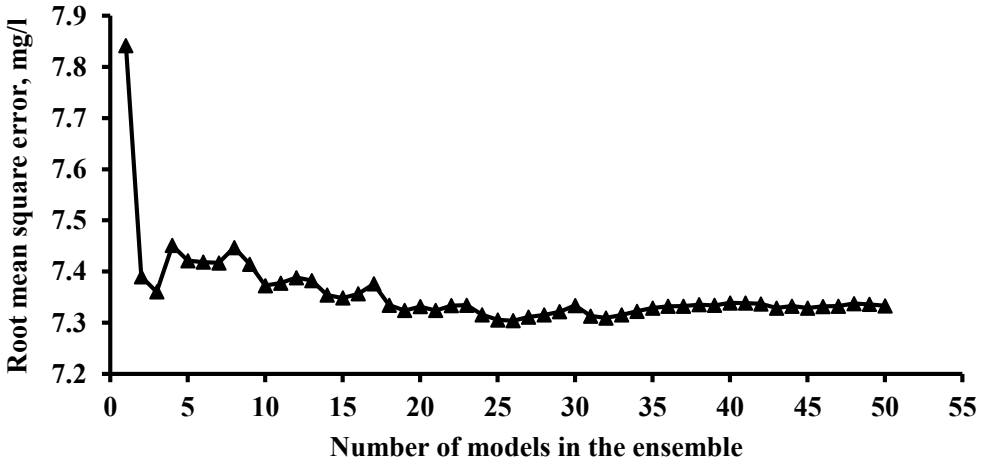


**Figure 7.5 Root-mean square errors for sequential addition of models in the ensemble at ML3**

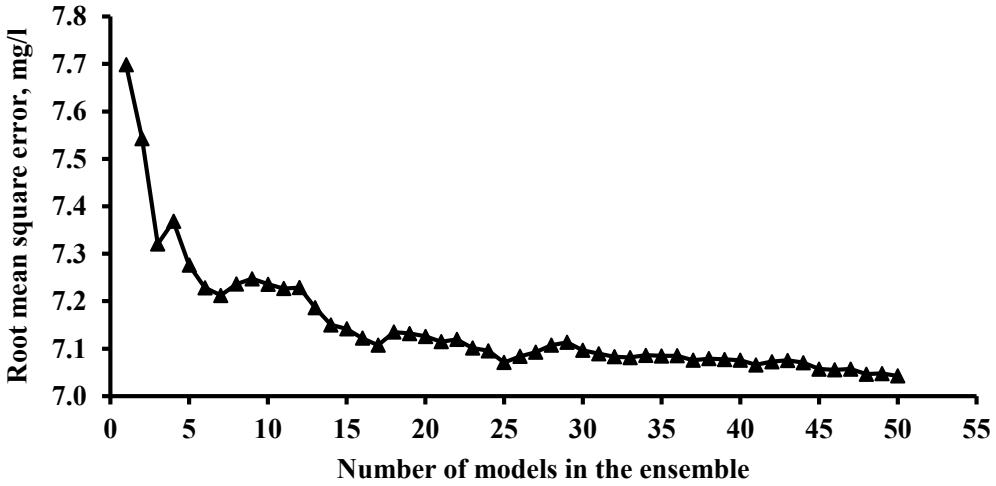


**Figure 7.6 Root-mean square errors for sequential addition of models in the ensemble at ML4**

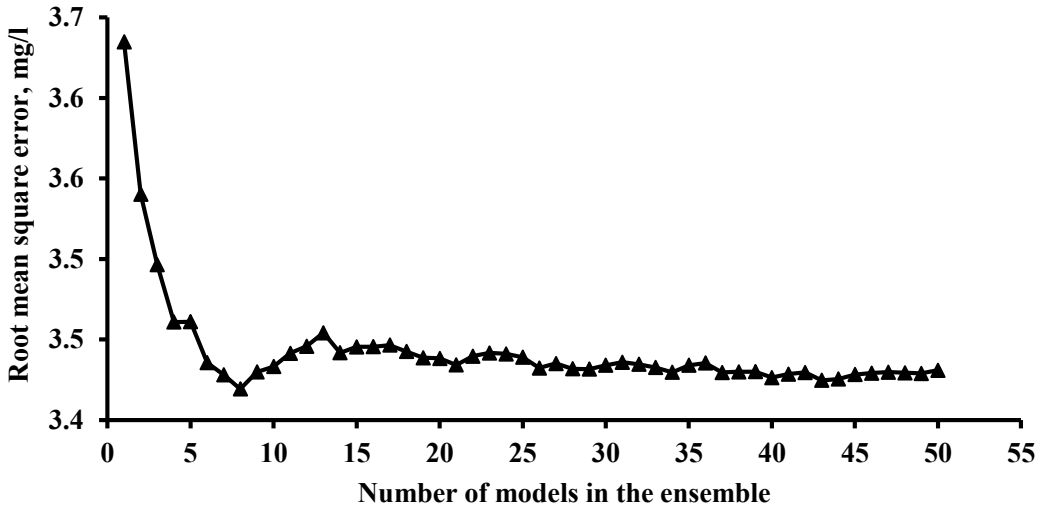


**Figure 7.7 Root-mean square errors for sequential addition of models in the ensemble at ML5**

The performance of the developed En-MARS models to approximate coupled flow and salt transport processes in a multilayered coastal aquifer system was evaluated and compared with those results obtained from the best performing models in the ensemble. The best performing model within the ensemble was also selected, based on the RMSE criteria on the test dataset. En-MARS models were used to predict salinity concentrations at specified MLs with respect to transient pumping stress applied to the aquifer. The accuracy of the models was evaluated based on their performance with a new realization of the test data set, the results of which are presented in Table 7.1. It is apparent from Table 7.1 that all the statistical performance indices were similar for both best individual MARS models and En-MARS models. Both models produced smaller values of RMSE and MAPRE as well as higher values of R and NS.

**Table 7.1 Performance evaluation results of MARS and En-MARS on testing data set**

| Statistical Indices | Models | ML1 | ML2 | ML3 | ML4 | ML5 |
|---|---|---|---|---|---|---|
| MAPRE, % | MARS | 1.16 | 0.45 | 0.53 | 0.08 | 0.04 |
| | En-MARS | 1.12 | 0.45 | 0.52 | 0.08 | 0.04 |
| RMSE, mg/l | MARS | 0.62 | 8.57 | 7.33 | 7.58 | 3.49 |
| | En-MARS | 0.63 | 8.56 | 7.30 | 7.04 | 3.42 |
| R | MARS | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| | En-MARS | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| NS | MARS | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 |
| | En-MARS | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 |

Note: MARS = Multivariate Adaptive Regression spline; En-MARS = Multivariate Adaptive Regression Spline Ensemble; MLs = Monitoring Locations; MAPRE = Mean Absolute Percentage Relative Error; RMSE = Root Mean Squared Error; R = Correlation Coefficient; NS = Nash-Sutcliffe Efficiency Coefficient

The MAPRE measure provides information on the distribution of errors, offers an indication as to whether a model tends to underestimate or overestimate, and is a measure for testing the robustness of the developed model (He et al., 2014). The developed models had relatively small MAPRE values (less than 1% for monitoring locations ML2-ML5 and close to 1% for monitoring location ML1). This indicates that both models are able to perform well from the relative error viewpoint. En-MARS provided slightly lower MAPRE values for all MLs. This is because En-MARS models were trained from a wide range of datasets arising from different realizations of the training datasets from the entire decision space. Therefore, these models are able to capture the trend of the data even from an entirely new realization of the test dataset. Khalil et al. (2015) also demonstrated the consistency and reliability of ensemble models constructed using different realizations of the training datasets in groundwater level predictions.

The prediction capability of any hydrological model can also be evaluated using the NS criterion. Models can be regarded as acceptable and reliable when they provide a NS value greater than or equal to 0.8 (Shu and Ouarda, 2008). Both best individual MARS and En-MARS models produced NS values very close to 1 for all MLs, which indicates an acceptable and reliable prediction capability of the developed models. The values of R for all MLs were very close to 1, which indicated that the model performance was also satisfactory from a correlation coefficient viewpoint. R values obviously exhibit a strong positive correlation between actual and predicted saltwater concentration values.

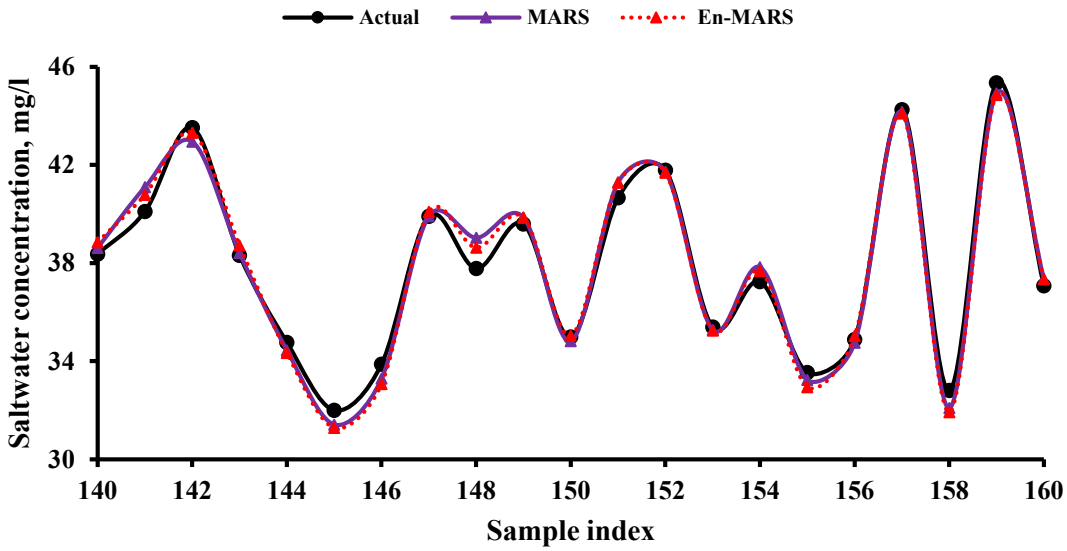


**Figure 7.8 Actual and predicted saltwater concentration values using the best MARS and En-MARS models at ML1**

Figure 7.8 presents a comparison of the actual and predicted saltwater concentration values by the best MARS and En-MARS models on the testing dataset at ML1 as an example.

Out of 300 test datasets, 21 datasets, taken from the middle of the 300 data points, are presented in Figure 7.8. The actual concentration values represent the FEMWATER simulated concentration values, which were assumed to be as accurate as can be obtained in real situations. The predicted concentration values are those obtained from the best individual MARS and the En-MARS models. As graphically illustrated in Figure 7.8, both models' estimates were very close to the corresponding values of actual salinity concentrations. Although both models' predictions were very accurate for all MLs, En-MARS provides even better estimates for some data points. For instance, test data points 141 and 148 of ML1 clearly indicate the superiority of the En-MARS model over the best model used in the ensemble. This was also true for most data points at all MLs. This is attributed to the superior learning capabilities of the En-MARS models from a wide variety of realizations of the actual datasets, obtained using random sampling without replacement technique.

**Table 7.2 Threshold statistics between actual and predicted saltwater concentration values**

| Models | RE threshold | ML1 | ML2 | ML3 | ML4 | ML5 |
|---|---|---|---|---|---|---|
| MARS | <1% | 54.33 | 90.00 | 86.67 | 100.00 | 100.00 |
| | <2% | 83.67 | 98.67 | 98.00 | – | – |
| | <5% | 99.00 | 99.00 | 100.00 | – | – |
| | <8% | 100.00 | 100.00 | – | – | – |
| En-MARS | <1% | 57.00 | 91.67 | 87.33 | 100.00 | 100.00 |
| | <2% | 86.00 | 99.33 | 98.00 | – | – |
| | <5% | 99.00 | 99.67 | 100.00 | – | – |
| | <8% | 100.00 | 100.00 | – | – | – |

Note: RE = Relative Error; MARS = Multivariate Adaptive Regression Spline; En-MARS = Multivariate Adaptive Regression Spline Ensemble; MLs = Monitoring Locations

Table 7.2 presents TS value differences between the actual and predicted saltwater concentrations. TS values were used to estimate the percentage of data points whose MAPRE values were less than the specified threshold values (Raghavendra and Deka, 2015). Four threshold values (1, 2, 5, and 8%) were selected in the present study. These results indicate that the performance of the En-MARS models was slightly better than the individual best MARS models.

### 7.3.2 Comparison of the performance of En-MARS with existing models

The comparison results with ANN models is not presented because the superiority of GP over ANN for predicting saltwater intrusion processes is well established (Sreekanth and

Datta, 2010; Sreekanth and Datta, 2011a). Therefore, the following paragraphs present a relative comparison of the proposed MARS based meta-model with GP, RBF, EPR, and FIS based meta-models. The study area, as well as the initial and boundary conditions, were different for each of the cases. Therefore, a direct comparison of the En-MARS testing result was not possible. Hence, the comparison was limited to the prediction capability of the models based on some error criteria. Sreekanth and Datta (2011b) used a simple single layered homogeneous coastal aquifer system with 33 input pumping variables in the development of the saltwater intrusion management model. However, the present study using En-MARS meta-model is implemented for a much more complex hydrogeological system with a larger number of layers which have different hydrogeological properties. In addition, the number of input variables was much larger in this study, e.g., the pumping variables totalled to 80 in space and time. Despite having a more complex study area with larger number of input variables, the proposed MARS based meta-model had relatively lower RMSE values compared to the GP models developed for a less complex study area having fewer input variables by Sreekanth and Datta (2011b).

RMSE values obtained for the best individual MARS meta-models were 0.62, 8.56, 7.33, 7.58, and 3.49 mg/l for monitoring locations ML1, ML2, ML3, ML4, and ML5, respectively. En-MARS produced RMSE values of 0.63, 8.56, 7.30, 7.04, and 3.42 mg/l for monitoring locations ML1, ML2, ML3, ML4, and ML5, respectively. These RMSE values were very close to the RMSE values obtained from the best individual MARS model in the ensemble. In contrast, the GP models developed by Sreekanth and Datta (2011b) produced RMSE values in the range of 9-27, 8-22, and 9-40 mg/l for monitoring locations C1, C2, and C3, respectively, obtained from Figures 4-6 in Sreekanth and Datta (2011b). Sreekanth and Datta (2011b) used an ensemble of 30 GP models at monitoring location C1 to produce a RMSE value of approximately 17 mg/l, which was almost double the RMSE value obtained from the best individual GP model in the ensemble. Therefore, it can be concluded that En-MARS performed better in the present study than the GP based ensemble meta-models presented by Sreekanth and Datta (2011b).

Christelis and Mantoglou (2016) presented a reduction in computational time of 96% for the variable density based optimization for a single objective management problem using a RBF based meta-model. In the present study, computational time savings with and without parallel processing facilities were computed. A computational time saving of 57.2% was achieved by using parallel processing facilities. However, the results obtained in the present study cannot be compared directly with those given by Christelis and Mantoglou (2016), because Christelis and Mantoglou (2016) presented a single objective optimization problem, which did not use ensemble and parallel processing facilities.

The values of Coefficient of Determination ($R^2$) obtained by MARS and En-MARS meta-models were better than the $R^2$ values obtained by EPR meta-model presented by Hussain et al. (2015) . Hussain et al. (2015) reported $R^2$ values of 95.5, 95.6, and 96.5% in testing datasets for the selected terms (to reduce complexity) T10, T50, and T90, respectively. MARS produced $R^2$ values of 98.2, 99.7, 99.7, 99.9, and 99.9% for monitoring locations ML1, ML2, ML3, ML4, and ML5, respectively. This indicates a relatively better fit of the MARS models than the EPR models. The En-MARS models also produced similar values of $R^2$ (98.16, 99.71, 99.73, 99.95, and 99.97% for monitoring locations ML1, ML2, ML3, ML4, and ML5, respectively).

FIS, as proposed by Roy and Datta (2017a) and presented in chapter 4, has the ability to predict saltwater concentrations at multiple locations in a single formulation for multiple output problems. The results obtained by the FIS at the end of the fifth time step are comparable to those obtained using MARS and En-MARS models in the present study. Only the RMSE values in the prediction set were used for comparison purposes. Roy and Datta (2017a) obtained RMSE values of 0.73, 23.96, 20.45, 15.79, and 4.58 mg/l for monitoring locations ML1, ML2, ML3, ML4, and ML5, respectively. Both MARS and En-MARS based meta-models performed better than the FIS based meta-model in terms of RMSE values on the test data set (Table 7.1).

### 7.3.3 Performance of En-MARS based management model

Each of the members (MARS meta-models) of the ensemble for all MLs was separately linked to the optimization algorithm (CEMOGA). Therefore, a total number of 139 constraints corresponding to five ensembles with 32, 23, 26, 50, and 8 MARS meta-models, respectively, for each ensemble were considered in this study. Other constraints of the management model included maximum permissible saltwater concentrations at the specified MLs. The optimum values of different parameters used in the optimization formulation were selected by running the optimization problem with several combinations of the different parameters. Population size is a crucial parameter that specifies the number of individuals in each generation. A large population size allows the genetic algorithm to search the solution space more thoroughly, thereby increasing the chances of obtaining a global solution. A smaller population narrows down the search space and affects the spread of solutions in the Pareto optimal front. This study used a population size of 1,600, considering both the diversity of the population and the computational requirements based on different trials in combination with other parameters.

The optimization algorithm used 2,494 generations to reach the global optimal solution. Hence, a total number of 1,600 × 2,494 = 3,990,401 evaluations of the aquifer

response to specific groundwater extraction patterns were executed by the optimization algorithm. The algorithm used a tournament selection with a tournament size of two. The crossover and mutation fractions used were 0.9 and 0.1, respectively. The migration parameter used a forward direction with a fraction and interval of 0.2 and 20, respectively. A Pareto front population fraction of 0.7 was used. Therefore, the Pareto front contained 1,600×0.7 = 1,120 non-dominated optimal solutions. Function and constraint tolerance were set as $1\times10^{-5}$ and $1\times10^{-3}$, respectively.

The optimization routine for this multiple objective saltwater intrusion management problem was carried out using the parallel computing platform in MATLAB. The parallel computation was performed by distributing the tasks among four workers (four physical cores) of the local cluster, using a seven-core PC [Intel Core i7-4790 CPU @ 3.60 GHz with 8 gigabytes (GB) RAM, Santa Clara, California]. The time required to obtain the optimal solution for the proposed management problem with and without parallel processing was computed for comparison purposes, using the same initial states of the optimization algorithm. Compared to 1.09875 hours of computational time with parallel processing, optimization without parallel processing took approximately 2.5674 hours. This indicates a considerable amount of computational time was saved by using the parallel processing capability.

The management problem developed and solved in this study considered 80 input variables, corresponding to water extraction from 16 locations for five time steps. It had five output variables that relate to saltwater concentrations at five MLs at the end of the management period of 5 years. The Pareto optimal front illustrating the trade-off between the two conflicting objectives is shown in Figure 7.9. The Pareto front illustrated in Figure 7.9 provides several alternate solutions that maximize water extraction from production wells and minimize barrier well extraction, while maintaining the salt concentrations within the indicated acceptable limits at specified monitoring locations.

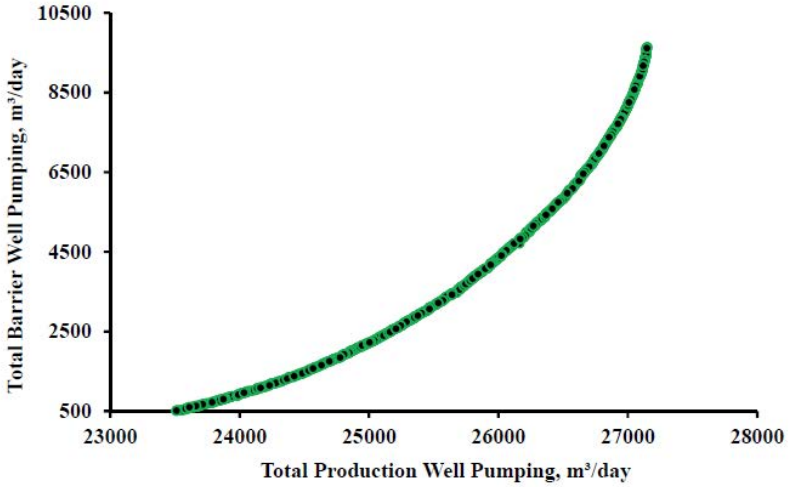


**Figure 7.9 Pareto Optimal front of the ensemble based management model**

It is observed from Figure 7.9 that a certain amount of water can be extracted from production wells without any barrier well pumping. However, as the amount of water extracted from production wells needs to be increased, an increased amount of barrier well extraction will also be required to maintain the water quality. However, the proposed management model did not look into the various disposal alternatives for water pumped from the barrier wells.

For simulation of the study area and to generate the required input-output patterns for training the MARS and En-MARS meta-models, pumping values for both production and barrier extraction wells were obtained from LHS within the practical range of 0-1,300 $m^3$/day. The same upper and lower limits of the decision pumping variables for both production and barrier wells were also specified in the management model. In the management model, the second objective was to minimize the rate of extraction from the barrier extraction wells so that the pre specified maximum allowable salt concentrations at all MLs were maintained. The Pareto optimal front shown in Figure 7.9 provides the optimized values of pumping from barrier extraction wells.

## 7.4 Conclusions

A linked S-O methodology using an ensemble meta-modelling approach was developed for a multiple objective groundwater extraction strategy in a multilayered coastal aquifer system. In this approach, ensembles of MARS meta-models were linked externally to a CEMOGA based optimization algorithm to derive the optimal groundwater extraction strategy. To achieve computational efficiency in the S-O methodology, the optimization routine was run on a parallel computing platform. The optimum number of MARS models in the ensemble was chosen based on the RMSE criteria. The ensemble that produced the minimum RMSE was chosen for the prediction of saltwater concentrations at specified MLs. The ensemble meta-modelling approach accounted for the predictive uncertainty of the meta-modelling approach. The MARS based ensemble meta-modelling technique using a parallel processing platform has the advantages of (1) addressing the predictive uncertainty of meta-modelling, and (2) achieving computational efficiency in searching for optimal solutions: 3,990,401 function evaluations in the optimization routine took only 76 minutes.

To ensure sufficient variability in the training dataset as well as to obtain diversified MARS meta-models for ensemble formation, different realizations of the training dataset were obtained from the entire dataset. The performance of En-MARS meta-models were compared with those of the best model in the ensemble. The ensemble MARS meta-models provided similar results when compared to the best individual model in the ensemble. The disadvantage of single meta-modelling approach is that it is not always possible to find the

best model a priori. The other disadvantage is that the best model may not provide a similarly successful result for another set of test data. In addition, a single meta-model does not consider predictive uncertainty. However, as the En-MARS learned from a diverse data range, it is expected to be more consistent for a newer test data set, thereby reducing predictive uncertainty. In addition, even with a larger number of input variables (groundwater extraction) and more complex hydrogeological configuration, En-MARS resulted in smaller prediction errors in terms of RMSE values when compared to existing meta-models. Another reason for adopting En-MARS is to develop new meta-models for more complex study areas. Therefore, the adaptive nature of MARS models, as well as use of ensembles and parallel processing, resulted in a computationally efficient, accurate, and reliable methodology for coastal aquifer management, which also incorporated uncertainties in modelling.

Generally, it takes many years to reach a steady state in terms of the flow field. Only then, may a steady state concentration field be attained. Any long-term spatial and temporal management strategy ultimately reaches a steady state if the extraction and the rates do not change substantially from year to year, although these can change seasonally. However, the performance evaluation reported here is for transient pumping values varying in space and time. In order to make the problem more unconstrained for evaluation purpose, the extraction rate is not limited to the steady state. Therefore, the specified management strategy obtained as a solution was not tested for achieving an ultimate steady state.

The present study considered a multilayered coastal aquifer system having four layers of materials, and the materials within each layer were considered homogeneous. Future research can be directed towards using heterogeneous coastal aquifer system using random fields of different aquifer parameters. In addition, seasonal effects and tidal fluctuations in the adjoining river within the management model can be incorporated in future research. In addition, the proposed management model did not look into the various disposal alternatives for water pumped from the barrier wells. Related economic and environmental considerations can be included in a more detailed management model. The proposed methodology potentially facilitates the computationally feasible and efficient determination of regional scale management strategies for coastal aquifers. One important practical aspect of developing a regional scale coastal aquifer management model is the plausible impact of climate change on the sustainable coastal aquifer management strategy. Therefore, the issue of sea level rise on the development of sustainable future management strategy is discussed in the next chapter.

# Chapter 8: Influence of sea level rise on the management of coastal aquifers

Partial contents of this chapter have been published and copyrighted, as outlined below:

Roy, D. K., & Datta, B. (2018c). Influence of sea level rise on multi-objective management of saltwater intrusion in coastal aquifers. *Journal of Hydrologic Engineering*, ASCE, *23*(8), 04018035.

This chapter covers the influence of climate change induced sea level rise on the development of sustainable future management strategy.

## 8.1 Summary

Relative sea level rise, providing an additional saline water head at the seaside, has a significant impact on an increase in the salinization process of the coastal aquifers around the globe. Sea level rise can accelerate saltwater intrusion processes in aquifer systems, and several centuries would be required to equilibrate this sea level rise induced saltwater intrusion, even if the sea level has fallen back to its original position. Although excessive groundwater withdrawal is considered to be the major cause of saltwater intrusion, relative sea level rise in combination with the effect of excessive groundwater pumping can exacerbate the already vulnerable coastal aquifers. Numerous studies have demonstrated the effect of relative sea level rise on the salinization process of the aquifer. However, the effect of relative sea level rise has not been incorporated in a saltwater intrusion management model for a multilayered coastal aquifer system under the influence of transient pumping stresses applied to the aquifer. This study intends to incorporate the effects of relative sea level on the optimized groundwater extraction values for a management period of 5 years. Variation of water concentrations of the tidal river and seasonal fluctuation of river water head were also incorporated in the model.

This chapter demonstrates the influence of climate change induced sea level rise on multiple objective saltwater intrusion management strategies in coastal aquifers. Three meta-models were developed from the solution results of a numerical code FEMWATER that simulates the coupled flow and salt transport processes in a coastal aquifer system. Results revealed that the proposed meta-models are capable of predicting density dependent coupled flow and salt transport patterns quite accurately. Based on the comparison results, the best meta-model was selected as a computationally cheap substitute for the simulation model in the coupled S-O based saltwater intrusion management model. To achieve computational efficiency, the optimization routine of the proposed management model was performed on a

parallel computing platform. The performance of the proposed methodology was evaluated for an illustrative multilayered coastal aquifer system in which the effect of climate change induced sea level rise was incorporated for the specified management period. Results showed that the proposed saltwater intrusion management model provides acceptable, accurate, and reliable solutions while significantly improving computational efficiency in a coupled S-O methodology.

## 8.2 Methodology

The proposed methodology includes selection of meta-models which are able to emulate physical processes of a coastal aquifer system under the influence of climate change induced sea level rise. Upon selection of the best performing meta-model, a saltwater intrusion management model was developed, using the best meta-model as a replacement for the simulation model in the coupled S-O approach. The meta-models were trained and validated from the input-output patterns generated by using a 3-D density reliant numerical simulation model FEMWATER.

### 8.2.1 Meta-models

Three different artificial intelligence based meta-models were evaluated and compared in order to identify an accurate, reliable, and computationally efficient substitute of the density reliant coupled flow and salt transport processes of a multilayered coastal aquifer system. The basic principles of the MARS and ANFIS based meta-models are presented in chapter 4 (Subsection 4.3) and chapter 6 (Subsection 6.2.1), respectively. Brief descriptions of the Gaussian Process Regression (GPR) are presented in this subsection.

GPR (Rasmussen and Williams, 2005) is a flexible, nonparametric, probability based approach in which the output variable, $Y$ is a function of input variables, $X(k)$ such that $Y = f\big(X(k)\big) + \varepsilon$, where $\varepsilon$ is a Gaussian noise with variance $\sigma_n^2$ (Bishop, 2006). The GPR approach of meta-modelling is based on the Gaussian process theory developed by Rasmussen and Williams (2005). Rasmussen and Williams (2005) defined Gaussian process (GP) as "*A Gaussian process is a collection of random variables, any finite number of which has a joint Gaussian distribution*". The definition inevitably suggests a prerequisite of '*consistency*', also known as marginalization property. This marginalization property implies that if the GP e.g. specifies $(y_1, y_2) \sim N(\mu, \Sigma)$, then it must also specify $y_1 \sim N(\mu_1, \Sigma_{11})$, where the relevant sub-matrix of $\Sigma$ is $\Sigma_{11}$.

A Gaussian process is entirely indicated by its mean and covariance functions. The mean function provides a description of the expected value of the function at any particular point within the input space, and is given by

$$Mean\ function:\ m(x_i) = E[f(x_i)] \tag{8.1}$$

On the other hand, the covariance function, considered as the most influential and important element of a GPR model, defines proximity (nearness) or resemblance (similarity) between the predictor values $x_i$ and response (target) value $y_i$ (Rasmussen and Williams, 2005). Covariance between the two latent variables $f(x_i)$ and $f(x_j)$ is specified by the covariance function, i.e. how response at one point $x_i$ is affected by responses at other points $x_j$, $i \neq j, i = 1,2,\dots,n$ is also determined by the covariance function. The covariance function is expressed as

$$\begin{aligned} &Covariance\ function:\ k(x_i, x_j) \\ &\quad = E\left[\left(f(x_i) - m(x_i)\right) \times \left(f(x_j) - m(x_j)\right)\right] \end{aligned} \tag{8.2}$$

Finally, the Gaussian process can be written as

$$f(x) \sim gp\left(m(x_i), k(x_i, x_j)\right) \tag{8.3}$$

The parameters associated with mean and covariance functions, commonly known as free parameters or hyperparameters, define the properties of predictive probability distribution. The values of the hyperparameters are obtained through maximizing log-likelihood function of the training data (Rasmussen and Williams, 2005), and is given by

$$\log p(Y|X) = -\frac{1}{2}Y^T(K + \sigma_n^2 I)^{-1}Y - \frac{1}{2}\log(|K + \sigma_n^2 I|) - \frac{n}{2}\log(2\pi) \tag{8.4}$$

where, n is the number of training samples.

### 8.2.2 Selection of optimum meta-model structures

#### *8.2.2.1 Optimum number of clusters for ANFIS*

Selecting optimum number of clusters using the FCM algorithm for the initial FIS structure, based on which the final ANFIS architecture is developed, provides an important pre-processing step. This procedure of deciding the right number of clusters is usually associated with the nature of problem, and types and size of input variables. In most cases, a model with simple architecture in terms of input variables is desirable (Mohammadi et al., 2016). However, the present study conducts a set of trials to select the right number of clusters for ANFIS models at individual MLs by observing RMSEs between the actual and predicted responses obtained using selected ANFISs. The selection process is illustrated in Figure 8.1. The optimum number of clusters was one that produced minimum RMSE values between the training and test datasets. The lowest variance in RMSE value between training and test datasets was also checked to protect against model overfitting. Figure 8.1 shows that for all MLs, two clusters produce a relatively simple model structure with reasonable accuracy.

Therefore, the number of input and output Membership Functions (MF) used in the present study for each individual ML were 160 (80 input variables × 2 clusters) and 2 (one output variable × 2 clusters), respectively. A Sugeno type ANFIS considers Gaussian type input MF, whereas the output MF considered was linear. All ANFIS models were developed using commands and functions of MATLAB and Fuzzy Logic Toolbox of MATLAB (MATLAB, 2017a).

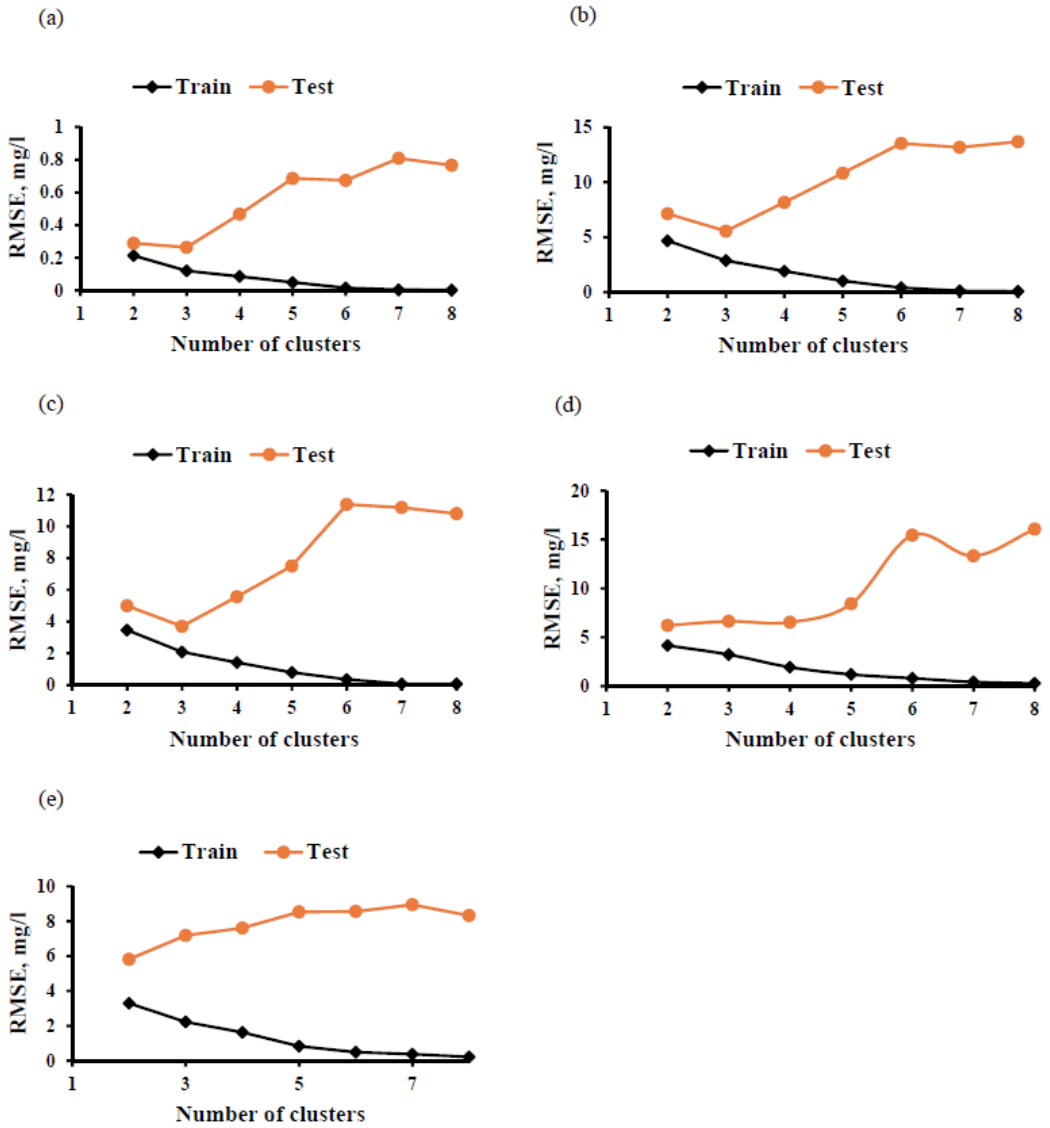


**Figure 8.1 Number of clusters vs. RMSE between the actual and predicted saltwater concentration values of training and test datasets predicted by ANFIS at: (a) ML1, (b) ML2, (c) ML3, (d) ML4, and (e) ML5**

*8.2.2.2 Optimum combination of Basis and kernel functions for GPR*

Optimum model structure for GPR based meta-models can be obtained by selecting the right combination of Basis and covariance (kernel) functions. The selection of Basis and kernel functions to obtain optimal GPR meta-model structure was conducted through a set of trials

by using several combinations of three Basis (constant, linear and quadratic) and six kernel functions (Squared Exponential, Matern 3/2, and Matern 5/2, ARD Squared Exponential, ARD Matern 3/2, and ARD Matern 5/2). RMSE values on training and test datasets with respect to the selection of Basis and kernel functions for ML1 are presented in Figure 8.2. A model producing smaller values of RMSE as well as the smallest difference in RMSEs between the training and test dataset was considered for further analysis. Based on RMSE values depicted in Figure 8.2, a combination of the quadratic Basis function and the ARD Squared Exponential kernel function was selected for the development of the GPR model at ML1. Similar results were obtained for other MLs, and the same combination of Basis and kernel functions were used. The hyperparameters were obtained through maximizing the log-likelihood function of the training dataset. An exact Gaussian process regression was chosen, i.e. an '*exact*' method was used to estimate the parameters of the developed GPR models. The initial value for the noise standard deviation of the GPR models was calculated by $\sigma = std(Y)/sqrt(2)$. The lower bound of this noise standard deviation was calculated as $Lower\ bound = 1e-2 * std(Y)$, where $Y$ represents the response variables (saltwater concentrations at specified MLs).

*8.2.2.3 Minimum observation between knots for selecting MARS models*

A minimum number of observations between knots was selected by varying this parameter ranging from 0-20 and observing the difference in RMSE values between the training and test datasets. The minimum number of observations between knots was chosen to develop MARS models, for which the difference in RMSE value was minimal. Based on these trial observations, as is seen in Figure 8.3, the minimum numbers of observations between the knots were chosen as 3, 11, 1, 12, and 4 respectively for monitoring locations ML1, ML2, ML3, ML4, and ML5.

For all considered meta-models, the training dataset consisted of 80% of the total input-output patterns generated by utilizing the numerical simulation model, FEMWATER. The remaining 20% of the generated patterns were used for validation of the meta-models. Once the training and validation steps were completed, the meta-models thus developed were presented with a very different realization of test dataset in order to check the prediction capability. This new realization of the test dataset was presented to all developed meta-models to maintain consistency and a fair comparison. Recently, researchers have focused on reducing the required number of training patterns to adaptively train the meta-model using a relatively small number of training patterns (Christelis and Mantoglou, 2016; Kourakos and Mantoglou, 2009; Papadopoulou et al., 2010; Sreekanth and Datta, 2010; Sreekanth and Datta, 2014a). Adaptive training of meta-models is able to significantly reduce the number of training patterns to develop a reasonably accurate approximate simulator of the density

dependent coupled flow and salt transport processes (Christelis and Mantoglou, 2016; Sreekanth and Datta, 2010).

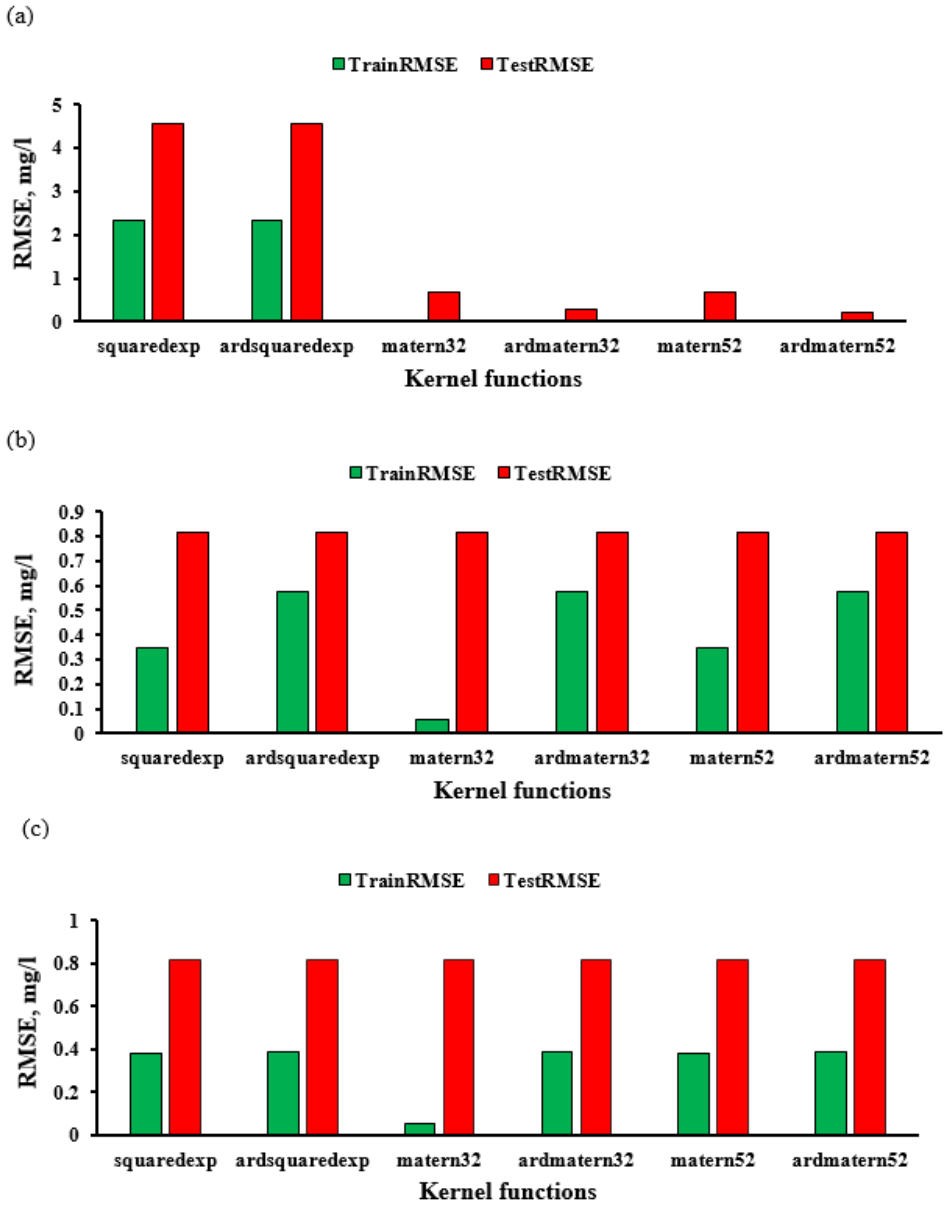


**Figure 8.2 Selection of Basis and covariance function for obtaining optimum GPR structure at ML1: (a) constant Basis function, (b) linear Basis function, (c) quadratic Basis function**

For adaptive training of the Modular Neural Network (MNN) and GP based meta-models, Sreekanth and Datta (2010) utilized an expanding set method in which more and more training patterns generated by the optimization strategy were added to the initial training pattern, based on the direction of search. Initially trained meta-models with a limited training pattern were used in conjunction with an optimization algorithm to find the near optimal solution. Once the near optimal solutions were obtained, a greater number of samples was generated in this near optimal space, and the meta-models were retrained for accurate predictions in the near optimal space. Christelis and Mantoglou (2016) developed an online

training scheme for the RBF based meta-model that was embedded within an optimization algorithm. Their approach also encompassed using the current best solutions found by the RBF model during the optimization operations to add infill points to the initial sampling plan. This infill strategy favoured a fast improvement of the RBF model at the region of the current optimum (local exploitation). However, this approach neglects the global improvement of the meta-model, and may fail to identify the region of the global optimum (Forrester et al., 2008).

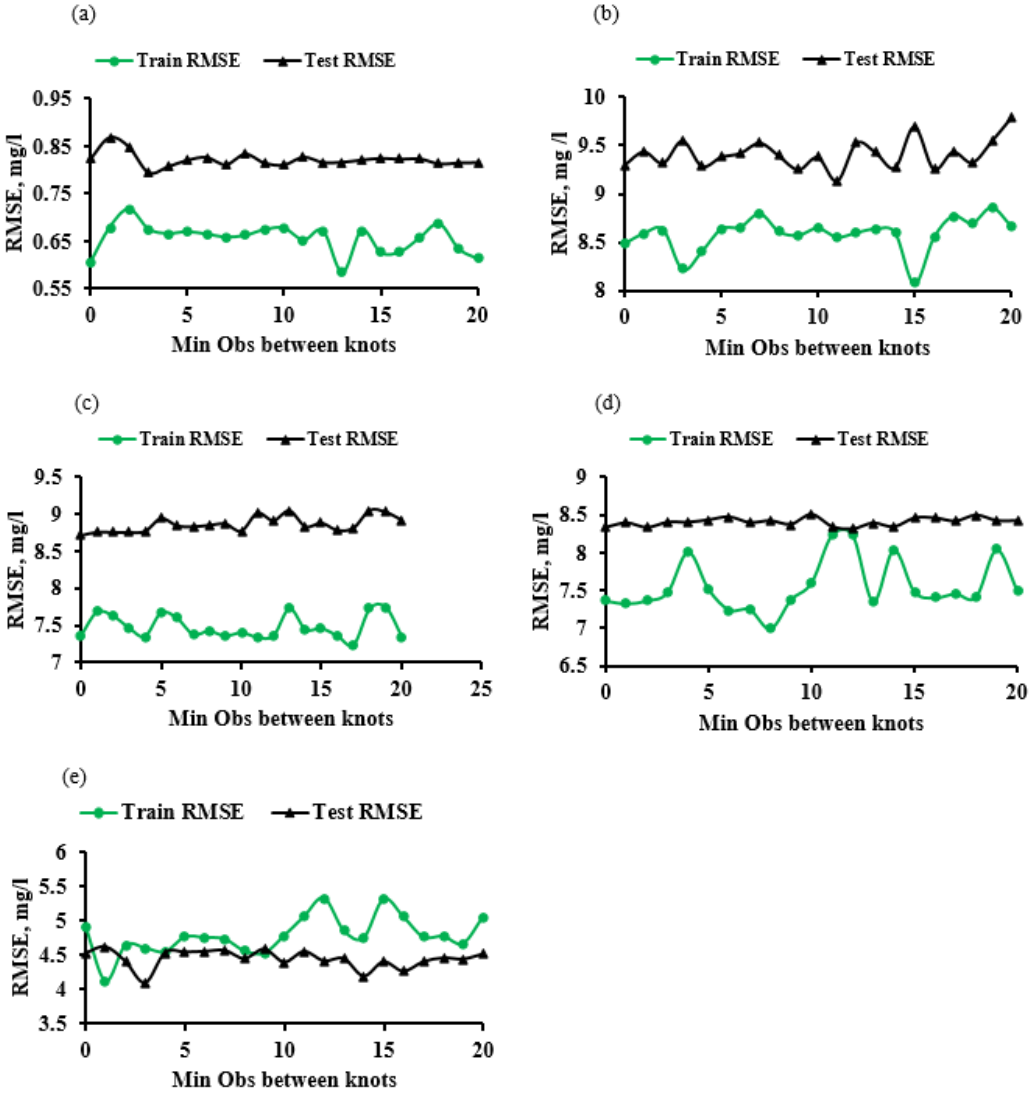


**Figure 8.3 Selection of optimum MARS model structure by varying minimum observation between knots and observing the resulting RMSE values between the training and test datasets: (a) ML1, (b) ML2, (c) ML3, (d) ML4, and (e) ML5**

Christelis and Mantoglou (2016) used a numerical model to generate input-output patterns for initial training of the RBF based meta-model. The trained RBF model was then used within the optimization algorithm to run the optimization algorithm until it found a local optimal solution. Subsequently, the numerical model was used to evaluate the current best

solution, and then added the new input-output data to the initial training pattern and retrained the RBF models. The process continued until it exceeded the pre-specified computational budget. While trying to achieve the computational efficiency associated with generating required input-output patterns for meta-model training, this approach ignores the accuracy and uncertainty of the meta-model predictions. In this approach, new meta-models are based on the local optimal solutions of the optimization process, and are only able to predict within limited regions of the total decision space of the input variables. Moreover, this approach requires repeatedly going back to the simulation model to evaluate the current best solutions. Therefore, the present study selects the optimal meta-model structures by conducting numerical experiments which vary different parameters and training over the entire feasible decision space. However, both approaches do have different advantages.

## 8.3 Application of the proposed methodology for an illustrative study area

An illustrative study area was used to evaluate performance of the proposed saltwater intrusion management model. The study area is similar to one presented in chapter 4 (Subsection 4.2.5) except that present study considers an average sea level rise of 3.6 mm/year, seasonal variation of river water stage, and varying saltwater concentration of river water near the sea. However, the present study did not consider the variation of head due to tidal fluctuations because several previous studies demonstrated that the effects of tidal fluctuation on salinization of coastal aquifers were negligible (Chen and Hsu, 2004; Heiss and Michael, 2014; Kuan et al., 2012; Narayan et al., 2007).

**Table 8.1 River water concentration**

| Distance from the sea side boundary, m | Concentration, mg/l |
|---|---|
| 0 - 193.77 | 12000 |
| 193.77 - 387.54 | 5000 |
| 387.54 - 581.31 | 1000 |
| 968.85 - 1356.39 | 100 |
| 1356.39 - 1743.93 | 0 |
| 1743.93 - 2906.55 | 0 |

The initial head of the seaside boundary was assumed as 0 m and allowed to increase incrementally for the management period of 5 years, during which the seaside head reached to 0.01667 m. Two stress periods, each having a duration of 6 months, were considered. Time varying specified heads were assigned to both ends of the seaside boundary. The upstream

end of the river was also assigned with a time varying specified head, taking into consideration a variation of head during the wet and dry seasons (1 m and 0.85 m respectively). These time varying specified heads varied linearly along the stream until they reached the assigned specified heads at the seaside boundary. The seaside boundary had a constant concentration of 35000 mg/l whereas the river boundary had varying concentrations. In the first 194 m length of the river from the sea, the river water concentration was assigned as 12000 mg/l, and assumed to be gradually decreasing until it reached 0 mg/l at the upstream end of the river. The assigned river water concentrations are given in Table 8.1.

A three-dimensional view of the aquifer system of the study area is shown in Figure 8.4.

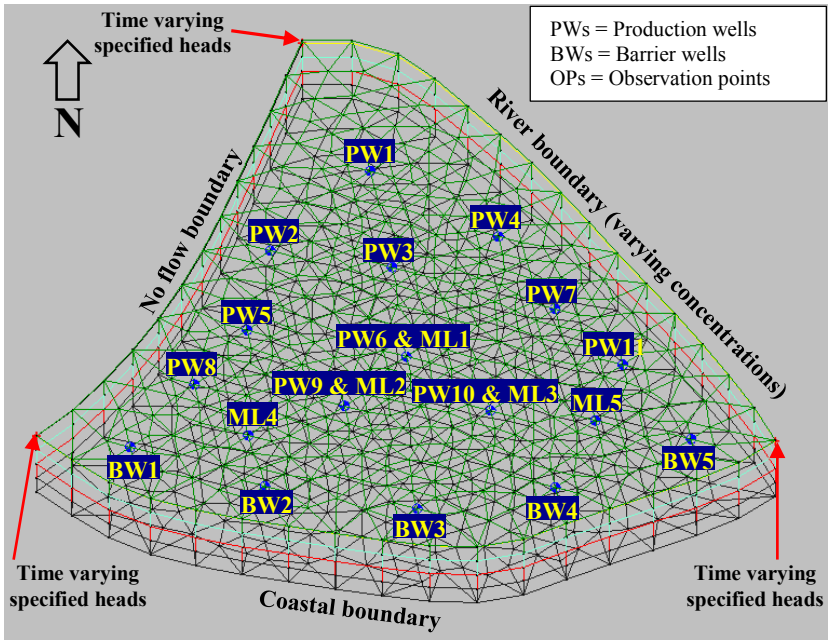


**Figure 8.4 Three-dimensional view of the model domain**

The study area had an aerial extent of 4.35 km$^2$ with a well spread pumping well field, having a well density of 3.68 wells/km$^2$ (16 wells/4.35 km$^2$). The unconfined aquifer had a total thickness of 80 m divided into four distinct layers of aquifer materials (materials within each layer was assumed homogeneous). An anisotropy ratio ($K_x/K_y$) = 2.0 was used, where $K_x$ represents the horizontal hydraulic conductivity in the $X$-direction. $K_y$ is the horizontal hydraulic conductivity in the $Y$-direction. $K_z$ is the vertical hydraulic conductivity in the $Z$-direction. The value of $K_z$ was taken as one tenth of the hydraulic conductivity values in the $X$-direction. Hydraulic conductivity values along with other aquifer parameters are same as presented in Table 3.1 (chapter 3).

The illustrative multilayered coastal aquifer study area considered 11 potential production wells that allowed water extraction for beneficial purposes (denoted by PW1-

PW11 in Figure 8.4). The study also considered 5 barrier extraction wells to create a hydraulic barrier near the sea side boundary for controlling saltwater intrusion (denoted by BW1-BW5 in Figure 8.4). Water was extracted from the 2nd and 3rd layer of the aquifer. Aquifer processes were simulated for a period of 5 years, divided into 5 uniform time steps of 1 year each, during which the rate of extraction of water from both the production and barrier wells were considered to be constant. Therefore, considering the spatial and temporal variation of water extraction, the present study considers 80 input variables (16 wells × 5 years) with lower and upper limits set as 0 and 1300 $m^3$/day, respectively. The resulting saltwater concentration due to water abstraction from the wells at the end of management period of 5 years was monitored at 5 MLs. The proposed management model was developed by optimizing water extraction from the production and barrier wells while keeping salinity concentrations at the specified MLs within maximum allowable limits. Consequently, the present study considers multiple objective type problem setting for developing the management model. MLs were located at three different salinity zones: ML1 was located in the low salinity zone, ML2 and ML3 were placed at the moderate salinity zone, and ML4 and ML5 were located in high salinity zones. MLs were placed at different salinity zones with a view to using the extracted water from different regions of the aquifer for different purposes.

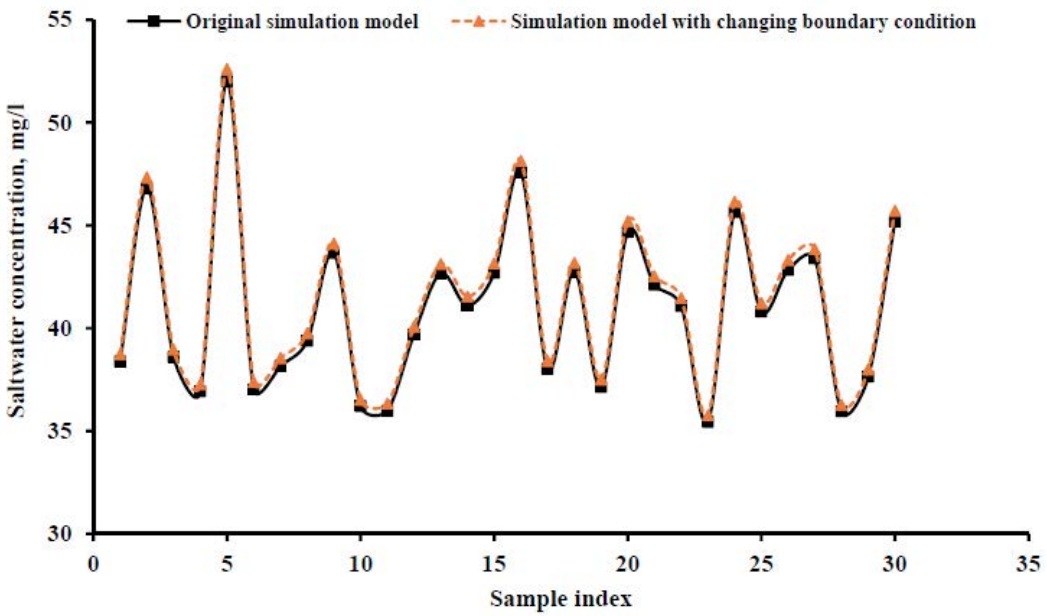


**Figure 8.5 Comparison of simulation results with and without sea level rise at ML1**

To demonstrate the effects of sea level rise, river water concentration and seasonal variation of river water stage on the migration of salt plume, a comparison of the simulation results was performed with and without these boundary conditions. However, the individual effects of these parameters were not considered in the present study. The Mean Absolute

Percentage Relative Difference (MAPRD) in saltwater concentration values with respect to the original simulation model (without the effects of sea level rise, river water concentration, and seasonal variation of river water stage) at monitoring locations ML1, ML2, ML3, ML4, and ML5 were 0.99%, 0.81%, 0.92%, 0.47%, and 0.82% respectively. These MAPRD values were compared with the MAPRE values obtained through proposed meta-models to make sure that the meta-models can accurately capture differences in saltwater concentration values at different MLs. Figure 8.5, Figure 8.6, Figure 8.7, Figure 8.8, and Figure 8.9 illustrate the difference in concentration values with and without these changing boundary conditions at five MLs. For brevity of presentation, only first 30 observations are presented in Figures 8.5-8.9.

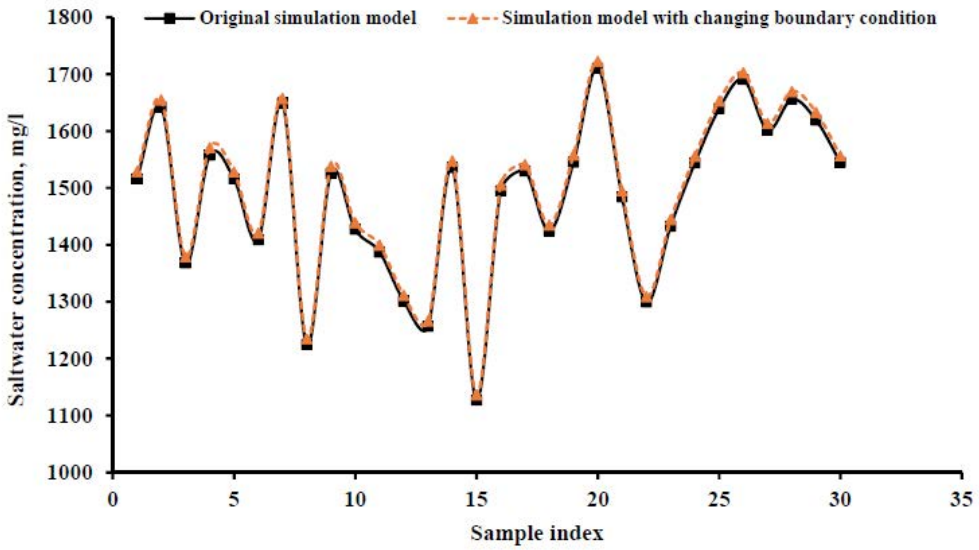


**Figure 8.6 Comparison of simulation results with and without sea level rise at ML2**

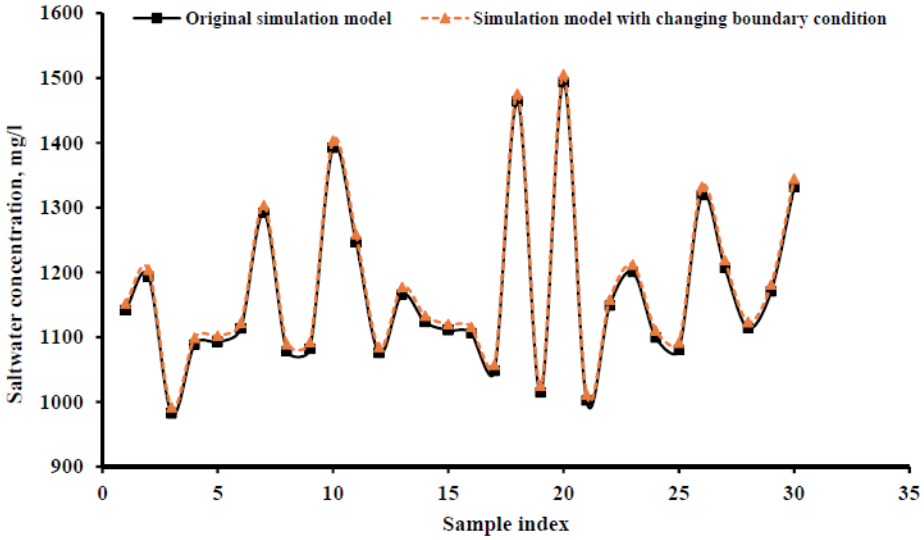


**Figure 8.7 Comparison of simulation results with and without sea level rise at ML3**

Modelling of groundwater flow and solute transport processes in real life coastal aquifers are influenced by uncertain model parameters (Sreekanth and Datta, 2014b). In a groundwater modelling system, the major sources of uncertainties arise from the associated aquifer characteristics such as hydraulic conductivity, compressibility, and bulk density (Ababou and Al-Bitar, 2004). Aquifer recharge is another source of uncertainty. To determine the effect of uncertain model parameter estimates on groundwater flow and saltwater intrusion processes in coastal aquifers, 30 different randomized realizations of hydraulic conductivity, compressibility, bulk density, and aquifer recharge were utilized in the present study. These uncertain model parameters were assumed to be homogeneous within a geological material layer, but heterogeneous within, with respect to each material layer. Different realizations of a representative set of hydraulic conductivity values were obtained from a lognormal distribution with a specified mean and standard deviation of the associated normal distribution. Aquifer recharge and compressibility realizations were generated from LHS uniform distributions within the parameter space with specific lower and upper bounds (Pebesma and Heuvelink, 1999). Realizations of bulk density were obtained from LHS technique from a $p$-dimensional multivariate normal distribution with specific mean and covariance. Groundwater flow and transport processes were simulated using this randomized set of uncertain model parameters for a fixed set of transient groundwater extraction values from both the production and barrier extraction wells. This procedure ensures a fair comparison between the resulting salinity concentrations from different randomized realizations of uncertain model parameters.

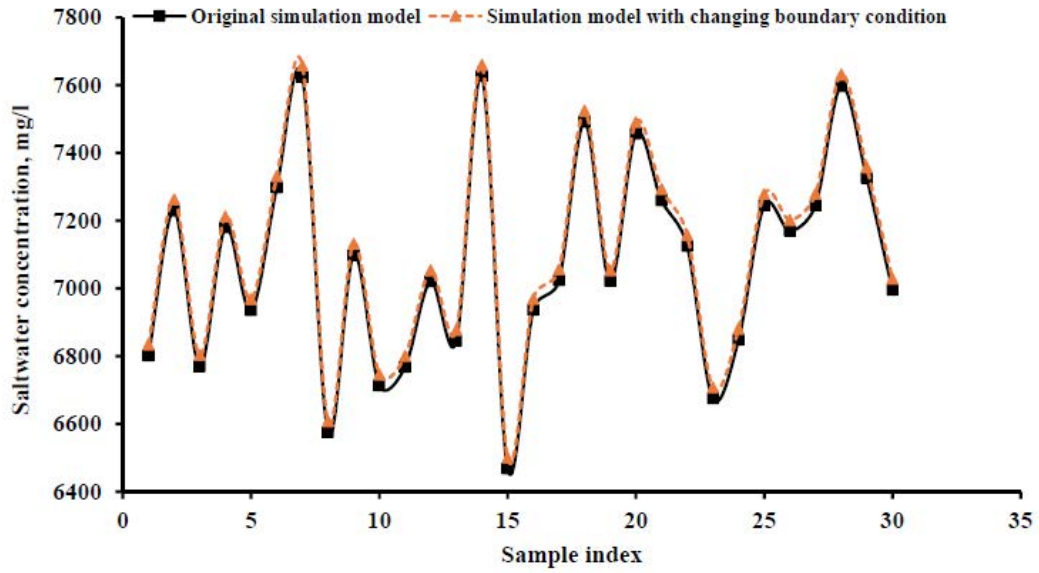


**Figure 8.8 Comparison of simulation results with and without sea level rise at ML4**

Salinity concentrations at different MLs obtained from these realizations of uncertain parameters were compared with those obtained when the mean values of these uncertain model parameters were used. MAPRD of salinity concentration values at different MLs were calculated for both cases, e.g. considering both with and without sea level rise scenarios. The MAPRD values were less than 6% for both cases and at different MLs. The MAPRD values at ML1, ML2, ML3, ML4, and ML5 were 2.43%, 5.0%, 5.43%, 4.18%, and 3.0%, respectively, when the sea level rise scenario was not considered. With the sea level rise scenario, the corresponding values of MAPRD were 2.51%, 5.15%, 5.63%, 4.22%, and 3.24%, respectively, at different MLs. It is also noted that more than 80% of the observations were below 9% relative difference for both sea level rise and no sea level rise scenarios. Therefore, incorporation of this type of model parameter uncertainty is important when developing a regional scale saltwater intrusion management model for real world case studies. However, in this limited evaluation study, the effects of the uncertainties arising from uncertain model parameters were not incorporated. These were omitted because the MAPRD values of salinity concentrations at all MLs were less than 6% for both with and without sea level rise scenarios. However, these issues of model parameter uncertainties, along with randomized multidimensional heterogeneity in terms of model parameters, need to be incorporated in a rigorous real world regional scale saltwater intrusion management model.

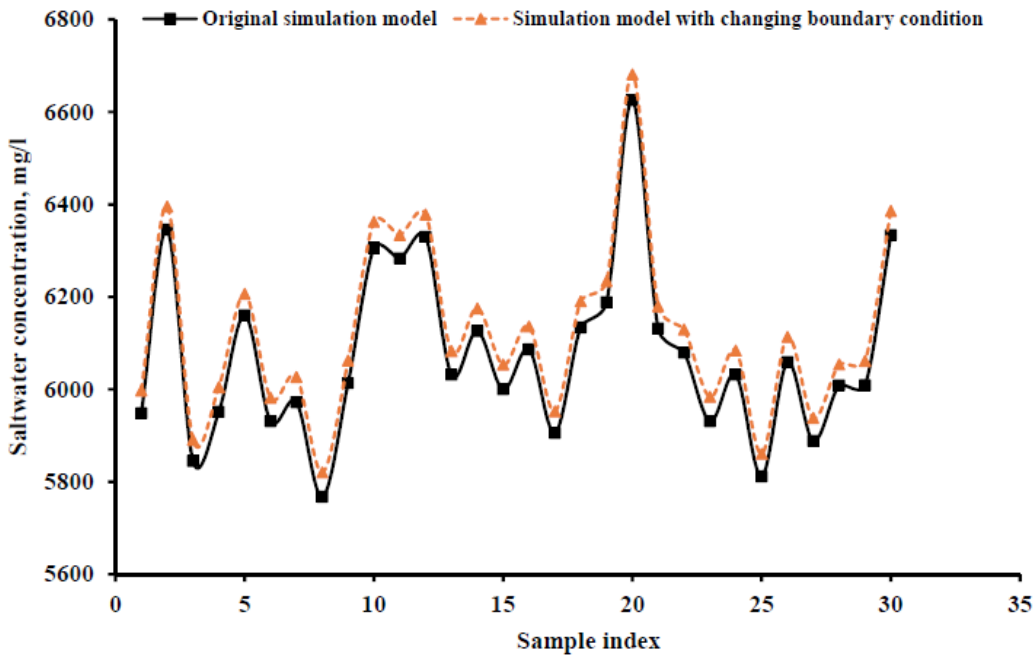


**Figure 8.9 Comparison of simulation results with and without sea level rise at ML5**

## 8.4 Comparison of the performance of ANFIS, GPR, and MARS based meta-models

The performance of the ANFIS, GPR, and MARS based meta-models to approximate density dependent coupled flow and salt transport processes in a multilayered coastal aquifer was evaluated. Each developed meta-model was utilized to predict the salinity concentration at specified MLs with respect to transient pumping stress applied to the aquifer. The performance of ANFIS, GPR, and MARS based meta-models were compared, based on their performances on a new realization of test dataset. Results are summarized in Figure 8.10, Figure 8.11 and Table 8.2, Table 8.3.

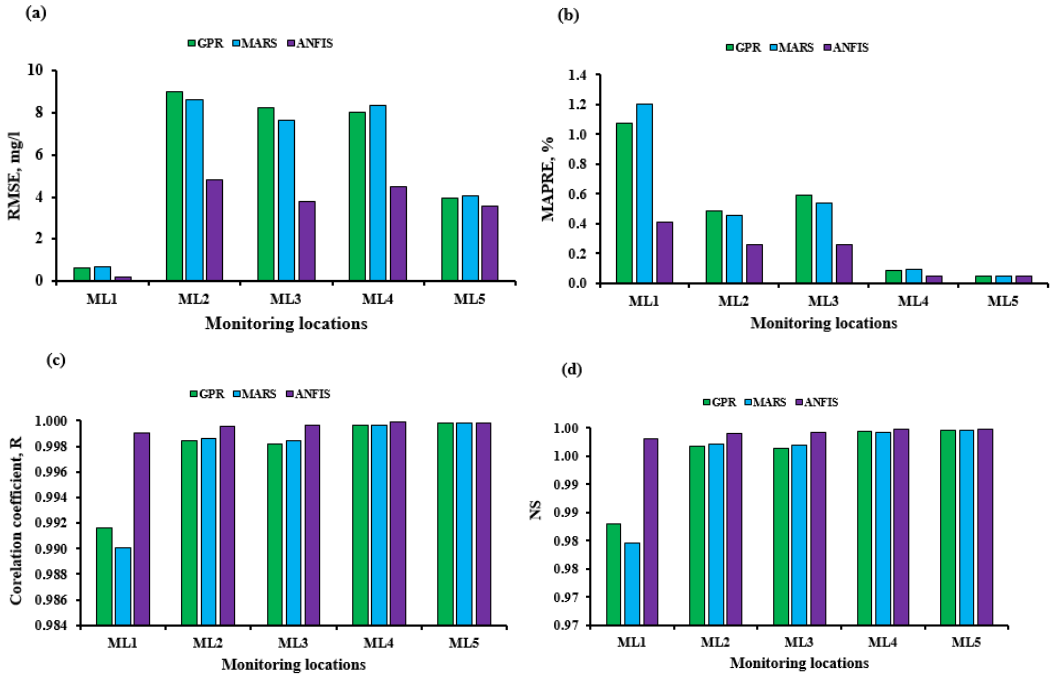


**Figure 8.10 Comparison of prediction performance based on different statistical indices: (a) RMSE, (b) MAPRE, (c) R, (d) NS**

Figure 8.10 illustrates the performance of the proposed meta-models at different MLs based on RMSE, MAPRE, R, and NS criteria, calculated from the actual and predicted saltwater concentration values. However, for the purpose of this performance evaluation, the actual concentrations are those synthetically obtained as the solution of the numerical simulation model in response to water abstraction from the aquifer. The predicted concentrations denote the concentrations predicted by the meta-models. Figure 8.10 demonstrates that all the meta-models produced lower values of RMSE ( Figure 8.10 a) and MAPRE (Figure 8.10 b) as well as higher values of R (Figure 8.10 c) and NS (Figure 8.10 d). Although ANFIS, GPR, and MARS based meta-models were sufficiently accurate in predicting the responses, ANFIS exhibited relatively better performance than GPR and

MARS based meta-models at all MLs, based on all the performance measures. Therefore, it is concluded that ANFIS models' prediction accuracy in terms of capturing the trends of responses at different regions was quite satisfactory. The prediction accuracy of GPR and MARS was different at different MLs: GPR produced better results at ML1, ML4, and ML5 whereas MARS provided better predictions at ML2 and ML3 for all statistical performance measures, as is shown in Figure 8.10.

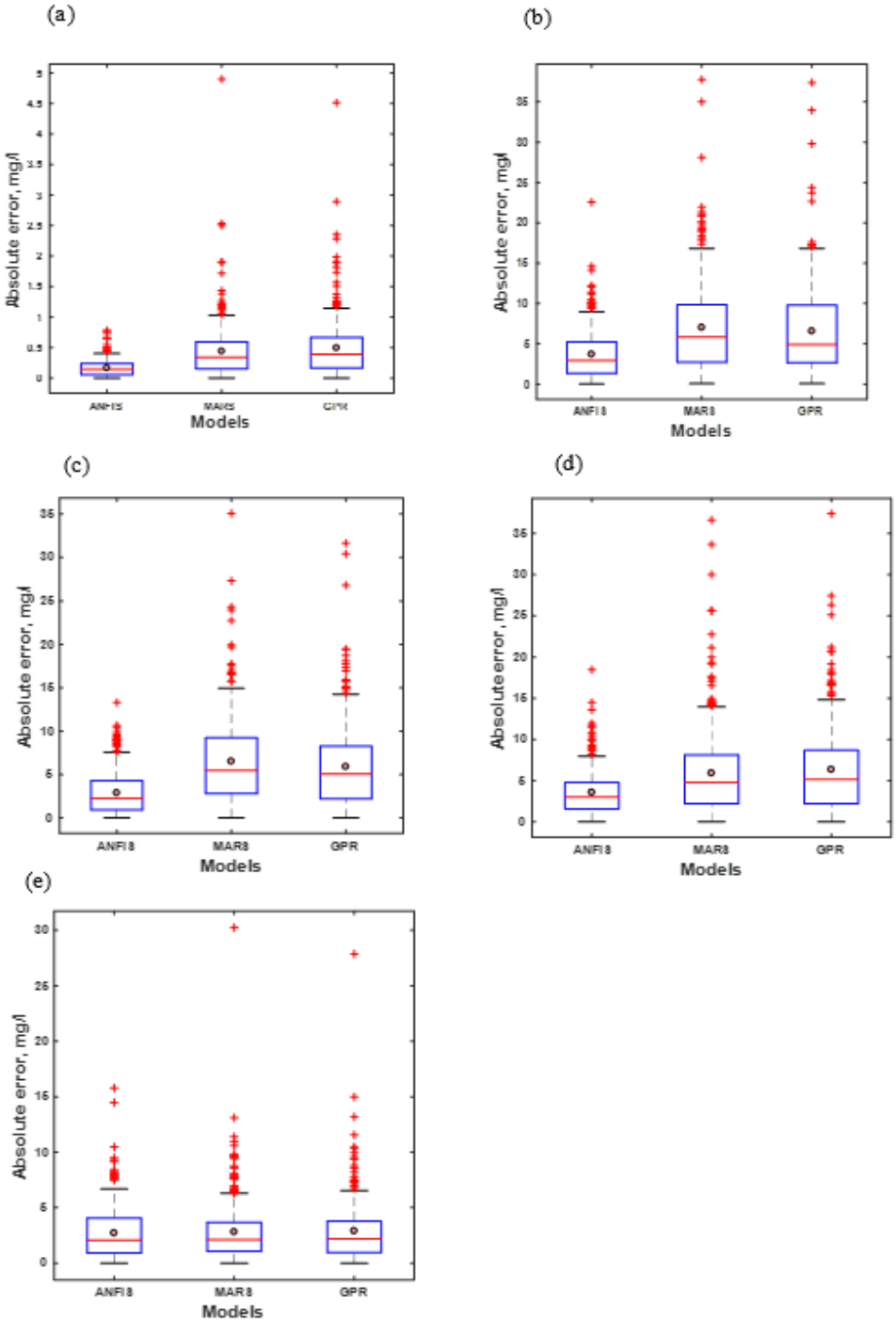


**Figure 8.11 Absolute prediction error boxplots at ML1 (b) ML2, (c) ML3, (d) ML4, and (e) ML5**

Figure 8.11 illustrates boxplots of the absolute errors between the actual and predicted saltwater concentration values obtained by ANFIS, GPR, and MARS meta-models at different MLs. In Figure 8.11, the red horizontal line indicates the median of absolute errors whereas mean of the absolute errors is represented by small circles. Figure 8.11 also demonstrates the superiority of ANFIS model over GPR and MARS models at all MLs.

**Table 8.2 Threshold statistics between actual and predicted concentration values predicted by ANFIS, GPR, and MARS meta-models at monitoring locations ML1–ML5**

| MLs | Selected thresholds | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | <1% | | | <2% | | | <5% | | |
| | ANFIS | GPR | MARS | ANFIS | GPR | MARS | ANFIS | GPR | MARS |
| ML1 | 94.33 | 56.33 | 52 | 100 | 87.33 | 84.33 | 100 | 99 | 98.67 |
| ML2 | 99.67 | 92 | 91 | 100 | 99.33 | 99.33 | 100 | 100 | 100 |
| ML3 | 99 | 84.33 | 88 | 100 | 98.33 | 98.67 | 100 | 100 | 100 |
| ML4 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| ML5 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

*MLs = Monitoring Locations

The TS values provide distribution of errors and give an estimation of the percentage of sample indices whose Relative Error (RE) values are smaller than the pre-defined threshold values. Three threshold values (<1%, <2%, and <5%) were chosen in the present study, so that predictive RE values obtained from all meta-models fall within the selected threshold values. It is apparent from Table 8.2 that ANFIS outperforms both GPR and MARS meta-models in terms of TS criteria.

Another important criterion that should be considered in the selection process of any meta-model for linked S-O methodology is the computational time required to train the model. Table 8.3 presents the computational time required to develop ANFIS, GPR, and MARS meta-models at all MLs. It is noted that the difference in the time requirement was not substantial among different MLs for any specific meta-model. However, for all MLs, MARS required the least training time, followed by ANFIS and GPR meta-models. All developed models are computationally efficient and can be suitable for linking within an optimization framework. However, considering both the prediction accuracy and computational efficiency criteria, the performance of ANFIS meta-model was considered to be superior. Therefore, it can be concluded that ANFIS is the most dependable prediction model to be used in a linked S-O methodology based saltwater intrusion management model at least based on the limited evaluations performed in this study.

**Table 8.3 Training time requirement (sec) to train different meta-models**

| Models | ML1 | ML2 | ML3 | ML4 | ML5 |
|---|---|---|---|---|---|
| ANFIS | 94 | 86 | 85 | 84 | 85 |
| GPR | 131 | 139 | 137 | 132 | 142 |
| MARS | 58 | 55 | 54 | 50 | 55 |

*MLs = Monitoring Locations

## 8.5 Performance of the proposed management model using the best performing meta-model (ANFIS)

Prediction accuracy, computational efficiency, and feasibility of incorporation within a linked S-O approach determine the suitability of any meta-model for the development of regional scale saltwater intrusion management model. Based on the above criteria, ANFIS was selected as a good candidate for replacing and approximating the numerical simulation model within a linked S-O approach to determine Pareto optimal groundwater extraction strategies. Five ANFIS meta-models predicting saltwater concentrations at five different MLs were individually linked to CEMOGA, and the optimization routine was run on a parallel computing platform. Several trials were conducted to choose the optimal parameter values of the optimization routine by using different combinations of these parameters. A population size of 2000, crossover fraction of 0.9, migration fraction of 0.2, and a Pareto front population fraction of 0.7 was found optimal for producing a reliable Pareto optimal front of the management problem. The value of function tolerance was set as $1\times10^{-5}$, while constraint tolerance was set to $1\times10^{-4}$. The optimization algorithm evaluated 4,496,001 different groundwater extraction patterns in 2247 iterations, in order to determine the global optimal solution of groundwater extraction patterns. The Pareto optimal front of the proposed saltwater intrusion management model is presented in Figure 8.12. The Pareto front provides 1400 non-dominated feasible groundwater extraction patterns, in which any one of the extraction patterns can be used without exceeding the maximum allowable saltwater concentrations at specified MLs. The optimization routine took about 6.87 hours to obtain the global optimal solution.

The Pareto optimal solutions shown in Figure 8.12 illustrate the conflicting nature of the two objectives, a necessary condition for a multiple objective optimal management. It can be seen from Figure 8.12 that an increased abstraction from production wells is associated with an increasing amount of barrier well extraction, the water extracted from which cannot be used for beneficial purposes due to its high salinity. Therefore, water withdrawal from

barrier extraction wells can be adjusted based on the demand for beneficial water use without compromising the pre-set maximum allowable salt concentrations at different MLs.

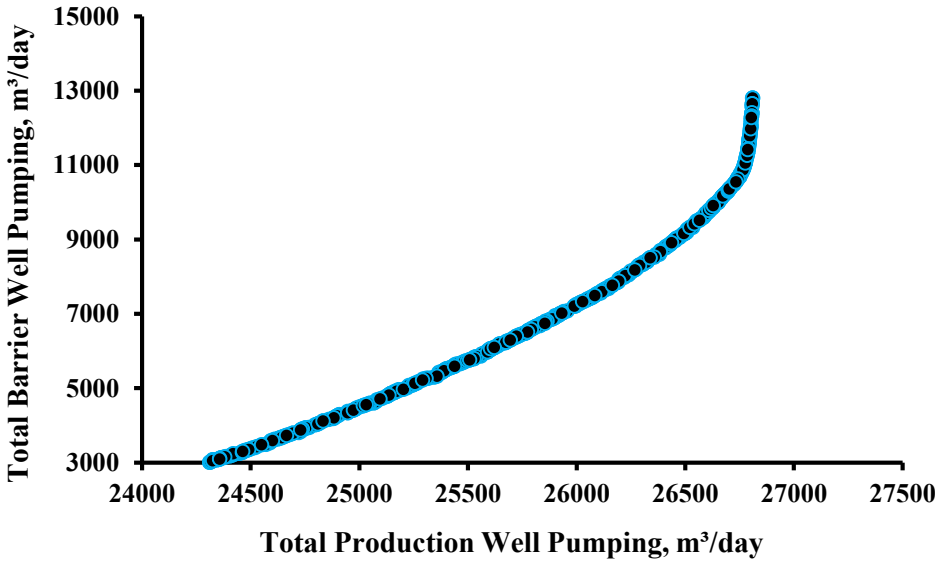


**Figure 8.12 Pareto optimal front of the management model**

The proposed linked S-O based optimization scheme was evaluated with two essential criteria: constraint violation, and closeness to the imposed constraints. The first criterion ensured the correctness of the optimization procedure, whereas the second criterion allowed maximum possible groundwater extraction within the imposed constraints. For all solutions in the Pareto optimal front, saltwater concentrations (computed by ANFIS within the optimization framework) were smaller than the pre-specified maximum allowable limits at all MLs. This confirmed the satisfaction of all the imposed constraints as per solution prediction, and as per actual concentrations. Moreover, for the prescribed groundwater extraction patterns, ANFIS meta-models (within the optimization framework) provided saltwater concentration values very close to the pre-specified limits. This implies that the optimization model converged to the upper limit of constraints. In addition, the prescribed management strategy in terms of optimized groundwater extraction provided by the ANFIS meta-model based optimization routine was verified with the results obtained from the original numerical simulation model. For this, 20 randomly selected groundwater extraction patterns obtained from different regions of the Pareto front were used to compare the ANFIS predicted and numerical model simulated saltwater concentration values. It was demonstrated that the numerical model simulation results were very close to the ANFIS predictions within the optimization model. Percentage Relative Error (PRE) values were very small for all selected solutions at all MLs, indicating the reliability of the proposed linked S-O based saltwater intrusion management model. At ML1, the PRE was less than 2%, whereas

the PRE values at ML2, ML3, ML4, and ML5 were close to 1%, less than 1%, less than 0.5%, and less than 0.1%, respectively. Therefore, the proposed ANFIS-CEMOGA based saltwater intrusion management model is capable of obtaining accurate solutions for optimal groundwater extraction from a set of beneficial pumping wells and barrier extraction wells for a multilayered coastal aquifer system.

## 8.6 Conclusions and recommendations

In this chapter, a linked S-O based saltwater intrusion management model is proposed for a multilayered coastal aquifer system subjected to climate change induced sea level rise. The proposed methodology considers seasonal variation of river water stage, river water concentration near the sea, and climate change induced sea level rise for the specified management period. ANFIS, GPR, and MARS based meta-models were developed as a computationally cheaper substitute of the density reliant coupled flow and salt transport processes in the proposed saltwater intrusion management model. These meta-models were trained from datasets of predictor-response arrays of groundwater withdrawal and resultant saltwater concentrations obtained through a density dependent 3-D coupled flow and salt transport numerical simulation model. The meta-models were compared based on their prediction accuracy, reliability, and computational efficiency. Results demonstrated that both models were capable of capturing the trend of coupled flow and salt transport processes of a multilayered coastal aquifer system. However, based on a closer look at prediction accuracy and computational efficiency, the ANFIS based meta-model was selected as a computationally cheap and feasible soft computing tool for incorporation into the linked S-O approach. The proposed methodology was employed to derive a solution to the multiple objective optimization formulation consisting of two conflicting objectives of groundwater extraction. The parallel computing platform of MATLAB was used to implement the proposed methodology to achieve further computational efficiency. For assessing the proposed methodology, an illustrative multilayered coastal aquifer system was selected as a synthetic case study. The optimal solution obtained utilizing the meta-model was verified using the actual simulation model, in which optimal pumping values derived from the optimal solution served as inputs to the simulation model. It was demonstrated for an illustrative coastal aquifer study area that extraction of water according to the prescribed management strategy can limit the salt concentrations at MLs to pre-specified limits.

Adaptive meta-models can be useful and efficient, especially when the initial data used for training the meta-models is small and the range is extensive. This option also requires going back to the numerical simulation model for retraining in a modified data range, which may not be computationally very efficient. In addition, the training can be restricted

to subdomains, and not encompass the entire decision space. Therefore, this approach may be largely dependent on the candidate or suboptimal solutions and may not result in a global optimal solution.

The present study considers a stratified coastal aquifer having four distinct layers of aquifer materials in which the materials within each layer are considered to be homogeneous. The issue of heterogeneity is also important, and the degree of heterogeneity may determine how this issue can be addressed in a real life study area. While the performance evaluation results presented here are based on spatial or layered heterogeneity, randomized heterogeneity, assuming the heterogeneity is a random field, has not been addressed in this study. Future research might be directed towards extending the application of this methodology to heterogeneous coastal aquifer systems incorporating random fields of different aquifer parameters. In addition, the present study utilizes a parallel pool of worker machines within the local clusters of a PC for efficiently executing the optimization routine. For more complex optimization settings, a parallel pool consisting of more PCs or any other high performance-computing platform might be considered. The next chapter will discuss an adaptive management strategy of coastal aquifers using three-dimensional sequential monitoring network design.

# Chapter 9: Adaptive management of coastal aquifers using entropy-set pair analysis based three-dimensional sequential monitoring network design

Partial contents of this chapter are under review for publication, as outlined below:

Roy, D. K., & Datta, B. (2018). Adaptive management of coastal aquifers using entropy-set pair analysis based three-dimensional sequential monitoring network design. *Journal of Hydrologic Engineering,* Under Review.

This chapter covers the adaptive management of coastal aquifers using entropy-set pair analysis based three-dimensional monitoring network design.

## 9.1 Summary

A 3-D compliance monitoring network design methodology is presented in order to develop an adaptive and sequentially modified management policy, which has as its goal to improve optimal and justifiable use of groundwater resources in coastal aquifers. In the first step, an ensemble meta-model based multiple objective prescriptive model was developed using a coupled S-O approach to derive a set of Pareto optimal groundwater extraction strategies. The prediction uncertainty of meta-models was addressed by utilizing a weighted average ensemble using Set Pair Analysis. In the second step, a monitoring network was designed for evaluating the compliance of the implemented strategies with the prescribed management goals due to possible uncertainties associated with field-scale application of the proposed management policy. Optimal monitoring locations were obtained by maximizing Shannon's entropy between the saltwater concentrations at the selected potential locations. The performance of the proposed 3-D sequential compliance monitoring network design was assessed for an illustrative multilayered coastal aquifer study area. The performance evaluations demonstrated that sequential improvements of optimal management strategy is possible by utilizing saltwater concentrations measured at the proposed optimal compliance monitoring locations.

The success of the developed management strategy largely depends on how accurately the prescribed management policy is implemented in real life situations. The actual implementation of a prescribed management strategy often differs from the prescribed planned strategy due to various uncertainties in predicting the consequences, as well as practical constraints including noncompliance with the prescribed strategy. This results in actual consequences of a management strategy differing from the intended results. To bring the management consequences closer to intended results, adaptive management strategies

can be sequentially modified at different stages of the management horizon, using feedback measurements from a deigned monitoring network. This feedback information can be the actual spatial and temporal concentrations resulting from the implementation of the actual management strategy. This feedback information can be obtained from an optimally designed compliance-monitoring network (Sreekanth and Datta, 2014a). The present study proposes a 3-D sequential compliance monitoring network design methodology for developing an adaptive management strategy in a multilayered coastal aquifer system. This methodology addresses both the uncertainties in prediction of concentrations due to uncertainties in parameter values, as well as the noncompliance of users with the prescribed management strategy.

## 9.2 Methodology

The proposed methodology consists of three components: an initial saltwater intrusion management model based on Weighted Average Ensemble of Multivariate Adaptive Regression Spline (WAE-MARS), an uncertainty based 3-D compliance monitoring network design and sequential modification of the initially prescribed management policy. The first step of the proposed adaptive management strategy encompasses developing a coupled S-O based multiple objective saltwater intrusion management model. In the second step, an uncertainty based 3-D monitoring network was designed, using the concept of information entropy. Third step involved modifying the initially prescribed optimal pumping strategies based on the information obtained from the designed monitoring network. Different components of the proposed methodology are briefly described in the following few subsections.

### 9.2.1 Generation of input-output training patterns incorporating parameter uncertainty

Input-output training patterns were generated by simulating aquifer flow and transport processes multiple times, using random sets of different transient groundwater pumping values also randomly paired with uncertain model parameters. The transient groundwater extraction values were obtained through Latin Hypercube Sampling (LHS) (Pebesma and Heuvelink, 1999), from a uniform distribution within the total input space of 0-1300 $m^3$/day. Five different realizations of four uncertain model parameters (hydraulic conductivity, aquifer recharge, bulk density, and compressibility of the aquifer material) were generated for pairing with the transient groundwater extraction values. Realizations of hydraulic conductivity values were obtained from a lognormal distribution with a specific mean and standard deviation of the associated normal distribution. Aquifer recharge, bulk density, and compressibility realizations were generated from LHS uniform distributions. These uncertain

model parameters, randomly paired with transient groundwater extraction values, were used as inputs to the numerical simulation model in order to obtain saltwater concentration values at specified Monitoring locations (ML) as outputs. Five different sets of input-output training patterns were generated using five randomized sets of uncertain model parameters to develop five MARS meta-models at specified MLs.

### 9.2.2 Weighted average ensemble: Set Pair Analysis (SPA)

An individual meta-model often fails to capture the true trend in the input-output patterns, especially when uncertainty in the model parameters is incorporated in the original simulation model to generate the input-output datasets. In such situations, an ensemble of meta-models is an efficient approach to representing the associated uncertainty of the system. Furthermore, an ensemble of meta-models protects against the prediction uncertainty of meta-models by utilizing each model's unique features in predicting the output. Individual members of the ensemble are combined by either simple averaging (Roy and Datta, 2017b; Roy and Datta, 2017c; Sreekanth and Datta, 2011b) or by a weighted averaging technique (Goel et al., 2007; Hou et al., 2017; Zerpa et al., 2005). The weighted average concept is based on assigning larger weights to more accurate meta-models, and vice versa, to equilibrate the prediction capabilities of the best and the worst meta-models within the ensemble. It can be represented by the following equation

$$Output_{WAE}(inputs) = \sum_{i=1}^{n} Weight_i\,(inputs) \times Output_{m_i}(inputs) \qquad (9.1)$$

where, $Output_{WAE}$ = weighted average prediction from all meta-models, $Output_{m_i}$ = predicted output from the $i^{th}$ meta-model, $Weight_i$ = weight assigned to $i^{th}$ meta-model, and $n$ is the number of meta-models in the ensemble. Here, the assigned weights are a function of $inputs$, and hence the ensemble thus obtained is regarded as adaptive in nature (Zerpa et al., 2005). The present study adopted weighted average ensemble of MARS based meta-models, WAE-MARS using SPA technique (Zhao and Xuan, 1996) to account for prediction uncertainty of meta-models.

The SPA concept considers certainties and uncertainties as a certainty-uncertainty system by researching the relationships between the certainty and uncertainty of an event using three aspects, e.g. identity, discrepant, and contradiction (Zhao and Xuan, 1996). Certainty and uncertainty are interconnected, influenced, mutually restrained and can be transmuted to each other under certain instances. A set pair is a pair of two sets that are connected to each other with a certain degree of connection. Suppose set $A$ and a relative set $B$ form a set pair $S(A, B)$, and the characteristics of set $A$ and the relative set $B$ are denoted by $N$ different terms in sets $A$ and $B$, respectively. Then, the three-element connection degree between these two sets is calculated as

$$\mu = \frac{I}{N} + \frac{D}{N} i + \frac{C}{N} j \tag{9.2}$$

where, $I$, $D$, and $C$ represents the numbers of identical, discrepant, and contrasting characteristics, respectively; $i$ denotes the coefficient of discrepancy and $i \in [-1, 1]$; $j$ is the coefficient of contrast that usually takes the value of $-1$. The identical, discrepant, and contrast degrees of the set pair are represented by $a = \frac{I}{N}$, $b = \frac{D}{N}$, and $c = \frac{C}{N}$, respectively. The connection degree can then be defined as follows

$$\mu = a + bi + cj; \quad a + b + c = 1 \tag{9.3}$$

In this study, a four-element connection degree based set pair was implemented. The set pair was divided into four groups based on the Percentage Absolute Relative Deviation (PARD) between the meta-model predictions and the simulation model outputs: identity, mild discrepancy, severe discrepancy, and contrast (Hou et al., 2017). The PARD values of less than 4% were considered as identical, PARD values between 4% and 8% were regarded as mild discrepancy, PARD values between 8% and 12% were regarded as severe discrepancy and PARD values greater than 12% were considered as contrast. The connection degree between the $m^{th}$ meta-model and the numerical simulation model is expressed as (Hou et al., 2017).

$$\mu_m = \frac{I^m}{N} + \frac{D_1^m}{N} i_1 + \frac{D_2^m}{N} i_2 + \frac{C^m}{N} j \tag{9.4}$$

where, $A = \{R_1, R_2, \dots, R_N\}$ is the set of responses in terms of saltwater concentrations obtained as outputs from the numerical simulation model; $B_m = \{R_1^m, R_2^m, \dots, R_n^m\} (m = 1, 2, \dots, k)$ is the set of responses in terms of saltwater concentrations obtained from the predictions of $m^{th}$ meta-model, and the associated set pair is denoted by $\{A, B_m\}$.

Typically, the values of discrepancy degree uncertainty coefficients are taken as $i_1 = 0.5$ and $i_2 = -0.5$ while the uncertainty coefficient of contrast takes the value of $j = -1$ (Hou et al., 2017). Then, the set pair weight of the $m^{th}$meta-model is calculated by

$$W_m = \frac{M_m}{\sum_{m=1}^{k} M_m} \tag{9.5}$$

where, $W_m$ = set pair weight of the $m^{th}$ meta-model, $M_m = 0.5 \times \mu_m + 0.5$ and $k$ = number of meta-models in the ensemble.

### 9.2.3 Uncertainty based 3-D monitoring network design

An optimal monitoring network is essential to obtain feedback information on the true impacts of the field level implementation of the prescribed management policy, adherence to

which is expected to deviate due to operational and groundwater parameter uncertainties. Based on the monitoring information from a specified number of monitoring locations, management policy for the future time steps can be adjusted. Typically, monitoring information should be collected from as many locations as possible in order to better understand the response of the system with respect to the field scale deviations of the prescribed strategies. This is seldom possible in real situations, in which the number of monitoring location installations is constrained by budgetary limitations. Therefore, locations for maximum allowable number of monitoring wells should be selected at points where the uncertainties in terms of deviations of saltwater concentrations among different scenarios are maximum. These scenarios are generated by simulating aquifer processes using different realizations of a randomized set of groundwater extraction values and other model parameter values. One hundred realizations of salinity concentrations were obtained by randomized pairing of five sets of uncertain model parameters (hydraulic conductivity, bulk density, compressibility, and aquifer recharge), with 20 sets of transient groundwater extraction values. These 20 sets of transient groundwater extraction values were selected by varying the pumping values obtained from the selected optimal pumping strategy from the Pareto optimal front of the prescribed management policy. A number of potential monitoring locations were selected from three different layers of the aquifer. From these potential locations, optimal locations were selected with the consideration of two constraints: budgetary limitations and uncertainty.

### 9.2.4 Shannon's entropy as a measure of uncertainty

Information entropy, also known as Shannon's entropy (Shannon, 1948) is a useful tool to measure uncertainties in a system. If a probabilistic system has $m$ probable outcomes that are equally probable, then each of the outcomes has a probability $p$ defined by $p = 1/m$. Due to the intrinsic uncertainties associated with the system, the uncertainty that a particular outcome truly occurs is an increasing function of $m$ denoted by $f(m)$ (Robinson, 2008). If $m = 1$, e.g. there is no uncertainty associated with the system, then $f(m = 1) = 0$. Now, if another independent system with $n$ possible outcomes is added to the previous system, then the united system has $m \times n$ probable outcomes. The uncertainties associated with these unified systems would be the sum of the individual systems' uncertainty expressed as

$$f(mn) = f(m) + f(n) \tag{9.6}$$

If this phenomenon is true for all positive integers, then $f(n) = \ln(n)$. Therefore, the uncertainty per outcome is given by

$$U_{outcome} = (1/n) * \ln(n) = -p * \ln(p) \tag{9.7}$$

Total uncertainty is given by

$$U_{total} = -\sum_{i=1}^{n} p_i * \ln(p_i) \tag{9.8}$$

Applied to a general probability distribution rather than discrete-valued outcomes, Shannon's entropy will take the form of the following equation

$$H(X) = -K * \sum_{i=1}^{n} p(x_i) * \ln\ p(x_i) \tag{9.9}$$

where, $K$ is a constant.

In this study, entropy was calculated at potential monitoring locations for different realizations of salinity concentration values. For the $i^{th}$ hydrogeological layer ($i = 1,2,\dots,n_i$; where, $n_i$ is total number of hydrogeological layers considered), entropy is measured from the following matrix of concentrations-realizations ($L$ realizations of $n_j$ potential locations; $l = 1,2,\dots,L$ and $j = 1,2,\dots,n_j$)

$$C_{ij}^{l} = \begin{bmatrix} C_{11}^{1} & C_{12}^{1} & \cdots & C_{1n_j}^{1} \\ C_{21}^{2} & C_{22}^{2} & \cdots & C_{2n_j}^{2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n_i1}^{L} & C_{n_i2}^{L} & \cdots & C_{n_in_j}^{L} \end{bmatrix} \tag{9.10}$$

The ratio of the salinity concentration values at the $j^{th}$ potential location in $l^{th}$ realization for $i^{th}$ layer is given by

$$P_{ij}^{l} = C_{ij}^{l} \Big/ \sum_{i=1}^{n_i} \sum_{l=1}^{L} \sum_{j=1}^{n_j} C_{ij}^{l} \tag{9.11}$$

The information entropy at $j^{th}$ potential location can then be expressed as

$$H_{ij} = -\frac{1}{\ln L} \sum_{l=1}^{L} P_{ij}^{l} \ln P_{ij}^{l} \tag{9.12}$$

### 9.2.5 Formulation of the uncertainty based 3-D monitoring network design

A single objective optimization model is considered for the design of optimal monitoring locations. The objective decides on optimal monitoring locations from the potential locations in which the uncertainty, measured based on information entropy is the highest. The mathematical formulation of the proposed optimal monitoring network design is as follows

$$Maximize \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} H_{ij}\, a_{ij} \tag{9.13}$$

$$H_{ij} = -\frac{1}{\ln L}\sum_{l=1}^{L} P_{ij}^{l} \ln P_{ij}^{l} \tag{9.14}$$

$$P_{ij}^{l} = C_{ij}^{l} \Big/ \sum_{i=1}^{n_i}\sum_{l=1}^{L}\sum_{j=1}^{n_j} C_{ij}^{l} \tag{9.15}$$

Subject to

$$\sum_{i=1}^{n_i}\sum_{j=1}^{n_j} a_{ij} \leq b \tag{9.16}$$

$$\sum_{j=1}^{n_j} a_{ij} \leq l_i \; \forall_i \tag{9.17}$$

$$\sum_{i=1}^{n_i} a_{ij} \leq 1 \; ; \forall_j \tag{9.18}$$

$$a_{ij} \in \text{either 0 or 1} \tag{9.19}$$

where, $n_i$ = total number of potential monitoring locations in hydrogeologic layer $i$; $n_j$ = total number of hydrogeologic layers considered; $a_{ij}$ = a binary decision variable taking the value of either 1 or 0: if $a_{ij}$ = 1 a monitoring location is selected as solution for $j^{th}$ location in layer $i$. $l_i$ = maximum number of monitoring locations to be selected in layer $i$. $b$ = maximum permissible total number of monitoring wells summed over all layers.

Equation 9.17 is an important criterion that can be implemented to impose a particular number of monitoring locations be chosen from each layer. However, in this study no constraint was imposed on the maximum number of monitoring locations chosen in a particular layer, in order to more emphatically emphasize the uncertainties.

$H_{ij}$ = salinity concentration based entropy at $j^{th}$ potential location for $i^{th}$ layer. Equation 9.18 imposes a spatial constraint that ensure that maximum one (1) monitoring location is selected in the vertical direction at any potential location in $(x, y)$ direction. The optimization formulation was solved using LINGO 17 (Schrage, 1999).

### 9.2.6 Field-scale compliance and subsequent optimal pumping strategy modification

Field level compliance with a proposed management policy can be evaluated by collecting information from a designed network of optimal compliance monitoring wells. In this approach, the entire management period is split into shorter time steps. Impacts in terms of the actually implemented strategy are identified in terms of measured concentrations at the designed monitoring locations. These deviations occur between the actual measured

concentrations at the designed monitoring location, and the predicted concentration at these locations at the end of the shorter management time steps as subsets of the management time horizon. These deviations may arise due to incorrect prediction of the impact of a designed management strategy due to parameter estimation uncertainties, or due to the fact that the actually implemented management strategy differs from the prescribed optimal management strategy, possibly due to noncompliance by users. These deviations between actual and predicted concentrations as measured at the designed monitoring locations are then incorporated in the management plan to modify future management strategies.

Initially, a designed network of monitoring wells was obtained as a solution of the monitoring network design model (Equations 9.13 to 9.19). The goal of the model was to locate monitoring locations optimally at different depths of the aquifer, by choosing those locations at which the uncertainty in terms of salinity concentration prediction was large. The prescribed optimal groundwater extraction values for the first time step of the management horizon was fed to the numerical simulation model to obtain the resulting salinity concentrations at the designed monitoring wells. These predicted salinity concentrations were compared with the measured concentrations resulting from the actually implemented groundwater extractions. This was accomplished to assess field level compliance to the implemented strategy. Based on the salinity concentration values obtained at the optimal monitoring wells and the already implemented groundwater extraction values for the first time horizon, the optimal groundwater extraction strategy for subsequent time horizons of the entire management policy was sequentially updated. This was obtained using the coupled S-O approach based on the information obtained from the designed monitoring wells at the end of each time step. The process was continued for the entire duration of the initially prescribed management period.

The procedure of sequential modification of the optimal pumping strategy at the end of each time step is analogous to that proposed in Sreekanth and Datta (2014a), except that the monitoring information in the present study was collected from a 3-D monitoring network design. In addition, unlike Sreekanth and Datta (2014a), the initially prescribed optimal pumping management strategy was developed using a multiple objective optimization procedure. More flexibility was provided in selecting a preferred management alternative from the Pareto optimal set of several alternatives. The 3-D monitoring network was designed for this selected optimal pumping management strategy. The multiple objective optimization formulation for the initially prescribed optimal pumping management strategy is described in the Subsection 5.2.1 (Equations 5.1 to 5.6) of chapter 5. The coupled S-O approach for field level compliance and sequential modifications of the optimal pumping strategy is performed using a single objective optimization formulation given by

$$Maximize: f(Q^{PW}) = \sum_{m=1}^{M}\sum_{t=1}^{T} {}_{m}^{t}Q^{PW} \tag{9.20}$$

Subject to

$$C_i = f(Q^{PW}, Q^{BW}) \tag{9.21}$$

$$C_i \leq C^{max} \tag{9.22}$$

$$\sum_{n=1}^{N}\sum_{t=1}^{T} {}_{n}^{t}Q^{BW} \leq Q^{BW}(selected) \tag{9.23}$$

$$Q^{PW}(min) \leq {}_{m}^{t}Q^{PW} \leq Q^{PW}(max) \tag{9.24}$$

$$Q^{BW}(min) \leq {}_{n}^{t}Q^{BW} \leq Q^{BW}(max) \tag{9.25}$$

where, $Q^{BW}(selected)$ is the total amount of water extraction from the barrier wells for the selected pumping management strategy. At the end of each implementation phase, the implemented pumping values from both the production and barrier wells were set as constant for deciding on the future optimal pumping scenarios. This single objective optimization formulation was solved using the GlobalSearch (MATLAB, 2017b) optimization technique. GlobalSearch optimization repeatedly runs a local solver to generate a global optimal solution. The present study utilized 'fmincon' solver (MATLAB, 2017b) as a local solver, which is able to find minimum value of a constrained nonlinear multivariable function. A maximum value of the objective function is attained by assigning a negative sign in the objective function.

The flow diagram of the proposed 3-D monitoring network design is presented in Figure 9.1.

## 9.3 Results and discussion

The performance of the proposed 3-D compliance monitoring network design and field level implementation induced sequential modification of the prescribed management policy was evaluated using an illustrative multilayered coastal aquifer system, which is described in Subsection 4.2.5 of chapter 4. To account for uncertainty in groundwater parameters, uncertain values of hydraulic conductivity, bulk density, compressibility of the aquifer materials, and aquifer recharge were considered in this work. Means and standard deviations of these uncertain model parameters are presented in Table 9.1. Five realizations of these uncertain parameters were randomly paired with 500 sets of transient pumping values to

generate input-output training patterns for meta-model training. Each MARS based meta-model was trained using 500 input-output training patterns obtained from simulation model outputs inputted with random pairing of the 500 input transient pumping values and one of these 5 realizations of uncertain model parameters. Predictions of each of these MARS meta-models were combined using weighted average ensemble approach using the SPA technique.

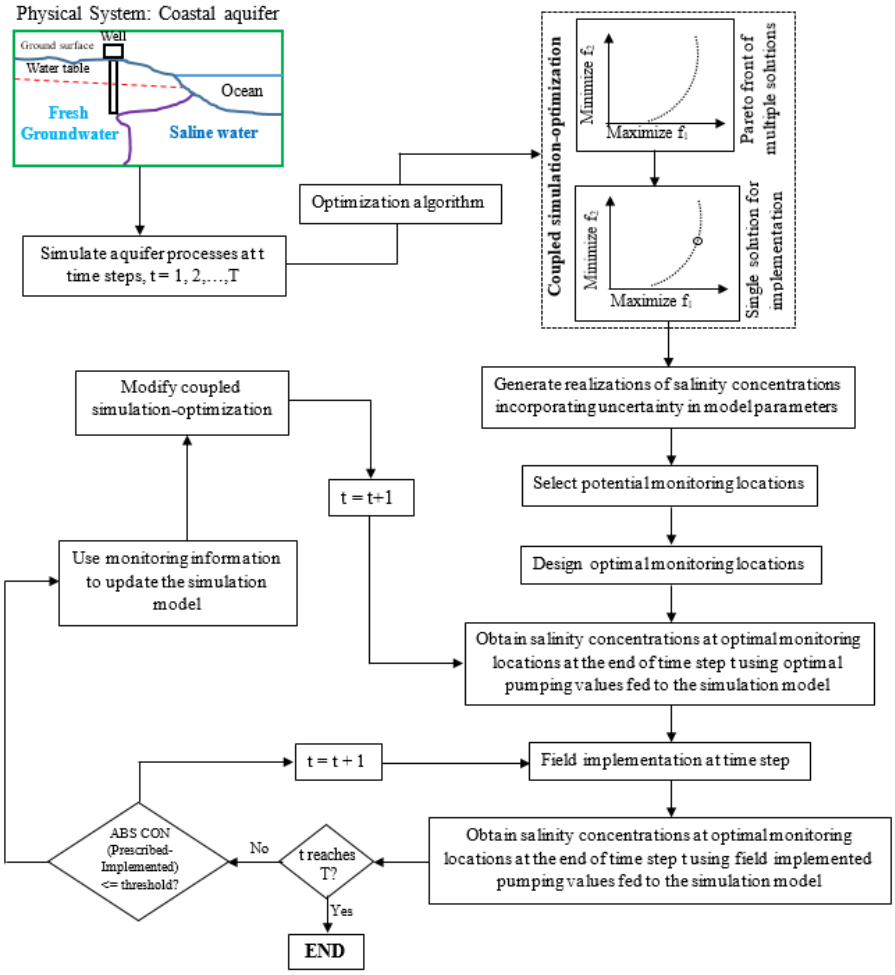


**Figure 9.1 Flow diagram of the proposed methodology**

The results of the WAE-MARS based multiple objective saltwater intrusion management model, optimal compliance monitoring network design, and sequential modification of the prescribed management strategy are presented and discussed in the following sub-sections.

**Table 9.1 Uncertain model parameters with mean and standard deviations**

| Parameters | Units | Material layer 1 | | Material layer 2 | | Material layer 3 | | Material layer 4 | |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| K in x-direction | m/d | 4.96 | 0.30 | 9.75 | 0.25 | 15.0 | 0.20 | 2.85 | 0.25 |
| K in y-direction | m/d | 2.5 | 0.15 | 4.87 | 0.13 | 7.5 | 0.10 | 1.42 | 0.13 |
| K in z-direction | m/d | 0.49 | 0.03 | 0.97 | 0.03 | 1.5 | 0.01 | 0.29 | 0.03 |
| Compressibility | $md^2/kg$ | $8.6 \times 10^{-16}$ | $4.3 \times 10^{-16}$ | $8.6 \times 10^{-18}$ | $4.3 \times 10^{-18}$ | $8.6 \times 10^{-18}$ | $4.3 \times 10^{-18}$ | $8.6 \times 10^{-17}$ | $4.3 \times 10^{-17}$ |
| Bulk density | $Kg/m^3$ | 1650 | 5 | 1600 | 5 | 1550 | 5 | 1700 | 5 |
| Aquifer recharge | m/d | 0.00016 | $2.31 \times 10^{-5}$ | - | | - | | - | |

*K = Hydraulic conductivity, SD = standard deviation

### 9.3.1 Ensemble of MARS meta-models

To incorporate uncertainties related to groundwater model parameters and uncertainties associated with meta-model predictions, the present study utilized WAE-MARS based meta-models in which the individual MARS meta-models were trained and tested from a dataset obtained from a randomized realization of uncertain model parameters and transient groundwater extraction values fed to the numerical simulation model. Set pair weights were assigned to individual MARS meta-models based on their performance on a new test dataset, which was not used for training of each of these models. Percentage Absolute Relative Deviation (PARD) values between the FEMWATER simulated and MARS predicted salinity concentration values were used as a basis for uncertainty criteria of the SPA approach. Five MARS meta-models were developed at each ML, making 25 individual MARS meta-models developed at five MLs. At each ML, the five meta-models developed from a set of uncertain model parameters were combined by assigning set pair weights to each model. The number of data points whose PARD values fall within the SPA uncertainty criteria at ML1 is presented in Table 9.2.

Percentage Absolute Relative Deviation (PARD) is calculated as

$$PARD = \left| \frac{C_{i,O} - C_{i,P}}{C_{i,O}} \right| \times 100 \tag{9.26}$$

Median Absolute Deviation (MAD) is calculated as

$$MAD(C_O, C_P) = median\left(\left|C_{1,O} - C_{1,P}\right|, \left|C_{2,O} - C_{2,P}\right|, \dots, \left|C_{n,O} - C_{n,P}\right|\right) \quad for\ i = 1,2,\dots,n \tag{9.27}$$

Relative Root Mean Squared Error (RRMSE) is calculated as

$$RRMSE = \frac{RMSE}{\frac{1}{n}\sum_i^n C_{i,O}} \tag{9.28}$$

where, $C_{i,O}$ = Observed salinity concentration values (mg/l), and $C_{i,P}$ = predicted salinity concentration values (mg/l)

**Table 9.2 Number of data points having PARD values within the uncertainty indexes of SPA at ML1**

| Models | SPA uncertainty indices | | | |
|---|---|---|---|---|
| | Identity < 4% | Mild discrepancy 4-8% | Severe discrepancy 8-12% | Contrast > 12% |
| MARS 1 | 68 | 28 | 4 | 0 |
| MARS 2 | 64 | 29 | 7 | 0 |
| MARS 3 | 47 | 34 | 15 | 4 |
| MARS 4 | 70 | 28 | 2 | 0 |
| MARS 5 | 81 | 19 | 0 | 0 |

Uncertainty based set pair weights assigned to each individual MARS meta-model are shown in Table 9.3.

**Table 9.3 Set pair weight of individual MARS meta-model**

| MLs | Set pair weight | | | | | Total weight |
|---|---|---|---|---|---|---|
| | MARS 1 | MARS 2 | MARS 3 | MARS 4 | MARS 5 | |
| ML1 | 0.202 | 0.200 | 0.184 | 0.204 | 0.210 | 1 |
| ML2 | 0.203 | 0.200 | 0.139 | 0.223 | 0.235 | 1 |
| ML3 | 0.209 | 0.200 | 0.141 | 0.208 | 0.243 | 1 |
| ML4 | 0.206 | 0.205 | 0.138 | 0.225 | 0.225 | 1 |
| ML5 | 0.206 | 0.208 | 0.181 | 0.203 | 0.203 | 1 |

Prediction performances of the WAE-MARS models at each ML were compared with those obtained from the best MARS models at the corresponding MLs. Table 9.4 presents the comparison results. Table 9.4 shows that WAE-MARS models outperformed single best MARS models at ML1 based on the correlation coefficient (R) criterion, and at ML2 and ML4 based on the Index of Agreement (IOA) criterion. When considering the MPARE criterion, WAE-MARS outperformed the best MARS model at all MLs except at ML5 at which the best MARS model's performance was slightly better than that of the WAE-MARS model. The Relative Root Mean Squared Error (RRMSE) criterion provided similar values for both WAE-MARS and the best MARS models at all MLs.

**Table 9.4 Performances of ensemble MARS models compared to the best MARS models**

| Statistical Indices | ML1 | | ML2 | | ML3 | | ML4 | | ML5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | MARS | WAE-MARS | MARS | WAE - MARS | MARS | WAE - MARS | MARS | WAE - MARS | MARS | WAE - MARS |
| IOA | 0.93 | 0.93 | 0.88 | 0.89 | 0.94 | 0.94 | 0.80 | 0.81 | 0.85 | 0.85 |
| R | 0.87 | 0.88 | 0.80 | 0.80 | 0.89 | 0.89 | 0.69 | 0.69 | 0.74 | 0.74 |
| MAD | 0.80 | 0.72 | 49.0 | 46.0 | 61.7 | 59.6 | 226 | 201 | 206 | 186 |
| MAPRE,% | 2.58 | 2.47 | 5.40 | 5.34 | 6.01 | 6.00 | 4.30 | 4.30 | 3.09 | 3.15 |
| RRMSE | 0.03 | 0.03 | 0.07 | 0.07 | 0.07 | 0.07 | 0.05 | 0.05 | 0.04 | 0.04 |

*IOA = Index of Agreement, R =Correlation Coefficient, MAD = Median Absolute Deviation, MAPRE = Mean Absolute Percentage Relative Error, RRMSE = Relative Root Mean Squared Error

### 9.3.2 Management model

WAE-MARS at five MLs were linked externally to the optimization algorithm CEMOGA to develop the saltwater intrusion management model. The management model provides the maximum amount of water that would be extracted from production wells through the least amount of water withdrawal from the barrier extraction wells, while keeping maximum permissible salinity concentrations at the MLs within the acceptable limits. The optimal sets of parameter values for CEMOGA were selected based on numerical experiments, by using several combinations of different optimization parameters. However, a detailed sensitivity analysis of the combined use of different optimization parameters was not carried out in this example problem. Based on the numerical experiments, the optimization algorithm used a population size of 3000, crossover fraction of 0.95, and a Pareto front population fraction of 0.7. The function and constraint tolerances were set to $1\times10^{-5}$ and $1\times10^{-5}$, respectively. The optimization algorithm used 1247 generations and 3,744,001 function evaluations to converge to the global Pareto optimal solution. The parallel computing platform of MATLAB (MATLAB, 2017c) was utilized to speed up the optimization procedure. The pumping optimization management model developed in this study to address saltwater intrusion in coastal aquifers considered 80 input pumping variables that correspond to water extraction from 16 locations (11 locations for production wells + 5 locations for barrier extraction wells) for a management time horizon of five years. Pumping from both wells at any particular time step of one year was assumed to be constant. The management model provides several alternate pumping strategies in the form of a Pareto optimal front, which shows a trade-off between the two conflicting objectives of the coastal aquifer management problem (Figure 9.2). The performance of the management model was verified by using the optimized pumping values as inputs to the numerical simulation model, and comparing the

resulting salinity concentrations at specified MLs obtained from both the numerical model and the WAE-MARS models. A particular management strategy was randomly selected from the Pareto optimal front of the proposed management strategy in order to develop the sequential compliance monitoring network design. This solution provides a total beneficial production well pumping at the rate of 31945 $m^3$/day at the expense of the total barrier extraction well pumping of 2382 $m^3$/day.

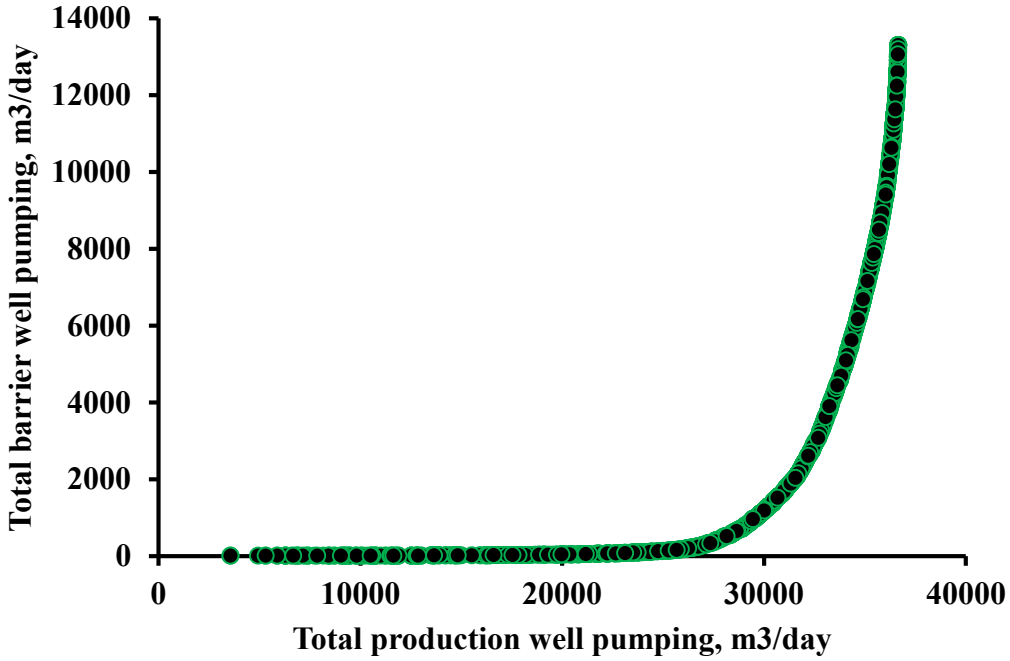


**Figure 9.2 Pareto optimal front for the management strategy**

9.3.3 Uncertainty based selection of potential locations for monitoring networks

The optimal monitoring network was designed by selecting locations from a set of potential locations at different spatial positions. Salinity concentrations at these potential locations were measured by considering the associated uncertainty of the groundwater system. Uncertainty was characterized by several probable realizations of salinity concentrations at different spatial locations, obtained from a set of uncertain model parameters paired with deviated prescribed optimal pumping strategy. This deviation of optimal pupping strategy relates to the actual field implementation of the strategy, which is expected to deviate from the prescribed values of pumping at different production and barrier wells. In this study, uncertainty resulting from hydraulic conductivity, bulk density, aquifer recharge, and compressibility of the aquifer material was considered.

Realizations of hydraulic conductivity were acquired from a lognormal distribution with a specific mean and standard deviation. Realizations of aquifer recharge, bulk density and compressibility were obtained from an LHS technique. Field level implementation

uncertainty of groundwater extraction ranged from –100 to + 100 $m^3$/day. Deviations of these groundwater extraction values were obtained from an LHS technique from a univariate normal distribution for this selected lower and upper bounds. Deviations thus obtained were added to the prescribed optimal pumping values to mimic the actual field level implementation of the pumping values. These uncertain pumping values were then randomly paired with the uncertain model parameters and used as inputs to the numerical simulation model in order to obtain 100 realizations of uncertain salinity concentrations. The potential locations were selected from the salinity intruded regions of the aquifer, at which the salinity concentrations of the aquifer water ranged from 100-28000 mg/l in at least one of these realizations. Three different depths of the aquifer (– 20 m, – 40 m, and – 60 m) were considered for selecting these arbitrary monitoring locations. A total number of 180 (60 spatial locations at three depths) potential locations were considered in this study. Figure 9.3 shows the spatial distribution ($X$ and $Y$ coordinates only) of potential monitoring locations.

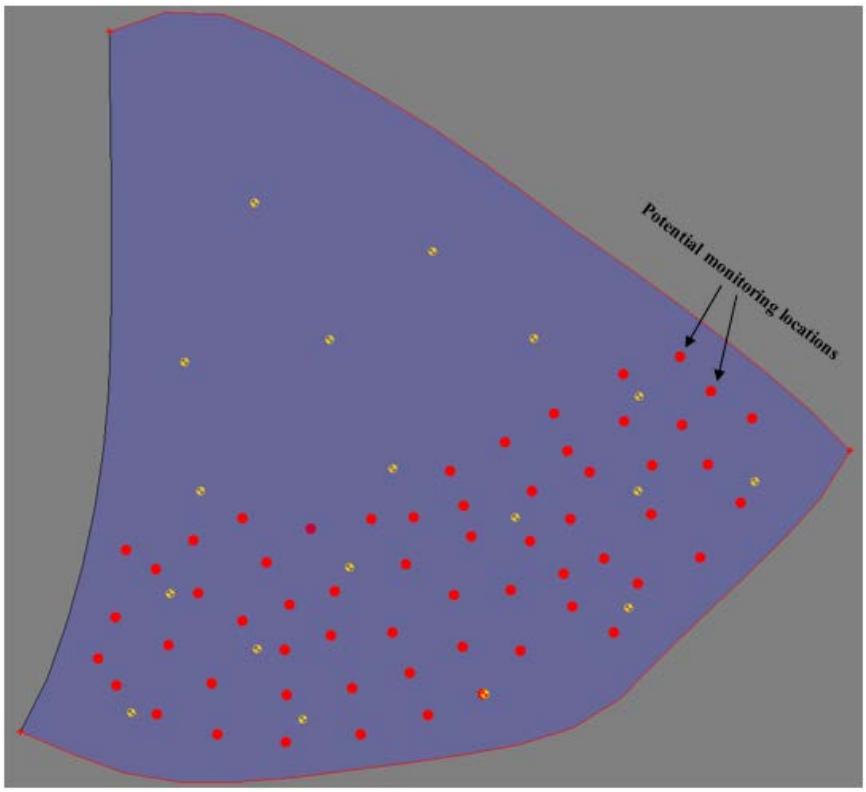


**Figure 9.3 Potential monitoring locations**

9.3.4 Optimal locations for monitoring networks

The optimal number of allowable monitoring locations is always constrained by financial limitations. This constraint limits the number of optimal locations to be designed for monitoring purposes. This constraint was incorporated in the design of the optimal

compliance monitoring network by specifying the maximum allowable monitoring locations for this illustrative study area as 15. The designed network of optimal monitoring locations for 15 allowable wells is shown in Figure 9.4. The design was based on the uncertainty in measured salinity concentrations characterized by maximizing Shannon's entropy in these locations. Shannon entropy was utilized in this study for the first time in determining the variance in salinity concentrations at different locations for the design of an optimal 3-D monitoring network. The optimal monitoring wells were installed at locations where uncertainty in terms of Shannon's entropy was high. The distribution of the monitoring wells was enumerated by summing the Euclidean distances of the monitoring wells from their centroid (Sreekanth and Datta, 2014a). The sum of the centroid distance for 15 wells was 9422 m, which is equivalent to 144.398 m per well per square km of the aquifer. This indicates a well spread network of monitoring wells that may avoid the redundancy of installation. Sreekanth and Datta (2014a) proposed a 2-D monitoring network design whose centroid distance was 35.897 m per well per square km of the aquifer. This indicates the suitability of Shannon's entropy as a measure of uncertainty in optimal design of a set of monitoring locations.

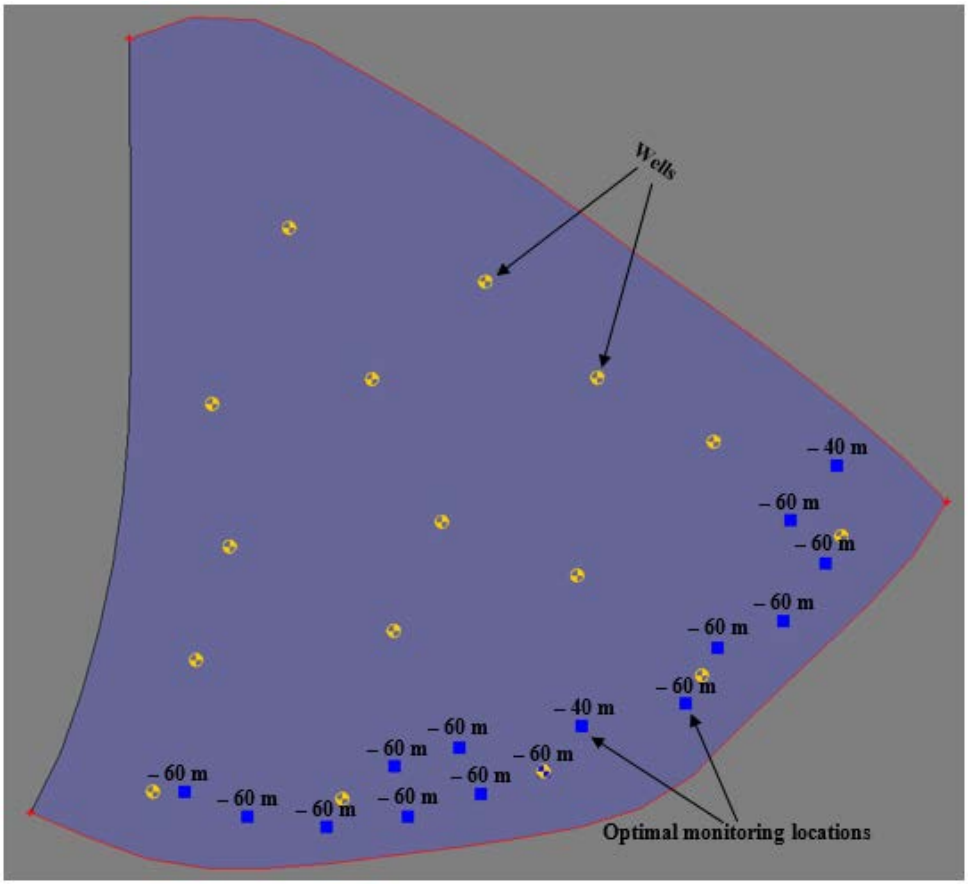


**Figure 9.4 Optimal monitoring locations**

### 9.3.5 Adaptive modifications of optimal pumping strategies

For performance evaluation purpose, a particular optimal extraction strategy selected from the Pareto front of the initial optimal groundwater extraction policy was used for the adaptive modification of the pumping strategies. For this selected extraction strategy, the total production well extraction was 31945 $m^3$/day, against a total barrier extraction well pumping of 2382 $m^3$/day. Salinity concentrations at the optimally designed monitoring locations were measured (for evaluation purposes, it was simulated for the pumping strategy) using these prescribed extraction values at the end of year 1. The corresponding salinity concentrations are presented under Case 1 of year 1 in Table 9.5. These concentration values were expected at the end of year 1 if the prescribed policy would be implemented exactly in the field and the predictions of the corresponding concentrations were correct.

**Table 9.5 Salinity concentrations at the optimal monitoring locations resulting from prescribed and implemented pumping strategies for five years of management period**

| MLs | Salinity concentrations, mg/l | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Year 1 | | Year 2 | | Year 3 | | Year 4 | | Year 5 | |
| | Case 1 | Case 2 | Case 1 | Case 2 | Case 1 | Case 2 | Case 1 | Case 2 | Case 1 | Case 2 |
| 1 | 16192 | 16194 | 16967 | 16976 | 17740 | 17750 | 18603 | 18614 | 19561 | 19573 |
| 2 | 20760 | 20765 | 21377 | 21384 | 22011 | 22019 | 22724 | 22733 | 23485 | 23495 |
| 3 | 26432 | 26462 | 26804 | 26833 | 27416 | 27435 | 28012 | 28031 | 28418 | 28444 |
| 4 | 11276 | 11274 | 12025 | 12024 | 12678 | 12681 | 13480 | 13484 | 14509 | 14513 |
| 5 | 25544 | 25556 | 26000 | 26014 | 26508 | 26523 | 27062 | 27077 | 27580 | 27595 |
| 6 | 21468 | 21488 | 22037 | 22057 | 22580 | 22603 | 23129 | 23154 | 23711 | 23735 |
| 7 | 9107 | 9103 | 9814 | 9812 | 10473 | 10474 | 11225 | 11227 | 12166 | 12168 |
| 8 | 19355 | 19441 | 19823 | 19924 | 20340 | 20421 | 20790 | 20893 | 21247 | 21355 |
| 9 | 20597 | 20647 | 21265 | 21313 | 21822 | 21870 | 22361 | 22409 | 22931 | 22980 |
| 10 | 20793 | 20767 | 21608 | 21588 | 22259 | 22288 | 22926 | 22956 | 23633 | 23663 |
| 11 | 22509 | 22515 | 23252 | 23260 | 23875 | 23889 | 24466 | 24479 | 25092 | 25106 |
| 12 | 16003 | 16030 | 16930 | 16967 | 17717 | 17755 | 18465 | 18504 | 19281 | 19320 |
| 13 | 6878 | 6893 | 8124 | 8154 | 9293 | 9282 | 10498 | 10488 | 11843 | 11835 |
| 14 | 8940 | 8929 | 9808 | 9796 | 10591 | 10586 | 11452 | 11443 | 12426 | 12418 |
| 15 | 7679 | 7673 | 8857 | 8871 | 9974 | 9983 | 11198 | 11203 | 12420 | 12423 |

*MLs = Monitoring Locations, Case 1 = Prescribed, Case 2 = Implemented

However, in order to evaluate the impact of the deviation of actual salinity levels from the expected or predicted salinity levels for the selected pumping strategy, this deviation was simulated based on perturbing the prescribed pumping values to reflect deviations in the

actual implementation. It was also assumed that the predicted average vertical recharge had also deviated by 5% of originally expected value. This scenario may also result in the originally predicted salinity concentrations differing from the measured ones after the implementation of the optimal pumping strategy. These evaluated scenarios reflect the fact that deviations from the prescribed pumping policy and possible changes in parameters such as average vertical recharge are plausible in the real life scenario. To incorporate these scenarios, this performance evaluation study incorporated random perturbations in groundwater extraction within 0-20% of the prescribed values, and a decrease in aquifer recharge of 5% in subsequent field level implementations of the management policy. The corresponding salinity concentrations at the end of year 1 with respect to this perturbation were obtained using the simulation model and are shown under Case 2 of year 1 in Table 9.5. It is evident from Table 9.5 that there were deviations in the concentration values at the designed monitoring locations as a result of the deviations in the field level implementation.

Based on the monitored concentrations at the end of year 1, the future pumping strategy for the rest of the management timeframe was updated, while keeping the original management constraints intact. A new optimal solution in terms of total production well pumping for the future management periods is presented in Table 9.6.

**Table 9.6 Matrix of prescribed and implemented total production well pumping during subsequent periods of entire management period**

| | Total production well pumping, $m^3$/day | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Year 1 | | Year 2 | | Year 3 | | Year 4 | | Year 5 | |
| | P | I | P | I | P | I | P | I | P | I |
| Year 1 | 5045 | 4939 | | | | | | | | |
| Year 2 | 5678 | | 7070 | 7076 | | | | | | |
| Year 3 | 5425 | | 8507 | | 8749 | 8754 | | | | |
| Year 4 | 7232 | | 11962 | | 11903 | | 11960 | 11966 | | |
| Year 5 | 8564 | | 13819 | | 13925 | | 13817 | | 13930 | - |
| SUM | 31945 | | 41358 | | 34577 | | 25778 | | 13930 | |

*P = Prescribed, I = Implemented

The objective function value (including year 1 implemented strategies = 4939 $m^3$/day and future years prescribed strategies = 41358 $m^3$/day) was 46297 $m^3$/day, which showed an improvement of the objective function value compared to the original optimal value. The corresponding salinity concentrations at the designed monitoring locations using the optimal prescribed solutions for year 2 were obtained from the numerical simulation model and are given under Case 1 of year 2 in Table 9.5. Case 2 of year 2 corresponds to the resulting

salinity concentrations obtained from perturbations of the prescribed strategies. This information from the optimal monitoring wells provides the optimal pumping strategies for the rest of the management periods (years 3, 4, and 5). The objective function value obtained was 46592 $m^3$/day, which was more than the original optimal solution as well as the second optimal solution (for years 2, 3, 4, and 5 including the year 1 implemented strategy). The salinity concentrations for prescribed and field implemented cases at the optimal monitoring locations and the modified objective function values for the rest of the management periods are presented in Table 9.5 and Table 9.6, respectively. The total objective function values for the fourth (for years 4 and 5) and fifth (for year 5) optimal solutions were 46547 $m^3$/day and 46666 $m^3$/day, respectively.

The solution results presented in Table 9.6 show that there is substantial deviation of the newly adopted optimal pumping strategy from the originally specified pumping strategy. These results show that the pumping strategy will need modification, as the actually implemented strategy may not be exactly the same as the one prescribed. Therefore, utilizing newly designed and implemented monitoring networks can help effectively and sequentially modify the pumping strategy to be implemented. The feedback information from actual measurements obtained, utilizing the newly designed monitoring networks, can also incorporate information on deviations from expected and actually realized consequences of a prescribed strategy due to prediction errors. It is also observed that the amount of total beneficial pumping may also increase because of the feedback based sequential management strategy implementation, while keeping the barrier well pumping at originally prescribed levels. It may be worthwhile to point out that these results are for performance evaluation purposes only, and therefore the field measurements are reflected in synthetic measurements obtained by simulation.

## 9.4 Conclusions

This chapter presents an adaptive management strategy to control saltwater intrusion in coastal aquifers with the aid of an optimally designed network of monitoring wells. The management policy explicitly incorporates the uncertainty of groundwater parameters, uncertainties arising from meta-model predictions, and operational uncertainty due to deviations from prescribed strategy during field level implementation. Different realizations of randomized parameters randomly paired with groundwater extraction values were used as inputs to the numerical simulation model to generate five realizations of input-output training patterns. These realizations were used to train five individual MARS based meta-models, which were integrated to develop an ensemble using set pair weights assigned to these meta-models. This ensemble, accounting for prediction uncertainty based on SPA theory, was

externally linked to the optimization algorithm in order to develop the initial management policy for a management period of five years. For the prescribed management strategy, a 3-D monitoring network was designed to monitor and modify the implemented saltwater intrusion management strategy. Several realizations of salinity concentrations, resulting from a set of uncertain model parameters and perturbed groundwater extraction values, were incorporated in the design of this uncertainty based 3-D monitoring network. Monitoring locations were selected from three different depths of the aquifer, at which the uncertainty in terms of salinity concentrations in these locations was higher. This was accomplished by proposing a new method of uncertainty quantification, i.e., Shannon's entropy, in the objective function of the optimal monitoring network design model for saltwater intrusion in coastal aquifers. The designed monitoring wells provide salinity concentrations resulting from the field level implementation deviations of the prescribed management strategies. Information collected from these optimally designed monitoring wells were used to sequentially update the optimal pumping management strategies for the future management periods. The results obtained from an illustrative multilayered coastal aquifer system demonstrated the potential applicability of the Entropy-SPA based 3-D monitoring network design to sequential and adaptive management of coastal aquifers. The effects of hydrogeological parameter estimation uncertainties were incorporated through the ensemble of meta-models used for the prediction of the resulting salinity concentrations. Sequential updating of these parameter estimations can be incorporated in a future study.

The next chapter presents an application of the developed methodology for developing a regional scale multiple objective management strategy for a coastal aquifer site in the tropical country, Bangladesh.

# Chapter 10: Modelling and management of salinity intrusion in a coastal aquifer system of Barguna District, Bangladesh

This chapter presents an application of the coupled S-O based multiple objective management strategy for a coastal aquifer site in the tropical country, Bangladesh.

## 10.1 Summary

Pumping induced saltwater intrusion in coastal aquifers is a challenging problem, due to the increased abstraction of groundwater resources to meet the growing demand for freshwater supplies. Sustainable beneficial water abstraction from coastal aquifers can be ensured by optimizing water abstraction from a set of production and barrier wells. An optimal pumping management strategy can be prescribed for a coastal aquifer system by utilizing an integrated simulation-optimization (S-O) approach. In this study, the integrated S-O approach was used to develop a saltwater intrusion management model for a real world coastal aquifer system in Barguna district of southern Bangladesh. The aquifer processes were simulated by using a calibrated and validated 3-D finite element based combined flow and solute transport numerical model using the code FEMWATER. The modelling and development of strategies for the management of seawater intrusion processes was performed based on the very limited quantity of available hydrogeological data. The model was calibrated with respect to hydraulic heads for a period of five years from April 2010 to April 2014. The calibrated model was validated for the next three years' period from April 2015 to April 2017. The calibrated and partially validated model was then used within the integrated S-O management model to develop optimal groundwater abstraction patterns to control saltwater intrusion in the study area. Computational efficiency of the management model was achieved by using a MARS based meta-model emulating the combined flow and solute transport processes of the study area. This limited evaluation demonstrates that a planned transient groundwater abstraction strategy, acquired as solution results of a meta-model based integrated S-O approach is useful for developing management strategy for optimized water abstraction, with saltwater intrusion control. This study shows the capability of the MARS meta-model, based an integrated S-O approach, to solve real-life complex coastal aquifer management problems in an efficient manner.

## 10.2 Methodology

The integrated S-O approach (Dhar and Datta, 2009a) is the core constituent of the present study. The study applies this integrated S-O approach in order to develop optimal pumping

management schemes for a real world coastal aquifer system. The basic components of the adopted S-O approach are: (1) a finite element based 3-D numerical simulation model to simulate the physical processes, (2) a properly trained and validated meta-model that approximates numerically simulated salinity concentrations at designated monitoring locations, and (3) an optimization algorithm to search for the optimal groundwater extraction patterns from the aquifer. Brief descriptions of each of these components are provided in the previous chapters of this thesis. The MARS based meta-model was used to achieve computational efficiency in the integrated S-O approach. A description of the MARS based meta-models is provided in Subsection 2.5.3 of chapter 2. Details of developmental steps of MARS based meta-models are presented in Subsection 4.3 of chapter 4. Performance evaluation indices of the MARS based meta-models are also mentioned in the previous chapters. However, the Kling–Gupta Efficiency (KGE) (Gupta et al., 2009) is introduced in this chapter as a new performance evaluation index. The KGE is calculated as

$$KGE = 1 - ED = 1 - \sqrt{(R-1)^2 + (\propto -1)^2 + (\beta - 1)^2} \tag{10.1}$$

$$\propto = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C_{i,P} - \overline{C_P}\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(C_{i,O} - \overline{C_O}\right)^2}} \tag{10.2}$$

$$\beta = \frac{\frac{1}{n}\sum_{i=1}^{n} C_{i,P}}{\frac{1}{n}\sum_{i=1}^{n} C_{i,O}} \tag{10.3}$$

where, $C_{i,O}$ = observed salinity concentration values (mg/l), $C_{i,P}$ = predicted salinity concentration values (mg/l), $\overline{C_O}$ = mean values of the simulated salinity concentration values (mg/l), $\overline{C_P}$ = mean values of the predicted salinity concentration values (mg/l), $n$ = number of data points, $ED$ = Euclidian distance from the ideal data points, $\propto$ = relative variability in the simulated and predicted salinity concentration values, $\beta$ = ratio between the mean predicted and mean simulated salinity concentration values representing the bias.

## 10.3 Study area

The study area is located in Barguna district, one of the coastal districts of Bangladesh. It is located in the southern part of Bangladesh, lying between 21º48′ and 22º29′ north latitudes and between 89º52′ and 90º22′ east longitudes, and is also within the tropics. The land area is nearly flat, having rivers and estuarine creeks with regular low and high tides. The Barguna district belongs to the Gangetics tidal floodplain that is highly susceptible to storms and tidal flooding. The coastal aquifer is contaminated by salinity intrusion

resulting from large amounts of extraction. The study area consists of two upazillas (administrative units) in the Barguna district: Patharghata (258.63 km$^2$) and Barguna sadar (339.54 km$^2$). The southern part of the study area is surrounded by the mangrove forest and the Bay of Bengal. The aerial map of the study area is presented in Figure 10.1.

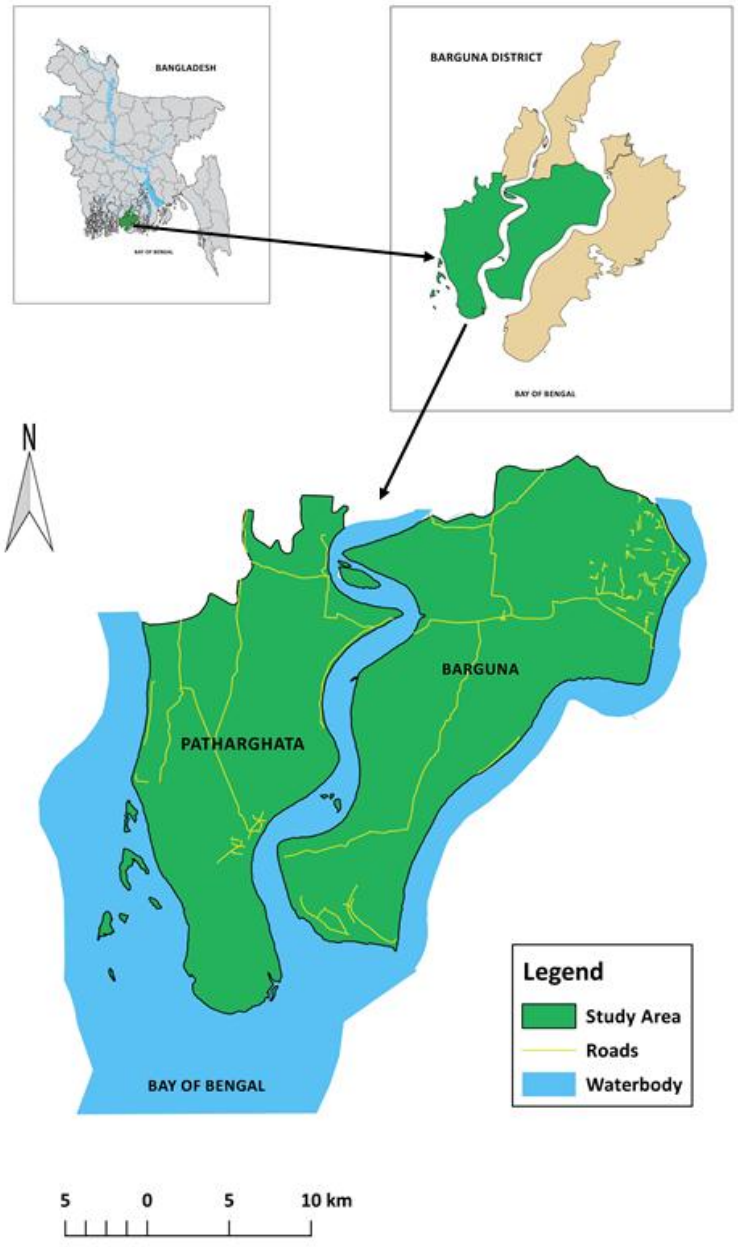


**Figure 10.1 Location and aerial map of the study area**

During the wet monsoon season, water enters the study area through recharge from rainfall and the water discharges into the river systems (Faneca Sanchez et al., 2015). In contrast, the study area receives water through infiltration from the rivers during the dry cold seasons, the quality of groundwater depending on the salt concentration of the surface water.

Groundwater abstraction has very little influence (only 4%) on the water balance of the study area in the wet monsoon season. In the dry season, groundwater abstraction accounts for about 50-70% of groundwater leaving the study area (Faneca Sanchez et al., 2015). Many wells are in use in each village, each shared by a single family or a group of several families, and 97% of the rural population uses groundwater for its water supply (Kinniburgh et al., 2003). Deep tube well depths range between 60-100 m whereas shallow tube well depths range between 10-60 m (Kinniburgh and Smedley, 2001; Mondal and Saleh, 2003).

## 10.4 Modelling of seawater intrusion processes

### 10.4.1 Model development

The major difficulty in developing a regional scale saltwater intrusion model is the scarcity of reliable data for hydrogeological parameters and groundwater use in the area. Data from various sources were utilized in the model development, based on the best possible subjective judgement. Hydraulic head data for a period of eight years (April 2010 to April 2017) were collected from the website of the Processing and Flood Forecasting Circle of Bangladesh Water Development Board (BWDB) (Bangladesh Water Development Board. Processing and Flood Forecasting Circle, 2018). Two hydraulic head values were obtained for the study area (one from each upazilla). The observation well at Patharghata upazilla is located at 22° 11' 8" north latitude and 89° 59' 21" east longitude. The other observation well at Barguna sadar upazilla is located at 22° 9' 26" north latitude and 90° 6' 2"east longitude. Surface water level and salinity data were collected from the website of BWDB (Bangladesh Water Development Board. Processing and Flood Forecasting Circle, 2018). Based on the data, a specified head was assigned at the upstream end of the river and interpolated over the lengths of the river. A constant concentration of river water salinity was assumed. There are three distinct seasons in Bangladesh: 1) a cold dry winter from November to February, 2) a humid hot summer from March to May and 3) a cool rainy monsoon season from June to October. These seasonal variations only affect the top layers of the aquifer (Faneca Sanchez et al., 2015). Therefore, the average values of the parameters for these three different seasons were used in this study.

Previous studies in the Bengal Delta modelled a very large area (Faneca Sanchez et al., 2015; Michael and Voss, 2009) by assuming groundwater abstraction per unit area of the model domain. The withdrawals were distributed based on estimates made for each administrative unit. Of note, it is difficult to represent pumping in the model domain as individual wells because of the large number of unreported wells and the large scale of the study area. However, the main aim of this study was to prescribe optimal groundwater

abstraction patterns to control saltwater intrusion. In addition, the implementation of hydraulic control measures in the form of barrier abstraction wells was also used as a measure of salinity control. Therefore, the exact location of the point pumping was approximated in the present study based on the land use pattern of the study area. In conformance with the total water abstraction and for simplicity in the model, total water abstraction was distributed among the individual wells during the calibration and validation processes. Groundwater abstractions were calculated from domestic, industrial, and agricultural water use. Total domestic and industrial pumping rates were based on estimates of population and per capita water consumption rates (Michael and Voss, 2009). Total domestic and industrial demand were taken as 50 l/day/capita (Faneca Sanchez et al., 2015; Michael and Voss, 2009) and assumed constant throughout the year. Total population for the study area were obtained by using data from the population census of the district statistics of Bangladesh (Bangladesh Bureau of Statistics (BBS), 2013). A population growth rate 2.09% per year (Michael and Voss, 2009) was used to calculate the number of population in subsequent years of simulation. Total domestic and industrial water abstraction was calculated from the following equation

$$Q_{domestic+industrial}(m^3/day) = Population\ size \times Annual\ growth\ rate \times Per\ capita\ water\ demand \quad (10.4)$$

The amount of water abstraction for irrigation purposes was computed from the total irrigated portion of the study area multiplied by the quantity of applied water to the irrigated area during the crop growing season (Michael and Voss, 2009). Total irrigated area in the study area was obtained from the district statistics for Barguna district (Bangladesh Bureau of Statistics (BBS), 2013). The total irrigated area was multiplied by an abstraction rate of 1 m/pumping season/$m^2$ of irrigated area (Harvey et al., 2006).

The type and thickness of aquifer material layers were chosen in accordance with the lithological data of the study area. It is noted that up to 300m depth of the study area falls under alluvium soil type composed mainly of clay, silt, sand, and occasional gravel (Faneca Sanchez et al., 2015). As most of the physical processes are occurred in the first few meters of the aquifer, an aquifer thickness of 150 m was chosen. The total thickness of the aquifer was divided into four layers of materials. First layer below the ground surface belongs to sandy silt with a thickness of 40 m, followed by a layer of sandy loam with 50 m thickness, followed by a soil type of sand with a thickness of 40 m. The bottom layer was specified as sandy clay with a thickness of 20 m. An average value of hydraulic conductivity was assigned to each model layer. The aquifer material within each model layer was assumed homogeneous, only vertical heterogeneity in terms of hydraulic conductivity was considered.

The hydraulic conductivity values used in this study were in accordance with previous studies conducted in the Bengal Delta (Faneca Sanchez et al., 2015; Michael and Voss, 2009). An anisotropy ratio ($kx/ky$) = 2.0 was used, where $kx$ represents the horizontal hydraulic conductivity in the $X$-direction. $ky$ is the horizontal hydraulic conductivity in the $Y$-direction. $kz$ is the vertical hydraulic conductivity in the $Z$-direction. The value of $kz$ was taken as one tenth of the hydraulic conductivity values in the $X$-direction. The 3-D view of the model domain with finite element meshes is illustrated in Figure 10.2.

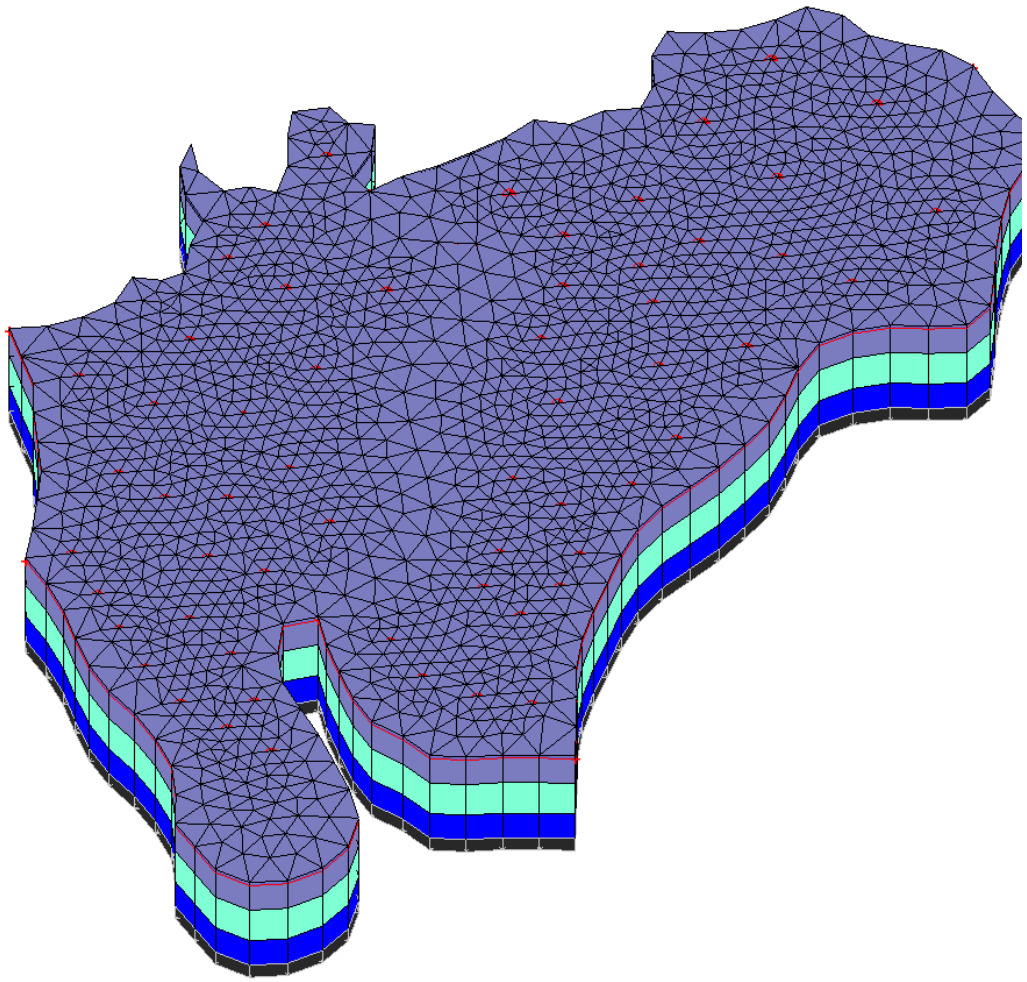

**Figure 10.2 Three dimensional view of the study area**

The study area is bounded by the Bay of Bengal in the southern side, Burishwar River in the eastern side, and Haringhata River in the western side. A river named Bishkhali is flowing in the middle part of the study area separating the two administrative upazillas. The Northern boundary is the administrative boundary, which was specified as the no flow boundary. Although it is very difficult to conclude that the northern boundary is a no flow boundary, the no flow boundary condition was assumed based on the consideration that the

study area has a negligible hydraulic gradient (around 1:20000) in the vicinity of this boundary. Therefore, lateral movement of groundwater across this boundary, as shown in Figure 10.1, can be considered as negligible (Faneca Sanchez et al., 2015). Therefore, for modelling purposes, this northern boundary as shown in Figure 10.1 was assumed as a no flow boundary. In addition, the northern boundary is relatively far away from the sea face boundary. Moreover, the calibration-validation process did not show that it was an unreasonable assumption. The seaside boundary was assigned as a constant head and constant concentration boundary. Constant head and constant concentration values in the seaside boundary were specified as zero (MSL) and 35000 mg/l, respectively. The upstream ends of the rivers were assigned specified head values that varied linearly along the stream and ended at zero meter at the seaside boundary. Specified heads of 0.8 m, 0.86 m, and 0.70 m were assigned at the upstream ends of Haringhata, Bishkhali, and Burishwar Rivers, respectively. The tidal river salinity concentrations were specified as 10000 mg/l and assumed to be constant throughout the simulation period. The fluid properties used in this study are presented in Table 10.1.

**Table 10.1 Fluid properties**

| Parameters | Units | Values |
|---|---|---|
| Density of freshwater | $Kg/m^3$ | 1000 |
| Density of seawater | $Kg/m^3$ | 1028 |
| Dynamic viscosity of water | Kg/m–day | 131.328 |
| Compressibility of water | $m–day^2/kg$ | $6.69796 \times 10^{-20}$ |
| Density reference ratio | dimensionless | 0.025 |

A plan view of the study area with boundaries and wells is shown in Figure 10.3. In Figure 10.3, the production wells, barrier wells, and monitoring locations are indicated by P1-P43, B1-B13, and M1-M16, respectively. For calibration and validation purposes, barrier well pumping was not considered and the hydraulic heads were observed at monitoring locations M1 and M2. Once proper calibration and validation were performed, barrier extraction wells were introduced as the hydraulic control measures of the saltwater intrusion processes. Moreover, an additional 14 monitoring locations were used to monitor salinity concentrations for developing the saltwater intrusion management model for the study area.

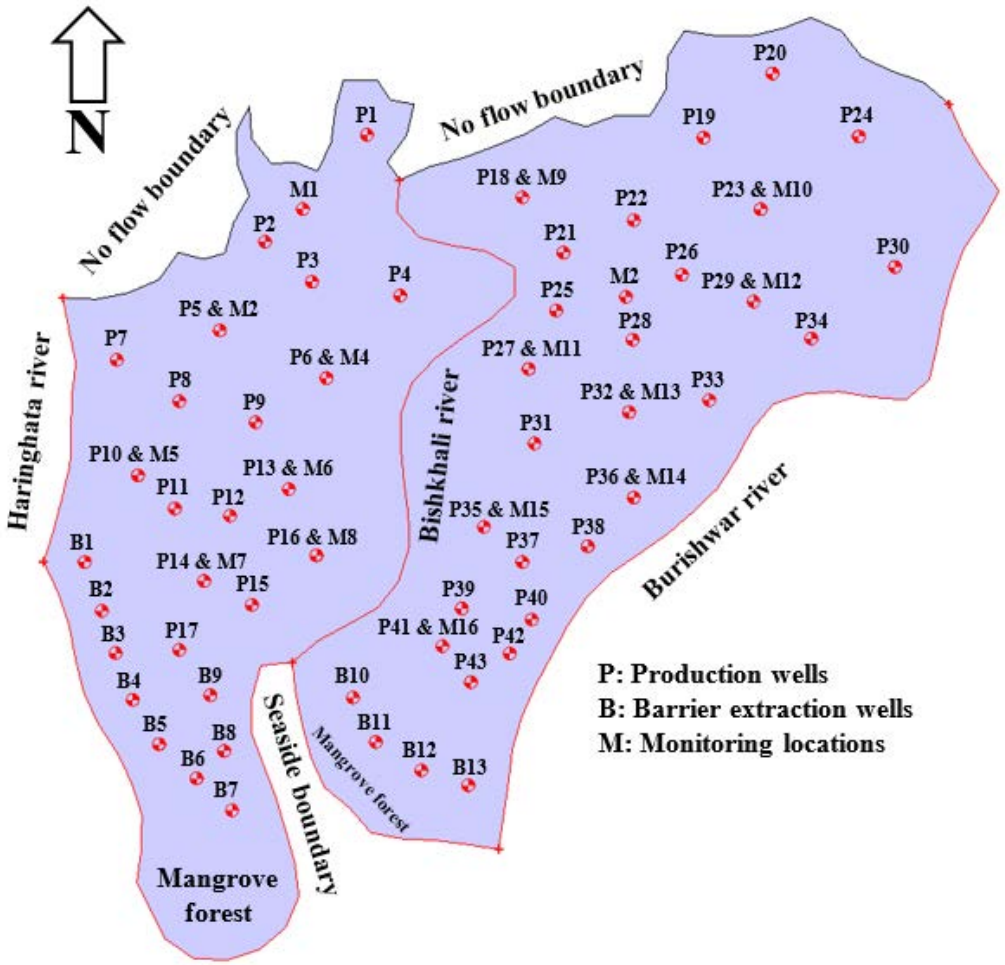


**Figure 10.3 Plan view of the study area showing the boundaries and wells**

10.4.2 Mesh dependency test

The accuracy of numerical simulation models often depends on the size of the finite elements. However, finer mesh size is associated with additional computational requirements. Therefore, modelling should be performed by maintaining a balance between accuracy and computational requirement. For this, the mesh dependency of the simulated hydraulic heads was determined by conducting numerical experiments, utilizing element sizes of 600m, 800m, 1000m, and 1200m. The simulation was performed for a period of 10 years, with a constant groundwater abstraction rate of 4000 $m^3$/day from each of the 43 production wells. Hydraulic heads and the time required for simulating the aquifer processes at the end of the simulation period were computed at two monitoring locations M1 and M2, and are presented in Table 10.2. Table 10.2 shows that the computed hydraulic heads did not vary significantly when different element sizes were used. However, there was a substantial increase in

simulation times with the increase in element size. Considering the computation time and accuracy of simulation, an element size of 1200 m was used in the present study.

**Table 10.2 Hydraulic heads and simulation times with different element sizes**

| Element size, m | Hydraulic heads, m | | Simulation time, min |
|---|---|---|---|
| | Patharghata (M1) | Barguna Sadar (M2) | |
| 1200 | 2.823 | 2.598 | 13.05 |
| 1000 | 2.825 | 2.597 | 18.57 |
| 800 | 2.826 | 2.598 | 29.58 |
| 600 | 2.825 | 2.598 | 47.95 |

10.4.3 Model calibration

The calibration process was initiated from a steady state condition of the hydraulic heads in the finite element nodes of the model domain. To achieve this condition, the transient simulation model was run for 80 years (April 1930 to April 2009). The simulation was performed in stages with an interval of 10 years. An average value of pumping was used during this simulation period. Outputs at the end of the $10^{th}$ year's simulation were used as initial conditions for the subsequent intervals of 10 years' period. The process was continued until April 2009, when a stable condition with respect to hydraulic head was achieved. These hydraulic head values at different nodes of the model domain were used as initial conditions for the calibration process. The calibration was performed from a period of five years (from April 2010 to April 2014). The hydraulic heads were monitored at the designated monitoring locations (M1 and M2) on April 2010, April 2011, April 2012, April 2013, and April 2014. Recharge and hydraulic conductivity values were adjusted to obtain the hydraulic heads closer to the actual hydraulic heads in the monitoring locations M1 and M2. Table 10.3 presents major parameter values used in the calibrated model.

**Table 10.3 Parameter values of the calibrated model**

| Parameters | Values | Units |
|---|---|---|
| Hydraulic conductivity in X-direction for soil layer 1 | 4 | m/day |
| Hydraulic conductivity in X-direction for soil layer 2 | 10 | m/day |
| Hydraulic conductivity in X-direction for soil layer 3 | 15 | m/day |
| Hydraulic conductivity in X-direction for soil layer 4 | 8 | m/day |
| Aquifer recharge applied on the top soil layer | 0.000689 | m/day |
| Longitudinal dispersivity | 80 | m |
| Lateral dispersivity | 30 | m |
| Molecular diffusion coefficient | 0.69 | $m^2$/day |

Actual and simulated hydraulic heads at two upazillas (M1 and M2 in Figure 10.3) during the calibration process are presented in Figure 10.4.

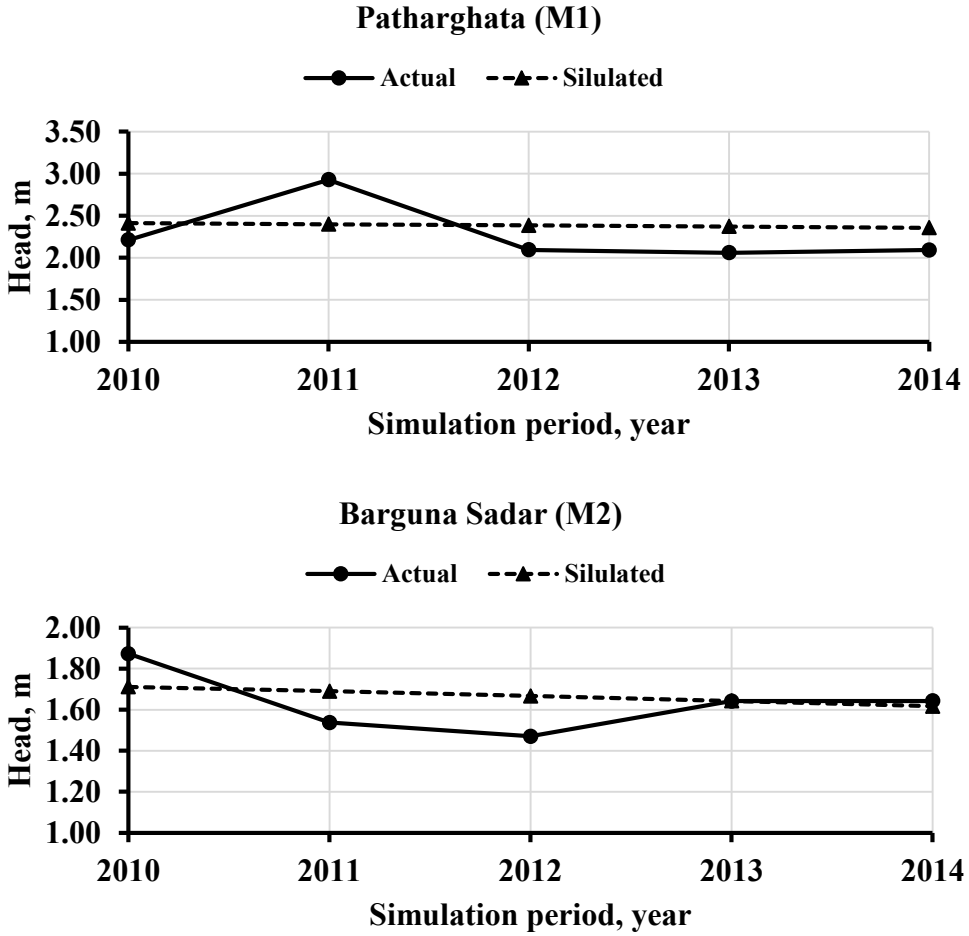


**Figure 10.4 Actual and simulated hydraulic heads at two upazillas during the calibration process**

The calibrated model was then validated for the next three years from April 2015 to April 2017. Outputs in terms of hydraulic heads on April 2014 were used as the initial condition for the simulation of the validation period. The model boundary conditions remained same as the calibrated model. At the end of the simulation period, the hydraulic heads were monitored at monitoring locations M1 and M2 in April 2015, April 2016, and in April 2017. Hydraulic heads during the validation process is presented in Figure 10.5.

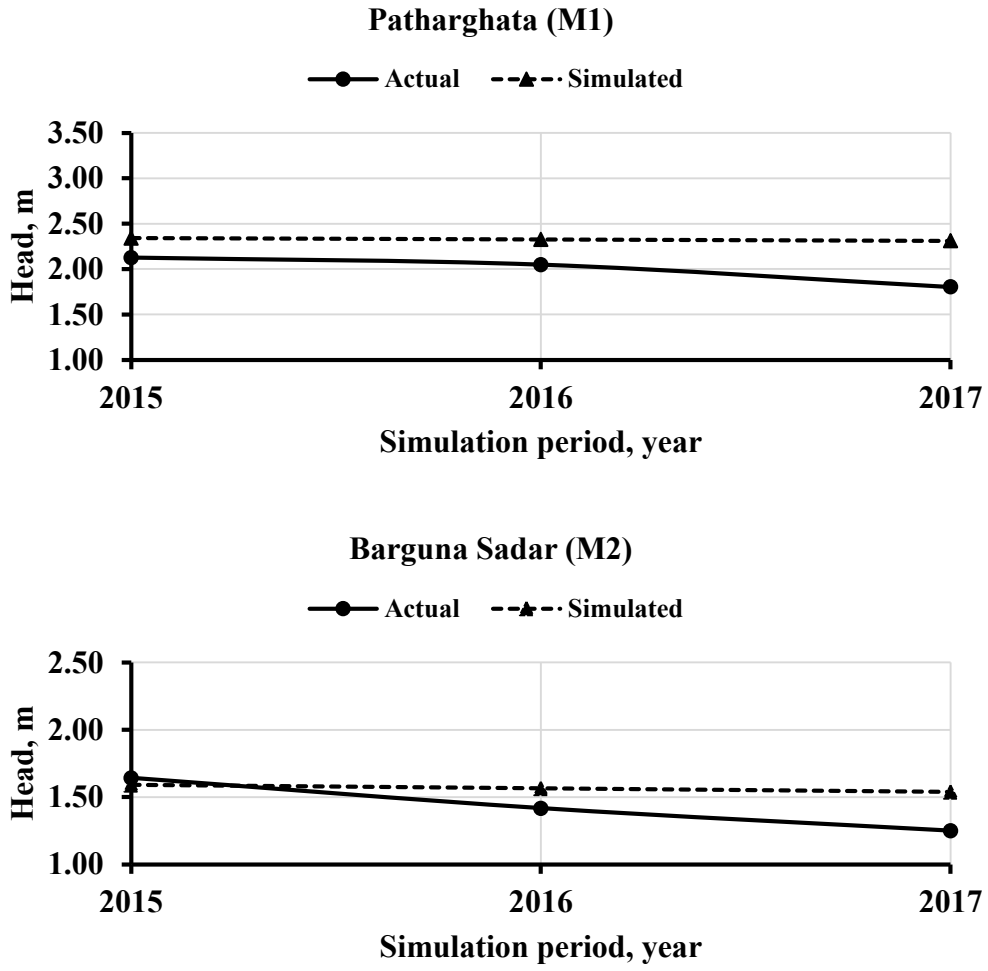


**Figure 10.5 Actual and simulated hydraulic heads at two upazillas during the validation period**

The calibration and validation processes were performed by using a uniform time step of 5 days' interval. A smaller time step is associated with higher computational time requirements and vice versa. Computational time is an issue in situations where multiple simulations of the aquifer processes with different sets of transient pumping values are required to train and test a meta-model for using in an integrated S-O approach. Therefore, a sensitivity analysis was performed to evaluate the effects of time steps of simulation on the computed hydraulic heads. Time steps of 1 day, 5 days, 10 days, and 73 days were used to

simulate the hydraulic heads during the calibration periods from April 2010 to April 2014. The results are presented in Table 10.4.

**Table 10.4 Sensitivity of simulation time steps to simulated hydraulic heads**

| Calibration period | Hydraulic heads, m | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Patharghata | | | | Barguna Sadar | | | |
| | 1 day | 5 days | 10 days | 73 days | 1 day | 5 days | 10 days | 73 days |
| 2010 | 2.41155 | 2.41105 | 2.41121 | 2.41094 | 1.71220 | 1.71135 | 1.71163 | 1.71120 |
| 2011 | 2.39968 | 2.39483 | 2.39844 | 2.39858 | 1.69179 | 1.68964 | 1.68968 | 1.68995 |
| 2012 | 2.38669 | 2.38463 | 2.38484 | 2.38503 | 1.66941 | 1.66588 | 1.66625 | 1.66661 |
| 2013 | 2.37317 | 2.37035 | 2.37030 | 2.37094 | 1.64606 | 1.64120 | 1.64112 | 1.64227 |
| 2014 | 2.35933 | 2.35583 | 2.35581 | 2.35645 | 1.62204 | 1.61599 | 1.61597 | 1.61712 |

Table 10.4 shows that the estimates of hydraulic heads did not differ substantially among different time steps during the calibration periods. However, a considerable amount of computational efficiency was achieved when a simulation time step of 73 days was used. Time taken to simulate the aquifer processes for the 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. Therefore, simulation time steps of 73 days were used in the multiple simulations with different transient pumping values. These multiple simulations were used to generate input-output training patterns for training of the meta-model.

## 10.5 Generation of input-output patterns for training of MARS meta-models

The physical processes of the aquifer with the transient groundwater abstraction patterns were simulated for the specified management period in order to generate input-output training patterns for MARS based meta-models. The spatial and temporal pumping stress applied to the aquifer was associated with water abstraction from a set of production, and barrier wells at specific locations and time steps. The transient groundwater abstraction values were obtained through Halton sequences (HA) (Halton, 1960), with specified lower and upper bounds. The HA are based on a deterministic algorithm, which uses prime numbers as bases for each dimension. This sampling technique was used in the current study because of its superiority over commonly used sampling techniques (Loyola R et al., 2016). The samples obtained by using the HA approach were found to be more uniform compared to Latin Hypercube Sampling (LHS) (Pebesma and Heuvelink, 1999). A comparison of the sampling sequences by utilizing both LHS and HA is provided in Figure 10.6.

The transient abstraction values were fed into the simulation model as inputs in order to obtain saltwater concentration values as outputs at specified monitoring locations. One set of input-output patterns was obtained from the transient groundwater abstraction values and the corresponding salinity concentrations. The salinity concentrations were measured at 16 monitoring locations. A number of such patterns were obtained by simulating the aquifer processes multiple times with different sets of transient groundwater abstraction values from a combination of production and barrier wells. The aquifer properties as well as the initial and boundary conditions remained constant for different simulations. Only the transient groundwater abstraction values varied in subsequent simulations to acquire different realizations of resulting saltwater concentration values obtained solely due to the pumping stress applied to the aquifer.

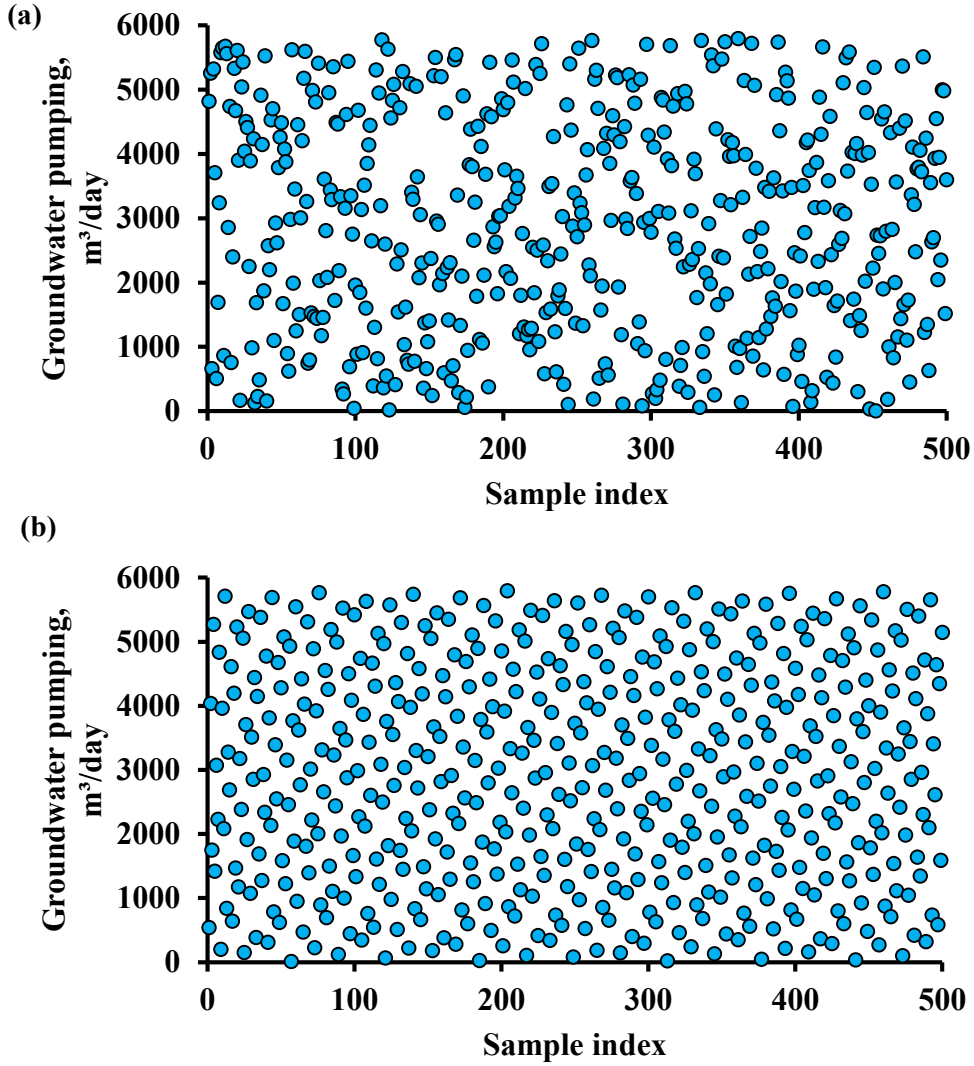


**Figure 10.6 Sampling sequences generated using (a) Latin hypercube sampling, (b) Halton sampling**

## 10.6 Performance of the MARS meta-models

The capability of the MARS meta-models in approximating coupled transient flow and solute transport processes in the coastal aquifer system was evaluated by using different statistical indices. After proper training and validation, the meta-models were presented with an unseen test dataset. For performance evaluation purposes, the R, IOA, KGE, RMSE, and MAPRE values were calculated for the developed meta-models on this new test dataset. The results are presented in Table 10.5. Generally, MARS meta-models produced higher values of R, IOA, KGE and lower values of MAPRE and RMSE at all 16 monitoring locations. These results demonstrate the ability of MARS meta-models for capturing the input-output patterns of the transient groundwater abstractions and the corresponding saltwater concentrations at specified monitoring locations.

**Table 10.5 Performance of the MARS meta-models on an unseen test dataset**

| Monitoring locations | Performance indices | | | | |
|---|---|---|---|---|---|
| | R | IOA | KGE | RMSE, mg/L | MAPRE, % |
| M1 | 0.99 | 0.99 | 0.99 | 0.005 | 0.0002 |
| M2 | 0.99 | 0.99 | 0.99 | 0.001 | 0.00004 |
| M3 | 0.99 | 0.99 | 0.99 | 0.211 | 0.006 |
| M4 | 0.99 | 0.99 | 0.99 | 0.187 | 0.005 |
| M5 | 0.99 | 0.99 | 0.99 | 0.393 | 0.006 |
| M6 | 0.99 | 0.99 | 0.99 | 0.270 | 0.006 |
| M7 | 0.99 | 0.99 | 0.99 | 0.234 | 0.005 |
| M8 | 0.99 | 0.99 | 0.99 | 0.232 | 0.005 |
| M9 | 0.99 | 0.99 | 0.99 | 0.150 | 0.004 |
| M10 | 0.99 | 0.99 | 0.99 | 0.146 | 0.005 |
| M11 | 0.99 | 0.99 | 0.99 | 0.168 | 0.005 |
| M12 | 0.99 | 0.99 | 0.99 | 0.146 | 0.005 |
| M13 | 0.99 | 0.99 | 0.99 | 0.157 | 0.004 |
| M14 | 0.99 | 0.99 | 0.99 | 0.189 | 0.005 |
| M15 | 0.99 | 0.99 | 0.99 | 0.171 | 0.004 |
| M16 | 0.99 | 0.99 | 0.99 | 0.216 | 0.004 |

As RMSE incorporates both variances and biases of the prediction error, the RMSE criterion was used to evaluate how well MARS meta-models fit the unseen test dataset.

Overall, MARS meta-models provide relatively lower RMSE values. However, one of the major drawbacks of RMSE is its tendency to give more weights to the outliers. Therefore, to obtain better information on the prediction capability of MARS meta-models by observing the distribution of errors, the MAPRE criteria were used as another performance measure. MARS meta-models at all monitoring locations provide very small values of MAPRE, which is acceptable in terms of meta-model based prediction accuracies. MARS meta-models' prediction accuracy was also verified from the R, IOA, and KGE viewpoint. All meta-models produced higher values of R, IOA, and KGE indicating the acceptable and reliable prediction accuracies of the developed meta-models.

## 10.7 Management of seawater intrusion

The calibrated model was used to develop a saltwater intrusion management model for a period of three years from April 2015 to April 2017. The proposed saltwater intrusion management model considered 168 decision variables of spatial and temporal groundwater abstraction patterns. Variables *X1–X129,* and *X130–X168* represent water abstraction from production and barrier wells, respectively. More specifically, for example, variables *X1–X3* denotes water abstraction from production well P1 in the first, second, and third time steps, respectively. Likewise, variables *X4–X6* denote water abstraction from the production well P2 and so on at three specified time steps. Variables *X130–X132, X133–X135, X136–X138…X166–X168* denote water abstraction from barrier wells *B1, B2, B3…B13*, at the first, second, and third time steps.

MARS based meta-models replaced the computationally expensive numerical simulation model within the integrated S-O framework for providing Pareto optimal groundwater abstraction strategies. Sixteen MARS meta-models predicting salinity concentrations at 16 designated monitoring locations were externally linked to CEMOGA in order to predict salinity concentrations. The MARS meta-models were also used to check constraint violations in terms of maximum allowable salinity concentrations at the specified monitoring locations. The CEMOGA used a population size of 2000, crossover fraction of 0.92, and Pareto front population fraction of 0.70. The function and constraint tolerances were set as $1\times10^{-5}$ and $1\times10^{-4}$, respectively. Although a detailed sensitivity analysis was not performed, several trials were conducted to select the optimal set of these parameters. The optimization routine evaluated 20528001 functions in 10263 generations to arrive at the global optimal solution. The optimization routine took 51 minutes to converge to optimal solutions. The Pareto optimal front shown in Figure 10.7 provided a set of different feasible solutions that show a trade-off between the two contradictory objectives of the saltwater intrusion management problem. The Pareto optimal front in Figure 10.7 demonstrates that

groundwater abstraction from production wells can be increased with increasing the water abstraction from the barrier wells.

## 10.8 Validation of the seawater intrusion management model

The validity of the proposed saltwater intrusion management model was assessed by observing the actual violation of the constraints. It is noted that the saltwater concentrations obtained from the optimization model solution (determined by MARS meta-models within the optimization framework) were smaller than the pre-specified maximum allowable saltwater concentrations at all monitoring locations. This implies that the imposed constraints were satisfied, and no constraint violation occurred during the search process. Moreover, obtained saltwater concentrations were very close to the prescribed values, which indicate that the optimization model converged to the upper limit of the imposed constraints.

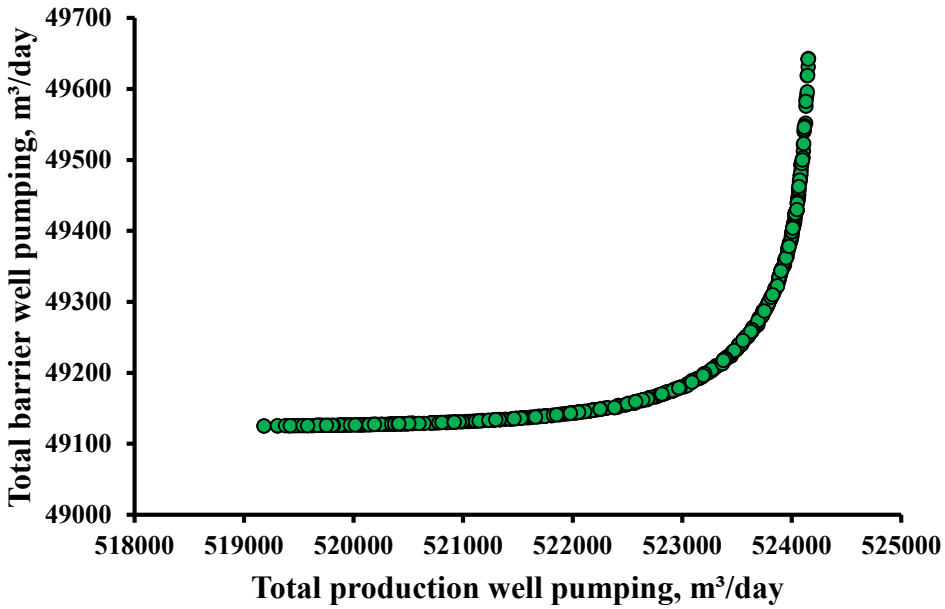


**Figure 10.7 Pareto optimal front of the management model**

In the second stage, the optimal groundwater abstraction strategies obtained from the optimization model were verified by comparing them with the numerical simulation results. To do this, 10 solutions were selected randomly from different regions of the Pareto optimal front. Table 10.6 shows the comparison of the MARS predicted and numerical model simulated saltwater concentration values obtained using optimal groundwater abstraction strategies prescribed by the proposed saltwater intrusion management model. Table 10.6 presents solution results from 5 monitoring locations. The results at other monitoring

locations followed a similar trend. It is observed from Table 10.6 that the solutions obtained from the numerical simulation model were very close to the MARS predictions.

The selection of the best solution from the Pareto optimal front was also proposed by applying decision theory. Gray Relational Analysis (GRA) was used as a decision theory to choose the best optimal solution. GRA is based on Gray system theory proposed by Deng (1982). Gray Relational Coefficient (GRC) as implemented in Wang and Rangaiah (2017) was computed to obtain the best optimal solution. The GRC approach was utilized to find the resemblance between the objective values of each candidate optimal solution and the ideal or best reference objective values.

**Table 10.6 Salinity concentrations calculated from optimal groundwater extraction strategy**

| Obs. | Saltwater concentration, mg/l | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | M1 | | M2 | | M3 | | M4 | | M5 | |
| | MARS | SM | MARS | SM | MARS | SM | MARS | SM | MARS | SM |
| 1 | 2584.76 | 2584.77 | 2346.81 | 2346.83 | 3033.77 | 3033.75 | 3234.85 | 3234.77 | 4995.95 | 4996.34 |
| 2 | 2584.76 | 2584.77 | 2346.82 | 2346.83 | 3033.95 | 3033.94 | 3235.04 | 3234.96 | 4995.96 | 4996.34 |
| 3 | 2584.76 | 2584.77 | 2346.82 | 2346.83 | 3033.94 | 3033.93 | 3235.02 | 3234.94 | 4995.96 | 4996.34 |
| 4 | 2584.76 | 2584.77 | 2346.82 | 2346.83 | 3033.94 | 3033.93 | 3235.02 | 3234.94 | 4995.96 | 4996.34 |
| 5 | 2584.76 | 2584.77 | 2346.82 | 2346.83 | 3034.02 | 3034.01 | 3235.05 | 3234.97 | 4995.96 | 4996.35 |
| 6 | 2584.76 | 2584.77 | 2346.81 | 2346.83 | 3033.83 | 3033.81 | 3234.91 | 3234.83 | 4995.95 | 4996.34 |
| 7 | 2584.76 | 2584.77 | 2346.82 | 2346.83 | 3033.95 | 3033.94 | 3235.07 | 3234.99 | 4995.96 | 4996.34 |
| 8 | 2584.76 | 2584.77 | 2346.82 | 2346.83 | 3034.01 | 3034.00 | 3235.06 | 3234.98 | 4995.96 | 4996.34 |
| 9 | 2584.76 | 2584.77 | 2346.81 | 2346.83 | 3033.75 | 3033.74 | 3234.83 | 3234.75 | 4995.94 | 4996.33 |
| 10 | 2584.76 | 2584.77 | 2346.82 | 2346.83 | 3033.94 | 3033.93 | 3235.02 | 3234.94 | 4995.96 | 4996.34 |

*M = Monitoring locations, SM = Simulation model

The GRC calculation is based on the following steps.

Step 1: Standardization of objective values to eliminate the effect of dimensionality. For objective 1 (maximization of total production well abstraction), the standardization was performed as

$$O_{ij} = \frac{O_{ij} - \min_{i \epsilon m} O_{ij}}{\max_{i \epsilon m} O_{ij} - \min_{i \epsilon m} O_{ij}} \tag{10.5}$$

The standardization of objective 2 (minimization of total barrier well abstraction) is expressed as

$$O_{ij} = \frac{\max_{i\epsilon m} O_{ij} - O_{ij}}{\max_{i\epsilon m} O_{ij} - \min_{i\epsilon m} O_{ij}} \tag{10.6}$$

Step 2: Searching for the ideal or best reference objective values

$$O_j^{Ideal} = \max_{i\epsilon m} O_{ij} \tag{10.7}$$

Step 3: Obtaining the deviation between the $O_j^{Ideal}$ and $O_{ij}$

$$\Delta D_{ij} = \left| O_j^{Ideal} - O_{ij} \right| \tag{10.8}$$

Step 4: Computation of GRC for each optimal solution

$$GRC_i = \frac{1}{m} \sum_{j=1}^{n} \frac{\Delta\min + \Delta\max}{\Delta D_{ij} + \Delta\max} \tag{10.9}$$

where, $i$ = index of the number of optimal solutions ($i = 1,2, \dots, m$), $j$ = the index of the number of objectives ($j = 1,2, \dots, n$), $\Delta\max = \max_{i\epsilon m, j\epsilon n}(\Delta D_{ij})$, $\Delta\min = \min_{i\epsilon m, j\epsilon n}(\Delta D_{ij})$.

Based on the concept of GRA, the larger the value of $GRC_i$ the more reliable the optimal solution is. Therefore, the largest value of $GRC_i$ is the recommended best optimal solution from the Pareto optimal solution.

The sustainability of the groundwater withdrawal as proposed in the management model was justified by observing the hydraulic heads at the monitoring locations and by calculating the amount of water entering and leaving the aquifer. Sustainability in terms of groundwater exploitation was evaluated by comparing the depth of water withdrawn from the proposed optimal pumping strategy with the existing pumping rates. Four solutions from the Pareto front were selected for this purpose. Three solutions were taken from the higher, average, and the lower ranges of total production well pumping. The fourth one was the optimal solution proposed as the best optimal solution based on the GRA method. For the optimal solutions, the total depth of water was calculated by considering pumping from both the production and barrier wells. The results are presented in Figure 10.8.

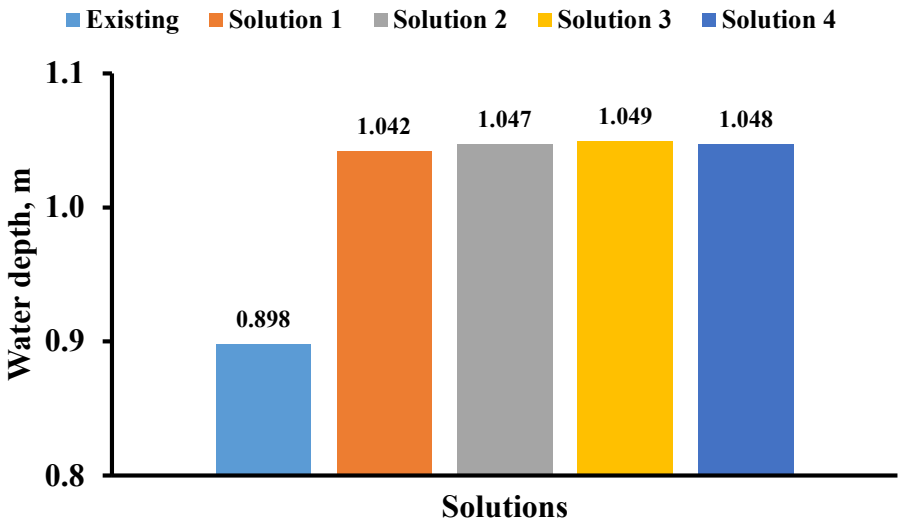


**Figure 10.8 Groundwater abstraction in terms of depth of water (m)**

Vertical recharge was assumed to be uniformly distributed over the top layer of the study area. The calibrated recharge of 0.000689 m/day amounted to 150430782 $m^3$/year. This amount of vertical recharge was 84.03% of the total average yearly pumping volume (179032829 $m^3$/year). The deficit amount of water comes into the system through the study area including streams and ocean boundaries, as the groundwater level fluctuates by only a small amount. The groundwater abstraction values obtained as solution of the management model were relatively higher than the existing withdrawals. This extra amount of water abstraction was possible while limiting the salinity to permissible levels due to the barrier wells now included. Although higher than the existing practice, the total optimal abstraction patterns are also safe and sustainable in terms of groundwater exploitation from the aquifer. For the selected optimal solutions, around 72% of the total abstracted groundwater comes from the vertical recharge. More specifically, vertical recharge accounts for 72.41%, 72.06%, 71.91%, and 72.03% of the total groundwater abstraction obtained from the optimal solutions 1, 2, 3, and 4 respectively. The rest of the water enters into the system through the model boundaries. It is expected that the change in pumping patterns should also affect the recharge coming into the aquifer, and therefore should be sustainable. The optimal pumping is again of the same order as the future recharge.

During the validation period (April 2015 to April 2017), the average drop in hydraulic head at the two monitoring locations was 0.062 m (observed by simulating the aquifer processes using the numerical code FEMWATER). Groundwater abstraction caused a 0.898 m of water loss from the system. The drop in hydraulic heads resulting from the proposed optimal groundwater abstraction rates were also estimated from the solution results of the numerical code FEMWATER. The average drop in hydraulic heads obtained by implementing the proposed optimal groundwater abstraction strategies were 0.062 m, 0.062

m, 0.069 m, and 0.067 m for the four optimal solutions, respectively. Water entering the model domain from vertical recharge was 0.754 m during the three years of the management period. Water deficit due to groundwater abstraction during the three years of management period was balanced by the vertical recharge to the aquifer and water entering the system from the model boundary (3 rivers and the seaside) aquifer.

The sustainability in terms of salinity concentration was ascertained by permitting groundwater abstraction within the limits of maximum permissible salinity concentrations at specified monitoring locations. Thus the proposed optimal groundwater abstraction patterns obtained as solutions of the management model was based on satisfying the constraints of maximum permissible salinity concentration at the specified monitoring locations. Sensitivity of the imposed constraints was ascertained by relaxing the constraints and observing the corresponding production and barrier well abstractions. The Pareto optimal front for the new sets of constraints in terms of salinity concentrations is presented in Figure 10.9.

It is observed from Figure 10.9 that the new relaxed constraints produced almost the same amount of production well abstraction for a decreased amount of barrier well abstraction. This implies that less amount of barrier well abstraction is required to obtain the same amount of production well abstraction when the maximum permissible salinity concentration is relaxed. In other words, production well abstraction can be increased with the same level of barrier well abstraction. Therefore, based on the guideline of maximum permissible salinity concentrations in the area of interest, the total amount of production and barrier well abstraction can be adjusted.

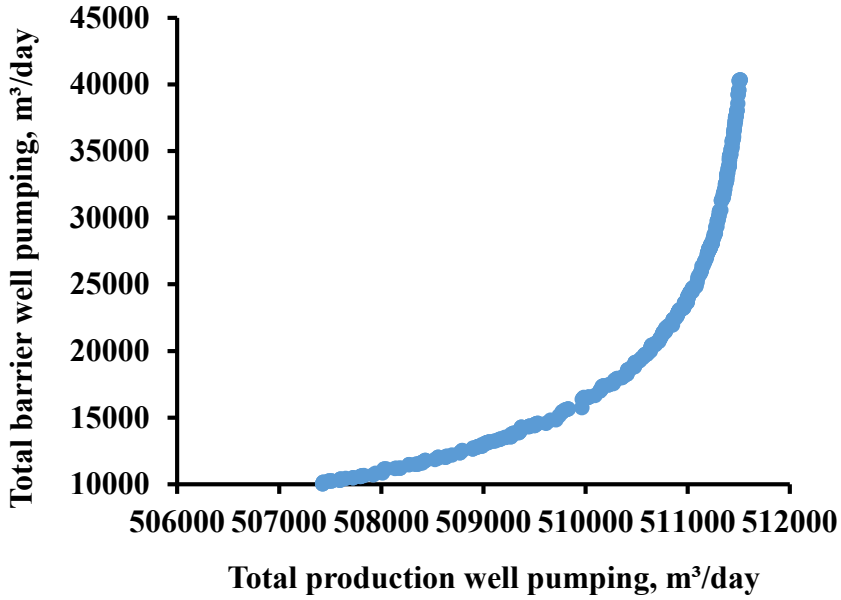


**Figure 10.9 Pareto front with relaxed constraint of maximum permissible salinity concentration**

It is noted that by implementing barrier well pumping, the total amount of beneficial pumping from the production wells can be enhanced, while maintaining the maximum permissible salinity concentrations within the permissible limits. Barrier wells aided in an additional groundwater abstraction rate of approximately 30000 $m^3$/day within the safe limit of maximum permissible salinity concentrations. For instance, for solution 1, 6.03% more water abstraction than the existing withdrawal could be achieved by implementing barrier wells. Approximately 0.65 $m^3$/day production well abstraction can be achieved from each $m^3$/day of barrier well abstraction. The values for individual optimal solutions are presented in Table 10.7. It is observed from Table 10.7 that the best optimal solution obtained by GRA provided better results than the two randomly selected solutions. However, it produced a slightly worse result than the other randomly selected optimal solution. The proposed GRA might serve as a quick decision making tool to select a relatively better solution from a large number of non-dominated solutions. However, managers can choose the optimal combination of production-barrier well pumping depending on the total demand of water for beneficial purposes.

Barrier wells are located near the seaside boundary of the study area. The seaside boundary is covered by mangrove forest. Therefore, the abstracted water from the barrier well can be efficiently disposed of into the ocean via the mangrove forest. Therefore, the proposed MARS-CEMOGA based saltwater intrusion management methodology is capable

of obtaining accurate solutions for optimal groundwater abstractions from a set of production bores and barrier wells in a real world coastal aquifer system.

**Table 10.7 Total amount of water withdrawal in different solutions**

| | PW pumping, $m^3$/day | BW pumping, $m^3$/day | Additional water abstraction from PW, $m^3$/day | Additional water abstraction from PW per unit water abstraction from BW |
|---|---|---|---|---|
| Existing | 490500.90 | - | - | - |
| Solution 1 | 520056.38 | 49126.85 | 29555.48 | 0.602 |
| Solution 2 | 522795.23 | 49168.69 | 32294.33 | 0.657 |
| Solution 3 | 523869.32 | 49322.38 | 33368.42 | 0.677 |
| Solution 4 | 523046.40 | 49181.84 | 32545.50 | 0.662 |

## 10.9 Conclusions

The optimal groundwater abstraction strategy using integrated an S-O approach has been considered an effective measure of controlling saltwater intrusion in coastal aquifers. Although the meta-model based integrated S-O approach has been widely applied in illustrative hypothetical example problems, only a few have focused on saltwater intrusion management in real world applications. This study demonstrates the applicability of the meta-model based integrated S-O approach in solving large-scale real world coastal aquifer management problems. A finite element based 3-D coupled flow and solute transport numerical code, FEMWATER, was utilized to simulate the saltwater intrusion processes in a coastal aquifer system in the Barguna district of southern Bangladesh. Input data for the selected study area of about 598 $km^2$ were collected from different sources. Scarcity and reliability of available data is a challenging issue in implementing regional scale saltwater intrusion models in this location. Therefore, the best possible subjective judgement was used in choosing the data for simulating the aquifer processes. The simulation model was calibrated with respect to hydraulic heads for a period of five years, from April 2010 to April 2014. With the selected hydrogeological parameters obtained from model calibration, the calibrated model was validated for a period of next three years, from April 2015 to April 2017. The calibrated and validated model was then used to develop a saltwater intrusion management model for the study area for a management period of three years, from April 2015 to April 2017. Computational efficiency in the integrated S-O approach was achieved by replacing the complex numerical simulation model with the properly trained MARS meta-

models. The limited evaluation results show that, using a carefully planned groundwater abstraction strategy, it is possible to modify saltwater intrusion processes and help in controlling spatial and temporal saltwater intrusion processes in a real world coastal aquifer study area.

This study utilized a planned pumping strategy from a set of production and barrier abstraction wells. Barrier wells were used as hydraulic control measures to control spatial and temporal distribution of saltwater concentrations. The limited evaluation results revealed that planned pumping strategy along with planned hydraulic control measures can be implemented to develop a saltwater intrusion management model in a realistic regional scale study area. Very limited data were available to calibrate and validate the numerical simulation model, and therefore the evaluation results are limited in scope. Additional reliable data will be required to develop a more acceptable calibrated model for the study area. Only then will the calibrated model be able to be used to prescribe actual implementation of the planned pumping strategy.

# Chapter 11: Summary and conclusions

## 11.1 Summary

Meta-model based coupled simulation-optimization (S-O) approaches were utilized in developing saltwater intrusion management models in order to obtain optimal groundwater abstractions from coastal aquifers. In order to develop coastal aquifer management models, a number of meta-models based on different algorithms were developed for integration within a coupled S-O approach. The coupled S-O approach was performed in a multiple objective problem setting using a Controlled Elitist Multiple Objective Genetic Algorithm. The meta-models developed were found to be more accurate and efficient than the existing meta-models in saltwater intrusion management models in coastal aquifers. Prediction uncertainty in meta-modelling approaches in the coupled S-O based methodology was addressed by proposing an ensemble of meta-models instead of a standalone meta-model. Ensembles were formed by varying the training dataset within the decision space, using random sampling without replacement approach. The ensemble meta-model based coupled S-O methodology was extended to incorporate the uncertainties in groundwater parameters. Uncertainties in hydraulic conductivity, bulk density, compressibility of the aquifer material, and aquifer recharge were considered in order to develop saltwater intrusion management models. Layered heterogeneity in hydraulic conductivity and porosity were considered for the illustrative multiple layered coastal aquifer system. Hydraulic conductivity and porosity were assumed to be uniform in each individual aquifer material layer, while the material layers were characterized with vertical heterogeneity. Climate change induced relative sea level rise was also considered in developing a multiple objective management model for a multiple layered coastal aquifer system. To address realistic scenarios, seasonal variation in the river water stage and salinity concentrations of the surface water in the tidal river were incorporated in the simulation of the aquifer processes.

A 3-D compliance monitoring network design was proposed in order to develop an adaptive management strategy, by sequentially modifying the groundwater abstraction patterns based on the field compliance of the implemented optimal abstraction strategies. A weighted average approach of ensemble meta-modelling using Set Pair Analysis was proposed in developing the adaptive management strategy. Entropy was proposed as a new measure of uncertainty in the saltwater intrusion management problems in the design of the uncertainty based 3-D compliance monitoring network design. A meta-model based coupled S-O approach was applied to develop a multiple objective saltwater intrusion management model for a realistic coastal aquifer system in the Barguna district of southern Bangladesh. The management model provides optimal groundwater abstraction patterns in the form of a

Pareto optimal front, which shows a trade-off between total beneficial water abstraction from the production wells and the abstraction for hydraulic control from the barrier wells. The selection of the best solution from the Pareto optimal front was proposed by applying a decision theory based on Gray Relational Analysis.

## 11.2 Conclusions

Sustainable beneficial water abstraction from coastal aquifers can be ensured by controlling saltwater intrusion through reversing the hydraulic gradient along the coast by using a set of barrier extraction wells. Another plausible way of controlling saltwater intrusion is to create a freshwater lens near the shoreline by utilizing a set of freshwater recharge wells. A comparative evaluation of these two types of hydraulic control measures for an illustrative 3-D multilayered coastal aquifer system demonstrated the superiority of barrier extraction wells over recharge wells in a shorter management time horizon for controlling saltwater intrusion of the illustrative coastal aquifer system.

FIS and MARS based meta-models were proposed that could be suitable for incorporation in a coupled S-O technique to develop optimum pumping strategy. Both FIS and MARS meta-models have unique advantages, for instance FISs are suitable for multiple output problems and MARS models are adaptive in nature and have very high computational efficiency. No doubt, further evaluations using field data from coastal aquifers are necessary to establish the applicability of such meta-models within the linked S-O methodology. Saltwater intrusion management models were developed by externally linking these meta-models to a Controlled Elitist Multiple Objective Genetic Algorithm (CEMOGA) within a computationally feasible coupled S-O approach in order to evolve optimal groundwater extraction strategies for a coastal aquifer system. Two conflicting objectives of groundwater extraction strategy were considered in this study. The first objective ensures the maximum withdrawal of groundwater for beneficial purposes. The second objective minimizes the water extraction from barrier pumping wells to control saltwater intrusion by establishing a hydraulic head barrier near the coastal boundary. The multiple objective management model provides a trade-off between these conflicting objectives in terms of a Pareto optimal front, which consists of several non-dominated feasible alternative groundwater extraction strategies that meet the pre-specified allowable saltwater concentration limits at specified monitoring locations. The results from the management models indicated that both FIS and MARS based meta-models provide acceptable, accurate, and reliable groundwater extraction patterns to limit the saltwater concentrations within the pre-specified maximum allowable limits. Therefore, the developed methodology utilizing the FIS and MARS based meta-models is potentially applicable for developing optimal groundwater extraction strategies for

sustainable regional scale management of coastal aquifers. Although the results presented here are limited in scope, these evaluation results may provide more generalized guidance for optimal extraction of groundwater resources from coastal aquifers.

The main advantage of meta-model based linked S-O models for regional scale management of coastal aquifers is attaining computational feasibility and efficiency. Computational efficiency can be significantly improved by implementing parallel computation procedures, in which objective function (s) and all other constraints of the optimization problem are evaluated in parallel by distributing them into physical cores of a PC. In this study, four physical cores of a standard seven core PC have been utilized to evaluate the objective functions and constraints of the multiple objective saltwater intrusion management problem. Integration of parallel processing capabilities within the optimization model was shown to improve computational efficiency and improve the feasibly of solving such large-scale multiple objective problems. It was also demonstrated that a considerable amount of reduction in the computational burden could be achieved by implementing a parallel computing platform for the solution of linked S-O methodology. The proposed methodology has the potential to improve the computational efficiency of aquifer management models, making it feasible to develop long-term sustainable optimal management strategies on a regional scale.

Uncertainties arising from estimating hydrogeological model parameters (hydraulic conductivity, compressibility, bulk density, and aquifer recharge), and in accurately estimating spatial and temporal variation of groundwater extraction patterns were addressed in developing the saltwater intrusion management models. In addition, the FIS based meta-models were advanced to a genetic algorithm (GA) tuned hybrid FIS model (GA-FIS) to emulate physical processes of coastal aquifers and to evaluate responses of the coastal aquifers to groundwater extraction under groundwater parameter uncertainty. The GA-FIS models thus obtained were linked externally to the CEMOGA in order to derive optimal pumping management strategies using a coupled S-O approach. The performance of the hybrid GA-FIS-CEMOGA based saltwater intrusion management model was compared with that of a basic Adaptive Neuro Fuzzy Inference System (ANFIS) based management model (ANFIS-CEMOGA). The performance evaluation results demonstrated the superiority of the GA-FIS-CEMOGA based management model over ANFIS-CEMOGA based saltwater intrusion management model. The proposed GA-FIS models did not include parameter uncertainty directly; however, they indirectly incorporated uncertainty because they were developed from the solution results of a numerical model that addressed parameter uncertainty.

An ensemble of MARS (En-MARS) based meta-models within a coupled S-O methodology was proposed to derive multiple objective optimal groundwater extraction

strategies in a multilayered coastal aquifer system. The prediction capability of En-MARS was compared with that of the best MARS model in the ensemble. Following the development of En-MARS meta-models, they were linked to an optimization algorithm within a regional scale saltwater intrusion management model. The results indicated that MARS based ensemble modelling approach is able to provide reliable solutions for a multilayered coastal aquifer management problem. The adaptive nature of MARS models, the use of ensembles, and the utilization of parallel processing resulted in a computationally efficient, accurate and reliable methodology for coastal aquifer management that also incorporated uncertainties in the modelling. The MARS-based ensemble meta-modelling technique using parallel processing platform has the advantages of (1) addressing the predictive uncertainty of meta-modelling, and (2) achieving computational efficiency in searching for optimal solutions: 3,990,401 function evaluations in the optimization routine took only 76 minutes. As the En-MARS learns from a diverse data range, it is expected to be more consistent for a newer test data set, thereby reducing predictive uncertainty.

The effects of relative sea level rise on the optimized groundwater extraction values for a management period of 5 years were incorporated. Variation of water concentrations of the tidal river and seasonal fluctuation of head at the upstream end of river were also incorporated. Aquifer processes were approximated by three meta-models, viz. ANFIS, GPR, and MARS. The meta-models were compared based on their prediction accuracy, reliability, and computational efficiency. The results demonstrated that both models are capable of capturing the trend of coupled flow and salt transport processes of a multilayered coastal aquifer system. However, based on a closer look at the prediction accuracy and computational efficiency, the ANFIS based meta-model was selected as a computationally cheap and feasible soft computing tool for incorporation into the linked S-O approach.

A 3-D sequential compliance monitoring network design methodology was proposed for developing an adaptive management strategy in a multilayered coastal aquifer system. This methodology addressed both uncertainties in prediction of concentrations due to uncertainties in parameter values, as well as the noncompliance of users with a prescribed management strategy. In the ensemble of MARS based meta-models, accounting for prediction uncertainty based on SPA theory was externally linked to the optimization algorithm to develop the initial management policy for a management period of five years. For the prescribed management strategy, a 3-D monitoring network was designed to monitor and modify the implemented saltwater intrusion management strategy. Monitoring locations were selected from three different depths of the aquifer at which the uncertainty in terms of salinity concentrations in these locations was highest. This is accomplished by proposing a new method of uncertainty quantification, i.e., Shannon's entropy, in the objective function of the optimal monitoring network design model for saltwater intrusion in coastal aquifers.

The results obtained from an illustrative multilayered coastal aquifer system demonstrated the potential applicability of the Entropy-SPA based 3-D monitoring network design for sequential and adaptive management of coastal aquifers. The effects of hydrogeological parameter estimation uncertainties were incorporated through the ensemble of meta-models used for prediction of the resulting salinity concentrations.

The application of a meta-model based integrated S-O approach in solving the large-scale real world coastal aquifer management problems was demonstrated. A finite element based 3-D coupled flow and solute transport numerical code, FEMWATER, was utilized to simulate the saltwater intrusion processes in a coastal aquifer system in the Barguna district of southern Bangladesh. Scarcity and reliability of available data was a challenging issue in implementing regional scale saltwater intrusion models. Therefore, the best possible subjective judgement was used in choosing the data for simulating the aquifer processes. The calibrated and validated model was then used to develop saltwater intrusion management model for the study area. Computational efficiency in the integrated S-O approach was achieved by replacing the complex numerical simulation model with the properly trained MARS meta-models. The limited evaluation results demonstrated that a carefully planned groundwater abstraction strategy in tandem with planned hydraulic control measures are able to modify the saltwater intrusion processes and help in controlling spatial and temporal saltwater intrusion processes in a real world coastal aquifer study area. This study shows the capability of the MARS meta-model based integrated S-O approach in solving real life complex management problems in an efficient manner.

## 11.3 Recommendations

The present study focuses on an illustrative coastal aquifer system in which aquifer material within each layer was assumed to be homogeneous. Future research may be directed towards considering aquifer heterogeneity by employing a random hydraulic conductivity field. In addition, the present study considered a parallel pool of worker machines within the local clusters of a PC for executing the optimization routines. For more complex optimization settings, a parallel pool consisting of more PCs or any other high performance computing platform could be used.

In addition, the proposed management model did not look into the various disposal alternatives for water pumped from the barrier wells. Incorporation of related economic and environmental considerations can be included in a more detailed management model. The proposed methodology potentially facilitates a computationally feasible and efficient determination of regional-scale management strategies for coastal aquifers.

Very limited data were available to calibrate and validate the numerical simulation model, and therefore, the evaluation results are limited in scope. Additional reliable data will be required to develop a more acceptable calibrated model for the study area. Only then can the calibrated model be used to prescribe actual implementation of the planned pumping strategy.

## References


Ababou, R., Al-Bitar, A., 2004. Salt water intrusion with heterogeneity and uncertainty: mathematical modeling and analyses. Devel. Water Sci., 55(2): 1559-1571. DOI:10.1016/S0167-5648(04)80166-7

Abd-Elhamid, H.F., Javadi, A.A., 2011. A cost-effective method to control seawater intrusion in coastal aquifers. Water Resour. Manag., 25(11): 2755-2780. DOI:10.1007/s11269-011-9837-7

Adamowski, J., Chan, H.F., Prasher, S.O., Sharda, V.N., 2012. Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in Himalayan micro-watersheds with limited data. J. Hydroinform., 14(3): 731-744. DOI:10.2166/hydro.2011.044

Allen, D.M., Schuurman, N., Zhang, Q., 2007. Using fuzzy logic for modeling aquifer architecture. J. Geogr. Syst., 9(3): 289-310. DOI:10.1007/s10109-007-0046-0

Altunkaynak, A., 2010. A predictive model for well loss using fuzzy logic approach. Hydrolog. Process., 24(17): 2400-2404. DOI:10.1002/hyp.7642

Araghi, S., Khosravi, A., Creighton, D. (2015). Design of an optimal ANFIS traffic signal controller by using cuckoo search for an isolated intersection. Paper presented at the 2015 IEEE International Conference on Systems, Man, and Cybernetics, 9-12 Oct 2015. DOI:10.1109/SMC.2015.363

Ataie-Ashtiani, B., Ketabchi, H., Rajabi, M.M., 2014. Optimal management of a freshwater lens in a small island using surrogate models and evolutionary algorithms. J. Hydrol. Eng., 19(2): 339-354. DOI:10.1061/(ASCE)HE.1943-5584.0000809

Ataie-Ashtiani, B., Werner, A.D., Simmons, C.T., Morgan, L.K., Lu, C., 2013. How important is the impact of land-surface inundation on seawater intrusion caused by sea-level rise? Hydrogeol. J., 21(7): 1673-1677. DOI:10.1007/s10040-013-1021-0

Ayvaz, M.T., Karahan, H., Aral, M.M., 2007. Aquifer parameter and zone structure estimation using kernel-based fuzzy c-means clustering and genetic algorithm. J. Hydrol., 343(3): 240-253. DOI:10.1016/j.jhydrol.2007.06.018

Baalousha, H., 2010. Assessment of a groundwater quality monitoring network using vulnerability mapping and geostatistics: A case study from Heretaunga Plains, New Zealand. Agr. Water Manage., 97(2): 240-246. DOI:10.1016/j.agwat.2009.09.013

Banerjee, P., Singh, V.S., Chatttopadhyay, K., Chandra, P.C., Singh, B., 2011. Artificial neural network model as a potential alternative for groundwater salinity forecasting. J. Hydrol., 398(3): 212-220. DOI:10.1016/j.jhydrol.2010.12.016

Bangladesh Bureau of Statistics (BBS). (2013). District statistics 2011: Barguna district Bangladesh Bureau of Statistics. Statistics and informatics division. Ministry of

planning. Government of the People's Republic of Bangladesh Retrieved from www.bbs.gov.bd

Bangladesh Water Development Board. Processing and Flood Forecasting Circle. (2018). Groundwater weekly data graphical view Vol. Accessed 12 March 2018. Retrieved from http://www.hydrology.bwdb.gov.bd/index.php?pagetitle=ground_water_weekly_data_view&sub2=215&_subid=132&id=140

Bashi-Azghadi, S.N., Kerachian, R., 2010. Locating monitoring wells in groundwater systems using embedded optimization and simulation models. Sci. Total Environ., 408(10): 2189-2198. DOI:10.1016/j.scitotenv.2010.02.004

Basser, H. et al., 2015. Hybrid ANFIS–PSO approach for predicting optimum parameters of a protective spur dike. Appl. Soft Comput., 30: 642-649. DOI:10.1016/j.asoc.2015.02.011

Bazi, Y., Alajlan, N., Melgani, F., 2012. Improved estimation of water chlorophyll concentration with semisupervised gaussian process regression. IEEE Geosci. Remote Sens. Lett., 50(7): 2733-2743. DOI:10.1109/TGRS.2011.2174246

Beebe, C., Ferguson, G., Kennedy, G. (2011). Analytical modeling of saltwater intrusion: tests from Nova Scotia and the eastern United States Saint Francis Xavier University and Nova Scotia Department of Natural Resources, Nova Scotia, Canada Retrieved from https://atlanticadaptation.ca/en/islandora/object/acasa%253A257

Behzadian, K., Kapelan, Z., Savic, D., Ardeshir, A., 2009. Stochastic sampling design using a multi-objective genetic algorithm and adaptive neural networks. Environ. Modell. Softw., 24(4): 530-541. DOI:10.1016/j.envsoft.2008.09.013

Bera, P. et al., 2006. Application of MARS in simulating pesticide concentrations in soil. T. ASABE, 49(1): 297-307. DOI:10.13031/2013.20228

Beuchat, X., Schaefli, B., Soutter, M., Mermoud, A., 2011. Toward a robust method for subdaily rainfall downscaling from daily data. Water Resour. Res., 47(9): W09524. DOI:10.1029/2010WR010342

Bezdek, J.C., Ehrlich, R., Full, W., 1984. FCM: The fuzzy c-means clustering algorithm. Comput. Geosci., 10(2): 191-203. DOI:10.1016/0098-3004(84)90020-7

Bhattacharjya, R.K., Datta, B., 2005. Optimal management of coastal aquifers using linked simulation optimization approach. Water Resour. Manag., 19(3): 295-320. DOI:10.1007/s11269-005-3180-9

Bhattacharjya, R.K., Datta, B., 2009. ANN-GA-based model for multiple objective management of coastal aquifers. J. Water Res. Plan. Man., 135(5): 314-322. DOI:10.1061/(ASCE)0733-9496(2009)135:5(314)

Bhattacharjya, R.K., Datta, B., Satish, M.G., 2007. Artificial neural networks approximation of density dependent saltwater intrusion process in coastal aquifers. J. Hydrol. Eng., 12(3): 273-282. DOI:10.1061/(ASCE)1084-0699(2007)12:3(273)

Bishop, C., 2006. Pattern recognition and machine learning, Springer-Verlag, New York.

Blanning, R.W., 1975. The construction and implementation of metamodels. Simulation, 24: 177-184.

Boschetti, T., González-Hernández, P., Hernández-Díaz, R., Naclerio, G., Celico, F., 2015. Seawater intrusion in the Guanahacabibes Peninsula (Pinar del Rio Province, western Cuba): effects on karst development and water isotope composition. Environ. Earth Sci., 73(9): 5703-5719. DOI:10.1007/s12665-014-3825-1

Bouzourra, H., Bouhlila, R., Elango, L., Slama, F., Ouslati, N., 2015. Characterization of mechanisms and processes of groundwater salinization in irrigated coastal area using statistics, GIS, and hydrogeochemical investigations. Environ. Sci. Pollut. R., 22(4): 2643-2660. DOI:10.1007/s11356-014-3428-0

Bower, J.W., Motz, L.H., Durden, D.W., 1999. Analytical solution for determining the critical condition of saltwater upconing in a leaky artesian aquifer. J. Hydrol., 221(1): 43-54. DOI:10.1016/S0022-1694(99)00078-5

Carneiro, J.F. et al., 2010. Evaluation of climate change effects in a coastal aquifer in Morocco using a density-dependent numerical model. Environ. Earth Sci., 61(2): 241-252. DOI:10.1007/s12665-009-0339-3

Chadalavada, S., Datta, B., 2008. Dynamic optimal monitoring network design for transient transport of pollutants in groundwater aquifers. Water Resour. Manag., 22(6): 651-670. DOI:10.1007/s11269-007-9184-x

Chang, S.W., Clement, T.P., 2012. Experimental and numerical investigation of saltwater intrusion dynamics in flux-controlled groundwater systems. Water Resour. Res., 48(9). DOI:10.1029/2012WR012134

Chang, S.W., Clement, T.P., Simpson, M.J., Lee, K.-K., 2011. Does sea-level rise have an impact on saltwater intrusion? Adv. Water Resour., 34(10): 1283-1291. DOI:10.1016/j.advwatres.2011.06.006

Chen, B.-F., Hsu, S.-M., 2004. Numerical study of tidal effects on seawater intrusion in confined and unconfined aquifers by time-independent finite-difference method. J. Waterw. Port. C., 130(4): 191-206. DOI:10.1061/(ASCE)0733-950X(2004)130:4(191)

Chen, C.-W., Wei, C.-C., Liu, H.-J., Hsu, N.-S., 2014. Application of neural networks and optimization model in conjunctive use of surface water and groundwater. Water Resour. Manag., 28(10): 2813-2832. DOI:10.1007/s11269-014-0639-6

Cheng, A.H.D., Halhal, D., Naji, A., Ouazar, D., 2000. Pumping optimization in saltwater-intruded coastal aquifers. Water Resour. Res., 36(8): 2155-2165. DOI:10.1029/2000WR900149

Cheng, J.-R., Strobl, R.O., Yeh, G.-T., Lin, H.-C., Choi, W.H., 1998. Modeling of 2D density-dependent flow and transport in the subsurface. J. Hydrol. Eng., 3(4): 248-257. DOI:10.1061/(ASCE)1084-0699(1998)3:4(248)

Christelis, V., Mantoglou, A., 2016. Pumping optimization of coastal aquifers assisted by adaptive metamodelling methods and radial basis functions. Water Resour. Manag., 30(15): 5845-5859. DOI:10.1007/s11269-016-1337-3

Colombani, N., Mastrocicco, M., Giambastiani, B.M.S., 2015. Predicting salinization trends in a lowland coastal aquifer: Comacchio (Italy). Water Resour. Manag., 29(2): 603-618. DOI:10.1007/s11269-014-0795-8

Comte, J.-C., Join, J.-L., Banton, O., Nicolini, E., 2014. Modelling the response of fresh groundwater to climate and vegetation changes in coral islands. Hydrogeol. J., 22(8): 1905-1920. DOI:10.1007/s10040-014-1160-y

Coulibaly, P., Baldwin, C.K., 2005. Nonstationary hydrological time series forecasting using nonlinear dynamic methods. J. Hydrol., 307(1): 164-174. DOI:10.1016/j.jhydrol.2004.10.008

Das, A., Datta, B., 1999a. Development of management models for sustainable use of coastal aquifers. J Irrig. Drain. E., 125(3): 112-121. DOI:10.1061/(ASCE)0733-9437(1999)125:3(112)

Das, A., Datta, B., 1999b. Development of multiobjective management models for coastal aquifers. J. Water Res. Plan. Man., 125(2): 76-87. DOI:10.1061/(ASCE)0733-9496(1999)125:2(76)

Das, A., Datta, B., 2000. Optimization based solution of density dependent seawater intrusion in coastal aquifers. J. Hydrol. Eng., 5(1): 82-89. DOI:10.1061/(ASCE)1084-0699(2000)5:1(82)

Das, A., Datta, B., 2001. Application of optimisation techniques in groundwater quantity and quality management. Sadhana, 26(4): 293-316. DOI:10.1007/bf02703402

Datta, B., Peralta, R.C., 1986. Interactive computer graphics-based multiobjective decision-making for regional groundwater management. Agr. Water Manage., 11(2): 91-116. DOI:http://dx.doi.org/10.1016/0378-3774(86)90023-5

Datta, B., Vennalakanti, H., Dhar, A., 2009. Modeling and control of saltwater intrusion in a coastal aquifer of Andhra Pradesh, India. J HYDRO-ENVIRON RES J. Hydro-environ. Res., 3(3): 148-159. DOI:10.1016/j.jher.2009.09.002

Dausman, A.M., Langevin, C.D., 2005. Movement of the saltwater interface in the surficial aquifer system in response to hydrologic stresses and water-management practices, Broward County, Florida (- ed.).

Deb, K., 1999. An introduction to genetic algorithms. Sadhana, 24(4): 293-315. DOI:10.1007/bf02823145

Deb, K., Agrawal, S., Pratap, A., Meyarivan, T., 2000. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In: Schoenauer, M. et al. (Eds.), Parallel problem solving from nature PPSN VI: 6th international conference Paris, France, September 18–20, 2000 proceedings. Springer, Berlin, pp. 849-858. DOI:10.1007/3-540-45356-3_83

Deb, K., Goel, T., 2001. Controlled elitist non-dominated sorting genetic algorithms for better convergence. In: Zitzler, E., Thiele, L., Deb, K., Coello Coello, C.A., Corne, D. (Eds.), Evolutionary multi-criterion optimization: first international conference, EMO 2001 Zurich, Switzerland, March 7–9, 2001 proceedings. Springer, Berlin, pp. 67-81. DOI:10.1007/3-540-44719-9_5

Deng, J.-L., 1982. Control problems of grey systems. Sys. Control. Lett., 1(5): 288-294. DOI:10.1016/S0167-6911(82)80025-X

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. DOI:10.1007/s10040-014-1214-1

Dhar, A., Datta, B., 2007. Multiobjective design of dynamic monitoring networks for detection of groundwater pollution. J. Water Res. Plan. Man., 133(4): 329-338. DOI:10.1061/(ASCE)0733-9496(2007)133:4(329)

Dhar, A., Datta, B., 2009a. Saltwater intrusion management of coastal aquifers. I: linked simulation-optimization. J. Hydrol. Eng., 14(12): 1263-1272. DOI:10.1061/(ASCE)HE.1943-5584.0000097

Dhar, A., Datta, B., 2009b. Saltwater intrusion management of coastal aquifers. II: Operation uncertainty and monitoring. J. Hydrol. Eng., 14(12): 1273-1282. DOI:10.1061/(ASCE)HE.1943-5584.0000155

Dhar, A., Datta, B., 2010. Logic-based design of groundwater monitoring network for redundancy reduction. J. Water Res. Plan. Man., 136(1): 88-94. DOI:10.1061/(ASCE)0733-9496(2010)136:1(88)

Dhar, A., Patil, R.S., 2011. Fuzzy uncertainty based design of groundwater quality monitoring networks. J. Environ. Res. Develop., 5(3A): 683-688.

Dokou, Z., Karatzas, G.P., 2012. Saltwater intrusion estimation in a karstified coastal system using density-dependent modelling and comparison with the sharp-interface approach. Hydrol. Sci. J., 57(5): 985-999. DOI:10.1080/02626667.2012.690070

Dougherty, D.E., 1991. Hydrologic applications of the connection machine CM- 2. Water Resour. Res., 27(12): 3137-3147. DOI:10.1029/91WR02026

Doulgeris, C., Zissis, T., 2014. 3D variable density flow simulation to evaluate pumping schemes in coastal aquifers. Water Resour. Manag., 28(14): 4943-4956. DOI:10.1007/s11269-014-0766-0

Emamgholizadeh, S., Moslemi, K., Karami, G., 2014. Prediction the groundwater level of Bastam Plain (Iran) by artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) Water Resour. Manag., 28(15): 5433-5446. DOI:10.1007/s11269-014-0810-0

Emch, P.G., Yeh, W.W.-G., 1998. Management model for conjunctive use of coastal surface water and ground water. J. Water Res. Plan. Man., 124(3): 129-139. DOI:10.1061/(ASCE)0733-9496(1998)124:3(129)

Faneca Sanchez, M. et al. (2015). SWIBANGLA: Managing saltwater intrusion impacts in coastal groundwater systems of Bangladesh Final Report. Deltares report number: 1207671-000-BGS-0016 doi:10.13140/2.1.2447.0721

Farhat, C., Roux, F.X., 1991. A method of finite element tearing and interconnecting and its parallel solution algorithm. Int. J. Numer. Meth. Eng., 32(6): 1205-1227. DOI:10.1002/nme.1620320604

Finney, B.A., Samsuhadi, Willis, W., 1992. Quasi-three-dimensional optimization model of Jakarta basin. J. Water Res. Plan. Man., 118(1): 18-31. DOI:10.1061/(ASCE)0733-9496(1992)118:1(18)

Fleet, M.E., Baird, J. (2001). The development and application of groundwater models to simulate the behaviour of groundwater resources in the Tripoli Aquifer, Libya. Paper presented at the First International Conference on Saltwater Intrusion and Coastal Aquifers-Monitoring, Modeling, and Management. Essaouira, Morocco, April 23–25.

Forrester, A.I.J., Sóbester, A., Keane, A.J., 2008. Constructing a surrogate, Engineering Design via Surrogate Modelling. John Wiley & Sons, Ltd, pp. 33-76. DOI:10.1002/9780470770801.ch2

Fred, G., 1989. Tabu search-part I. ORSA J. Comput., 1: 190-206.

Friedman, J.H., 1991. Multivariate adaptive regression splines (with Discussion). Ann. Stat., 19(1): 1-67. DOI:10.1214/aos/1176347963

Giambastiani, B.M.S., Antonellini, M., Essink, G.H.P.O., Stuurman, R.J., 2007. Saltwater intrusion in the unconfined coastal aquifer of Ravenna (Italy): A numerical model. J. Hydrol., 340(1–2): 91-104. DOI:10.1016/j.jhydrol.2007.04.001

Goel, T., Haftka, R.T., Shyy, W., Queipo, N.V., 2007. Ensemble of surrogates. Struct. Multidiscip. O., 33(3): 199-216. DOI:10.1007/s00158-006-0051-9

Goel, T., Stander, N. (2009). Adaptive simulated annealing for global optimization in LS-OPT. Paper presented at the Seventh European LS-DYNA Conference, LSTC, Salzburg, Austria.

Goldberg, D.E., 1989. Genetic algorithms in search, optimization and machine learning, Addison-Wesley Longman, Boston.

Gong, Y., Zhang, Y., Lan, S., Wang, H., 2016. A Comparative study of artificial neural networks, support vector machines and adaptive neuro fuzzy inference system for forecasting groundwater levels near lake Okeechobee, Florida. Water Resour. Manag., 30(1): 375-391. DOI:10.1007/s11269-015-1167-8

Gorelick, S.M., 1983. A review of distributed parameter groundwater management modeling methods. Water Resour. Res., 19(2): 305-319. DOI:10.1029/WR019i002p00305

Grabow, G.L., Mote, C.R., Sanders, W.L., L Smoot, J., Yoder, D.C., 1993. Groundwater monitoring network design using minimum well density. Water Sci. Technol., 28(3-5): 327-335.

Green, N.R., MacQuarrie, K.T.B., 2014. An evaluation of the relative importance of the effects of climate change and groundwater extraction on seawater intrusion in coastal aquifers in Atlantic Canada. Hydrogeol. J., 22(3): 609-623. DOI:10.1007/s10040-013-1092-y

Green, P.J., Silverman, B.W., 1993. Nonparametric regression and generalized linear models: A roughness penalty approach, Taylor & Francis, Abingdon U.K.

Grundmann, J., Schütze, N., Lennartz, F., 2012. Sustainable management of a coupled groundwater–agriculture hydrosystem using multi-criteria simulation based optimisation. Water Sci. Technol., 67(3): 689-698. DOI:10.2166/wst.2012.602

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol., 377(1): 80-91. DOI:10.1016/j.jhydrol.2009.08.003

Halton, J.H., 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer. Math., 2(1): 84-90. DOI:10.1007/bf01386213

Harvey, C.F. et al., 2006. Groundwater dynamics and arsenic contamination in Bangladesh. Chem. Geol., 228(1): 112-136. DOI:10.1016/j.chemgeo.2005.11.025

Hastie, T., Tibshirani, R., Friedman, J., 2008. The elements of statistical learning (2nd ed.), Springer, New York.

He, Z., Wen, X., Liu, H., Du, J., 2014. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. J. Hydrol., 509: 379-386. DOI:10.1016/j.jhydrol.2013.11.054

Heiss, J.W., Michael, H.A., 2014. Saltwater-freshwater mixing dynamics in a sandy beach aquifer over tidal, spring-neap, and seasonal cycles. Water Resour. Res., 50(8): 6747-6766. DOI:10.1002/2014WR015574

Hemker, T., Fowler, K.R., Farthing, M.W., von Stryk, O., 2008. A mixed-integer simulation-based optimization approach with surrogate functions in water resources management. Optim. Eng., 9(4): 341-360. DOI:10.1007/s11081-008-9048-0

Holland, J.H., 1992. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control and artificial intelligence, MIT Press, Cambridge, MA, 228 pp.

Holman, D. et al., 2014. Gaussian process models for reference ET estimation from alternative meteorological data sources. J. Hydrol., 517: 28-35. DOI:10.1016/j.jhydrol.2014.05.001

Hou, Z., Lu, W., Xue, H., Lin, J., 2017. A comparative research of different ensemble surrogate models based on set pair analysis for the DNAPL-contaminated aquifer remediation strategy optimization. J. Contam. Hydrol., 203: 28-37. DOI:10.1016/j.jconhyd.2017.06.003

Hsieh, W.W., Tang, B., 1998. Applying neural network models to prediction and data analysis in meteorology and oceanography. Bull. Am. Meteorol. Soc., 79(9): 1855-1870. DOI:10.1175/1520-0477(1998)079<1855:annmtp>2.0.co;2

Hugman, R., Stigter, T.Y., Monteiro, J.P., Costa, L., Nunes, L.M., 2015. Modeling the spatial and temporal distribution of coastal groundwater discharge for different water use scenarios under epistemic uncertainty: case study in South Portugal. Environ. Earth Sci., 73(6): 2657-2669. DOI:10.1007/s12665-014-3709-4

Hussain, M.S., Javadi, A.A., Ahangar-Asr, A., Farmani, R., 2015. A surrogate model for simulation–optimization of aquifer systems subjected to seawater intrusion. J. Hydrol., 523: 542-554. DOI:10.1016/j.jhydrol.2015.01.079

Ingber, L., 1989. Very fast simulated re-annealing. Math. Comput. Model., 12(8): 967-973. DOI:10.1016/0895-7177(89)90202-1

Ishigami, H., Fukuda, T., Shibata, T., Arai, F., 1995. Structure optimization of fuzzy neural network by genetic algorithm. Fuzzy Set. Syst., 71(3): 257-264. DOI:10.1016/0165-0114(94)00283-D

Jacobs, J.P., Koziel, S., 2015. Reduced-cost microwave filter modeling using a two-stage Gaussian process regression approach. Int. J. RF. Microw. C. E., 25(5): 453-462. DOI:10.1002/mmce.20880

Jalalkamali, A., 2015. Using of hybrid fuzzy models to predict spatiotemporal groundwater quality parameters. Earth Sci. Inform., 8(4): 885-894. DOI:10.1007/s12145-015-0222-6

Jang, J.-S.R. (1991). Fuzzy modeling using generalized neural networks and Kalman filter algorithm. Paper presented at the Ninth National Conference on Artificial Intelligence, July 14-19, 1991, Anaheim, California. Published by The AAAI Press, Menlo Park, California. Volume 2 pp 762-767.

Jang, J.-S.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern., 23(3): 665-685. DOI:10.1109/21.256541

Jang, J.-S.R., Sun, C.-T., Mizutani, E., 1997. Neuro-fuzzy and soft computing: A computational approach to learning and machine intelligence, Prentice Hall, Upper Saddle River, NJ.

Javadi, A., Hussain, M., Sherif, M., Farmani, R., 2015. Multi-objective optimization of different management scenarios to control seawater intrusion in coastal aquifers. Water Resour. Manag., 29(6): 1843-1857. DOI:10.1007/s11269-015-0914-1

Jin, Y., 2005. A comprehensive survey of fitness approximation in evolutionary computation. Soft Comput., 9(1): 3-12. DOI:10.1007/s00500-003-0328-5

Jin, Y., Branke, J., 2005. Evolutionary optimization in uncertain environments-a survey. IEEE Trans. Evol. Comput., 9(3): 303-317. DOI:10.1109/TEVC.2005.846356

Jovanović, R.Ž., Sretenović, A.A., Živković, B.D., 2015. Ensemble of various neural networks for prediction of heating energy consumption. Energ. Buildings, 94: 189-199. DOI:10.1016/j.enbuild.2015.02.052

Kennedy, J., Eberhart, R. (1995, Nov/Dec 1995). Particle swarm optimization. Paper presented at the ICNN'95 - International Conference on Neural Networks, 27 Nov.-1 Dec. 1995, Perth, WA. IEEE, vol.4, 1942-1948.

Kerrou, J., Renard, P., Tarhouni, J., 2010. Status of the Korba groundwater resources (Tunisia): observations and three-dimensional modelling of seawater intrusion. Hydrogeol. J., 18(5): 1173-1190. DOI:10.1007/s10040-010-0573-5

Ketabchi, H., Ataie-Ashtiani, B., 2015. Evolutionary algorithms for the optimal management of coastal groundwater: A comparative study toward future challenges. J. Hydrol., 520: 193-213. DOI:10.1016/j.jhydrol.2014.11.043

Ketabchi, H., Mahmoodzadeh, D., Ataie- Ashtiani, B., Werner, A.D., Simmons, C.T., 2014. Sea- level rise impact on fresh groundwater lenses in two- layer small islands. Hydrolog. Process., 28(24): 5938-5953. DOI:10.1002/hyp.10059

Khaki, M., Yusoff, I., Islami, N., 2015. Application of the artificial neural network and neuro-fuzzy system for assessment of groundwater quality. Clean (Weinh), 43(4): 551-560. DOI:10.1002/clen.201400267

Khalil, B., Broda, S., Adamowski, J., Ozga-zielinski, B., Donohoe, A., 2015. Short-term forecasting of groundwater levels under conditions of mine-tailings recharge using wavelet ensemble neural network models. Hydrogeol. J., 23(1): 121-141. DOI:10.1007/s10040-014-1204-3

Khashei-Siuki, A., Sarbazi, M., 2015. Evaluation of ANFIS, ANN, and geostatistical models to spatial distribution of groundwater quality (case study: Mashhad plain in Iran). Arab. J. Geosci., 8(2): 903-912. DOI:10.1007/s12517-013-1179-8

Khashei, M., Bijari, M., 2011. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. Appl. Soft Comput., 11(2): 2664-2675. DOI:10.1016/j.asoc.2010.10.015

Kim, K.-Y., Han, W.S., Park, E., 2013. The impact of highly permeable layer on hydraulic system in a coastal aquifer. Hydrolog. Process., 27(22): 3128-3138. DOI:10.1002/hyp.9437

Kinniburgh, D.G., Smedley, P.L. (2001). Arsenic contamination of groundwater in Bangladesh British geological survey technical report WC/00/19 Retrieved from http://nora.nerc.ac.uk/id/eprint/11986/1/WC00019.pdf

Kinniburgh, D.G. et al., 2003. The scale and causes of the groundwater arsenic problem in Bangladesh. In: Welch, A.H., Stollenwerk, K.G. (Eds.), Arsenic in ground water: geochemistry and occurrence. Springer, Boston, MA, pp. 211-257. DOI:10.1007/0-306-47956-7_8

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by Simulated Annealing. Science, 220(4598): 671-680. DOI:10.1126/science.220.4598.671

Kollat, J.B., Reed, P.M., 2007. A computational scaling analysis of multiobjective evolutionary algorithms in long-term groundwater monitoring applications. Adv. Water Resour., 30(3): 408-419. DOI:10.1016/j.advwatres.2006.05.009

Kord, M., Asghari Moghaddam, A., 2014. Spatial analysis of Ardabil plain aquifer potable groundwater using fuzzy logic. J. King Saud Univ. Sci., 26(2): 129-140. DOI:10.1016/j.jksus.2013.09.004

Kourakos, G., Mantoglou, A., 2009. Pumping optimization of coastal aquifers based on evolutionary algorithms and surrogate modular neural network models. Adv. Water Resour., 32(4): 507-521. DOI:10.1016/j.advwatres.2009.01.001

Kourakos, G., Mantoglou, A., 2011. Simulation and multi-objective management of coastal aquifers in semi-arid regions. Water Resour. Manag., 25(4): 1063-1074. DOI:10.1007/s11269-010-9677-x

Kourakos, G., Mantoglou, A., 2013. Development of a multi-objective optimization algorithm using surrogate models for coastal aquifer management. J. Hydrol., 479: 13-23. DOI:10.1016/j.jhydrol.2012.10.050

Kuan, W.K. et al., 2012. Tidal influence on seawater intrusion in unconfined coastal aquifers. Water Resour. Res., 48(2): W02502. DOI:10.1029/2011WR010678

Kurtulus, B., Razack, M., 2010. Modeling daily discharge responses of a large karstic aquifer using soft computing methods: Artificial neural network and neuro-fuzzy. J. Hydrol., 381(1–2): 101-111. DOI:10.1016/j.jhydrol.2009.11.029

Langevin, C.D., Zygnerski, M., 2013. Effect of sea-level rise on ralt water intrusion near a coastal well field in Southeastern Florida. Groundwater, 51(5): 781-803. DOI:10.1111/j.1745-6584.2012.01008.x

Legates, D.R., McCabe, G.J., 1999. Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation. Water Resour. Res., 35(1): 233-241. DOI:10.1029/1998WR900018

Lin, H.-C.J. 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 Waterways Experiment Station Coastal and Hydraulics Laboratory, Vicksburg, MS.

Liu, Y., Shang, S.-H., Mao, X.-M., 2012. Tidal effects on groundwater dynamics in coastal aquifer under different beach slopes. J. Hydrodynam., 24(1): 97-106. DOI:10.1016/S1001-6058(11)60223-0

Llopis- Albert, C., Pulido- Velazquez, D., 2015. Using MODFLOW code to approach transient hydraulic head with a sharp-interface solution. Hydrolog. Process., 29(8): 2052-2064. DOI:10.1002/hyp.10354

Loáiciga, H.A., Pingel, T.J., Garcia, E.S., 2012. Sea water intrusion by sea- level rise: Scenarios for the 21st century. Groundwater, 50(1): 37-47. DOI:10.1111/j.1745-6584.2011.00800.x

Loyola R, D.G., Pedergnana, M., Gimeno García, S., 2016. Smart sampling and incremental function learning for very large high dimensional data. Neural Net., 78: 75-87. DOI:10.1016/j.neunet.2015.09.001

Lu, C., Xin, P., Li, L., Luo, J., 2015. Seawater intrusion in response to sea-level rise in a coastal aquifer with a general-head inland boundary. J. Hydrol., 522: 135-140. DOI:10.1016/j.jhydrol.2014.12.053

Lu, W., Yang, Q., Martin, J.D., Juncosa, R., 2013. Numerical modelling of seawater intrusion in Shenzhen (China) using a 3D density-dependent model including tidal effects. J. Earth Syst. Sci., 122(2): 451-465. DOI:10.1007/s12040-013-0273-3

Luo, J., Lu, W., 2014. Comparison of surrogate models with different methods in groundwater remediation process. J. Earth Syst. Sci., 123(7): 1579-1589. DOI:10.1007/s12040-014-0494-0

Marin, L.E., Perry, E.C., Essaid, H.I., Steinich, B. (2001). Hydrogeological investigations and numerical simulation of groundwater flow in the karstic aquifer of northwestern Yucatan, Mexico Paper presented at the First International Conference on Saltwater Intrusion and Coastal Aquifers-Monitoring, Modeling, and Management, Essaouira, Morocco, April 23-25.

Marryott, R.A., Dougherty, D.E., Stollar, R.L., 1993. Optimal groundwater management: 2. Application of simulated annealing to a field- scale contamination site. Water Resour. Res., 29(4): 847-860. DOI:10.1029/92WR02801

MATLAB, 2017a. Fuzzy logic toolbox: MATLAB version 2017a The Mathworks Inc, Mathworks, Natick, MA.

MATLAB, 2017b. MATLAB Version R2017a. The Mathworks Inc, Mathworks, Natick, MA.

MATLAB, 2017c. Parallel computing toolbox: MATLAB version R2017a The Mathworks Inc, Mathworks, Natick, MA.

MATLAB. (2017d). What is sugeno-type fuzzy inference? MATLAB documentation. Retrieved from http://au.mathworks.com/help/fuzzy/what-is-sugeno-type-fuzzy-inference.html

Mazi, K., Koussis, A.D., Destouni, G., 2013. Tipping points for seawater intrusion in coastal aquifers under rising sea level. Environ. Res. Lett., 8(1): 014001. DOI:10.1088/1748-9326/8/1/014001

McKinney, D.C., Lin, M.-D., 1994. Genetic algorithm solution of groundwater management models. Water Resour. Res., 30(6): 1897-1906. DOI:10.1029/94WR00554

Melloul, A., Collin, M., 2006. Hydrogeological changes in coastal aquifers due to sea level rise. Ocean Coast. Manage., 49(5): 281-297. DOI:10.1016/j.ocecoaman.2006.03.009

Meyer, P.D., Brill Jr., E.D., 1988. A method for locating wells in a groundwater monitoring network under conditions of uncertainty. Water Resour. Res., 24(8): 1277-1282. DOI:10.1029/WR024i008p01277

Meyer, P.D., Valocchi, A.J., Eheart, J.W., 1994. Monitoring network design to provide initial detection of groundwater contamination. Water Resour. Res., 30(9): 2647-2659. DOI:10.1029/94WR00872

Michael, H.A., Voss, C.I., 2009. Controls on groundwater flow in the Bengal Basin of India and Bangladesh: regional modeling analysis. Hydrogeol. J., 17(7): 1561. DOI:10.1007/s10040-008-0429-4

Mohammadi, K., Shamshirband, S., Petković, D., Yee, P.L., Mansor, Z., 2016. Using ANFIS for selection of more relevant parameters to predict dew point temperature. Appl. Therm. Eng., 96: 311-319. DOI:10.1016/j.applthermaleng.2015.11.081

Mondal, M.S., Saleh, A.F.M., 2003. Evaluation of some deep and shallow tubewell irrigated schemes in Bangladesh using performance indicators. Agr. Water Manage., 58(3): 193-207. DOI:10.1016/S0378-3774(02)00087-2

Mugunthan, P., Shoemaker, C.A., 2004. Time varying optimization for monitoring multiple contaminants under uncertain hydrogeology. Bioremediat. J., 8(3-4): 129-146. DOI:10.1080/10889860490887509

Muhammetoglu, A., Yardimci, A., 2006. A fuzzy logic approach to assess groundwater pollution levels below agricultural fields. Environ. Monit. Assess., 118(1): 337-354. DOI:10.1007/s10661-006-1497-3

Nakayama, H., Arakawa, M., Sasaki, R. (2001). A computational intelligence approach to optimization with unknown objective functions. Paper presented at the ICANN 2001: Artificial Neural Networks, International Conference on Artificial Neural Networks.Lecture Notes in Computer Science book series (LNCS, volume 2130), pp.73-80, Springer, Berlin.

Namura, N., Shimoyama, K., Jeong, S., Obayashi, S. (2011). Kriging/RBF-hybrid response surface method for highly nonlinear functions. Paper presented at the 2011 IEEE Congress of Evolutionary Computation (CEC). DOI:10.1109/CEC.2011.5949933

Narayan, K.A., Schleeberger, C., Bristow, K.L., 2007. Modelling seawater intrusion in the Burdekin Delta Irrigation Area, North Queensland, Australia. Agr. Water Manage., 89(3): 217-228. DOI:10.1016/j.agwat.2007.01.008

Nocchi, M., Salleolini, M., 2013. A 3D density-dependent model for assessment and optimization of water management policy in a coastal carbonate aquifer exploited for water supply and fish farming. J. Hydrol., 492: 200-218. DOI:10.1016/j.jhydrol.2013.03.048

Nunes, L.M., Cunha, M.C., Ribeiro, L., 2004. Groundwater monitoring network optimization with redundancy reduction. J. Water Res. Plan. Man., 130(1): 33-43. DOI:10.1061/(ASCE)0733-9496(2004)130:1(33)

Oliveira, M.V., Schirru, R., 2009. Applying particle swarm optimization algorithm for tuning a neuro-fuzzy inference system for sensor monitoring. Prog. Nucl. Energ., 51(1): 177-183. DOI:10.1016/j.pnucene.2008.03.007

Oude Essink, G.H.P., 2001a. Improving fresh groundwater supply—problems and solutions. Ocean Coast. Manage., 44(5): 429-449. DOI:10.1016/S0964-5691(01)00057-6

Oude Essink, G.H.P., 2001b. Salt water intrusion in a three-dimensional groundwater system in the Netherlands: A numerical study. Transport Porous. Med., 43(1): 137-158. DOI:10.1023/a:1010625913251

Oude Essink, G.H.P., van Baaren, E.S., de Louw, P.G.B., 2010. Effects of climate change on coastal groundwater systems: A modeling study in the Netherlands. Water Resour. Res., 46(10): W00F04. DOI:10.1029/2009WR008719

Papadopoulou, M.P., Nikolos, I.K., Karatzas, G.P., 2010. Computational benefits using artificial intelligent methodologies for the solution of an environmental design problem: saltwater intrusion. Water Sci. Technol., 62(7): 1479–1490.

Pebesma, E.J., Heuvelink, G.B.M., 1999. Latin hypercube sampling of gaussian random fields. Technometrics, 41(4): 303-312. DOI:10.2307/1271347

Pham, V.H., Lee, S.-I., 2015. Assessment of seawater intrusion potential from sea-level rise and groundwater extraction in a coastal aquifer. Desalin. Water Treat., 53(9): 2324-2338. DOI:10.1080/19443994.2014.971617

Pillay, N., 2004. An investigation into the use of genetic programming for the induction of noviceprocedural programming solution algorithms in intelligent programming tutors. Computer Science PhD Thesis, University of KwaZulu-Natal, Durban, South Africa.

Piret, C., 2007. Analytical and numerical advances in radial basis functions. Department of Applied Mathematics, PhD Thesis, University of Colorado, Boulder, CO.

Prakash, O., Datta, B., 2014. Multiobjective monitoring network design for efficient identification of unknown groundwater pollution sources incorporating genetic programming-based monitoring. J. Hydrol. Eng., 19(11): 04014025. DOI:10.1061/(ASCE)HE.1943-5584.0000952

Qi, S.-Z., Qiu, Q.-L., 2011. Environmental hazard from saltwater intrusion in the Laizhou Gulf, Shandong Province of China. Nat. Hazards., 56(3): 563-566. DOI:10.1007/s11069-010-9686-3

Raghavendra, N.S., Deka, P.C., 2015. Forecasting monthly groundwater level fluctuations in coastal aquifers using hybrid Wavelet packet–Support vector regression. Cogent Eng., 2(1): 999414. DOI:10.1080/23311916.2014.999414

Rajabi, M.M., Ketabchi, H., 2017. Uncertainty-based simulation-optimization using Gaussian process emulation: Application to coastal groundwater management. J. Hydrol., 555(Supplement C): 518-534. DOI:10.1016/j.jhydrol.2017.10.041

Rao, S.S., 1996. Engineering optimization: theory and practice, John Wiley and Sons, Hoboken, NJ.

Rasmussen, C.E., Williams, C.K.I., 2005. Gaussian processes for machine learning, The MIT Press, Cambridge, MA.

Razavi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. Water Resour. Res., 48(7): W07401. DOI:10.1029/2011WR011527

Reddy, M.J., Nagesh, K.D., 2007. Multi- objective particle swarm optimization for generating optimal trade- offs in reservoir operation. Hydrolog. Process., 21(21): 2897-2909. DOI:10.1002/hyp.6507

Reed, P., Minsker, B., Valocchi, A.J., 2000. Cost- effective long-term groundwater monitoring design using a genetic algorithm and global mass interpolation. Water Resour. Res., 36(12): 3731-3741. DOI:10.1029/2000WR900232

Reed, P.M., Minsker, B.S., 2004. Striking the balance: long-term groundwater monitoring design for conflicting objectives. J. Water Res. Plan. Man., 130(2): 140-149. DOI:10.1061/(ASCE)0733-9496(2004)130:2(140)

Reeves, C.R. (Ed.), 1993. Modern heuristic techniques for combinatorial problems. John Wiley and Sons, New York, 320 pp.

Rini, D.P., Shamsuddin, S.M., Yuhaniz, S.S., 2016. Particle swarm optimization for ANFIS interpretability and accuracy. Soft Comput., 20(1): 251-262. DOI:10.1007/s00500-014-1498-z

Ritzel, B.J., Eheart, J.W., Ranjithan, S., 1994. Using genetic algorithms to solve a multiple objective groundwater pollution containment problem. Water Resour. Res., 30(5): 1589-1603. DOI:10.1029/93WR03511

Robinson, D.W., 2008. Entropy and uncertainty. Entropy, 10: 493-506. DOI:10.3390/e10040493

Rosenwald, G.W., Green, D.W., 1974. A method for determining the optimum location of wells in a reservoir using mixed-integer programming. Soc. Pet. Eng. J., 14(1): 44-54. DOI:10.2118/3981-PA

Roy, D.K., Datta, B., 2017a. Fuzzy c-mean clustering based inference system for saltwater intrusion processes prediction in coastal aquifers. Water Resour. Manag., 31(1): 355-376. DOI:10.1007/s11269-016-1531-3

Roy, D.K., Datta, B., 2017b. Multivariate adaptive regression spline ensembles for management of multilayered coastal aquifers. J. Hydrol. Eng., 22(9): 04017031. DOI:10.1061/(ASCE)HE.1943-5584.0001550

Roy, D.K., Datta, B. (2017c). Optimal management of groundwater extraction to control saltwater intrusion in multi-layered coastal aquifers using ensembles of adaptive neuro-fuzzy inference system. Paper presented at the World Environmental and Water Resources Congress 2017, May 21–25, 2017, American Society of Civil Engineers, Sacramento, CA, pp.139-150.

Rozell, D.J., Wong, T.-F., 2010. Effects of climate change on groundwater resources at Shelter Island, New York State, USA. Hydrogeol. J., 18(7): 1657-1665. DOI:10.1007/s10040-010-0615-z

Sahu, M., Mahapatra, S.S., Sahu, H.B., Patel, R.K., 2011. Prediction of water quality index using neuro fuzzy inference system. Water Qual. Expos. Hea., 3(3): 175-191. DOI:10.1007/s12403-011-0054-7

Sakr, S.A., 1999. Validity of a sharp-interface model in a confined coastal aquifer. Hydrogeol. J., 7(2): 155-160. DOI:10.1007/s100400050187

Salford-Systems, 2016. SPM users guide: introducing MARS, Salford Systems, San Diego, CA.

Samadi, M., Jabbari, E., Azamathulla, H.M., Mojallal, M., 2015. Estimation of scour depth below free overfall spillways using multivariate adaptive regression splines and artificial neural networks. Eng. Appl. Comp. Fluid, 9(1): 291-300. DOI:10.1080/19942060.2015.1011826

Sanford, W.E., Pope, J.P., 2010. Current challenges using models to forecast seawater intrusion: lessons from the Eastern Shore of Virginia, USA. Hydrogeol. J., 18(1): 73-93. DOI:10.1007/s10040-009-0513-4

Sbai, M.A., Larabi, A., De Smedt, F., 1998. Modelling saltwater intrusion by a 3-D sharp interface finite element model. WIT Trans. Ecol. Envir., 24.

Schrage, L., 1999. Optimization modelling with Lingo, LINDO Systems Inc, Chicago, IL.

Shalabey, M.E., Kashyap, D., Sharma, A., 2006. Numerical model of saltwater transport toward a pumping well. J. Hydrol. Eng., 11(4): 306-318. DOI:10.1061/(ASCE)1084-0699(2006)11:4(306)

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J., 27(3): 379-423. DOI:10.1002/j.1538-7305.1948.tb01338.x

Sharda, V.N., Patel, R.M., Prasher, S.O., Ojasvi, P.R., Prakash, C., 2006. Modeling runoff from middle Himalayan watersheds employing artificial intelligence techniques. Agr. Water Manage., 83(3): 233-242. DOI:10.1016/j.agwat.2006.01.003

Sharda, V.N., Prasher, S.O., Patel, R.M., Ojasvi, P.R., Prakash, C., 2008. Performance of Multivariate Adaptive Regression Splines (MARS) in predicting runoff in mid-Himalayan micro-watersheds with limited data. Hydrol. Sci. J., 53(6): 1165-1175. DOI:10.1623/hysj.53.6.1165

Sharkey, A.J.C., 1999. Combining artificial neural nets: Ensemble and modular multi-net systems, Springer-Verlag, New York.

Sherif, M., Sefelnasr, A., Ebraheem, A.A., Javadi, A., 2014. Quantitative and qualitative assessment of seawater intrusion in Wadi Ham under different pumping scenarios. J. Hydrol. Eng., 19(5): 855-866. DOI:10.1061/(ASCE)HE.1943-5584.0000907

Sherif, M.M., Singh, V.P., 1999. Effect of climate change on sea water intrusion in coastal aquifers. Hydrolog. Process., 13(8): 1277-1287. DOI:10.1002/(SICI)1099-1085(19990615)13:8<1277::AID-HYP765>3.0.CO;2-W

Shiri, J., Kişi, Ö., 2011. Comparison of genetic programming with neuro-fuzzy systems for predicting short-term water table depth fluctuations. Comput. Geosci., 37(10): 1692-1701. DOI:10.1016/j.cageo.2010.11.010

Shrivastava, G.S., 1998. Impact of sea level rise on seawater intrusion into coastal aquifer. J. Hydrol. Eng., 3(1): 74-78. DOI:10.1061/(ASCE)1084-0699(1998)3:1(74)

Shu, C., Burn, D.H. (2004). Artificial neural network ensembles and their application in pooled flood frequency analysis. Water Resour. Res., 40(9). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2003WR002816 DOI:10.1029/2003WR002816

Shu, C., Ouarda, T.B.M.J. (2007). Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space. Water Resour. Res., 43(7), W07438. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2006WR005142 DOI:10.1029/2006WR005142

Shu, C., Ouarda, T.B.M.J., 2008. Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system. J. Hydrol., 349(1): 31-43. DOI:10.1016/j.jhydrol.2007.10.050

Simpson, M.J., Clement, T.P., 2003. Theoretical analysis of the worthiness of Henry and Elder problems as benchmarks of density-dependent groundwater flow models. Adv. Water Resour., 26(1): 17-31. DOI:10.1016/S0309-1708(02)00085-4

Simpson, T.B., Holman, I.P., Rushton, K.R., 2010. Understanding and modelling spatial drain–aquifer interactions in a low-lying coastal aquifer—the Thurne catchment, Norfolk, UK. Hydrolog. Process., 25(4): 580-592. DOI:10.1002/hyp.7845

Sóbester, A., Forrester, A.I.J., Toal, D.J.J., Tresidder, E., Tucker, S., 2014. Engineering design applications of surrogate-assisted optimization techniques. Optim. Eng., 15(1): 243-265. DOI:10.1007/s11081-012-9199-x

SPM, 2016. SPM® (Version 8.2), Salford Predictive Modeller, Salford Systems, San Diego, CA.

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(3–4): 245-256. DOI: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(13): 3201-3218. DOI: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., 47(4): W04516. DOI:10.1029/2010WR009683

Sreekanth, J., Datta, B., 2011c. Optimal combined operation of production and barrier wells for the control of saltwater intrusion in coastal groundwater well fields. Desalin. Water Treat., 32(1-3): 72-78. DOI:10.5004/dwt.2011.2680

Sreekanth, J., Datta, B., 2014a. Design of an optimal compliance monitoring network and feedback information for adaptive management of saltwater intrusion in coastal aquifers. J. Water Res. Plan. Man., 140(10): 04014026. DOI:10.1061/(ASCE)WR.1943-5452.0000406

Sreekanth, J., Datta, B., 2014b. Stochastic and robust multi-objective optimal management of pumping from coastal aquifers under parameter uncertainty. Water Resour. Manag., 28(7): 2005-2019. DOI:10.1007/s11269-014-0591-5

Sreekanth, J., Datta, B., Mohapatra, P.K., 2012. Optimal short-term reservoir operation with integrated long-term goals. Water Resour. Manag., 26(10): 2833-2850. DOI:10.1007/s11269-012-0051-z

Sreekanth, P.D., Sreedevi, P.D., Ahmed, S., Geethanjali, N., 2011. Comparison of FFNN and ANFIS models for estimating groundwater level. Environ. Earth Sci., 62(6): 1301-1310. DOI:10.1007/s12665-010-0617-0

Strack, O.D.L., 1976. A single- potential solution for regional interface problems in coastal aquifers. Water Resour. Res., 12(6): 1165-1174. DOI:10.1029/WR012i006p01165

Subagadis, Y.H., Grundmann, J., Schütze, N., Schmitz, G.H., 2014. An integrated approach to conceptualise hydrological and socio-economic interaction for supporting management decisions of coupled groundwater–agricultural systems. Environ. Earth Sci., 72(12): 4917-4933. DOI:10.1007/s12665-014-3238-1

Sudheer, C., Mathur, S., 2010. Modeling uncertainty analysis in flow and solute transport model using Adaptive Neuro Fuzzy Inference System and particle swarm optimization. KSCE J. Civ. Eng., 14(6): 941-951. DOI:10.1007/s12205-010-0865-2

Sugeno, M., 1985. Industrial applications of fuzzy control, Elsevier Science, New York, 269 pp.

Sugeno, M., Yasukawa, T., 1993. A fuzzy-logic-based approach to qualitative modeling. IEEE Trans. Fuzzy Syst., 1(1): 7. DOI:10.1109/tfuzz.1993.390281

Sun, A.Y., Wang, D., Xu, X., 2014. Monthly streamflow forecasting using Gaussian process regression. J. Hydrol., 511: 72-81. DOI:10.1016/j.jhydrol.2014.01.023

Takagi, T., Sugeno, M., 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. Syst. Man Cybern., SMC-15(1): 116-132. DOI:10.1109/TSMC.1985.6313399

Todd, D.K., 1974. Salt-water intrusion and its control. J. Am. Water Works Ass., 66(3): 180-187.

Torczon, V., Trosset, M.W., 1998. From evolutionary operation to parallel direct search: pattern search algorithms for numerical optimization. Comp. Sci. Stat., 29: 396-401.

Tu, J.V., 1996. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J. Clin. Epidemiol., 49(11): 1225-1231. DOI:10.1016/S0895-4356(96)00002-9

Tutmez, B., Hatipoglu, Z., Kaymak, U., 2006. Modelling electrical conductivity of groundwater using an adaptive neuro-fuzzy inference system. Comput. Geosci., 32(4): 421-433. DOI:10.1016/j.cageo.2005.07.003

Unsal, B., Yagbasan, O., Yazicigil, H., 2014. Assessing the impacts of climate change on sustainable management of coastal aquifers. Environ. Earth Sci., 72(6): 2183-2193. DOI:10.1007/s12665-014-3130-z

Virginia, M.J., Leah, L.R., 2000. Accuracy of neural network approximators in simulation-optimization. J. Water Res. Plan. Man., 126(2): 48-56. DOI:10.1061/(ASCE)0733-9496(2000)126:2(48)

Wada, Y. et al., 2010. Global depletion of groundwater resources. Geophys. Res. Lett., 37(20). DOI:10.1029/2010GL044571

Wagner, B.J., 1995. Sampling design methods for groundwater modeling under uncertainty. Water Resour. Res., 31(10): 2581-2591. DOI:10.1029/95WR02107

Walther, M. et al., 2014. Assessing the saltwater remediation potential of a three-dimensional, heterogeneous, coastal aquifer system: Model verification, application and visualization for transient density-driven seawater intrusion. Environ. Earth Sci., 72: 3827–3837. DOI:10.1007/s12665-014-3253-2

Wang, Z., Rangaiah, G.P., 2017. Application and analysis of methods for selecting an optimal solution from the Pareto-optimal front obtained by multiobjective optimization. Ind. Eng. Chem. Res., 56(2): 560-574. DOI:10.1021/acs.iecr.6b03453

Watson, T.A., Werner, A.D., Simmons, C.T., 2010. Transience of seawater intrusion in response to sea level rise. Water Resour. Res., 46(12). DOI:10.1029/2010WR009564

Webb, M.D., Howard, K.W.F., 2011. Modeling the transient response of saline intrusion to rising sea-levels. Groundwater, 49(4): 560-569. DOI:10.1111/j.1745-6584.2010.00758.x

Werbos, P.J., 1974. Beyond regression: new tools for prediction and analysis in the behavioral sciences. Ph.D. Thesis, Harvard University, Cambridge, MA.

Werner, A.D., Simmons, C.T., 2009. Impact of sea-level rise on sea water intrusion in coastal aquifers. Groundwater, 47(2): 197-204. DOI:10.1111/j.1745-6584.2008.00535.x

Werner, A.D. et al., 2012. Vulnerability indicators of sea water intrusion. Groundwater, 50(1): 48-58. DOI:10.1111/j.1745-6584.2011.00817.x

Willis, R., Finney, B.A., 1988. Planning model for optimal control of saltwater intrusion. J. Water Res. Plan. Man., 114(2): 163-178. DOI:10.1061/(ASCE)0733-9496(1988)114:2(163)

Willmott, C.J., 1984. On the evaluation of model performance in physical geography. In: Gaile, G.L., Willmott, C.J. (Eds.), Spatial statistics and models. Springer Netherlands, Dordrecht, pp. 443-460. DOI:10.1007/978-94-017-3048-8_23

Wolpert, D.H., 1992. Original contribution: Stacked generalization. Neural Net., 5(2): 241-259. DOI:10.1016/s0893-6080(05)80023-1

Yang, C.-C., Prasher, S.O., Lacroix, R., Kim, S.H., 2003. A multivariate adaptive regression splines model for simulation of pesticide transport in soils. Biosyst. Eng., 86(1): 9-15. DOI:10.1016/S1537-5110(03)00099-0

Yang, J., Graf, T., Ptak, T., 2015. Sea level rise and storm surge effects in a coastal heterogeneous aquifer: a 2D modelling study in northern Germany. Grundwasser, 20(1): 39-51. DOI:10.1007/s00767-014-0279-z

Yang, Y.S., Kalin, R.M., Zhang, Y., Lin, X., Zou, L., 2001. Multi-objective optimization for sustainable groundwater resource management in a semiarid catchment. Hydrol. Sci. J., 46(1): 55-72. DOI:10.1080/02626660109492800

Yechieli, Y., Shalev, E., Wollman, S., Kiro, Y., Kafri, U., 2010. Response of the Mediterranean and Dead Sea coastal aquifers to sea level variations. Water Resour. Res., 46(12). DOI:10.1029/2009WR008708

Zabihi, M., Pourghasemi, H.R., Pourtaghi, Z.S., Behzadfar, M., 2016. GIS-based multivariate adaptive regression spline and random forest models for groundwater potential mapping in Iran. Environ. Earth Sci., 75(8): 665. DOI:10.1007/s12665-016-5424-9

Zaier, I., Shu, C., Ouarda, T.B.M.J., Seidou, O., Chebana, F., 2010. Estimation of ice thickness on lakes using artificial neural network ensembles. J. Hydrol., 383(3): 330-340. DOI:10.1016/j.jhydrol.2010.01.006

Zanaganeh, M., Mousavi, S.J., Etemad Shahidi, A.F., 2009. A hybrid genetic algorithm–adaptive network-based fuzzy inference system in prediction of wave parameters. Eng. Appl. Artific.Intel., 22(8): 1194-1202. DOI:10.1016/j.engappai.2009.04.009

Zerpa, L.E., Queipo, N.V., Pintos, S., Salager, J.-L., 2005. An optimization methodology of alkaline–surfactant–polymer flooding processes using field scale numerical simulation and multiple surrogates. J. Petrol. Sci. Eng., 47(3): 197-208. DOI:10.1016/j.petrol.2005.03.002

Zhang, G.P., Berardi, V.L., 2001. Time series forecasting with neural network ensembles: an application for exchange rate prediction. J. Oper. Res. Soc., 52(6): 652-664. DOI:10.1057/palgrave.jors.2601133

Zhang, W., Chen, X., Tan, H., Zhang, Y., Cao, J., 2015. Geochemical and isotopic data for restricting seawater intrusion and groundwater circulation in a series of typical volcanic islands in the South China Sea. Mar. Pollut. Bull., 93(1): 153-162. DOI:10.1016/j.marpolbul.2015.01.024

Zhang, W., Goh, A.T.C., 2016. Multivariate adaptive regression splines and neural network models for prediction of pile drivability. Geosci. Frontiers, 7(1): 45-52. DOI:10.1016/j.gsf.2014.10.003

Zhao, K.Q., Xuan, A.L., 1996. Set pair theory-a new theory method of non-define and its applications (in Chinese). Syst. Eng., 14(1): 18-23.

Zhou, Z.-H., Wu, J., Tang, W., 2002. Ensembling neural networks: many could be better than all. Artif. Intell., 137(1): 239-263. DOI:10.1016/S0004-3702(02)00190-X