

TABLE OF CONTENT

CERTIFICATE	iii
ACKNOWLEDGEMENT.....	ix
ACRONYMS.....	17
ABSTRACT.....	19
CHAPTER 1: INTRODUCTION.....	23
1.1 Background and motivation.....	23
1.1.1 River-Aquifer Exchanges and its significance.....	23
1.1.2 The transdisciplinary Water-Energy-Food-Environment nexus.....	26
1.2 Problem statements.....	30
1.3 Research design.....	31
1.4 Research aims and objectives	32
1.5 Thesis structure and organization	33
CHAPTER 2: LITERATURE REVIEW.....	35
2.1 Introduction.....	35
2.2 Simulation-Optimization framework.....	37
2.2.1 Simulation models for groundwater-surface water interactions	38
2.2.1.1 Numerical classification	39
2.2.1.2 Applicability/Use-case Classification.....	39
2.2.1.3 Use of surrogate models in Simulation Optimization	45
2.2.2 Optimization models	47
2.2.2.1 Classification	47
2.2.3 Large Scale multi-objective problems	51
2.2.3.1 Objective functions for nexus management	54
2.2.3.2 Constraints.....	58

2.2.3.3 <i>Uncertainty analysis</i>	59
2.2.4 Decision-making strategy	61
2.2.4.1 <i>Visualization methods</i>	61
2.2.4.2 <i>Weighted statistical methods</i>	61
2.3 Research gaps	63
2.4 Summary	65
CHAPTER 3: DEVELOPING AN INTEGRATED SIMULATION- OPTIMIZATION MODEL FOR MANAGING GROUNDWATER PUMPING- INDUCED R-A EXCHANGES	67
3.1 Introduction	67
3.2 Methods	68
3.2.1 Study area	68
3.2.1.1 <i>Hydrogeological settings</i>	70
3.2.1.2 <i>Groundwater and River stage-discharge</i>	72
3.2.1.3 <i>Recharge and Evapotranspiration</i>	73
3.2.1.4 <i>Agriculture and Land Use</i>	74
3.2.1.5 <i>Water demand and supply</i>	75
3.2.2 Numerical modeling	78
3.2.3 MODFLOW model	79
3.2.4 Model Conceptualization	80
3.2.5 Boundary Conditions	82
3.2.6 Sensitivity analysis	83
3.2.7 MT3DMS model	85
3.2.7.1 <i>Advection</i>	87
3.2.7.2 <i>Dispersion</i>	87
3.2.7.3 <i>Streamflow depletion and nitrate loading</i>	88
3.2.8 Model Calibration	90
3.2.9 Optimization Model	94

3.2.10 Multi-Objective Groundwater Simulation-Optimization.....	94
3.2.11 Metaheuristic algorithms.....	96
3.2.11.1 Multi-Objective Particle Swarm Optimization.....	97
3.2.11.2 Non-Dominated Sorting Genetic Algorithm - II.....	99
3.2.11.3 Multi-Objective Evolutionary Algorithm based on Decomposition.....	101
3.2.11.4 Pareto Search.....	103
3.2.11.5 Hyperparameter tuning.....	105
3.3 Linking groundwater model to optimization algorithms	105
3.3.1 PyGWMO module	105
3.3.1.1 Structure of PyGWMO.....	106
3.3.1.2 Software testing.....	110
3.4 Results	112
3.4.1 Flow budget and R-A Exchanges hotspots	113
3.4.2 Hydrogeological Factors Influencing Head and R-A Exchanges	118
3.5 Summary.....	121
CHAPTER 4: OPTIMAL GROUNDWATER MANAGEMENT STRATEGIES FOR IMPROVED R-A EXCHANGES	123
4.1 Introduction.....	123
4.2 Methods.....	125
4.2.1 Linked S-O-based management model	127
4.2.1.1 Objective functions	128
4.2.2 Pareto Performance Indicators	132
4.2.3 Optimal Management Scenarios and multi-criteria decision making.....	134
4.3 Results	136
4.3.1 Calibration and Validation Results.....	136
4.3.2 Pareto Optimal Trade-offs.....	137
4.3.3 Pareto Front Performance Evaluation	143

4.3.4 Reduced PF and decision-making.....	146
4.3.5 Management scenarios.....	147
4.4 Summary.....	150
CHAPTER 5: APPLICABILITY OF SURROGATE MODELS IN LARGE SCALE GW OPTIMIZATION.....	153
5.1 Introduction.....	153
5.2 Methods.....	155
5.2.1 Physics-Informed Neural Networks.....	156
5.2.2 Machine Learning techniques	157
5.2.2.1 <i>XGBoost Regressor</i>	157
5.2.2.2 <i>Support vector machine regressor</i>	158
5.2.2.3 <i>Artificial neural networks</i>	159
5.2.3 ANN- SO Model Development.....	160
5.3 Results	161
5.3.1 Head prediction.....	161
5.3.2 ANN-SO model.....	165
5.4 Summary.....	167
CHAPTER 6: EVALUATING DECISION VARIABLE IMPORTANCE IN LARGE SCALE GROUNDWATER OPTIMIZATION	169
6.1 Introduction.....	169
6.2 Methods.....	172
6.2.1 Data preprocessing.....	172
6.2.2 Statistical analysis.....	173
6.2.2.1 <i>Correlation analysis</i>	173
6.2.2.2 <i>Principal Component Analysis</i>	174
6.2.3 Explainable Artificial Intelligence Techniques	175
6.2.3.1 <i>Ensemble model development</i>	175
6.2.3.2 <i>SHAPely and LIME analysis</i>	176

6.2.4 Generalized Additive Modeling	178
6.3 Results and discussions.....	179
6.3.1 Correlation Analysis and Variable Loading	179
6.3.2 Variable importance	182
6.3.3 Variable dependency	185
6.3.4 Environmental and Development Zones.....	186
6.4 Summary.....	187
CHAPTER 7: DECISION VARIABLE REDUCTION IN LARGE SCALE MULTI- OBJECTIVE GROUNDWATER OPTIMIZATION PROBLEMS	189
7.1 Introduction.....	189
7.2 Methods.....	190
7.2.1 Clustering Algorithms	193
7.2.1.1 <i>K-means Clustering</i>	193
7.2.1.2 <i>Ward method</i>	194
7.2.1.3 <i>Affinity Propagation</i>	195
7.2.2 Clustering Performance Metrics	196
7.2.2.1 <i>Calinski Harabasz Value</i>	196
7.2.2.2 <i>Davis Bouldin Index</i>	197
7.2.2.3 <i>Silhouette index</i>	198
7.3 Case study.....	200
7.4 Results	200
7.5 Summary.....	207
CHAPTER 8: MANY OBJECTIVE OPTIMIZATION FOR RIVER-AQUIFER EXCHANGES: PERCEPTIONS FROM WATER-ENERGY-FOOD- ENVIRONMENT NEXUS	209
8.1 Introduction.....	209
8.2 Methods.....	210
8.2.1 WEFEnexus management model	212

8.2.1.1 <i>The Energy Objective</i>	212
8.2.1.2 <i>Crop yield and Economic benefit objective</i>	213
8.2.1.3 <i>Water Supply Objective</i>	214
8.2.1.4 <i>Environmental Objective</i>	215
8.2.1.5 <i>Constraints</i>	215
8.3 Results	217
8.3.1 Pareto optimal solutions.....	217
8.3.2 Correlation between WEFE nexus elements and R-A exchanges.....	220
8.3.2.1 <i>LeakageIn (Downwelling)</i>	221
8.3.2.2 <i>Leakage Out (Upwelling)</i>	221
8.3.3 Nitrate contamination and WEFE nexus.....	224
8.3.4 Decision-making for optimized WEFE objective.....	225
8.3.5 Scenario comparison.....	227
8.3.5.1 <i>Scenario 1 Low GW extraction with balanced R-A exchanges and nitrate load</i>	227
8.3.5.2 <i>Scenario 2 Moderate GW extraction with high R-A exchanges and high Nitrate load</i>	228
8.3.5.3 <i>Scenario 3 High GW extraction with emphasis on reducing Nitrate load</i> ..	228
8.4 Discussions	229
8.5 Summary	233
CHAPTER 9: CONCLUSION	235
9.1 Introduction	235
9.2 Specific conclusion of each objective	236
9.3 Research findings	238
9.3.1 Implications for Integrated Water Management	239
9.4 Limitations	240
9.5 Future scope	241
9.6 Concluding remarks	243

REFERENCES.....	245
PUBLICATIONS	267

LIST OF FIGURES

Figure 1.1 Conceptualization of (a) hydraulically disconnected from the regional groundwater system and (b) hydraulically connected to the regional groundwater system Source: Wilson et al. (2024)	24
Figure 2.1 A schematic representation of the simulation-optimization framework for groundwater management.	38
Figure 2.2 Large-scale variable reduction techniques	54
Figure 3.1 Lower Ain River basin with pumping wells, observation wells, and river stations.	69
Figure 3.2 Hydrogeological properties of the aquifers in the study area: Hydraulic conductivity (HK), specific yield (SY), aquifer thickness (thickness), saturated thickness (saturated thickness), transmissivity and vertical conductivity (VK).	71
Figure 3.3 Monthly discharge variability at three river stations (Chazey-Sur-Ain, Point de Ain, and Lagnieu) from 2002 to 2019.....	73
Figure 3.4 Workflow of evolutionary algorithms in groundwater S-O problems.....	78
Figure 3.5 A 3-D conceptualization of the groundwater model.....	81
Figure 3.6 Different domains conceptualized to study the effect of varying boundary conditions on GW flow budget and R-A exchanges.	83
Figure 3.7 GW calibration at 4 observation wells presents in the study area.....	93
Figure 3.8 Different statistical measures used in comparative analysis of the error across four observation wells.....	94
Figure 3.9 A schematic diagram of software structure and code dependency	107
Figure 3.10 PyGWMO graphical user interface	111
Figure 3.11 Groundwater head contours and Nitrate dispersion at the end of the simulation period.....	113
Figure 3.12 Flow Budget of lower Ain River basin. The colors represent different components of the flow budget, while the colored band shows the 90% confidence interval.	115
Figure 3.13 Variation of R-A exchange along the stretch of the river Ain at the simulation's start, mid, and end.	92

Figure 3.14 Distribution of R-A exchange metrics at the simulation's start, mid, and end.....	116
Figure 3.15 Three zones of R-A exchanges and their temporal variation at the simulation's start, mid, and end.....	118
Figure 3.16 First-order sensitivity indices with 90% Confidence Interval.....	119
Figure 3.17 Sobol first-order sensitivity indices for various input aquifer parameters across three different model domains	119
Figure 3.18 Sobol T indices for various input aquifer parameters across three model domains.	121
Figure 4.1 Methodology chart of the S-O framework	126
Figure 4.2 Pseudo code for S-O model.....	128
Figure 4.3 Calibration results for 4 observation wells.....	137
Figure 4.4 All feasible and nondominated solutions obtained by NSGA-II and MOPSO.....	138
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.	141
Figure 4.6 Decision Variable distribution in model 1 of NSGA-II, MOEA/D, and MOPSO algorithm	143
Figure 4.7 Values of pareto front evaluation parameters	144
Figure 4.8 Normalized radar plot for pareto performance indicators for model 1	145
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.....	147
Figure 4.10 The GW pumping rates for the management period for the chosen solution.....	149
Figure 5.1 Types of surrogates used for groundwater optimization	155
Figure 5.2 PINN architecture.....	156
Figure 5.3 RMSE loss over epochs for the ANN model.....	162
Figure 5.4 Predicted head values by PINN with 2000 and 5000 epochs at three-time steps. MODFLOW results are presented in the third row.....	164
Figure 5.5 Predicted head values by SVM, XgBoost, and ANN.....	165
Figure 5.6 Pareto performance metrics after introducing ANN-SO	166

Figure 5.7 Percentage (a) increase in hypervolume (b) decrease in IGD after introduction of ANN	166
Figure 5.8 Pareto fronts for ANN-SO	167
Figure 6.1 Correlation distribution among decision variables.....	180
Figure 6.2 Correlation between decision variables and objective functions	180
Figure 6.3 Clustered heatmap of correlations	181
Figure 6.4 Variable loading plot using PCA of decision variables	182
Figure 6.5 (A) Training and validation loss over the epochs (B) mean absolute weights for all the variables in the algorithm and (C) LIME weights	183
Figure 6.6 SHAP value plot for each decision variable in two objectives.....	184
Figure 6.7 Partial dependency plot of decision variables for each objective.....	186
Figure 6.8 Hotspot analysis based on SHAP, LIME, variable loading, and partial dependency.....	187
Figure 7.1 Schematic diagram of the workflow.....	192
Figure 7.2 Case study of Lower Ain River basin.....	200
Figure 7.3 Optimal number of clusters using three clustering metrics. Figure A and B represents the 40 well clusters by K-means and Ward algorithm.....	201
Figure 7.4 Percentage (a) increase in hypervolume (b) decrease in IGD, after introduction of aquifer parameter-based clustering.	202
Figure 7.5 Variation of Pareto performance metrics for different clustering algorithms after decision variable reduction.	203
Figure 7.6 Pareto fronts obtained after convergence of S-O model based upon aquifer parameter clustering. The X-axis and Y-axis represent the objective functions 1 and 2 respectively.	204
Figure 7.7 Differences in total discharge of all communes with and without aquifer parameter-based clustering.....	205
Figure 8.1 Parallel coordinate plot showing the conflicting objectives with the three indicators of LeakageIn, leakage out and total conc. The color of the solution represents the downwelling within the LARB.....	219
Figure 8.2 The trade-offs among three objectives of the WEFEX nexus in the scatterplot while the fourth objective of total concentration loading is denoted by the color. The pareto optimal solutions represent the different synergies. The arrows represent the direction preferable for each objective.....	219

Figure 8.3 Variation of the decision variables. The pumping wells for each commune represent the corresponding discharge decision variables marked by box. The colors are for representation purpose only	220
Figure 8.4 The correlation heatmap of all the objectives and the R-A exchanges parameters (LeakageIn and Leakage out)	222
Figure 8.5 Scatterplot representing the conflict between energy cost objective and water supplied categorized by LeakageIn and the colors represent the leakage out categorization.....	223
Figure 8.6 Parallel plot representing all the four objectives of the WEFE nexus and their categorization based on the total concentration load.....	225
Figure 8.7 The 2D contour plot representing the density of non-dominated solution for objective functions. The color bar represents the density of solutions obtained during optimization problem.....	226
Figure 8.8 Spatial distribution of decision variables obtained from the 10 solutions within LARB.....	229
Figure 8.9 Stacked bar plot of selected solutions	229

LIST OF TABLES

Table 2.1 Types of the Simulation model for groundwater and R-A exchanges modeling.....	43
Table 2.2 Different types of metaheuristic algorithms used in S-O.....	48
Table 2.3 Different objective functions and decision variables in large-scale S-O Problems.....	56
Table 3.1 Different Crop Types and GW Demand in the Lower Ain River Basin.....	76
Table 3.2 Different properties used for the simulation model	82
Table 4.1 Pareto front performance indicators	133
Table 4.2 Pareto performance indicators	145
Table 5.1 Summary statistics of GW prediction dataset	155
Table 5.2 ANN optimization results for different combinations of layer size for all three domains. The highlighted row indicates the best parameters for the ANN model.	160
Table 5.3 RMSE table for different algorithms.....	162
Table 7.1 The total combination of S-O models run.....	195
Table 7.2 The commune-wise discharge is calculated before and after parameter clustering.....	206
Table 8.1 The control parameters of AMOPSO for the optimization process	211
Table 8.2 Crop pattern in agricultural zones with constants	213
Table 8.3 Different discretization of solutions for R-A exchanges and total nitrate concentration and their range	223
Table 8.4 Table represents the 10 solutions obtained by uniform sampling and weighted TOPSIS method.....	227

ACRONYMS

GW: Groundwater
HK: Hydraulic Conductivity
LARB: Lower Ain River Basin
GMS: Groundwater Modeling System
R-A Exchanges: River-Aquifer Exchanges
WEFE nexus: water, energy, food, and environment nexus
SDGs: Sustainable Development Goals
S-O: Simulation-Optimization
PSO: Particle Swarm Optimization
MOPSO: Multi-Objective Particle Swarm Optimization
NSGA-II: Non-dominated Sorting Genetic Algorithm II
MOEA/D: Multi-Objective Evolutionary Algorithm Based on Decomposition
SVM: Support Vector Machines
SVR: Support Vector regression
RBF: Radial Basis Function
MLP: multi-layer Perceptron
PyGWMO: Python-Based Groundwater Many Objective Optimization Module
GUI: Graphical User Interface
PINNs: Physics-informed neural networks
DeepONet: Deep Neural Operator
SSE: Sum of Squared Error
VRC: Variance Ratio Criterion
FDM: Finite Difference Method
AEM: Analytical Element Method
FVM: Finite Volume Technique
BEM: Boundary Element Method
FEM: Finite Element Method
RM: Response matrix

EA: Evolutionary algorithm
DV(s): Decision Variable(s)
MOPs: Multi-objective problems
MaOPs: Many Objective Problems
Many-LSOPs: Large-Scale Many Objective Problems
GLUE: Generalized Likelihood Uncertainty Estimation
MCMC: Markov Chain Monte Carlo
BaRE: Bayesian Recursive Estimation
PDF: probability density function
TOPSIS: Technique for Order Preference by Similarity to Ideal Solution
LINMAP: Linear Programming Technique for Multi-dimensional Analysis of Preference
ELECTRE III: Elimination and Choice Translating Priority III
xAI: Explainable Artificial Intelligence
SHAP: SHapley Additive exPlanations
LIME: Local Interpretable Model-agnostic Explanations
GAMs: Generalized Additive Models
PCA: Principal Component Analysis
SOM: Self-organizing maps
RMSE: Root Mean Square Error
R2: R-squared
MAPE: Mean Absolute Percentage Error
MXAE: Maximum Absolute Error
AMOPSO: Adaptive Multi-Objective Particle Swarm Optimizer
PDEs: partial differential equations
AD: automatic differentiation
FD: Finite-difference
TVD: total variation diminishing
STR: stream package