Chapter 2: LITERATURE REVIEW

2.1 Introduction

This chapter offers an in-depth overview of the simulation-optimization models and a thorough review of the key concepts and methodologies central to this thesis. First, it provides a systematic understanding of the use of numerical models to quantify the complex interactions between surface and subsurface water systems, focusing on the role of R-A exchanges in sustaining hydrological balance and addressing nitrate dynamics in rivers resulting from agricultural activities. It also incorporates contemporary trends in coupled hydrogeological models and key findings. Secondly, the chapter focuses on applying S-O models in different nexus management strategies and enables decision-makers to derive solutions that maximize resource use while minimizing adverse environmental impacts. Third, it gives several key methodologies and lists different metaheuristic algorithms and objective functions used in S-O models that are scalable to other GW optimization problems where several decision variables become large and complex. It discusses using surrogate models and machine learning techniques in GW management problems. Lastly, by integrating WEFE nexus considerations, the research offers insights into the broader implications of groundwater management on food production, energy generation, and environmental sustainability.

GW is a critical resource for meeting industrial, irrigation, and household demands, accounting for approximately one-third of global freshwater usage (Moreaux et al., 2006). Modern pumping technology and increased demand have extracted excessive water from the aquifers. These unplanned withdrawals have over-exploited major aquifers and severely polluted groundwater (Jhaa et al., 2011; Remonti et al., 2016). With increasing demand and quality concerns, sustainable GW management is essential to regulate water flow and control pollutants, benefiting people and ecosystems. Effective management practices can help balance human water needs with the ecological requirements of groundwater-dependent ecosystems (GDEs). This involves integrated approaches considering hydrological, ecological, and socioeconomic factors to ensure long-term water sustainability.

The S-O technique has been at the core of discussion while deciding the suitable GW management strategy for several years (Ayvaz et al., 2008; Gong et al., 2016; Vedula et al., 2005). S-O's popularity comes from its quick decision support system, aided by an auxiliary black box model, which includes a conceptual model of the desired area, objective functions, design variables, and optimization algorithms. It calculates management plans more accurately and quickly than trial-and-error simulation modeling. A strategy obtained by S-O modeling will be at least 20% more effective than one produced through trial and error solely with a GW simulation model (Peralta, 2012).

In the S-O framework, the simulation model solves governing GW flow equations and their prediction. In contrast, the optimization model calculates the optimal management strategy from a collection of feasible alternatives (Kamali et al., 2017). The management objectives determine whether simulation models solve for the governing flow and solute transport equations separately or concurrently (Das et al., 2001). Hydraulic heads, source/sink rates, and source/sink solute concentrations at each node/grid/mesh are unknown simulation model variables that become decision variables in the optimization task. These variables are solved iteratively until optimized management goals are achieved. Finally, running the simulation model is necessary to recheck the optimal management approach that satisfies the constraints. As a result, the groundwater simulation model functions as a "multi-functional agent" within the S-O framework, acting both as a stimuli response calculator and as a management model checker. (Jean-Baptiste et al., 2020). However, real-world management problems include several objectives in a high-dimensional search space, making them non-convex, nonlinear, and multi-modal (Heydari et al., 2016). Therefore, understanding the S-O approach requires thoroughly examining all the elements that comprise its framework, including conceptualizing predictive models, types of management issues addressed, uncertainty and interactions between the elements, and decision-making.

Before (Gorelick, 1983) work, the role of S-O was largely unknown. He classified the S-O models into hydraulic management (pumping and recharge optimization) and policy evaluation and allocation (agricultural yield, water allocation problems). Later, (Wagner, 1995) reviewed the developments in S-O, focusing on mixed integer programming and the stochastic nature of aquifer properties. However, solver approaches were limited to linear and quadratic programming due to the computational resource constraints and uncertainties within simulation parameters. Stakeholders' social and economic goals were ignored. Another work (Yeh, 2015) compared various optimization techniques in management problems such as parameter estimation, experimental design, and groundwater planning. The analysis did not provide information on other challenges of S-O, such as the suitability of simulation models and design variables, nor did it examine the use of meta-heuristic algorithms. In the latest review by (Norouzi Khatiri et al., 2023), the authors discussed the uncertainty associated with S-O models in decision-making. However, the paper did not cover the applicability of surrogate models and objective functions for other management problems.

2.2 Simulation-Optimization framework

S-O is an integrated system that solves hydrogeological processes using a simulation model and employs an optimization technique to select the best plan among potential solutions.

Figure 2.1 illustrates the general methodology of a groundwater S-O model. The following sections discuss the objectives mentioned above.

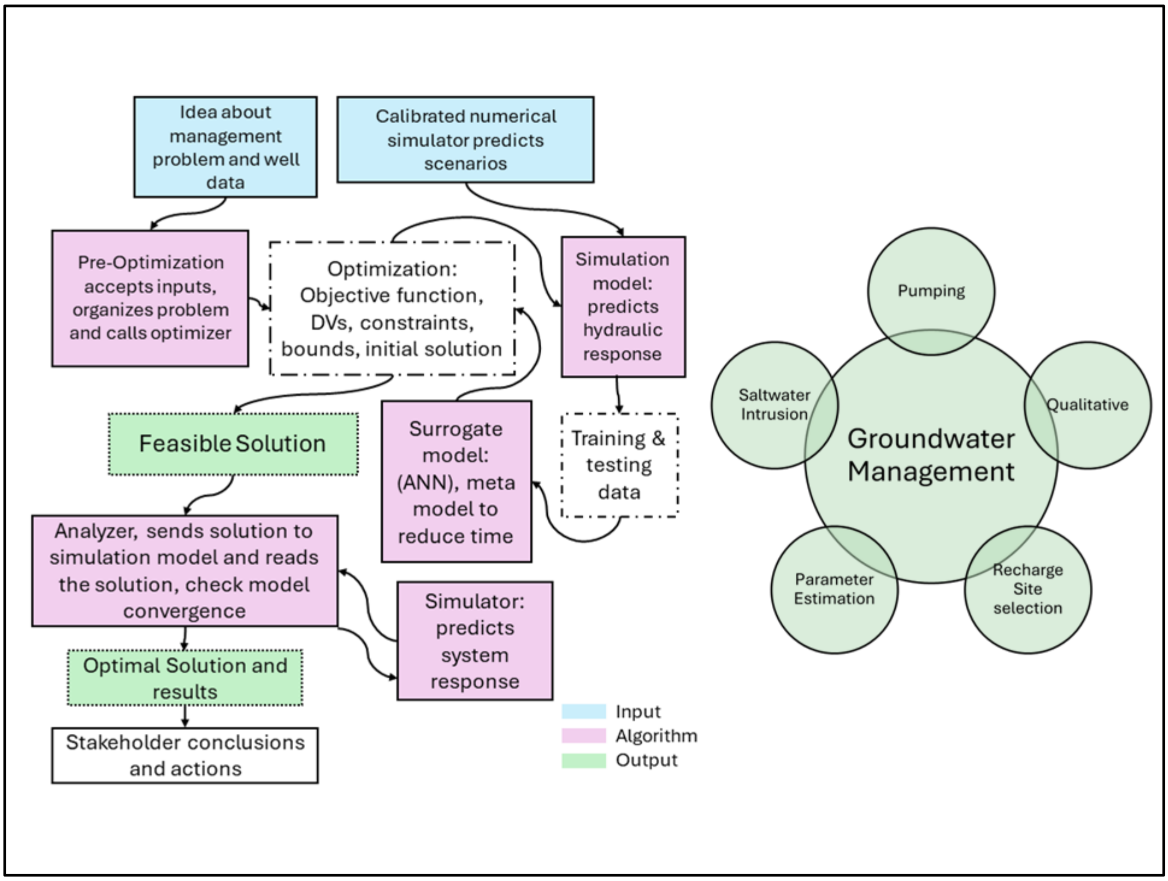

Figure 2.1 A schematic representation of the simulation-optimization framework for groundwater management.

2.2.1 Simulation models for groundwater-surface water interactions

In the S-O approach, simulation models predict state variable values, which provide information about the system's hydraulic condition (head, concentration, flux). These models can be numerical (process-based) or non-numerical (experimental) models. The scope of the article is limited to process-based numerical groundwater flow models. These models commonly simulate natural flow processes using mathematical equations, like Theis and other sources (Nasiri-Gheidari et al., 2018; P. Zhang et al., 2022). They consider the effects of various hydraulic stimuli, such as source/sink, pumping, and recharge, which are crucial in predictive modeling. Simulation models have primarily three goals in the S-O framework:

- Hindcasting/Forecasting the natural/anthropogenic changes in the aquifer regarding transient GW flow, pollutant transfer, and salt-water intrusion.
- Predicting state variable values for scenario-based analyses such as climate change (Dong, 2014) and land subsidence (M. Pang et al., 2022).
- As a management model only for simple case studies (real/hypothetical). Simple case studies = model involving only a few wells + calculation based on trial and error.

2.2.1.1 Numerical classification

The numerical classifications of simulation include but are not limited to the analytical element method (AEM), finite volume technique (FVM), boundary element method (BEM), finite element method (FEM), and finite difference method (FDM) (Maghsoudi et al., 2023). Each technique has its own set of benefits and flaws. By constructing a control network with structured rectangular grids, FDM systematically solves the governing equations of groundwater flow; however, it struggles to conceptualize complicated geological features and steep stratigraphic gradients (Shadab et al., 2023). While dealing with uncertain topologies, FEM's triangular element mesh is more accurate and adaptable than other techniques (Omar et al., 2019). However, FEM uses a grid approach to estimate the groundwater flow.

On the other hand, AEM solves groundwater flow equations without relying on a grid. It combines multiple analytical elements' exact solutions by depicting groundwater flow as a complex potential. Highly heterogeneous media and transient flow conditions are still beyond the AEM's simulation capabilities (Majumder et al., 2016).

2.2.1.2 Applicability/Use-case Classification

Groundwater is managed by controlling hydraulic stresses' location, timing, and magnitude. Most hydraulic stresses are anthropogenic, e.g., groundwater extractions, recharge from pits, and injection wells (Gerey et al., 2021). While some are unintentional, e.g., recharge from excessive irrigation (Gaur et al., 2015), controlling the entry, transport, and fate of contaminants in groundwater flow into the aquifer are some GW sustainable goals. Simulation models can vary in usability and proposed hydraulic stimuli that the user intends to predict. By its usage/applicability, simulation models are divided into qualitative, quantitative, and miscellaneous for the scope of this thesis.

A.) **Quantitative simulators:** According to (Peralta, 2012), simulation models calculate the head value(s) at the center of the cell before implementing an optimal strategy, determine influence factors from past stimuli, and, following the system nonlinearity, calculate system response to the optimal strategy. Flow optimization S-O models call only flow simulators. MOFLOW is a standard flow simulator (Brunner et al., 2010; Kumar, 2019; Meng et al., 2018; Soleimani et al., 2021). MODFLOW is a quantitative simulator for groundwater flow and contaminant transport. It employs advanced numerical techniques to solve complex flow equations, making it suitable for various environmental evaluations, including site investigations and aquifer balance simulations. The software's flexibility allows for simulating saturated and unsaturated flow conditions and integrating multiple hydrologic processes within a single framework. Additionally, MODFLOW's capability to model variable-density groundwater flow and its efficient computational performance make it a preferred choice for large-scale and detailed hydrogeological studies. The recent advancements in MODFLOW 6 further enhance its accuracy and efficiency, particularly in handling complex sedimentary structures and axisymmetric conditions. MODFLOW-response is compared with constraints to determine whether further iterations are needed. Models like WEAP and AquaCrop simulate agricultural demands and surface water consumption while solving conjunctive surface and GW resources (Zarriello et al., 2009).

Other simulators like GSFLOW are integrated SW-GW models created by combining 2-D surface hydrology with the 3-D groundwater flow model. HydroGeoSphere (HGS) is an excellent alternative to MODFLOW because of its capacity to simulate unsaturated flow separately for surface and GW flow (An et al., 2018), (Kayhomayoon, Milan, et al., 2022). HGS is robust and capable of modeling the entire terrestrial portion of a hydrologic cycle. However, it is complex to couple with an optimization model and has many assumptions.

B.) Qualitative simulators: Quality goals of sustainable GW management include minimizing the quality concentration (Yang et al., 2022), identifying contaminant sources (Cousquer et al., 2019), pump and treatment systems (Singh et al., 2016), and minimizing annual quality changes (Yin et al., 2022). To simulate the transport of multiple pollutants and their reactions together, researchers use MT3D or PHT3D as a standard simulator (Ayvaz, 2010). MT3D operates on the assumption that change in pollutant concentration does not affect the GW flow field. Quality optimization S-O models often call flow and transport simulators (such as MODFLOW and MT3DMS) iteratively until specific criteria are met. The flow simulator provides velocities and flow rates to the transport model.

In the past, researchers have used numerous models to simulate the behavior of contaminant flow, such as QUAL2K (Babamiri et al., 2021) to simulate temperature, pH, EC, DO, BOD, N-NH_4, and N-NO_3 parameters and their transport and fate; (Mirghani et al., 2012) used parallel groundwater transport and remediation (PGREM3D) codes, a finite element method based suite of codes for contaminant transport simulation. RT3D, a derivative of MT3DMS, extends its capabilities by incorporating reactive transport simulations in environments where biological, chemical, or physical reactions significantly affect contaminant transport. Other include HPx, PHT3D, OpenGeoSys (OGS), HYTEC, ORCHESTRA, TOUGHREACT, eSTOMP, HYDROGEOCHEM, CrunchFlow, MIN3P, and PFLOTRAN (Mirzaie et al., 2021; Naghdi et al., 2021; Paul et al., 2022; B. Wu et al., 2015). The models mentioned above are suitable for handling multiple species of contaminants; however, these models are computationally expensive and much more complex to couple with optimization algorithms. They also require extensive data, particularly for large, complex systems (P. S. Huang et al., 2018; Soleimani et al., 2021). Additionally, they typically assume homogeneity and isotropy in the aquifer properties, which may not always be accurate.

C.) **Miscellaneous:** These are the types of simulators used in the GW S-O technique, apart from quantity and quality. Some of them include seawater intrusion, irrigation water distribution, reservoir operations, parameter estimation, and pump and treat models (Al-Maktoumi et al., 2021; Bao et al., 2023; Fatkhutdinov et al., 2019; Mo et al., 2019; Zhao et al., 2015). One example is (Cheng et al., 2016), who used SUTRA, a 3-D saturated-unsaturated density-dependent flow and solute transport model, to simulate the seawater intrusion numerically. SUTRA utilizes a finite element estimate of the governing equations in space and an absolute finite difference approximation in time. (Christelis et al., 2019a) Used a variable density and salt transport (VDST) numerical model to provide high-fidelity simulations of seawater intrusion. Flopy is an alternative to GUI-based simulators (Ou, 2020). It is a set of Python scripts to run MODFLOW, MT3D, SEAWAT, and other MODFLOW-related GW simulators (Larsen et al., 2022). One significant advantage of using Flopy is the reproducibility of codes and models. Python has many scientific packages, making data exploration, iterative model calculations, and analysis quick and robust. However, it struggles with constantly editing the input file, making it inconvenient to create complex models. Table 1 lists the mathematical models utilized for simulation models in S-O.

Table 2.1 Types of the Simulation model for groundwater and R-A exchanges

Numerical method	Simulation Model	Usability	Advantages	Disadvantages
Finite Difference Method (FDM)	MODFLOW	Quantity	Popular & and widely used. Easy to define grid structure (rectangular or cubic lattice), which follows mass conservation. It can be easily integrated with decision support systems.	Arbitrary boundary stratigraphy cannot be modeled. Horizontal anisotropy is not handled well. The interpolation of values between nodes is not a primary consideration.
	MT3DMS	Quality		
	PHT3D	Quality (Reaction between pollutants)		
	OWHM	Human and Water demand-supply model		
	GWM	GW flow with optimization		
	SEAWAT	3D variable-density groundwater-flow		
	CFP	Turbulent GW flow model		
	Surface Water Routing (SWR)	SW-GW interaction		
	PMWIN	GW flow model		
	mflab	GW flow model (MATLAB-based)		
	Visual MODFLOW Flex	GW flow model		
	GMS	GW flow model + quality		

	Flopy	GW flow model (Python-based)		
	Groundwater Vistas	GW flow model + quality		
	MODFLOW-SURFACT	Subsurface flow model		
Finite element method (FEM)	HydroGeoSphere (HGS)	Integrated SW+ GW flow modeling	Better grid/mesh adaptability than FDM consisting of elements with variable node distribution. Employs piecewise continuous basis functions defined over elements, allowing for local approximation of the solution.	Mass imbalance at grid refinement
	FEFLOW	Saturated GW flow modeling		
	SUTRA	Saturated/ unsaturated model		
	3DFEMWATER			
Analytical Element Method (AEM)	flexAEM	GW flow modeling in saturated and steady state	Grid-independent analysis using divergence theorem. More accurate flow budget calculation.	Transient data handling still needs improvement. Transport models are cumbersome.
	Visual AEM			
	AnAqSim			
Finite Volume Method (FVM)	MPFA method	GW flow modeling in complex aquifers.	Can handle irregular grids on anisotropic heterogeneous domains.	Other packages, like transport particle tracking, are missing.
	MODFLOW (USG)			

	VCFVM (Qian et al., 2023)			

2.2.1.3 Use of surrogate models in Simulation Optimization

Optimization involves long, repetitive fitness calculations of the objective function(s) and constraint(s)) Achieving feasible solution(s)makes using S-O techniques computationally challenging (Gaur et al., 2013). This is particularly true in complex hydrogeological models with uncertain lithological layers and aquifer properties. For this problem, (Rajabi et al., 2018) proposed several potential strategies. Some strategies are (1) employment of efficient optimization algorithms, (2) parallelization and grid computing, (3) use of efficient Monte Carlo methods, and (4) employment of approximation techniques, such as lower-fidelity or response surrogate modeling. The last method has been widely explored in the work of (Chau, 2006; Seifi et al., 2020; Y. Wang et al., 2022).

For simplicity, simulators that use approximate ways to generate flow values are categorized into surrogate/approximate simulators. Some of the examples include response matrix simulators, statistically derived regression models, power functions, and artificial neural networks (ANNs). The most popular approximate simulators in the earlier years were the response matrix (RM) techniques (Kerebih et al., 2021). A response matrix of quantities describes how a physical system reacts to a specified hydraulic stimulus. Theoretically, such coefficients influence discrete kernels, Dirac delta, and response functions. This matrix elucidates the impact of decision variables, such as groundwater pumping, on state variables like head, stream stage, and stream-aquifer interactions. For instance, in the context of stream-aquifer interactions, closed-form solutions expressed as convolution integrals of impulse response and unit step response functions can be derived to relate channel reach discharge, stream-aquifer exchange rates, and associated flow volumes to hydrologic processes and management controls. Additionally, a discrete kernel generator can model aquifer behavior without stream interaction, expressing any aquifer response as an explicit function of pumping rates. Furthermore, a coupled simulation-optimization model using a response matrix approach can optimize the conjunctive use of groundwater and surface water, ensuring groundwater sustainability and maintaining environmental flow in rivers (Bajpai et al., 2022).

Response matrices have also been utilized as management models, including simulation and optimization models. The main advantage of RM is the low fidelity, approximate computationally faster simulator. It can solve several linear, nonlinear, and mixed binary linear GW management problems (Toews et al., 2016). However, the total impact on outcomes (the result of the S-O model) and the error associated with such models have not been discussed in much detail.

In recent years, numerous researchers have used ANN models or surrogate models as a means of reducing simulation time in S-O problems (X. Wang et al., 2020; Zibo Wang et al., 2022). The simulation model is replaced by training the black box ANN model; thus, there is no need to call the simulation model repeatedly during the entire iteration, saving considerable time. Several architectures have been reported for predicting simulated values, such as nonlinear autoregressive networks with exogenous input (NARX), long-short-term memory (LSTM), and convolutional neural networks (CNNs) (Hosseini et al., 2019). By far, feed-forward neural networks are the standard neural network architecture, followed by multilayer perceptron neural networks (MLP) and recurrent neural networks (RNNs) (Xu et al., 2022). Novel neural structures, like modular neural networks and generative adversarial networks, still lack application in GW flow simulators.

However, ANN simulation models have two main limitations: accuracy and versatility (D. Chen et al., 2016). The accuracy of the ANN simulator is directly related to its simulation data training, data redundancy, inaccurate training hyperparameters, and the architecture and design of the network. The versatility of the ANN model is also another limitation. For each state variable, a separate ANN needs to be trained. For example, ANN trained for one hydraulic stimulus cannot be replicated for another. However, in recent years, a class of deep neural networks, namely, physics-informed neural networks (PINNs), has emerged (Ghaseminejad et al., 2020). They are designed to satisfy training data and the concerned governing equations.

Some authors have applied hybrid simulators. One example is (Borah and Bhattacharjya, 2016), who developed a modified S-O approach. The best fitness values of the population were computed using a numerical simulation model, while the fitness of other individuals was simulated using ANN. This hybrid method significantly reduces the model's computing time while maintaining enough accuracy (Kourakos et al., 2023; Lykkegaard et al., 2021). The substitute (surrogate) model offers fast solutions at the cost of slight inaccuracy; therefore, it must be chosen based on problem definition, computational efficiency, and stakeholder requirements.

2.2.2 Optimization models

Optimization models consist of the parameters of the managed aquifer system, the management goals, and constraints. S-O can employ a variety of optimization algorithms to identify optimal design and extraction strategies, provided the problem is adequately defined and formulated. This section goes through components of optimization in the S-O framework.

2.2.2.1 Classification

Optimization models for GW management are classified based on the nature of the optimization problem, its objectives, and decision variable behavior. In the first type of classification, optimization models are classified into classical, non-classical, and hybrid approaches (Ayvaz et al., 2018; Chrispell et al., 2014). Classical optimization methods include simplex, quadratic, steepest descent, conjugate gradient, and dynamic programming (Beça et al., 2023; Gamache et al., 2014; He et al., 2021; Zhenchen Wang et al., 2022a). These methods are suited for linear GW systems with hydraulic problems (Peralta, 2012). However, these methods struggle when the gradient or Hessian matrix of the objective function is unknown (Chadalavada et al., 2011; Wehrens et al., 2000).

Evolutionary algorithms (EA) constitute heuristic optimization that uses biological learning and natural processes as inspiration to address complex optimization problems. As (Yeh, 2015) noted, one of the primary advantages of this approach is its ability to tackle complex multi-modal optimization problems in single and multi-objective domains. Some of the famous and commonly used EA(s) used in GW optimization and their pros and cons are listed in Table 2.

Table 2.2 Different types of metaheuristic algorithms used in S-O

EA Algorithm	Advantages	Drawbacks	Case study(s)
Non-Dominated Sorting Genetic Algorithm-II (NSGA-II)	Robust and requires less hyperparameter tuning	Slow convergence rate in many-objective-large scale optimization problems	Hydraulic (water supply, GW yield planning, conjunctive use) (Lalehzari et al., 2020; L. M. Pang et al., 2020; Triki et al., 2020; Yi et al., 2020)
Multi-objective Particle Swarm Optimization (PSO)	Rapid convergence, easy to comprehend, suitable for many objective GW optimization problems in quantity and quality optimization	Premature Convergence, Hyperparameter tuning in high dimensional problems	Hydraulic and transport (constraints on the contaminant, salt-water intrusion, contamination remediation(Mahmod et al., 2021; Saghi-Jadid & Ketabchi, 2021) (Ghaseminejad & Shourian, 2019) (Mohtashami et al., 2020)(Sabzzadeh et al., 2020; Zare et al., 2021)

Diversity-enhanced fuzzy multi-objective particle swarm optimization (fMOPSO/Div)	Diversity-based Gbest selection and mutation enabled global search.	Not suitable for high dimensional pumping optimization	(Mirzaee et al., 2021; Rezaei et al., 2020)
Predator-Prey PSO	Improved exploration and exploitation, Suitable for parameter estimation problems, effective constraint handling	Complex parameter setting, Slow convergence speed, Poor scalability	(Higashitani et al., 2006)
Firefly Algorithm (FA)	Strong searchability, adaptive and diverse, suitable for coastal management problems	Slow convergence, Inferior quality results	Kerebih and Keshari (2021)(Kazemzadeh-Parsi et al., 2015)
Chaotic Ant Colony Optimiser (CACO)	High robustness suitable for taking uncertainty problems, Effective for multi-objective problems	Extremely sensitive to hyperparameters, less application in GW optimization	(Ghadimi & Ketabchi, 2019) (Emami Skardi et al., 2015)(Zibo Wang et al., 2022)
Grey Wolf Optimizer (GWO)	Simple implementation, suitable for remediation of problems	Less application in GW, Poor results	(Majumder et al., 2020)
Multi-objective Evolutionary Algorithm by Decomposition (MOEA/D)	Suitable for large-scale optimization problems, robust and diverse solutions	Computationally complex, parameter sensitive, lacks rigorous theoretical analysis	(Bajpai et al., 2022)

The number of objective functions and decision variables (DVs) divides GW optimization problems into six categories (Tian et al., 2022). These include optimization problems with:

1. Single Objective Problems (SOPs): These are the most straightforward optimization problems, focusing on a single objective function, such as minimizing pumping costs or maximizing aquifer recharge. SOPs are helpful for situations where the decision-making process revolves around one clear, dominant goal.
2. Multi-Objective Problems (MOPs): MOPs involve two or more conflicting objectives that must be optimized simultaneously. In GW systems, examples include balancing the maximization of groundwater extraction against minimizing environmental impacts or ensuring sufficient river-aquifer exchanges while maintaining agricultural water supply. These problems often rely on Pareto-based optimization methods to identify trade-offs and support decision-making.
3. Many Objective Problems (MaOPs): These extend MOPs to scenarios with more objectives (typically more than three). In GW management, MaOPs might optimize water allocation for agriculture, industry, and domestic use while minimizing groundwater contamination and energy consumption. Such problems require advanced optimization techniques to handle the increased complexity and interdependencies between objectives.
4. Large-Scale Many Objective Problems (Many-LSOPs): The most complex category, Many-LSOPs, involves many objectives (more than three) and decision variables (DVs > 100). These problems are common in integrated Water-Energy-Food-Environment nexus management at a regional or basin scale. The objective functions might include groundwater sustainability, energy efficiency, food production, and ecosystem preservation. Many LSOPs demand robust

computational frameworks, such as surrogate modeling and decomposition-based multi-objective algorithms, to manage their scale and complexity effectively.

In practicality, the GW management problem is large-scale and has numerous objectives (Allam, 2017). Convergence issues, huge computational burdens, and irregular results are some of the challenges in high-dimensional problems. Few articles derived methodologies for reducing large DV(s) (number of wells) in optimization by clustering them based on administrative boundaries (Zhenchen Wang et al., 2022b); or by constructing a well zone within some threshold from natural boundaries (rivers, reservoirs). Reducing significant decision variables in GW optimization problems requires further exploration and research in decision variable space. (Tan et al., 2023) have recommended three ways to reduce DVs in any optimization problems: a) Decision variable grouping (clustering), b) decision space reduction, and c) novel search strategy-based algorithms. The first is a decomposition method using various variable grouping techniques, including random differential grouping (Mulligan et al., 2016). The decomposition approach splits the decision variables into multiple groups and optimizes each group of decision variables in succession. The second method, decision space reduction, reduces the high-dimensional decision space to small-scale optimization problems using problem transformation or dimensionality reduction approaches (Tang et al., 2020), addressed using traditional evolutionary algorithms (X. Zhang et al., 2018). The third approach suggests efficient reproduction operators, such as parameter adaptation-based differential evolution and new learning strategy-based particle swarm optimization, to provide well-converged solutions.

2.2.3 Large Scale multi-objective problems

In large-scale multi-objective optimization problems (MOPs) in groundwater management, many real-world applications involve equal to or more than 100 decision variables, making them highly complex and challenging to solve using traditional multi-objective evolutionary algorithms (MOEAs). Existing MOEAs often struggle with scalability in such problems, as the performance of these algorithms tends to deteriorate significantly as the dimensionality of the decision variable space increases. This challenge, known as extensibility in the variable space, has received limited attention in research addressing large-scale optimization problems in GW management. Three substantial approaches have been explored within other domains of the optimization field (Tian et al., 2022), as shown in Figure 2.2. These approaches include a) DV decomposition by grouping, b) DV space reduction, and c) Novel search strategy-based algorithms (non-decomposition). The decomposition method has gained popularity and is widely adopted in the case of GSOPs. It involves dividing the decision variables into smaller groups and optimizing each group separately.

On the other hand, DV space reduction techniques aim to reduce the dimensionality of high-dimensional decision variable spaces using problem transformation and dimensionality reduction methods such as random embedding and unsupervised neural networks. As for the non-decomposition approach, it focuses on achieving convergence through efficient search strategies employed by emerging evolutionary algorithms. For instance, the utilization of new adaptive reproduction operators in the differential evolution algorithm (Tanabe and Fukunaga, 2013) and the parallel dynamic search with surrogate-based optimization (Pang et al., 2020). However, it is essential to note that these strategies have mainly been tested on design functions, limiting their applicability to broader contexts. Efficient handling of huge decision variables in groundwater simulation optimization problems (GSOPs) has been relatively underexplored, with only a limited number of authors delving into this area. For instance, (A. J. Siade et al., 2020) utilized truncated singular value decomposition to effectively reduce the number of decision variables. Similarly, (Fienen et al., 2018) employed a straightforward clustering technique to group wells based on proximal distance. However, it is worth noting that the impact of this clustering approach on the quality of the obtained optimal solutions has yet to be comprehensively assessed.

Decision variable grouping adopts a divide-and-conquer approach, where the decision variables are grouped so that both convergence (toward the Pareto front) and diversity (spread across the front) are prioritized. This method facilitates handling high-dimensional decision spaces by breaking them into smaller, more manageable subcomponents, allowing for easier optimization.

Alternatively, decision space reduction reduces the dimensionality of the high-dimensional decision space, simplifying the search process for optimal solutions. This is typically achieved using problem transformation or dimensionality reduction techniques, where GSOPs are transformed into smaller-scale optimization problems that can be solved using conventional evolutionary algorithms. Different dimensionality reduction methods, such as random embedding, principal component analysis (PCA), unsupervised neural networks, and other machine learning techniques, have been explored. However, it is essential to note that many machine learning-based dimensionality reduction techniques do not apply to GSOPs, as they produce vectors that cannot be recovered to their original form, which is crucial in evolutionary optimization. In problem transformation, the decision space is reduced by optimizing weight vectors rather than the complete set of decision variables. This approach can efficiently find local optima in the reduced decision space, although it may struggle to identify global optima, even with extensive function evaluations.

In contrast to these methods, a newer category of MOEAs focuses on directly solving LSMOPs by introducing novel search strategies within the original decision space without reducing its dimensionality beforehand. These approaches incorporate advanced techniques such as novel reproduction operators and probability models, which generate offspring solutions in the high-dimensional decision space. By evolving populations in the original decision space, these MOEAs aim to maintain the full complexity of the optimization problem, potentially improving the chances of finding global optima in large-scale groundwater management problems. These new MOEAs have potential, but to effectively manage the complexity of high-dimensional landscapes, they usually need additional computer power and complex evolutionary methods.

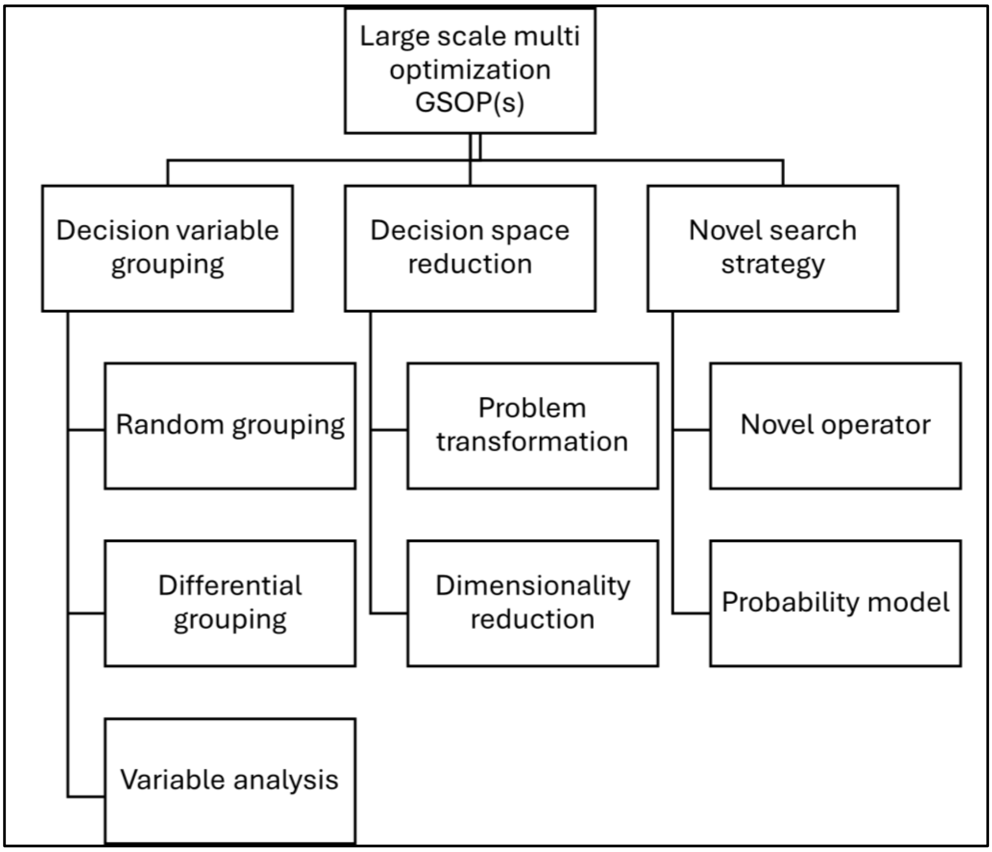

Figure 2.2 Large-scale variable reduction techniques

2.2.3.1 Objective functions for nexus management

The objective functions represent the crucial management goal(s) mathematically. In multi-objective goals, a conflict or trade-off exists. A solution that minimizes or maximizes all objectives does not exist (Siade et al., 2020). In other words, there are multiple trade-offs among all the objectives, where the improvement of one objective cannot be achieved without the deterioration of at least another. Researchers incorporate the simultaneous goals in the objective functions or transform them as constraints.

Some quantitative goals include maximizing the pumping while maintaining a permissible head (Norouzi Khatiri et al., 2023; Papadopoulou et al., 2007), minimizing the pumping energy/well diversion and installation cost (González Perea et al., 2016), maximizing river-aquifer exchanges (Bajpai et al., 2022; Hernandez et al., 2015); minimize the drawdown and energy consumption (Yidana et al., 2008), maximize the total yield from the area (Kerebih et al., 2021), maximizing the water supply to the irrigated area (S. Chen et al., 2016), and minimize the total volume of water applied during the whole irrigation season and optimal conveyance of irrigation water (Babamiri et al., 2021; Majedi et al., 2021a).

However, current GW management problems demand an interdisciplinary approach (Govender et al., 2021). For example, the management of coastal aquifers needs changes in coastal groundwater environmental parameters such as groundwater level, water demand, salt-water wedge volume, and salt-water intrusion area, which are usually considered during optimization. This is shown by the study of (Ayaz et al., 2021; Majedi et al., 2021b; Zekri et al., 2015). Managing agricultural water through S-O demands socio-economic, hydrodynamic, and environmental disciplines, while designing managed aquifer recharge (MAR) structures for GW recharge through S-O demands qualitative, policymaking, and hydrogeology aspects.

The conflicting objectives are solved, and the non-dominated solutions from objective vectors are called Pareto vectors. Consequently, these results are presented as a Pareto front, where stakeholders can find various solutions showing different trade-offs and complements. However, only a few authors ((Bajpai et al., 2022; Perea et al., 2020; Sreekanth et al., 2016)) have discussed the quality of Pareto fronts regarding convergence and diversity of solutions. Table 2.3 lists objective functions and their corresponding decision variables for various management goals.

Table 2.3 Different Objective Functions and Decision Variables in large-scale S-O Problems

Problem Type	Common objective functions	Constraints	Decision Variable	Authors
Quantity	$\text{Min } f(q,t) = Penalty + \sum_{p=1}^{r}\left[\sum_{k=1}^{timep}\left(oflag_{pk} * \sum_{i=1}^{n} a_i * q_{ip}\right) + \sum_{i=1}^{n} a_{ip}^{0} * (1 - \exp(-b * q_{ip})\right]$ Qip= component vector of pumping rates at wells i =1, ..., n during management period p =1, ..., r.	$q^{min} \le q_{ip} \le q *_{ip}, i = 1, \ldots, n$ q^{min} minimum allowable pumping rate. $q*_{ip}$ maximum allowable pumping rate for well "i" at the period p.	Pumping Diversion Pumping energy cost Well installment Cost Monetary value Aquifer head	Papadopoulou et al. (2007)
	$Max\ Q = \sum_{m}^{M}\sum_{i}^{P}\sum_{j}^{S} q_{m,i,j}$	$\sum_{m}^{M} Q_{demand} \le \sum_{m}^{M}\sum_{i}^{P}\sum_{j}^{S} q_{m,i,j}$	Head difference, gradient, or velocity	Yidana (2008)
	$Max \begin{cases} \sum_{i} L_i - Penalty \\ \sum_{i=1}^{nz} N_i * Q_i - Penalty \end{cases}$	$(Q_i)lb \le Q_i \le (Q_i)_{ub}$	Stream aquifer seepage Stream Flow	Bajpai et. al., (2022)
	$Min = \sum_{t-0}^{T} Q_{f,t}\ (Minimization\ of\ freshwater\ footprint)$	$dd_t \le dd_{acc}$	User-defined variable	Hernandez & Uddameri, (2015)
Quality	$Min\ F1 = \sum_{k=1}^{nk}\sum_{i=1}^{N}(C_i^k - Cobs_i^k)\text{^}2$ where Ci is calculated conc. Furthermore, Cobs is observed conc of pollutants.	$C_i \le C_{permissible}$ $(Q_i)lb \le Q_i \le (Q_i)_{ub}$	Particles captured	Jha & Datta (2012)
			Maximum Concentration	Zhao et al. (2015)
			Residual	Avyaz (2016)

	$Min = \frac{1}{12}\sum_{i=1}^{I}\sum_{k=1}^{K}(\frac{C_i(k)}{C_{max}}) /(\frac{a_k}{A})$		Contaminant mass remaining	Heydari et al. (2016)
Agriculture	$Max\ Q = \sum_{m=1}^{M}\sum_{i=1}^{P}(\Delta t_p)_k (Q_{gw})_k + \sum_{m=1}^{M}\sum_{i=1}^{P}(\Delta t_d)_k (Q_{SD})_k$ Where Q_{SD} is the stream diversion, $(\Delta t_p)_k$ is the pumping time, $(Q_{gw})_k$ is GW extraction.	0<= $(Q_{gw})_k$<= $(Q_{gw})_{up}$ 0 <= Q_{SD} <=Qup	Pumping and Diversion Stream-aquifer seepage	Kerebih & Keshari (2021)
	$F = max\sum_{m=1}^{M}\sum_{i=1}^{I} Y_{j,i} C_i A_{j,i}$ where, $A_{j,i}$ the area under crop, $Y_{j,i}$ actual crop yield, C_i current market price.	$h_{min} \leq h_{i,m} \leq h_{max}$	Crop yield Groundwater supply	Chen (2016)
	$F2 = Min\ \sum_{h=1}^{H}[(E_t) - DI_h]$ Where, E_t is seasonal energy cost, DI_h deep percolation loss	-	Monetary cost	Perea et al., (2016)
Coastal Management	$F = Min\ \sum_{j=1}^{J} D(F_j, F_j *)$ where, $D(F_j, F_j *)$ is the energy distance between Fj and Fj*, J total number of observation wells. Fj = cumulative distribution function (CDF) of ith well. Fj* = CDF of ith well at zero discharge.	$E = \sum_{j=1}^{N} E_i * A_i \geq 1.44Mm^3 /year$	Installation cost Concentration difference	Rajabi & Ketabchi (2017)
	$\min f = [(mass_{end} - mass_{ini}) /mass_{ini}] * 100\%$ $mass_{ini}$ and $mass_{end}$ total solute mass in the coastal aquifer at the beginning and end of the management period Q_j^{in} Is the water injection rate of the *j*th injection well; Q_i^{ex} is the pumping rate of the *i*th pumping well.	$\sum_{i=1}^{m} Q_i^{ex} = Q_0 + \sum_{j=1}^{n} Q_j^{in}$	Salt-water concentration Water diversion	Han et al., (2020)
	$\min \sum_{i=1}^{I}\sum_{j=1}^{J} Q_{i,j}$ where, $Q_{i,j}$ is injection rate $C_{i,T}$ simulated conc. and C_i^{max} max allowable conc.	$C_{i,T} \leq C_i^{max}$ $0 \leq Q_{n,m} \leq Q_{max}$	Residual concentration	Huang & Chiu (2018)

Parameter Estimation/ Uncertainty Analysis	$Min\ \sqrt{\frac{1}{v}\sum_{i=1}^{I} e^2}$ where, e represents the difference between the observed and estimated parameters.		Head difference	Zheng et al., (2019)
	$Min = \sum_{i=1}^{I}(AMD)$ $AMD = \left(\frac{1}{NC}\sum_{j=1}^{J}\lvert H_j\rvert - H_y\right)$ where, AMD= annual mean drawdown NC= number of active cells		User-defined Variable	Zekri et al., (2015)

2.2.3.2 Constraints

The accuracy of a simulation model is limited to the data range used in calibration and validation. In optimization problems, constraints help to address these limits beyond models. Constraints for GW management are broadly divided into the following categories. They:

- Ensure the simulator adheres to physical laws and establishes relationships between DVs and system variables (SVs) (permissible drawdown, maximum discharge, yield capacity). (Javan et al., 2024; Shi et al., 2023)
- Offer transient initial and background conditions for specified scenarios (including climatic, hydrologic, anthropomorphic, infrastructural, and socio-economic). (Cook et al., 2022; Ren et al., 2022)
- Meet sociopolitical, legal, environmental, and economic objectives. (Cai et al., 2021a; Lin et al., 2022; Uen et al., 2018)

In classical optimization problems, constraint handling is done via the Lagrangian approach that converts the constraint problem into an unconstrained one using the Lagrangian function (Fienen et al., 2013); projection techniques, where each infeasible solution is projected to its nearest feasible solution in the search space (Alexander and Ndambuki, 2021); and sequential unconstrained techniques, which convert the main problem into sequential optimization with linear constraints.

In the GW problems with the heuristic approach, constraints are handled by introducing a penalty to the total cost based on constraint violations (Crevillén-García, 2018). The penalty functions can be either static (penalties do not change during optimization), dynamic (functions vary during each or group of iterations), or adaptive (functions that penalize infeasible solutions based on search space). Some authors have compared constraint handling techniques, e.g., (Hilton and Culver, 1998), who compared additive and multiplicative penalty methods and recommended multiplicative penalty methods for more robustness and better solutions.

The number of constraints plays a crucial role in the convergence and computational time of the management problems. Most of the S-O approaches can handle non-linear systems if there are a smaller number of constraints. However, when the situation becomes large-scale (regarding constraint number), the S-O approach struggles with its convergence (Christelis et al., 2019b).

2.2.3.3 Uncertainty analysis

Uncertainty in S-O models can result from the inherent characteristics of the applied numerical model and the parameters of the input aquifer. These parameters, which encompass subsurface lithology, stratigraphy, and physical and chemical properties, contribute to the complexity of the system analysis and optimization procedure (Teixeira Parente et al., 2019). Moreover, a single deterministic model does not consider the researcher's uncertain knowledge of the accurate aquifer parameters. This uncertainty is addressed in operational research by employing a probability density function (PDF) for the uncertain parameters. Nonetheless, accurate PDFs are seldom available for groundwater optimization problems.

However, there exists a trade-off between reliability and computational time. Experts handle aquifer parameter uncertainty using a stochastic approach to generate numerous parameter combinations. In S-O, we can handle the issue of uncertainty in three methods: generalized likelihood uncertainty estimation (GLUE), Markov Chain Monte Carlo (MCMC), and Bayesian recursive estimation (BaRE) (Luciano Raso, 2013). The central concept of the GLUE method is to abandon the idea of a global optimal parameter solution. Instead, GLUE assumes that the performance of a groundwater model depends on the interaction of multiple model parameters. Markov Chain Monte Carlo (MCMC) uses random sampling to explore the probability distribution of the objective function. It builds a Markov chain with the same distribution as the objective function and moves along the chain to generate samples. Examples of the MCMC method include differential evolution adaptive metropolis (DREAM (zs)), delayed acceptance, metropolis-hastings (Keating et al., 2010; Wu and Zeng, 2013). Meanwhile, the BaRE approach provides a probabilistic framework with beliefs about unknown parameters or states based on observed data and dynamically adapts predictions for making decisions in the presence of evolving uncertainties.

However, some of the challenges are as follows: Nonetheless, any simulation model can fail to capture crucial characteristics of a problem. Hence, prediction needs to involve alternate solutions that may inherit essential characteristics. Secondly, mathematical formulation cannot prove absolute correctness; therefore, uncertainty prediction should aim to eliminate extreme and incorrect solutions. Lastly, it involves hidden and poorly characterized subsurface; consequently, it is incredibly challenging to determine which method will capture the necessary information (Amirabdollahian et al., 2019).

2.2.4 Decision-making strategy

Since GW-related goals often conflict with each other, the stakeholders need to understand the gains and sacrifices between multiple solutions before making a management plan. Researchers and practitioners have actively employed various techniques to select a single optimal solution from the pareto frontier in recent years. One can broadly categorize these techniques into visualization and statistical methods relying on weighted ranking.

2.2.4.1 Visualization methods

Visualization techniques help us understand and explore management plans with trade-offs in multiple or many objectives, especially in large-scale optimization problems. In recent years, visualization techniques have gained popularity in GW optimization problems. Visualization techniques present a clear picture of the objective functions along with DV space, which helps in its better exploration and choosing key decision variables. This ability to explore DV space permits the decision maker to make more informed, sustainable decisions. Some of the popular visualization techniques are Scatterplot Matrix, parallel coordinate plot (PCP), heatmap, RadViz-3D, t-SNE plot, principal component analysis (PCA), self-organizing maps (SOM), and hyperspace Pareto frontier (HPF).

Few visualization techniques, like the scatterplot matrices and tile plots, are easy to implement. However, they can only represent a limited number of dimensions. Meanwhile, PCP, heatmap, RadViz, and t-SNE plots are used for any dimension. (Nagar et al., 2023) introduced a new method, iSOM, to understand insights among DV interactions.

2.2.4.2 Weighted statistical methods

Researchers prefer statistical methods that involve weight/rank for each objective. This method considers the preferences of multiple stakeholders in GW management. However, it has challenges like method selection, weight assignment, and DV interaction. Some standard statistical methods for choosing an optimal management plan are listed below.

Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS)

TOPSIS is a multi-criteria decision-making method that ranks the solutions based on their similarity to an ideal solution and their dissimilarity to a negative ideal solution. The ideal solution is the one that has the best value for each objective function, while the negative-ideal solution is the one that has the worst value for each objective function. TOPSIS calculates the Euclidean distance of each solution from the ideal and the negative ideal solutions and then calculates a relative closeness coefficient that reflects how close each solution is to the ideal solution and how far it is from the negative ideal solution. The solution with the highest relative closeness coefficient is the best compromise among the Pareto front. Some of the examples include identifying GW remediation strategies in naphthalene-contaminated sites (Qin et al., 2009), reducing water demands from stakeholders (Mansour et al., 2023; Yu et al., 2024), and conjunctive use of treated wastewater and GW (Yu et al., 2024).

Linear Programming Technique for Multi-dimensional Analysis of Preference (LINMAP)

LINMAP shares a similar methodology to TOPSIS, except that TOPSIS calculates both positive and negative ideal points. On the other hand, LINMAP considers distance only to be a positive ideal. LINMAP uses a linear programming model to estimate the coordinates of the ideal point and the weights of the attributes by minimizing a measure of poorness of fit. LINMAP can handle crisp and fuzzy data and perform external and internal analysis. Both TOPSIS and LINMAP are widely used in quantitative and qualitative optimization problems.

Elimination and Choice Translating Priority III (ELECTRE III)

ELECTRE III is a powerful multi-criteria decision-making method for selecting a single solution from the Pareto front in groundwater optimization challenges. Its strengths lie in its ability to handle complex, real-world problems by incorporating various criteria and preferences. However, ELECTRE III may encounter challenges in larger-scale applications due to computational demands and potential subjectivity in assigning weights to criteria, which can impact the final solution. Additionally, while it can effectively navigate trade-offs among conflicting objectives, selecting specific parameters might influence the method's robustness, necessitating careful calibration for optimal outcomes.

A little study has compared the applicability of these weighted statistical methods. (Gaur et al., 2021) applied these methods to identify optimal GW management strategy and found TOPSIS performs better with GW problems. Another study by (Dong, 2014) highlighted comparing different weighted methods used in GW quality studies. Although each method provides different levels of accuracy and flexibility, stakeholder preference in the decision-making strategy choice is always crucial.

2.3 Research gaps

Adequate groundwater management is critical for sustainable water resources, particularly in regions where groundwater and surface water systems interact dynamically. Despite significant advancements in groundwater modeling and optimization, several critical gaps remain, particularly in simulation-optimization frameworks and their application to R-A systems within the broader WEFE nexus.

- Comparative analysis of evolutionary algorithms for groundwater management:

  While evolutionary algorithms have shown promise in solving complex multi-objective groundwater optimization problems, there is a noticeable lack of comparative studies that evaluate their performance in different groundwater management contexts. This gap limits our ability to determine the most effective algorithm for specific hydrogeological conditions, particularly in R-A exchange scenarios.

- Multi-Criteria decision-making based on decision variable space: Decision-making in groundwater management often involves trade-offs between competing objectives, such as maximizing groundwater extraction while sustaining R-A exchanges. However, limited research has focused on leveraging the decision variable space to facilitate multi-criteria decision-making (MCDM). Exploring the decision variable space can provide deeper insights into how specific management strategies influence optimization outcomes and enable the identification of optimal solutions that balance economic, environmental, and social objectives.
- Neglect of R-A exchanges in groundwater optimization models: R-A exchanges play a critical role in maintaining hydrological balance, yet they are often overlooked in groundwater optimization models. The exclusion of these exchanges can lead to unsustainable management practices. Addressing this gap requires the integration of R-A exchange dynamics into optimization frameworks to ensure the sustainability of both surface and subsurface water systems.
- Handling large decision variable sets in optimization problems: Groundwater management problems often involve large and complex decision variable sets, particularly in regional-scale models. However, no standardized methodology exists for efficiently handling these large decision spaces in optimization frameworks.
- Novel Surrogate models for Transient Data Prediction: Surrogate models, such as PINNs, have shown potential for improving computational efficiency in S-O frameworks. However, their application in predicting transient groundwater data, particularly in R-A systems, remains limited.
- Lack of a harmonized framework for nexus trade-offs and synergies: The WEFE nexus approach emphasizes the interconnectedness of water, energy, food, and

environmental systems. However, no standardized or harmonized framework systematically explores these nexus components' trade-offs and potential synergies. This gap limits our ability to develop integrated management strategies that optimize resource use while minimizing adverse environmental impacts.

2.4 Summary

The literature review on the management of GW and its nexus suggests that the simulation optimization technique is suitable for handling complex management goals while adhering to different constraints. However, several pertinent components for its application in real-life scenarios still require improvements. Addressing these gaps is critical for advancing the science and practice of sustainable groundwater management. This thesis aims to develop a comprehensive S-O framework that integrates advanced computational tools, such as evolutionary algorithms and surrogate models, to optimize groundwater use while sustaining R-A exchanges. By bridging these research gaps, the study seeks to provide a robust decision-support framework that balances the competing demands of the WEFE nexus, contributing to sustainable water resource management at regional and global scales.