Chapter 4: OPTIMAL GROUNDWATER MANAGEMENT STRATEGIES FOR IMPROVED R-A EXCHANGES

4.1 Introduction

This study systematically compared different Pareto fronts generated by solving multi-objective optimization problems focused on maximizing the R-A exchanges and GW extraction within a conjunctive surface water-groundwater system. A simulation model for the Ain River was developed to calculate GW flux and R-A exchanges. Additionally, metaheuristic evolutionary optimization techniques were employed for multi-objective optimization. The resulting Pareto fronts were quantitatively and qualitatively compared to determine the best algorithm for balanced trade-off and accuracy. The set of optimal solutions was finally interpreted within the context of groundwater management, emphasizing the challenges inherent in selecting the most appropriate algorithm and model domain boundary conditions. The comparison of optimal results plays a crucial role in devising physically meaningful management strategies that enhance our understanding of R-A exchanges and promote more efficient river health-inclusive management practices.

This chapter delves into the results of the S-O approach, which was employed to identify the most effective algorithm for simultaneously optimizing GW extraction and R-A exchanges. The Pareto fronts generated by various algorithms are compared, and the best-performing algorithm is determined based on multiple criteria. The results are instrumental in developing strategies that ensure sustainable utilization of groundwater while maintaining the ecological integrity of interconnected surface water systems.

In line with the European Water Framework Directive (2000), which treats water resources as a heritage site, recognizing the intrinsic value of aquifers is imperative for preserving these vital resources. These management frameworks must balance the demands for development with the aquifer systems' ecological, environmental, and hydrological integrity, reflecting the essence of sustainable groundwater exploitation (Lucca et al., 2023). The GW over-extraction disrupts the equilibrium between inflows and outflows, posing severe risks to groundwater and surface water systems. To this end, (Fleckenstein et al., 2006) emphasized the necessity of understanding and quantifying R-A exchanges to optimize water resource conservation. These interactions are integral to numerous contemporary issues, such as ensuring sustainable drinking water supplies, managing environmental flow regimes, safeguarding riverine ecosystems, and mitigating the effects of pollutants.

In multi-objective optimization problems, a comprehensive comparative analysis of Pareto fronts—each representing a spectrum of trade-off solutions—is essential for understanding the performance and behavior of optimization models. These Pareto fronts, generated through S-O models under diverse management scenarios, offer valuable insights into the trade-offs between conflicting objectives. By evaluating and comparing these fronts, it becomes possible to identify patterns, similarities, and variations in the optimal solutions and assess the uncertainties associated with decision-making. Such analyses are pivotal for informed water resource management, as they help pinpoint robust solutions that balance competing priorities effectively.

4.2 Methods

The proposed methodology integrates a comprehensive S-O framework to address groundwater-stream management challenges. The process begins with identifying the GW-stream management problem, wherein maximum discharge and R-A exchange capacities are determined. Subsequently, transient groundwater pumping patterns are defined within the MODFLOW model domains, which are used to simulate groundwater head values and corresponding R-A exchanges. These exchanges are recorded in the LARB as input data for the optimization algorithm. The S-O model operates through an iterative cycle, where a validated groundwater model is coupled with a surrogate model to enhance computational efficiency. The optimization algorithm generates an initial random population using MOPSO, NSGA-II, MOEA/D, and Pareto search techniques. The fitness function evaluates each solution based on its ability to achieve feasible R-A exchanges and groundwater extraction targets. Optimal solutions are then identified and clustered to form a reduced Pareto front, representing a diverse set of trade-off solutions.

After optimization, a comparative analysis of the Pareto fronts is conducted to evaluate the model's sensitivity to different management scenarios. A reduced pareto fronts are obtained using DBSCAN clustering to generate five solutions from a larger set of solutions. The five solutions represent scenarios facilitating decision-making and analysis for sustainable water resource management. The methodology is explained in Figure 4.1.

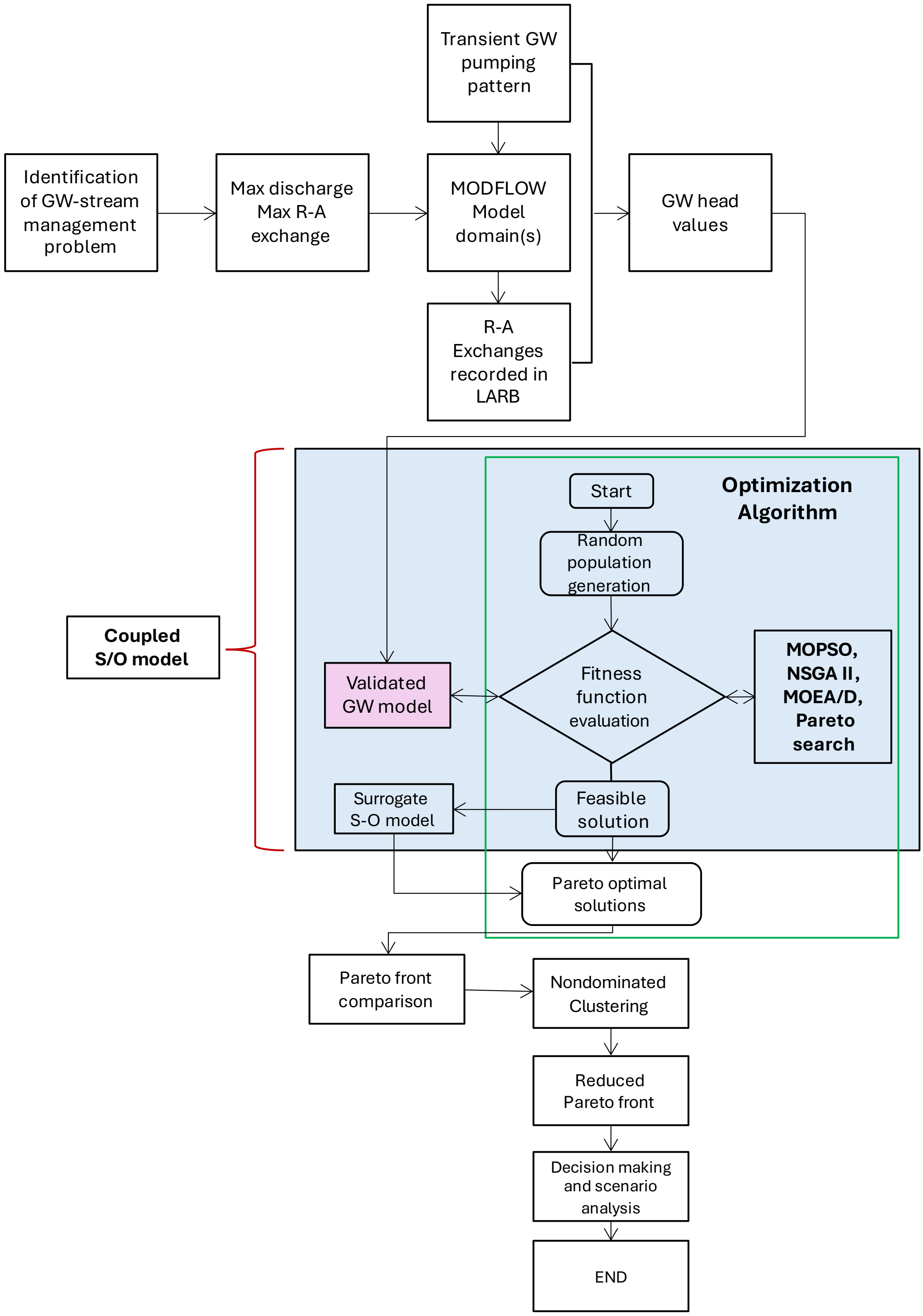


Figure 4.1 Methodology chart of the S-O framework

A regional groundwater flow model was subsequently developed and calibrated to transient-state conditions, using a four-month stress period to capture the system's temporal variability. The calibration process involved adjusting the recharge and boundary inflow values, with calibration targets being the observed groundwater head values at twenty evenly distributed points within the study area. Initial calibration, conducted from 2008 to 2010, included all available piezometers, while subsequent calibration from 2010 to 2012 utilized data from the four wells. Model validation, performed using data from 2012 to 2015, demonstrated the robustness of the model in replicating observed hydrological conditions.

A critical component of the calibration process was refining boundary conditions to accurately represent groundwater inflows from adjacent aquifers. The PEST tool optimized the recharge rate, allowing for precise calibration of the model to observed conditions. Finally, the calibrated model was employed to quantify river-aquifer exchanges, providing valuable insights into the region's complex interactions governing groundwater-surface water dynamics. This calibrated and validated model is a powerful tool for developing sustainable groundwater management strategies, especially in areas where R-A exchanges are critical to the hydrological cycle.

4.2.1 Linked S-O-based management model

The decision variables within the simulators (MODFLOW) represent the pumping rates at each well. The objective function measures the performance of the different pumping solutions while satisfying the set constraints. A seven-year management horizon was chosen for this evaluation, considering the number of stress periods.

To deal with many wells in the domain, they were grouped based on the municipal zones (communes) and the distance of the well centroid from the river. If any well in the municipal zone had a distance from the river greater than a threshold of 1 km, it was classified as a new well zone. 319 wells were reduced to 31 decision variables demarcated by the administrative boundary and distance from the river. These decision variables correspond to pumping rates from the centroid of the well zone equivalent to the sum of discharges of all pumping wells in that well zone for the management period (23 time steps). As discussed in the previous chapter, the management model was formulated by coupling the simulators to four optimization algorithms: NSGA-II, MOPSO, Pareto Search, and MOEA/D.

The Pareto front developed by different coupled simulation-optimization models were compared. Optimized costs were evaluated and compared using convergence, diversity, and uniformity criteria and metrics like hypervolume, spread, and Inverted Generational Distance IGD. The decision variables associated with these Pareto fronts were further analyzed, focusing on their spatial and quantitative distribution across different model domains.

Python scripts were developed to implement the coupling between the two systems and decision-making, which carried out the tasks shown in Figure 4.2.

Initialize the Parameters:

Set up initial decision variables.
Define boundary conditions and input

Extract and Post-process the Simulation Output:

Read the output file generated by MODFLOW.
Calculate the Cost for the Optimization Model.
Read and Write Decision Variables:

Modify the '.h5' input file for MODFLOW:

Read the current set of decision variables.
Write the updated decision variables to the '.h5' input file based on initial cost.

Run the Optimization Model:

Evaluate the MODFLOW model with new decision variables as part of the optimization loop.

Check for Convergence:

If the optimization criteria are met (e.g., cost below a threshold, maximum iterations reached), exit the loop.

- If not, return to Step 2 with the updated decision variables.

Figure 4.2 Pseudo code for S-O model

4.2.1.1 Objective functions

The primary objective of the optimization model was to maximize the water withdrawal rate from the aquifer, represented by the total discharge (Q), while enhancing the water contribution from the aquifer to the river, indicated by the leakage outflow (L), while minimizing drawdown at the wells within the model domain. Here, the leakage outflow (L) quantifies explicitly the rate at which water discharges from the aquifer into the river, which is critical for maintaining river-aquifer interactions. On the one hand, increasing the withdrawal rate from the aquifer (total discharge, Q) lowers groundwater levels, leading to a drawdown effect around the wells. This reduction in groundwater levels can diminish the hydraulic gradient that drives water from the aquifer into the river, potentially reducing or reversing the leakage out (L) into the river. On the other hand, maintaining or increasing the leakage rate into the river requires sustaining higher groundwater levels, which often necessitates limiting the water withdrawal rate. Thus, while one objective seeks to maximize the utility of the groundwater resource, the other aims to preserve the ecological function of the river system, making these objectives inherently conflicting. The optimization process involves adjusting the pumping rate for each stress period for the well zones. These rates are expressed as single values representing various management scenarios within designated well zones over specified time steps. For optimization purposes, the key parameters considered include the water withdrawal rate from the well zones, the leakage rate from the aquifer into the river, and the drawdown at the wells during the final time step. The watershed is divided into different well zones, defined based on the proximity to the river and data from municipal zones. This zonal differentiation allows for a more precise adjustment of discharge rates by each zone's specific characteristics and requirements.

$$\text{Maximise} \begin{cases} \sum_{i \in R} L_i - P \\ \sum_{i=1}^{n_z} N_i * Q_i - P \end{cases} \tag{4.1}$$

Subject to:

$$Q_i \leq (Q_i)_{demand} \tag{4.2}$$

$$(Q_i)_{lb} \leq Q_i \leq (Q_i)_{ub} \tag{4.3}$$

Penalty is introduced as:

$$P = C_{model} \times d_{distance} \tag{4.4}$$

$$the\ d_{dist} = \sqrt{\sum_{i \in W} (d_i - d_{threshold})^2} \tag{4.5}$$

where, L_i = rate of leakage out of the aquifer to the river; Q_i = rate of discharge of i^{th} well zone; R = set of the river grid cells with leakage out; n_z = 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. d_i represents the drawdown at the i^{th} well, W is the set of well zones under consideration, and $d_{threshold}$ is the acceptable threshold for drawdown, set at 2 meters. In this study, a maximum drawdown limit of 2 meters was imposed. It was derived based on the hydrogeological constraint which ensures that the aquifer remains within a safe operational range, preventing potential adverse effects such as aquifer compaction and reduced recharge capacity. Multiple groundwater model simulations determined the limit by considering the aquifer's transmissivity, storage capacity, and natural recharge rates. The term d_{dist} serves as a distance-based metric for the drawdown, measuring the deviation of the actual drawdown from the threshold across all wells. The penalty P is then calculated by multiplying this metric by a constant C_{model} chosen to ensure that the penalty is of the same order of magnitude as the optimization cost. It is determined by evaluating the expected value of ddist through multiple (500) random model iterations, then setting C_{model} to the ratio of 1×10^5 (the order of the cost) to this expected value, d_{dist}. This approach ensures the penalty effectively discourages excessive drawdown, aligning the optimization process with the desired drawdown constraints.

The imposed constraints, represented as inequalities, are crucial in ensuring sustainable groundwater management by defining the permissible range of pumping rates from the 31 well zones. These constraints establish upper and lower bounds for the pumping rates, which prevent overexploitation of the aquifer while still meeting the water demand requirements. Specifically, the upper and lower bounds for the pumping rates were set at 1500 m^3/day and 500 m^3/day, respectively. These values were determined based on the estimated yield potential of the lower Ain River basin. The upper bound of 1500 m^3/day is a theoretical maximum limit, restricting the search space within the optimization framework. It is important to note that these bounds are not prescriptive but serve as guidelines for the optimization algorithm.

The final optimal solutions are determined based on the trade-offs between the competing objective functions while adhering to the set constraints. Consequently, the actual pumping rates in the optimal solutions may fall well below the upper bound of 1500 m^3/day, depending on the balance between groundwater sustainability and demand fulfillment. A dynamic soft distance-based penalty was applied to increase the number of feasible solutions in the search space. In GW optimization, penalty handling techniques are crucial for managing constraints and ensuring feasible solutions. Common approaches include static penalty functions, dynamic weight-based penalty functions, and adaptive penalty methods. Static penalty functions impose a fixed penalty on constraint violations, which can lead to suboptimal performance as the balance between objective optimization and constraint satisfaction may not be well-adjusted throughout the optimization process. On the other hand, dynamic weight-based penalty functions adjust the penalty weights dynamically based on the evolution of the optimization process. This allows for a more flexible approach where the penalties can be reduced as the solution approaches feasibility, leading to a more balanced and effective optimization. Adaptive penalty methods modify the penalties based on the performance and progress of the algorithm, aiming to improve convergence to feasible and optimal solutions.

4.2.2 Pareto Performance Indicators

The solution set in multi-objective optimization algorithms approximates the actual Pareto front. The approximated fronts are compared based on convergence, diversity, and uniformity. Various performance metrics that account for one, two, or all three aspects of the Pareto fronts are available. These metrics can be unary or binary. This work uses six metrics: epsilon, spread, generalized spread, generational distance, inverted generational distance, and hypervolume, as presented in Table 4.1. The matrices require a reference point or reference Pareto front. The non-dominated set from the union of the points obtained from the compared Pareto front is used as the reference Pareto front, and a value lower than that of the anti-utopia point of this reference front is taken as a reference point ($1e^4$, $5e^4$). A detailed discussion of the indicators is given in the following paragraphs.

Unary Epsilon-Indicator Metric quantifies the smallest distance by which an approximate Pareto front must be translated to entirely dominate the reference Pareto set. Mathematically, if P represents the reference Pareto front and P^* the approximate Pareto front, the unary epsilon-indicator is defined as:

$$\epsilon = min\{\nexists \in \boldsymbol{R} \mid \forall \boldsymbol{x} * \in \boldsymbol{P}^*, \exists \boldsymbol{x} \in \boldsymbol{P} \; \boldsymbol{such\ that}\ \boldsymbol{f_i}(\boldsymbol{x}^*) \leq \boldsymbol{\epsilon}' \cdot \boldsymbol{fi}(\boldsymbol{x}), \boldsymbol{i} = \boldsymbol{1}, \boldsymbol{2}\} \tag{4.6}$$

In other words, the unary ε-indicator metric represents the minimum distance that an approximate Pareto front must be moved up or down so that all the solutions of approximate pareto sets completely dominate the reference Pareto set. A lower value of ε indicates better convergence of the approximate Pareto front to the reference set.

Table 4.1 Pareto front performance indicators

Metric	**Evaluation criteria**		
	Convergence	Diversity (Total)	Uniformity (Subcomponent of diversity)
Epsilon	✓	-	-
Spread	-	-	✓
Generalized Spread	-	-	✓
Generational Distance	✓	-	-
Inverted Generational Distance	✓	✓	-
Hypervolume	✓	✓	-

Generational distance (GD) represents the average distance between all the optimal solutions of the reference Pareto set and the approximate Pareto solutions. This metric measures the average Euclidean distance between the solutions on the approximate Pareto front and the reference Pareto front. It is calculated as:

$$GD = \frac{1}{|\boldsymbol{P}^*|}\sqrt{\sum_{\boldsymbol{x}*\in\boldsymbol{P}^*} min_{\boldsymbol{x}\in\boldsymbol{P}}\left(\sum\nolimits_{i=1}^{2}(f_i(x) - f_i(x^*))^2\right)} \tag{4.7}$$

Both epsilon and generational distance measure the convergence. The lower the epsilon value and generational distance, the better the convergence.

Spread (S) and Generalized Spread (GS) represent the distribution of the solution set (uniformity) and its extent. It quantifies the non-uniformity of the approximate Pareto front. The small value of these matrices indicates a better and more diverse set of approximate Pareto fronts.

Inverted GD (IGD) is the average distance from each solution in the reference Pareto front to the nearest solution in the approximate Pareto front. The distribution of the approximate Pareto front highly influences it.

Hypervolume (HV) measures the volume of the space dominated by the approximate Pareto front relative to a reference point.

$$HV = Vol\left(\cup_{\boldsymbol{x}*\in\boldsymbol{P}^*}[f_1(x^*), z1)] \times [f_2(x^*), z2]\right) \quad (4.8)$$

where $z1$ and $z2$ are the reference point coordinates, a higher hypervolume value indicates better overall performance in terms of convergence and diversity.

Both IGD and HV are a unique metric that considers convergence and diversity. In addition, HV is the most widely used performance metric. (Auger et al., 2009) has provided a method to choose the value of the reference point for calculating HV.

In summary, metrics such as ϵ and GD focus primarily on convergence, where lower values indicate a closer approximation to the true Pareto front. Spread and generalized spread (S and GS) emphasize the diversity of solutions, with lower values indicating a more uniform and extensive solution set. IGD and HV combine convergence and diversity, with lower IGD and higher HV indicating a superior approximate Pareto front. Among these, hypervolume is particularly noted for its wide use in assessing the performance of Pareto fronts, as it captures both convergence and diversity effectively.

4.2.3 Optimal Management Scenarios and multi-criteria decision making

In the decision-making framework, visualizing trade-offs and understanding the interrelationships between the objective functions is necessary for the policymakers to comprehend the results. The clustering was performed using the DBSCAN algorithm, which groups data points based on density rather than centroids-based, as in the k-means approach (Frey and Dueck, 2007). Below is a concise outline of the process for reducing the Pareto-optimal solution set using DBSCAN.

Step 1: Derive the Pareto-optimal solution by addressing the multi-objective optimization problem.

Step 2: Utilize the DBSCAN clustering algorithm to categorize the optimal solutions along the Pareto front according to density. DBSCAN clusters solutions that are closely packed in space and marks those that do not fit the density criteria as noise.

Step 3: Once the DBSCAN algorithm is executed, identify each cluster's core points and representative solutions.

Step 4: Select the solution within each cluster that best represents the density-based grouping, reducing the number of solutions to a more manageable set.

Step 5: The decision-maker receives a condensed selection of representative solutions from the initial Pareto-optimal set. A conclusive solution can subsequently be selected according to preferences to attain the optimization goal.

The DBSCAN clustering implementation was done in the Python environment. In this algorithm the optimal number of clusters is not predetermined but emerges from the algorithm's ability to detect regions of high point density. The key parameters influencing clustering are the neighborhood radius (ε) and the minimum number of points (minPts) required to form a dense region. These parameters were selected through a process of trial and error. A common approach to determine an appropriate ε value is to analyze the k-distance graph, which plots the distance to the k-nearest neighbor (often k = minPts). A sharp change in the slope of this plot—referred to as the "elbow"—can guide the choice of ε, ensuring clusters are formed around regions with sufficient point density while minimizing noise. Unlike other clustering techniques, DBSCAN does not rely on centroids. Instead, it groups solutions based on density, and the representative solution is typically chosen from among the "core points," which are points that have enough neighboring points within the radius ε. These core points are the most densely connected and, thus, best represent the cluster. By selecting a representative solution from each cluster, the decision-maker is provided with a reduced set of solutions that encapsulate the key characteristics of the entire Pareto front, simplifying decision-making without sacrificing solution quality. This approach avoids predefining the number of clusters, making DBSCAN more suitable for non-linear and irregularly shaped solution spaces.

4.3 Results

4.3.1 Calibration and Validation Results

The computed (orange line in Figure 4.3) and observed (black line in Figure 4.3) groundwater heads follow similar trends for all three domains, suggesting a good agreement between the model outputs and the actual field data. There are minor deviations between the computed and observed values, but overall, the alignment is consistent across all wells, indicating that the models have been well-calibrated. This level of agreement supports the accuracy of the models in simulating groundwater dynamics for the respective domains, ensuring their reliability for future predictions or decision-making related to groundwater management. The calibration results are already explained in Chapter 3.

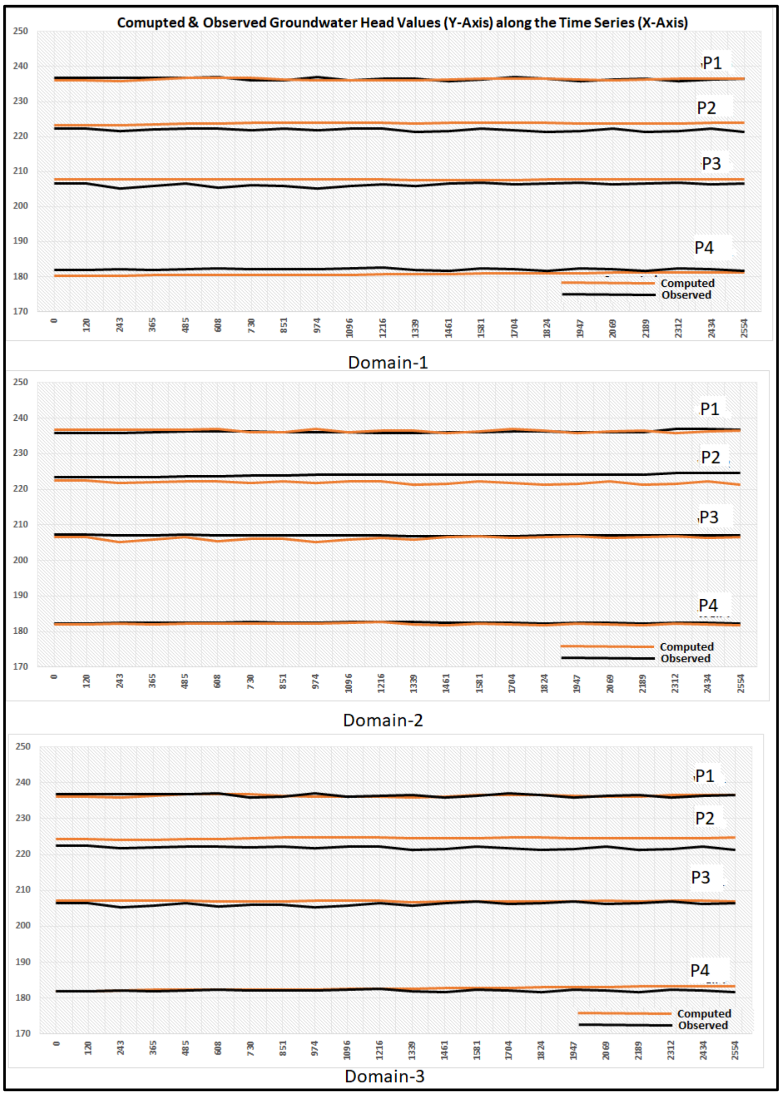


Figure 4.3 Calibration results for 4 observation wells.

4.3.2 Pareto Optimal Trade-offs

Each combination of the simulation model and optimization technique was evaluated up to 5000 points depending upon the optimization algorithm and its parameters. MOPSO and NSGA-II gave the largest number of feasible solutions. Non-dominated sorting was done to identify the pareto solutions, as shown in Figure 4.4. In all three models, the Pareto fronts by MOPSO were more converged, whereas the results obtained from NSGA-II were more diverse for model 1 and model 3. The Pareto front analysis for the three models reveals distinct optimization performance across the techniques. For *Model 1,* NSGA-II achieves the highest f_1 values, reaching approximately 2.2 $\times 10^5$ with a corresponding f_2 around 1.0

$\times 10^5$ However, MOEAD offers a better balance between the objectives, with f_1 ranging from 1.5 $\times 10^5$ to 2.5 $\times 10^5$ and f_2 values 1.0 $\times 10^5$ and 1.5 $\times 10^5$. In *Model 2*, the Pareto fronts shift, with MOEAD maintaining a solid performance, reaching an f_1 maximum of approximately 2.8 $\times 10^5$ and f_2 values above 1.2 $\times 10^5$. The MOPSO technique, though more dispersed, reaches up to $f_2 = 1.4 \times 10^{5,}$ suggesting it explores solutions favoring higher river-aquifer exchanges. Finally, in *Model 3*, the NSGA-II again dominates the upper f_1 range, reaching 1.6 $\times 10^4$ with corresponding f_2 values around 4 $\times 10^3$. However, MOPSO covers most of the Pareto front, providing solutions that balance f_1 and f_2 more evenly. This analysis shows that while NSGA-II explores a broader solution space, MOEAD and MOPSO offer more balanced and potentially optimal trade-offs, particularly in complex scenarios where both objectives are crucial. Therefore, the choice of optimization technique must consider not just the range of solutions but also the specific trade-offs and practical constraints relevant to the modeling scenario.

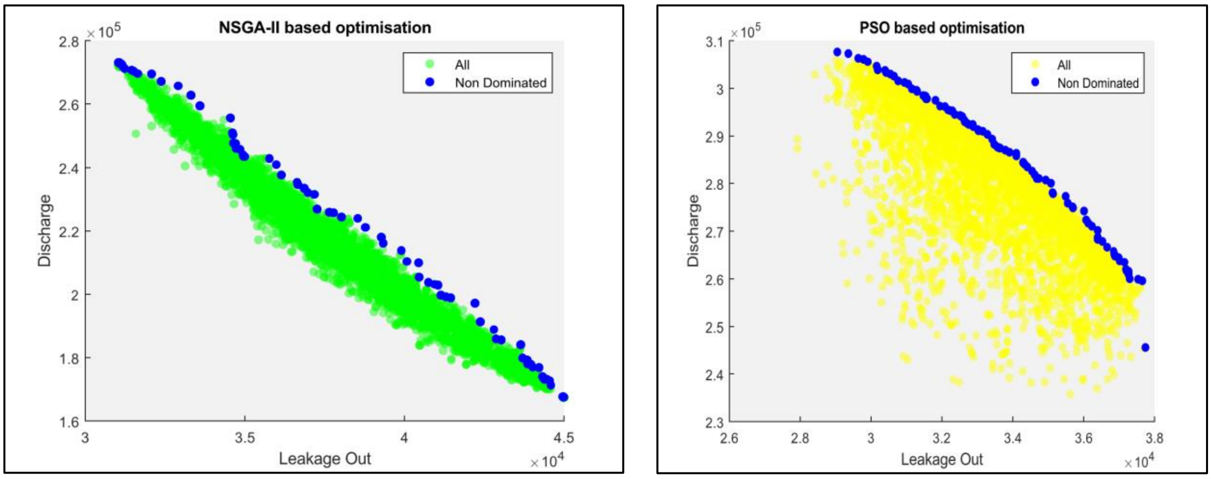
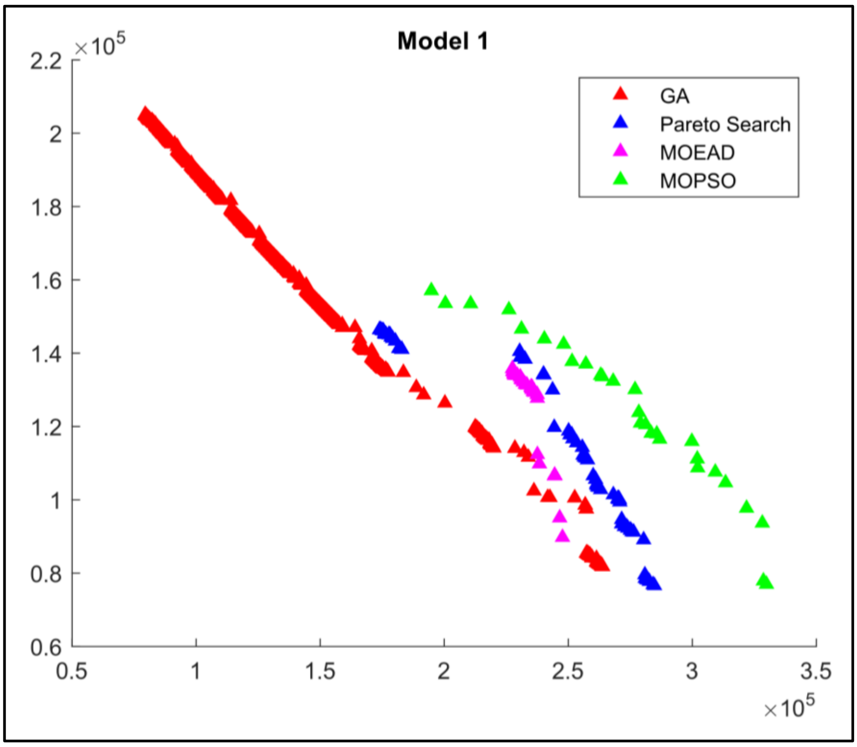
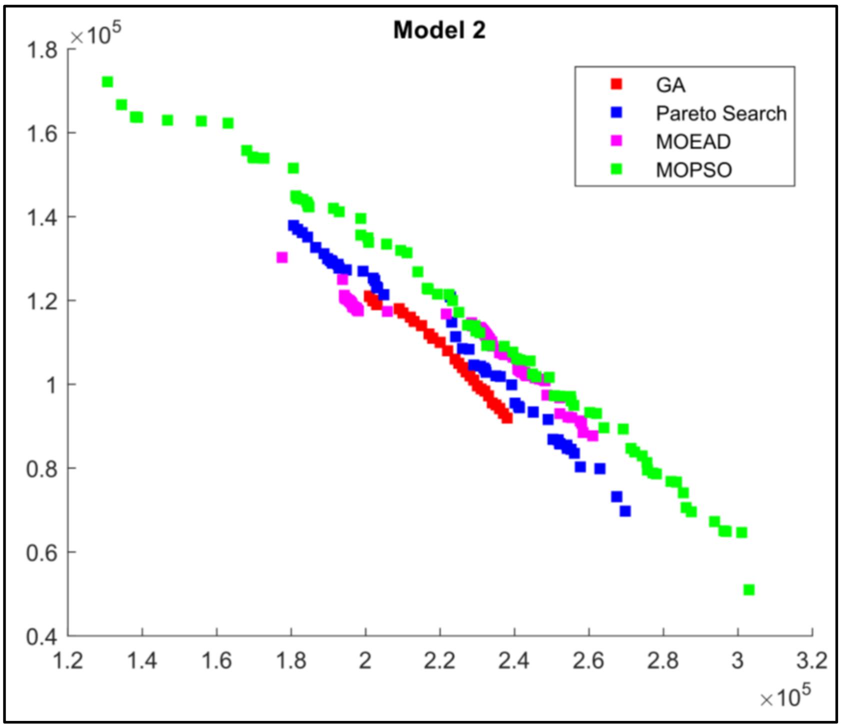


Figure 4.4 All feasible and nondominated solutions obtained by NSGA-II and MOPSO.

To understand the impact of different model domains on optimal outputs, the best-performed MOPSO front was studied. Pareto fronts depicted in Figure 4.5 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 ex. for discharge value of 2000000 m^3/day leakage out in Domain-1, Domain-2, and Domain-3 are as follows 153560 m^3/day, 134940 m^3/day and 100800 m^3/day. It shows that adding the Rhone River to the model alters the groundwater flow correspondingly to water budgeting of the area and gives the most negligible groundwater flow into the river due to excess pumping. Results show that Model-2, consisting of the shortest area along with the dominated constant boundary inflow, provides less groundwater to the river compared to Domain-1, which consists of a larger area in the simulation and a one-sided watershed boundary. The Pareto fronts also depict the Effect of the model domain on the optimal discharge and river gain relation. 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 in 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. It can be observed that even though all three domains extract the same amount of groundwater, the modeled. The 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. This relation will not only influence the simulation results but also the decision-making in groundwater management.


Model 1
GA
Pareto Search
MOEAD
MOPSO
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.5
1
1.5
2
2.5
3
3.5
×10^5
×10^5


Model 2
GA
Pareto Search
MOEAD
MOPSO
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
3.2
×10^5
×10^5

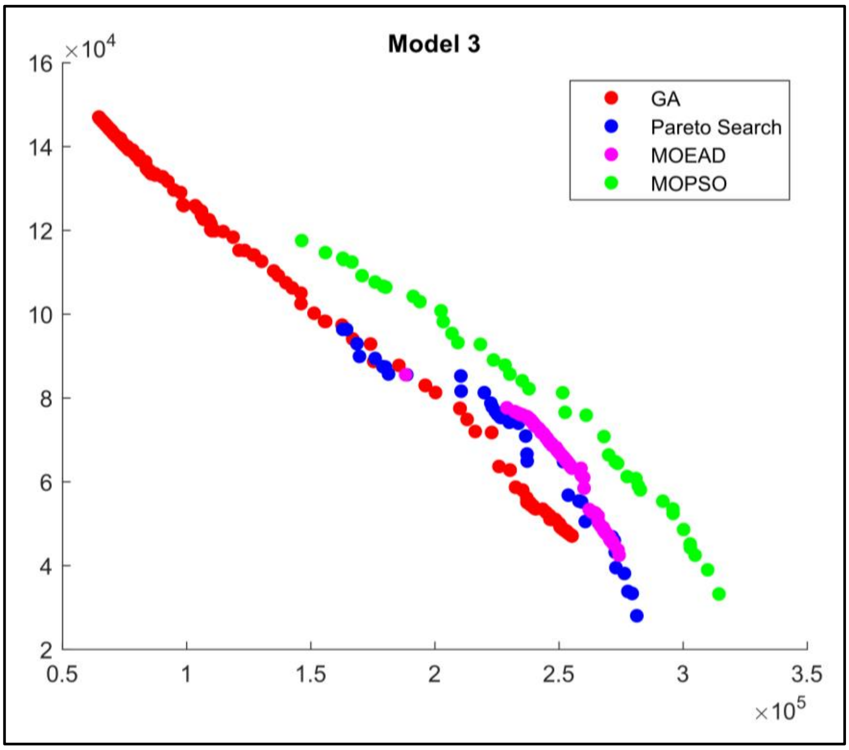
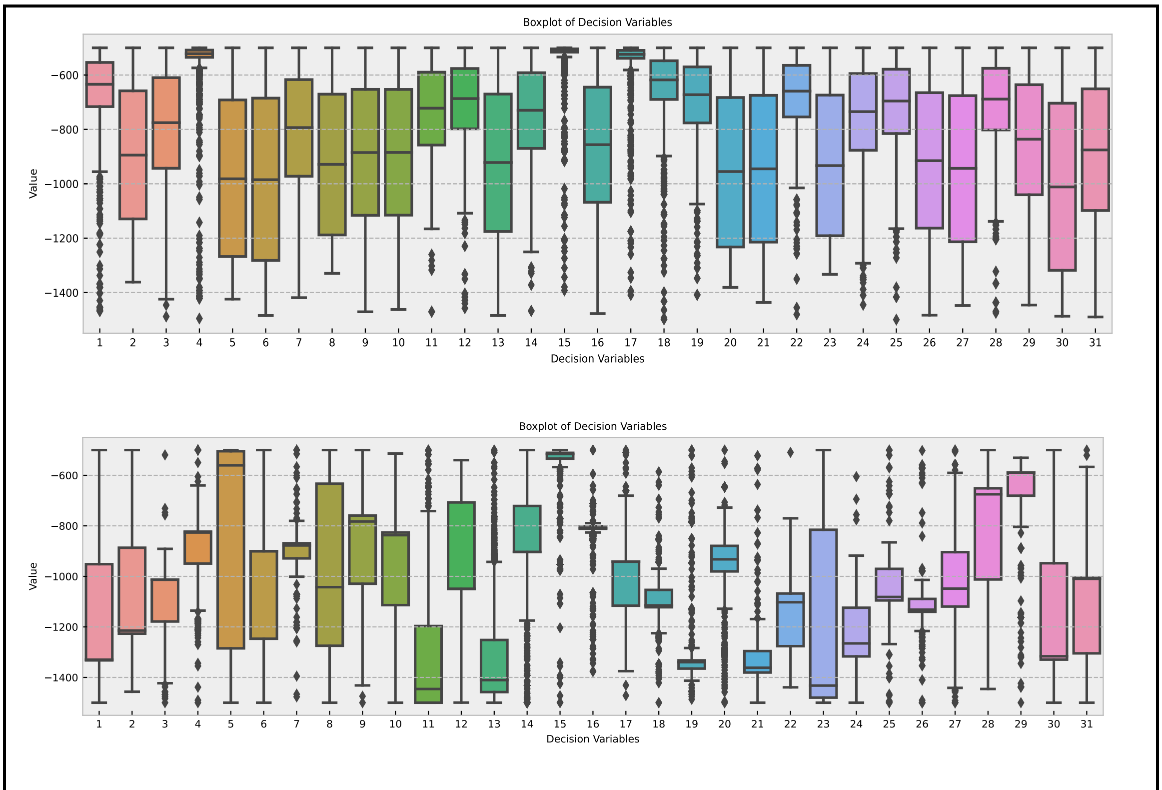


Figure 4.5 Pareto fronts for all three models. The colors show different algorithms. The x-axis has maximum discharge, while the y-axis values are maximum R-A exchanges.

Figure 4.6 shows the box-whisker plot of the decision variables from the Pareto solutions obtained from each optimization technique and two (1 and 3) model domain sizes. It can be observed that the distribution of decision variables for the solutions is dependent on the optimization algorithm rather than the model domain size. Further, in NSGA-II, the solution decision variables have fewer outliers compared to well zone distribution. A similar trend is shown in MOPSO and Pareto searches, with more outliers. MOEA/D shows a broader distribution with the box length (25 and 75 percentiles bound) greater than other optimization techniques. MOPSO solutions have many outliers, depicting many solutions focused on the quartile range. The mean discharge for NSGA-II solutions is less than or close to 1000 for nearly all the zones. In MOEA/D, the mean values are close to the bounds of the box (25 or 75 percentile) of the solutions, and the mean discharge largely varies with the decision variable number (i.e., location of the well zone). MOPSO solutions primarily consist of outliers with either too-small or too-large box bounds. MOEA/D solutions are mainly uniform and cover decision variables' lower and upper bounds.

In total, there are 31 well zones in both domains and half of them are classified as “far away” or “close” depending upon their shortest distance from the river. The average optimal discharge of faraway wells is slightly higher than those near the river.

In Model 3, the value means optimal discharge of faraway wells is 860.8 m^3/day, and that of wells close to the river is 702.9 m^3/day. However, in Model 1, the gap between the two values narrows further, with a faraway well average discharge of 975.4 m^3/day and a close well average discharge of 901.2 m^3/day. Regardless of the slight difference, the overall trend is that 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.



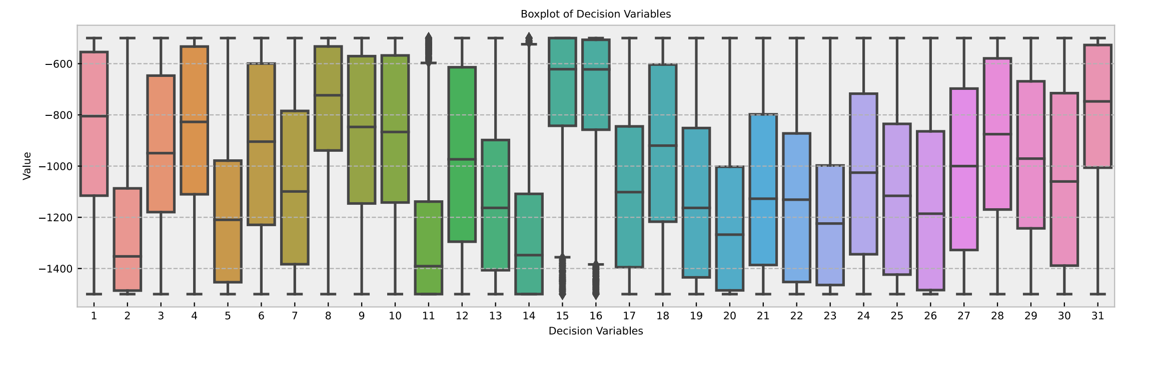


Figure 4.6 Decision Variable distribution in model 1 of NSGA-II, MOEA/D, and MOPSO algorithm

4.3.3 Pareto Front Performance Evaluation

All six metrics for the three model domains and four optimization techniques are shown in Figure 4.7 and described in Table 4.2. The Epsilon indicator, which measures the convergence of solutions, shows the lowest value for MOPSO at 1,980 across all models, indicating the best convergence among the techniques. In contrast, Pareto Search has the highest Epsilon value of 82,220, indicating less effective convergence. For the Spread indicator, which assesses the distribution of solutions, Pareto Search consistently achieves the highest value of 1.347, while MOPSO shows the lowest spread at 0.5618 across all models. This suggests that Pareto Search provides better solutions distribution along the Pareto front. The Generalized Spread, like Spread but considering overall distribution, is highest for Pareto Search at 1.3512 across all models, indicating it again offers a well-distributed solution set. For Generational Distance, which measures how far the solutions are from the true Pareto front, MOEA/D and MOPSO both achieve a value of 0.0, indicating perfect convergence, while NSGA-II shows a slightly higher GD at 0.015448, suggesting less accurate convergence. The Inverted GD, which evaluates the diversity of the Pareto front, is best for MOPSO with a value of 0.054306, indicating high diversity. Further, the epsilon metric and generational distance suggest that convergence of MOPSO is best. The parameters, epsilon, and generational distance show conflict compared to other optimization techniques.

Regarding uniformity of solutions, spread, and generalized spread suggest the superiority of NSGA-II for model 1. In model 3, even though the Pareto front of NSGA-II is diverse, the solutions are not uniform due to the spread of NSGA-II is high. For model 3, MOPSO is much more uniformly distributed and hence has a low value in the spread. The Spread and Generalized spread parameters depict that the MOPSO and NSGA-II have uniform solution distribution. In comparison, the MOEA/D shows a polar and non-uniform distribution of solutions.

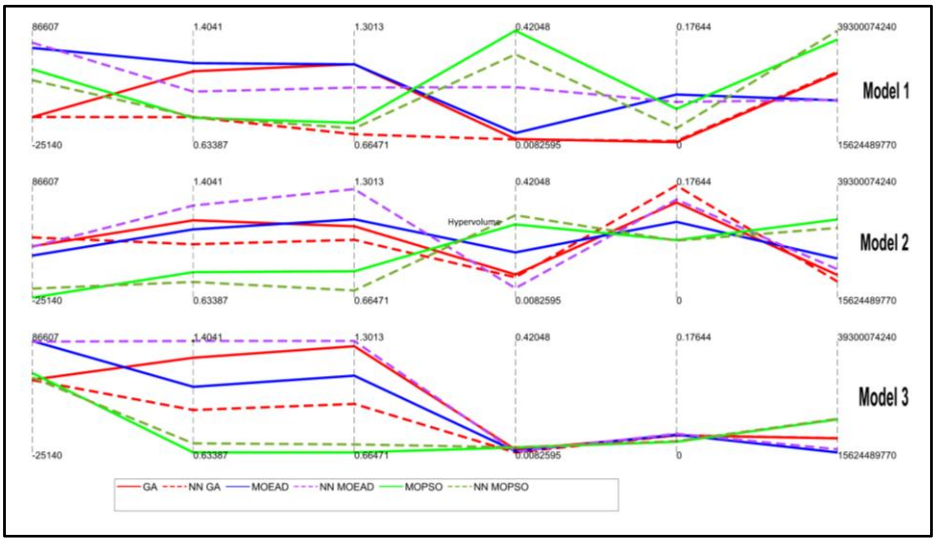


Figure 4.7 Values of pareto front evaluation parameters

Hypervolume, a measure combining convergence and diversity, is consistently highest for NSGA-II and MOPSO at 3.7×10^{10} across all models, indicating that these algorithms perform best in convergence and diversity. In contrast, MOEAD and MOPSO achieve a lower hypervolume of 1.5×10^{10}, implying less effective performance in capturing the Pareto front. Based on these indicators, MOPSO emerges as the best algorithm for solution distribution (Spread and Generalized Spread) and diversity (Inverted GD), making it a strong candidate for groundwater management scenarios requiring well-distributed and diverse solutions. However, MOEA/D shows superior convergence (Epsilon and GD), which is crucial for accurate optimization in groundwater management where precise trade-offs between objectives like discharge and river-aquifer exchanges are critical. NSGA-II and Pareto Search lead in Hypervolume, suggesting their overall robustness in balancing convergence and diversity. The normalized values for all the parameters for the model domain 1 are shown in Figure 4.8 below.

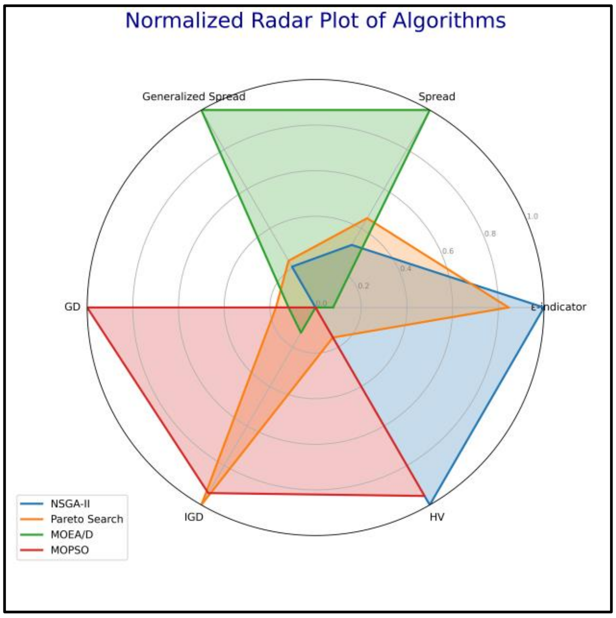


Figure 4.8 Normalized radar plot for pareto performance indicators for model 1

Table 4.2 Pareto performance indicators

	NSGA-II	**Pareto Search**	**MOEA/D**	**MOPSO**	**NSGA-II**	**Pareto Search**	**MOEA/D**	**MOPSO**	**NSGA-II**	**Pareto Search**	**MOEA/D**	**MOPSO**
Epsilon	66130	64970	59210	58636.8	34270.71	50623.9	48020	1980	29450	82220	41970	61467
Spread	1.15	1.19	1.34	1.06659	0.971	0.844	0.66866	0.609	0.59944	1.24544	1.14770	1.07673
Generalized Spread	1.059	1.07	1.35	0.98300	1.0819	0.883	0.723	0.678	0.654	1.132	1.124	1.086
Generational distance	0.002	0.005	0.004	0.015	0.0073	0.0145	0	0.0002	0	0.00379	0.003	0.005
Inverse generational distance	0.001	0.014	0.003	0.0139	0.0074	0.0272	0.014	0.0005	0.018	0.021	0.009	0.035
Exact hypervolume (x 1e10)	3.026	2.039	1.862	2.973	2.5375	1.7028	3.221	3.749	2.269	2.449	2.386	1.562

The Pareto search and MOEAD were more converged than NSGA-II, but they lacked diversity. Only in model 2 the diversity of the PF of MOPSO outperforms that of NSGA-II. The solutions of NSGA-II are highly diverse in model 1 and model 3, even at a low number of simulation runs. Other than the optimization algorithm, the Pareto front trends are determined by the model demarcation and its boundary conditions. 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. NSGA-II, PS, and MOEAD show a close value of hypervolume and inverted GD, indicating a nearly equivalent performance in terms of convergence and diversity.

4.3.4 Reduced PF and decision-making

To facilitate decision-making, 74 nondominated solutions from the Pareto front were grouped into 5 clusters using the DBSCAN clustering algorithm, which groups solutions based on density rather than predetermined centroids. These clusters, defined by their density structure, provided representative solutions that serve as reference points for each group. The solution closest to the core points within each cluster forms part of the reduced Pareto front. Figure 4.9 presents the clustered Pareto front with the highlighted representative solutions from each cluster. In this case, five representative solutions effectively summarise the Pareto front. The decision-maker now has a set of five optimal solutions that best represent the diverse trade-offs on the Pareto front, offering a more manageable set of choices for developing a strategy for optimal groundwater extraction from the LARB aquifer.

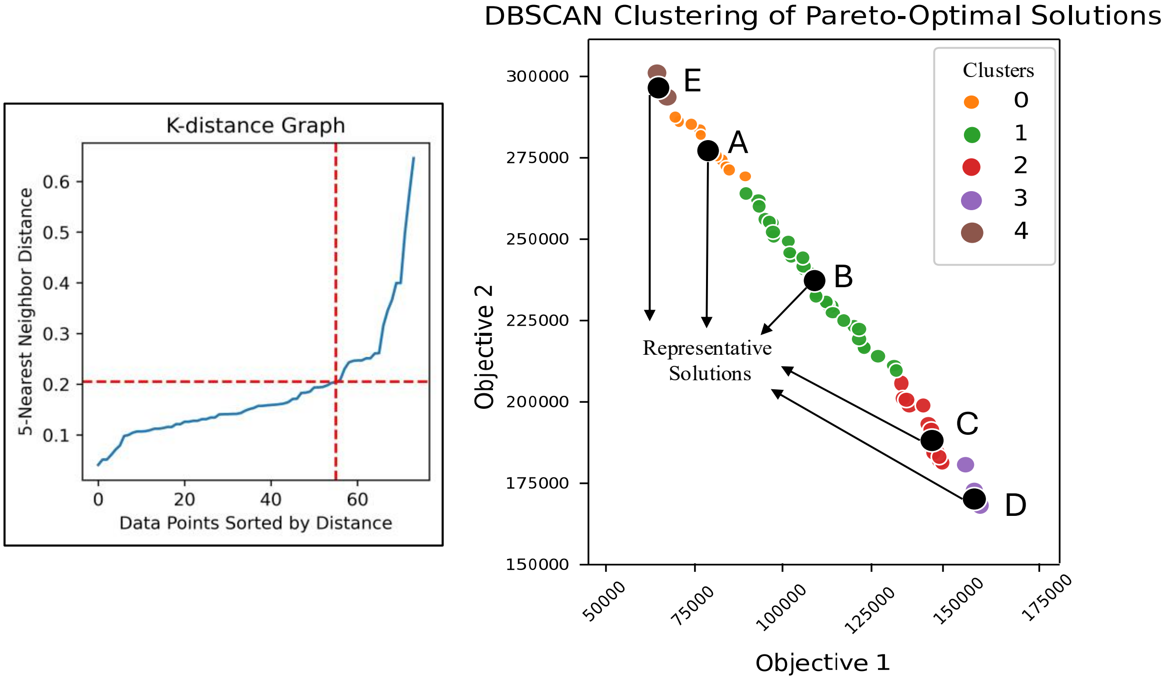


Figure 4.9 Part A represents the K-distance graph for optimal cluster of 5. Part B is the DBSCAN clustering of Pareto front solutions and representative solutions for decision makers

4.3.5 Management scenarios

This study shows that a GW multi-objective optimization is often conflicting, and using different metaheuristic optimizers can produce different high dimensional pareto fronts involving the non-dominated solutions. The reduced Pareto front containing five optimal solutions makes it relatively easier for decision-makers to weigh up the various trade-offs in the Pareto front. Implementing a solution from the reduced Pareto front can help balance these competing objectives, thereby ensuring the long-term sustainability of groundwater resources and river ecosystems in the LARB. Different pumping rates from the zones are shown in Figure 4.10.

Upon calibration of the simulation model with adequate precision, the optimization algorithm generates and assesses all management possibilities as potential solutions. Subsequently, one might be chosen that achieves the optimal objective function values.

Five solutions represent different amounts of trade-offs and management scenarios, which are explained below:

Solution A: High Groundwater Extraction, Low River-Aquifer Interaction

The pumping rates are relatively high and consistent across most decision variables, with values generally ranging between 1000 and 1600 m^3/d. A few variables show slightly lower values, indicating a more uniform approach to groundwater extraction. Most decision variables are operated near their upper limits, suggesting an aggressive extraction strategy.

Solution B: Balanced Groundwater Extraction and Moderate River-Aquifer Interaction

In this scenario, the solution balances maximizing groundwater extraction and moderately increasing river-aquifer exchanges like Solution A, but with slightly more variability. There are noticeable reductions in pumping rates for certain decision variables, particularly around variables 10 and 20. Peaks are overall, still high, but with more fluctuations compared to Solution A. This solution might be aimed at optimizing certain wells while reducing stress on others.

Solution C: Moderate Groundwater Extraction, High River-Aquifer Interaction

This solution shows a broader range of pumping rates, from very low (close to 400 m^3/d) to high values (up to 1600 m^3/d). The distribution indicates a more targeted approach, possibly to balance groundwater extraction and R-A exchanges more effectively. The R-A exchanges are maximized here through penalty introduction. There are distinct peaks around decision variables 5, 15, and 25, with very low values in between, suggesting a minimal extraction strategy.

Solution D: Minimal Groundwater Extraction, High River-Aquifer Interaction

This solution is ideal for environmentally sensitive areas where maintaining the ecological integrity of rivers and aquifers is paramount. It could be applied in regions where groundwater levels are already stressed, and the focus is on recharge and recovery of the aquifer system rather than extraction. It exhibits moderate to high pumping rates with fewer variables at lower extraction rates than Solution C. This solution appears to have a more balanced extraction strategy, with higher variability in the middle variables suggesting high R-A exchanges. Peaks are observed at variables 12 and 26, with consistent medium-range pumping rates across the rest.

Solution E: Extreme extraction with minimal R-A exchanges

It has high pumping rates across almost all variables, like Solutions A and B, but with even more uniformity. It demonstrates a consistent high-rate extraction strategy. Almost all variables are operated near the upper limit, indicating maximum utilization of available pumping capacities.

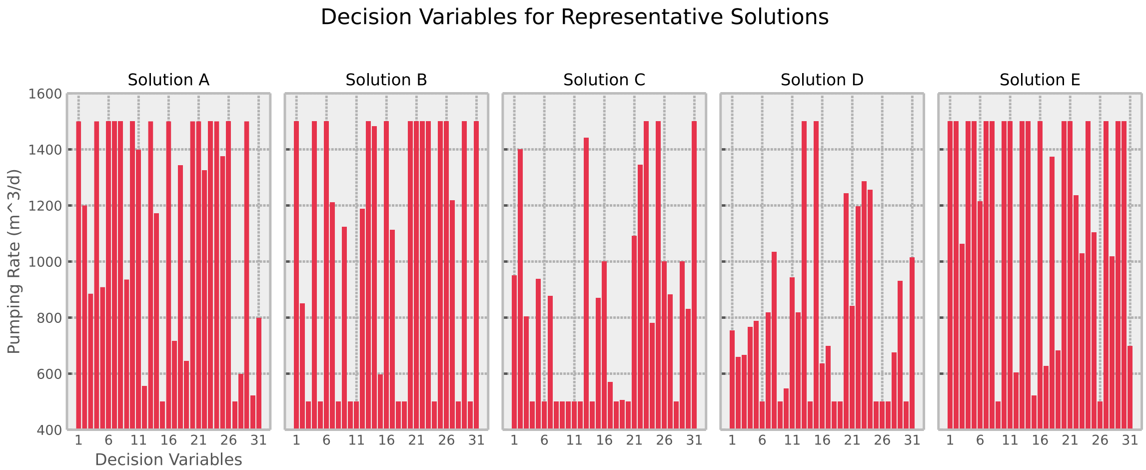


Figure 4.10 The GW pumping rates for the management period for the chosen solution

Solutions A, B, and E show a high and relatively uniform extraction strategy, potentially maximizing groundwater withdrawal. At the same time, solutions C and D demonstrate more variation, with selective lower pumping rates, possibly aimed at reducing the impact on the aquifer and improving R-A exchanges. The different patterns suggest that each solution represents a trade-off between maximizing extraction and other objectives like minimizing the environmental impact or preserving R-A exchanges.

4.4 Summary

The optimization of groundwater withdrawal and river-aquifer interaction was explored by applying three distinct boundary domains within a GW model, utilizing four optimization algorithms. The multi-objective linked S-O framework generated an extensive Pareto front, representing a range of optimal trade-offs. Performance metrics were employed to evaluate the resulting Pareto fronts. The findings underscored the influence of boundary conditions and domain size on the outcomes of simulation optimization models. Specifically, domain 1 exhibited superior efficiency, yielding a higher rate of river-aquifer exchange to the discharge from pumping wells. The comparison of Pareto fronts indicated MOPSO as the most effective algorithm, producing dominant solutions across the Pareto front. While NSGA-II generated diverse Pareto solutions, even with limited evaluations, these solutions outperformed those obtained from MOPSO.

In the context of the LARB, the management model developed in this study employed three simulation models and linked with four optimizers to optimize groundwater extraction patterns while maximizing R-A exchanges. Selecting a single solution from this large set was challenging, particularly for stakeholders with varying priorities and objectives. To address this issue, a DBSCAN clustering technique was applied to categorize the solutions based on similar characteristics, resulting in 5 distinct clusters. Each cluster, representing unique extraction strategies, was defined by its centroid, which served as a reference point. The solution from the Pareto front closest to each centroid was selected to form a reduced Pareto front. This method enhances the interpretability of the solution space and facilitates the selection of a focused and informed management strategy. This management plan can enhance SDAGE's goal of making good of the condition Rhône-Mediterranean basin. Good condition allows aquatic environments to provide the population with sustainable services: water supply, regulation of hydrological cycles (floods/droughts), fishing, swimming and water sports, and biodiversity. It contributes to the preservation of human health. If the management plan is implemented in the LARB, the GW can be sustainably developed in the region without affecting the river ecosystem while meeting every stakeholder's demand. Future research can extend the approach to incorporate uncertainties in aquifer parameters and boundary conditions inherent in real-world groundwater systems. By integrating stochastic methods, the model can be refined to develop more reliable and resilient management strategies for the LARB. This would further enhance the model's capability to support sustainable groundwater management under varying climatic and anthropogenic stressors.