Water Resources Management (2022) 36:1863–1878
https://doi.org/10.1007/s11269-022-03112-5

Optimization of Groundwater Pumping and River-Aquifer Exchanges for Management of Water Resources

Mayank Bajpai[1] · Shreyansh Mishra[1] · Shishir Gaur[1] · Anurag Ohri[1] · Hervé Piégay[3] · Didier Graillot[2]

Received: 24 September 2021 / Accepted: 11 March 2022 /Published online: 13 April 2022

Abstract

Multi-objective optimization problems can be solved through Simulation-Optimization (S-O) techniques where the pareto front gives the optimal solutions in the problem domains. During the selection of different modelling methods, optimization techniques and management scenarios, several pareto fronts can be generated. In the present work, an attempt has been made by performing intensive comparisons between different pareto fronts to compare the efficiency and convergence of different S-O models. In this process, groundwater models were developed to simulate the River-Aquifer (R-A) exchanges for the study area as groundwater pumping influences the rate of R-A exchanges and alters the flow dynamics. The developed models were coupled with optimization models and were executed to solve the multi-objective optimization problems based on the maximization of discharge through pumping wells and maximization of groundwater input into the river through R-A exchanges. The distinctive features of the paper include a pareto front comparison where fronts developed by different S-O models were compared and analysed based on various parameters. The results show the dominance of Multi-Objective Particle Swarm Optimization (MOPSO) over other optimization algorithms and concluded that the maximization of pumping rate significantly changes after considering the R-A exchanges-based objective functions. This study concludes that the model domain also alters the output of simulation–optimization. Therefore, model domain and corresponding boundary conditions should be selected carefully for the field application of management models. The artificial neural network (ANN) models have been also developed to deal with the computationally expensive simulation models by reducing the processing time and found efficient.

Keywords Simulation-optimization · Multi-objective optimization · Artificial neural network · River-Aquifer exchanges

✉ Shreyansh Mishra
shreyanshmishra.rs.civ20@itbhu.ac.in

[1] Department of Civil Engineering, Indian Institute of Technology (BHU), Varanasi, India

[2] CNRS, UMR 5600 Environnement Ville Et Société, Géosciences et environnement, Ecole Nationale Supérieure Des Mines de Saint-Etienne, 42023 Saint-Etienne Cedex 2, France

[3] CNRS-UMR 5600, Université de Lyon, ENS de Lyon, Plateforme ISIG. 15 Parvis René Descartes, BP 7000, 69342 Lyon Cedex 07, France

1 Introduction

The R-A exchange influences both the quantity and the quality aspects of groundwater and surface water systems significantly. Consequently, proper and effective quantification and representation of R-A exchange are very important for the management of water resources and aquatic ecosystems (Gomo 2011). The climate and groundwater pumping for irrigation have caused rapid groundwater depletion in India and other parts of the world (Dangar et al. 2021). For the best possible conservation of water resources, Conant et al. (2019) suggested that it is essential to understand and quantify the exchange activities between rivers and groundwater. The exchange between rivers and groundwater is significant in a variety of current concerns, including providing drinking water, characterizing and managing environmental flow regimes, preserving or restoring riverine ecosystem health and functioning, and reducing the effect of toxins (Paran et al. 2012). Several works have been carried out for the identification and the quantification of groundwater and surface water interactions i.e., R-A exchanges (Atwell et al. 1971; Kay et al. 2005; Handcock et al. 2006; Loheide and Gorelick 2006; Cristea and Burges 2009; Wawrzyniak et al. 2012; Constantz 1998; Sophocleous 2002; Becker et al. 2004; Anderson 2005; Kalbus et al. 2006; Keery et al. 2007; Lowry et al. 2007). However, some studies also demonstrate the effects of R-A exchange on river temperature (Westhoff et al. 2007; Burkholder et al. 2008; Hebert et al. 2011).

The Simulation–Optimization (S–O) techniques are often used in identifying the optimal management practices of groundwater for the selected area (Etsias and Katsifarakis 2017; Mao et al. 2017). In multi-objective optimization problems, result comparison becomes more important since various Pareto fronts can be developed using S–O models and management scenarios, with some similarities and perhaps few differences. Multi-objective optimization problems related to water resources have been studied and solved using different algorithms. (Hernández-Lobato et al. 2016; Emmerich and Deutz 2018; Abd-Elmaboud et al. 2021; Marjit and Hopfe 2009; Asadzadeh et al. 2014; Audet et al. 2020; Jha et al. 2020; Li et al. 2010).

The qualitative assessment of Pareto fronts which was addressed by Calandra et al. (2014) and Belakaria and Deshwal (2019) brings up an interesting point to qualitatively analyse the pareto front developed by multi-objective optimization (Horn et al. 2018; Höllermann and Evers 2019). Moreover, calculating the bias and its statistical analysis associated with the solution of multi-objective problems becomes very decisive in understanding the uncertainties concerning Pareto fronts (Binois et al. 2015; Cao et al. 2017; Bassi et al. 2018; Avent et al. 2020; Asadzadeh et al. 2014; Marjit and Hopfe 2009).

This present work was carried out to compare the different pareto fronts quantitatively by addressing the issues of R-A exchanges. Different model domain demarcations for the same river system were considered for the development of the groundwater model, and their impact on R-A exchanges was analysed. In addition, various optimization techniques were also used and their outputs (i.e., Pareto fronts) were compared. Finally, this set of optimal result was interpreted in terms of groundwater management.

2 Study Area

The Ain River has an oceanic hydrologic regime and drains a basin of about 3630 km^2. It is located in the southern Jura Mountains, France. The length of the Ain River is about 200 km and it is the right bank tributary of the Rhône River. The lower part of the Ain River, which is located between the Allement dam and the confluence with the Rhône

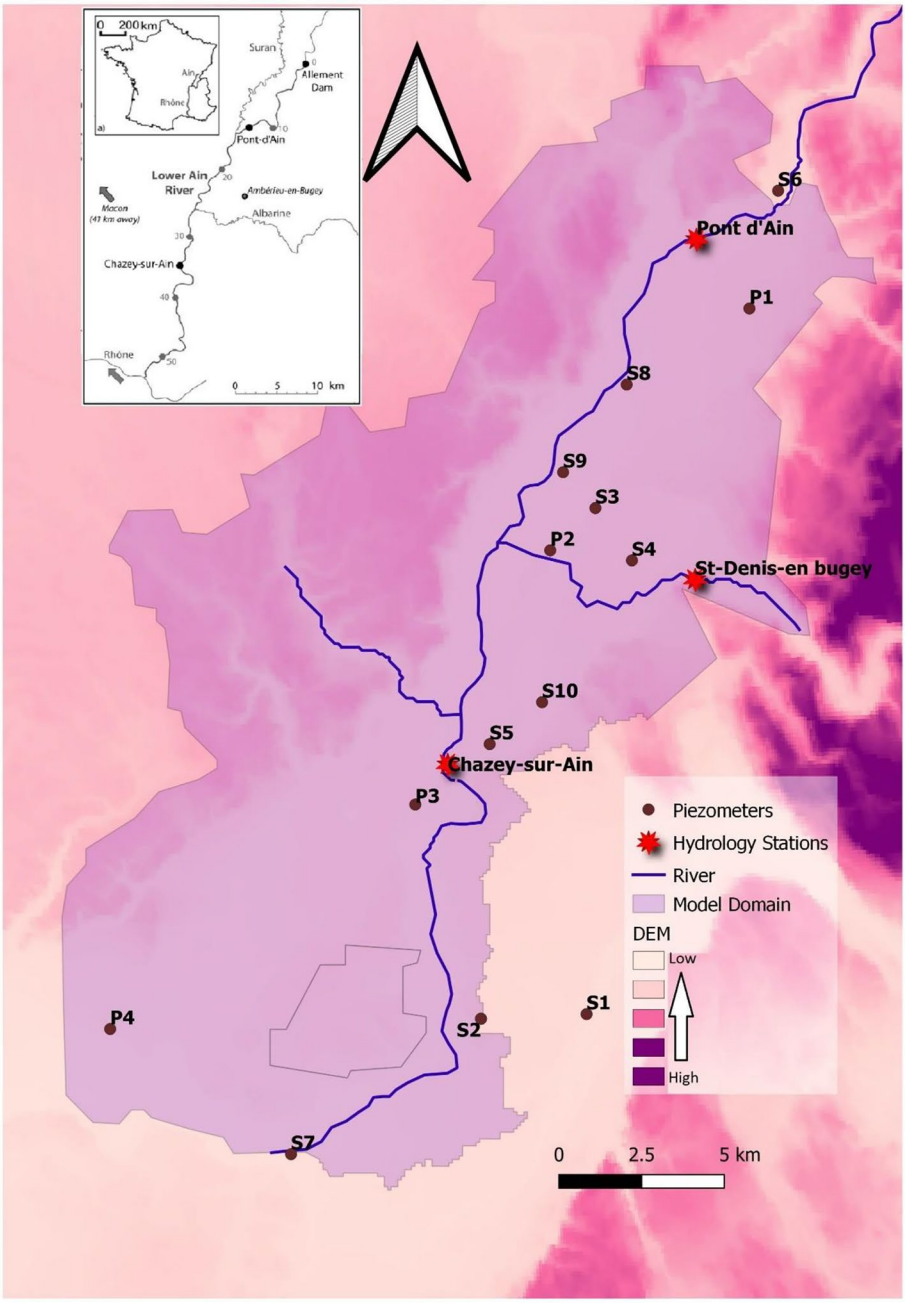

Fig. 1 Study Area

River, was taken for the present study purpose. Figure 1 shows the location of the lower Ain River area. Along this section, the average channel width is about 60 m and the slope is about 1.3%. The mean annual discharge of the river, between 1959 to 2019, was evaluated

as 120 m^3/s at the measuring station of Chazey-sur-Ain. Most floods occur between October and March and summer is characterized by low flow conditions. Further details about the study area are provided in the groundwater model section.

3 Methodology

In the objective functions, the first aim was to identify the maximum pumping rates of the wells by considering the R-A exchanges component. This also includes examining the impact of different types of model domains on the output of R-A exchanges and to find the best model domain and boundary conditions. Whereas, the second objective was to compare the Pareto front developed by different optimization techniques using coupled simulation–optimization models. Total four optimization techniques, (Multi-objective Genetic Algorithm (MOGA), MOPSO, Pareto Search (PS), and Multi-objective Evolutionary Algorithm Based on Decomposition (MOEA/D)) were compared. To deal with a large number of wells, the wells in the area were grouped based on the municipal zones and distance of the wells from the river. If any well in the municipal zone had a distance from the river less than a threshold of 1 km, then it was classified as a new well zone. Figure 2 shows the flow chart of the methodology.

Groundwater models were developed to simulate the R-A exchanges for the lower part of the River Ain, France. These models were coupled with optimization algorithms in the MATLAB platform. To implement the coupling between the two systems, MATLAB scripts were developed which carried out the following tasks, (i) run the simulation model, extract and post-process the output of the simulation model to calculate a cost for optimization model, (ii) read and write the decision variables from the '.h5' input file for MODFLOW, (iii) run the optimization model with MODFLOW as cost function.

Further, the set of optimized costs were compared using various criteria (convergence, diversity, and uniformity) and metrics (hypervolume, spread, and Inverted Generational Distance).

3.1 Groundwater Model Development

The groundwater flow modelling was performed using MODFLOW (McDonald and Harbaugh 1988). A conceptual model was developed by creating the different geospatial data-based input layers for defining the surface recharge, boundary conditions, and various aquifer parameters. The initial piezometric surface of groundwater was created using hydrograph data of 15 wells in the study area.

The top, bottom and surface of the study area was found in the range of 240 m to 550 m where Shuttle Radar Topography Mission (SRTM) data was used. The bottom surface was prepared with the help of well log data obtained by Bureau de Recherches Géologiques et Minières (BRGM). The two-layered model was developed for the underlying sediments where each layer was assumed to be horizontal, homogeneous, and isotropic. The mean thickness of the layer was taken as 25 m. The initial values of hydraulic conductivity were taken from 0.0018 m/s for the older sediment and 0.003 m/s for the younger sediment. Specific yield (Ss) values for the alluvial deposits were found in the range from 1 to 17%.

Boundary Conditions The model was developed by defining two types of external model boundaries i.e., constant head and constant flow boundaries. Figure 3 shows the

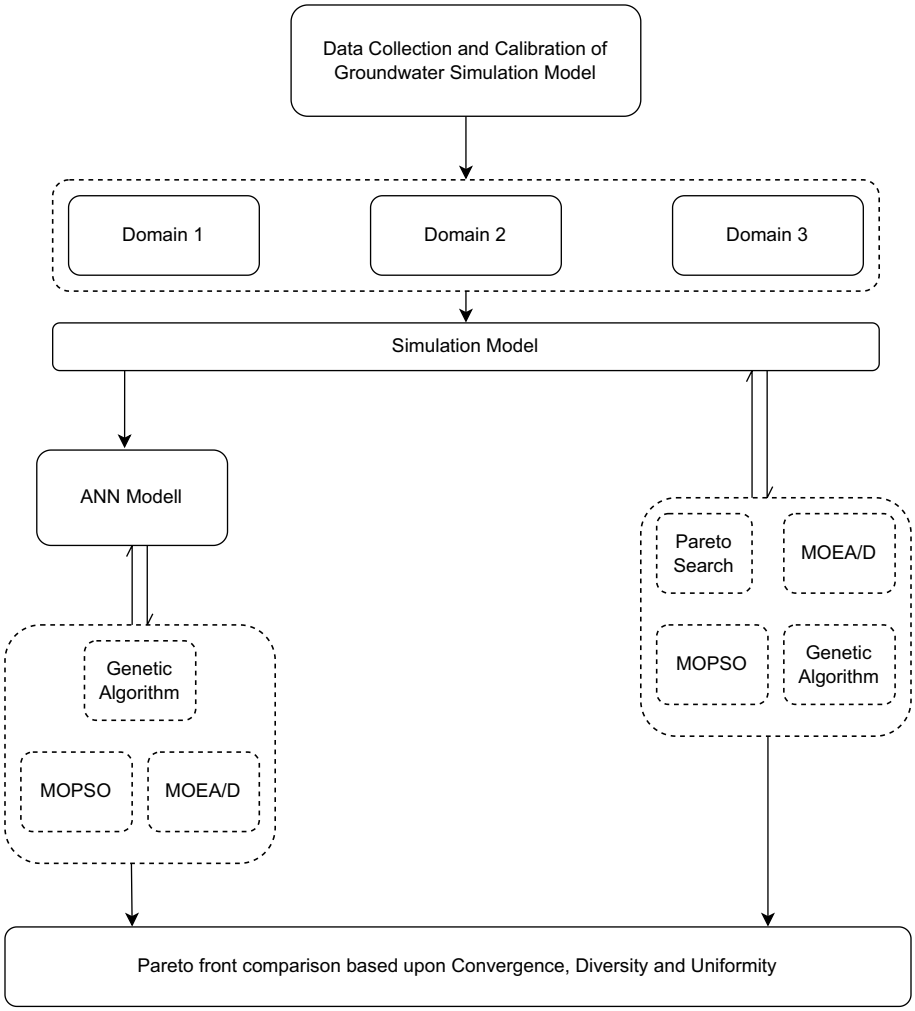

Fig. 2 Flow chart showing the 2 main steps in the optimization process considering (1) different in boundary conditions and (2) with and without speeding up calculation using ANN model

different types of domains that were considered to perform the groundwater modelling. In the Domain-1, eastern & western sides of the model were defined based on the watershed divide line i.e. no-flow boundary respectively. In Domain-2, alluvium plain of the river were considered and the eastern & western side of the model domain was defined on the basis of constant flow boundaries. Domain-3 was chosen as same as the Domain-1 with one modification in the lower eastern part. The Rhône River was introduced in the model domain through the constant head boundary. This modification helped to calculate the impact of incorporating the Rhône River in the modelling area.

Piezometer and River Water Level Data In this study, a total of 15 piezometers were available, where a few piezometers had the data for the period of 2008 to 2010 and the remaining piezometers consisted of the data from 2002 to 2015. The collected data demonstrated that

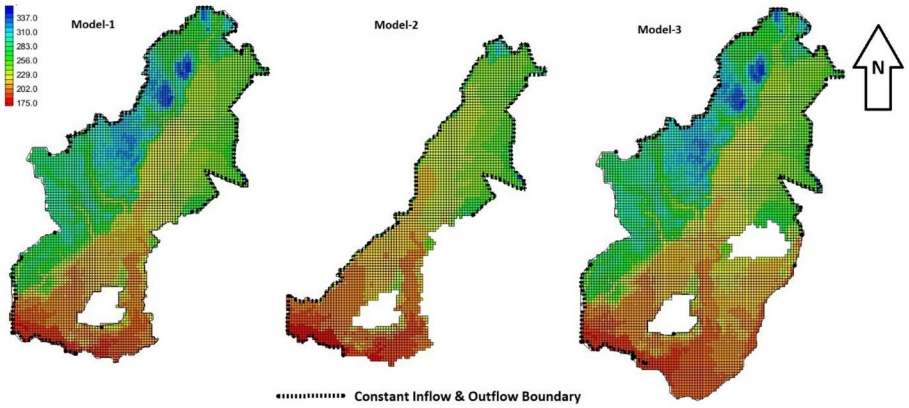

Fig. 3 Different Groundwater Modelling Domains

the piezometer doesn't show very high fluctuation in head values. The head ranged from 3.7 m to 1.2 m, which showed that the groundwater table, and correspondingly groundwater scenario, is stable in the region. The trend analysis of the data showed a fluctuation range of 2.54 m at Chezy and 1.67 m at point de'Ain for the river Ain.

Recharge & Evapotranspiration The rainfall was considered as the source of groundwater recharge. The rainfall data from the Météo France database was used to calculate the recharge input values and applied uniformly over the polygons constructed in the model domain. The different recharge polygons were developed based on a land-use map created from satellite imageries. The initial value of recharge was taken as 10%, 50%, 50%, 80%, and 60% for built-up, agriculture, vegetation, sand, and fellow land respectively. The estimated evapotranspiration was taken as 638 mm/year with an extinction depth of 2 m. It is considered that the potential evapotranspiration was uniformly distributed over the study area. The recharge and evapotranspiration rates were further calibrated.

Water Demand and Supply The water consumption in the study area is mainly done by agriculture (27,000,000 m^3 annual draft). The rest is used for domestic purposes (8,340,700 m^3 annual draft) and some for industrial usage (5,063,204 m^3 annual draft). The field survey was carried out to collect the information of water demand as per the cropping pattern and domestic water consumption. The quantitative data, is obtained from three sources—the Rhône-Méditerranée-Corse Water Agency (AERMC), the Directorate Department of Agriculture and Forestry (DDAF), and the Association Syndicale of Ain Irrigation (ASIA).

Calibration and Validation A regional groundwater flow model was constructed and calibrated to the transient-state condition with a stress period of four months. In calibration of the model, the value of recharge and boundary inflow was taken. Calibration was performed based on the computed and observed values of groundwater head, at 20 m evenly distributed points in the study area. The model was calibrated from 2008 to 2010 based on all piezometers and further from 2010 to 2012 on the basis of remaining wells. Further the model was validated based on data from 2012 to 2015. Initially, the constant flow boundary was calibrated in the steady-state condition to incorporate the groundwater in-flow from

the adjacent aquifer. The recharge rate was calibrated with the help of PEST (parameter estimation tool). The R-A exchanges were calculated with the help of a calibrated model.

3.2 Optimization Problem

The objectives of the optimization model are to maximize the water withdrawal rate from the aquifer (Total discharge, Q) and water gain of the river from the aquifer (Leakage out, L). The pumping rate is assumed constant for the well zones and specified time steps. The water withdrawal rate from the well zones, leakage rate out of the aquifer into the river, of the last time step, are considered for optimization. The drawdown at the wells was considered a constraint. The domain contains different zones of the wells, based on the distance from the river and municipal zone data. To achieve the objective, the discharge in these zones are adjusted by the following optimization algorithm:

$$Maximise\begin{cases} \sum_{i\in R} L_i - P \\ \sum_{i=1}^{n_z} N_i * Q_i - P \end{cases} \tag{1}$$

$$\left(Q_i\right)_{lb} \le Q_i \le \left(Q_i\right)_{ub} \tag{2}$$

where L_i = rate of leakage out of the aquifer to the river; Q_i = rate of discharge of ith well zone; R = set of the river grid cells with leakage out; n_z = the total number of well zones; N_i = number of wells in the i^{th} zone; P = penalty imposed due to drawdown constraint violation; $(Q_i)_{lb}$ = lower bound of the discharge for the i^{th} well zone; $(Q_i)_{ub}$ = upper bound of the discharge for the i^{th} well zone.

The penalty (P) is a static distance-based penalty, which is calculated as:

$$d_{dist} = \sqrt{\sum_{i\in W} \left(d_i - d_{threshold}\right)^2} \tag{3}$$

$$P = C_{\text{model}} * d_{dist} \tag{4}$$

where, d_i = drawdown at i^{th} well; W = set of wells; $d_{threshold}$ = threshold value of drawdown, taken as 2 m; d_{dist} = distance-based metric for drawdown; C_{model} = constant to amplify the penalty. The value of C_{model} is calculated such that the penalty is of the same order as that of the cost. Both the cost has values of order 10^5 and from the multiple (500) random model evaluations, an expected value of d_{dist} is obtained.

3.3 Optimization Techniques

Total four techniques viz. MOGA, MOPSO, PS, and MOEA/D were compared on three groundwater model domain sizes. The following subsections discuss the details of the optimization algorithms.

3.3.1 Multi-Objective Genetic Algorithm

In MATLAB, 'gamultiobj' is the function used for the Multi-objective GA. This function uses a controlled and elitist genetic algorithm (The MathWork Inc.), which is a variant of NSGA-II, (Deb 2001). This algorithm increases the diversity by favouring a variety of individuals even if they have a lower fitness value. After the initialization, generic operations (crossover, mutation, and selection) are iteratively performed. Consequently, crossover fraction, mutation rate, number of generations, and population size are major parameters affecting the performance of the algorithm. For the ANN model, these parameters were set to 0.85, 0.015, 2000, and 200 respectively, and for the simulation model, these were set as 0.85, 0.015, 150, and 25 respectively.

3.3.2 Multi-Objective Particle Swarm Optimization

This heuristic algorithm is inspired by the movement of a bird flock. Each bird or particle has a position and local velocity in the feasible solution domain. The exploration of the domain by these particles is governed by both local and global velocity, which are computed based on the personal best and the global best solution obtained in the domain. Coello et al. (2004), have used the implementation of MOPSO by introducing a mutation operator that enriches the exploration capability of the algorithm. Martínez-Cagigal (2020) provides a MATLAB implementation of Coello et al. (2004), which is used in this work. The c1 and c2 were set to 2 for both ANN and simulation models. Maximum iterations and number of particles, for the ANN model, were set to 1000 and 100 respectively. For the simulation model, these were 150 and 25.

3.3.3 Pareto Search

The Pareto search algorithm by MATLAB, finds the non-dominated solution by the use of Pareto search in a set of points (archives and iterates). The algorithm uses the poll to find better solutions and if better solution is not available, then in the next iteration it multiplies the mesh size by half. Theoretically, the algorithm converges to points near the true Pareto front (The MathWork Inc.). In MATLAB, *paretosearch* function is used for its implementation. No parameter tuning is required for the Pareto search algorithm, though stopping criteria based upon tolerance and time can be adjusted.

3.3.4 Multi-objective Evolutionary Algorithm Based on Decomposition

MOEA/D solves multi-objective problems by decomposing them into multiple scalar sub-problems and solving them simultaneously. The solution of these sub-problems is evaluated based on its neighbouring sub-problems, which makes its computational cost lower in each generation in comparison to NSGA-II (Zhang and Li 2007). The polynomial mutation (order n) is used as the mutation function. Several sub-problems, maximum iterations, percentage of the neighbourhood, and mutation rate are the major parameters for MOEA/D. For the final ANN model, these parameters were set to 1500, 100, 40%, and 1/n respectively. For the simulation model, these were set to 100, 20, 40%, and 1/n respectively.

The use of these algorithms is computationally expensive with the groundwater simulation models, which require a few seconds for a single run and require days to complete the single solution (Tigkas et al. 2016). Therefore, an ANN model is also developed to speed up the evaluation of cost function. The efficiency of the ANN-Optimization model is measured with the performance of the actual simulation–optimization algorithm by performing a feasible number of simulations. The ANN model was trained with a dataset generated by simulation models itself.

3.4 ANN-Optimization Model

The ANN is a widely used technique to map non-linear and black box functions for fast evaluations. The Feed Forward Neural Network (FFNN) is the most commonly used ANN. It consists of multiple parallel layers of memory units (neurons). Each layer of neuron is fully connected with its adjacent layer and the strength of the connection is defined as its weight. The backpropagation technique is used to find out these weights such that error between the actual and predicted values, is minimized.

For this work, nearly 10^4 random data points, with zone-wise discharge as input and leakage out (L with penalty) and total discharge (Q with penalty) as output, were generated using a calibrated groundwater simulation model. The total data is divided into three subsets: training (70%), validation (15%), and testing (15%). Since, the ANN training time increases with an increase in the size of the training dataset and layer size (neurons), this makes the combination of a large training data set and layer size infeasible. Different layer sizes (20, 10, 25) were tested and the best of the ANN model was selected based on R-square and Mean Absolute Percentage Error (MAPE) of Leakage out because the total discharge predictions were accurate. Along with these matrices, the values of Maximum Absolute Error (MXAE) and Root Mean Squared Error (RMSE) are shown in Table 1.

3.5 Performance Metrics

In multi-objective optimization algorithms, the solution set generated approximates the actual Pareto front (Salomon et al. 2018). The approximated fronts can be compared based on convergence, diversity, and uniformity. In this work, six metrics are used, namely, epsilon, spread, generalized spread, generational distance, inverted generational distance, and hypervolume (Table 2). A detailed comparison of various performance metrics can be found in Zitzler et al. (2003).

The unary e-indicator metric represents the smallest distance that an approximate Pareto front must be translated to completely dominate the reference Pareto set (Kollat and Reed 2005). The average Euclidean distance between the reference Pareto set and the approximate Pareto solutions is called the Generational distance (GD). Both, epsilon and generational distance measure the convergence. The smaller the value of epsilon and generational distance, the better is the convergence. The diversity of the Pareto front can be compared based on the distribution of the solution set (uniformity) and its extent. Spread and generalized spread quantify the non-uniformity of approximate Pareto front. The small values of these matrices indicate, a better and more diverse set of approximate Pareto front. The inverted GD is the average distance between each member of the reference Pareto front and the approximate Pareto front. Hypervolume measures the size of the space enclosed by the approximate Pareto front. Both, inverted generational distance and hypervolume are a unary metric that considers convergence and diversity.

Table 1 ANN optimization results

Model 1								
	Leakage out				**Total discharge**			
Layer size	**MXAE**	**RMSE**	**R square**	**MAPE (%)**	**MXAE**	**RMSE**	**R square**	**MAPE (%)**
20	11172.1	1856.1	0.984	1.010	1496.2	323.5	1.000	0.105
10	12573.0	1970.5	0.982	1.140	1622.5	351.1	1.000	0.114
25	12713.2	1872.8	0.984	1.035	1735.5	312.7	1.000	0.095
Model 2								
	Leakage out				**Total discharge**			
Layer size	**MXAE**	**RMSE**	**R square**	**MAPE (%)**	**MXAE**	**RMSE**	**R square**	**MAPE (%)**
20	92676.2	5677.8	0.974	25.963	91579.4	5513.9	0.980	1.601
10	93669.8	5482.7	0.976	27.137	91244.6	5219.5	0.982	1.715
25	94107.8	6822.7	0.963	28.047	94727.1	6663.2	0.971	2.038
Model 3								
	Leakage out				**Total discharge**			
Layer size	**MXAE**	**RMSE**	**R square**	**MAPE (%)**	**MXAE**	**RMSE**	**R square**	**MAPE (%)**
20	4033.8	456.6	0.999	0.456	832.3	220.0	1.000	0.070
10	3676.8	584.6	0.998	0.630	897.0	241.8	1.000	0.080
25	3017.9	435.3	0.999	0.433	940.4	225.4	1.000	0.072

In addition, hypervolume is the most widely used performance metric (Riquelme et al. 2015). Auger et al. (2009) has provided a method to choose the value of reference point for the calculation of hypervolume. Unlike inverted generational distance, the higher the hypervolume better is the combined effect of convergence and diversity of the approximate Pareto front.

Table 2 Evaluation criteria

Metric	**Evaluation criteria**		
	Convergence	Diversity (Total)	Uniformity (Sub component of diversity)
Epsilon	✓	-	-
Spread	-	-	✓
Generalized Spread	-	-	✓
Generational distance	✓	-	-
Inverted generational distance	✓	✓	-
Hypervolume	✓	✓	-

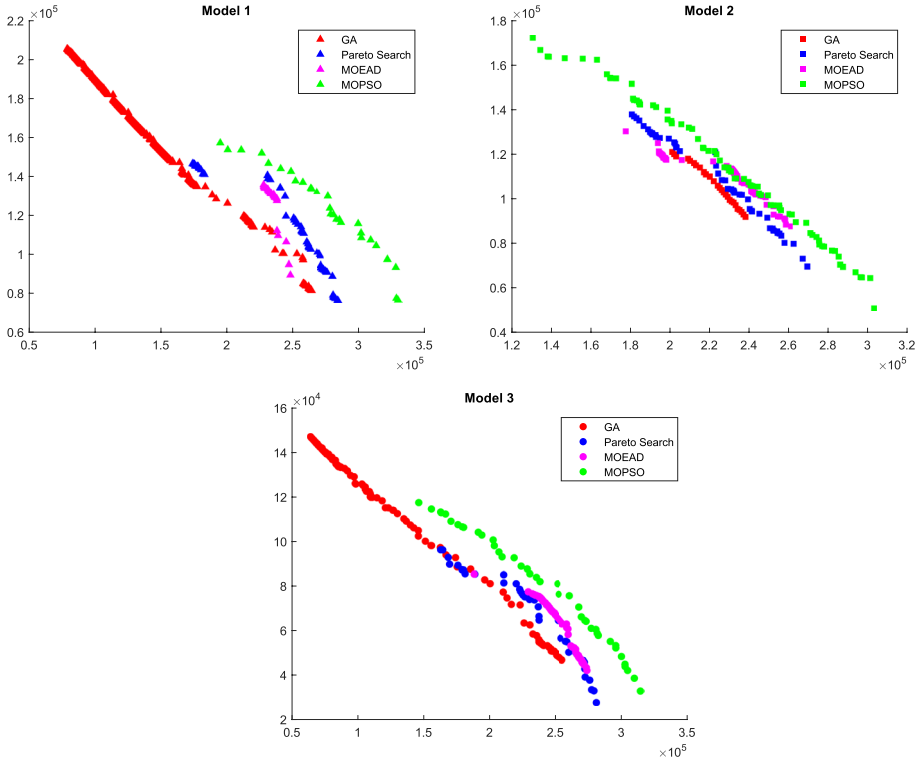

Fig. 4 Pareto front comparison of optimization techniques through Simulation-Optimization model

4 Results

4.1 Simulation–Optimization Model

Each combination of the simulation model and optimization technique was evaluated up to 5000 points depending upon the optimization algorithm and its parameters. In all the three models, the Pareto fronts by MOPSO are more converged, whereas the results obtained from GA are more diverse for model 1 and model 3 (Fig. 4). The Pareto search and MOEA/D are more converged in comparison to GA, but they lack diversity. Only in model 2, the diversity of the Pareto front of MOPSO outperforms that of GA.

To understand the impact of different model domains on optimal outputs, the best-performed MOPSO front was studied. The Pareto fronts show that Domain-1 is giving a high value of leakage out in comparison to the optimal value of groundwater discharge through wells whereas the output of Domain-2 is higher than Domain-3. The Pareto fronts also depict the effect of a model domain on the optimal discharge and river gain relation (Fig. 4). For instance, in the case of MOPSO, an increase of discharge by 50,000 m^3/day (from 200,000 m^3/day to 250,000 m^3/day) leads to a decrease in river gain by 15,820 m^3/day, 37,657 m^3/day, and 19,535 m^3/day for the model domain 1, 2, and 3 respectively. Correspondingly, the Pareto search demonstrates a decrease of river gain by 22,278.9 m^3/day, 40,116 m^3/day, and 20,457.4 m^3/day for the previously mentioned discharge increase.

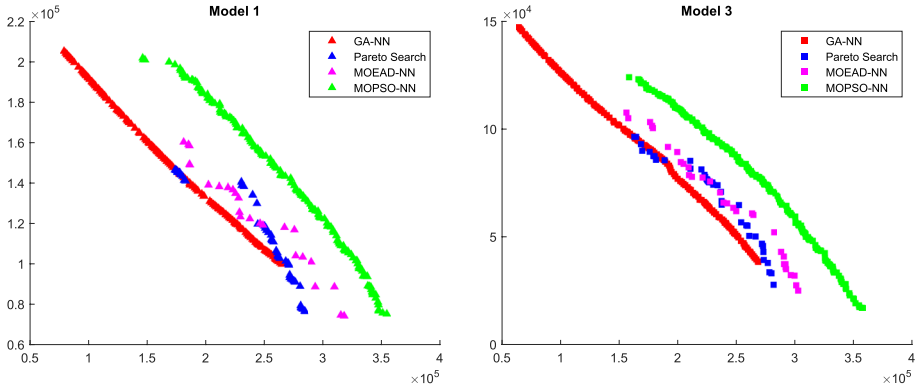

Fig. 5 Pareto front comparison of optimization techniques with ANN models

In model 3, even though the Pareto front of GA is diverse, the solutions are not uniform due to which the spread of GA is high. For model 3, MOPSO is much more uniformly distributed and hence has a low value of the spread.

4.2 ANN-Optimization Model

MOPSO showed significant improvement in terms of convergence, diversity, and uniformity of the solutions (Fig. 5). MOEA/D showed improved diversity and uniformity whereas convergence was still similar to the Pareto search (without ANN). The Pareto search solution without the use of ANN can be taken as the reference for the comparison (Figs. 4 and 5). The diversity and uniformity of the GA and MOPSO are close and comparable. Spread and generalized spread depict that the MOPSO and GA have a uniform distribution of the solutions. Whereas the MOEA/D shows a polar and non-uniform distribution of solutions. The MOPSO is superior to other algorithms, in terms of convergence, suggested by epsilon and generational distance. The combined effect of convergence and diversity, measured with hypervolume and inverted GD, is also best for MOPSO. GA, PS, and MOEA/D show a close value of hypervolume and inverted GD indicating a nearly equivalent performance in terms of convergence and diversity. The percentage increase in hypervolume and decrease in inverted GD, both show that MOPSO in model 1 and MOEA/D in model 3, has significant improvement. Whereas GA and MOEA/D in model 1 show the least improvement. However, the GA and MOPSO (in Model 3) showed conflict in the change in values. Inverted GD is less reliable because it is strongly influenced by the distribution of the approximate Pareto front.

5 Discussions

The results show that solutions of GA are highly diverse in model 1 and model 3, even at a low number of simulations. Other than the optimization algorithm, the Pareto fronts trends are determined by the model demarcation and its boundary conditions. It shows

that adding the Rhône River in the model alters the groundwater flow and correspondingly water budget of the area. The results show that Domain-2 which consists of the shortest area along with the dominated constant boundary inflow is providing less groundwater to the river in comparison to Domain-1 which consists of a larger area in the simulation along with a one-sided watershed boundary. It can be observed that even though all three domains extract the same amount of groundwater, the modelled effect on the river varies for each domain. In addition, the data from the Pareto front suggest that in model 2, the extraction from the well and the river gain are heavily related to each other. This relation will not only influence the simulation results but will also affect the decision-making in groundwater management. In ANN Optimization model, it is evident that in terms of convergence, the solution points of MOPSO dominate every other solution, and solution points of GA are dominated by most of the other points. PS and MOEA/D solutions lie between GA and MOPSO front for both models. MOPSO solutions largely consist of outliers with either too small or too large box bounds. MOEA/D solutions are largely uniform and cover the lower and upper bound of decision variables. Regardless of the small difference, the overall trend is: wells close to rivers extract less groundwater than far away. Another factor, that can contribute to the high discharge of a close well is spatial heterogeneity in the aquifer properties.

6 Conclusions

Three different domain sizes of groundwater model for optimization of withdrawal and gain in the river from the aquifer, by four different optimization algorithms were conferred. The groundwater model development, optimization models, ANN model to reduce evaluation time, performance metrics for Pareto fronts, and the comparison of the distribution of optimal decision variables for different optimization techniques, were discussed. The ANN model for three domains shows the difference in accuracy, suggesting a change in domain size and consequently boundary conditions can alter the performance of ANN. The results show that the boundary conditions and domain size influence the result of simulation–optimization models. Domain-1 was found more efficient and was capable to give a higher value of R-A exchange corresponding to the discharge of pumping wells.

The result of the simulation–optimization model suggests that GA produced a diverse set of Pareto solutions even for a low number of evaluations but the solution of other algorithms was dominating, with MOPSO being the best. Among the three domains, model 2 consistently implied a greater interrelation between the two costs, i.e., total discharge and river gain. In the ANN-Optimization model, MOPSO showed significant improvement in both diversity and convergence. GA solutions did not show major improvement in convergence. This suggested the use of GA to obtain a tentative solution set that can be obtained without the use of the ANN model. The use of ANN can significantly improve the performance of MOPSO by allowing it to have a greater number of evaluations, which was limited in the raw simulation model due to the infeasible computation time.

In case of decision-making problems, this study can be helpful to have a good distribution of decision variables, even though they provide a better solution but not the best one. Apart from the statistical distribution of solutions, the spatial distribution of optimized decision variables (well discharge) provided a lot of information regarding the physical interpretation and conceptual verification of the solutions. These solutions followed an expected trend of higher discharge at faraway wells and lower discharge at the close wells.

From this study, the influence of domain demarcation, boundary conditions, and optimization algorithm has been observed on the Pareto fronts and optimized decision variables. The conceptual verification of the Pareto solutions was also explained. Besides this, the use of ANN and handling the constraints of the optimization problem is also considered for a real-world groundwater problem. Still, several challenges in R-A exchange-related groundwater decision-making exist. Accurate R-A exchange modelling, precise surrogate models to reduce S–O time, and dealing with high heterogeneity of the domain, to name a few.

Author's Contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Mayank Bajpai and Shreyansh Mishra. The first draft of the manuscript was written by Shreyansh Mishra while Dr Shishir gaur finalized the draft. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declarations

Competing Interest The authors have no conflicts of interest to declare that are relevant to the content of the article.

References

Abd-Elmaboud M, Abdelgawad HAA, El-Alfy K, Ezzeldin M (2021) Estimation of groundwater recharge using simulation-optimization model and cascade forward ANN at East Nile Delta aquifer, Egypt. J Hydrol Reg Stud 34. https://doi.org/10.1016/j.ejrh.2021.100784

Anderson MP (2005) Heat as a ground water Tracer. Ground Water 43(6):951–968

Asadzadeh M, Tolson BA, Burn DH (2014) A new selection metric for multiobjective hydrologic model calibration. Water Resour Res 50(9):7082–7099

Atwell BH, MacDonald RB, Bartolucci LA (1971) Thermal mapping of streams from airborne radiometric scanning. Water Resour Bull 7(2):228–243

Audet C, Bigeon J, Cartier D, Le Digabel S, Salomon L (2020) Performance indicators in multiobjective optimization. Eur J Oper Res 292. https://doi.org/10.1016/j.ejor.2020.11.016

Auger A, Bader J, Brockhoff D, Zitzler E (2009) Theory of the hypervolume indicator: optimal μ-distributions and the choice of the reference point. In Proceedings of the tenth ACM SIGEVO Workshop on Foundations of genetic algorithms, 87–102

Avent B, González J, Diethe T, Paleyes A, Balle B (2020) Automatic discovery of privacy-utility pareto fronts. Proc Priv Enh Technol 2020(4):5–23

Bassi M, Cursi ESD, Pagnacco E, Ellaia R (2018) Statistics of the Pareto front in Multi-objective Optimization under Uncertainties. Lat Am J Solids Struct 15:11

Becker MW, Georgian T, Ambrose H, Siniscalchi J, Fredrick K (2004) Estimating flow and flux of groundwater discharge using water temperature and velocity. J Hydrol 296(1–4):221–233

Belakaria S, Deshwal A (2019) Max-value entropy search for multi-objective bayesian optimization. In International Conference on Neural Information Processing Systems (NeurIPS)

Binois M, Ginsbourger D, Roustant O (2015) Quantifying uncertainty on Pareto fronts with Gaussian process conditional simulations. Eur J Oper Res 243(2):386–394

Burkholder BK, Grant GE, Haggerty R, Khangaonkar T, Wamper PJ (2008) Influence of hyporheic flow and geomorphology on the temperature of a large, gravel-bed River, Clackamas River, Oregon, USA. Hydrol Process 22:941–953

Calandra R, Peters J, Deisenrothy MP (2014) Pareto front modelling for sensitivity analysis in multi-objective bayesian optimization. In NIPS Workshop on Bayesian Optimization, vol 5

Cao P, Fan Z, Gao R, Tang J (2017) A manufacturing-oriented single point search hyper-heuristic scheme for multi-objective optimization. In ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers Digital Collection

Coello CAC, Pulido GT, Lechuga MS (2004) Handling multiple objectives with particle swarm optimization. IEEE Trans Evol Comput 8(3):256–279

Conant B, Robinson CE, Hinton MJ, Russell HAJ (2019). A framework for conceptualizing groundwater-surface water interactions and identifying potential impacts on water quality, water quantity, and ecosystems. J Hydrol 574:609–627

Constantz J (1998) Interaction between stream temperature, streamflow, and groundwater exchanges in alpine streams. Water Resour Res 34:1609–1615

Cristea NC, Burges SJ (2009) Use of thermal infrared imagery to complement monitoring and modeling of spatial stream temperature. J Hydrol Eng 14(10):1080–1090

Dangar S, Asoka A, Mishra V (2021) Causes and implications of groundwater depletion in India: A review. J Hydrol. https://doi.org/10.1016/j.jhydrol.2021.126103

Deb K (2001) Multi-objective optimization using evolutionary algorithms. John Wiley & Sons, Ltd, Chichester, England

Emmerich MTM, Deutz AH (2018) A tutorial on multiobjective optimization: fundamentals and evolutionary methods. Nat Comput 17(3):585–609

Etsias G, Katsifarakis KL (2017) Combining pumping flowrate maximization from polluted aquifers with cost minimization. Environ Process 4:991–1012

Gomo M (2011) A groundwater-surface water interaction study of an alluvial channel aquifer. Doctoral dissertation, Institute for Groundwater Studies - University of the Free State, South Africa

Handcock RN, Gillespie AR, Cherkauer KA, Kay JE, Burges SJ, Kampf SK (2006) Accuracy and uncertainty of thermal infrared remote sensing of stream temperatures at multiple spatial scales. Remote Sens Environ 100:427–440

Hebert C, Caissie D, Satish MG, El-Jabi N (2011) Study of stream temperature dynamics and corresponding heat fluxes within Miramichi River catchments (New Brunswick, Canada). Hydrol Process 25:2439–2455

Hernández-Lobato D, Hernandez-Lobato J, Shah A, Adams R (2016) Predictive entropy search for multi-objective bayesian optimization. In International Conference on Machine Learning, pp 1492–1501. PMLR

Höllermann B, Evers M (2019) Coping with uncertainty in water management: Qualitative system analysis as a vehicle to visualize the plurality of practitioners' uncertainty handling routines. J Environ Manage 235:213–223

Horn D, Demircioğlu A, Bischl B et al (2018) A comparative study on large-scale kernelized support vector machines. Adv Data Anal Classif 12:867–883

Jha MK, Peralta RC et al (2020) Simulation-optimization for conjunctive water resources management and optimal crop planning in Kushabhadra-Bhargavi river delta of Eastern India. Int J Environ Res Public Health 17(10):3521

Kalbus E, Reinstorf F, Schirmer M (2006) Estimating flow and flux of groundwater discharge using water temperature and velocity. Hydrol Earth Syst Sci 10:873–887

Kay JE, Kampf SK, Handcock RN, Cherkauer KA, Gillespie AR, Burges SJ (2005) Accuracy of lake and stream temperatures estimated from thermal infrared images. J Am Water Resour Assoc 41:1161–1175

Keery J, Binley A, Crook N, Smith JWN (2007) Temporal and spatial variability of groundwater-surface water fluxes: Development and application of an analytical method using temperature time series. J Hydrol 336:1–16

Kollat JB, Reed PM (2005) The value of online adaptive search: a performance comparison of NSGAII, ε-NSGAII, and εMOEA. International Conference on Evolutionary Multi-Criterion Optimization. Springer, Berlin, Heidelberg

Li Y-F, Ng S, Xie M, Goh T (2010) A systematic comparison of metamodeling techniques for simulation optimization in Decision Support Systems. Appl Comput 10:1257–1273. https://doi.org/10.1016/j.asoc.2009.11.034

Loheide SP, Gorelick SM (2006) Quantifying stream – aquifer interactions through the analysis of remotely sensed thermographic profiles and in situ temperature histories. Environ Sci Technol 40(10):3336–3341

Lowry CS, Walker JF, Hunt RJ, Anderson MP (2007) Identifying spatial variability of groundwater discharge in a wetland stream using a distributed temperature sensor. Water Resour Res 43:W10408

Mao W, Yang J, Zhu Y, Ye M, Jingwei Wu (2017) Loosely coupled SaltMod for simulating groundwater and salt dynamics under well-canal conjunctive irrigation in semi-arid areas. Agric Water Manag 192:209–220

Marjit R, Hopfe CJ (2009) Multi-objective robust optimization algorithms for improving energy consumption and thermal comfort of buildings

Martínez-Cagigal V (2020) Multi-objective particle swarm optimization (MOPSO). MATLAB Central File Exchange 5–14

McDonald MG, Harbaugh AW (1988) A modular three-dimensional finite-difference ground-water flow model, 6. U.S. Geological Survey Techniques of Water-Resources Investigations
Paran F, Arthaud F, Bornette G, Graillot D, Lalot E, Marmonier P, Novel M, Piscart C (2012) Characterization of exchanges between Rhône River and groundwater., Research report ZABR Phase 4, AERMC
Riquelme N, Von Lücken C, Baran B (2015) Performance metrics in multi-objective optimization. Latin American Computing Conference (CLEI). IEEE
Salomon L, Audet C, Bigeon J, Le Digabel S (2018) Review of the quality of approximated Pareto fronts in multiobjective optimization. JOpt
Sophocleous M (2002) Interactions between groundwater and surface water: the state of the science. Hydrogeol J 10:52–67
Tigkas D, Christelis V, Tsakiris G (2016) Comparative study of evolutionary algorithms for the automatic calibration of the Medbasin-D conceptual hydrological model. Environ Process 3:629–644
Wawrzyniak V, Piégay H, Poirel A (2012) Longitudinal and temporal thermal patterns of the French Rhône River using Landsat ETM+ thermal infrared images. Aquat Sci 74(3):405–414
Westhoff MC, Savenije HHG, Luxemburg WMJ, Stelling GS, van de Giesen NC, Selker JS, Pfister L, Uhlenbrook S (2007) A distributed stream temperature model using high resolution temperature observations. Hydrol Earth Syst Sci 11:1469–1480
Zhang Q, Li H (2007) MOEA/D: A multiobjective evolutionary algorithm based on decomposition. IEEE Trans Evol Comput 11(6):712–731
Zitzler E, Thiele L, Laumanns M, Fonseca CM, da Fonseca VG (2003) Performance assessment of multiobjective optimizers: An analysis and review. IEEE Trans Evol Comput 7(2):117–132