Groundwater management in megacities: Advancing multi-objective decision analysis under model uncertainty and spatially distributed impacts

By

MARINA REYES LOPEZ MAUTNER

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Hydrologic Science

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Jonathan D. Herman, Chair

Jay R. Lund

Thomas Harter

Committee in Charge

2022

Contents

List of Figures vi

List of Tables xi

Abstract xii

Acknowledgements xiv

Preface xvii

1 Introduction 1
1.1 Background and Motivation . . . 1
1.2 Research Objectives . . . 3
Bibliography . . . 5

2 Urban growth and groundwater sustainability: evaluating spatially distributed recharge alternatives in the Mexico City Metropolitan Area 6
2.1 Abstract . . . 6
2.2 Introduction . . . 7
2.3 Case Study . . . 10
2.4 Methods . . . 12
2.4.1 Groundwater Model Inputs . . . 12
2.4.1.1 Model Layers and Hydrostratigraphic Units . . . 14
2.4.1.2 Active Model Extent . . . 15
2.4.1.3 Precipitation . . . 15
2.4.1.4 Land Use . . . 15
2.4.1.5 Water Supply Distribution System Leaks . . . 16
2.4.1.6 Well Pumping . . . 18
2.4.2 Sensitivity Analysis and Model Calibration . . . 19
2.4.2.1 Observations . . . 20
2.4.2.2 Parameters . . . 22
2.4.3 Application of Management Alternatives . . . 22
2.4.3.1 Infiltration Basins . . . 24
2.4.3.2 Wastewater Reuse at Existing Treatment Plants . . . 25
2.4.3.3 Repair to Leaks in Potable Supply Distribution Network . . . 25

2.4.4 Calculation of Planning Objectives . . . 26
2.4.4.1 Minimize Energy Use . . . 27
2.4.4.2 Minimize Water Quality Risks in Subsidence-Prone Areas . . . 28
2.4.4.3 Minimize Groundwater Flooding in Urban Areas . . . 28
2.5 Results and Discussion . . . 29
2.5.1 Calibration and Parameter Sensitivities . . . 29
2.5.2 Simulation Agreement with Historical Observations . . . 31
2.5.3 Performance of Aquifer Management Alternatives . . . 33
2.5.4 Model Uncertainty . . . 36
2.6 Conclusions . . . 39
2.7 Appendix . . . 40
2.7.1 Land Use Determination . . . 40
2.7.2 Selected Infiltration Basins . . . 43
2.7.3 Leak Calculations . . . 45
2.7.3.1 Estimated non-groundwater supply sources . . . 45
2.7.3.2 Derivation of new groundwater pumping relationship from repair leaks scenario . . . 45
2.7.4 Model Parameters and their Calibrated Values . . . 45
2.7.5 Simulated and Observed Hydraulic Head . . . 47
2.8 Data Availability . . . 48
2.9 Acknowledgments . . . 48
Bibliography . . . 60

3 Coupled effects of observation and parameter uncertainty on urban groundwater infrastructure decisions 61
3.1 Abstract . . . 61
3.2 Introduction . . . 62
3.3 Methodology . . . 66
3.3.1 Urban Groundwater Model . . . 66
3.3.2 Uncertain Parameters . . . 68
3.3.3 Spatially Clustered Observations . . . 69
3.3.4 Model Error . . . 71
3.3.5 Parameter Set Selection . . . 71
3.3.6 Sensitivity Analysis . . . 72
3.3.7 Evaluation of Decision Uncertainty . . . 73
3.4 Results and Discussion . . . 76
3.4.1 Cluster Behavioral Parameter Sets . . . 76
3.4.2 Sensitivity of Error Metric and Management Objectives . . . 79
3.4.3 Decisions Under Observational Uncertainty . . . 82
3.4.4 Limitations and Future Work . . . 87
3.5 Conclusions . . . 89
3.6 Appendix . . . 90
3.7 Data Availability . . . 94
3.8 Acknowledgements . . . 94
Bibliography . . . 99

4 Socially informed spatial analysis: evaluating the role of aggregation scale in modeling differential impacts of urban groundwater pumping policies across socioeconomic indicators **100**
4.1 Abstract . . . 100
4.2 Introduction . . . 101
4.3 Methodology . . . 106
4.3.1 Study Area . . . 106
4.3.2 Urban Groundwater Model . . . 107
4.3.3 Pumping Alternatives . . . 108
4.3.4 Socioeconomic Indicators . . . 109
4.3.5 Spatial resolution . . . 111
4.3.6 Objectives . . . 112
4.3.7 Preferred policies . . . 115
4.4 Results . . . 116
4.4.1 Spatial resolution . . . 116
4.4.2 Policy performance under management objectives . . . 118
4.4.3 Alternative preference . . . 120
4.4.4 Limitations and future work . . . 126
4.5 Conclusion . . . 127
4.6 Appendix . . . 129
4.7 Data Availability . . . 132
4.8 Acknowledgements . . . 132
Bibliography . . . 136

5 Conclusion **137**

List of Figures

2.1 The Mexico City Metropolitan Area (MCMA) is made up of urban areas in Mexico City and the State of Mexico. The model area is shown in black with population density shown by municipality. 10

2.2 Hydrologic components and flows in the Mexico City Metropolitan Area with historical flows in blue and managed aquifer recharge alternative flows in green. The components of the groundwater model developed in this study include fixed inputs (blue), calibrated inputs (yellow), recharge alternatives (green), and model output (orange), with processes not modeled in this study shown in gray. 11

2.3 Workflow and datasets used in the 1. Pre-processing of model inputs, 2. Sensitivity analysis and calibration of the model, 3. Application of the management alternatives, and 4. Calculation of planning objectives. System processes are defined by both datasets (italics), and the MODFLOW packages or software (bold) used where appropriate. 13

2.4 Input data for the numerical groundwater model: a) geologic formations with overlain lacustrine layer and side view of A-A', b) 2005 pumping data from the Public Register of Water Rights (REPDA), and Mexico City Water System and State of Mexico service areas, c) adapted INEGI and Landsat datasets showing urban, natural, and wetland/water land use types for 1985 and 2015, d) average July meteorological station data during the model period interpolated over the model area. 17

2.5 Aquifer management alternatives evaluated: a) implementation of five infiltration basins (INEGI, 2010), b) repair of 20% of potable supply leaks (DGCOH, 1997), and c) increased recharge at all wastewater treatment plants (Riveros-Olivares, 2013; CONAGUA, 2015) 23

2.6 Spatial set of cells or points measured in each objective: pumping wells for energy (black), lacustrine layer for water quality (red), and urban cells for flooding (blue). . 27

2.7 Composite scaled sensitivity values of 15 most sensitive parameters, out of a total of 33 adjustable model parameters: zonal geologic (orange), time varied infrastructure (blue), zonal recharge (green), and leak infiltration (gray). Calculated for a) stage 1 of the calibration process with a confined system, b) stage 2 of the calibration process with a convertible confined/unconfined system with wetting, and c) final calibrated parameters. 31

2.8 Selected hydrographs of simulated and observed change in hydraulic head over time for representative observations from the MODFLOW model using calibrated parameters. 32

2.9 Change in head for each aquifer management alternative as difference from historical change in head over the simulation period of 1984 to 2013. 34
2.10 Performance of aquifer management alternatives a) according to cumulative change in storage over time, and b) under three spatially aggregated objectives: minimize pumping energy use, minimize depth to groundwater in clay layer, and minimize urban groundwater mounding. 37
2.11 (a) Landsat image from the Polanco neighborhood of Mexico City; (b) final shape representing natural and urban land uses from the three data sources: INEGI land use, time specific Landsat imagery, and INEGI topographically determined greenspace; (c) percentage of natural land use type by model cell. 41
2.12 Cell by cell percentages for (a) urban, (b) water, and (c) natural land use types for 1990 42
2.13 Proposed location for infiltration basin 1, randomly selected from all potential basin locations discussed in Section 2.4.3.1 Infiltration Basins. 43
2.14 Proposed location for infiltration basin 2, randomly selected from all potential basin locations discussed in Section 2.4.3.1 Infiltration Basins. 43
2.15 Proposed location for infiltration basin 3, randomly selected from all potential basin locations discussed in Section 2.4.3.1 Infiltration Basins. 44
2.16 Proposed location for infiltration basin 4, randomly selected from all potential basin locations discussed in Section 2.4.3.1 Infiltration Basins. 44
2.17 Proposed location for infiltration basin 5, randomly selected from all potential basin locations discussed in Section 2.4.3.1 Infiltration Basins. 44
2.18 First simulated head for each observation well versus first observed head for each observation well. The shaded regions represent the 95% confidence interval for the linear regression performed. 47
2.19 Simulated drawdown for each observation well versus observed drawdown for each observation well. The shaded regions represent the 95% confidence interval for the linear regression performed. 48

3.1 Flowchart of methods. 65
3.2 (Top) A 3-dimensional visualization of the 5 clusters of observations used in this study. (Bottom left) Observation clusters shown with the geologic formations within the model area. (Bottom right) Observation clusters shown with the land use types for the model period covering 2010. 70
3.3 The distributions along the parameter ranges of the filtered samples using the sum of squared error metric. The distributions are colored according to the observation cluster used to filter the dataset. The prior distribution (not shown) is uniform for all parameters. Parameter abbreviations given in Table 3.1. 75

3.4 A representative view of the four model output metrics for the historical alternative, plotted against the parameter range for the hydraulic conductivity of the alluvial formation (the most sensitive parameter from Figure 3). These include the error metric (sum of squared weighted residuals in m^2), energy objective (kWh), water quality risk objective (percent of cells not meeting the objective), and urban flooding objective (percent of cells not meeting the objective). Gray points represent all parameter sets, while colors represent behavioral parameter sets meeting the error threshold. 77
3.5 δ sensitivity of the energy objective according to the 5,000 filtered samples for the 33 model parameters (columns). The sensitivity is shown by cluster (rows) and by the four alternatives from left to right (light to dark): historical, wastewater reuse, infiltration basins, and repair leaks. The bootstrapped 95% confidence interval for each sensitivity value is shown as a red line. 79
3.6 δ sensitivity of the error metric and three management objectives (rows) according to the 5,000 filtered samples for the 8 model parameters (columns) with the largest differences in sensitivity between clusters for the historical management alternative. The sensitivity is shown by cluster in order from left to right: C-00001, C-00002, C-00003, C-00004, C-00005, C-12345. 81
3.7 Alternative performance across the observation cluster parameter sets shown as heatmaps of the count of sets where the alternative performance was ranked as (1) best to (4) worst. Within each heatmap, the rows are the rank and the columns are the cluster behavioral parameter sets. The subplots are organized by the three management objectives as the rows and the aquifer management alternatives as columns. 83
3.8 Alternative performance across the parameter range of the alluvial hydraulic conductivity (one of the most sensitive parameters) shown as heatmaps of the count of sets where the alternative performance was ranked as (1) best to (4) worst. Within each heatmap, the rows are the rank and the columns are the parameter value from minimum (1.00E-1) to maximum (1.00E+2). The subplots are organized by the three management objectives as the rows and the aquifer management alternatives as columns. 84
3.9 Alternative performance in the water quality objective. Shown as heatmaps of the count of parameter sets where the alternative performance was ranked as (1) best to (4) worst. Within each heatmap, the rows are the rank and the columns are the parameter value from minimum (1.00E-1) to maximum (1.00E+2). The subplots are organized by the observation cluster used for behavioral parameter set selection as the rows and the aquifer management alternatives as columns. 85
3.10 The difficulty of the decision represented by the relative performance of the alternatives within the samples evaluated for each objective (columns). The top row shows the distribution of the percent difference in each sample between the 1st and 2nd ranked alternatives within the cluster datasets. The bottom row shows the distribution of the percent difference in each sample between the 1st and 4th ranked alternatives within the cluster datasets. 87

3.11 A representative view of the four model output metrics for the historical alternative, plotted against the parameter range for the vertical anisotropy of the hydraulic conductivity of the alluvial formation. These include the error metric (sum of squared weighted residuals), energy objective (kWh), water quality risk objective (percent of cells not meeting the objective), and urban flooding objective (percent of cells not meeting the objective). Gray points represent all parameter sets, while colors represent behavioral parameter sets meeting the error threshold. 91
3.12 δ sensitivity of the water quality risk objective according to the 5,000 filtered samples for the 33 model parameters (columns). The sensitivity is shown by cluster (rows) and by the four alternatives from left to right (light to dark): historical, wastewater reuse, infiltration basins, and repair leaks. 92
3.13 δ sensitivity of the urban flooding risk objective according to the 5,000 filtered samples for the 33 model parameters (columns). The sensitivity is shown by cluster (rows) and by the four alternatives from left to right (light to dark): historical, wastewater reuse, infiltration basins, and repair leaks. 93

4.1 Problem formulation and summary of socially informed spatial analysis implemented. 105
4.2 Pumping quantities (blue circles) of all known pumping wells in the Valley of Mexico shown as either the average over the model period (1984-2013) for municipal wells or the concession limit for private wells. Additionally, municipal pumping wells used in the aquifer management alternatives separated into four groups of approximately equal pumping quantities are indicated by colored diamonds. 107
4.3 Water access indicator for each of the municipality units (polygons) and centroids of the census block units (points). Darker indicates a lower fraction of households with piped water within the dwelling unit. 109
4.4 Close-up of census block polygons, rural census block points, and numerical groundwater model 500x500 meter grid. 111
4.5 spatial units for the municipal, cluster, and census block scales. 116
4.6 The four social indicators (education, health, water access, and poverty) determined at the cluster scale. 117
4.7 The water quality objective plotted against the health indicator (percentage of the population with public or private health insurance) for the five pumping policies compared to the historical management alternative. 118
4.8 The urban flooding objective plotted against the poverty indicator (percentage of the population with an income above the level necessary to satisfy basic needs) for the five pumping policies compared to the historical management alternative as a baseline. 119
4.9 The groundwater availability objective plotted against the water access indicator (percentage of households with piped water inside the dwelling unit) for the five pumping policies compared to the historical management alternative as a baseline. . 120
4.10 Nondominated solutions viewed at the three spatial unit scales as three categories: units with a single nondominated policy, units with multiple nondominated policy, or units with all policies among the nondominated solutions. 123

4.11 Spatial distribution of the spatial units at the cluster scale that contain the indicated policy within the nondominated solutions. A layer with all units is shown below the spatial extent of the preference of each of the policies for easier comparison. The P2 policy (black) is nondominated across all spatial units, leaving no gray visible. 124
4.12 Health indicator for each of the municipality units (polygons) and centroids of the census block units (points). Darker indicates a smaller fraction of the population with private or public health insurance. 129
4.13 Education indicator for each of the municipality units (polygons) and centroids of the census block units (points). Darker indicates lower average education level among the adult population. 130
4.14 Poverty indicator for each of the municipality units. Darker indicates a lower fraction of the population above the threshold for basic needs used to indicate poverty in Mexico. 130
4.15 Spatial distribution of the spatial units at the census block scale that contain the indicated policy within the nondominated solutions. A layer with all units is shown below the spatial extent of the preference of each of the policies for easier comparison. The P2 policy (black) is nondominated across all spatial units, leaving no gray visible. 131
4.16 Spatial distribution of the spatial units at the municipal scale that contain the indicated policy within the nondominated solutions. A layer with all units is shown below the spatial extent of the preference of each of the policies for easier comparison. The P2 policy (black) is nondominated across all spatial units, leaving no gray visible. 131

List of Tables

2.1 Model Hydrostratigraphic Units 14
2.2 Correlation coefficients for parameters that were chosen for calibration according to composite scaled sensitivity values. 30
2.3 Aquifer management alternative increase in recharge 33
2.4 Estimated values for non-groundwater supply 45
2.5 Model Parameters 45
2.6 Mean absolute error (MAE) between weighted and unweighted simulated and observed hydraulic head at observation wells for each geologic formation 47

3.1 Model Parameters and Sampling Ranges 67

4.1 Proportion of average pumping attributed to each pumping group P1-4 out of the total reported municipal pumping over each of the three model phases. 108
4.2 Objective values measured at the regional scale. 120
4.3 Summary of nondominated alternatives by spatial unit scale. H indicates the historical alternative, P1-4 are the pumping policies where groups 1-4 are eliminated from pumping, and AllP is the pumping policy where pumping is reduced in all four groups by 25%. The percentage of units that contain the policy within the nondominated solutions is shown out of the total number of spatial units by scale of aggregation. 121
4.4 Percentage of spatial units within the first quartile of each social indicator that contain the policy within the nondominated solution set for that unit shown for each spatial unit scale. On the left hand side, a value of 100% indicates that all spatial units that are within the lowest quartile for that socioeconomic indicator contain the policy within the nondominated solution set. On the right hand side is the difference between the values on the left hand side and those in Table 4.3, representing the preference rates of the full set of spatial units. Positive values are shown in blue, while negative values are show in red. 125

Abstract

Groundwater dependent mega-cities exert enormous pressure on the watersheds in which they are situated, creating a demand for complex hydrogeologic modeling and analysis that provides decision-makers with the information they need. However, uncertainties about the characterization of the subsurface, hydrologic fluxes, and other model input, along with assumptions made during the modeling process can create difficulties in developing the models and interpreting the results in a way that ensures sustainable and equitable regional aquifer management. This dissertation uses a multiobjective analysis approach to determine the performance of spatially distributed aquifer management alternatives for the case study of the Valley of Mexico, where the Mexico City Metropolitan Area is situated. First, a three-dimensional finite-difference groundwater model is developed alongside an initial set of managed aquifer recharge alternatives and planning objectives in Chapter 2. The management alternatives tested show spatially distinct reductions in drawdown over the historical period for equivalent changes in storage and that combining multiple alternatives results in a more than an additive risk of groundwater flooding. These results point to the importance of including planning objectives calculated over diverse spatial extents in capturing impacts to groundwater security in urban basins. To address the potential for endogenous uncertainties in models to influence decision-making, Chapter 3 carries out a global sensitivity analysis across model parameter values and subsets of well observations using the model and alternatives developed in Chapter 2. Error metrics generally used to calibrate groundwater models are found to be sensitive to distinct parameters from those to which management objectives are sensitive and the coupled effects of the endogenous uncertainties have amplifying effects on the ranking of management alternatives. Both of these findings highlight the importance of performing sensitivity analyses that are carried through to the decision-making stage and not just for calibration purposes. Finally, Chapter 4 builds on the findings from Chapters 2 and 3 first by testing improvements on the best performing alternative from the previous chapters and second by examining the limitations of regional aggregation of management objectives in providing equitable groundwater supply planning. A method is developed to alleviate scale issues that are found to arise when attempting to use existing subregional spatial units to evaluate socioeconomic indicators and

groundwater management objectives simultaneously. Findings suggest that the most marginalized spatial units experience deviations from the more advantaged spatial units in performance of groundwater pumping policies, and can diverge from both the regional preference regime and the preference regime of the full set of spatial units depending on the scale of the spatial unit definition. Overall, this dissertation develops methods to represent anthropogenic impacts, assess aquifer recharge policies, and evaluate the effects of model uncertainties on decisions at the regional scale to improve the tools available needed to approach the challenges inherent in long-term urban groundwater supply planning.

Acknowledgements

First and foremost, a huge thanks to Jon Herman for the countless hours he dedicated to advising and supporting me through this journey. No matter what, Jon always had my back and worked to make sure that I had the resources that I needed. Whether it was meeting (excessively) tight deadlines for just-out-of-reach grant applications, counseling me when I was at my lows, or preparing material to teach our research group the hidden graduate school curriculum, I knew I could always ask him about what was worrying me, which was usually quite a lot. Jon cultivated a hub for novel ideas by putting together an all-star research group that I am so proud to have been a part of and challenged me to take on a vast array of complex concepts. A special thanks to Jay Lund for creating an open atmosphere in his lab group and courses that always made me and my ideas feel welcome and appreciated. Thank you to Carlos Puente (and Marta!) and Laura Foglia (and the girls!) for making my time living in Davis full of late evenings, good people, and great food. I had a great experience working with many professors at UC Davis, and would like to recognize Helen Dahlke for her valuable input on our book chapter that made me a better writer and a better academic, and to Sam Sandoval Solis, gracias por ser el primero en darme la bienvenida a Davis. Many thanks to my dissertation and qualifying exam committee members Laura Foglia, Josué Medellín Azuara, and Thomas Harter for their feedback and commitment to my formation as an academic. I would be remiss to leave out Brad Barrett and Joe Kasprzyk who taught me that caring about grad students as people is cool and that being an academic is for the cool kids too.

My path to the PhD was influenced by many mentors along the way, including Alma Chávez Mejía, Heather Pohl, Christy Kennedy, and Kara Nelson, all admirable professionals and academics who also happen to be women who are at the forefront of their areas of expertise and who showed me what was achievable. Completing this degree would not have been possible without my funding sources, the UC Davis Center for Watershed Sciences, Department of Civil and Environmental Engineering Johannes “Joe” DeVries Graduate Student Award, Henry A. Jastro Graduate Research Award, and the Ford Foundation Predoctoral Fellowship Program of the National Academies of Science, Engineering, and Medicine. Additionally, this research could not have been completed without collaboration from contacts at the Organismo

de Cuencas: Aguas del Valle de México (OCAVM) of the Comisión Nacional del Agua (CONAGUA), and the Instituto de Geofísica and the Instituto de Ingeniería of the Universidad Nacional Autónoma de México (UNAM), especially Graciela Herrera and Rosa Galán.

Next, what is grad school if not a practice in teamwork with your fellow grad students during times of extreme stress and overwhelming accomplishment? I would most certainly not have even applied for this PhD without the encouragement and pathbreaking of Alejandro Schuler da Costa Ferro and Gabi Kirk, who have been my peers the longest. They both set a high standard for what it means to practice and promote solidarity among student workers and are both going to make incredible professors. Of course, making this journey would have been a hell of a lot harder without Samira Ismaili and Claire Kouba along for the ride, both who made me want to be a better student and better friend. I could write another thesis for all the thanks I have for my many peers, but a list will have to suffice here, thank you to the students of Ghausi 1011, Jon's research group, and the groundwater peer group, especially, Nusrat Molla, Liam Ekblad, Aaron Alexander, Beth Robinson, Natalie Mall, Asli Kol, Janis Patino, Arturo Palomino, Gaby Castrellon Romero, Fede Zabaleta, Yara Pasner, Joaquin Meza, Darcy Bostic, Jason Wiener, Gus Tolley, Sebastien Poore, Noelle Patterson, and Bill Rice. And a super big thank you to Jon Cohen, Rich Pauloo, and Stephen Maples, whose qualifying exam and dissertation materials, and boundless camaraderie, have been invaluable. Thank you to the students in the Land, Air, and Water Resources Diversity, Equity, and Inclusion Graduate Student Subcommittee, and UC Davis Student Researchers United, who I have had the privilege of organizing alongside to make our university a better place for graduate students. Finally, I am very grateful for my non-UC peers Dean Chahim and Lorelay Mendoza, I cannot wait to continue exploring and bettering this field of water resources engineering together.

As an extracurricular addict, I would like to thank those who have contributed to my non-dissertation activities. I am incredibly thankful for my experience at the Stockholm Environment Institute, which saved me from my mid-PhD slump by giving me purpose and deadlines, and teaching me the art of being concise. I would especially like to thank Marisa Escobar for giving me an opportunity to join the SEI family, and Laura Forni for her hours of support through our many projects and my life's ups and downs. A heartfelt thanks to Nancy Erbstein and

Chuck Walker, who supported me in my interdisciplinary academic pursuits and allowed me to continue my parallel path into the world of global education and human rights. So many thanks to all the students in Aggies Near and Far and the Global Education for All fellows program who put up with our endless meetings and all my emails to create our one-of-a-kind podcast.

To my many friends and family over the last five years who have helped me step outside the academic realm for shared meals, shared misery, and shared fun, you will never know how much you helped. In no particular order muchisimas gracias a Karen Hernandez, Alissa Koeppe, Dafne Uscunga, Abi Muñoz, Nidia Bautista, Duja Michael, Elias Izpisua Rodriguez, Margaret Doyle, Kati Ventura, and the Cool Kids. Thank you since birth to all my aunts, uncles, cousins, cousins' cousins, in-laws, and family friends who make this world worth saving.

Lastly, thank you to my namesake and all around firecracker, my Nana Adelina "Nina" Ruiz, for always approaching life with love and laughs, even in the face of such adversity. Equally beloved, my grandma Marjie Mautner, who is the reason for my love of reading and passion for justice, and who has been a blessing of a roommate over the last two years. I wish Tata and grandpa Fred could have been here to share in my latest accomplishments, they were both so important in all our lives. Thank you to my dad who started me on this path by being the first to show me that grad school was possible and the first to teach me a for loop. I will forever be grateful for how he has always encouraged me to study hard and mosh harder. My mom is the best professor I know and worked damn hard to be one. The first in her family to go to college, the most caring and knowledgeable nurse I know, my mom is the person I call when I need a second opinion and I am so proud to be her daughter. Thank you to Katrina for teaching me how to read, being the rock that settles this crazy family, and always getting me, even before I even get myself. Thank you to Allegra for sharing the womb and being the best friend a person could have. Thank you Lee, Alejandro, Atalanta, and Camilo for making our family complete. Omar Reyes Lopez, gracias por los nombres extras, las vueltas al mundo conmigo, por ser el tutor de español más dedicado que he tenido, y por todos los años que vienen.

Preface

About this dissertation

The idea for this dissertation began in discussion with colleagues at the Instituto de Ingeniería of the Universidad Nacional Autónoma de México (UNAM) while completing a research stay there from 2015 to 2017. The connections developed in Mexico City during the 9th International Symposium on Managed Aquifer Recharge in 2016 and furthered at numerous international conferences, organized research trips, and virtual encounters since, have been absolutely vital to the final product recorded here.

Notes on Chapter 2

Chapter 2 develops a physically-based groundwater model that incorporates land use change, water supply infrastructure, and pumping trends for the Mexico City Metropolitan Area (MCMA) as a case study to examine the effects of the choice of spatial and temporal dimensions for the calculation of objectives on preferred aquifer management alternatives. It was coauthored with Jonathan Herman, Laura Foglia, Graciela Herrera, and Rosa Galán. Chapter 1 was published in a special issue on "Water Scarcity, Security and Sustainability" in *Journal of Hydrology*, (https://doi.org/10.1016/j.jhydrol.2020.124909).

Notes on Chapter 3

Chapter 3 implements a global sensitivity analysis of the groundwater model developed in Chapter 2 and was coauthored with Jonathan Herman and Laura Foglia. Chapter 3 was published in *Hydrology and Earth System Sciences* (https://doi.org/10.5194/hess-26-1319-2022).

Notes on Chapter 4

Chapter 4 explores the social implications of the scale over which groundwater pumping policies are evaluated and proposes a novel method for adapting heterogenous socioecnomic indicator data for use in analysis of groundwater model output. Chapter 4 is coauthored with Jonathan Herman and is currently in preparation for submission to *Journal of Water Resources Planning and Management*.

Chapter 1

Introduction

1.1 Background and Motivation

Throughout history attempts to tame the chaotic nature of urban hydrogeological systems have largely been met with failure, whether through the construction of flood control infrastructure and the draining of lakes and wetlands that have yet to keep our cities dry, or through the expanding exploitation of the blue gold flowing in our rivers or buried beneath our feet that has yet to quench our ever-growing thirst (Reisner, 1993; Orsi, 2005; Nelson, 2017). Today, more than 30 cities worldwide are characterized as "megacities", with over 10 million inhabitants, and over half of the world's population lives in cities of all sizes (United Nations, 2016). Megacities pose particular threats to the efficient and equitable management of surface and groundwater supply quantity and quality, wastewater treatment, and stormwater management resulting from their size, fast growth, inter-sectoral lack of adequate infrastructure, and diverse users; however, they also represent an opportunity for incredible innovation as they serve as economic, political, and technological centers (Howard and Gelo, 2003; Varis et al., 2006; Foster et al., 1998). Urban hydrogeology research during the second half of the 20th century has focused on characterizing the processes typical of such systems including the description of fluxes such as leakage from and into water distribution and sewage networks, abstraction from public and private wells, and irrigation of landscaping and agriculture; system structure such as the hardening of the ground surface with pavement and the creation of preferential flow paths through "urban karst" from water infrastructure in the subsurface;

biogeochemical setting such as sources and sinks of contaminants from industrial, domestic, and agricultural activities; and negative consequences of unregulated or underregulated groundwater exploitation such as subsidence, drying of wells, and damage to underground structures due to mounding from artificial recharge (Lerner, 1990; Schirmer et al., 2013; Bonneau et al., 2017; Wittenberg and Aksoy, 2010; Osmanoğlu et al., 2011; Ohgaki et al., 2007; Vázquez-Suñé et al., 2005). As our understanding of the threats to sustainable groundwater supply in megacities has grown, so to has the body of literature developing potential solutions to these consequences and simulation modeling tools to test such solutions.

Managed aquifer recharge has been promoted as a way to diversify water sources in urban areas and address the challenges created by the pressures exerted by megacities (Page et al., 2018; Dillon et al., 2019). In particular, centralized options such as infiltration basins, conjunctive use, in-lieu recharge, and infiltration or injection of treated wastewater, as well as decentralized options such as low impact development and capture and reuse projects are well adapted for urban settings and have shown promise in cities in California, Australia, and Mexico, among others (Dillon et al., 2010; Dahlke et al., 2018; Cruz-Ayala and Megdal, 2020). However, the implementation of MAR can be costly and slow, and the regional impacts of such measures uncertain, creating a need for hydrogeologic simulation modeling to develop optimal spatial and temporal implementation of such actions and evaluate potential positive and negative consequences of their use (Wada et al., 2017; Singh, 2014; Locatelli et al., 2017). In particular, multi-objective analysis can be key to simulating and weighing policy solutions across many competing needs in increasingly complex groundwater systems (Reed et al., 2013). At the same time, wrangling the myriad inputs necessary to construct an urban hydrogeologic model and distilling a complex human-natural system into a computer readable input is an art, that comes with all the creativity and biases present in every modeler and planner (Melsen, 2022). Walker et al. (2003) define five sources of uncertainty within the model-based water management process: context and framing, input uncertainty, model structure uncertainty, parameter uncertainty, and model technical uncertainty. To expand upon the urban groundwater management and multi-objective decision-making literature, the research herein uses the case study

of the Mexico City Metropolitan Area to work through the modeling process while tracing the propagation of modeling decisions made regarding spatially defined model components, including observation choice, parameter uncertainty, and management objective calculation, through to the selection of equitable and effective urban groundwater recharge policies.

1.2 Research Objectives

This dissertation develops a number of methods tailored to the unique challenges posed by the shifting fluxes characteristic of urban groundwater basins experiencing excessive overdraft. The methods explored in this research are:

- (Chapter 2) Develop a physically based regional groundwater model of the Valley of Mexico that incorporates land and water use trends, water supply infrastructure, precipitation recharge, and geologic characteristics to test four managed aquifer recharge alternatives (historical, demand management, increased wastewater recharge, and infiltration basins) across three spatially determined planning objectives (pumping energy, subsidence risk, and urban flooding risk).

- (Chapter 3) Evaluate the effects of endogenous model uncertainty on management alternative performance and ranking by performing global sensitivity analysis on model error and objective values, and developing novel visualizations of alternative selection.

- (Chapter 4) Improve upon the best performing managed aquifer recharge alternative (demand management) from Chapters 2 and 3 by disaggregating municipal pumping schemes and groundwater management objectives, and comparing with socioeconomic indicators at a selection of subregional scales. This study proposes a grid cell clustering method that leverages the understanding of the local social context of groundwater planning to assess distributive justice when calculating policy performance.

Bibliography

Bonneau, J., Fletcher, T. D., Costelloe, J. F. and Burns, M. J. (2017). Stormwater infiltration and the 'urban karst' – A review, *Journal of Hydrology* **552**: 141–150.

Cruz-Ayala, M. B. and Megdal, S. B. (2020). An overview of managed aquifer recharge in mexico and its legal framework, *Water* **12**(2): 474.

Dahlke, H. E., LaHue, G. T., Mautner, M. R., Murphy, N. P., Patterson, N. K., Waterhouse, H., Yang, F. and Foglia, L. (2018). Managed Aquifer Recharge as a Tool to Enhance Sustainable Groundwater Management in California: Examples From Field and Modeling Studies, *in* J. Friesen and L. Rodríguez-Sinobas (eds), *Advances in Chemical Pollution, Environmental Management and Protection*, Elsevier, chapter 8.

Dillon, P., Stuyfzand, P., Grischek, T., Lluria, M., Pyne, R. D. G., Jain, R. C., Bear, J., Schwarz, J., Wang, W., Fernandez, E., Stefan, C., Pettenati, M., van der Gun, J., Sprenger, C., Massmann, G., Scanlon, B. R., Xanke, J., Jokela, P., Zheng, Y., Rossetto, R., Shamrukh, M., Pavelic, P., Murray, E., Ross, A., Bonilla Valverde, J. P., Palma Nava, A., Ansems, N., Posavec, K., Ha, K., Martin, R. and Sapiano, M. (2019). Sixty years of global progress in managed aquifer recharge, *Hydrogeology Journal* **27**(1): 1–30.

Dillon, P., Toze, S., Page, D., Vanderzalm, J., Bekele, E., Sidhu, J. and Rinck-Pfeiffer, S. (2010). Managed aquifer recharge: rediscovering nature as a leading edge technology, *Water Science and Technology* **62**(10): 2338–2345.

Foster, S., Lawrence, A. and Morris, B. (1998). Groundwater in urban development, *World Bank Technical Paper* **390**: 1–55.

Howard, K. and Gelo, K. (2003). Intensive groundwater use in urban areas: the case of megacities, *in* R. Llamas and E. Custodio (eds), *Intensive Use of Groundwater: Challenges and Opportunities*, A.A. Balkema Publishing, chapter Chapter 2, pp. 35 – 58.

Lerner, D. N. (1990). Groundwater recharge in urban areas, *Atmospheric Environment. Part B. Urban Atmosphere* **24**(1): 29–33.

Locatelli, L., Mark, O., Mikkelsen, P. S., Arnbjerg-Nielsen, K., Deletic, A., Roldin, M. and Binning, P. J. (2017). Hydrologic impact of urbanization with extensive stormwater infiltration, *Journal of Hydrology* **544**: 524–537.

Melsen, L. A. (2022). It Takes a Village to Run a Model—The Social Practices of Hydrological Modeling, *Water Resources Research* **58**(2).

Nelson, B. (2017). 21st Century Water Storage Strategies, *San Francisco Estuary and Watershed Science* **15**(4): 0–14.

Ohgaki, S., Takizawa, S., Kataoka, Y., Kuyama, T., Herath, G., Hara, K., Kathiwada, N. R. and Moon, H.-j. (2007). *Sustainable Groundwater Management in Asian Cities: A final report of Research on Sustainable Water Management Policy*, Institute for Global Environmental Strategies.

Orsi, J. (2005). *Hazardous Metropolis*, University of California Press.

Osmanoğlu, B., Dixon, T. H., Wdowinski, S., Cabral-Cano, E. and Jiang, Y. (2011). Mexico City subsidence observed with persistent scatterer InSAR, *International Journal of Applied Earth Observation and Geoinformation* **13**(1): 1–12.

Page, D., Bekele, E., Vanderzalm, J. and Sidhu, J. (2018). Managed Aquifer Recharge (MAR) in Sustainable Urban Water Management, *Water* **10**(3): 239.

Reed, P. M., Hadka, D. M., Herman, J. D., Kasprzyk, J. R. and Kollat, J. B. (2013). Evolutionary multiobjective optimization in water resources: The past, present, and future, *Advances in Water Resources* **51**: 438–456.

Reisner, M. (1993). *Cadillac desert: The American West and its disappearing water*, Penguin.

Schirmer, M., Leschik, S. and Musolff, A. (2013). Current research in urban hydrogeology – A review, *Advances in Water Resources* **51**: 280–291.

Singh, A. (2014). Groundwater resources management through the applications of simulation modeling: A review, *Science of the Total Environment* **499**: 414–423.

United Nations (2016). The World's Cities in 2016, *Technical report*.

Varis, O., Biswas, A. K., Tortajada, C. and Lundqvist, J. (2006). Megacities and water management, *International Journal of Water Resources Development* **22**(2): 377–394.

Vázquez-Suñé, E., Sánchez-Vila, X. and Carrera, J. (2005). Introductory review of specific factors influencing urban groundwater, an emerging branch of hydrogeology, with reference to Barcelona, Spain, *Hydrogeology Journal* **13**(3): 522–533.

Wada, Y., Bierkens, M. F. P., de Roo, A., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., Konar, M., Liu, J., Müller Schmied, H., Oki, T., Pokhrel, Y., Sivapalan, M., Troy, T. J., van Dijk, A. I. J. M., van Emmerik, T., Van Huijgevoort, M. H. J., Van Lanen, H. A. J., Vörösmarty, C. J., Wanders, N. and Wheater, H. (2017). Human–water interface in hydrological modelling: current status and future directions, *Hydrology and Earth System Sciences* **21**(8): 4169–4193.

Walker, W., Harremoës, P., Rotmans, J., van der Sluijs, J., van Asselt, M., Janssen, P. and Krayer von Krauss, M. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support, *Integrated Assessment* **4**(1): 5–17.

Wittenberg, H. and Aksoy, H. (2010). Groundwater intrusion into leaky sewer systems, *Water Science and Technology* **62**(1): 92–98.

Chapter 2

Urban growth and groundwater sustainability: evaluating spatially distributed recharge alternatives in the Mexico City Metropolitan Area[1]

2.1 Abstract

Groundwater-dependent cities face increasing population and changes in urban land use, threatening long-term aquifer sustainability by simultaneously increasing pumping demand while modifying recharge rates. Aquifer management involves a portfolio of alternatives to balance the sustainability of multiple water supply sources, including novel recharge alternatives to supplement traditional groundwater and surface water sources. However, potential managed aquifer recharge solutions can lead to unintended consequences as modern infrastructure and management alter natural hydrology, changing the underlying groundwater paradigm. This study uses a multiobjective analysis approach to evaluate the performance of spatially distributed aquifer management alternatives for the Mexico City Metropolitan Area (MCMA), a region where the population has increased by 48% and urban land use by 35% in the last 30 years. Recharge alternatives are evaluated using a physically based groundwater model that incorporates land and water use trends,

[1]This chapter has been published: Mautner, M. R. L., Foglia, L., Herrera, G. S., Galán, R., and Herman, J. D. (2020). "Urban growth and groundwater sustainability: modeling spatially distributed recharge alternatives in the Mexico City Metropolitan Area". *Journal of Hydrology* ***586***: 124909.

and include: repairs to the water supply distribution network, increased wastewater treatment and infiltration, and increased infiltration of imported water supplies. These alternatives are compared according to planning objectives aggregated over three overlapping subregions within the model to minimize energy use for pumping, minimize water quality risks in subsidence-prone areas, and minimize groundwater flooding in urban areas. Results indicate that management alternatives resulting in equivalent changes in storage can exhibit spatially distinct reductions in drawdown over the historical period. Additionally, combining multiple alternatives results in a more than an additive risk of groundwater flooding, demonstrating the need for mitigation of rising water tables. This model will serve as the basis for future spatial optimization studies under future climate and land use scenarios. These approaches have broad applicability in other rapidly urbanizing regions dependent on groundwater supplies.

2.2 Introduction

Globally, urban areas have grown rapidly compared to rural areas, particularly in megacities with populations greater than ten million (United Nations, 2016). This, in combination with climate variability over time, has led to overexploitation of both surface water and groundwater worldwide (Vörösmarty et al., 2000) Groundwater is increasingly relied on for water supply during times of shortage, placing stress on urban aquifers both in developing and industrialized countries. Urban groundwater is threatened by unsustainable pumping and contamination from surface and subsurface sources, reducing its ability to meet rising demands (Foster et al., 1998; McDonald et al., 2014). Expanding urban land cover amplifies these risks by increasing impervious surface area affecting infiltration, runoff, evapotranspiration, recharge, and, ultimately, regional groundwater capture (Alley and Leake, 2004). These hydrologic changes may require megacities to seek additional sources of local surface water or consider costly imports (Jacobson, 2011). As conventional water supply and sanitation infrastructure with high capital and operating costs have failed to adapt to changing climate and demographic forces, managed aquifer recharge infrastructure and water

management solutions have been advocated to mitigate local flooding vulnerabilities, regulate surface water runoff and groundwater quality, and increase groundwater infiltration to urban aquifers (Dahlke et al., 2018; Roy-Poirier et al., 2010).

As watersheds globally become more urbanized, hydrologic and hydrogeologic models used to support water resources planning must incorporate human impacts on several key dynamics of the water cycle and water quality (Brath et al., 2006; Thompson et al., 2013). For example, urban expansion can increase recharge because losses from decreased precipitation are offset by increases from irrigation with imported water, water supply distribution system leakage, and infiltration from unlined wastewater systems (Lawrence et al., 1998). Alternatively, sewer systems have also been shown to remove significant precipitation recharge (Braud et al., 2013), which can result in an overall negative effect on the urban groundwater budget depending on site-specific characteristics. Increased groundwater abstraction in urban areas can also cause ancillary impacts including subsidence (Ortiz-Zamora and Ortega-Guerrero, 2010) and subsequent ground ruptures (Frigo et al., 2019), as well as flood risks resulting from damaged infrastructure and settling (Jagoon et al., 2009). Finally, urbanization also poses water quality threats from concentrated pollutants in stormwater, wastewater, and landscape and agricultural runoff (Lerner, 2004), which can enter shallow unconfined aquifers via infiltration and confined aquifers via fractures or well boreholes (Huizar-Alvarez et al., 2004). The potential for aquifer management to provide solutions for these risks to urban groundwater security is high given the diversity of alternative water supplies and management tools available (Page et al., 2018). However, the interactions between such solutions and the natural hydrogeologic system may lead to unintended consequences (Li et al., 2017; Maliva, 2020).

As regional-scale hydrologeologic modeling has progressed in large metropolitan areas, experts have stressed that a simple determination of safe yield from an aquifer rather than a more detailed analysis of the dynamic processes governing groundwater availability within a basin is necessary to properly inform decision making (Dendrou, 1982; Bredehoeft, 2002). More recent urban water modeling efforts have included extensive integrated modeling to provide robust groundwater

models (Hanson et al., 2012; Hanson, Lockwood and Schmid, 2014; Hanson, 2015). In these frameworks, a hydrologic model can incorporate economic, climate, energy, ecological, and land use components, and represent distinct anthropogenic processes such as impervious surfaces, groundwater storage, residential water use, urban irrigation, wastewater treatment and storage, and leakage from water distribution and collection infrastructure (Bach et al., 2014; Mitchell et al., 2001; Hanson, Boyce, Schmid, Hughes, Mehl, Leake, Maddock and Niswonger, 2014). However, though regional-scale planning efforts can alleviate general sustainability concerns, they may not consider distributional impacts, particularly in economically marginalized districts where urban growth is often concentrated (Baker, 2012). Specifically, while regional planning decisions may provide water supply needed to support urban growth in the aggregate, stresses may be unevenly distributed across finer spatial scales, creating challenging trade-offs (Daw et al., 2011).

Understanding the coupled dynamics of subsurface hydrology, urban land use, and water consumption is key to ensuring sustainable groundwater management in metropolitan areas (Tellman et al., 2018). In cities facing rapid urbanization, spatially distributed analysis of planning objectives at multiple scales can help to achieve reliable water access across districts (Varis et al., 2006). For example, Reichard et al. (2010) evaluated potential groundwater management options in emergency groundwater use based on the costs associated with negative impacts from intensive groundwater extraction. However, the problem often requires a multiobjective comparison of spatially distributed aquifer management alternatives to demonstrate the conflicts that arise in planning for complex systems. In this study, we develop a physically-based groundwater model that incorporates land use change, water supply infrastructure, and pumping trends for the Mexico City Metropolitan Area (MCMA) as a case study to examine the effects of the choice of spatial and temporal dimensions for the calculation of objectives on preferred aquifer management alternatives.

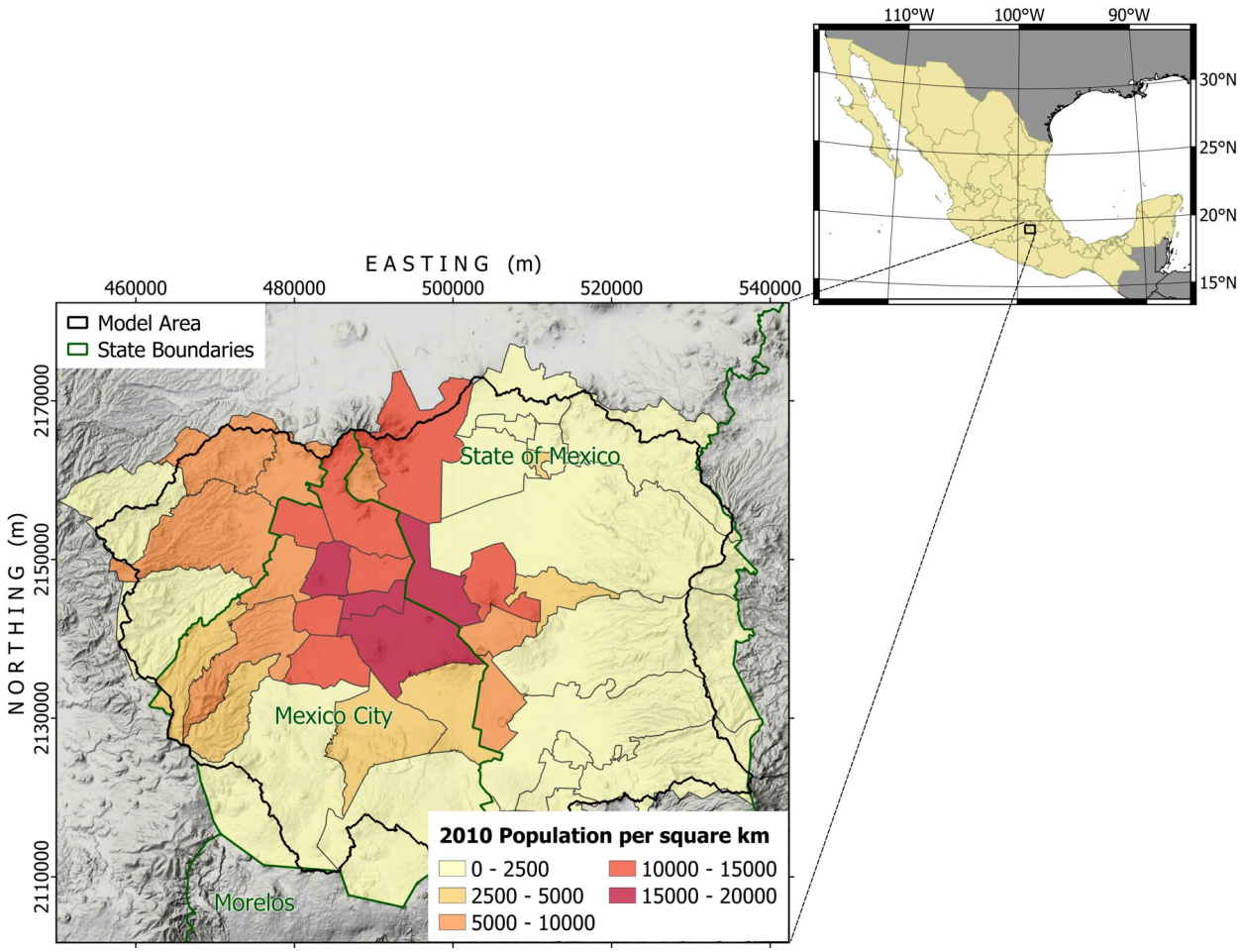

Figure 2.1: The Mexico City Metropolitan Area (MCMA) is made up of urban areas in Mexico City and the State of Mexico. The model area is shown in black with population density shown by municipality.

2.3 Case Study

The MCMA lies within the southwestern portion of the Valley of Mexico watershed (Figure 2.1), characterized by volcanic peaks surrounding a high plains basin (OCAVM, 2014). Until roughly 300 years ago, there were four natural lakes present in the valley, forming an aquifer system which comprised a saturated lakebed overlying alluvial sediment and volcanic rock formations (González-Morán et al., 1999). The first large scale drainage project, el Gran Canal del Desagüe, was completed in 1900. Today, along with the Deep Drainage System, it carries untreated stormwater and wastewater out of the expanding urban area (SACM, 2012). Like many metropolitan areas around the world, the hydrologic cycle of the MCMA has been altered substantially with the

introduction of inter-basin water and wastewater conveyance infrastructure, the alteration of land surface properties in high-density human settlements, and the overdraft of groundwater resources.

From 1980 to 2015, urban land cover expanded significantly (Figure 2.4c), driven by a population that nearly doubled during the same period (CONAPO, 2014). Exploitation of the productive aquifers in the center of the metropolitan area has occurred at varying rates since the discovery of potable groundwater in the valley in 1846 (National Research Council et al., 1995). Today, the city depends on underlying alluvial and volcanic aquifers for 58% of its water supply, which has caused the water table to decline throughout the city, leading to increased energy costs for pumping, water quality concerns, and land subsidence (OCAVM, 2014). A number of studies have estimated various components of the total water budget in both the larger Valley of Mexico watershed as well as the smaller MCMA (DGCOH, 1997; Birkle et al., 1998; Carrera-Hernández and Gaskin, 2008; Gómez-Reyes, 2013). The major flows in the MCMA urban hydrologic system are shown in Figure 2.2 along with how they are represented in the modeling framework developed for this study.

The drinking water supply to the Valley of Mexico includes a complex administrative structure for water management, in which regional, state and local agencies interact. The National Water Commission (CONAGUA) administrates water supply at the national level. It acts at the regional level through the "Organismo de Cuenca Aguas del Valle de México" (OCAVM). Water supply within Mexico City is managed by the Sistema de Aguas de la Ciudad de México (SACMEX), the operating agency of the Mexico City government. On the other hand, the State of Mexico Water Commission (CAEM) is the agency that provides water for

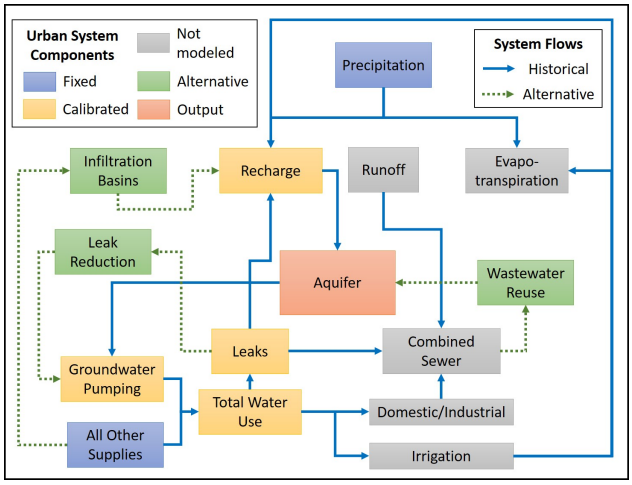

Figure 2.2: Hydrologic components and flows in the Mexico City Metropolitan Area with historical flows in blue and managed aquifer recharge alternative flows in green. The components of the groundwater model developed in this study include fixed inputs (blue), calibrated inputs (yellow), recharge alternatives (green), and model output (orange), with processes not modeled in this study shown in gray.

the State of Mexico municipalities, some of which also have local operating agencies (Banco Interamericano de Desarrollo (BID), 2012; Escolero et al., 2016). In practice, decisions are made at the federal level within OCAVM to distribute supplies within the basin, including wells in the Immediate Action Plan program (PAI), and find and manage water sources from outside of the basin, including the Lerma well system and the Cutzamala reservoir system (OCAVM, 2014; CONAGUA, 2005).

The characteristics of the basin allow for the testing of novel parameterizations of urban aquifer components under traditional groundwater modeling frameworks. Previous basin-wide groundwater modeling efforts have included local water infrastructure such as pumping rates and water supply distribution system leaks, as well as zonal geologic and recharge quantities (Herrera-Zamarrón et al., 2005). However, sensitivity analysis and calibration efforts did not incorporate the effects of time varied land use change, annual and intra-annual precipitation, or pumping data (Lopez-Alvis, 2014; Galán-Breth, 2018). Palma Nava et al. (2015) proposed an integrated, spatially distributed groundwater model that incorporates both subsidence and climate modeling for the MCMA system, but did not consider the siting and evaluation of recharge infrastructure alternatives.

2.4 Methods

2.4.1 Groundwater Model Inputs

One of the most commonly used groundwater modeling code libraries is MODFLOW, a modular finite-difference, quasi-3D flow model developed by the US Geological Survey, which relies on a number of packages to represent different boundary conditions (Harbaugh, 2006). MODFLOW provides a suitable environment to simulate historical and future conditions, perform sensitivity analyses, uncertainty estimations, and generate alternative scenarios via a comprehensive suite of modules. Given its widespread use and modular nature, MODFLOW serves as an appropriate tool to construct a complex urban groundwater model, with the difficulty in the modeling process arising

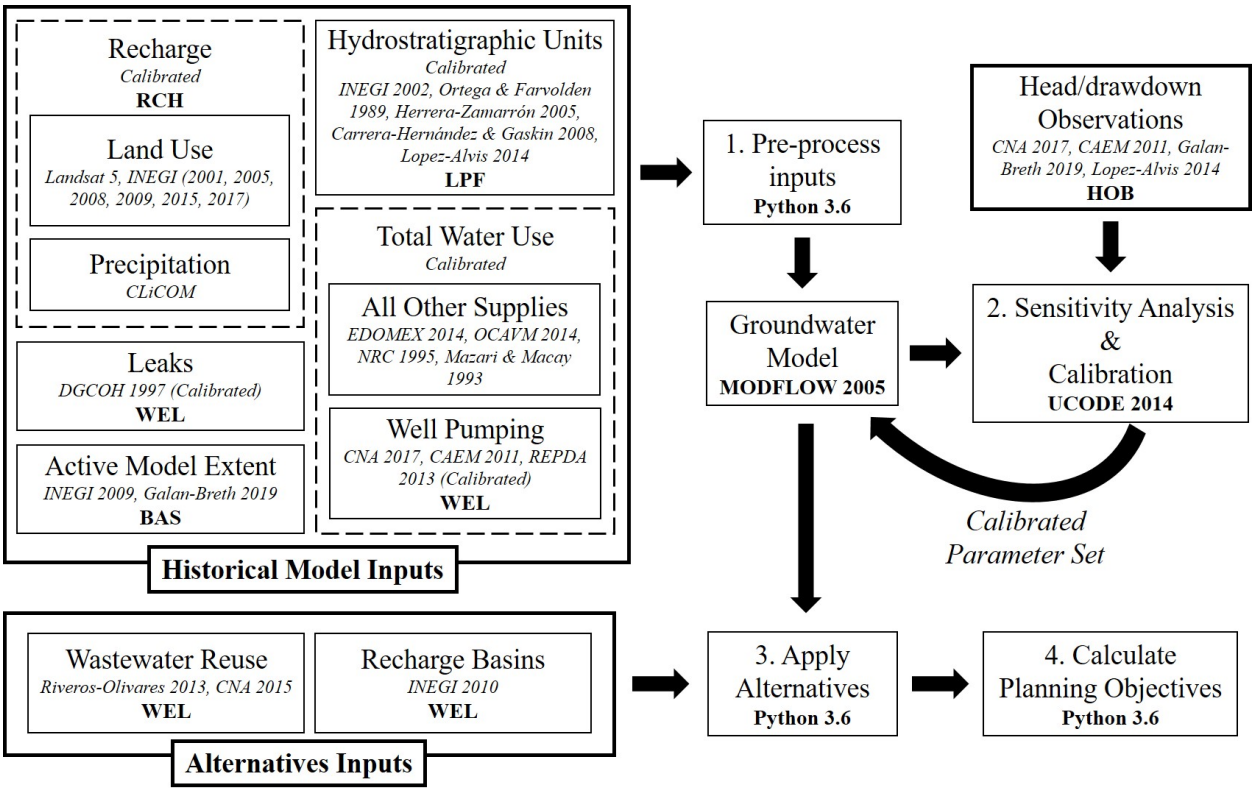

Figure 2.3: Workflow and datasets used in the 1. Pre-processing of model inputs, 2. Sensitivity analysis and calibration of the model, 3. Application of the management alternatives, and 4. Calculation of planning objectives. System processes are defined by both datasets (italics), and the MODFLOW packages or software (bold) used where appropriate.

from the representation of the various natural and anthropogenic fluxes. The scale and complexity of the model can lead to dramatic increases in runtime, making groundwater models challenging to run iteratively, as required for many optimization routines. However, the modular nature of MODFLOW allows researchers to add complexity incrementally, evaluating the trade-offs between the accuracy of the physical representation and the runtime of the simulation (Zhou and Li, 2011).

The numerical model was developed in Python and solved using MODFLOW (model datasets and workflow shown in Figure 2.3), based on the conceptual model of the aquifers in the south-western portion of the Valley of Mexico watershed proposed by (Galán-Breth, 2018; Lopez-Alvis, 2014; Herrera-Zamarrón et al., 2005; Herrera-Revilla et al., 1994). The model contains two vertical layers, and 500 by 500 meter grid cells in a model grid space 84 km East to West and 67 km North to South. The simulation period is from 1984 to 2013 with daily time steps, and monthly or annually

varied data, using the MODFLOW stress period set to a monthly format. The modular nature of MODFLOW allows each physical component of the groundwater system to be represented using a generalized package type, with most components defined in specified flux packages. Modular packages were populated using publicly available and institutional datasets described below, which were preprocessed in Python and QGIS (Bakker et al., 2016).

2.4.1.1 Model Layers and Hydrostratigraphic Units

The model has two numerical layers. Table 2.1 and Figure 2.4 show the relation of these layers with the hydrostratigraphic units and the geologic structure of the model area. The surface layer represents the clay unit, which ranges from 5 to 150 meters thick and has the hydraulic properties of an aquitard. It covers about 25% of the total model surface area. The second numerical layer represents the primary aquifer, which underlies the first layer where it is present (Herrera-Zamarrón et al., 2005). Geologic data was obtained from the Mexican National Institute of Statistics and Geography (INEGI) Geology Inventory (INEGI, 2002). The dataset was simplified into six main geologic formations based on work done by Lopez-Alvis (2014) and Herrera-Zamarrón et al. (2005). Hydraulic conductivities and model layers were initialized based on a review of the literature and are indicated in Table 2.1 (Ortega G. and Farvolden, 1989; Herrera-Zamarrón et al., 2005; Carrera-Hernández and Gaskin, 2008; Lopez-Alvis, 2014).

Table 2.1 Model Hydrostratigraphic Units

Layer	Unit	Hydraulic Conductivity ($\frac{m}{s}$)
1	Lacustrine clay	5E-9
2	Alluvial deposits	5E-4
	Fractured basaltic lavas	5E-5
	Tarango formation (volcanoclastic deposits)	5E-6
	Andesitic fractured rocks (stratovolcanoes of andesitic to dacitic composition)	1E-6

2.4.1.2 Active Model Extent

Previous groundwater models of the valley use an approximation of the watershed boundary of the southwestern portion of the Valley of Mexico watershed as the natural groundwater basin delineation based on the assumption that the surrounding mountain ranges create a natural groundwater divide (Ortega G. and Farvolden, 1989; González-Morán et al., 1999). The INEGI-delineated watershed was used in this study (INEGI, 2010), with corrections for a naturally occurring closure at the northern and southeastern edges of the basin, with the resulting active model extent shown in black as the Model Area in 2.1 (Herrera-Zamarrón et al., 2005; Lopez-Alvis, 2014; Galán-Breth, 2018).

2.4.1.3 Precipitation

Mexico City experiences a marked wet and dry season, with the wet season predominating from May to October and producing annual averages of 600 to 1200 mm of precipitation. Precipitation data was taken from the National Meteorologic Service CLimateCOMputing project which provides daily meteorological station data for precipitation, temperature, and evaporation (SMN, 2015). Of the stations in the region, 97 were chosen for their proximity to the model extent and availability of data. Daily data was summed to monthly and interpolated using cubic-spline interpolation in QGIS for each month over the 30-year period from 1984 to 2013. Rainfall is concentrated in the southwestern area due to orographic effects of the volcanic mountain range, demonstrated in 2.4d, which shows the average rainfall at each station for the month of July interpolated over the model area.

2.4.1.4 Land Use

Spatially distributed, time varied land use data is used in this model to assign both recharge percentages from precipitation (**Zonal Recharge** in 2.5) and incidental recharge from leaks in the water supply distribution system (described in Section 2.4.1.5). Based on growth trends and data availability, land use data was chosen for 1990, 2000, and 2010 to represent 1984 – 1993, 1994 –

2004, and 2005 – 2013, respectively. These model periods are used throughout time-varied data in the model. An extensive GIS analysis was applied to classify land use types into three categories, urban, natural, and water/wetland. Land use percentages for each model cell were calculated based on remote sensing images as described in Appendix B. Each land use category was then assigned an initial infiltration multiplier to be calibrated. The monthly precipitation was multiplied by the percentage of each land use by cell, and the multiplier for each land use type, then summed over all land uses to achieve 100% coverage. Potential monthly infiltration was then divided by the days in the month and applied as recharge for each daily time step as 100% on all geologic formations except over the lacustrine aquitard, where it is applied at 1% of that value.

2.4.1.5 Water Supply Distribution System Leaks

Potential recharge from water infrastructure is a well known component of the urban water cycle, but can vary significantly in magnitude for each component depending on local infrastructure maintenance and conditions (Lerner, 1990). In the MCMA, leaks in the water supply distribution network are estimated to account for between 20% and 55% of water use by municipality or delegation (DGCOH, 1997; National Research Council et al., 1995). A portion of these leaks is assumed to recharge the underlying aquifers depending on local water use, land use type, and geologic subsurface material. Interviews with engineers and administrators at the National Water Commission have suggested that there is a large potential for reductions in groundwater pumping if leaks were controlled using improved operation and maintenance. Thus, including leaks as parameterized objects during the calibration process was necessary to allow for a reduction in leaks as an aquifer management alternative. We considered leaks in the metropolitan area to be a function of total water supply distributed through the public system (Herrera-Zamarrón et al., 2005). Since groundwater supply was calibrated in the model, the total quantity of water use was calculated as the sum of the calibrated pumping and an estimation of other historical water sources in the basin. These sources are listed in 2.4 and summed to 16.7 m^3/s, 28.0 m^3/s, and 33.5 m^3/s for the three model periods defined above (Appendix D) (National Research Council et al., 1995; Mazari and

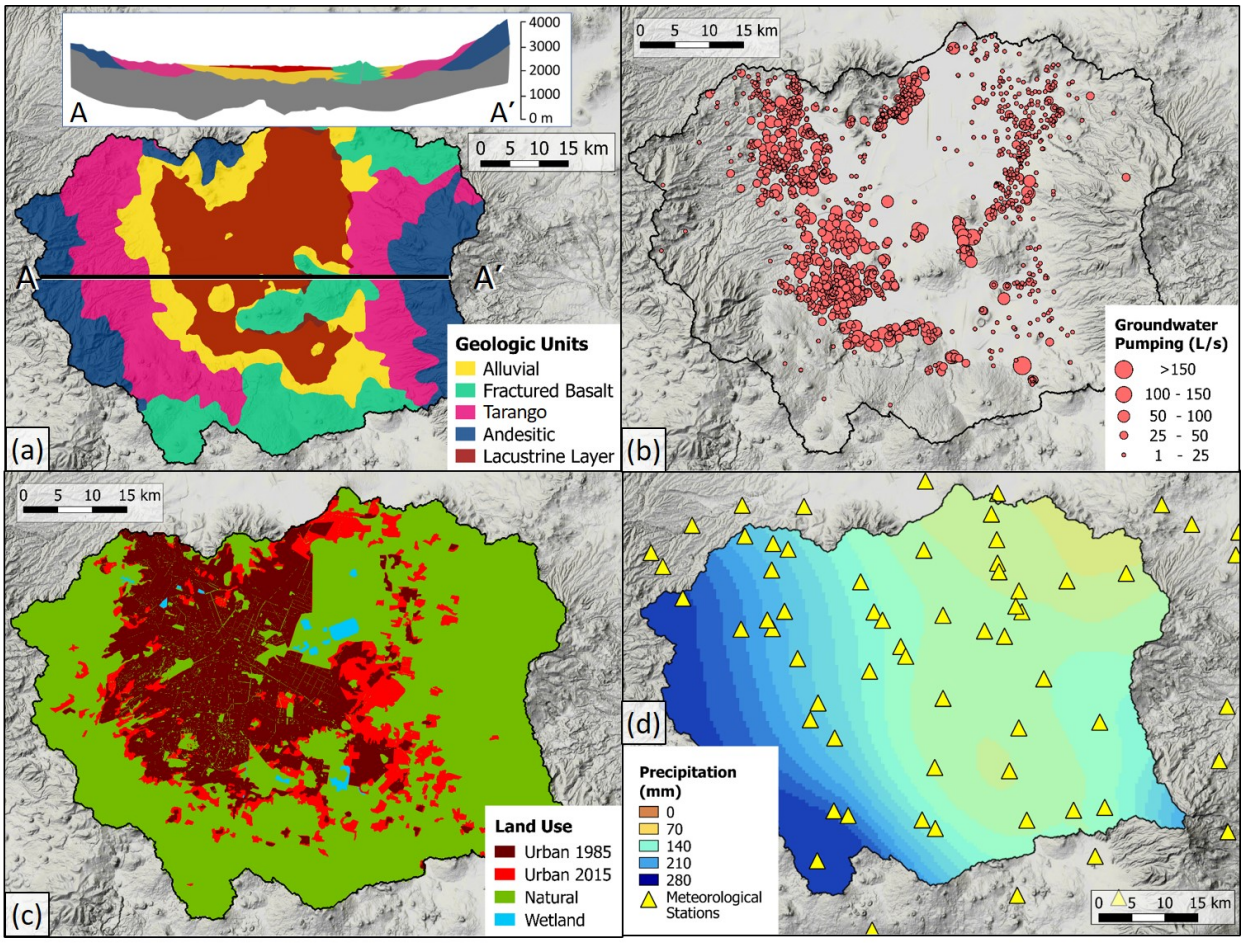

Figure 2.4: Input data for the numerical groundwater model: a) geologic formations with overlain lacustrine layer and side view of A-A', b) 2005 pumping data from the Public Register of Water Rights (REPDA), and Mexico City Water System and State of Mexico service areas, c) adapted INEGI and Landsat datasets showing urban, natural, and wetland/water land use types for 1985 and 2015, d) average July meteorological station data during the model period interpolated over the model area.

Mackay, 1993; OCAVM, 2014; Torres-Bernardino, 2014). These values were used to calculate groundwater pumping as $G = T - O$, where G is local groundwater, T is total water use, and O is other supplies, for each of the model phases 1990, 2000, and 2010.

The calibrated value of T was then multiplied by the percentage of water usage per municipality reported in DGCOH (1997) to get a total water use by municipality. Previous research by the Mexico City Department of Hydraulic Construction and Operations estimated leaks to be between 27% and 58% of total water distribution for 25 of the 41 municipalities and delegations. Total water distribution of municipalities in the State of Mexico was estimated to have losses of 35% based on

the average losses to leaks in Federal District delegations with similar characteristics. These values for leaks from the system, shown in 2.5, were applied to the total water use for each municipality. Any urban leaks occurring over the lacustrine layer were considered to not contribute to recharge of the underlying aquifer based on the low permeability of the confining layer (DGCOH, 1997) and the presence of preferential flow paths provided by the urban drainage system (Bonneau et al., 2017). A multiplier of 1 was applied to each model period, which would allow for calibration of the multiplier if the parameters were found to be sensitive in the model. An infiltration parameter of 10% was included as a calibration parameter, assuming that preferential flow paths to the sewer and evaporation reduce the total infiltration of potential leak recharge.

2.4.1.6 Well Pumping

Well pumping data for the region can be separated into two main categories: municipal pumping rates gathered on a monthly or yearly basis, and well concessions data indicating a single maximum allowable pumping rate. The combined pumping data for 2005 are shown in Figure 2.4b. Municipal data consists of one dataset of monthly pumping data for municipal wells in the Mexico City service area and one dataset of yearly pumping data for municipal wells in the State of Mexico (CONAGUA, 2017; CAEM, 2011). Both municipal datasets were assumed to be accurate and were not modified in the calibration process. Georeferenced well pumping concessions data for the study site is available beginning in 1983 from the Public Register of Water Rights (REPDA) hosted by the National Water Commission (CONAGUA, 2013). Since the REPDA data does not include information about time varied quantities nor when the wells were installed, the total quantity and the distribution of pumping was assumed to be unknown, but able to be calibrated from the properties of the dataset. To determine the total quantity, a Total Water Use multiplier was applied, one for each of the three model periods, to an initial estimate for pumping. Then the other water supplies defined in Section 2.4.1.5 plus the known municipal pumping were subtracted from this estimate to obtain the remaining groundwater pumping (2.5). Pumping distribution in the basin has shifted from high pumping in the historical downtown region to pumping in the surrounding areas (National Research

Council et al., 1995). To mimic this trend, the wells in the REPDA dataset were separated into two groups: historical downtown and peripheral. Next, a multiplier was applied to only the downtown REPDA wells, with the peripheral wells calculated as the remaining quantity from the difference of the calibrated groundwater pumping and the urban REPDA set.

2.4.2 Sensitivity Analysis and Model Calibration

A fundamental aspect of groundwater model construction is evaluation of the model against observed conditions; inverse modeling is generally considered a valuable tool to tune model characteristics using observations (Hill et al., 2016). Calibration with inverse modeling is based on the principle that model parameters are either difficult to determine in the field at the scale that has been selected, or cannot be directly measured or estimated from the information available (Foglia et al., 2009; McLaughlin and Townley, 1996). The first step is the parameterization of model characteristics so that the potentially infinite number of values to be adjusted over time and space are reduced to a limited number of variables to evaluate. We performed sensitivity analyses and parameter estimation using UCODE 2014 (Poeter et al., 2014), which employs a weighted least-squares objective function to compare model simulated values to observations provided.

As in any calibration process, the quality of the observations is important in determining which observations will lead to the most accurate representation of the system. These errors or uncertainties are represented in the modeling process by applying a weight to the observation in the objective function. Once the weights are determined, a sensitivity analysis is conducted on the model error with respect to perturbation of the initial parameter values to select a smaller group of parameters that are most sensitive (Hill and Tiedeman, 2007). The composite scaled sensitivity combined with the evaluation of parameters correlations is used to evaluate the parameters that contribute the largest changes to the weighted sum of squared error between the observations and their simulated equivalent. Finally, an iterative process is completed whereby the parameters are perturbed locally and the weighted sum of squared error is calculated for each parameter calibration until an acceptable tolerance is met.

2.4.2.1 Observations

An initial set of 669 historical groundwater observation wells were obtained from the National Water Commission. After an initial review, the data was determined to contain various irregularities that would impede the proper calibration of the model. As such, wells that exhibited a difference in water level of 20 m or more from one year to the next were reviewed to see if the fluctuation was persistent and/or representative of behavior in nearby wells. The wells were filtered to contain only one observation well per model grid cell and exclude any well observation sets with erratic behaviour as compared to nearby wells. Wells were chosen first by highest number of observations and then proximity to cell center. Each well dataset has between one and 30 years of yearly water level measurements, however, there is limited information on the time of year or conditions in which the observations were taken. This yielded a total of 531 wells, with 6896 unique observations over the 30-year historical model period. To set the initial hydraulic head throughout the model area, the 132 observations from 1984 were interpolated using ordinary kriging with a spherical variogram model. Of the remaining 6764 observations, 531 were used as initial head values and the rest were used as drawdown observations to perform the sensitivity analysis and calibration.

Weights for each observation were assigned based on the relative confidence in the quality of the observation. Following Lopez-Alvis (2014) and Galán-Breth (2018), error is introduced by uncertainties associated with well elevation, well position, non-simulated transitory processes, sampling, and model discretization according to the following calculations:

1. Error in the elevation of the well

$$\sigma_1 = \sqrt{\left(\frac{\nabla_{DEM} \times r_h}{4}\right)^2 + \left(\frac{r_v}{4}\right)^2 + \left(\frac{\nabla_{DEM} \times a_m}{4}\right)^2} \tag{2.1}$$

where ∇_{DEM} is the gradient of the digital elevation model (DEM), r_h is the horizontal resolution of the DEM (5 m), r_v is the vertical resolution of the DEM (1 m), and a_m is the accuracy of the GPS measuring device ($\approx \pm 90m$).

2. Error in the horizontal position of the well

$$\sigma_2 = \frac{\nabla_H \times a_m}{4} \tag{2.2}$$

where ∇_H is the hydraulic gradient at the position of the well.

3. Error in the transitory processes not simulated

$$\sigma_3 = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(d_i - \overline{d}\right)^2} \tag{2.3}$$

where d_i is the ith difference between subsequent observations for a single well (m), $\overline{d}$ is the mean of the differences for all observations for a single well (m), and n is the number of differences for the well.

4. Error in observation measurement accuracy

$$\sigma_4 = \frac{(d \times 0.001)}{4} \tag{2.4}$$

where d is the depth of the measurement (m) and 0.001 refers to an accuracy of 0.1% of the well depth.

5. Error in the discretization of the model

$$\sigma_5 = \frac{(s_{nodal} \times \nabla_H)}{4} \tag{2.5}$$

where s_{nodal} is the nodal spacing of the mesh.

The total standard deviation is assumed to be additive and was determined as the root sum of squares of each of the five errors,

$$\sigma_T = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 + \sigma_5^2} \tag{2.6}$$

2.4.2.2 Parameters

The model parameters evaluated can be separated into four main groups: zonal geologic, time varied infrastructure, zonal recharge, and leak infiltration. The 33 model parameters are shown in 2.5, including: horizontal hydraulic conductivity [HK], vertical anisotropy of hydraulic conductivity [$VANI$], specific storage [S_s, specific yield [S_y], pumping [Q], leak percentage [LK], total water use [TWU], land use recharge [RCH], and leak infiltration [IN]. Lower and upper bounds for the zonal geologic parameters were selected based on previous studies in the area and accepted values in the literature, which were used as constraints for those parameters during the calibration process (Singhal and Gupta, 1999; Galán-Breth, 2018; Lopez-Alvis, 2014; Herrera-Zamarrón et al., 2005). Temporal variations in groundwater pumping and leaks from the water supply distribution system were separated into three model phases: 1984-1993, 1993-2005, and 2005-2014, chosen based on the availability of land use imagery. Land use multipliers were assigned to urban, natural, or water/wetland land use types.

Parameters were selected for perturbation in two stages: first as a confined system to determine approximate parameter values and then as a convertible confined/unconfined system to determine final parameter values. For each stage, a sensitivity analysis was performed to evaluate the most sensitive parameters and any parameter correlations, parameters to be calibrated were selected, and the weighted sum of squares was minimized to determine the final calibrated parameters.

2.4.3 Application of Management Alternatives

Managed aquifer recharge offers a variety of infrastructure and policy options to increase infiltration into over-drafted aquifers and promote conjunctive use of alternative water supply sources (Dahlke et al., 2018). In the MCMA a number of aquifer management alternatives have been proposed to mitigate the effects of intensive groundwater use and the alteration of the natural hydrologic regimes including demand reduction schemes as well as managed aquifer recharge. Demand reduction within the Valley of Mexico could include exchange of fresh water for treated wastewater for agricultural irrigation, reduction of leaks in the water supply distribution network,

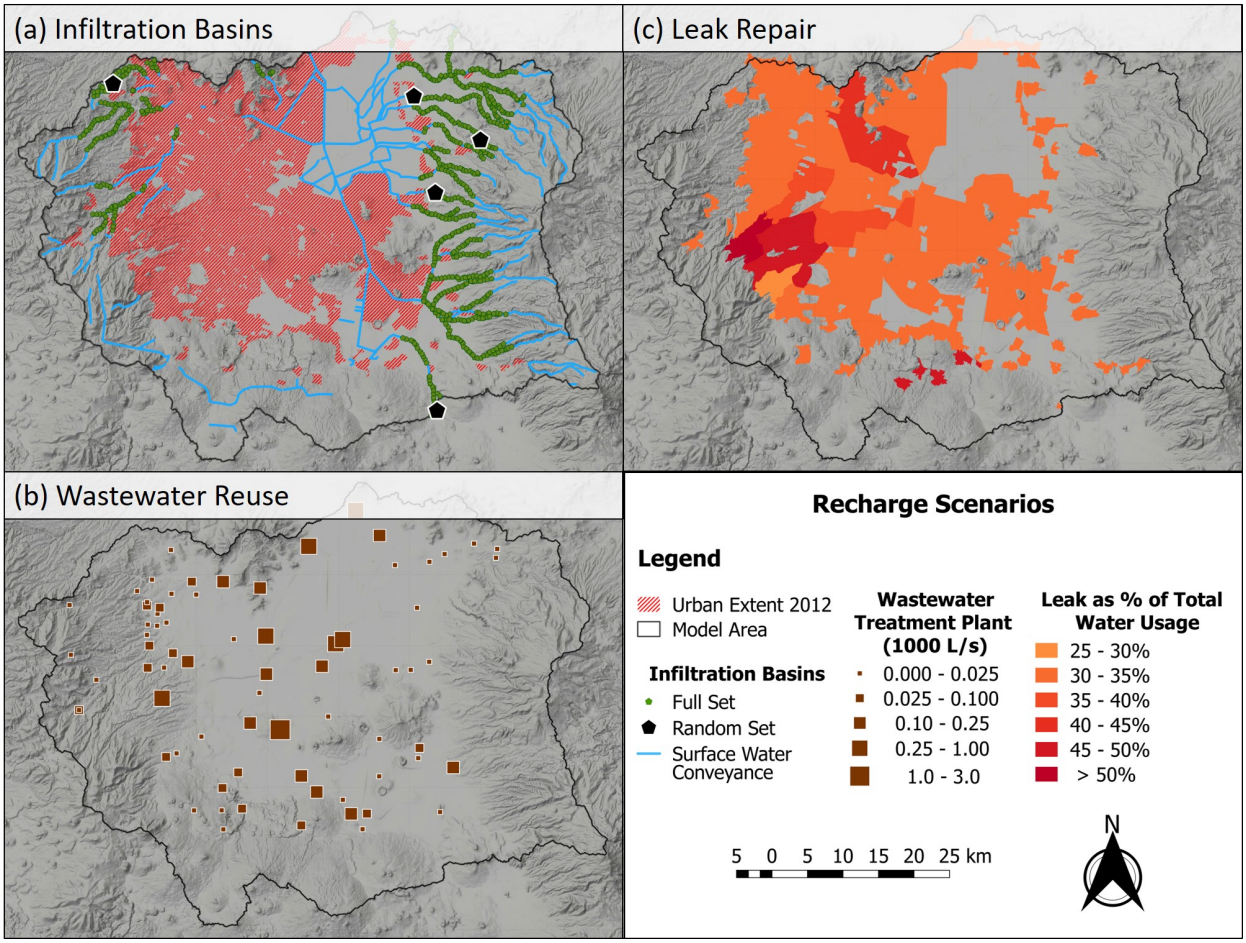

Figure 2.5: Aquifer management alternatives evaluated: a) implementation of five infiltration basins (INEGI, 2010), b) repair of 20% of potable supply leaks (DGCOH, 1997), and c) increased recharge at all wastewater treatment plants (Riveros-Olivares, 2013; CONAGUA, 2015)

substitution of imported sources for local groundwater sources, and redistribution of pumping to reduce pumping at problem zones (Palma Nava et al., 2015; OCAVM, 2014; Herrera-Zamarrón et al., 2005). Alternatively, there are many possible managed aquifer recharge projects for the MCMA including infiltration basins, injection of treated wastewater, and decentralized infiltration at the household scale with low impact development design. For this study three alternatives with varying spatial extents were chosen for multiobjective analysis: increased infiltration of imported water supply (Figure 2.5a), increased wastewater treatment and infiltration (Figure 2.5b), and repairs to the water supply distribution network (Figure 2.5c), representing centralized, regionally centralized, and decentralized interventions, respectively.

The spatial dimension of the recharge problem adds a degree of complexity; the planning and design problem requires selecting not only infrastructure capacities, but also their locations. Numerical representation of the recharge alternatives requires data describing infrastructure, including the location and capacity of existing wastewater treatment facilities and the location of existing surface water conveyance routes. These data were collected from publicly available planning documents and GIS databases and implemented as described in this section (INEGI, 2017a; OCAVM, 2014; Riveros-Olivares, 2013). In the development of this model, the management alternatives and supply policies are assumed to be implemented by centralized decision-makers with no dynamic model for interaction between basin agencies. The historical condition, the three management alternatives described below, and an alternative comprising all three management alternatives were analyzed by calculating the cumulative change in storage as well as the spatially aggregated planning objectives defined in Section 2.4.4. Cumulative change in storage (ΔS) was calculated as the sum of the difference between model inflows, including precipitation recharge (R) and water supply network leaks (L), and model outflow from groundwater well pumping (W), over all stress periods (t):

$$\Delta S = \sum_{t=1}^{360} [R(t) + L(t)] - W(t) \tag{2.1}$$

2.4.3.1 Infiltration Basins

The Cutzamala reservoir system is the main source of imported water for the MCMA (CONAGUA, 2005). De Nys et al. (2015) suggest that up to 5 m^3/s of supply could be yielded from reoperation of the reservoir system, while Birkle et al. (1998) calculate a potential of 17.6 m^3/s of local surface water for use within the basin. Other studies of potential import supplies show that there is economic interest in bringing more external sources into the system, however, new surface reservoirs to store such sources may be limited given the growing urban areas (OCAVM, 2014). This alternative proposes the implementation of five randomly placed surface infiltration basins to recharge the underlying aquifers at 1 m3/s based on similar basins installed in other urban aquifers (Dahlke et al., 2018). Sites were selected from a complete set of evenly spaced points along existing surface

water conveyance infrastructure that connects to the Cutzamala reservoir system (Figure 2.5a). The points were filtered such that only one potential site was allowed per model cell and no sites were allowed in cells exceeding the elevation at the Cutzamala inlet (2600 meters above sea level), in urban areas, nor above the lacustrine layer.

2.4.3.2 Wastewater Reuse at Existing Treatment Plants

There are 74 existing wastewater treatment plants within the model area, a majority of which are operating at less than their installed capacity due to substandard maintenance and operation practices (Riveros-Olivares, 2013). The 2015 inventory for wastewater treatment plants was used to determine the installed capacity and the current operating capacity (CONAGUA, 2015). The difference between these values was then added as recharge to the aquifer layer assuming all repairs and operations modifications were made to improve the amount of treated water produced at each plant. This simple difference does not account for infiltration inefficiencies, evapotranspiration, nor conveyance losses, making this estimate an upper bound for this recharge potential. However, less than 10% of wastewater is currently treated within the valley (Sosa-Rodriguez, 2010), indicating that there is a indeed a much higher overall potential for wastewater reuse within the valley than what is represented in the wastewater reuse management alternative implemented in this model.

2.4.3.3 Repair to Leaks in Potable Supply Distribution Network

In the final alternative, the recharge from leaks was reduced while simultaneously decreasing pumping. This was achieved by applying a leak persistence percentage factor between 0 and 1. A value of 1 indicates that the leaks are present at 100% of the historical value while a value of 0 indicates that all the leaks have been fixed. The pumping quantities were reduced at the municipal or delegation scale by applying Equation 2.2 depending on the base leak percent per municipality of the total water pumped as described in Section 2.4.1.5.

$$T_N = T[1 - P_L(1 - L_P)] \tag{2.2}$$

where T_N is the new total pumping, T is the initial total pumping, P_L is the percentage of leak as a total of the total pumping by municipality, and L_P is the leak persistence percentage.

All water use savings were assumed to reduce groundwater pumping, representing a form of demand management, and resulting in an overall increase in storage within the basin. The difference between the new groundwater pumping G_N and the original groundwater pumping G and the difference between T_N and T are equivalent:

$$G_N - G = T_N - T \tag{2.3}$$

Thus, using substitution and algebraic methods as described in Appendix B, the ratio of the new groundwater pumping to the original groundwater pumping is:

$$\frac{G_N}{G} = \frac{T(1 + P_L(1 - L_P)) - s}{T - s} \tag{2.4}$$

Which is then applied to all pumping wells within a municipality or delegation. This application could later be used to apply varying degrees of repairs to each municipality or delegation individually.

2.4.4 Calculation of Planning Objectives

Urban water supply planning is inherently a multiobjective problem, as long-term sustainability generally conflicts with net present cost, as well as future revenue streams for stakeholders reliant on the water sector. In such cases, no single optimal solution exists, but rather a set of Pareto-optimal solutions, where one objective can only be improved by worsening another (Banzhaf, 2009). These challenges motivate the need for novel methods to identify the tradeoffs between multiple conflicting objectives that arise in the design of aquifer management portfolios in urban settings. In this case, three management objectives are defined to demonstrate the reliance of any given management objective on spatial context: minimize water supply pumping energy use, minimize water quality risks in subsidence-prone areas, and minimize flooding potential in urban

areas. This multiobjective approach allows several advantages. First, objectives are not required to be aggregated using commensurate units of utility, which may be difficult in systems with many stakeholders with conflicting preferences (Cohon and Marks, 1975). Second, a weighted aggregation approach results in a single preferred alternative, without fully capturing the range of performance in all objectives that may lead to a change in preference (Bond et al., 2008). The following planning objectives were calculated separately for the historical condition, the three management alternatives, and the combined management alternatives.

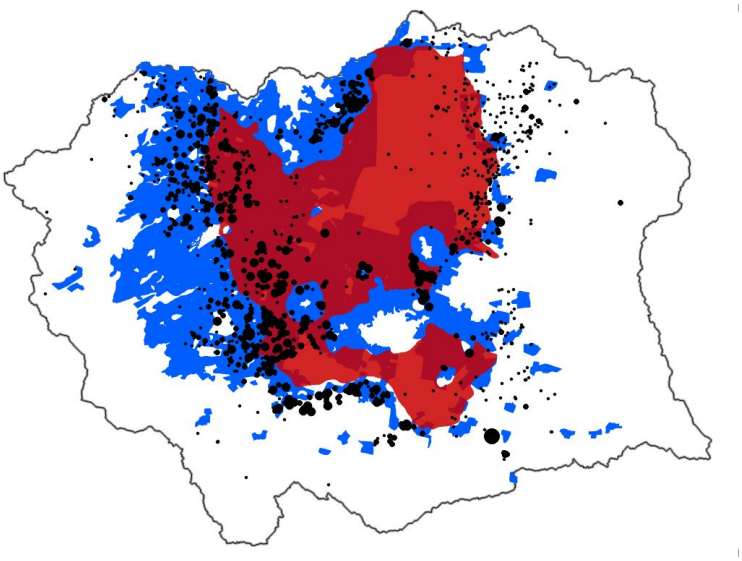

Figure 2.6: Spatial set of cells or points measured in each objective: pumping wells for energy (black), lacustrine layer for water quality (red), and urban cells for flooding (blue).

2.4.4.1 Minimize Energy Use

Falling groundwater tables result in costs relating to increased energy costs to pump deeper water and/or the cost of drilling deeper wells. This objective estimates a lower bound for energy use associated with pumping in wells by calculating the amount of energy required to lift the quantity of groundwater pumped per day from the head elevation in the cell of the pumping well to the land surface over all wells. This objective was calculated in kilowatt hours (kWh). Given n_{wells} is the set of all pumping wells, ϵ is the conversion factor from lift to kWh, d is the number of days in month t, $h_{surface}$ is the elevation of the land surface, and h_{well} is the hydraulic head at the location

of well w, then the energy measure E is,

$$E = \sum_{t=25}^{360} \sum_{w=1}^{n_{wells}} \epsilon d \left(h_{surface} - h_{well} \right) \tag{2.1}$$

2.4.4.2 Minimize Water Quality Risks in Subsidence-Prone Areas

Subsidence in Mexico City contributes to many potential problems for water supply planners in the region. In addition to infrastructure damage, there are also major water quality concerns that have arisen as the water table underneath the confining clay layer has fallen (Huizar-Alvarez et al., 2016). The reversal of the hydraulic gradient as the confined aquifer beneath the clay layer becomes unconfined threatens water security via two primary mechanisms. First, as the potentiometric surface of the confined aquifer falls below the bottom of the confining layer, discharge from the chloride-rich aquitard increases to the good quality aquifer (Ortega-Guerrero et al., 1993). Secondly, in a confined aquifer with a potentiometric surface above the land surface, the flow in the system is upward, however, as the hydraulic head drops below the surface, poor quality water from the surface, including sewage and stormwater runoff, flows downward through fractures caused by subsidence and poorly managed boreholes (Huizar-Alvarez et al., 2004). The water quality risk measure was defined as the percent of confined cells with a potentiometric surface below the elevation of the bottom of the confining layer during the last year of the model period. Given n_{clay} is the set of lacustrine model cells, $h_{surface}$ is the elevation of the bottom of the confining layer, and h_{clay} is the hydraulic head in the aquifer below the confining payer at time t, then the water quality risk measure W is:

$$W = 12 * \frac{1}{n_{clay}} \sum_{c=1}^{n_{clay}} \sum_{t=348}^{360} \left(h(t,c)_{surface} - h(c)_{clay} \right) \tag{2.2}$$

2.4.4.3 Minimize Groundwater Flooding in Urban Areas

Mexico City, like other urban areas built in low-lying regions, suffers from flooding that arises from a variety of mechanisms. These mechanisms can be separated into two main categories:

overland flooding and groundwater flooding (Abboud et al., 2018). The former, overflow of rivers or canals, has been studied and addressed extensively, while the latter, the emergence of groundwater through rising water tables, has largely been overlooked both worldwide and in Mexico City (Abboud et al., 2018; Aragón-Durand, 2007). Groundwater flooding can occur in three main forms: short-term rising water tables in shallow, yet permeable aquifers; long-term rising water tables in unconfined aquifers; and return of water tables to preindustrial levels in historically depleted aquifers (Macdonald et al., 2008). In addition to exacerbating surface flooding conditions, groundwater flooding can cause significant structural damages to buildings and public infrastructure (Toll et al., 2012).

To address this risk, the groundwater flooding objective was defined as the percent of the total urban area over the model period that contained a hydraulic head above the land surface. Given c_{urban} is the set of model cells with urban land cover, c_{flood} is the set of model cells with urban land cover that are flooded at time t, and A_{urban} is the percent urban area in the model cell, then the urban flooding measure F is,

$$F = \frac{\sum_{t=25}^{360} \sum_{u=1}^{c_{flood}} A_{urban}}{\sum_{t=25}^{360} \sum_{u=1}^{c_{urban}} A_{urban}} \tag{2.3}$$

2.5 Results and Discussion

2.5.1 Calibration and Parameter Sensitivities

Composite scaled sensitivities are shown for the fifteen most sensitive of the thirty-three model parameters in Figure 2.7 for stage 1 and stage 2 of the calibration process. The parameter correlation coefficients for stage 1 and stage 2 are shown in Table 2.2. Based on the parameter correlations, the vertical anisotropy of the clay formation was not included in the calibration process because it was highly correlated with the horizontal hydraulic conductivity of the clay layer. The eight most sensitive parameters were calibrated in stage 1 (Figure 2.7a), while in stage 2 the thirteen most sensitive parameters were calibrated (Figure 2.7b). The composite scaled sensitivities of the

Table 2.2 Correlation coefficients for parameters that were chosen for calibration according to composite scaled sensitivity values.

Parameter A	Parameter B	Correlation Coeff.
Stage 1		
HK_{clay}	$Z - Ani_{clay}$	0.999
Q_{1990}	Q_{2000}	0.966
TWU_{1990}	TWU_{2000}	0.963
TWU_{1990}	TWU_{2010}	0.929
TWU_{2000}	TWU_{2010}	0.925
Q_{2000}	Q_{2010}	0.916
Q_{1990}	Q_{2010}	0.900
Stage 2		
Q_{1990}	Q_{2000}	0.932
$HK_{Tarango}$	$S_{y,Tarango}$	0.910
TWU_{1990}	TWU_{2000}	0.880
HK_{basalt}	$S_{y,basalt}$	0.857
Q_{1990}	Q_{2010}	0.852

resulting parameter set is shown in 2.7c.

The total water use parameters are the three most sensitive parameters in the set, indicating the importance of incorporating time varied water use factors (blue) into hydrogeologic modeling. This could be a result of these parameters acting as scale factors that exert influence over many pumping cells together rather than individually, thus making them more dominant parameters. The zonal geologic parameters (orange) follow, with the clay, basalt, alluvial, and Tarango formations being the most sensitive. This may be a factor of the location of the observations available, given that there are no observations located in the andesitic formation.

Next, of the remaining time varied infrastructure parameters (blue) we can see that all three urban pumping parameters are among the most sensitive parameters for stages 1 and 2, with the third urban pumping parameter dropping out of the top fifteen for the final sensitivity analysis, which highlights the spatial factor of the pumping and the expected shift from urban to periurban pumping. Finally, the recharge infiltration multiplier is the only zonal recharge parameter (green) among the most sensitive parameters. This is likely a result of a combination of the concentration of precipitation in natural areas, the higher infiltration rates experienced under natural land cover,

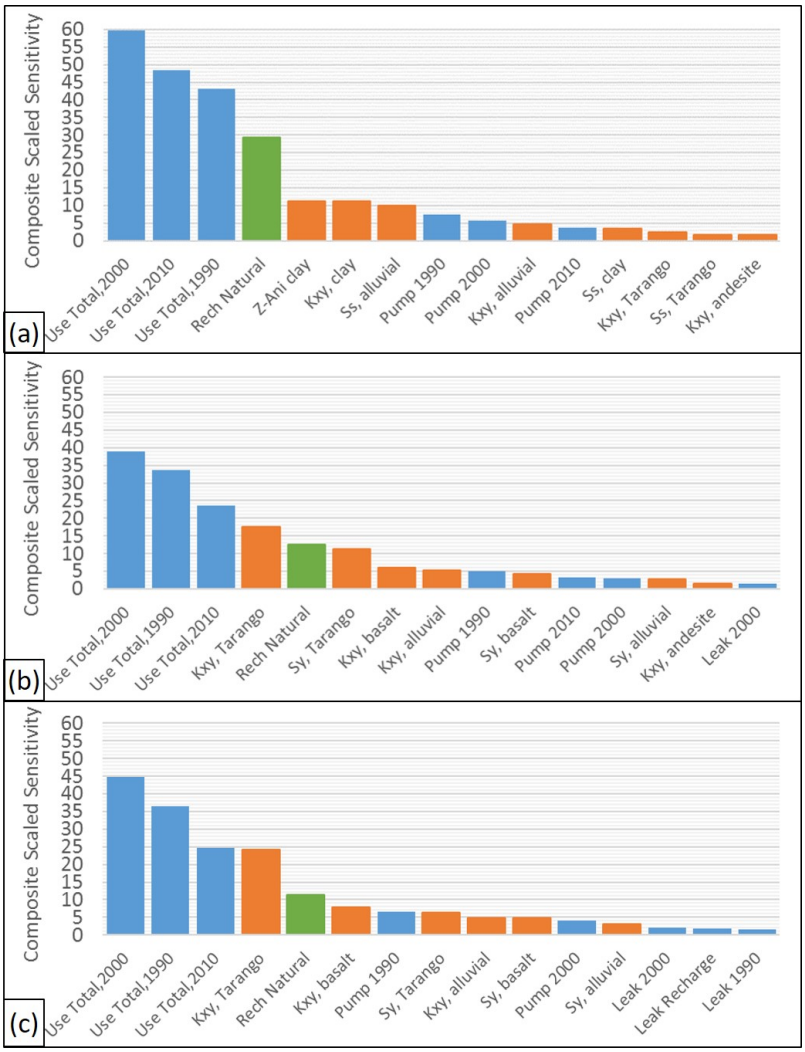

Figure 2.7: Composite scaled sensitivity values of 15 most sensitive parameters, out of a total of 33 adjustable model parameters: zonal geologic (orange), time varied infrastructure (blue), zonal recharge (green), and leak infiltration (gray). Calculated for a) stage 1 of the calibration process with a confined system, b) stage 2 of the calibration process with a convertible confined/unconfined system with wetting, and c) final calibrated parameters.

and the large percentage of area represented by the natural land cover.

The fifteen most sensitive parameters chosen for parameter estimation were horizontal hydraulic conductivity [HK] for the clay, alluvial, Tarango, and fractured basalt formations; specific storage for the alluvial formation; all three urban pumping multipliers; all three total water use multipliers; the recharge multiplier for the natural land use type; and specific yield for the alluvial, Tarango, and fractured basalt formations. The parameters initial values and calibrated results are shown in Table 2.5. The values calculated for pumping and recharge fluxes are consistent with the ranges reported in the literature (Gómez-Reyes, 2013; Birkle et al., 1998; OCAVM, 2014).

2.5.2 Simulation Agreement with Historical Observations

Figure 2.8 shows the simulated changes in hydraulic head over time for selected wells in the basin, which follow the general downward trend of the corresponding sets of observations. Overall model

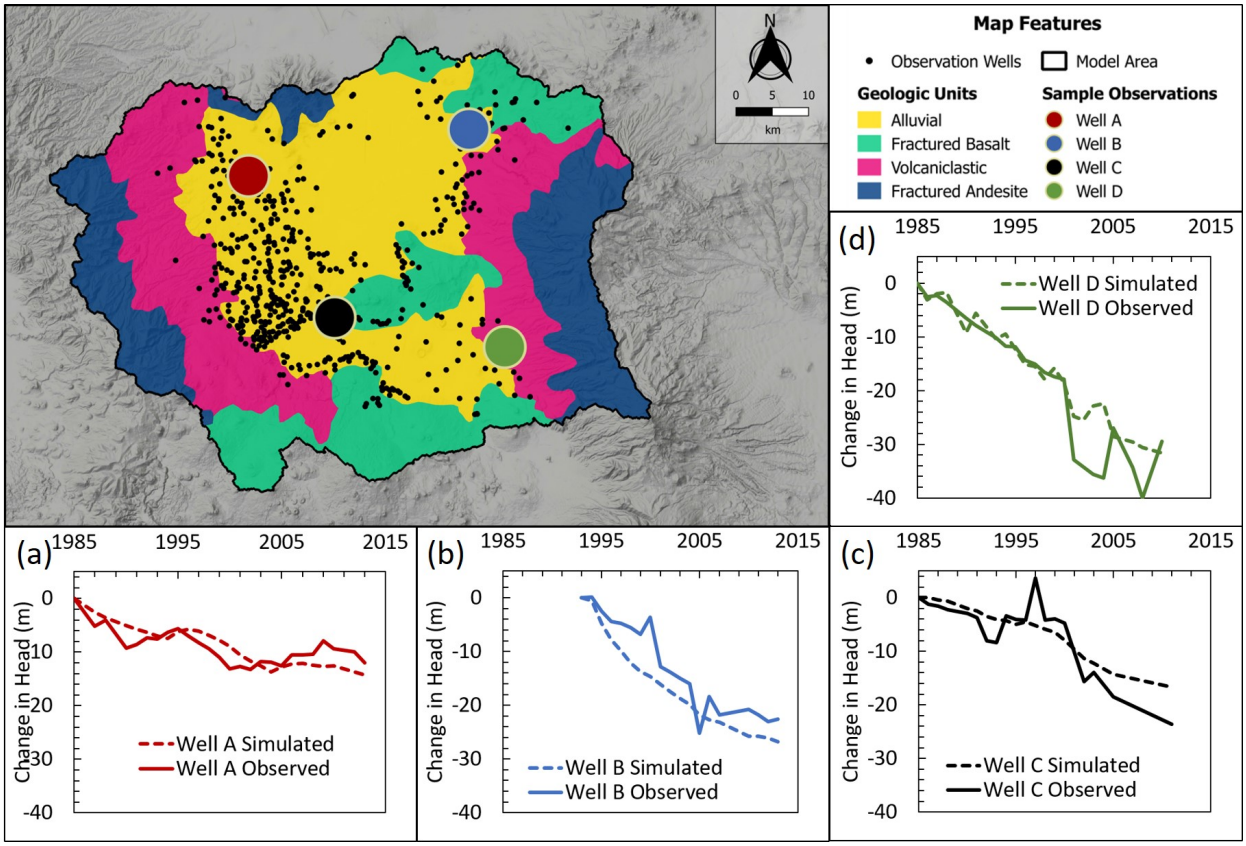

Figure 2.8: Selected hydrographs of simulated and observed change in hydraulic head over time for representative observations from the MODFLOW model using calibrated parameters.

performance compares well with historical observations (Figure 2.18), however, a comparison between all simulated drawdown with observed drawdown shows better agreement in the alluvial formation with observed values than the other two formations Figure 2.19. This is confirmed in Table 2.6, in which the alluvial formation performs better than the other two formations in both the weighted and unweighted mean absolute error. The reason for this discrepancy is twofold: first, the Tarango and fractured basalt formations are more heterogeneous in composition than the alluvial formation, and second, the boundaries of the formation have some error associated with their placement, which could indicate that a number of observation wells along the edges of the formations were incorrectly classified. These results highlight the importance of calibrating spatially distributed variables within the modeling framework. In the drawdown time-series for the four wells shown in Figure 2.8, the simulated changes in groundwater levels exhibit distinct differences across the geologic formations, with the two wells in the alluvial formation (A and

B) and the well in the fractured basalt formation (C) showing more gradual annual changes as compared to the well in the Tarango formation (D). In addition, comparison of model fluxes, to more detailed studies of recharge discussed in Section 2.3 confirm that the simulated groundwater recharge is in agreement with estimates for the area. For example, calibrated pumping rates were 36.4 m^3/s on average, which is 13.4% lower than that reported by Gómez-Reyes (2013).

2.5.3 Performance of Aquifer Management Alternatives

The three alternatives resulted in an average increase in storage over the basin of 5.00 m^3/s, 5.10 m^3/s, 5.31 m^3/s, and 15.41 m^3/s, for infiltration basins, wastewater reuse, repair leaks, and combined, respectively over the 30-year period (Table 2.3). Palma Nava et al. (2015) report an overall basin deficit of 23 m^3/s for the Valley of Mexico watershed, of which this study covers the southwest portion, which would not be entirely overcome by the combined alternative change in storage of 15.41 m^3/s. However, as shown in Figure 2.10, the year-to-year deficit in the southwest portion of the basin is overcome by the combined alternative, although long-term storage deficits within the basin were not addressed in this study. It is important to note that the model period (1984 - 2013) has been part of an increasingly wet period, which is uncertain to continue (Ibarraran, 2011). Therefore, this model may not capture a likely drier climate scenario that would further exacerbate decreasing storage over time, thus this model scenario may not represent a conservative estimate in terms of water supply reliability in the future.

Table 2.3 Aquifer management alternative increase in recharge

Alternative	**Average Increase in Recharge** (m^3/s)
Infiltration Basins	5.00
Wastewater Reuse	5.10
Leak Repair	5.33

When viewing the distribution of recharge throughout the basin (Figure 2.9) it is apparent that each aquifer recharge alternative has a unique spatial signature. The wastewater reuse and repair leaks alternatives result in a more distributed change in hydraulic head over the basin, while the recharge basins result in more localized changes in head. These results also highlight the importance of accounting for geologic heterogeneity in regional scale supply planning as the

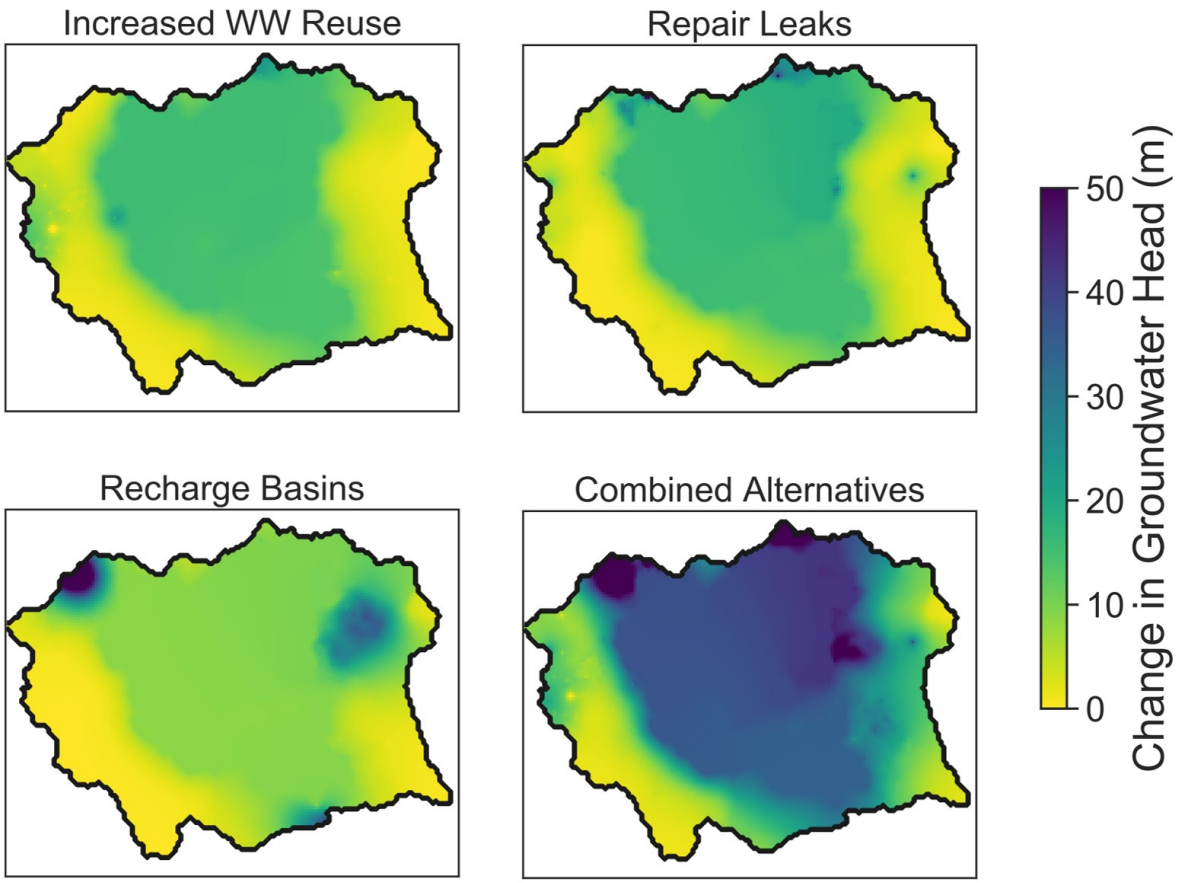

Figure 2.9: Change in head for each aquifer management alternative as difference from historical change in head over the simulation period of 1984 to 2013.

Tarango and andesitic formations experience much greater response when compared to the alluvial and fractured basalt formations. Despite the relatively equal recharge rates for each alternative in Table 2.3, Figure 2.9 indicates that the alternatives may result in spatially varied reductions in drawdown, with corresponding impacts on management objectives.

An important aspect of this model is the incorporation of diverse spatial and temporal scales used to measure the planning objectives (Figure 2.6). Each objective uses hydraulic head and a spatial filter to measure conflicting management priorities. Cumulative change in storage over the model period can be seen in Figure 2.10a. The three management alternatives show similar total reduction in storage deficit over time as compared with the historical cumulative groundwater storage deficit, with exception of the combined alternative, which essentially sums the deficit savings of the other three alternatives. However, in Figure 2.9 it is apparent that the repair leaks

and increased wastewater infiltration alternatives exhibit more distributed increases in hydraulic head during the 1984 to 2014 model period than the recharge basin alternative, with the combined alternative resulting in approximately two-fold increases in the hydraulic head in comparison with the other three alternatives throughout the majority of the model area.

These differences in the distribution of hydraulic head are captured to varying degrees in the spatially and temporally computed objectives (Figure 2.10b). Among these objectives, energy use and water quality maximize groundwater head, while urban flooding avoidance minimizes hydraulic head. Additionally, both the energy use and water quality objectives are spatially concentrated over the alluvial aquifer, which has fairly evenly distributed hydraulic head changes in both the wastewater infiltration and repair leaks alternatives, causing them to both outperform the historical and recharge basins alternatives. However, the combined alternative performs much better in relation to the other alternatives in the water quality objective when compared with its performance in the energy use objective. This is likely a factor of the depth at which the water quality objective is being measured. Specifically, the clay layer becomes thinner at the edges, particularly in the northwest corner of the clay formation. Therefore, hydraulic head must increase much more to surpass its elevation in those areas, requiring the combined head increases from all three alternatives to improve further in this objective. Alternatively, the pumping alternative is mostly concentrated along the western portion of the basin, where there are more modest increases in hydraulic head compared to the historical conditions. The repair leaks alternative should be expected to perform better than the wastewater and basin infiltration alternatives in the energy use objective because pumping is reduced at the exact locations where the energy use objective is measured.

The historical scenario minimizes urban flooding, but only because it causes the greatest drawdown in urban areas. The recharge basins alternative performs slightly better than both the wastewater reuse and leak repair alternatives in urban flooding, likely a result of the proximity of wastewater treatment plants and pumping locations to urban areas, while recharge basins are located away from urban areas as required by the model. However, the combined alternative performed much worse than any of the other alternatives. Locatelli et al. (2017) found that stormwater

infiltration, when combined with increasing recharge from urbanization, led to long term increases in the groundwater table and decreased the effectiveness of infiltration basins. This is consistent with the warnings raised by Bonneau et al. (2017) that recharge efforts may lead to localized rises in the groundwater table. A deficit in the current model is the lack of drainage points throughout the system to simulate discharge to surface water routes or ponds. This could be contributing to an overestimate of groundwater mounding in certain areas of recharge, which in turn may be causing the nonlinear increase in the urban groundwater flooding objective. However, this objective reveals that the implementation of managed aquifer recharge must be accompanied by mitigation of negative effects of flooding in vulnerable urban areas.

All three objectives are based on simulated head values across the model, but they differ in the how they are aggregated spatially and temporally. This analysis demonstrates that the spatial and temporal choices of how planning objectives are calculated can have significant impacts on infrastructure decisions.

2.5.4 Model Uncertainty

In the current model, there is uncertainty associated with assumptions made to represent certain physical processes and uncertainty propagated from a lack of representation of important physical processes. Examples of assumptions made in the modeling process that could introduce large uncertainties include the calibration of parameters governing pumping flows, leak infiltration, precipitation recharge percentages, and characteristics of the geologic formations within the basin. These uncertainties would benefit from improved measurement or estimation of these properties going forward.

To resolve uncertainty associated with missing physical processes this model requires improved representation of existing surface and subsurface boundary conditions. For example, future models could incorporate the presence of seasonal springs in mountainous regions through the use of streamflow routing networks along known surface water routes. These are particularly important, given that marginalized populations have historically relied on surface sources for water supply

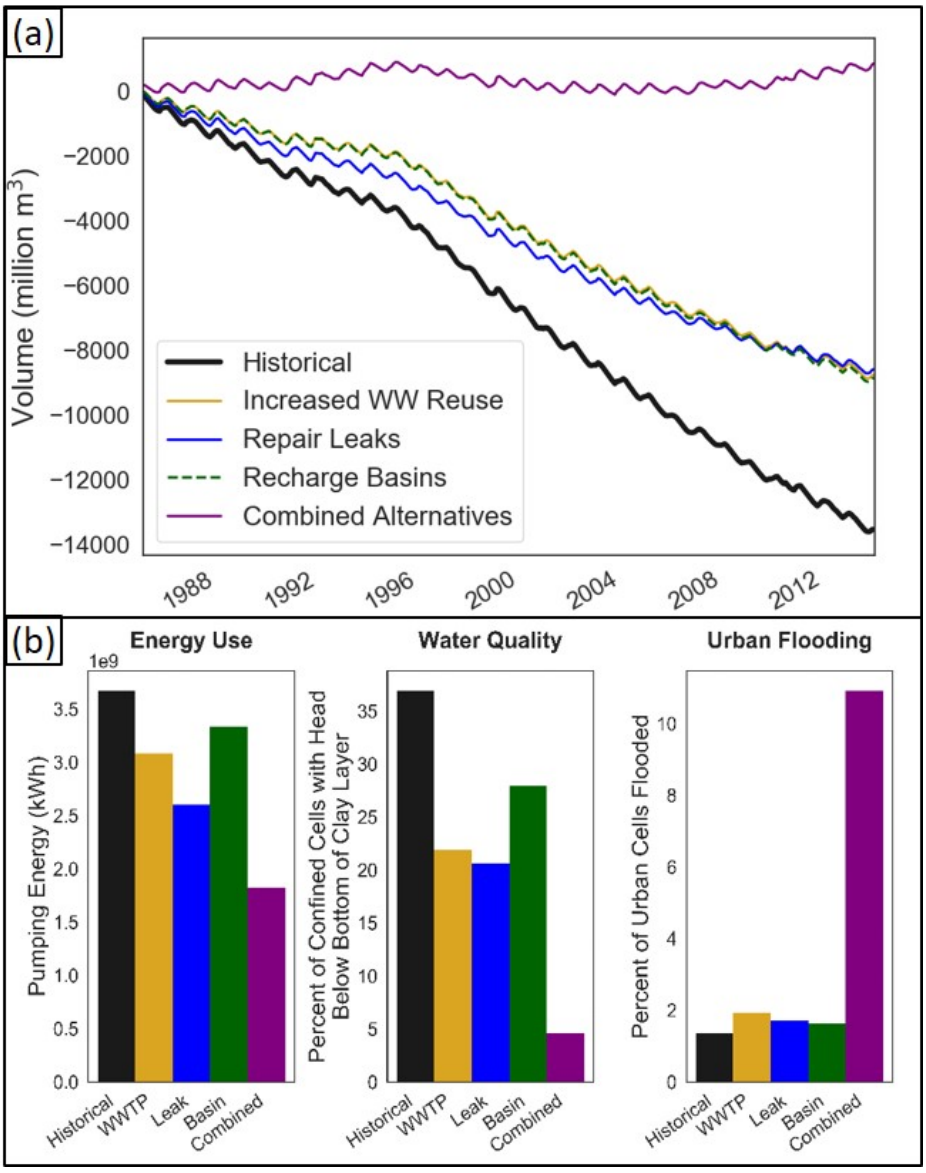

Figure 2.10: Performance of aquifer management alternatives a) according to cumulative change in storage over time, and b) under three spatially aggregated objectives: minimize pumping energy use, minimize depth to groundwater in clay layer, and minimize urban groundwater mounding.

in high population, low-income lakebed areas that are also prone to seasonal flooding. The inclusion of such flow observations are important in any inverse modeling application and would help to better constrain the model, particularly along the higher elevation boundaries for intra-annual timescales. Similarly, representation of subsidence processes and inclusion of subsidence observations would improve the physics and constraints of the model. However, such improvements to the modeled representation of physical processes and may significantly increase model run-time.

Additionally, simulations of future land use and precipitation to plan going forward would need to include projected population and climate scenarios. Such integrated modeling has been used to model aquifer management alternatives (Hanson, Boyce, Schmid, Hughes, Mehl, Leake, Maddock and Niswonger, 2014; Boyce et al., 2019) and would allow for a more accurate multiobjective comparison for this basin.

Uncertainties in the observations used to measure model error and calibrate the model are important to accurately representing historical groundwater and surface water flows. While these uncertainties were accounted for in the weighting methods of calibration process, there is always potential to not fully represent all possible sources of error in observation measurements. One such source of uncertainty arises from the partial penetration of both observation and pumping wells to multiple hydrostratigraphic units. Such partial penetration of wells is common with the deep wells that are present in the basin and could lead to either false representation or lack of representation of local anomalies. Another source is the lack of information on the timing of the measurement of well water levels during the year. Given the intra-annual changes in groundwater head at any point within the basin, this uncertainty can cause the calibration effort to try to match values that are either too high or low based on the time of year that is being represented. Each of these uncertainties has an unknown effect on the final management objectives that would in turn affect decisions made with such a model.

Finally, this model does not dynamically represent decision-making processes at the local level, and instead relies on the assumption of a top-down implementation of managed aquifer recharge alternatives through a centralized agency. While this may be a realistic assumption given the current governance framework of OCAVM, uncertainties exist concerning available resources and level of cooperation at the local agency level. Such uncertainties could limit the full implementation of any of the alternatives proposed, for example, leaks may only be repaired in a single municipality or subsection of the water supply distribution network rather than evenly across the entire system.

2.6 Conclusions

This study develops a modeling framework to analyze the effects of urban infrastructure on hydrogeological processes, focusing on the case study of the Mexico City Metropolitan Area (MCMA). The inclusion of planning objectives that span varying spatial and temporal extents to properly understand the local effects of groundwater impact mitigation schemes can be a valuable addition to modeling efforts in urban hydrogeologic systems like the MCMA. A new accounting of the possible impact of leaks and land use types on recharge identifies the potential for managed recharge actions to improve aquifer sustainability. Results indicate that management alternatives improve groundwater storage and drawdown, though with spatially distinct impacts over the historical period. Additionally, the spatial distribution of the management objectives captures conflicting performance between the alternatives, demonstrating the importance of not only the choice of management objectives, but also the sub-basin and temporal scale at which they are evaluated. The measurement of the objectives using spatial components can help to tease out the effects of the spatial signature of management alternatives on the overall groundwater security of the basin.

Future work will consider each of these objectives at the local municipality level to further analyze the spatial distribution of management impacts. Such an experiment would also allow for the application of objective weights based on other characteristics such as socio-economic status or rural versus urban areas. The set of objectives could also be expanded for an optimization problem to incorporate costs of infrastructure alternatives or the difficulty of implementing a given alternative. There are limitations to evaluating alternatives using precipitation data from a thirty-year model period that may leave out long-term cycles of wet and dry climate patterns. Future studies could implement statistically or general circulation model (GCM)-derived climate trends to evaluate recharge alternatives across ranges of reasonable precipitation.

The application of multiobjective analysis to models that include a more rigorous representation of hydrologic processes in the MCMA, including subsidence and surface water runoff, would help to further test the limits and benefits of spatially defined planning objectives. The continued refinement of this groundwater model will provide future research opportunities particularly in

advancing our understanding of the behavior of urban aquifer systems at a regional scale and in the context of water supply risks faced by vulnerable populations within and at the periphery of urban areas. The current model has been formulated to enable future spatial optimization studies in which combinations of interventions comprise a portfolio of management options, an approach with broad applicability in other rapidly urbanizing regions dependent on groundwater supplies. However, a first step in providing valuable decision-making information will be to account for the propagation of parameter uncertainties into groundwater management objectives calculated using these types of hydrologic models.

2.7 Appendix

2.7.1 Land Use Determination

Land use shape files were obtained from INEGI for data collected in 1984, 1997, 2003, 2010, 2012, and 2016 (INEGI, 2001, 2005, 2008, 2009, 2015, 2017b). The INEGI land use files include over 30 land use and vegetation markers that vary from year to year which were homogenized into three overarching categories based on their potential for allowing infiltration of precipitation: urban, natural, and water/wetland. The land use polygons were then compared to preprocessed Landsat images to determine the extent of urban land use cover at 1990, 2000, and 2010.

Landsat images were chosen from the full set of images available for the model area for one year prior to and during the year selected, prioritizing images with low percentages of cloud cover (USGS, 2013). Images were chosen in both the wet and dry season to compare visually determined land use type between the seasons. The Landsat images were preprocessed before classification as follows. All images overlaying the model area were mosaicked into one image selecting the minimum value for any pixels which were represented in more than one image. Next, any pixels with no data or data equal to 255 were filled to match surrounding data points that were between 0 and 255. The single bands were then converted to top of atmosphere (TOA) reflectance and the dark object subtraction 1 (DOS1) method was applied as further atmospheric correction.

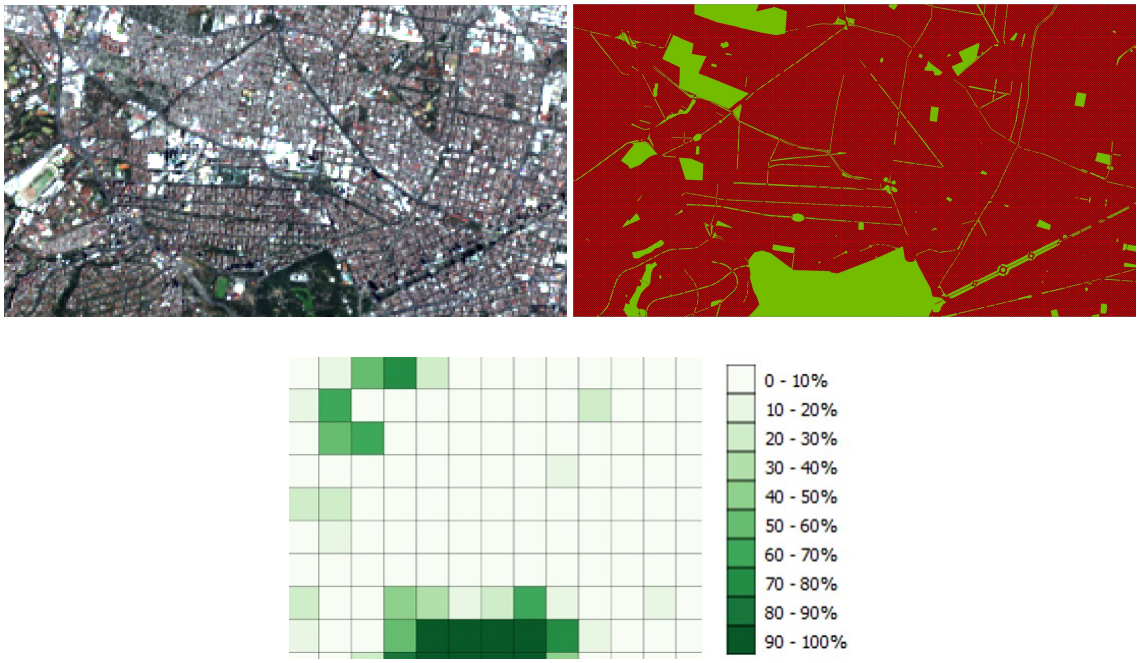

Figure 2.11: (a) Landsat image from the Polanco neighborhood of Mexico City; (b) final shape representing natural and urban land uses from the three data sources: INEGI land use, time specific Landsat imagery, and INEGI topographically determined greenspace; (c) percentage of natural land use type by model cell.

Next, the topographic shape map for urban infrastructure including parks, medians, and other greenspace was superimposed over the resulting map to exclude these features from the urban land use type (2.11). Finally, the percentage of each of the three land use types was determined according to the 500 m by 500 m model cells used in the groundwater model. The full model extent for each of the three land use types is shown for 1990 in 2.12 below.

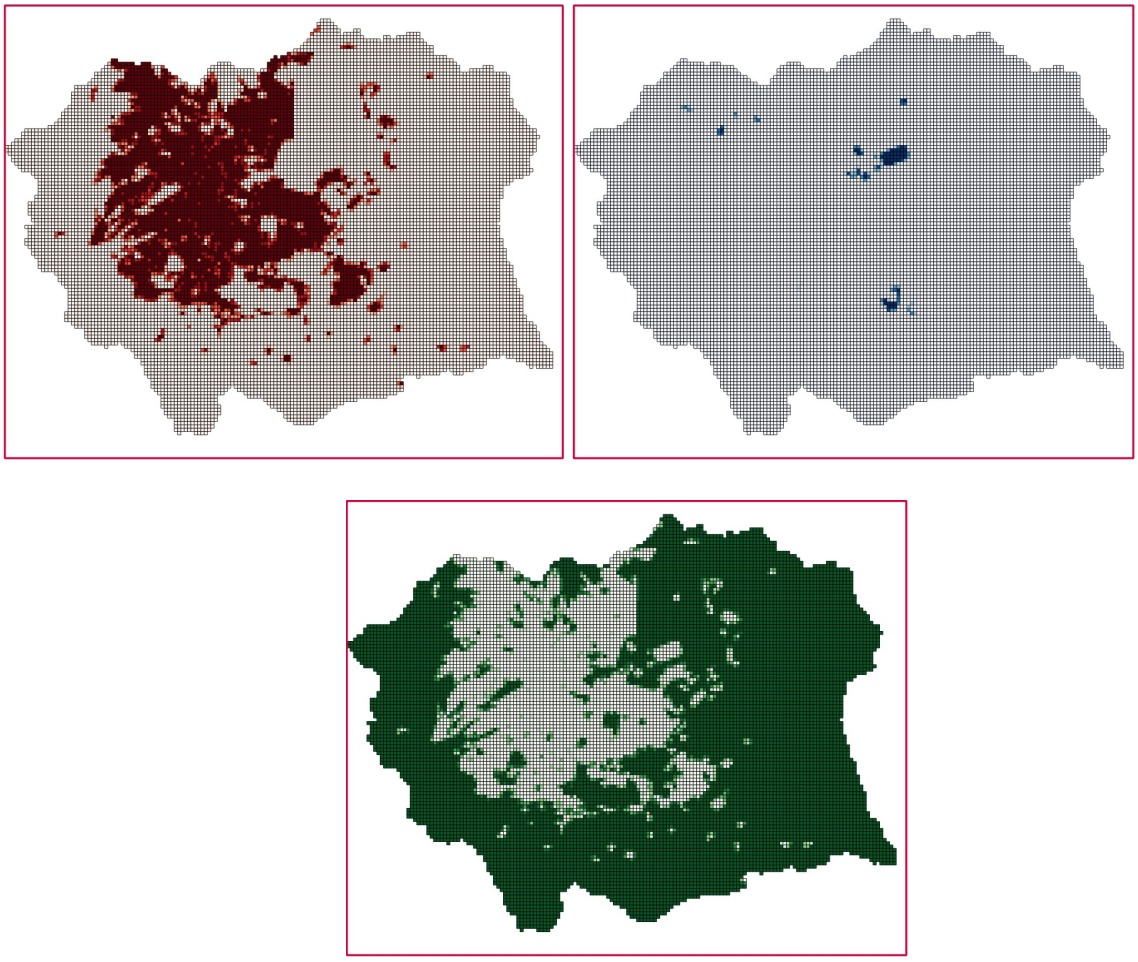

Figure 2.12: Cell by cell percentages for (a) urban, (b) water, and (c) natural land use types for 1990

2.7.2 Selected Infiltration Basins

Figure 2.13: Proposed location for infiltration basin 1, randomly selected from all potential basin locations discussed in Section 2.4.3.1 Infiltration Basins.

Figure 2.14: Proposed location for infiltration basin 2, randomly selected from all potential basin locations discussed in Section 2.4.3.1 Infiltration Basins.

Figure 2.15: Proposed location for infiltration basin 3, randomly selected from all potential basin locations discussed in Section 2.4.3.1 Infiltration Basins.

Figure 2.16: Proposed location for infiltration basin 4, randomly selected from all potential basin locations discussed in Section 2.4.3.1 Infiltration Basins.

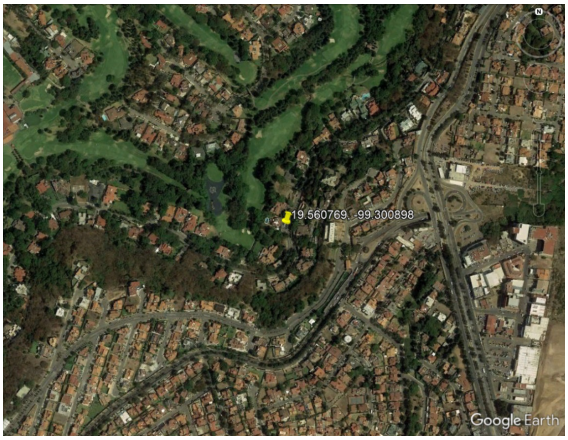

Figure 2.17: Proposed location for infiltration basin 5, randomly selected from all potential basin locations discussed in Section 2.4.3.1 Infiltration Basins.

2.7.3 Leak Calculations

2.7.3.1 Estimated non-groundwater supply sources

Table 2.4 Estimated values for non-groundwater supply

Source	1990	2000	2010
Cutzamala	6.5	15	15.5
Lerma	6.6	9.4	6.2
PAI	0	0	4.5
Internal Surface	1	1	1.5
Internal WW Reuse	2.6	2.6	5.8

2.7.3.2 Derivation of new groundwater pumping relationship from repair leaks scenario

$$G = T - s \tag{2.1}$$

$$G_n - G = T_n - T \tag{2.2}$$

$$G_n = G - (T_n - T) \tag{2.3}$$

$$\frac{G_n}{G} = \frac{G - (T_n - T)}{G} \tag{2.4}$$

$$\frac{G_n}{G} = \frac{T - s - (T[1 - P_L(1 - L_P)] - T)}{T - s} \tag{2.5}$$

$$\frac{G_n}{G} = \frac{T(1 + P_L(1 - L_P)) - s}{T - s} \tag{2.6}$$

2.7.4 Model Parameters and their Calibrated Values

Table 2.5: Model Parameters

Parameter	Units	Lower Bound	Upper Bound	Initial Value	Calibrated Value
Zonal Geologic					
$HK_{lacustrine}$	m/d	8.64E-05	5.00E-02	5.00E-03	5.00E-02
$HK_{alluvial}$	m/d	1.00E-01	1.00E+02	3.00E+01	1.00E+02
$HK_{fracturedbasalt}$	m/d	3.46E-02	1.50E+02	5.00E+0	9.32E-01
$HK_{Tarango}$	m/d	4.32E-02	4.32E+01	5.00E-01	3.47E-01
$HK_{andesitic}$	m/d	4.32E-03	8.64E+01	5.00E-02	-
$S_{s,lacustrine}$	1/m	9.19E-04	2.03E-02	2.00E-02	-
$S_{s,alluvial}$	1/m	4.92E-05	1.05E-03	1.00E-04	5.59E-05
$S_{s,fracturedbasalt}$	1/m	1.00E-07	6.89E-05	2.00E-06	-

Parameter	Units	Lower Bound	Upper Bound	Initial Value	Calibrated Value
$S_{s,Tarango}$	1/m	1.00E-07	1.02E-04	2.00E-06	-
$S_{s,andesitic}$	1/m	1.00E-07	6.89E-05	2.00E-06	-
$S_{y,lacustrine}$	-	0.001	0.08	0.06	-
$S_{y,alluvial}$	-	0.05	0.4	0.15	0.13
$S_{y,fracturedbasalt}$	-	0.01	0.2	0.10	0.14
$S_{y,Tarango}$	-	0.05	0.4	0.15	0.27
$S_{y,andesitic}$	-	0.001	0.1	0.01	-
$VANI_{lacustrine}$	-			5	-
$VANI_{alluvial}$	-			10	-
$VANI_{fracturedbasalt}$	-			1	-
$VANI_{Tarango}$	-			10	-
$VANI_{andesitic}$	-			10	-
Time Varied Infrastructure	**multiplier**				
$Q_{urban,1990}$	-			1.5	2.653
$Q_{urban,2000}$	-			1.2	1.467
$Q_{urban,2010}$	-			1.0	0.464
LK_{1990}	-			1	-
LK_{2000}	-			1	-
LK_{2010}	-			1	-
TWU_{1990}	-			0.8	0.835
TWU_{2000}	-			1.0	1.175
TWU_{2010}	-			1.2	1.202
Zonal Recharge	**recharge**				
RCH_{urban}	%			1	-
$RCH_{natural}$	%			30	35.8
$RCH_{water/wetland}$	%			50	-
Leak Infiltration	**infiltration**				
IN	%			10	-

2.7.5 Simulated and Observed Hydraulic Head

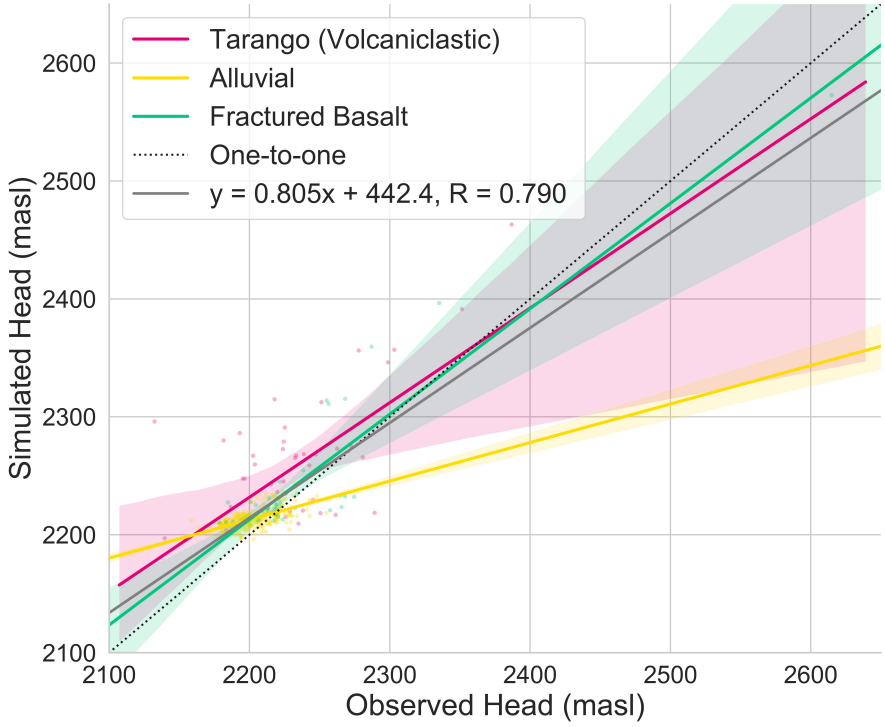

Figure 2.18: First simulated head for each observation well versus first observed head for each observation well. The shaded regions represent the 95% confidence interval for the linear regression performed.

Table 2.6 Mean absolute error (MAE) between weighted and unweighted simulated and observed hydraulic head at observation wells for each geologic formation

		Mean Absolute Error (m)	
Formation	**RMSE**	**Unweighted**	**Weighted**
Overall	18.33 m^2	7.20	5.72
Alluvial		5.75	4.78
Fractured Basalt		9.61	7.55
Tarango		14.20	10.01

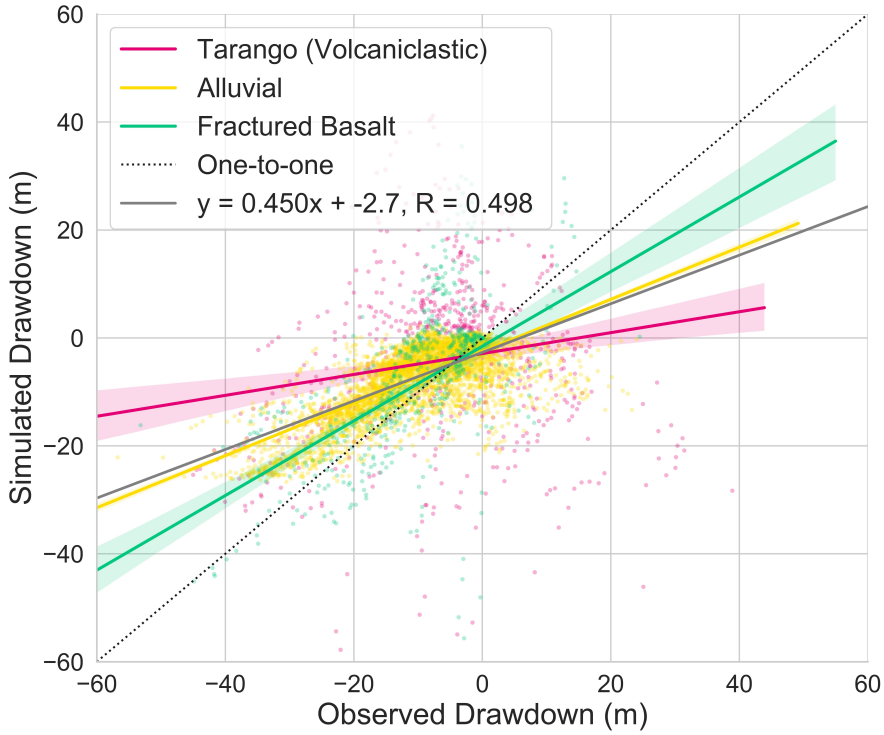

Figure 2.19: Simulated drawdown for each observation well versus observed drawdown for each observation well. The shaded regions represent the 95% confidence interval for the linear regression performed.

2.8 Data Availability

The model with input datasets, observations, results, and postprocessing scripts are available in a GitHub repository at https://github.com/mrlmautner/UrbanGW.

2.9 Acknowledgments

This work has been supported in part by the Ford Foundation Predoctoral Fellowship Program of the National Academies of Science, Engineering, and Medicine. Research trips to Mexico City were funded in part by the University of California Davis Henry A. Jastro Graduate Research Award. We would also like to thank the Organismo de Cuencas: Aguas del Valle de México (OCAVM) of the National Water Commission (CONAGUA) of Mexico and the Instituto de Geofísica of the Universidad Nacional Autónoma de México (UNAM) for providing pumping and observation data,

the conceptual groundwater model, and input on potential management alternatives. Dr. Graciela Herrera would like to thank the Instituto de Geofísica for providing a MODFLOW model that was developed by members of this Institute from 1994 to 2018.

Bibliography

Abboud, J., Ryan, M. and Osborn, G. (2018). Groundwater flooding in a river-connected alluvial aquifer, *Journal of Flood Risk Management* **11**(4): e12334.

Alley, W. M. and Leake, S. A. (2004). The Journey from Safe Yield to Sustainability, *Ground Water* **42**(1): 12–16.

Aragón-Durand, F. (2007). Urbanisation and flood vulnerability in the peri-urban interface of Mexico City, *Disasters* **31**(4): 477–494.

Bach, P. M., Rauch, W., Mikkelsen, P. S., McCarthy, D. T. and Deletic, A. (2014). A critical review of integrated urban water modelling – Urban drainage and beyond, *Environmental Modelling & Software* **54**: 88–107.

Baker, J. L. (ed.) (2012). *Climate Change, Disaster Risk, and the Urban Poor*, The World Bank.

Bakker, M., Post, V., Langevin, C. D., Hughes, J. D., White, J. T., Starn, J. J. and Fienen, M. N. (2016). Scripting MODFLOW Model Development Using Python and FloPy, *Groundwater* **54**(5): 733–739.

Banco Interamericano de Desarrollo (BID) (2012). Modelo de Marco Institucional para la Gestión de los Recursos Hídricos en el Valle de México: Primer Informe, *Technical report*, Banco Interamericano de Desarrollo (BID).

Banzhaf, H. S. (2009). Objective or Multi-Objective? Two Historically Competing Visions for Benefit-Cost Analysis, *Land Economics* **85**(1): 3–23.

Birkle, P., Torres Rodríguez, V. and González Partida, E. (1998). The water balance for the Basin of the Valley of Mexico and implications for future water consumption, *Hydrogeology Journal* **6**(4): 500–517.

Bond, S. D., Carlson, K. A. and Keeney, R. L. (2008). Generating Objectives: Can Decision Makers Articulate What They Want?, *Management Science* **54**(1): 56–70.

Bonneau, J., Fletcher, T. D., Costelloe, J. F. and Burns, M. J. (2017). Stormwater infiltration and the 'urban karst' – A review, *Journal of Hydrology* **552**: 141–150.

Boyce, S. E., Hanson, R. T., Ferguson, I. M., Schmid, W., Reihmann, T., Mehl, S. M. and Henson, W. (2019). One-Water Hydrologic Flow Model Version 2 – MODFLOW-OWHM: U.S. Geological Survey Techniques and Methods 6-Axx, v. p., *Technical report*, U.S. Geologic Survey.

Brath, A., Montanari, A. and Moretti, G. (2006). Assessing the effect on flood frequency of land use change via hydrological simulation (with uncertainty), *Journal of Hydrology* **324**(1-4): 141–153.

Braud, I., Breil, P., Thollet, F., Lagouy, M., Branger, F., Jacqueminet, C., Kermadi, S. and Michel, K. (2013). Evidence of the impact of urbanization on the hydrological regime of a medium-sized periurban catchment in France, *Journal of Hydrology* **485**: 5–23.

Bredehoeft, J. D. (2002). The Water Budget Myth Revisited: Why Hydrogeologists Model, *Ground Water* **40**(4): 340–345.

CAEM (2011). Well observations in the service area of the Comisión del Agua del Estado de México.

Carrera-Hernández, J. J. and Gaskin, S. J. (2008). Spatio-temporal analysis of potential aquifer recharge: Application to the Basin of Mexico, *Journal of Hydrology* **353**(3-4): 228–246.

Cohon, J. L. and Marks, D. H. (1975). A review and evaluation of multiobjective programing techniques, *Water Resources Research* **11**(2): 208–220.

CONAGUA (2005). *Sistema Cutzamala: agua para millones de mexicanos*, 4th edn, Comision Nacional del Agua (CONAGUA), México.

CONAGUA (2013). Well pumping concessions between 2005 and 2013.

CONAGUA (2015). Inventario nacional de plantas municipales de potabilización y de tratamiento de aguas residuales en operación, *Technical report*, Subdirección General de Agua Potable,

Drenaje y Saneamiento, Comisión Nacional del Agua, Mexico City.

URL: https://www.gob.mx/cms/uploads/attachment/file/197610/Inventario_2015.pdf

CONAGUA (2017). Well observations in the service area of the Sistema de Aguas de la Ciudad de México (SACM).

CONAPO (2014). Proyecciones de la población de los municipios que componen las zonas metropolitanas, 2010-2030, *Technical report*, Consejo Nacional de Población (CONAPO), Mexico City.

URL: http://www.conapo.gob.mx/work/models/CONAPO/Resource/1206/2/images/Proyecciones_de_Poblacion_ZM.xlsx

Dahlke, H. E., LaHue, G. T., Mautner, M. R., Murphy, N. P., Patterson, N. K., Waterhouse, H., Yang, F. and Foglia, L. (2018). Managed Aquifer Recharge as a Tool to Enhance Sustainable Groundwater Management in California: Examples From Field and Modeling Studies, *in* J. Friesen and L. Rodríguez-Sinobas (eds), *Advances in Chemical Pollution, Environmental Management and Protection*, Elsevier, chapter 8.

Daw, T., Brown, K., Rosendo, S. and Pomeroy, R. (2011). Applying the ecosystem services concept to poverty alleviation: The need to disaggregate human well-being, *Environmental Conservation* **38**(4): 370–379.

De Nys, E., Funes, E. S., Contijoch-Escontria, M., González-Reynoso, A. E., Ramos-Bustillo, L. E., Calderon-Bartheneuf, J. L., Menendez-Martínez, C., Aparicio, F. J., Uyytendaele, G. P., González, L. M., Simas, J., Torres-Millan, R. N., Hernandez, A. and Villanueva-Moreno, J. A. (2015). Cutzamala - Diagnostico integral.

URL: http://documents.worldbank.org/curated/en/309801468189248037/Cutzamala-Diagnostico-integral

Dendrou, S. A. (1982). *Overview of Urban Stormwater Models*, Vol. 7.

DGCOH (1997). Plan Maestro de Agua Potable del Distrito Federal 1997-2010, *Technical report*, Dirección General de Construcción y Operación Hidráulica (DGCOH), Distrito Federal, Mexico.

Escolero, O., Kralisch, S., Martínez, S. E. and Perevochtchikova, M. (2016). Diagnóstico y análisis de los factores que influyen en la vulnerabilidad de las fuentes de abastecimiento de agua potable a la Ciudad de México, México, *Boletín de la Sociedad Geológica Mexicana* **68**(3): 409–427.

Foglia, L., Hill, M. C., Mehl, S. W. and Burlando, P. (2009). Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function, *Water Resources Research* **45**(6).

Foster, S., Lawrence, A. and Morris, B. (1998). Groundwater in urban development, *World Bank Technical Paper* **390**: 1–55.

Frigo, M., Ferronato, M., Yu, J., Ye, S., Galloway, D., Carreón-Freyre, D. and Teatini, P. (2019). A Parametric Numerical Analysis of Factors Controlling Ground Ruptures Caused by Groundwater Pumping, *Water Resources Research* **55**(11): 9500–9518.

Galán-Breth, R. I. (2018). *Modelación matemática de nitratos en el agua subterránea en la región Sur de la Ciudad de México*, PhD thesis, Universidad Nacional Autónoma de México (UNAM).

Gómez-Reyes, E. (2013). Valoración de las componentes del balance hídrico usando información estadística y geográfica: la cuenca del Valle de México, *REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA* **4**(3): 4–27.

González-Morán, T., Rodríguez, R. and Cortes, S. (1999). The Basin of Mexico and its metropolitan area: water abstraction and related environmental problems, *Journal of South American Earth Sciences* **12**(6): 607–613.

Hanson, R. T. (2015). Hydrologic framework of the Santa Clara Valley, California, *Geosphere* **11**(3): 606–637.

Hanson, R. T., Boyce, S. E., Schmid, W., Hughes, J. D., Mehl, S. M., Leake, S. A., Maddock, T. I. and Niswonger, R. G. (2014). One-Water Hydrologic Flow Model (MODFLOW-OWHM). U.S. Geological Survey Techniques and Methods 6–A51, *Technical report*, USGS.

Hanson, R. T., Flint, L. E., Flint, A. L., Dettinger, M. D., Faunt, C. C., Cayan, D. and Schmid, W. (2012). A method for physically based model analysis of conjunctive use in response to potential climate changes, *Water Resources Research* **48**(6): 1–23.

Hanson, R. T., Lockwood, B. and Schmid, W. (2014). Analysis of projected water availability with current basin management plan, Pajaro Valley, California, *Journal of Hydrology* **519**(PA): 131–147.

Harbaugh, A. W. (2006). MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model—the Ground-Water Flow Process, *U.S. Geological Survey Techniques and Methods: Book 6*, USGS, Reston, Virginia, chapter 16, pp. 1–253.

Herrera-Revilla, I., Medina-Bañuelos, R., Carrillo-Rivera, J. and Vazquez-Sánchez, E. (1994). Diagnóstico del Estado Presente de las Aguas Subterráneas de la Ciudad de México y Determinación de sus Condiciones Futuras, *Technical Report 3–33–1–6684*, DGCOH, DDF, Instituto de Geofísica, UNAM., México, D.F.

Herrera-Zamarrón, G., Cardona-Benavides, A., González-Hita, L., Gutiérrez-Ojeda, C., Hernández-Calero, R., Hernández-García, G., Hernández-Laloth, N., López-Hernández, R. I., Martínez-Morales, M., Pita de la Paz, C., Sánchez-Díaz, L. F., Báez-Durán, J. A., Cruickshank-Villanueva, C. and Herrera-Revilla, I. (2005). Estudio para obtener la disponibilidad del acuífero de la Zona Metropolitana de la Ciudad de México, *Technical Report Contract No. 06-CD-03-1O-0272-1-06*, Secretaría del Medio Ambiente del Gobierno del Distrito Federal, Sistema de Aguas de la Ciudad de México (SACM), and Instituto Mexicano de Tecnología del Agua (IMTA), Mexico City.

Hill, M. C., Kavetski, D., Clark, M., Ye, M., Arabi, M., Lu, D., Foglia, L. and Mehl, S. (2016). Practical Use of Computationally Frugal Model Analysis Methods, *Groundwater* **54**(2): 159–170.

Hill, M. C. and Tiedeman, C. R. (2007). *Effective groundwater model calibration : with analysis of data, sensitivities, predictions, and uncertainty*, Wiley-Interscience.

Huizar-Alvarez, R., Carrillo-Rivera, J., Ángeles-Serrano, G., Hergt, T. and Cardona, A. (2004). Chemical response to groundwater extraction southeast of Mexico City, *Hydrogeology Journal* **12**(4): 436–450.

Huizar-Alvarez, R., Ouysse, S., Espinoza-Jaramillo, M. M., Carrillo-Rivera, J. J. and Mendoza-Archundia, E. (2016). The effects of water use on Tothian flow systems in the Mexico City conurbation determined from the geochemical and isotopic characteristics of groundwater, *Environmental Earth Sciences* **75**(13): 1060.

Ibarraran, M. E. (2011). Cities and Climate Change: Global Report on Human Settlements 2011, Climate's Long-term Impacts on Mexico's City Urban Infrastructure, *Technical report*, United Nations Human Settlements Programme.

INEGI (2001). Land Use and Vegetation, 1:250000 (national extent).
URL: *http://www.conabio.gob.mx/informacion/metadata/gis/usv250kcs1agw.xml?_httpcache=yes&_xsl=/db/metadata/xsl/fgdc_html.xsl&_indent=no*

INEGI (2002). Conjunto de Datos Vectoriales Geológicos. Continuo Nacional. Scale 1:1,000,000.

INEGI (2005). Land Use and Vegetation, 1:250000 (national extent).
URL: *http://www.conabio.gob.mx/informacion/metadata/gis/usv250ks2gw.xml?_httpcache=yes&_xsl=/db/metadata/xsl/fgdc_html.xsl&_indent=no*

INEGI (2008). Land Use and Vegetation, 1:250000 (national extent).
URL: *http://www.conabio.gob.mx/informacion/metadata/gis/usv250ks3gw.xml?_httpcache=yes&_xsl=/db/metadata/xsl/fgdc_html.xsl&_indent=no*

INEGI (2009). Land Use and Vegetation, 1:250000 (national extent).

URL: *http://www.conabio.gob.mx/informacion/metadata/gis/usv250ks4gw.xml?_httpcache=yes&_xsl=/db/metadata/xsl/fgdc_html.xsl&_indent=no*

INEGI (2010). Subcuenca Hidrográfica RH26Dp 1. Texcoco y Zumpango/Cuenca R. Moctezuma/R.H. Pánuco.

INEGI (2015). Land Use and Vegetation, 1:250000 (national extent).

URL: *http://www.conabio.gob.mx/informacion/metadata/gis/usv250s5ugw.xml?_httpcache=yes&_xsl=/db/metadata/xsl/fgdc_html.xsl&_indent=no*

INEGI (2017a). Conjuntos de datos vectoriales de información topográfica por Entidad Federativa serie V 2016-2017. Scale 1:250000.

URL: *http://www.beta.inegi.org.mx/temas/mapas/topografia/*

INEGI (2017b). Land Use and Vegetation, 1:250000 (national extent).

URL: *http://www.conabio.gob.mx/informacion/metadata/gis/usv250s6gw.xml?_httpcache=yes&_xsl=/db/metadata/xsl/fgdc_html.xsl&_indent=no*

Jacobson, C. R. (2011). Identification and quantification of the hydrological impacts of imperviousness in urban catchments: A review, *Journal of Environmental Management* **92**(6): 1438–1448.

Jago-on, K. A. B., Kaneko, S., Fujikura, R., Fujiwara, A., Imai, T., Matsumoto, T., Zhang, J., Tanikawa, H., Tanaka, K., Lee, B. and Taniguchi, M. (2009). Urbanization and subsurface environmental issues: An attempt at DPSIR model application in Asian cities, *Science of The Total Environment* **407**(9): 3089–3104.

Lawrence, A. R., Morris, B. L. and Foster, S. S. D. (1998). Hazards induced by groundwater recharge under rapid urbanization, *Geological Society, London, Engineering Geology Special Publications* **15**(1): 319–328.

Lerner, D. N. (1990). Groundwater recharge in urban areas, *Atmospheric Environment. Part B. Urban Atmosphere* **24**(1): 29–33.

Lerner, D. N. (2004). *Urban groundwater pollution*, International contributions to hydrogeology ; v. 24, A.A. Balkema, Lisse, The Netherlands ; Exton, PA.

Li, C., Fletcher, T. D., Duncan, H. P. and Burns, M. J. (2017). Can stormwater control measures restore altered urban flow regimes at the catchment scale?, *Journal of Hydrology* **549**: 631–653.

Locatelli, L., Mark, O., Mikkelsen, P. S., Arnbjerg-Nielsen, K., Deletic, A., Roldin, M. and Binning, P. J. (2017). Hydrologic impact of urbanization with extensive stormwater infiltration, *Journal of Hydrology* **544**: 524–537.

Lopez-Alvis, J. (2014). *Calibración de un modelo de flujo del Acuífero de la Zona Metropolitana de la Ciudad de México (AZMCM)*, PhD thesis, Universidad Nacional Autónoma de México (UNAM).

Macdonald, D., Bloomfield, J., Hughes, A., MacDonald, A., Adams, B. and McKenzie, A. (2008). Improving the understanding of the risk from groundwater flooding in the UK, *Flood Risk Management: Research and Practice*, CRC Press, pp. 1071–1080.

Maliva, R. G. (2020). Unmanaged and Unintentional Recharge, *in* J. C. Santamarta Cerezal (ed.), *Anthropogenic Aquifer Recharge*, wsp method edn, Springer, Fort Meyes, Florida, chapter 24, pp. 827–854.

Mazari, M. and Mackay, D. M. (1993). Potential for groundwater contamination in Mexico City, *Environmental Science & Technology* **27**(5): 794–802.

McDonald, R. I., Weber, K., Padowski, J., Flörke, M., Schneider, C., Green, P. A., Gleeson, T., Eckman, S., Lehner, B., Balk, D., Boucher, T., Grill, G. and Montgomery, M. (2014). Water on an urban planet: Urbanization and the reach of urban water infrastructure, *Global Environmental Change* **27**: 96–105.

McLaughlin, D. and Townley, L. R. (1996). A reassessment of the groundwater inverse problem.

Mitchell, V. G., Mein, R. G. and McMahon, T. A. (2001). Modelling the urban water cycle, *Environmental Modelling & Software* **16**(7): 615–629.

National Research Council, The Joint Academies Committee on the Mexico City Water Supply, Water Science and Technology Board, Commission on Geosciences, E. and Resources (1995). *Mexico City's Water Supply*, National Academies Press, Washington, D.C.

OCAVM (2014). Programa Hídrico Regional 2014-2018: Region Administrativo Hidrológico XIII, Aguas del Valle de México, *Technical report*, Comisión Nacional del Agua, Tlalpan, Mexico, D.F.

Ortega G., A. and Farvolden, R. N. (1989). Computer analysis of regional groundwater flow and boundary conditions in the basin of Mexico, *Journal of Hydrology* **110**(3-4): 271–294.

Ortega-Guerrero, A., Cherry, J. A. and Rudolph, D. L. (1993). Large-Scale Aquitard Consolidation Near Mexico City, *Ground Water* **31**(5): 708–718.

Ortiz-Zamora, D. and Ortega-Guerrero, A. (2010). Evolution of long-term land subsidence near Mexico City: Review, field investigations, and predictive simulations, *Water Resources Research* **46**(1).

Page, D., Bekele, E., Vanderzalm, J. and Sidhu, J. (2018). Managed Aquifer Recharge (MAR) in Sustainable Urban Water Management, *Water* **10**(3): 239.

Palma Nava, A., Cruickshank Villanueva, C., González Villarreal, F., Hanson, R. T. and Boyce, S. E. (2015). A New Integrated Hydrologic Model for Mexico Valley, Mexico City, Mexico, *MODFLOW and More 2015: Modeling a Complex World – Advances in Integrated Hydrologic Modeling*, Golden, Colorado.

Poeter, E. P., Hill, M. C., Lu, D., Tiedeman, C. and Mehl, S. W. (2014). UCODE_2014, with new capabilities to define parameters unique to predictions, calculate weights using simulated values,

estimate parameters with SVD, evaluate uncertainty with MCMC, and more, *Technical report*, Integrated Groundwater Modeling Center (IGWMC), of the Colorado School of Mines.

Reichard, E. G., Li, Z. and Hermans, C. (2010). Emergency use of groundwater as a backup supply: Quantifying hydraulic impacts and economic benefits, *Water Resources Research* **46**(9): 1–20.

Riveros-Olivares, B. (2013). *Tratamiento de Aguas Residuales Municipales en la Ciudad de México*, PhD thesis, Universidad Nacional Autónoma de México (UNAM).
URL: *http://www.ptolomeo.unam.mx:8080/xmlui/bitstream/handle/132.248.52.100/3205/Tesis_BrunoRiveros.pdf?sequence=1*

Roy-Poirier, A., Champagne, P. and Filion, Y. (2010). Review of Bioretention System Research and Design: Past, Present, and Future, *Journal of Environmental Engineering* **136**(9): 878–889.

SACM (2012). El gran reto del agua en la Ciudad de México, *Technical report*, Sistema de Aguas de la Ciudad de México (SACM), Mexico City.
URL: *https://agua.org.mx/wp-content/uploads/2013/02/El-gran-reto-del-agua-en-la-Ciudad-de-Mexico.pdf*

Singhal, B. B. S. and Gupta, R. P. (1999). Hydraulic properties of rocks, *Applied Hydrogeology of Fractured Rocks*, Springer Netherlands, Dordrecht, pp. 151–168.

SMN (2015). Climate Computing System (CLICOM-CICESE), Daily Climatic Data.
URL: *http://clicom-mex.cicese.mx/*

Sosa-Rodriguez, F. S. (2010). Impacts of Water-management Decisions on the Survival of a City: From Ancient Tenochtitlan to Modern Mexico City, *International Journal of Water Resources Development* **26**(4): 675–687.

Tellman, B., Bausch, J. C., Eakin, H., Anderies, J. M., Mazari-Hiriart, M., Manuel-Navarrete, D. and Redman, C. L. (2018). Adaptive pathways and coupled infrastructure: Seven centuries of

adaptation to water risk and the production of vulnerability in Mexico city, *Ecology and Society* **23**(1): art1.

Thompson, S. E., Sivapalan, M., Harman, C. J., Srinivasan, V., Hipsey, M. R., Reed, P. M., Montanari, A. and Blöschl, G. (2013). Developing predictive insight into changing water systems: use-inspired hydrologic science for the Anthropocene, *Hydrology and Earth System Sciences* **17**(12): 5013–5039.

Toll, D. G., Abedin, Z., Buma, J., Cui, Y.-j., Osman, A. S. and Phoon, K. K. (2012). The impact of changes in the water table and soil moisture on structural stability of buildings and foundation systems, *Technical report*.

Torres-Bernardino, L. (2014). *Sistema Lerma: una visión política en la gestión pública del agua, ¿solución Estatal o Federal?*, Instituto de Administración Pública del Estado de México, A.C., Toluca, México.

United Nations (2016). The World's Cities in 2016, *Technical report*.

USGS (2013). Landsat 5.

Varis, O., Biswas, A. K., Tortajada, C. and Lundqvist, J. (2006). Megacities and water management, *International Journal of Water Resources Development* **22**(2): 377–394.

Vörösmarty, C. J., Green, P., Salisbury, J. and Lammers, R. B. (2000). Global water resources: vulnerability from climate change and population growth., *Science (New York, N.Y.)* **289**(5477): 284–8.

Zhou, Y. and Li, W. (2011). A review of regional groundwater flow modeling, *Geoscience Frontiers* **2**(2): 205–214.

Chapter 3

Coupled effects of observation and parameter uncertainty on urban groundwater infrastructure decisions[1]

3.1 Abstract

Urban groundwater management requires complex environmental models to represent interactions between hydrogeological processes and infrastructure systems. While the impacts of external uncertainties, such as climate and population growth, have been widely studied, there is limited understanding of how decision support is altered by endogenous uncertainties arising from model parameters and observations used for calibration. This study investigates (1) the importance of observation choice and parameter values on aquifer management objectives when controlling for model error, and (2) how the relative performance of management alternatives varies when exposed to endogenous uncertainties, individually and in combination. We use a spatially distributed groundwater model of the Valley of Mexico, where aquifer management alternatives include demand management, targeted infiltration, and wastewater reuse. The effects of uncertainty are evaluated using global sensitivity analysis, performance ranking of alternatives under a range of human-natural parameters, and identification of behavioral parameter sets filtered with an error metric calculated from varying subsets of observations. Results show that the parameters govern-

[1]This chapter has been published: Mautner, M. R. L., Foglia, L., and Herman, J. D. (2022). "Coupled effects of observation and parameter uncertainty on urban groundwater infrastructure decisions". *Hydrology and Earth System Sciences* ***26*** (5): 1319-1340.

ing hydraulic conductivity and total water use in the basin have the greatest effect on management objectives. Error metrics (i.e., squared residuals of piezometric head) are not necessarily controlled by the same parameters as the head based objectives needed for decision-making. Additionally, observational and parameter uncertainty each play a larger role in objective variation than the management alternatives themselves. Finally, coupled endogenous uncertainties have amplifying effects on decision-making, leading to larger variations in the ranking of management alternatives than each on their own. This study highlights how the uncertain parameters of a physically-based model and their interactions with uncertain observations can affect water supply planning decisions in densely populated urban areas.

3.2 Introduction

Groundwater resource planning and management requires increasingly complex models to represent interactions between hydrogeological and infrastructure systems to achieve sustainability (Megdal et al., 2015; Singh, 2014; Wada et al., 2017; Peters-Lidard et al., 2017). A key challenge for model-based decision support is understanding the influence of multiple sources of uncertainty on the choice of infrastructure alternatives. In particular, the role of external uncertainties such as future climate, population, and land use change, have been investigated extensively in the systems analysis field (Hadka et al., 2015; Maier et al., 2016; Kwakkel and Haasnoot, 2019). Similar approaches have been applied in groundwater systems to analyze the combined effects of perturbations in external forcing (Dams et al., 2008, 2012; Mustafa et al., 2019; Fletcher et al., 2019). However, the endogenous uncertainties arising from physically-based hydrologic and hydrogeologic models are often neglected in infrastructure planning studies, despite often influencing predictions as much or more than external drivers (Mendoza et al., 2016; Qiu et al., 2019; Herman et al., 2020). Further, the effects of endogenous model uncertainties on model error may be different from their effects on the ranking of alternatives, and therefore on decision making. This difference has been largely understudied and is the focus of this paper.

Physically-based groundwater models can support infrastructure decisions by ranking alternatives according to their performance under stakeholder-defined management objectives. Global sensitivity analyses of the ranking of alternatives have generally focused on the influence of objective values and weights in multi-criteria decision models, without providing a physical basis for the determination of such variations (Hyde and Maier, 2006; Ganji et al., 2016). As a result, these decision models often do not account for uncertainty in hydrologic processes, leaving an opportunity to relate processes to the criteria values that are produced for a given management alternative. For example, Ravalico et al. (2009, 2010) analyze the effects of parameter changes in on the optimal policy ranking by determining the minimum, median, and maximum parameter values that change the ranking of alternatives based on a single management objective; however, their implementation did not address model error. Specifically, none of the existing approaches explicitly evaluates the relationship between uncertain endogenous model characteristics used to determine model error and ranking of management alternatives for decision-making based on model output.

In hydrogeologic models, endogenous uncertainty is contributed by model parameters describing natural and human components of the system, and the set of historical observations used to calibrate or constrain the parameters (Moore and Doherty, 2005; Doherty and Simmons, 2013). Parameters provide the flexibility to represent complex systems on a broader scale, and in some cases can encapsulate differences in model structure as well (Guillaume et al., 2016). The propagation and attribution of parameter uncertainty has been the topic of numerous hydrologic modeling studies, using a combination of uncertainty analysis and sensitivity analysis (Razavi et al., 2021; Pianosi et al., 2016), though generally without considering the influence of this uncertainty on model-based decision support (Jing et al., 2019), or only focusing on local sensitivity analyses (Tolley et al., 2019). Global sensitivity analysis in particular has seen growing usage with advances in computing power (Razavi and Gupta, 2015), including sensitivity varying over time and/or space (Herman et al., 2013; Şalap-Ayça and Jankowski, 2016; Reinecke et al., 2019; Zhang and Liu, 2021), and model structure (Mai et al., 2020). Observational uncertainty is typically also excluded, except in the case of inverse modeling (Refsgaard et al., 2007).

The choice of observations to support parameter identification is often complicated by a number of factors, including: temporal and spatial representation of the model area, data quantity and quality, and resolution of datasets that determine model structure with respect to observation locations (McMillan et al., 2018; Refsgaard et al., 2012; Lehr and Lischeid, 2020). This is especially true for groundwater modeling in urban environments, where infrastructure, monitoring practices, and pumping patterns can complicate groundwater data collection procedures meant to ensure accurate and repeatable results (Foster et al., 1998; Vázquez-Suñé et al., 2010; Bhaskar et al., 2016). Uncertainty in the selection of observations will alter the parameter calibration (Montanari and Di Baldassarre, 2013), and in turn, the planning problem (Brunner et al., 2012). Similarly, Rojas et al. (2010) explore the availability and variety of observations in characterising the choice of conceptual models in multimodel analysis, again focusing on effects on model error.

When developing groundwater models for planning purposes, calibration is often carried out by selecting a "best" parameter set by minimizing one or more error metrics while adjusting parameter values, using parameter sensitivity or expert evaluation to determine which parameters to adjust. Alternatively, some calibration frameworks use observations and the resulting behavioral model space of a selected error metric to refine the distribution of parameter values, rather than optimizing a single one (Wagener et al., 2003; Bárdossy, 2007; Beven, 2016). In such calibration frameworks, a behavioral parameter set comprises a sample of parameter sets from the behavioral model space through minimization of the error metric. A number of studies have focused on improving behavioral parameter set analysis by including regional datasets and expert knowledge in addition to parameters and inputs (Kelleher et al., 2017) or evaluating sets that perform poorly with respect to a given error metric in addition to acceptable simulations (Reusser and Zehe, 2011). However, beyond prior studies of model error, there remains a need to understand the coupled effect of uncertainty in hydrogeologic model parameters and observations on the relative performance of decision alternatives (Razavi et al., 2021).

This study aims to evaluate the sensitivity of groundwater model error and decision-relevant management objectives to uncertain parameters and observations, and to determine the effects

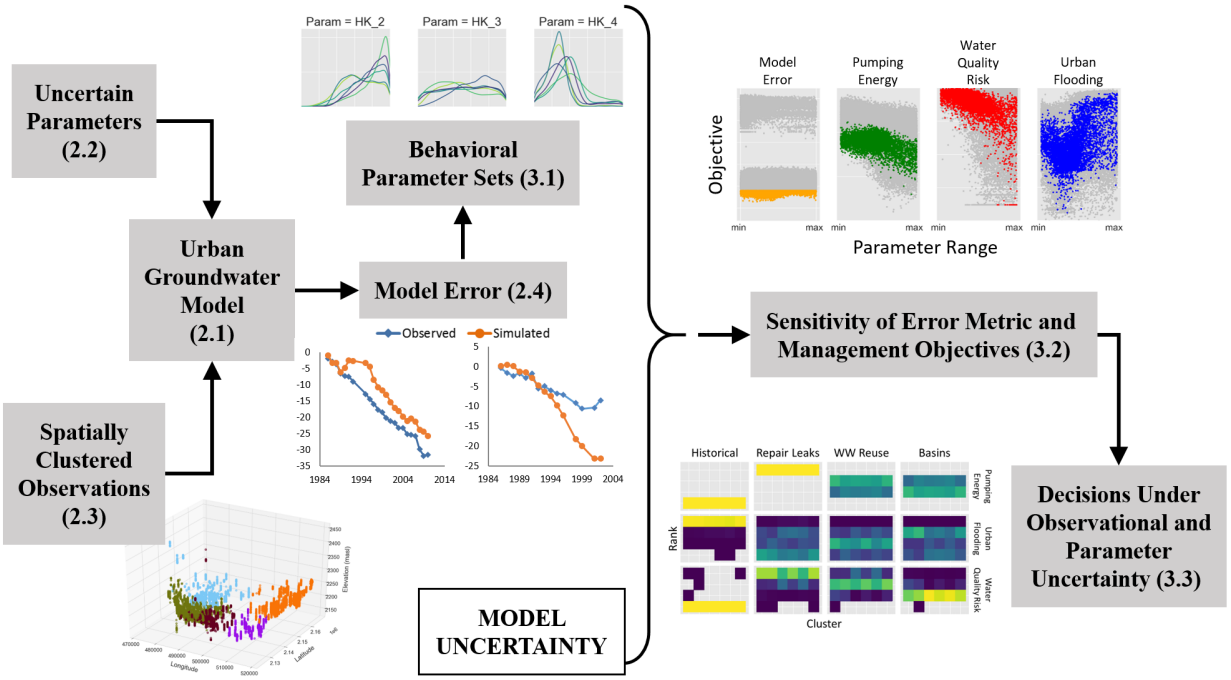

Figure 3.1: Flowchart of methods.

of this coupled uncertainty on the infrastructure planning problem. The result is a planning-driven evaluation of uncertainty to support groundwater management, with the goals of identifying parameters to improve the accuracy of the hydrogeologic model as well as those that should be better constrained to support the selection of management alternatives. This is done through a combination of a global sensitivity analysis and a performance ranking under a range of human-natural parameters, and with the identification of behavioral parameter sets based on multiple possible subsets of historical observations (Figure 3.1). These diagnostic methods aim to evaluate two main consequences of these decision-relevant uncertainties: first, the importance of observation choice and parameter values on the absolute objective performance when controlling for model error, and second, how the relative performance of management alternatives varies when exposed to endogenous uncertainties, individually and in combination. This approach exemplifies how the propagation of multiple endogenous uncertainties throughout the modeling process can ultimately affect the outcomes of regional groundwater supply planning.

3.3 Methodology

This study focuses on the Mexico City Metropolitan Area to evaluate the effects of parameter and observation uncertainty on multi-objective groundwater modeling and decision-making. The Mexico City Metropolitan Area lies within the southwestern portion of the Valley of Mexico watershed, characterized by volcanic peaks surrounding a high plains basin (OCAVM, 2014). This paper uses a case study of the urban aquifer management problem in the Valley of Mexico using a spatially distributed groundwater model adapted from prior work (Herrera-Zamarrón et al., 2005; Lopez-Alvis, 2014; Galán-Breth, 2018; Mautner et al., 2020). This type of complex three-dimensional model is required to approximate the interactions between physical hydrogeologic properties and managed aquifer recharge interventions. This model complexity makes uncertainty analysis difficult, but also critical to understand how spatially and temporally aggregated management objectives vary across many parameter combinations.

3.3.1 Urban Groundwater Model

The Valley of Mexico model is written in Python using the *flopy* package to preprocess data and run the model in MODFLOW, a widely used software which solves the groundwater flow equation (Bakker et al., 2016), as presented in Mautner et al. (2020). The following is a brief overview of the Valley of Mexico test case. A set of model parameters govern model representation of geologic setting, land use and land cover, and water resource infrastructure in Mexico City, including artificial and natural recharge, time varied groundwater pumping, and heterogeneous subsurface characteristics (Table 3.1). This model covers an area of 84 km by 67 km on a 500x500 m spatial grid, and the time period from 1984 to 2013. All model inflows and outflows are applied at a daily time-step, varied according to a monthly stress-period, meaning that data is provided at the monthly time-scale, although data availability may cause some fluxes to vary at the annual or decadal time-scale. Four management alternatives designed to increase groundwater recharge within the basin while avoiding flooding are drawn from Mautner et al. (2020). The alternatives were chosen based

Table 3.1 Model Parameters and Sampling Ranges

Param	Units	Lower Bound	Upper Bound	Param	Units	Lower Bound	Upper Bound
Zonal Geologic				**Time Varied Infrastructure**			
HK_1	m/d	8.64E-7	5.00E-2	Q_{1990}	-	0.3	2.25
HK_2	m/d	1.00E-1	1.00E+2	Q_{2000}	-	0.45	3.5
HK_3	m/d	3.46E-2	1.50E+2	Q_{2010}	-	0.5	4
HK_4	m/d	4.32E-2	4.32E+1	LK_{1990}	-	0.5	2
HK_5	m/d	4.32E-4	8.64E+1	LK_{2000}	-	0.5	2
$S_{s,1}$	1/m	9.19E-4	2.03E-2	LK_{2010}	-	0.5	2
$S_{s,2}$	1/m	4.92E-5	1.05E-3	TWU_{1990}	-	0.75	2
$S_{s,3}$	1/m	1.00E-7	6.89E-5	TWU_{2000}	-	0.95	2
$S_{s,4}$	1/m	1.00E-7	1.02E-4	TWU_{2010}	-	1.1	2
$S_{s,5}$	1/m	1.00E-7	6.89E-5				
$S_{y,1}$	-	0.001	0.08	**Zonal Recharge**			
$S_{y,2}$	-	0.05	0.4	RCH_{urban}	%	0	10
$S_{y,3}$	-	0.01	0.2	$RCH_{natural}$	%	1	80
$S_{y,4}$	-	0.05	0.4	RCH_{water}	%	10	50
$S_{y,5}$	-	0.001	0.1				
$VANI_1$	-	1	1000	**Leak Infiltration**			
$VANI_2$	-	1	1000	IN	%	5	50
$VANI_3$	-	0.1	100				
$VANI_4$	-	0.1	100				
$VANI_5$	-	0.1	100				

[1]Lacustrine, [2]Alluvial, [3]Fractured basalt, [4]Volcaniclastic, [5]Andesitic

on conversations with local practitioners and previous modeling efforts. The alternatives are: the implementation of spatially distributed infiltration basins, demand management through repair of leaks in the water supply network, injection of treated wastewater at existing wastewater treatment plants, and the status quo historical alternative.

Each alternative is then evaluated according to three aquifer management objectives: pumping energy use, water quality risk, and urban flood risk (Equations 1-3). The management objectives evaluated are drawn from Mautner et al. (2020), modified to avoid outlier values that would occur when parameter combinations led to high quantities of model error that would affect the sensitivity analysis. The pumping energy objective (Y_E) is governed by the energy required to pump a daily quantity of groundwater (p) from the water table (h) to the ground surface (s) across all time periods (t) of varying length in days (d) and across all pumping wells (w), converted to kilowatt hours using

an efficiency and conversion term (ϵ). Y_E is calculated starting in the 3rd year of the model period to avoid spin-up effects. In the Valley of Mexico, the lacustrine aquitard in the center of the valley serves as a barrier to contamination of the underlying productive alluvial aquifer; ensuring that the hydraulic head remains above the confining layer reduces water quality impacts in the long-term. The water quality risk objective (Y_W) indicates the number of cells not meeting the groundwater levels below the confining lacustrine layer necessary to maintain water quality (l) divided by the total number of lacustrine cells in the model (L) during the time periods (t) in the last year of the model period. In conflict with the previous two objectives, certain parts of the city lie in areas that are affected by seasonal flooding resulting from medium-term groundwater mounding, which is particularly damaging in urban areas. To take into account these possible negative effects from increasing groundwater head within the valley, the urban flood risk (Y_F) is the sum of the urban area in cells with groundwater mounding (a) divided by the total urban area in the model (A) during the time periods (t) in the last year of the model period.

$$Y_E = \sum_{t=25}^{360} \sum_{w=1}^{n_{wells}} \epsilon p d \left(s_{t,w} - h_{t,w} \right) \tag{3.1}$$

$$Y_W = \frac{\sum_{t=348}^{360} l_t}{\sum_{t=348}^{360} L} \tag{3.2}$$

$$Y_F = \frac{\sum_{t=348}^{360} a_t}{\sum_{t=348}^{360} A} \tag{3.3}$$

3.3.2 Uncertain Parameters

The 33 model parameters include zonal geologic, time varied infrastructure, zonal recharge, and infiltration characteristics (Table 3.1). There are four zonal geologic parameters for each of the five geologic formations; one parameter for each of the three decades during the model period for the total water use, ratio of urban to periurban pumping, and distribution system leak multiplier; a recharge percentage for each land use type; and an infiltration parameter for leaked water. Parameter

ranges are adapted from Mautner et al. (2020), adding or adjusting maxima and minima where necessary based on literature and physical relationships. The calibration carried out in Mautner et al. (2020) used a local sensitivity analysis in which some parameters were not assigned sampling ranges. In this study, a global sensitivity analysis is used and thus some combinations of parameter values had to be avoided based on the structure of the model (Table 3.1). For example, estimated pumping in the region is determined by subtracting historical non-pumping water source quantities from the total regional water use derived from the total water use multiplier [TWU], thus, the TWU must result in a regional water use greater than the historical non-pumping water sources. Similarly, the urban pumping multiplier [Q] acts on a historical dataset, and must result in total pumping less than the estimated pumping determined by the combination of regional water use, historical other supply sources, and the TWU multiplier. The selected parameter ranges are shown in Table 3.1. Using these ranges, 100,000 unique parameter sets are generated using Latin Hypercube sampling. Simulations for each of the management alternatives using the parameter sets were carried out on 296 processors over a total of 107,814 CPU hours. A single model run is on the order of 5 minutes, depending on the combination of parameters and the processor speed.

3.3.3 Spatially Clustered Observations

Uncertainties introduced throughout the groundwater modeling process propagate through to decision-making based on the simulated performance of management alternatives. Parameter uncertainty is reduced by calibration against observations. However, there is also error in, and uneven representation of the model area by, the observations used for calibration. Ideally, observation data can be filtered according to knowledge of collection methods, characteristics of monitoring wells, and distribution across the model area. However, increasingly, modelers face unwieldy and incomplete observation data sets that have greater degrees of freedom and limited or uncertain boundary conditions with which to calibrate models (Tiedeman et al., 2004; Tonkin et al., 2007; Hrachowitz et al., 2013; Nearing et al., 2021). These uncertainties can include lack of data on geologic formation boundaries, placement and magnitude of cones of influence from pumping

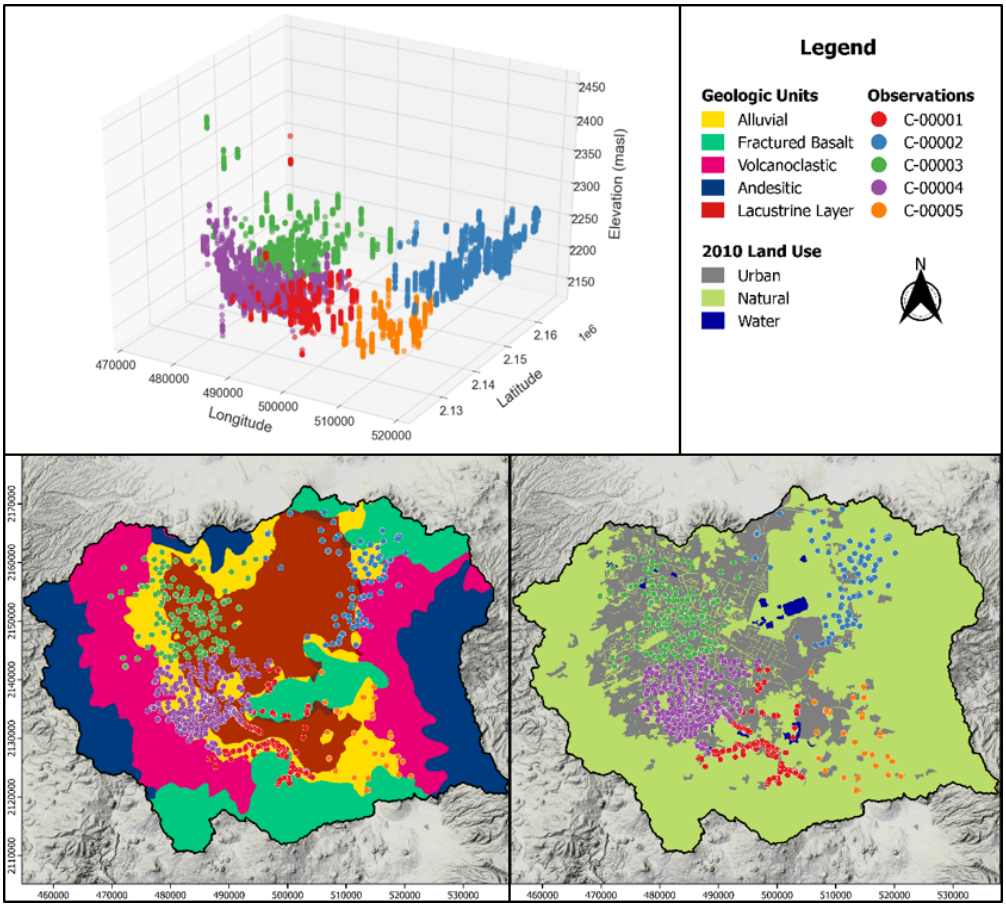

Figure 3.2: (Top) A 3-dimensional visualization of the 5 clusters of observations used in this study. (Bottom left) Observation clusters shown with the geologic formations within the model area. (Bottom right) Observation clusters shown with the land use types for the model period covering 2010.

wells, and the effects of urban karst and land cover types on natural and artificial recharge near monitoring wells.

A set of 8,181 observations from 676 monitoring wells is available for the area and time period of the urban groundwater model used in this study. Well observations vary from 1 to 29 data points per well over the 30 year model period, with a maximum of one observation per year. Multiple interacting uncertainties (e.g. land use, pumping wells, geologic formations) can have unpredictable effects on the relevance of certain observations, thus observation uncertainty is represented in this study by separating the full set of observations into randomly selected and spatially distinct subsets of observations to be act as proxies for incomplete historical records (Figure 3.2). The full set was separated into five clusters using centroid initialization of K-means clustering normalized within

the 3-dimensional space of the set (Figure 3.2), resulting in a total of six clusters when the full set is included as a control.

3.3.4 Model Error

Doherty and Moore (2020) propose the selection of a decision-critical prediction when assimilating observed data into a model for calibration. In this model, the three groundwater planning objectives are based on various spatial and temporal aggregations of groundwater head values, thus, an error metric that assesses model agreement with piezometric head, the decision-critical prediction, through space and time was selected. As in Mautner et al. (2020), model error is captured by the sum of squared weighted residuals (SSWR) between historical head observations ($h_{obs,i}$) and simulated values ($h_{sim,i}$), using weights (ω) determined in Lopez-Alvis (2014) and Galán-Breth (2018):

$$SSWR = \sum_{i=1}^{n} \frac{1}{\omega^2}(h_{obs,i} - h_{sim,i})^2 \tag{3.1}$$

The model error is calculated under the status quo scenario to characterize model agreement with historical hydraulic head observations. Given the inclusion of multiple observations for a single well over time, the error metric captures both spatial and temporal variability in hydraulic head. Higher values of this metric indicate poor model agreement with observations, with larger disagreements amplified in the metric as a result of the squared residual.

3.3.5 Parameter Set Selection

In complex systems with uncertain inputs, model processes can be difficult to parameterize and even more difficult to constrain. While perfect monitoring and representation is the ideal, in reality, simplifying assumptions must be calibrated to create models that can better inform policy and management. In such cases, it is common to have multiple viable parameter sets that produce simulations with acceptable or equivalent model error. Changes in the observations used to evaluate

error can lead to differences in the behavioral parameter sets that are chosen as the best performing simulations. Here, we calculate the error metric for each of the six observation clusters, inclusive of the full set of observations, and choose the 5% best performing parameter sets according to that metric. This gives a sample of 5,000 parameter sets that perform relatively well with respect to the full sample of 100,000. We refer to these as the cluster behavioral parameter sets for each of the six observation clusters.

3.3.6 Sensitivity Analysis

To better ensure robust management alternatives under uncertain model inputs, global sensitivity analysis has been increasingly explored as a decision support tool (Razavi et al., 2021). The sensitivity of the management objectives across the parameter space with respect to both management alternatives and cluster behavioral parameter sets indicates the variability of uncertainty with respect to individual physical model parameters. Using the cluster behavioral parameter sets, a global sensitivity analysis is performed using the Delta Moment-Independent Measure (δ) Borgonovo (2007); Plischke et al. (2013),

$$\delta_i = \frac{1}{2} E_{X_i}[\int |d\mu_Y - d\mu_{Y|X_i}|] \tag{3.1}$$

where the moment independent sensitivity indicator of parameter X_i with respect to the output Y (δ_i) represents the normalized expected shift in the distribution of Y, as a function of μ_Y and $\mu_{Y|X_i}$, the unconditional and conditional measures of Y, respectively. In this study, the parameters (X_i) are the 33 parameters shown in Table 3.1 and the output (Y) are the three management objectives described in Equations 1-3. This method was selected for two reasons. First, δ provides a better representation of sensitivity with respect to model structure when parameters are correlated, often true in complex human-natural systems (e.g. increased groundwater pumping during periods of reduced recharge and surface supplies from drought), when compared to variance-based methods (Borgonovo and Plischke, 2016). Second, the Delta method does not require a specific structure of

parameter samples, allowing for the sub-selection of 5,000 samples from the initial set. By only evaluating objective sensitivity across the solution space of the cluster behavioral parameter sets rather than the entire solution space, we remove objective values of simulations that do not agree with observations, which have the potential to introduce further uncertainties. Parameter sensitivity is calculated for four model outputs: the error metric, and three management objectives. A total of 72 sensitivity analyses on 33 parameters are performed across combinations of: 4 alternatives, 3 objectives, and 6 cluster behavioral parameter sets, resulting from filtering based on the error metric among each of the 6 observation clusters. The sensitivity analyses were performed on 72 processors over a total of 4.9 CPU hours.

3.3.7 Evaluation of Decision Uncertainty

To understand the extent to which uncertainty in observations and parameters can affect decision-making analyses, we compare alternative performance across cluster behavioral parameter sets. First, management alternatives are ranked within each objective for each of the parameter sets to view differences in the alternative ranking across the cluster behavioral parameter sets. We evaluate the model results to visualize changes in ranking according to three types of comparisons using sets of heatmaps that summarize the ranking of the alternatives across all the parameter sets. The comparisons evaluated are: (1) all three objectives across the observation cluster behavioral sets; (2) all three objectives for observation cluster behavioral set C-12345 across the range of the alluvial hydraulic conductivity parameter [HK2]; and (3) the water quality objective (Y_W) across the observation cluster behavioral sets and parameter HK2, simultaneously.

In all three comparisons, the first step is to rank the alternatives according to the objective(s) from lowest (1) to highest (4) in each parameter set. Then, the ranking data for all the parameter sets in each comparison are summarized as follows:

Three objectives across observation cluster behavioral sets - The count of rankings for each alternative. Each column (cluster) in each objective row will sum to 5,000.

Three objectives for C-12345 across parameter HK2 - The 5,000 sample set is separated into ten bins along the parameter value range from Table 3.1. The ranking count in each bin is divided by the total number of parameter samples in each bin to allow direct comparison across all bins. This is necessary because the distribution of behavioral parameters can be non-uniform. Each column (parameter value bin) in each objective row will sum to 100%, or null if there are no parameter sets in that bin.

Y_W **across the observation cluster behavioral sets and HK2, simultaneously -** Same as in the previous comparison, but for only the water quality objective. This is repeated for the remaining observation cluster behavioral sets (C-00001 to C-00005). Each column (parameter value bin) in each cluster row will sum to 100%, or null if there are no parameter sets in that bin.

Second, the difficulty of the decision was measured by evaluating the percent difference between the first and second ranked alternatives, and between the first and worst ranked alternatives. The distribution of these differences indicate the relative performance between the alternatives, with a distribution concentrated among lower values indicating a more difficult decision because the relative differences between the objective measures of the options are smaller. While alternative ranking can provide some information on the relative performance of aquifer management alternatives with respect to each other, it does not provide information on the difference between the performance in each simulation. More importantly, by not knowing the range of objective values between the management alternatives in a given simulation, decision-makers might incorrectly infer the difficulty of a decision. For example, take the case of two simulations where the performance in the urban flood risk objective of the historical, infiltration basin, wastewater reuse, and repair leaks alternatives are 1.5%, 1.7%, 1.8%, and 2%, respectively, in the first simulation, and 2%, 15%, 32%, and 40% in the second simulation. These two simulations may produce the same alternative ranking: historical (1), infiltration basins (2), wastewater reuse (3), and repair leaks (4). However, it is clear that second simulation produces a much "easier" decision than the first because the absolute and relative differences between the objective values are larger in the second simulation.

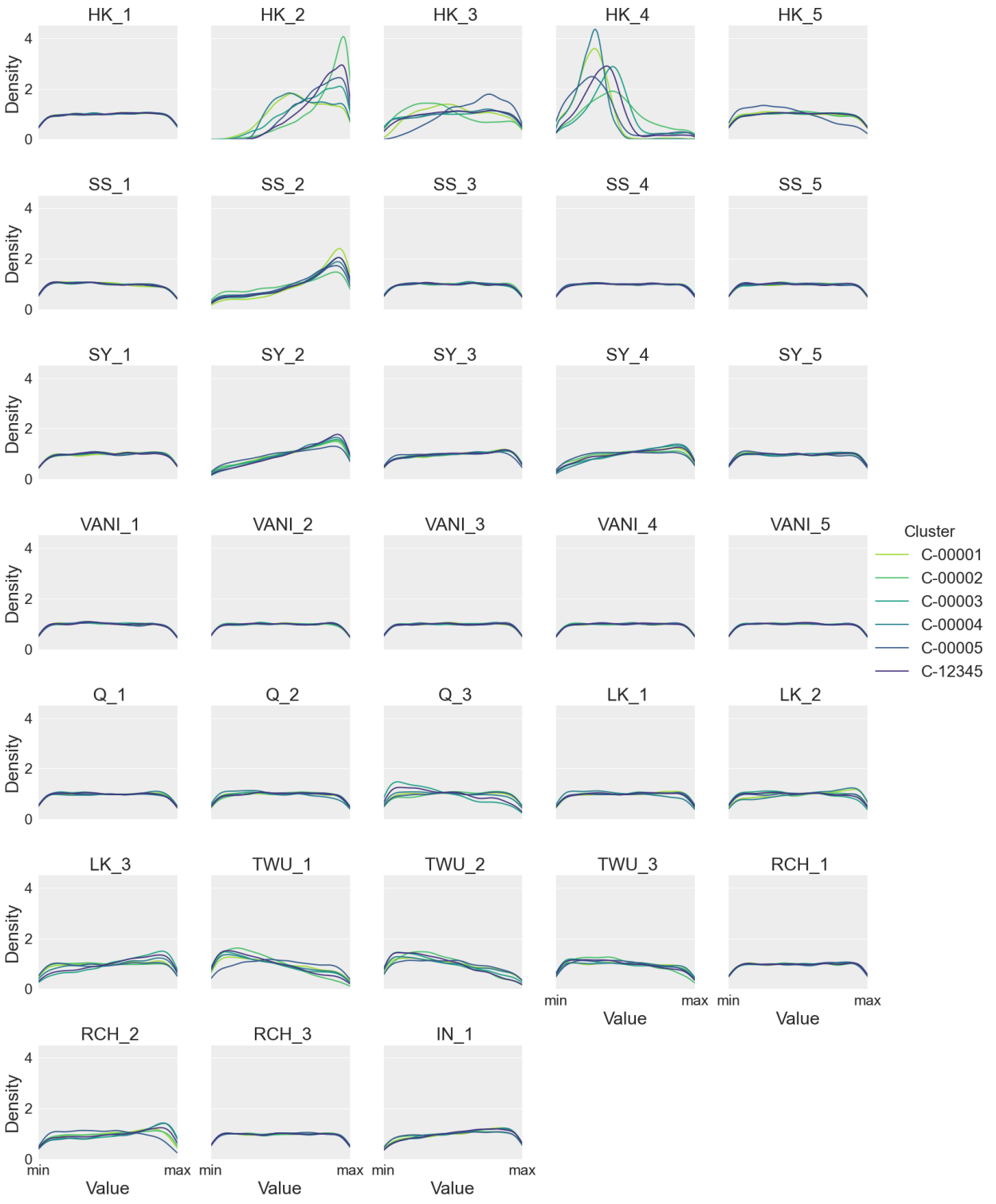

Figure 3.3: The distributions along the parameter ranges of the filtered samples using the sum of squared error metric. The distributions are colored according to the observation cluster used to filter the dataset. The prior distribution (not shown) is uniform for all parameters. Parameter abbreviations given in Table 3.1.

3.4 Results and Discussion

3.4.1 Cluster Behavioral Parameter Sets

Figure 3.3 shows the kernel density estimations (KDEs) for the resulting parameter distributions when selecting the 5,000 samples with the lowest error using each of the observation clusters (C-00001, C-00002, C-00003, C-00004, C-00005) and the entire observation set (C-12345). The initial distribution (not shown) is uniform for all parameters. These distributions indicate the parameters that have the greatest influence on model error, defined here as those with the greatest deviation in distribution from the prior uniform distribution, namely the horizontal hydraulic conductivity [HK] parameters. The higher parameter values for the geologic characteristics (horizontal hydraulic conductivity [HK], specific storage [SS], and specific yield [SY]) of the alluvial formation (formation number 2) are preferentially represented in the low-error parameter sets. For hydraulic conductivity, this indicates that an alluvial formation [HK2] that allows for more rapid flow of groundwater, and thus greater dispersion of groundwater throughout the model area, results in lower error. When combined with high values for flux parameters such as the total water use [TWU] (governing groundwater pumping) and recharge [RCH], this could signal that these models avoid extreme mounding or drawdown that would increase model error. Similarly, the selection of a larger number of high values for specific storage [SS2] and specific yield [SY2] in the alluvial formation confirms that the selected parameters would tend to mitigate the effects of higher flux values. Alternatively, all distributions of the horizontal hydraulic conductivity show a concentration of lower values for the volcaniclastic formation [HK4]. This indicates parameter values that encourage higher groundwater retention in the mountainous, volcaniclastic areas, which could be a result of observations in perched or mountainous regions having an outsized effect on the error metric.

In terms of flux parameters, the total water use [TWU], the recharge percentage of the natural land use type [RCH2], and the leak [LK3] and pumping [Q3] multipliers for the third decade of the model period all show a small redistribution toward the extremes of the parameter ranges.

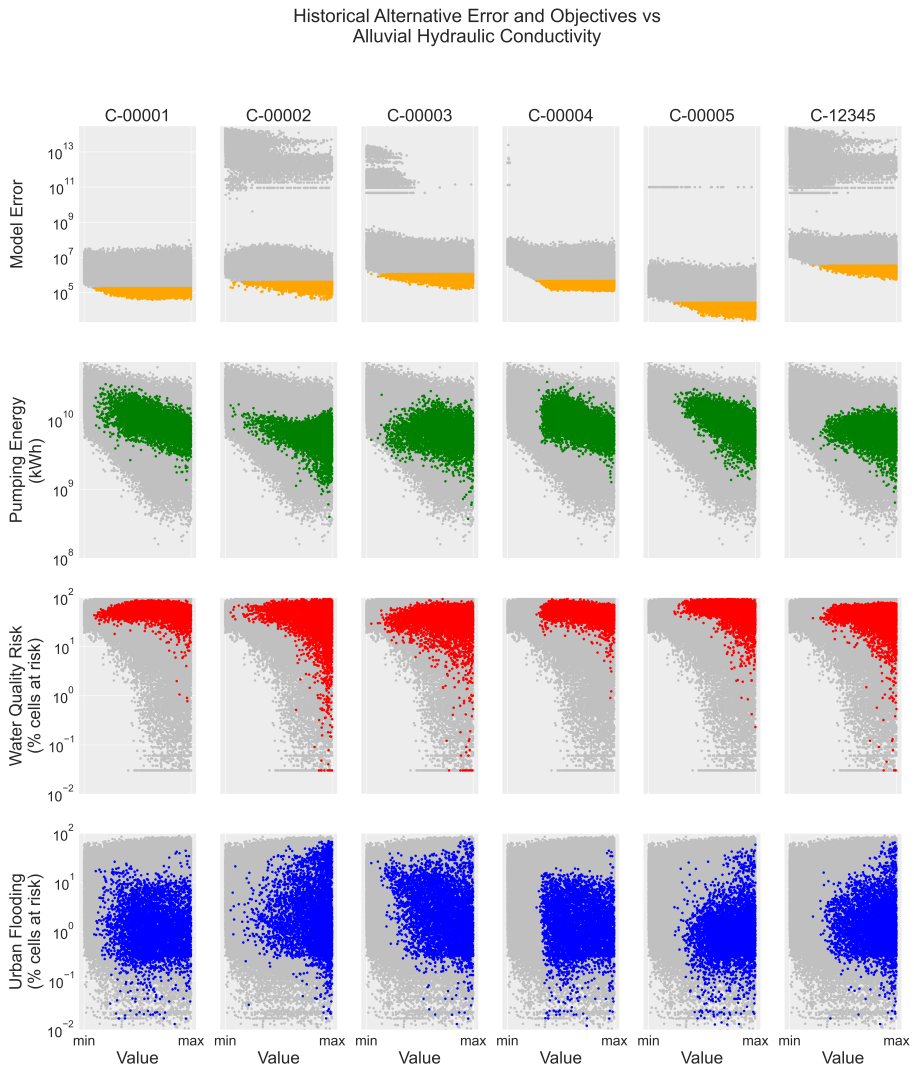

Figure 3.4: A representative view of the four model output metrics for the historical alternative, plotted against the parameter range for the hydraulic conductivity of the alluvial formation (the most sensitive parameter from Figure 3). These include the error metric (sum of squared weighted residuals in m^2), energy objective (kWh), water quality risk objective (percent of cells not meeting the objective), and urban flooding objective (percent of cells not meeting the objective). Gray points represent all parameter sets, while colors represent behavioral parameter sets meeting the error threshold.

The preference for lower values of total water use, particularly in the first decade [TWU1], could confirm that mitigated drawdown in the model leads to lower error. At the same time, the slight tendency toward increased recharge in the natural land use type agrees with the tendency toward low hydraulic conductivity in the volcaniclastic formation that, combined, would indicate a preference

for groundwater mounding along the model edges. Finally, the higher values for the leak parameter in the last decade of the model period [LK3] further confirms the preference for increased hydraulic head in the urban areas.

In isolation, these findings reveal information about the model representation and how to improve parameterizations to minimize error given the existing observations. However, there are visible differences between the distributions of the parameter values from the various cluster behavioral parameter sets. This is particularly evident in the hydraulic conductivities of the alluvial [HK2], fractured basalt [HK3], and volcaniclastic formations [HK4]. Behavioral parameter sets tend to focus on sub-ranges of the horizontal hydraulic conductivity depending on the subset of observations used to calculate the error metric, highlighting the importance of observational uncertainty on parameter identification.

Error reduction through parameter selection is an important consideration for model use. However, we are also interested in how management objectives produced by the model respond to uncertainty in model parameters. Figure 3.4 shows the error metric and the three management objectives for all parameter sets in gray and the behavioral parameter sets in color. Here we visualize how the choice of observation cluster affects the sample of parameter sets and subsequently the range of performance among the pumping energy (green), water quality risk (red), and urban flooding objectives (blue). This example yields noticeable differences between the observation cluster choices, while other parameters (Figure S1) result in fairly uniform sampling across the parameter ranges, following Figure 3.3. The three objectives are to be minimized, thus, in certain objectives, higher alluvial hydraulic conductivity [HK2] results in better performance, particularly for the energy and water quality objectives, while in the flooding objective the performance is more variable across the parameter range. This performance is not consistent across clusters for the alluvial hydraulic conductivity, indicating the impact of observational uncertainty on the performance evaluation of the system.

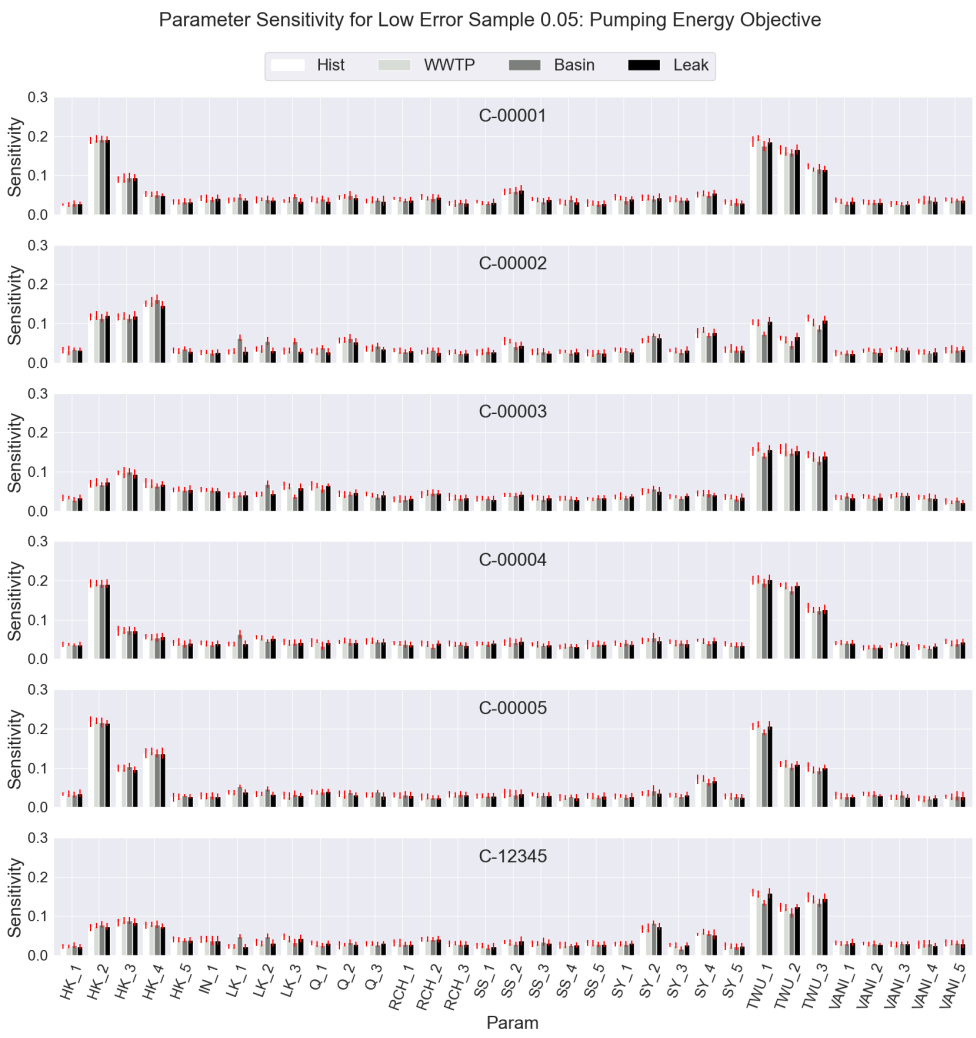

Figure 3.5: δ sensitivity of the energy objective according to the 5,000 filtered samples for the 33 model parameters (columns). The sensitivity is shown by cluster (rows) and by the four alternatives from left to right (light to dark): historical, wastewater reuse, infiltration basins, and repair leaks. The bootstrapped 95% confidence interval for each sensitivity value is shown as a red line.

3.4.2 Sensitivity of Error Metric and Management Objectives

To better understand the effects of parameter values on management objectives, the Moment-Independent Sensitivity Measures, δ, are shown for the energy objective in Figure 3.5 (Figure S2 for water quality risk and Figure S3 for urban flooding risk). The value of δ can range from 0, indicating that the output is independent of the parameter in question, to 1. There is not a standard value for δ that is considered to be highly sensitive because parameter sensitivities should be evaluated in

relation to each other and in the context of each case study. Based on the sensitivity values for this system, we consider a δ of roughly 0.2 and above to be highly sensitive. As in Figure 3.4, the patterns of objective sensitivity to the parameters vary across the samples chosen using different observation clusters. However, in Figure 3.5 we can also compare the sensitivity of the objectives across management alternatives. With few exceptions, the sensitivities of the objectives across the alternatives within each cluster sample are fairly consistent. This suggests that the performance of the system with respect to the management objectives is minimally affected by the choice of alternative. This has two main implications. First, this could signify that the relative performance of the alternatives is similar across a range of parameter values and indicate that the decisions made are robust across many parameter combinations. Second, if decision makers are using a sensitivity analysis to choose parameters for further study, they can be relatively confident that the choice of parameters to monitor will not favor a given alternative. The most notable exception is the sensitivity of the pumping energy objective with respect to the leak multiplier (LK1, and to a lesser extent LK2 and LK3) for the repair leaks alternative. This is expected given the reliance of the leak repair alternative on the quantity of leaks present–essentially, the more leaks available to be repaired indicates a larger water savings and thus a higher water table from which to pump to the ground surface.

It is also valuable to understand how the sensitivities of the three management objectives compare to those for the error metric. Many numerical groundwater models are constructed with a specific management purpose, but the model itself is calibrated to error metrics that represent available data, and these may not necessarily rely on the same mechanisms driving the performance of management alternatives. Figure 3.6 shows the δ values for the parameters with the largest differences in sensitivity between clusters. The sensitivities of the error metric across the filtered sample are relatively small because they include only the parameter sets with the lowest error. While the sensitivities of the error metric to the parameters are smaller overall than those of the objective values, the effects seen on the distributions in Figure 3.3 are mirrored to some extent here, with slight increases in the sensitivity of the error metric to the horizontal hydraulic conductivity

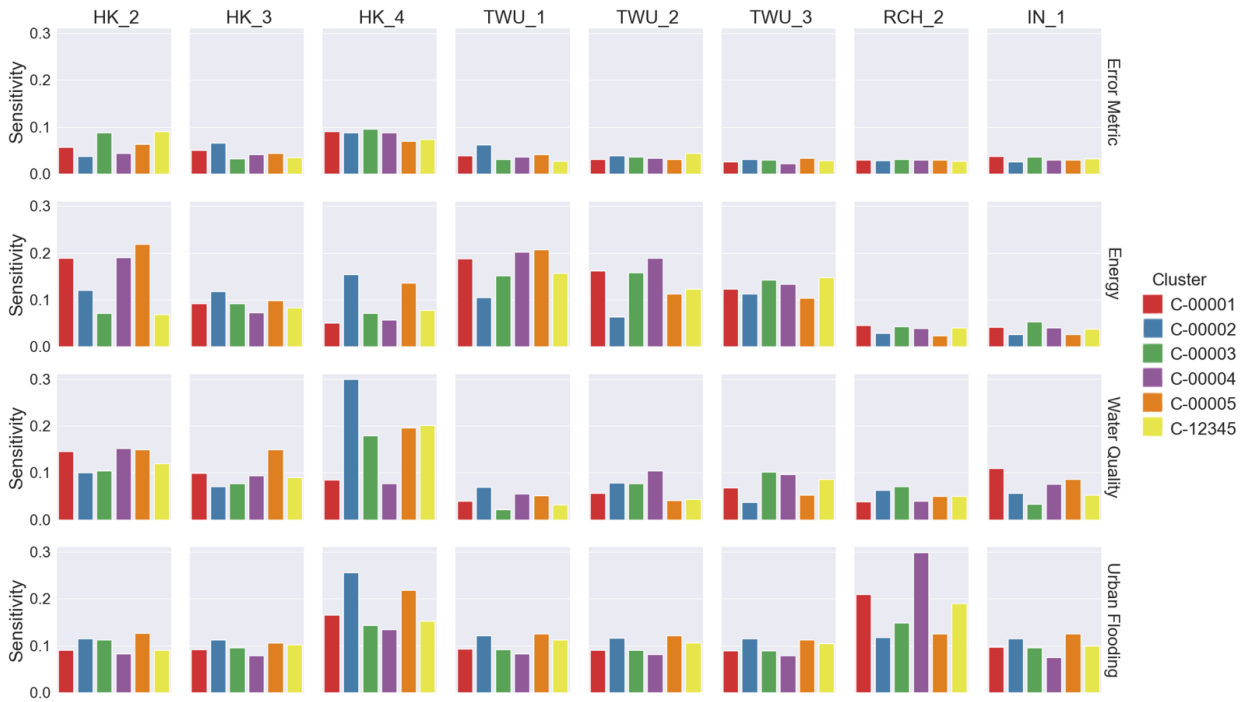

Figure 3.6: δ sensitivity of the error metric and three management objectives (rows) according to the 5,000 filtered samples for the 8 model parameters (columns) with the largest differences in sensitivity between clusters for the historical management alternative. The sensitivity is shown by cluster in order from left to right: C-00001, C-00002, C-00003, C-00004, C-00005, C-12345.

of the alluvial [HK2] and volcaniclastic [HK4] formations.

However, the patterns of the sensitivity of the error metric generally do not align with the patterns seen in the management objectives. Objectives are more or less sensitive to specific parameters depending on the cluster behavioral parameter sets. For example, the sensitivity of all three management objectives to the volcaniclastic hydraulic conductivity [HK4] is largest for the C-00002 samples, and is most pronounced for the water quality risk objective. Figure 3.4 shows that the parameter sets selected using the observation cluster C-00002 result in a much broader set of values for the hydraulic conductivity of the volcaniclastic formation than the other objective cluster samples, particularly for the water quality risk and urban flooding indicators. Similarly, samples C-00001 and C-00004 result in much higher sensitivities of the urban flooding objective to the recharge parameter of the natural land cover [RCH2].

Higher sensitivities for certain cluster behavioral parameter sets may indicate that the chosen

observations do not properly constrain the model with respect to the given parameter, resulting in a number of non-unique solutions. Alternatively, higher sensitivities may occur when the spatial extent of the parameter and the management objective calculation are conincident, as in the case of the total water use parameters [TWU], which act upon the pumping wells, and the energy objective, which is calculated at the location of the pumping wells. Finally, the sensitivities are also affected by the physical processes governed by a given parameter, as in the case of the high sensitivity of the urban flooding objective to the recharge percentage parameter [RCH]. Understanding which parameters contribute most to objective uncertainty indicates opportunities for data collection to improve model representation of those processes. The δ values show that uncertainties in the observations used in calibration can result in appreciable changes in the distribution of the performance in management objectives. These findings underline the importance of high quality, well distributed, and diverse observation data for calibration. Additionally, decision-making often depends on the behavior of spatially and temporally aggregated indicators or objectives whose sensitivity to model parameters may or may not be aligned with the sensitivity of the error metric to those same parameters.

3.4.3 Decisions Under Observational Uncertainty

Parameter sensitivities provide information about improvements that can be made in the modeling and calibration process to reduce error. However, it is also important to understand how these uncertainties propagate into the decision-making process, particularly whether they contribute to changes in potential decisions informed by the simulation model. Figure 3.7 shows the relative performance of the aquifer management alternatives according to the cluster behavioral parameter set and management objective. In the heatmaps, a lighter (yellow) color indicates more parameter sets where that alternative is ranked at that value and a darker (purple) color indicates fewer parameter sets that are ranked at that values. If no parameter sets result in a given rank for that alternative, the space is left gray.

For the pumping energy objective, the historical and repair leaks alternatives rank worst (4)

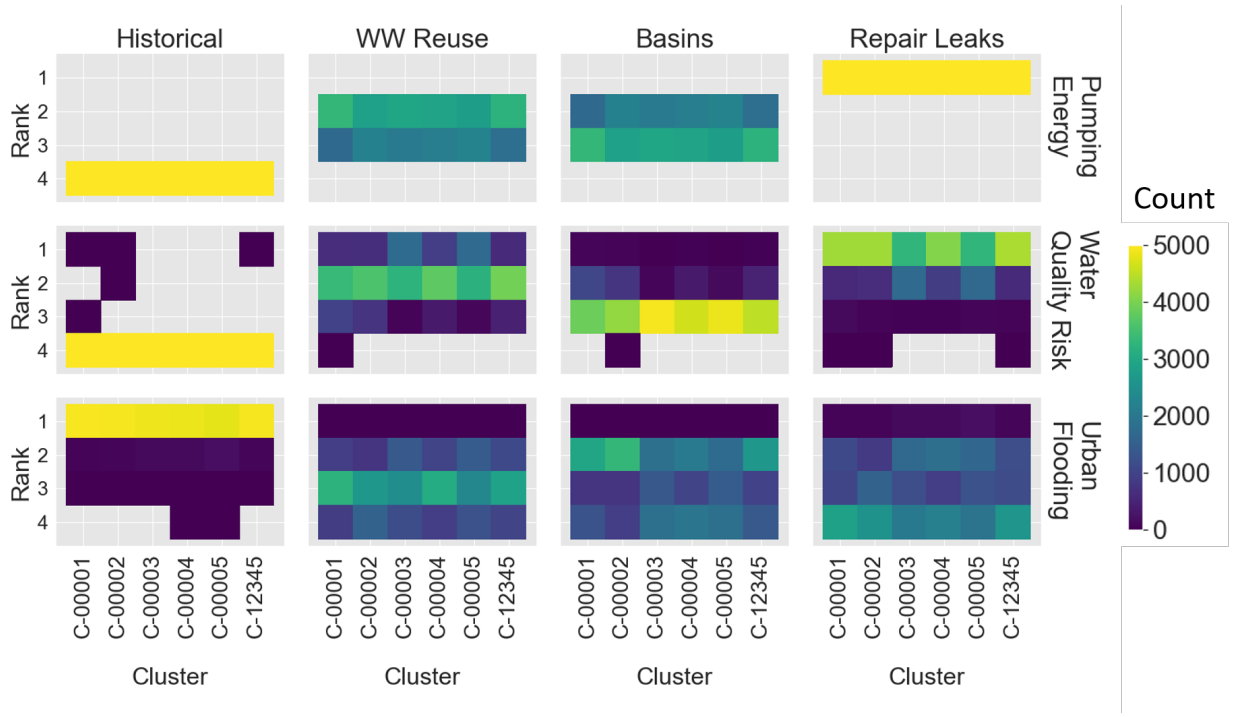

Figure 3.7: Alternative performance across the observation cluster parameter sets shown as heatmaps of the count of sets where the alternative performance was ranked as (1) best to (4) worst. Within each heatmap, the rows are the rank and the columns are the cluster behavioral parameter sets. The subplots are organized by the three management objectives as the rows and the aquifer management alternatives as columns.

and best (1), respectively, across all simulations in all parameter set samples, while the wastewater reuse and infiltration basin alternatives rank 2nd and 3rd almost evenly across the simulations. The wastewater reuse alternative ranks 2nd slightly more often (lighter) in the pumping energy objective than the infiltration basin alternative, particularly in the C-00001 cluster behavioral parameter set and the full observation sample set (C-12345). In the water quality risk objective, the historical alternative ranks 4th across practically all the cluster behavioral parameter sets. Similarly, the infiltration basins alternative ranks 3rd in almost all behavioral parameter sets. The 1st and 2nd ranked alternatives, while less definitive are still fairly clear, with the repair leaks alternative ranking 1st and the wastewater reuse alternative 2nd across most of the cluster behavioral parameter sets. Here, C-00003 and C-00005 have less difference in the number of parameter sets where the repair leaks alternative ranks 1st and the wastewater reuse alternative 2nd when compared to the other cluster behavioral parameter sets (C-00001, C-00002, C-00004, C-12345). Finally, in the urban

Figure 3.8: Alternative performance across the parameter range of the alluvial hydraulic conductivity (one of the most sensitive parameters) shown as heatmaps of the count of sets where the alternative performance was ranked as (1) best to (4) worst. Within each heatmap, the rows are the rank and the columns are the parameter value from minimum (1.00E-1) to maximum (1.00E+2). The subplots are organized by the three management objectives as the rows and the aquifer management alternatives as columns.

flooding objective, the best performing alternative is the historical alternative in the vast majority of the parameter sets across all cluster behavioral parameter sets. This is expected given that the urban flooding objective measures groundwater mounding in the model, and since the remaining three management alternatives all increase recharge in the model, the status quo alternative experiences the least amount of mounding. However, the relative ranking between the other three alternatives is much less clear, particularly in the C-00003 and C-00005 cluster behavioral parameter sets.

Here, it is apparent that the choice of observations by spatial clusters would have a minimal effect on decision-making, making this type of comparison of the alternatives robust across behavioral parameter sets chosen using observations from many different regions within the model area. This reveals two main points: first, the apparent agreement between sensitivities of performance to parameters across the alternatives may indicate relative stability of the performance of alternatives across the cluster behavioral parameter sets, even though parameter sensitivities are not consistent

across those same sets. Second, the comparison of rankings across the observational clusters may not capture the full interplay of absolute performance under observational uncertainties.

Next, in Figure 3.8 we compare the ranking across one of the most sensitive parameters, the hydraulic conductivity of the alluvial formation [HK2], looking only at the behavioral parameter sets chosen using the C-12345 (full) observation cluster. Similar to the comparison across observation clusters, the ranking of the management alternatives across the range of parameter values is stable for the pumping energy objective. The wastewater treatment and infiltration basin alternatives show a roughly even split between the 2nd and 3rd ranking. However, in the other two objectives, the ranking changes depending on the value of the alluvial hydraulic conductivity. This is particularly apparent in the water quality risk objective, although it also occurs to a lesser degree in the urban flooding objective. Notably, the repair leaks alternative ranks 1st in the water quality risk objective except at lower values of the parameter range, where the wastewater reuse objective is preferred. There are many competing factors that could contribute to this outcome. For example, lower hydraulic conductivity in the alluvial aquifer would indicate higher groundwater retention and could thus favor parameter sets with lower urban leak and total water use values, to reduce model error by avoiding local mounding and cones of depression. In those

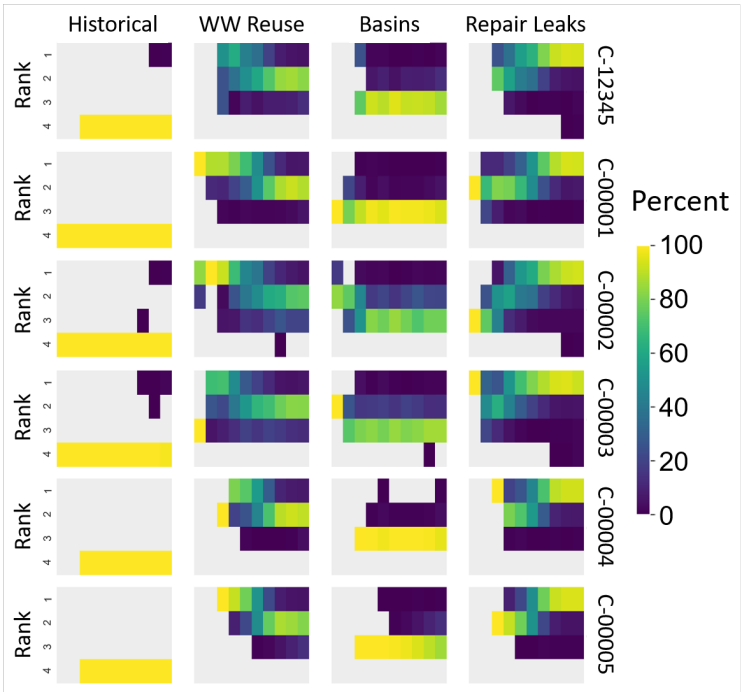

Figure 3.9: Alternative performance in the water quality objective. Shown as heatmaps of the count of parameter sets where the alternative performance was ranked as (1) best to (4) worst. Within each heatmap, the rows are the rank and the columns are the parameter value from minimum (1.00E-1) to maximum (1.00E+2). The subplots are organized by the observation cluster used for behavioral parameter set selection as the rows and the aquifer management alternatives as columns.

cases, the wastewater treatment alternative would increase groundwater recharge more than the repair leak alternative, and thus improve groundwater levels within the clay layer that influences water quality risk in the basin. Additionally, some of the fluctuations in the ranking result from the sample counts in each bin after the behavioral parameter set filtering. For example, the lowest bin in the water quality risk objective of the historical and infiltration basin alternatives shows a large difference because the sample count is low. In this case, a change in the bin size could change the relationship between the parameter values and the alternative ranking.

Finally, Figure 3.9 shows the combined effects of the observation and parameter uncertainty on alternative performance in the water quality risk objective. Here, it is apparent that the observation cluster choice has an effect on the ranking patterns of the management alternatives across the parameter range. While the pattern of favoring the wastewater reuse alternative at the lower alluvial hydraulic conductivity values and the repair leaks at the higher conductivity values is consistent across all the observation cluster behavioral parameter sets, the point along the parameter values at which this occurs changes between the clusters used to evaluate model error. There is even a case, at low alluvial hydraulic conductivity in the C-00002 set, where the wastewater reuse, infiltration basins, and repair leaks alternatives are ranked 1st, 2nd, and 3rd, respectively, in contrast with the findings from Figure 3.7 and, to some extent, Figure 3.8. This makes clear the importance of evaluating the coupled effects of multiple types of endogenous uncertainties on management outcomes in concert, rather than in isolation.

To visualize the effects of the cluster behavioral parameter set on the difficulty of the decision, Figure 3.10 shows the distributions of the percent differences between the 1st and 2nd ranked alternatives in each sample (row 1) and between the best (1st) and worst (4th) ranked alternatives in each sample (row 2) for each cluster behavioral parameter set. In this figure, a distribution that is clustered near the origin of the graph indicates a more difficult decision because the percent difference between the objective values of each of the alternatives is smaller.

In the pumping energy objective, the minimal differences in the distributions confirm the conclusions from Figure 3.7, that the alternative rankings are not affected by which cluster behavioral

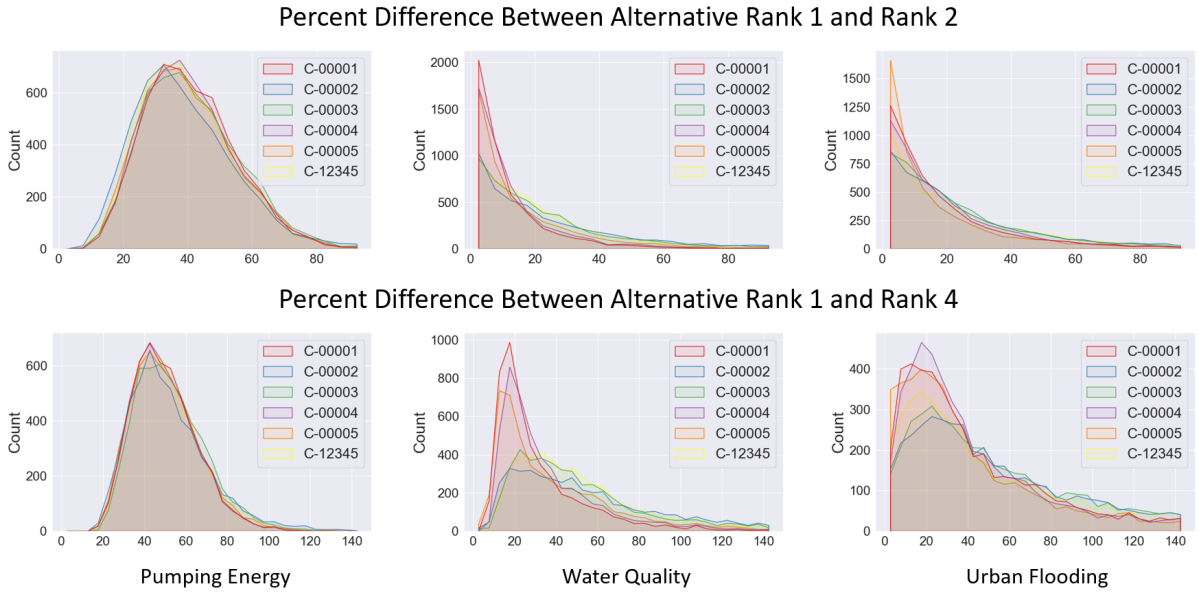

Figure 3.10: The difficulty of the decision represented by the relative performance of the alternatives within the samples evaluated for each objective (columns). The top row shows the distribution of the percent difference in each sample between the 1st and 2nd ranked alternatives within the cluster datasets. The bottom row shows the distribution of the percent difference in each sample between the 1st and 4th ranked alternatives within the cluster datasets.

parameter set was used for calibration. However, in the water quality risk objective, and to a lesser extent in the urban flooding objective, the cluster behavioral parameter set has an effect on the distribution of the percent difference between the 1st and 2nd ranked alternatives as well as the best and worst ranked alternatives. In the water quality risk objective, C-00001, C-00004, and C-00005 show more instances of difficult decisions. These same cluster behavioral parameter sets also showed more difficult decisions in the urban flooding objective. This indicates that the availability of observational data could contribute to changes in the decision-making process when using the urban flooding and water quality risk objectives in this system.

3.4.4 Limitations and Future Work

Uncertainty analyses face limitations from model complexity and the sample size needed to capture multiple interacting forms of uncertainty. This study can be extended in several ways to address the challenge of propagating uncertainties throughout the groundwater infrastructure

modeling and planning process. For example, this study did not consider multiple model structures and their effects on objective sensitivity and alternative ranking. Such changes could include varying representations of model geology and feedbacks after the implementation of management alternatives. Similarly, the use of observation clusters to reveal spatially dependent sensitivities may obscure the role of outlier observations on parameter sensitivity. Future work could identify the individual observations that contribute the most to sensitivity in each objective across the various parameters to understand better the limitations of the available observations, as has been achieved in other local sensitivity analysis approaches (Poeter et al., 2014; Matott, 2005; Tonkin et al., 2007).

As previous studies have applied space-time optimization for groundwater monitoring networks to reduce the variance of water quality estimates, future studies can apply similar techniques combined with the δ sensitivity measure of groundwater management objectives to determine optimal sampling locations. Additionally, uncertainty in the proper weighting of observations could be simulated using Monte Carlo selection of weights. Finally, clusters were chosen spatially in this study to simulate over-representation of certain areas in monitoring; however, future research may compare clusters based on physical properties such as land use and geologic formation, or other factors, such as the time period or the agency collecting the data. Similarly, bootstrapping or random selection instead of clustering could reveal the outsized influence of certain individual observations on parameter calibration and decision-making.

Additionally, while this study investigates parameter sensitivity and the effects of parameter uncertainty on ranking decisions, it does not explicitly quantify the relationship between the two. The results do not show a clear relationship between the magnitude of the sensitivity of the objectives to changes in the parameters. However, relative differences in the sensitivities of the objectives under different management alternatives may play a role in the alternative ranking. This relationship could be further investigated by developing a metric to capture the fluctuations in ranking driven by each parameter, to be compared with the differences in sensitivity of a given parameter under each alternative. Similarly, while this study considers the impact of coupled uncertainties on three different management objectives, future work could implement a multiobjective approach

evaluating Pareto optimality to consider all three objectives simultaneously.

Finally, the implementation of groundwater recharge alternatives could be modified to improve the accuracy of the simulations. One option would be to include costs of the management alternatives as either an additional objective or as a constraint to the implementation. Similarly, combinations of the various management alternatives or varying degrees of implementation may give further multi-objective benefits beyond those of each management alternative implemented individually.

3.5 Conclusions

In this study, we explore how observation and parameter uncertainty propagate through a hydrogeologic model to influence the ranking of decision alternatives. Using global sensitivity analysis and evaluation of aquifer management objectives across behavioral parameter sets filtered from a global sample, we evaluate how physical properties of the model and choice of observations for calibration can lead to variations in decision-relevant model outputs. We find that metrics that are generally used to determine predictive ability, such as the sum of squared weighted residuals, are not necessarily aligned with the decision-making applications for which models are applied. The management objective values in the behavioral parameter samples show a much greater range of sensitivity than those demonstrated by the model error. This underlines the importance of carrying through sensitivity analyses to the decision-making stage of the modeling process, beyond just the parameter calibration stage.

Additionally, results show that observational uncertainty plays a much larger role in the sensitivity of the objectives than the management alternatives themselves. This suggests that the performance of the system with respect to the management objectives is minimally affected by the choice of alternative when compared to the variability produced by endogenous model uncertainties. Under certain conditions, the relative performance of the alternatives under some of the objectives is consistent across many combinations of parameters and observation clusters–particularly for

the pumping energy objective. This confirms that the performance of the demand management represented by the leak repair alternative is robust across many realizations of uncertainty.

The choice of observations shows a minimal effect on decision-making, with almost no differences in alternative ranking between the behavioral parameter sets. In contrast, the ranking of the leak repair and wastewater reuse alternatives showed fluctuations in ranking across the range of one of the most sensitive model parameters, the alluvial hydraulic conductivity. However, when combined with the parameter uncertainty, the observational uncertainty does contribute to greater fluctuations in alternative ranking. This makes clear the importance of evaluating the coupled effects of multiple types of endogenous uncertainties on management outcomes in concert, rather than in isolation.

Finally, the selection of alternatives becomes more or less difficult according to the relative performance of management objectives. Specifically, the distribution of the difficulty metric in each of the objectives changes based on the observation cluster used to select the behavioral parameter sets. These methods could be leveraged to determine which additional observations would help to more easily identify the best performing alternative under multiple management objectives. This study highlights the importance of understanding how the uncertain parameters of a physical model and their interactions with the observations used to calibrate them can affect water supply planning decisions in densely populated urban areas.

3.6 Appendix

Figure 3.11 is an additional figure similar to Figure 3.4 in the main text, but with the x-axis representing the vertical anisotropy of the hydraulic conductivity of the alluvial formation, a parameter that maintained uniform sampling across the parameter range.

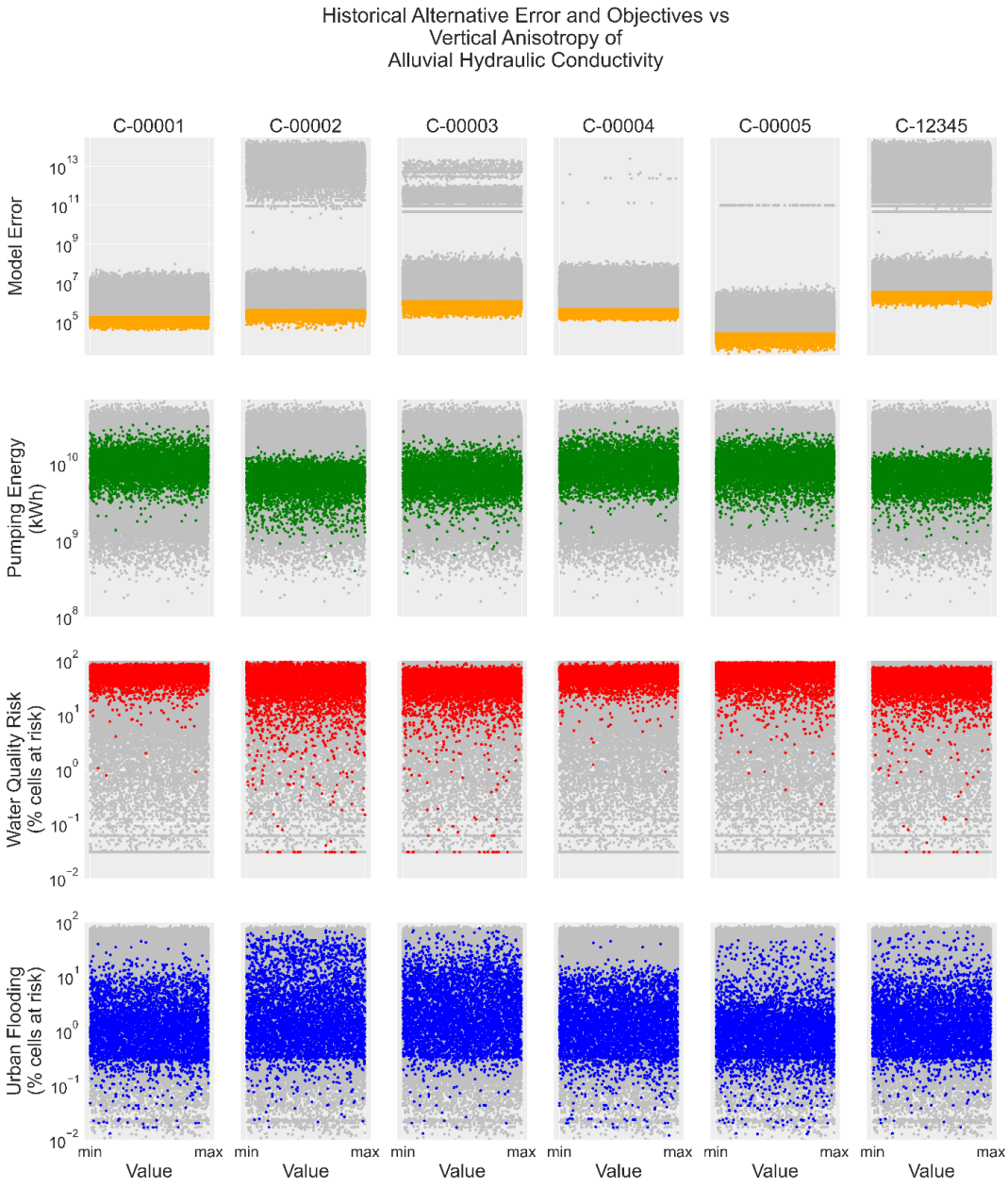

Figure 3.11: A representative view of the four model output metrics for the historical alternative, plotted against the parameter range for the vertical anisotropy of the hydraulic conductivity of the alluvial formation. These include the error metric (sum of squared weighted residuals), energy objective (kWh), water quality risk objective (percent of cells not meeting the objective), and urban flooding objective (percent of cells not meeting the objective). Gray points represent all parameter sets, while colors represent behavioral parameter sets meeting the error threshold.

Figures 3.12 and 3.13 are additional figures similar to Figure 3.5 in the main text, but with

the δ sensitivity values for the water quality risk objective and the urban flooding risk objective, respectively.

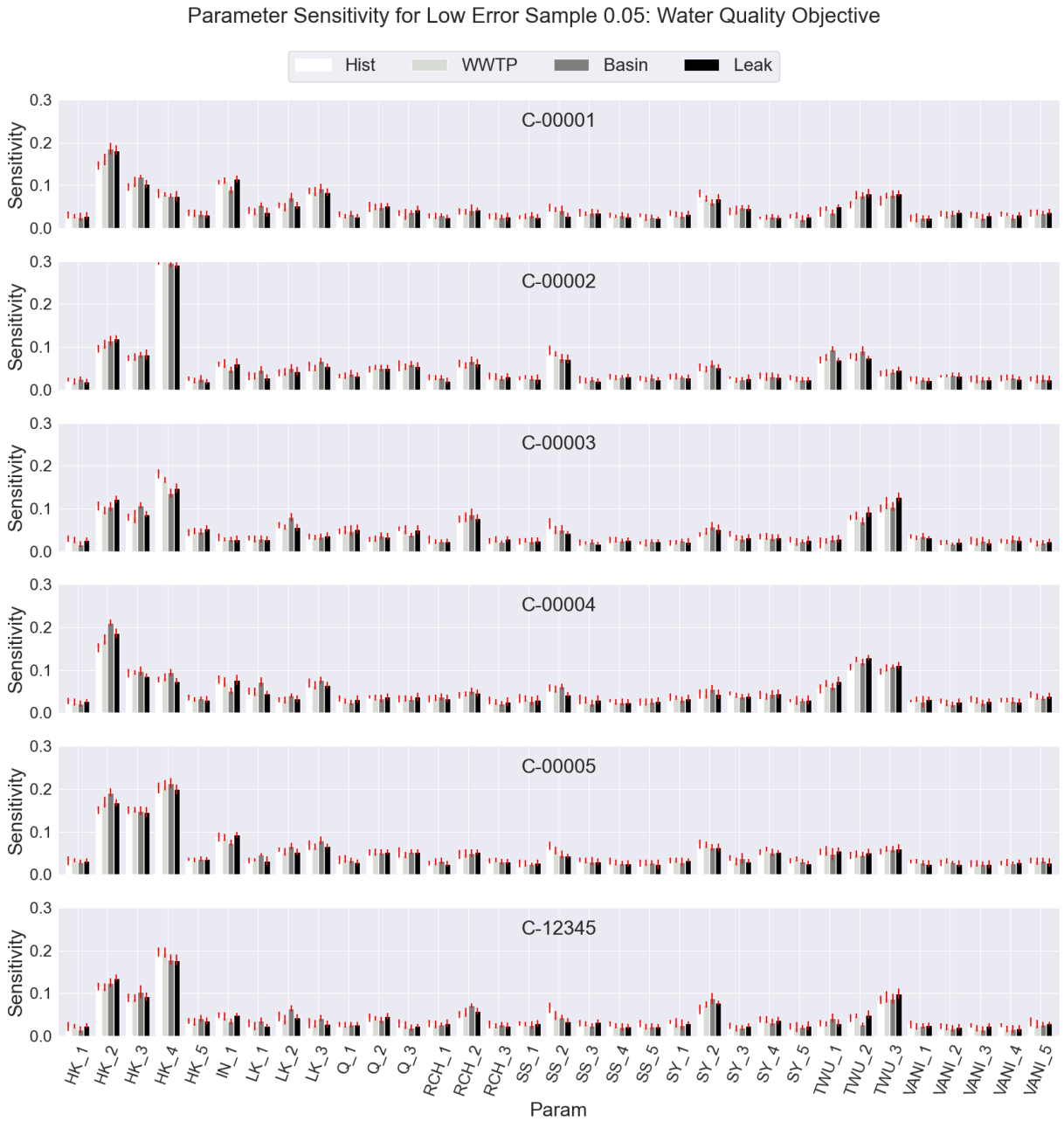

Figure 3.12: δ sensitivity of the water quality risk objective according to the 5,000 filtered samples for the 33 model parameters (columns). The sensitivity is shown by cluster (rows) and by the four alternatives from left to right (light to dark): historical, wastewater reuse, infiltration basins, and repair leaks.

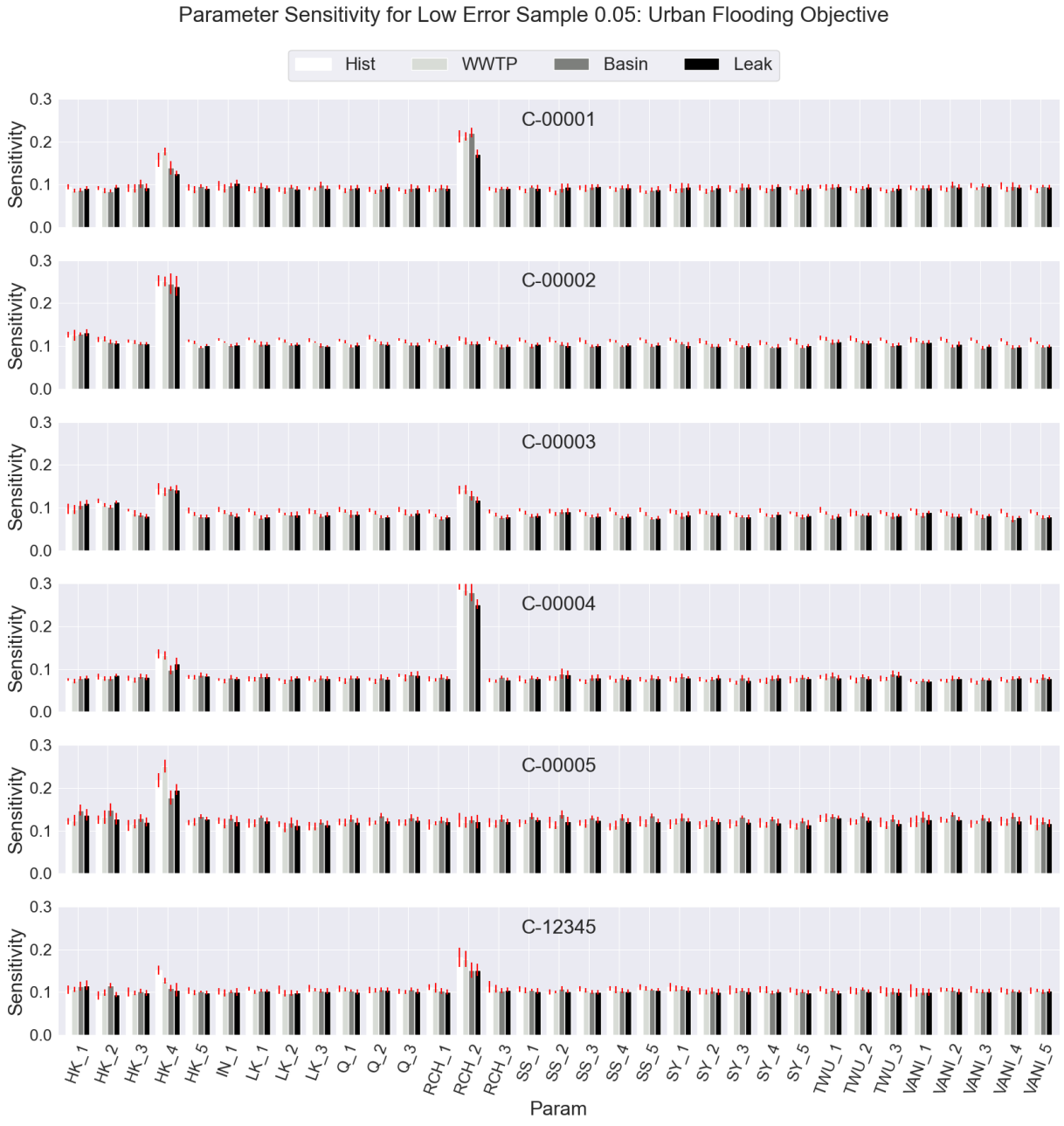

Figure 3.13: δ sensitivity of the urban flooding risk objective according to the 5,000 filtered samples for the 33 model parameters (columns). The sensitivity is shown by cluster (rows) and by the four alternatives from left to right (light to dark): historical, wastewater reuse, infiltration basins, and repair leaks.

3.7 Data Availability

The model with input datasets, observations, results, and postprocessing scripts are available in a GitHub repository at https://github.com/mrlmautner/UrbanGW/tree/sensitivityanalysis (DOI: 10.5281/zenodo.6039830).

3.8 Acknowledgements

This work has been supported in part by the Ford Foundation Predoctoral Fellowship Program of the National Academies of Science, Engineering, and Medicine. Research trips to Mexico City were funded in part by the University of California Davis Henry A. Jastro Graduate Research Award. We thank the Organismo de Cuencas: Aguas del Valle de México (OCAVM) of the National Water Commission (CONAGUA) of Mexico and the Instituto de Geofísica of the Universidad Nacional Autónoma de México (UNAM) for providing pumping and observation data, the conceptual groundwater model, and input on potential management alternatives.

Bibliography

Bakker, M., Post, V., Langevin, C. D., Hughes, J. D., White, J. T., Starn, J. J. and Fienen, M. N. (2016). Scripting MODFLOW Model Development Using Python and FloPy, *Groundwater* **54**(5): 733–739.

Bárdossy, A. (2007). Calibration of hydrological model parameters for ungauged catchments, *Hydrology and Earth System Sciences* **11**(2): 703–710.

Beven, K. (2016). Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication, *Hydrological Sciences Journal* **61**(9): 1652–1665.

Bhaskar, A. S., Beesley, L., Burns, M. J., Fletcher, T. D., Hamel, P., Oldham, C. E. and Roy, A. H. (2016). Will it rise or will it fall? Managing the complex effects of urbanization on base flow, *Freshwater Science* **35**(1): 293–310.

Borgonovo, E. (2007). A new uncertainty importance measure, *Reliability Engineering & System Safety* **92**(6): 771–784.

Borgonovo, E. and Plischke, E. (2016). Sensitivity analysis: A review of recent advances, *European Journal of Operational Research* **248**(3): 869–887.

Brunner, P., Doherty, J. and Simmons, C. T. (2012). Uncertainty assessment and implications for data acquisition in support of integrated hydrologic models, *Water Resources Research* **48**(7): 1–18.

Dams, J., Salvadore, E., Van Daele, T., Ntegeka, V., Willems, P. and Batelaan, O. (2012). Spatio-temporal impact of climate change on the groundwater system, *Hydrology and Earth System Sciences* **16**(5): 1517–1531.

Dams, J., Woldeamlak, S. T. and Batelaan, O. (2008). Predicting land-use change and its impact on the groundwater system of the Kleine Nete catchment, Belgium, *Hydrology and Earth System Sciences* **12**(6): 1369–1385.

Doherty, J. and Moore, C. (2020). Decision Support Modeling: Data Assimilation, Uncertainty Quantification, and Strategic Abstraction, *Groundwater* **58**(3): 327–337.

Doherty, J. and Simmons, C. T. (2013). La modélisation de nappe comme support de décision: Réflexions sur un cadre conceptuel unifié, *Hydrogeology Journal* **21**(7): 1531–1537.

Fletcher, S., Strzepek, K., Alsaati, A. and de Weck, O. (2019). Learning and flexibility for water supply infrastructure planning under groundwater resource uncertainty, *Environmental Research Letters* **14**(11): 114022.

Foster, S. S. D., Lawrence, A. and Morris, B. (1998). *Groundwater in urban development: assessing management needs and formulating policy strategies*, number no. 390 in *World Bank technical paper*, World Bank, Washington, D.C.

Galán-Breth, R. I. (2018). *Modelación matemática de nitratos en el agua subterránea en la región Sur de la Ciudad de México*, PhD thesis, Universidad Nacional Autónoma de México (UNAM).

Ganji, A., Maier, H. R. and Dandy, G. C. (2016). A modified Sobol' sensitivity analysis method for decision-making in environmental problems, *Environmental Modelling and Software* **75**: 15–27.

Guillaume, J. H. A., Hunt, R. J., Comunian, A., Blakers, R. S. and Fu, B. (2016). Methods for Exploring Uncertainty in Groundwater Management Predictions, *in* A. Jakeman, O. Barreteau, R. Hunt, J. Rinaudo and A. Ross (eds), *Integrated Groundwater Management*, Springer International Publishing, chapter 28, pp. 711–737.

Hadka, D., Herman, J., Reed, P. and Keller, K. (2015). An open source framework for many-objective robust decision making, *Environmental Modelling and Software* **74**: 114–129.

Herman, J. D., Quinn, J. D., Steinschneider, S., Giuliani, M. and Fletcher, S. (2020). Climate Adaptation as a Control Problem: Review and Perspectives on Dynamic Water Resources Planning Under Uncertainty, *Water Resources Research* **56**(2).

Herman, J. D., Reed, P. M. and Wagener, T. (2013). Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior, *Water Resources Research* **49**(3): 1400–1414.

Herrera-Zamarrón, G., Cardona-Benavides, A., González-Hita, L., Gutiérrez-Ojeda, C., Hernández-Calero, R., Hernández-García, G., Hernández-Laloth, N., López-Hernández, R. I., Martínez-Morales, M., Pita de la Paz, C., Sánchez-Díaz, L. F., Báez-Durán, J. A., Cruickshank-Villanueva, C. and Herrera-Revilla, I. (2005). Estudio para obtener la disponibilidad del acuífero de la Zona Metropolitana de la Ciudad de México, *Technical Report Contract No. 06-CD-03-1O-0272-1-06*, Secretaría del Medio Ambiente del Gobierno del Distrito Federal, Sistema de Aguas de la Ciudad de México (SACM), and Instituto Mexicano de Tecnología del Agua (IMTA), Mexico City.

Hrachowitz, M., Savenije, H., Blöschl, G., McDonnell, J., Sivapalan, M., Pomeroy, J., Arheimer, B., Blume, T., Clark, M., Ehret, U., Fenicia, F., Freer, J., Gelfan, A., Gupta, H., Hughes, D., Hut, R., Montanari, A., Pande, S., Tetzlaff, D., Troch, P., Uhlenbrook, S., Wagener, T., Winsemius, H., Woods, R., Zehe, E. and Cudennec, C. (2013). A decade of Predictions in Ungauged Basins (PUB)—a review, *Hydrological Sciences Journal* **58**(6): 1198–1255.

Hyde, K. M. and Maier, H. R. (2006). Distance-based and stochastic uncertainty analysis for multi-criteria decision analysis in Excel using Visual Basic for Applications, *Environmental Modelling and Software* **21**(12): 1695–1710.

Jing, M., Heße, F., Kumar, R., Kolditz, O., Kalbacher, T. and Attinger, S. (2019). Influence of input and parameter uncertainty on the prediction of catchment-scale groundwater travel time distributions, *Hydrology and Earth System Sciences* **23**(1): 171–190.

Kelleher, C., McGlynn, B. and Wagener, T. (2017). Characterizing and reducing equifinality by constraining a distributed catchment model with regional signatures, local observations, and process understanding, *Hydrology and Earth System Sciences* **21**(7): 3325–3352.

Kwakkel, J. H. and Haasnoot, M. (2019). Supporting DMDU: A Taxonomy of Approaches and Tools, *in* V. Marchau, W. Walker, P. Bloemen and S. Popper (eds), *Decision Making under Deep Uncertainty*, Springer International Publishing, pp. 355–374.

Lehr, C. and Lischeid, G. (2020). Efficient screening of groundwater head monitoring data for anthropogenic effects and measurement errors, *Hydrology and Earth System Sciences* **24**(2): 501–513.

Lopez-Alvis, J. (2014). *Calibración de un modelo de flujo del Acuífero de la Zona Metropolitana de la Ciudad de México (AZMCM)*, PhD thesis, Universidad Nacional Autónoma de México (UNAM).

Mai, J., Craig, J. R. and Tolson, B. A. (2020). Simultaneously determining global sensitivities of model parameters and model structure, *Hydrology and Earth System Sciences* **24**(12): 5835–5858.

Maier, H., Guillaume, J., van Delden, H., Riddell, G., Haasnoot, M. and Kwakkel, J. (2016). An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together?, *Environmental Modelling & Software* **81**: 154–164.

Matott, L. S. (2005). *OSTRICH: An optimization software tool: Documentation and users guide*.

Mautner, M. R., Foglia, L., Herrera, G. S., Galán, R. and Herman, J. D. (2020). Urban growth and groundwater sustainability: Evaluating spatially distributed recharge alternatives in the Mexico City Metropolitan Area, *Journal of Hydrology* **586**(April): 124909.

McMillan, H. K., Westerberg, I. K. and Krueger, T. (2018). Hydrological data uncertainty and its implications, *WIREs Water* **5**(6): 1–14.

Megdal, S. B., Gerlak, A. K., Varady, R. G. and Huang, L. Y. (2015). Groundwater Governance in the United States: Common Priorities and Challenges, *Groundwater* **53**(5): 677–684.

Mendoza, P. A., Clark, M. P., Mizukami, N., Gutmann, E. D., Arnold, J. R., Brekke, L. D. and Rajagopalan, B. (2016). How do hydrologic modeling decisions affect the portrayal of climate change impacts?, *Hydrological Processes* **30**(7): 1071–1095.

Montanari, A. and Di Baldassarre, G. (2013). Data errors and hydrological modelling: The role of model structure to propagate observation uncertainty, *Advances in Water Resources* **51**: 498–504.

Moore, C. and Doherty, J. (2005). Role of the calibration process in reducing model predictive error, *Water Resources Research* **41**(5): 1–14.

Mustafa, S. M. T., Hasan, M. M., Saha, A. K., Rannu, R. P., Van Uytven, E., Willems, P. and Huysmans, M. (2019). Multi-model approach to quantify groundwater-level prediction uncertainty using an ensemble of global climate models and multiple abstraction scenarios, *Hydrology and Earth System Sciences* **23**(5): 2279–2303.

Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame, J. M., Prieto, C. and Gupta, H. V. (2021). What Role Does Hydrological Science Play in the Age of Machine Learning?, *Water Resources Research* **57**(3).

OCAVM (2014). Programa Hídrico Regional 2014-2018: Region Administrativo Hidrológico XIII, Aguas del Valle de México, *Technical report*, Comisión Nacional del Agua, Tlalpan, Mexico, D.F.

Peters-Lidard, C. D., Clark, M., Samaniego, L., Verhoest, N. E. C., van Emmerik, T., Uijlenhoet, R., Achieng, K., Franz, T. E. and Woods, R. (2017). Scaling, similarity, and the fourth paradigm for hydrology, *Hydrology and Earth System Sciences* **21**(7): 3701–3713.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B. and Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow, *Environmental Modelling & Software* **79**: 214–232.

Plischke, E., Borgonovo, E. and Smith, C. L. (2013). Global sensitivity measures from given data, *European Journal of Operational Research* **226**(3): 536–550.

Poeter, E. P., Hill, M. C., Lu, D., Tiedeman, C. and Mehl, S. W. (2014). UCODE_2014, with new capabilities to define parameters unique to predictions, calculate weights using simulated values, estimate parameters with SVD, evaluate uncertainty with MCMC, and more, *Technical report*, Integrated Groundwater Modeling Center (IGWMC), of the Colorado School of Mines.

Qiu, J., Yang, Q., Zhang, X., Huang, M., Adam, J. C. and Malek, K. (2019). Implications of water management representations for watershed hydrologic modeling in the Yakima River basin, *Hydrology and Earth System Sciences* **23**(1): 35–49.

Ravalico, J. K., Dandy, G. C. and Maier, H. R. (2010). Management Option Rank Equivalence (MORE) - A new method of sensitivity analysis for decision-making, *Environmental Modelling and Software* **25**(2): 171–181.

Ravalico, J. K., Maier, H. R. and Dandy, G. C. (2009). Sensitivity analysis for decision-making using the MORE method-A Pareto approach, *Reliability Engineering and System Safety* **94**(7): 1229–1237.

Razavi, S. and Gupta, H. V. (2015). What do we mean by sensitivity analysis? the need for comprehensive characterization of "global" sensitivity in Earth and Environmental systems models, *Water Resources Research* **51**(5): 3070–3092.

Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Lo Piano, S., Iwanaga, T., Becker, W., Tarantola, S., Guillaume, J. H., Jakeman, J., Gupta, H., Melillo, N., Rabitti, G., Chabridon, V., Duan, Q., Sun, X., Smith, S., Sheikholeslami, R., Hosseini, N., Asadzadeh, M., Puy, A., Kucherenko, S. and Maier, H. R. (2021). The Future of Sensitivity Analysis: An essential discipline for systems modeling and policy support, *Environmental Modelling and Software* **137**(December 2020).

Refsgaard, J. C., Christensen, S., Sonnenborg, T. O., Seifert, D., Højberg, A. L. and Troldborg, L. (2012). Review of strategies for handling geological uncertainty in groundwater flow and transport modeling, *Advances in Water Resources* **36**: 36–50.

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L. and Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process - A framework and guidance, *Environmental Modelling and Software* **22**(11): 1543–1556.

Reinecke, R., Foglia, L., Mehl, S., Herman, J. D., Wachholz, A., Trautmann, T. and Döll, P. (2019). Spatially distributed sensitivity of simulated global groundwater heads and flows to hydraulic conductivity, groundwater recharge, and surface water body parameterization, *Hydrology and Earth System Sciences* **23**(11): 4561–4582.

Reusser, D. E. and Zehe, E. (2011). Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity, *Water Resources Research* **47**(7): 1–15.

Rojas, R., Feyen, L., Batelaan, O. and Dassargues, A. (2010). On the value of conditioning data to reduce conceptual model uncertainty in groundwater modeling, *Water Resources Research* **46**(8): 1–20.

Şalap-Ayça, S. and Jankowski, P. (2016). Integrating local multi-criteria evaluation with spatially explicit uncertainty-sensitivity analysis, *Spatial Cognition and Computation* **16**(2): 106–132.

Singh, A. (2014). Groundwater resources management through the applications of simulation modeling: A review, *Science of the Total Environment* **499**: 414–423.

Tiedeman, C. R., Ely, D. M., Hill, M. C. and O'Brien, G. M. (2004). A method for evaluating the importance of system state observations to model predictions, with application to the Death Valley regional groundwater flow system, *Water Resources Research* **40**(12): 1–14.

Tolley, D., Foglia, L. and Harter, T. (2019). Sensitivity Analysis and Calibration of an Integrated Hydrologic Model in an Irrigated Agricultural Basin With a Groundwater-Dependent Ecosystem, *Water Resources Research* **55**(9): 7876–7901.

Tonkin, M. J., Tiedeman, C. R., Ely, D. M. and Hill, M. C. (2007). OPR-PPR, a Computer Program for Assessing Data Importance to Model Predictions Using Linear Statistics, *Techniques and Methods* p. 115.

Vázquez-Suñé, E., Carrera, J., Tubau, I., Sánchez-Vila, X. and Soler, A. (2010). An approach to identify urban groundwater recharge, *Hydrology and Earth System Sciences* **14**(10): 2085–2097.

Wada, Y., Bierkens, M. F. P., de Roo, A., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., Konar, M., Liu, J., Müller Schmied, H., Oki, T., Pokhrel, Y., Sivapalan, M., Troy, T. J., van Dijk, A. I. J. M., van Emmerik, T., Van Huijgevoort, M. H. J., Van Lanen, H. A. J., Vörösmarty, C. J., Wanders, N. and Wheater, H. (2017). Human–water interface in hydrological modelling: current status and future directions, *Hydrology and Earth System Sciences* **21**(8): 4169–4193.

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S. and Gupta, H. V. (2003). Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis, *Hydrological Processes* **17**(2): 455–476.

Zhang, X. and Liu, P. (2021). A time-varying parameter estimation approach using split-sample calibration based on dynamic programming, *Hydrology and Earth System Sciences* **25**(2): 711–733.

Chapter 4

Socially informed spatial analysis: evaluating the role of aggregation scale in modeling differential impacts of urban groundwater pumping policies across socioeconomic indicators[1]

4.1 Abstract

Groundwater supply planning in urban settings is often achieved at the regional scale while the effects of excessive aquifer exploitation are experienced on a local scale. In the Mexico City Metropolitan Area, land and water use changes have caused massive overdraft and subsidence, threatening water supply, particularly in marginalized communities. Regional-scale, spatially distributed groundwater models have been developed in prior work to test model uncertainty and policy interventions. However, there is little understanding of how regional aggregation of spatially distributed planning objectives can obscure or exacerbate existing inequalities. This study uses socially informed spatial analysis to investigate the differential impacts of urban groundwater pumping policies at multiple scales and understand how decision-making is affected by spatial aggregation of system performance. Specifically, we compare four planning objectives calculated

[1]This chapter will be submitted to *Journal of Water Resources Planning and Management*: Mautner, M. R. L. and Herman, J. D. as "Socially informed spatial analysis: evaluating the role of aggregation scale in modeling differential impacts of urban groundwater pumping policies across socioeconomic indicators".

at three spatial resolutions (municipality, census block, and cluster) to determine whether the preferred groundwater pumping policies show a relationship with socioeconomic indicators. Three major findings result from this comparison. First, the multi-objective performance of pumping policies can neglect marginalized communities according to multiple indicators. Second, clustering of model grid cells according to socioeconomic indicators for the evaluation of management objectives at a model-relevant scale can allow for more accurate analysis of the effects of policies on marginalized populations than when using existing spatial unit scales. Finally, the determination of policy preference at the subregional scale and considering the socioeconomic characteristics of subregional units can aid in equitable and reasoned selection of aquifer management alternatives at the regional scale. This study highlights the importance of understanding how the interaction between spatially distributed management objectives and local hydrogeologic and socioeconomic characteristics can affect water supply decision-making processes in densely populated urban areas, and shows the potential for adverse impacts to be concentrated in the communities least able to manage them.

4.2 Introduction

Groundwater resources are becoming increasingly stressed as the effects from extreme climatic events and historical pumping regimes lead to greater competition between current and future water users, including domestic, industrial, agricultural, and environmental (Megdal et al., 2015; de Graaf et al., 2019). Urban regions in particular are affected as the problems of dwindling groundwater resources are further exacerbated by negative impacts to groundwater quality, concurrent flooding from artificial recharge, and infrastructure damages from subsidence, among other issues (Schirmer et al., 2013; Foster, 2020). Over the past three decades, integrated groundwater models of urban areas have greatly improved our understanding of how infrastructure and site-specific hydrogeologic characteristics interact, giving researchers the ability to more accurately identify, parameterize, and simulate the dominant processes involved, as well as test future aquifer recharge strategies

to enhance regional groundwater planning (Lerner, 2002; Jacobson, 2011). Regional-scale urban groundwater supply planning must encapsulate heterogeneous hydrogeological settings, a vast array of competing interests, and spatially distributed management responses, making it a paragon of the "wicked problems" defined by Rittel and Webber (1973). As a response, multi-objective analysis has become a necessity in simulating and weighing policy solutions designed to assuage competing needs in increasingly complex groundwater systems (Reed et al., 2013). This is especially true given that policies may produce undesirable outcomes at the subregional scale despite aggregate improvements at the city or regional scale. Concurrently, social scientists have stressed the need for the incorporation of distributive justice theory within regional aquifer management to ensure both equitable access to high-quality groundwater and the correction of, or compensation for, damages caused by excessive pumping (Zhou, 2009; Neal et al., 2016).

Walker et al. (2003) define five sources of uncertainty within the model-based water management process: context and framing, input uncertainty, model structure uncertainty, parameter uncertainty, and model technical uncertainty. In regional water resources modeling, context and framing uncertainty is often explored as it relates to defining boundary conditions and indicators that can be aggregated across the model area. Recent advances in the field of multi-objective decision-making have begun to evaluate how uncertainty in objective framing propagates through the modeling and analysis process to affect measured performance and selection of water resources management policies by using rival framings to compare across a variety of competing problem formulations (Quinn et al., 2017). Such methods include the comparison of policy performance under a selection of robustness metrics in the face of a changing climate and evolving demands across a diverse set of sectors and users, showing that stakeholders within a basin can experience vastly different effects from the same uncertain futures as shown by Hadjimichael et al. (2020). These differences are particularly important in the context of equity-informed decision-making that aims to incorporate the distributional effects of policy options on stakeholders into the modeling and analysis process (Yang et al., 2022). The advent of such research underscores the need to understand the social implications of the generation of model structure in an ethical context, approaching sensitivity

analysis not only from a quantitative standpoint, but also an epistemological framing that considers how different communities can be unevenly affected by modeling choices (Razavi et al., 2021). Nonetheless, calls to consider the social equity implications of groundwater policy implementation in hydrogeologic analysis, which go back decades (Barthel et al., 2017; Re et al., 2022) have gone relatively unaddressed. Thus, the evaluation of groundwater management alternatives in conjunction with spatially determined social indicators, such as poverty and healthcare access, is still nascent (Zagonari, 2010; Bach et al., 2014).

Many current socio-hydrological studies promote the introduction of human decisions and actions into the modeling process using tools to approximate human behavior, such as agent-based modeling and coupled differential equations (Gain et al., 2021). However, many socio-hydrology models carry high levels of uncertainty in the human behavior components of such integrated models and, additionally, often lack explicit spatial representation, instead aggregating regional indicators into a single scalar value to be dynamically updated within the modeling framework (Wada et al., 2017). Despite the difficulty in calibrating such models to real-world situations, such studies can be useful in answering questions about equity between populations, as in Zagonari (2010), where optimal control is used to compare the performance of a groundwater exploitation strategy between current and future generations. Unfortunately, the complexity of such comparisons can create difficulties in also representing spatially distributed performance, and is still unexplored. Recent interdisciplinary research on the implications of spatially neutral groundwater management postulated that the spatial disaggregation of management objectives, such as the water budget and safe yield, could be critical to developing equitable aquifer recharge policies, without implementing nor evaluating such disaggregation (Cabello et al., 2012). Thus, the impact of spatial boundary selection on policy performance under management objectives, in addition to potential recharge quantities, remains an open question although we may be able to look to other fields of spatially distributed environmental modeling for examples.

Jafino, Kwakkel and Taebi (2021) define a list of twelve requirements to enable the evaluation of distributive justice in model-based support for climate planning, which are also applicable to

regional-scale water resources planning. Among the requirements are value-based disaggregated metrics, postprocessing of metrics to account for distributive principles, and disaggregated representation of values. Even among climate change models, the presence of these requirements is scant and it is often difficult to quantify the many intersections of infrastructural, meteorological, and social components. Similarly, a number of flood risk studies have generated important findings that are more readily visualized regarding equity between stakeholders at the subregional scale. For example, Ciullo et al. (2020) found that when using strict cost-benefit aggregation over subregions, aggregated risk reduction came at the expense of specific subregions that experienced increase in risk by performing multi-objective optimization for a flood risk management problem. These issues were mitigated by reformulating the problem using egalitarian and prioritarian social-welfare principles to formulate objective functions, which led to better performance under Gini-based indicators that evaluate equity. Similarly, a study from La Rosa and Pappalardo (2020) incorporates the number of children and elderly into analysis of sustainable urban drainage system ability to reduce flood risk in urban areas. Specifically, the study mentioned is one of very few that attempts to visualize how the social structure of a city is related to the benefits achieved through the implementation of water resources management policies. Yang et al. (2022) formulate rival framings designed with and without using an equity index within the optimization problem to address marginalized objectives in an attempt to improve upon conventional hydrologic modeling. However, evaluation of the effects of scale, particularly in the context of the definition of the areas over which water management objectives are calculated, in multi-objective hydrologic analysis is sparse. While analysis of spatially distributed impacts of uncertainties in policy analysis has been recognized to some extent in economic modeling, the use of such analysis in numerical groundwater modeling and the evaluation of non-monetary based performance objectives to promote equity are limited (Kind et al., 2017; Jafino, Kwakkel, Klijn, Dung, van Delden, Haasnoot and Sutanudjaja, 2021).

Thus, this study aims to fill the gap in evaluating rival spatial framings of a groundwater management problem in the context of social marginalization according to multiple socioeconomic

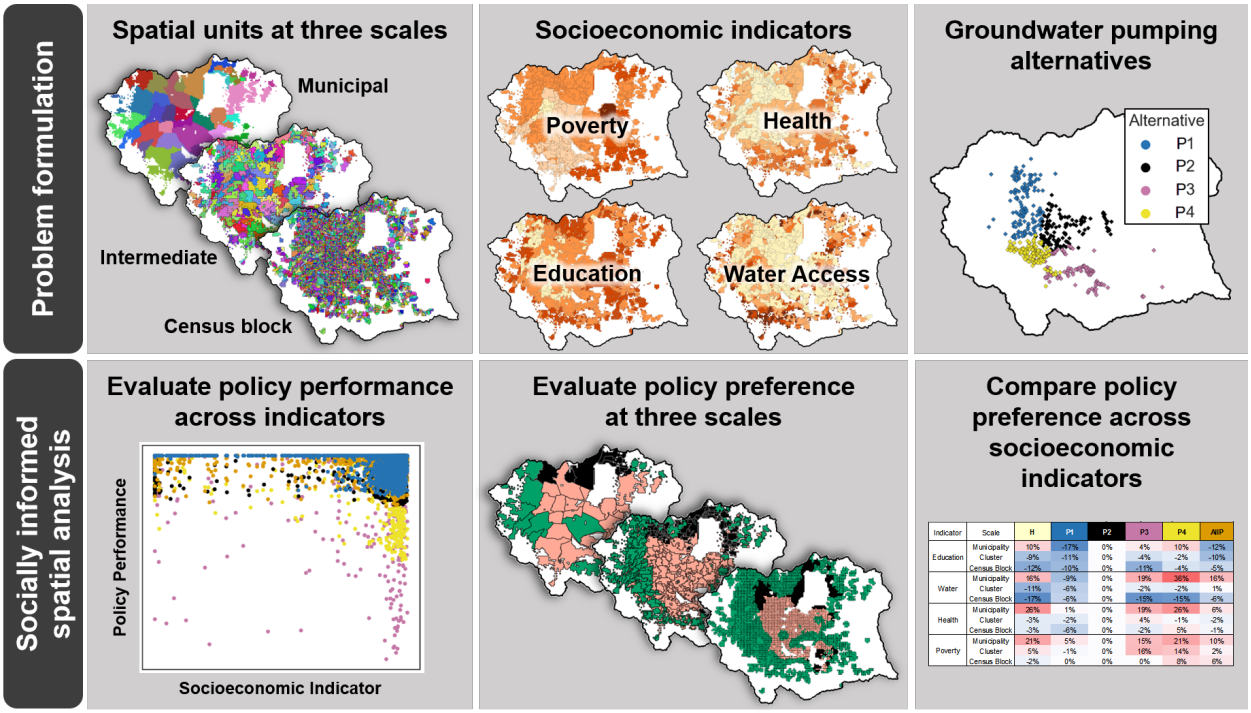

Figure 4.1: Problem formulation and summary of socially informed spatial analysis implemented.

indicators. There are complex challenges of incorporating equity into spatially distributed groundwater planning, creating a number of research questions that we undertake. First, what limitations arise when attempting to compare spatially heterogeneous social indicators across gridded groundwater model output at the subregional scale, and how can these limitations be mitigated through modeling practices? Second, how does groundwater pumping policy performance interact with social indicators at subregional scales, and does it perpetuate the challenges faced by marginalized populations in the region? And, finally, how do spatial differences in preference for groundwater pumping policies change based on the scale of spatial units evaluated? We address these questions by evaluating the performance of five spatially distributed urban groundwater pumping policies against a baseline historical policy at three subregional scales to understand how decision-making is affected by the spatial aggregation of system performance in the problem framing.

4.3 Methodology

4.3.1 Study Area

The Mexico City Metropolitan Area (MCMA) serves as a valuable case study to evaluate the interaction between spatially distributed aquifer management and community-level socioeconomic indicators. The MCMA lies within the southwestern portion of the Valley of Mexico watershed, characterized by volcanic peaks surrounding a high plains basin (OCAVM, 2014). Once a lake system, the Valley of Mexico has been transformed over the last 700 years through the development of the region, first by the Mexica, then Spanish settlers, and finally in the modern era. Over time, the spatial externalization of risks by those in power have led to the marginalization of communities both in terms of the concentration of long term flooding and the lack of access to municipal water systems (Tellman et al., 2018). The effects of an overburdened groundwater system, on the one hand from sinking groundwater tables and on the other hand from novel water sources (e.g. distribution system leaks and landscape irrigation), have created a need for a change in course from historical groundwater management, or lack thereof. Notably, engineers within the city have and continue to make decisions based on political expediency, destining socioeconomically disadvantaged communities to flooding for the good of the basin as a whole and allowing for relatively unimpeded abstraction of groundwater to spur economic growth (Chahim, 2022). This has occurred through policies that prioritize fast-track development of the urban core while simultaneously designating peri-urban areas for slow-track regularization, resulting in limited proliferation of urban services such as tap water in disadvantaged regions throughout the basin (Wigle, 2020). While the effects of these policies on marginalized populations has been described through ethnographic and historical analysis, the evaluation of such relationships through the use of hydrologic modeling remains untapped. This paper uses a case study of the urban aquifer management problem in the Valley of Mexico using a spatially distributed groundwater model adapted from prior work (Herrera-Zamarrón et al., 2005; Lopez-Alvis, 2014; Galán-Breth, 2018; Mautner et al., 2020) to evaluate promising managed aquifer recharge alternatives while taking into

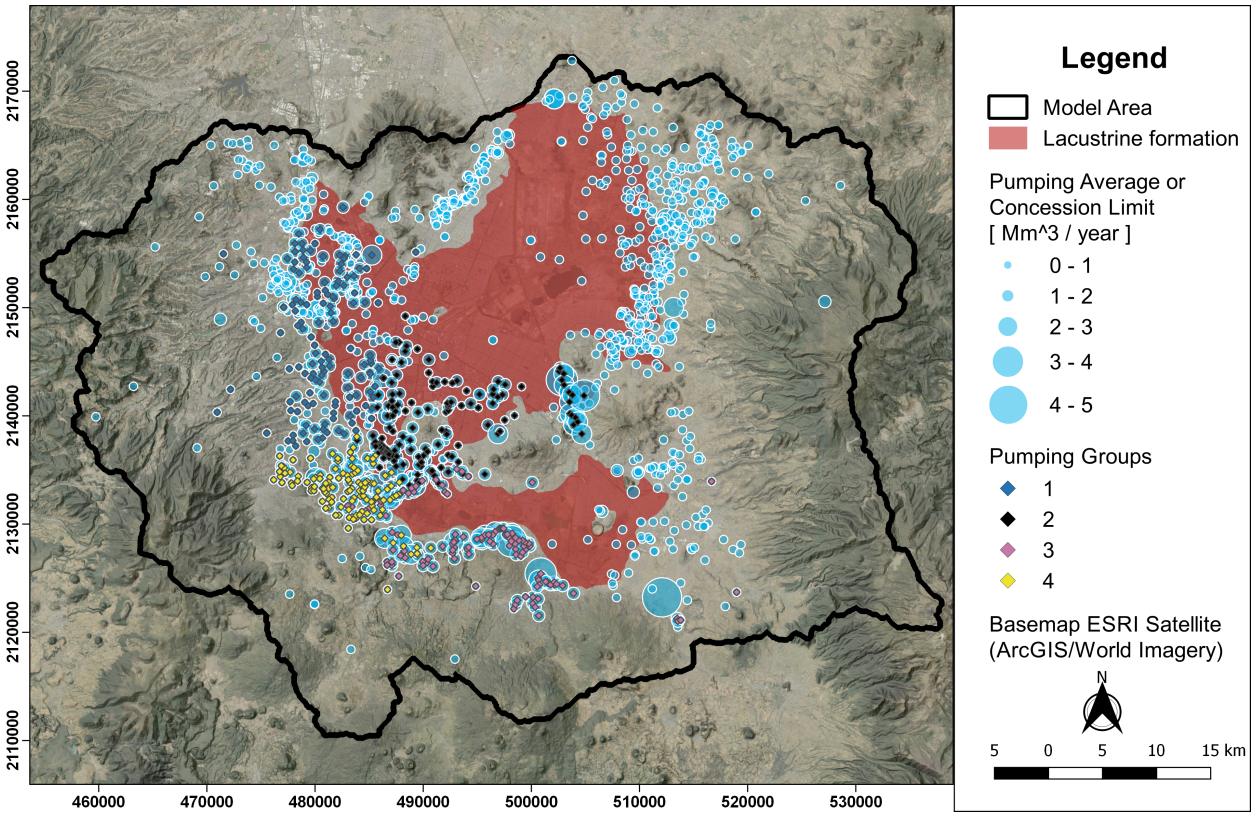

Figure 4.2: Pumping quantities (blue circles) of all known pumping wells in the Valley of Mexico shown as either the average over the model period (1984-2013) for municipal wells or the concession limit for private wells. Additionally, municipal pumping wells used in the aquifer management alternatives separated into four groups of approximately equal pumping quantities are indicated by colored diamonds.

account the social context of the impacts of such policies at a localized scale.

4.3.2 Urban Groundwater Model

The Valley of Mexico model is written in Python using the *flopy* package to preprocess data and run the model in MODFLOW, a widely used software which solves the groundwater flow equation (Bakker et al., 2016), as presented in Mautner et al. (2020). The following is a brief overview of the Valley of Mexico test case. A set of model parameters govern model representation of geologic setting, land use and land cover, and water resource infrastructure in Mexico City, including artificial and natural recharge, time varied groundwater pumping, and heterogeneous subsurface characteristics. This model covers an area of 84 km by 67 km on a 500x500 m spatial grid, and the time period from 1984 to 2013 on a daily time-step and a monthly stress-period.

Leaks are represented in the system by distributing a proportion of total water use over the urban area according to historical estimates for leaks and multiplying by an infiltration factor.

4.3.3 Pumping Alternatives

In previous studies, demand management resulting from repairing leaks in the distribution system performed the best according to objectives that were positively correlated with increased recharge within the valley (Mautner et al., 2020, 2022). Therefore, this study evaluates spatially grouped reduction in pumping at the order of magnitude of the demand reduction in the repair leaks alternative from the previous studies. Approximately 45% of the total pumping in the valley is attributed to reported municipal pumping. The alternatives in this study assume that all pumping reductions are done within the municipal pumping dataset. This is done by selecting and eliminating approximately one quarter of the municipal dataset pumping by average volumetric pumping rate as follows.

Average annual pumping for each well is calculated for each of the three model phases (1984-1994, 1995-2004, 2005-2013) by summing the monthly pumping over each of the model phases and dividing by the number of months in each phase (132, 120, and 108 months, respectively). The groups of pumping wells were then selected first based on geographic proximity, and second to ensure an even quantity of pumping in the groups during each model phase and over the entire model period. The final groups are shown in Figure 4.2 along with the distribution of total groundwater pumping within the model area. The distribution of average pumping in the final four groups is shown in Table 4.1.

Table 4.1 Proportion of average pumping attributed to each pumping group P1-4 out of the total reported municipal pumping over each of the three model phases.

Model phase	**Group proportion**				**Municipal**
	1	**2**	**3**	**4**	**pumping** $[m^3/s]$
Phase 1 (1984-1994)	0.248	0.245	0.256	0.251	13.26
Phase 2 (1995-2004)	0.249	0.256	0.247	0.248	13.78
Phase 3 (2005-2013)	0.250	0.252	0.250	0.248	13.29
Entire model period	0.249	0.251	0.251	0.249	13.43

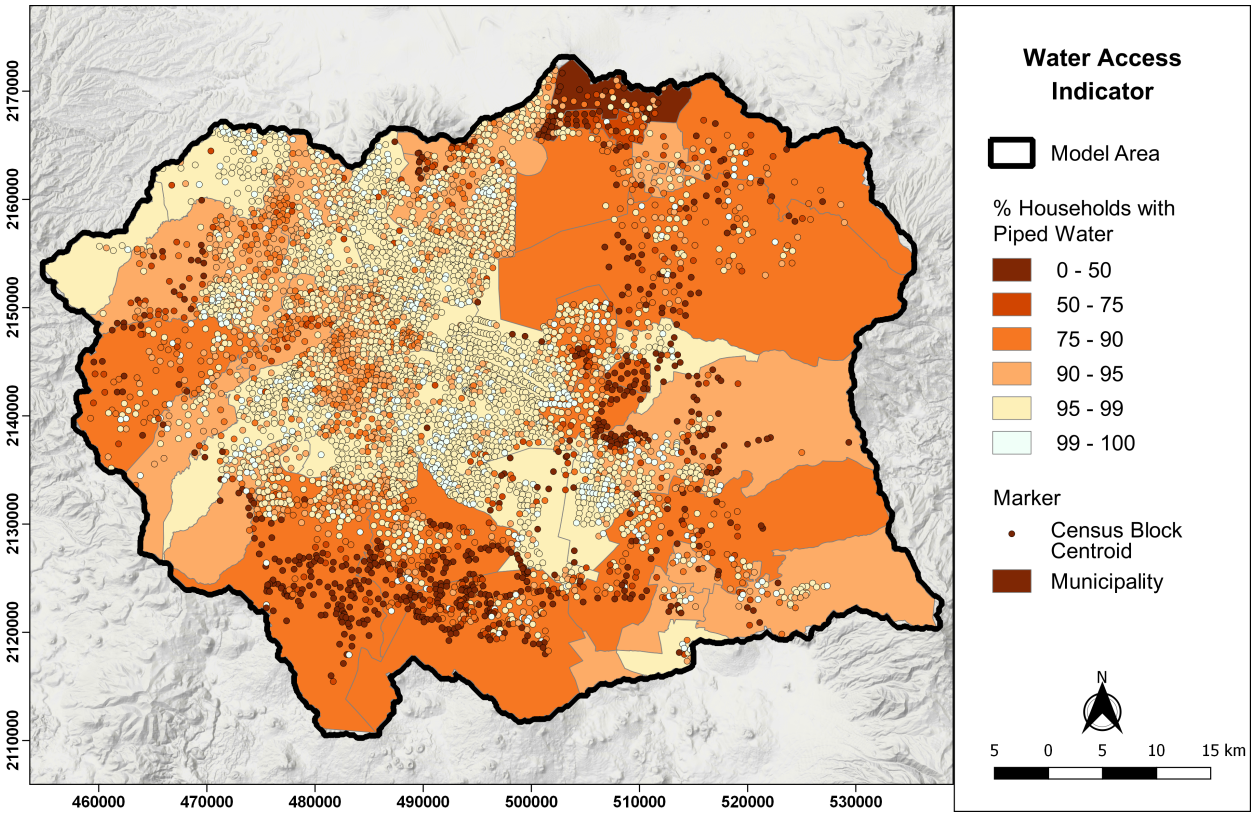

Figure 4.3: Water access indicator for each of the municipality units (polygons) and centroids of the census block units (points). Darker indicates a lower fraction of households with piped water within the dwelling unit.

Six management pumping policies were developed based on modifications to the pumping across the four groups shown in Figure 4.2. In the historical alternative (H), municipal pumping is unmodified from the reported pumping during the model period. In alternatives P1-P4, historical pumping is eliminated altogether in each of the four groups, respectively. In the final alternative (AllP), pumping is reduced in all four groups by 25% simultaneously. For each of the alternatives except the historical (business as usual), 25% of the total municipal pumping is also subtracted from the total amount that is infiltrated as leaks.

4.3.4 Socioeconomic Indicators

In Mexico City, surface and groundwater decision-making occurs at the state and federal level, while effects from these decisions may be felt unequally at the municipal or community

level. The vulnerability of populations in a given municipality can be determined by a number of socioeconomic factors. For example, the fraction of the population with an income below the level for basic needs can be considered to be unable to adapt to challenges to water access, such as retrofitting and repair to damages caused by land subsidence, unreliable supply, and flooding impacts to transportation infrastructure.

Social marginalization is caused by a number of intersecting factors. To encompass a selection of indicators that are important for communities trying to mitigate or adapt to negative consequences of the overexploitation of groundwater within the basin, four social indicators are developed: Health, Education, Water Access, and Poverty. These indicators were derived from the 2010 National Census and a lower value indicates increased marginalization INEGI (2016). The poverty indicator is available at the municipal level, while the other three indicators are available at the census block level as polygons for urban census blocks and points for rural census blocks. Health is measured by the percentage of the population with public or private health insurance. Education is measured as the average educational grade level attained by the adult population. Water access is measured by the percentage of households with piped water inside the dwelling unit. Finally, poverty is measured as the percentage of the population with an income above the level necessary to satisfy basic needs.

The example of the census block and municipal level distribution of the water access indicator is shown in Figure 4.3. As a large metropolitan area with rural subregions, there is a wide range in the presence of tap water within dwelling units throughout the model area. In Figure 4.3 (as well as Figure 4.12 and Figure 4.13) the census block information is shown as the centroids of the census block polygons and the municipal level information is shown as polygons. By visualizing both scales simultaneously it is apparent that a low resolution at the municipal level leads to a smoothing of the social indicator range, creating poor representation of particularly marginalized regions in the south and east. With the availability of the high resolution census block data, there is a much greater potential to understand the interaction between marginalization and the effects of groundwater management alternatives. However, Figure 4.4 shows the difficulty of representing

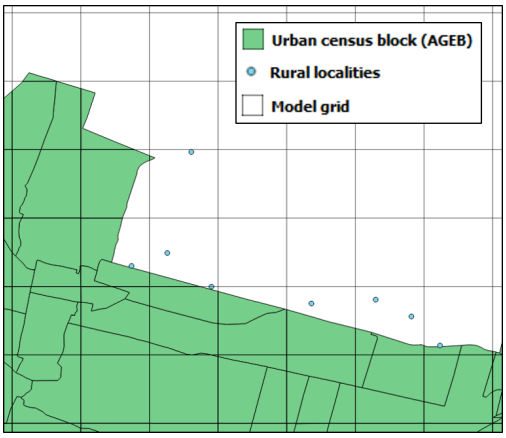

Figure 4.4: Close-up of census block polygons, rural census block points, and numerical groundwater model 500x500 meter grid.

the census block units at the model grid scale, as there are multiple census blocks represented in a single 500 m by 500 m grid cell. As in any regional groundwater model, this model is limited by the understanding of the underlying geology and its properties, availability of historical observations throughout the model area, and uncertainty associated with each of the boundary conditions and flux inputs that are required to construct the model. Increasing the resolution of the spatial discretization of a model can help to improve model representation in many cases where the data warrants (White et al, 2020). However, increasing model complexity or making conclusions based on results from individual grid cells in a regional model may mislead regional decision-makers on the confidence in the spatial information content of the model at that scale (Schöniger et al, 2015). Thus, given that the model discretization resolution cannot defensibly be increased to accommodate the scale of the census block, a method to measure objectives at the census block scale and to make the census block scale indicator information available at the grid scale is necessary.

4.3.5 Spatial resolution

To match the planning objective calculations (described below), the social indicators for the municipal and census block spatial units are resampled at the model grid scale. For the census block scale, the census block with the largest area is selected to represent each grid cell, while the municipal polygons are simply resampled at a raster resolution of 500x500 meters. At the census

block scale, grid cells with less than 5% of the area covered by a census block are removed to avoid edge effects. Any census blocks with a population of zero are also removed to avoid divide by zero errors in indicator calculations.

An intermediate-scale spatial unit is also developed so as to not obscure the distribution of marginalized areas while also providing a large enough unit size to match the scale of the groundwater model **spatial information**. For the medium scale unit, the social indicators are calculated at the model grid cell resolution by splitting the population for each census block by area across all intersecting grid cells, and then performing a population-weighted average of the indicator for each census block represented in a grid cell. Then spectral clustering using k-nearest neighbor affinity ($k = 10$) with the *scikit-learn* Python package is performed for the grid cells. Spectral clustering uses an affinity matrix to determine the best distribution of points between n clusters (n=200). We use social indicator and spatial information to quantify the relative similarity, or affinity, between each pair of all grid cells. Specifically, the four social indicators and the x, y, and z coordinates result in seven total features for each grid cell. All feature values are rescaled from 0 to 1, with the minimum and maximum values for each of the features set to 0 and 1, respectively, resulting in an equal weighting across all features. This intermediate scale is referred to as the cluster scale throughout the study.

4.3.6 Objectives

The planning objectives are modeled for each spatial unit at each of the three scales described above for each of the six pumping policies. The four aquifer planning objectives are water availability, change in average depth to groundwater, water quality risk, and urban flood risk (Equations 1-5).

In the Valley of Mexico, many residents receive water via tankers or "pipas", which are filled at nearby groundwater pumping stations. To represent the importance of proximity to groundwater pumping sources in water access, grid cell water availability (A_g) is calculated as the total groundwater pumped ($pump$) during all model periods (m) at all groundwater pumping wells (w) within

10 grid cells (5 km) in each direction to the cell (21 x 21 = 441), to model the local availability. The total groundwater pumping available is then divided by the sum of the population (pop) in all grid cells (g) within 10 grid cells in each direction to determine a per capita availability based on that cell's location. A_g is then summed over all cells within each spatial unit (s) and divided by the number of cells (c) in that unit to determine the average water availability (A_s) in each spatial unit. A is measured in m^3 per capita and is calculated starting in the third year of the model period to avoid spin-up effects. Pumping policies with a higher per capita groundwater availability are preferred.

$$A_g = \left(\frac{1}{n_{months} - 24}\right) \times \left(\sum_{m=25}^{n_{months}} \sum_{w=1}^{n_{wells}} pump_{w,m}\right) \times \left(\frac{1}{\sum_{g=1}^{441} pop_g}\right) \tag{4.1}$$

$$A_s = \frac{1}{c} \sum_{g=1}^{n_{cells}} A_g \tag{4.2}$$

Falling water tables in the region have been correlated with damaging subsidence and costs to deepen pumping wells, among other impacts. Therefore, the second aquifer management objective measures the change in average depth to groundwater from the beginning of the model period to the end of the model period. Change in depth to groundwater (D) is calculated as the sum of depths to groundwater ($depth$) over all active cells (c) in the subregion over all months (m) in the end year minus the sum of depths to groundwater in the start year, each of which are divided by the number of active cells in the subregion (C) times 12 months in a year. D is measured in m, is averaged over each of the years evaluated to avoid seasonal impacts, and is calculated using the third year of the model period for the initial year to avoid spin-up effects. This objective is minimized for better performance, meaning a decrease in depth to groundwater is preferred.

$$D = \frac{\sum_{m=349}^{360} \sum_{c=1}^{n_{active}} depth_{m,c}}{C \times 12} - \frac{\sum_{m=25}^{36} \sum_{c=1}^{n_{active}} depth_{m,c}}{C \times 12} \tag{4.3}$$

Water quality in the lacustrine layer (see Figure 4.2) is known to be poorer than that in the underlying alluvial aquifer resulting from both fine particulate matter and salts present. Therefore,

water quality is at risk of contamination in areas where the piezometric level falls below the bottom of the confining lacustrine layer. Thus, the water quality risk objective (W_s) indicates the number of cells (l) where groundwater levels remain below the bottom elevation of the confining lacustrine layer divided by the total number of lacustrine cells in the spatial unit (L_s) during the time periods (t) in the last year of the model period. A lower value of the water quality risk metric is preferred.

$$W_s = \frac{\sum_{t=348}^{360} l_{t,s}}{\sum_{t=348}^{360} L_s} \tag{4.4}$$

Finally, seasonal flooding resulting from groundwater mounding is common in the region, particularly in parts of the Valley that were once lacustrine. To address this risk, the urban flooding objective (F_s) is defined as the sum of the urban area in cells with groundwater mounding (u) above the land surface divided by the total urban area in each spatial unit (U) during the months (m) in the last year of the model period. The last year was chosen given that the objective is meant to measure the long term effects of lowering or rising water tables. This objective is minimized for better performance, meaning a decrease in the percentage of urban cells with groundwater mounding is preferred.

$$F_s = \frac{\sum_{m=349}^{360} u_m}{\sum_{m=348}^{360} U_m} \tag{4.5}$$

Performance under each of the four objectives are plotted against the four social indicators across the policies to understand how the effects from management alternatives are experienced by populations of varying marginalization. This is achieved by focusing on the difference between performance in the objectives of the historical and the five other alternatives, calculated by subtracting the historical objective value from the objective value calculated for policies P1-4 and AllP for each respective spatial unit.

In addition to calculating objectives for each spatial unit, objectives were measured at the regional scale by calculating across all populated grid cells, active grid cells, lacustrine grid cells, and urban grid cells for the availability, depth to groundwater, water quality, and urban flood risk

objectives, respectively, for comparison.

4.3.7 Preferred policies

In addition to performance under the objectives, it is also important to analyze the dynamics between local preference for management alternatives and social indicators. We determine spatial unit preference for management alternatives by selecting the management alternatives that are Pareto optimal or nondominated for each spatial unit based on the objective values. Nondominated solutions perform better than or equal to dominated solutions in all objectives and perform strictly better than dominated solutions in at least one objective. Pareto optimality in the context of the four objectives in this study requires a minimization of the depth to groundwater, water quality, and flooding objectives, and a maximization of the groundwater availability objective. The Pareto front for each spatial unit can contain from one to all six of the management alternatives, with a preferred alternative defined as any alternative that is nondominated. A visual comparison is evaluated by mapping all units in which each policy is among the nondominated set compared to the full set of spatial units at the cluster scale. Additionally, differences between policy preference at the three scales are explored by separating the spatial units into three groups for visual comparison: units with one nondominated policy, units with multiple nondominated policies, and units with all policies among the nondominated solutions.

Finally, to capture the scale-dependent nature of the relationship between policy preference and marginalization, this relationship is evaluated at each scale by comparing the social indicator information across each type of preferred policy. To do this, the percentile value is calculated for each socioeconomic indicator at the spatial unit level for each scale. We are interested in understanding the relationship for the most disadvantaged communities, therefore, the first quartile, which represents the lowest 25% of the spatial units in each socioeconomic indicator, are chosen for analysis. The number of units that prefer each policy are then divided by the total number of units in the first quartile to determine the percentage of disadvantaged communities that prefer a given policy according to each socioeconomic indicator. In this analysis, 100% indicates that all spatial

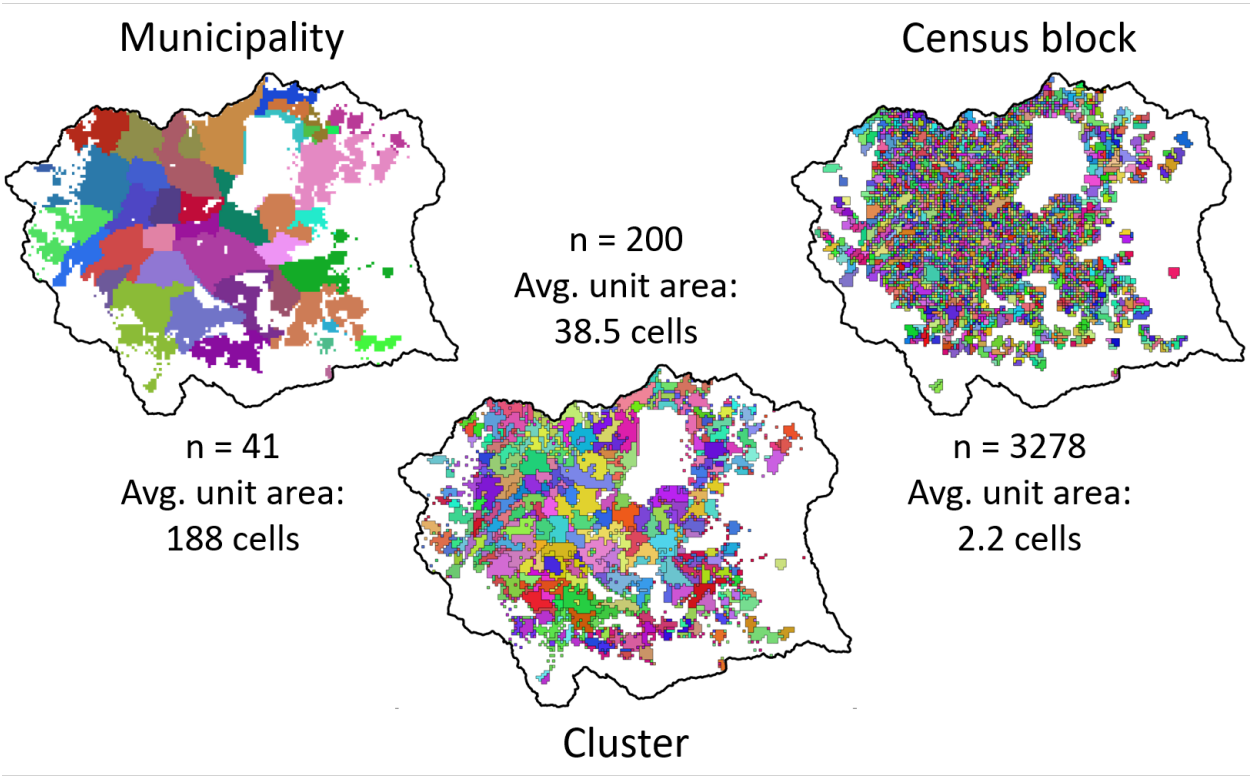

Figure 4.5: spatial units for the municipal, cluster, and census block scales.

units that are within the lowest quartile for that socioeconomic indicator contain the policy within the nondominated solution set, while 0% indicates that no spatial units within the lowest quartile contain the policy within the nondominated solution set. The percentage of the most marginalized units that prefer the policy is then compared to the percentage in the full set of spatial units to see if the policy is more or less preferred among the most marginalized set.

4.4 Results

4.4.1 Spatial resolution

Figure 4.5 shows the spatial distribution of each of the spatial unit scales. The municipal scale has the smallest number of units (41) and the highest average unit area, at 188 grid cells per unit. Next, the cluster scale has 200 units with an average area of 38.5 grid cells per unit. Finally, the

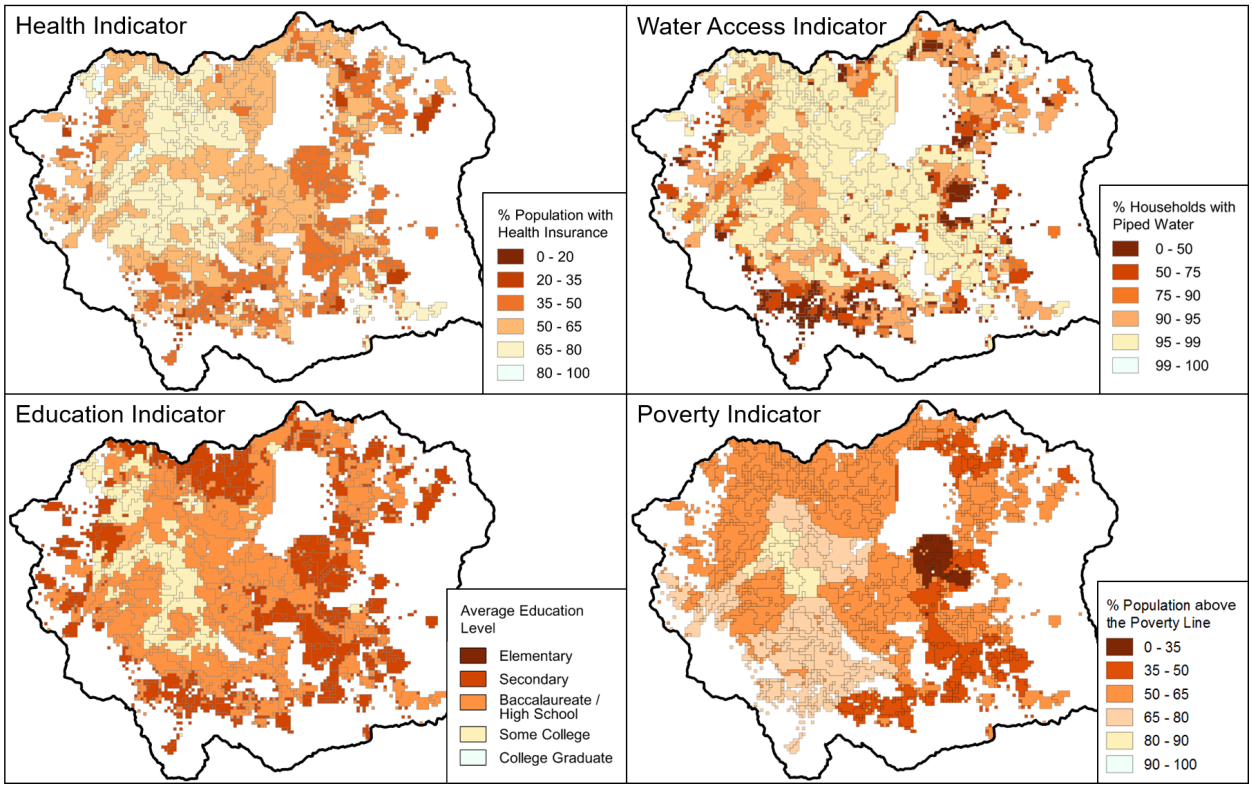

Figure 4.6: The four social indicators (education, health, water access, and poverty) determined at the cluster scale.

census block scale has 3,278 units with an average area of 2.3 cells. To assess how well the cluster scale represents the diversity of marginalization within the basin, Figure 4.6 shows each of the three indicators available at the census block scale (education, water access, and health) represented at the new intermediate "cluster" scale. Each of the indicators shown for the cluster scale capture the range of values from the census block scale, indicating a higher-resolution representation of the social indicators. Specifically, they represent many of the peri-urban areas along the external borders, particularly when compared to the municipal unit scale. Additionally, the larger scale provides the evaluation areas necessary for meaningful calculation of management objectives, based on the limitations of data availability for a regional scale groundwater model, as compared to the census block scale. Note that the poverty indicator is only available at the municipal scale and not at the census block scale, as shown in 4.14.

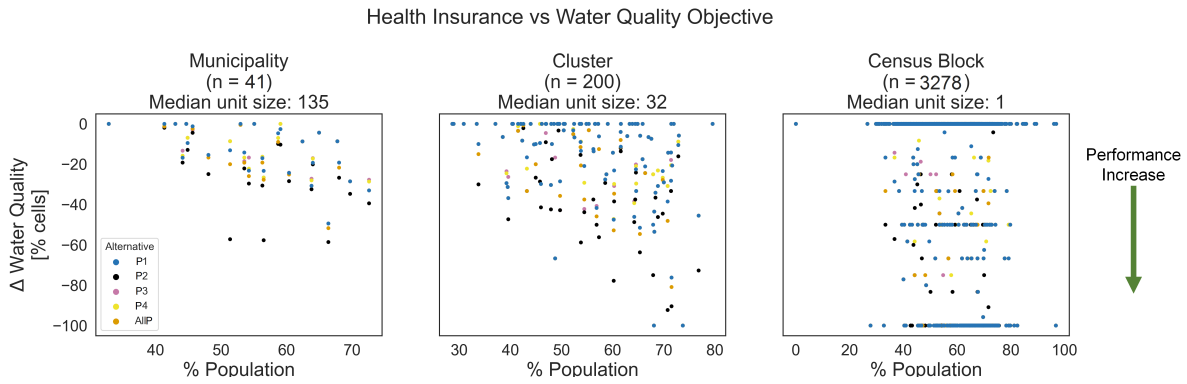

Figure 4.7: The water quality objective plotted against the health indicator (percentage of the population with public or private health insurance) for the five pumping policies compared to the historical management alternative.

4.4.2 Policy performance under management objectives

To examine the relationship between the social indicators and performance under the management objectives, the four management objectives were plotted against each of the four social indicators across the three spatial unit scales for each of the pumping policies except the historical alternative, which is used as the baseline. Three selected cases of these comparisons will be discussed here to point out the most pertinent findings.

Figure 4.7 shows the percentage of the population with health insurance plotted against the difference between the pumping policies (P1-4 and AllP) and the historical management alternative (H) in the water quality objective. Since the water quality objective measures the number of cells that do not meet the water quality risk threshold, a lower percentage indicates better performance. The municipal and cluster scales show that improved performance between the pumping policies and the historical alternative is concentrated in the higher portion of the social indicator range. This implies that the management alternatives create more improvements in this objective for communities that may be more prepared to handle the consequences of poor water quality—specifically, with health insurance, health effects from water quality issues are more easily treated. This relationship is more pronounced at the cluster scale for two reasons. First, the units have fewer cells on average and have the potential to experience a larger change in percentage of cells that do not meet the threshold. Second, the health indicator spans a wider range at this scale, which better highlights

the relationship. However, the census block scale in Figure 4.7 also reveals the limitations of a higher resolution. When the spatial unit contains only a few cells, particularly in an objective that is percentage based, there are artefacts within the visualization at the 100, 50, 30, and 0 percentage change levels.

Next, Figure 4.8 shows the percentage of the population below the income level for basic needs plotted against the difference between the pumping policies (P1-4 and AllP) and the historical management alternative (H) in the urban flooding objective. An increase in the percentage of cells that experience groundwater mounding indicates an increase in the risk of urban flooding. In all three scales, the units with increased flooding risk occur at lower percentages of the population that are above the poverty line. However, only a small fraction of units experience this change in the objective as a result of the management alternatives, with the vast majority showing no change. It is relevant to note that even if only a few communities experience increased flooding, those below the poverty line are much more likely to experience financial hardship as a result. This emphasizes the importance of mitigating such potential effects of regional-scale management.

Finally, Figure 4.9 shows the percentage of households with piped water plotted against the difference between the pumping policies (P1-4 and AllP) and the historical management alternative (H) in the groundwater availability objective. This comparison demonstrates the differences between the performance of the management alternatives across the spatial units. Specifically, the

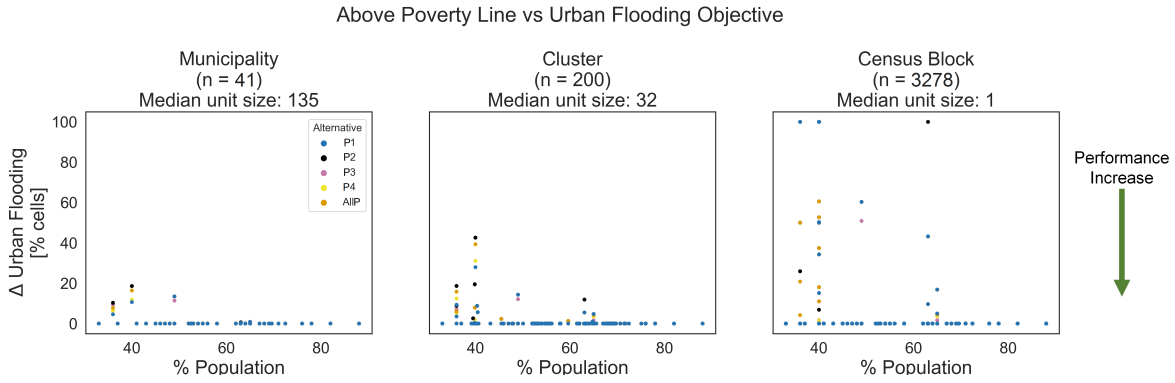

Figure 4.8: The urban flooding objective plotted against the poverty indicator (percentage of the population with an income above the level necessary to satisfy basic needs) for the five pumping policies compared to the historical management alternative as a baseline.

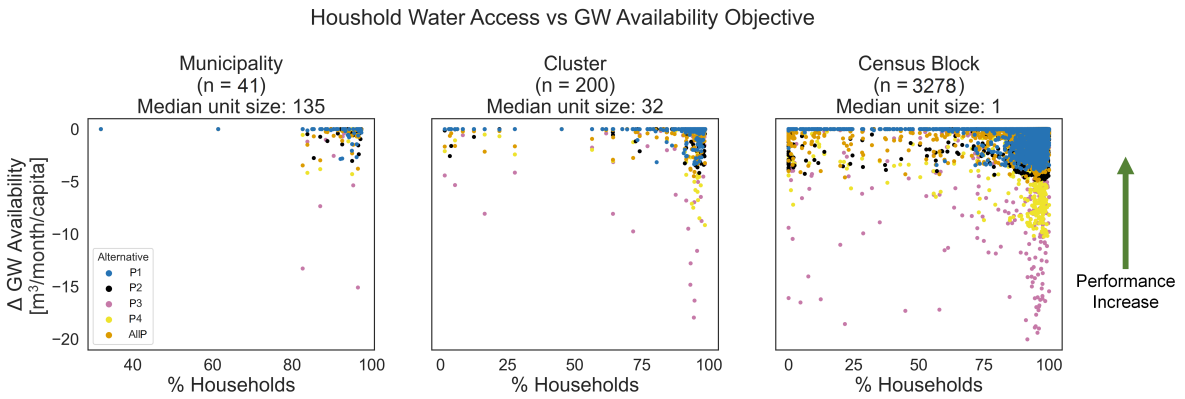

Figure 4.9: The groundwater availability objective plotted against the water access indicator (percentage of households with piped water inside the dwelling unit) for the five pumping policies compared to the historical management alternative as a baseline.

P3 policy (pink) and, to a lesser degree, the P4 policy, show an overall greater decrease in the groundwater availability objective than the other three policies (P1-2 and AllP), indicating poorer performance. This effect is slightly more pronounced among households with higher levels of water access. Additionally, as in Figure 4.7, the consequence of scale is apparent in this figure. While at the municipal scale it appears that the more marginalized populations (lower percentage of water access) are unaffected in this planning objective, at the other two scales, there are a number of spatial units at the lower range of the indicator that experience a decrease in the groundwater available within an accessible distance.

4.4.3 Alternative preference

To determine the differences in the performance of each alternative under the four objectives, multi-obejctive preference is measured using Pareto optimality for each individual spatial unit.

Table 4.2 Objective values measured at the regional scale.

		Policy					
Objective	**Units**	**H**	**P1**	**P2**	**P3**	**P4**	**AllP**
Groundwater availability	m^3 / month / capita	6.14	5.73	5.65	5.07	5.52	5.49
Depth to groundwater	m	30.05	21.73	19.00	21.39	21.69	20.95
Water quality risk	% cells	35.69	21.91	17.89	22.39	22.38	21.09
Urban flooding risk	% cells	0.60	0.99	1.31	1.26	1.05	1.18

Table 4.2 contains the four measured objectives at the regional scale. In the regional spatial extent, without considering the spatial units, all six policies, including the historical alternative, were among the nondominated set, indicating that there was no preference for any of the policies over the others based on the performance under the four objectives evaluated. At the regional scale, H performs the best in the groundwater availability and urban flooding risk objectives because these objectives favor more groundwater pumping. Alternatively, P2 performs the best in the depth to groundwater and water quality risk objectives, likely a result of the location and concentration of pumping in pumping group P2. Specifically, the water quality objective is calculated according to the extent of the lacustrine layer, making it sensitive to changes in pumping regimes in P1 and P2. Additionally, the pumping wells located in a vertical line on the eastern edge of P2 are located in an area with a predominance of basaltic formations that are prone to larger fluctuations in groundwater level, causing reductions in groundwater pumping in P2 to result in a greater rebound of groundwater levels than in the other pumping groups. The remaining policy alternatives lie along the Pareto front for the four objectives improving upon certain objectives at the cost of others between H and P2.

Table 4.3 shows the summary of the preferred management alternative(s) at the three subregional spatial unit scales. The number in each column in the top half of the table indicates the number of units where that alternative is nondominated. The only policy within the nondominated solution

Table 4.3 Summary of nondominated alternatives by spatial unit scale. H indicates the historical alternative, P1-4 are the pumping policies where groups 1-4 are eliminated from pumping, and AllP is the pumping policy where pumping is reduced in all four groups by 25%. The percentage of units that contain the policy within the nondominated solutions is shown out of the total number of spatial units by scale of aggregation.

Scale	Spatial Units	H	P1	P2	P3	P4	AllP
Municipality	41	27	25	41	20	27	27
Cluster	200	119	135	200	124	135	141
Census Block	3278	1960	2369	3278	2161	2388	2667
Municipality	100%	66%	61%	100%	49%	66%	66%
Cluster	100%	60%	68%	100%	62%	68%	71%
Census Block	100%	60%	72%	100%	66%	73%	81%

set in all spatial units at all scales is the P2 alternative. This is likely the result of two main reasons, first, the urban flooding risk objective in which P2 performs the worst is only relevant in low lying urban spatial units that represent a small number out of all units (see Figure 4.8). Second, and most importantly, as stated above, policy P2 performs well in both the water quality risk and depth to groundwater objectives, placing it along the Pareto optimal front. The next most preferred policies depend on the scale at which the alternatives are evaluated. At the municipal scale, H, P4, and AllP are the next most preferred among the same number of units, albeit not the among same specific units, while at the cluster and census block scale, AllP, then P4 and P1 are the next more preferred policies. In fact, almost all policies see an increase in the percentage of spatial units in which they are preferred with a higher resolution. This indicates that the effects of the policies on the management objectives are fairly localized because a higher resolution allows for those specific locations that deviate from the regional preference regime to be better characterized.

Figure 4.10 demonstrates how the localized preference is better represented at a higher resolution of spatial unit. At the municipal, cluster, and census block scales, the P2 policy is the only nondominated alternative in 17%, 16%, and 12% of spatial units, respectively; multiple, but not all, policies are nondominated in 44%, 34%, and 31% of spatial units, respectively; and, in agreement with the regional preference regime, all policies are nondominated in 39%, 51%, and 57% of spatial units, respectively. As the spatial unit goes from a lower resolution (larger area units) at the municipal scale (see also Figure 4.15) to a higher resolution at the cluster (see also Figure 4.11) and census block scale (see also Figure 4.16), the areas that are consistent with the regional scale, meaning all policies are nondominated, cover a larger portion of the basin. Similarly, the specific areas within the basin that deviate from the regional preference are better defined. In the southeastern part of the basin at the cluster and census block scale, a small area that prefers P2 to all other policies emerges that was not visible at the municipal scale. This type of analysis helps to identify communities that experience the greatest benefits or the greatest consequences from the implementation of aquifer management alternatives.

Using the cluster scale as an example, we can further explore policy preference across spatial

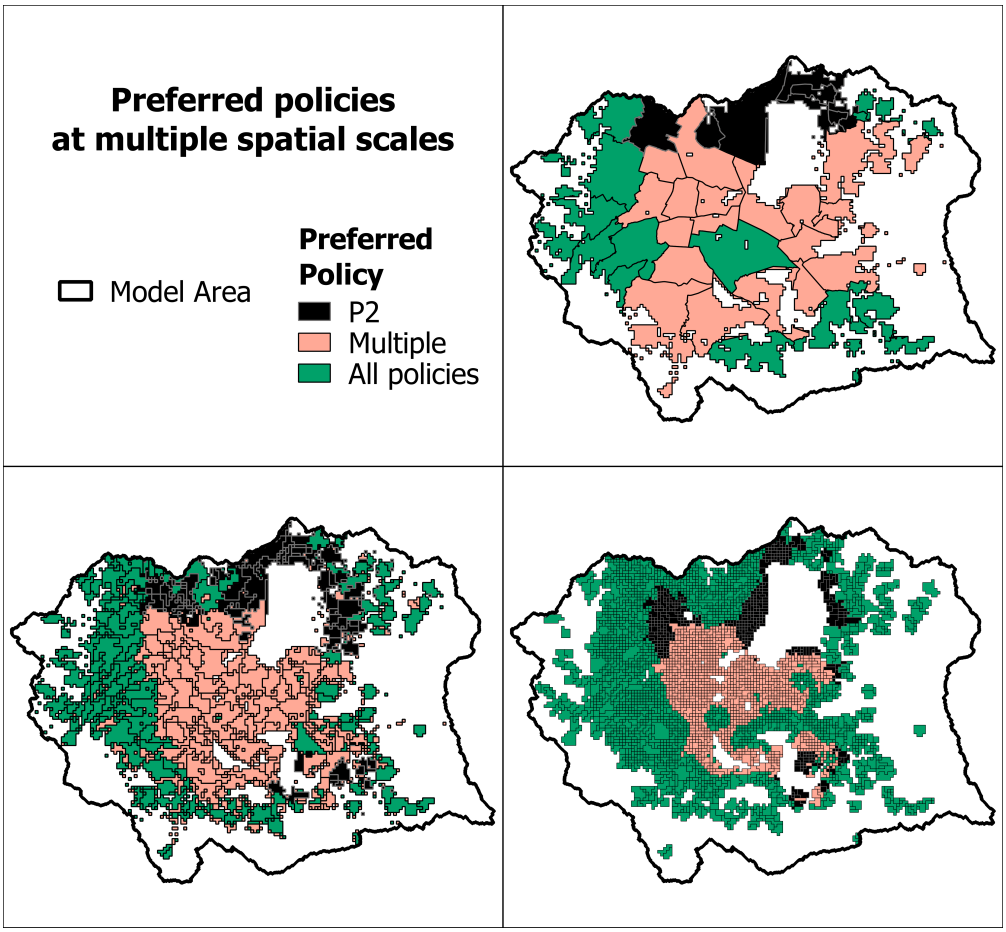

Figure 4.10: Nondominated solutions viewed at the three spatial unit scales as three categories: units with a single nondominated policy, units with multiple nondominated policy, or units with all policies among the nondominated solutions.

units visually in Figure 4.11. The first important trend is that the a higher proportion of spatial units surrounding the pumping groups 1, 3, and 4, do not include policies P1, P3, and P4, respectively, in the nondominated solution set. This is expected given that these policies reduce groundwater availability in the units directly surrounding the corresponding pumping groups. Another observation is that spatial units close to, but not directly surrounding, the pumping groups include the corresponding policies in the nondominated solution set. This is likely a result of the rising groundwater table that improves performance in the depth to groundwater and water quality risk objectives. Thus, the trade-off between the positive effects of localized reduction in groundwater pumping on the water table and the negative effects of such reductions on water access are visible in the extents of P1, P3, and P4. Finally, the historical alternative remains within the

nondominated set in low lying flood areas in the southeastern portion of the basin and in the areas along the borders of the basin where groundwater availability is lower.

To understand the relationship between policy preference and socioeconomic marginalization, the values on the left side of Table 4.4 show the portion of the lowest ranked spatial units that contain a policy in the nondominated set. Since all units at all scales include the P2 policy, this indicates that the performance of P2 produces an acceptable tradeoff among the management objectives for each spatial unit within the region, including the most marginalized. Alternatively, among the spatial units in the lowest quartile of socioeconomic indicator values in the region, P3 and P4 are only in the nondominated set for less than half of the units at the municipal scale in the water, health, and poverty indicators. The same is true for the P3 policy in the education indicator and the historical policy in the health and povery indicators, also at the municipal scales. This means that the most vulnerable communities deviate from the regional preference regime in accepting these policies.

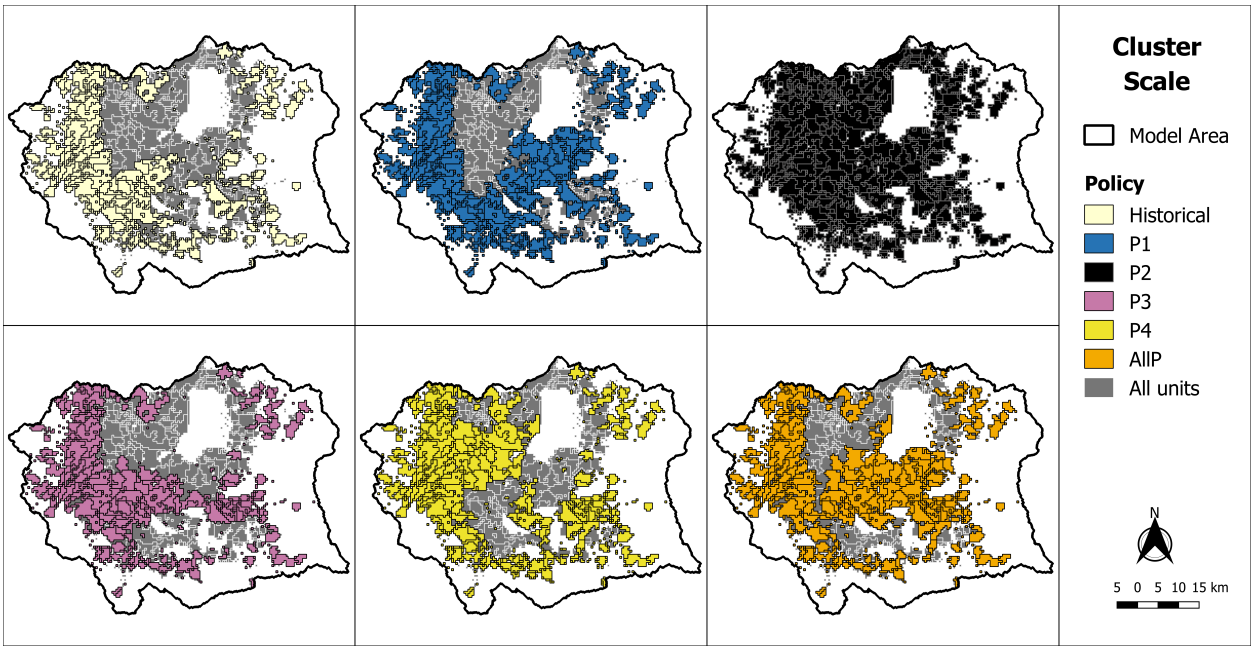

Figure 4.11: Spatial distribution of the spatial units at the cluster scale that contain the indicated policy within the nondominated solutions. A layer with all units is shown below the spatial extent of the preference of each of the policies for easier comparison. The P2 policy (black) is nondominated across all spatial units, leaving no gray visible.

Table 4.4 Percentage of spatial units within the first quartile of each social indicator that contain the policy within the nondominated solution set for that unit shown for each spatial unit scale. On the left hand side, a value of 100% indicates that all spatial units that are within the lowest quartile for that socioeconomic indicator contain the policy within the nondominated solution set. On the right hand side is the difference between the values on the left hand side and those in Table 4.3, representing the preference rates of the full set of spatial units. Positive values are shown in blue, while negative values are show in red.

			Percentage of lowest quartile units with policy preferred						Difference from percentage among all units					
Indicator	Scale	Count Q1	H	P1	P2	P3	P4	AllP	H	P1	P2	P3	P4	AllP
Education	Municipality	9	56%	78%	100%	44%	56%	78%	-10%	17%	0%	-4%	-10%	12%
	Cluster	50	68%	78%	100%	66%	70%	80%	9%	11%	0%	4%	2%	10%
	Census Block	822	72%	82%	100%	77%	77%	87%	12%	10%	0%	11%	4%	5%
Water	Municipality	10	50%	70%	100%	30%	30%	50%	-16%	9%	0%	-19%	-36%	-16%
	Cluster	50	70%	74%	100%	64%	70%	70%	11%	6%	0%	2%	2%	-1%
	Census Block	821	76%	78%	100%	81%	87%	88%	17%	6%	0%	15%	15%	6%
Health	Municipality	10	40%	60%	100%	30%	40%	60%	-26%	-1%	0%	-19%	-26%	-6%
	Cluster	50	62%	70%	100%	58%	68%	72%	3%	2%	0%	-4%	1%	2%
	Census Block	817	62%	78%	100%	68%	68%	82%	3%	6%	0%	2%	-5%	1%
Poverty	Municipality	9	44%	56%	100%	33%	44%	56%	-21%	-5%	0%	-15%	-21%	-10%
	Cluster	50	54%	68%	100%	46%	54%	68%	-5%	1%	0%	-16%	-14%	-2%
	Census Block	840	61%	72%	100%	65%	64%	76%	2%	0%	0%	0%	-8%	-6%

The above finding is further supported by comparing the percentage of units preferring the policy in Table 4.3 to the percentage of marginalized units preferring the policy in Table 4.4. A higher percentage in Table 4.3 indicates that the policy is preferred at a higher rate in the general population than among the lowest quartiles of each indicator, while a higher percentage in Table 4.4 indicates an increased preference among the more marginalized units. The right hand side of Table 4.4 shows the difference between the values in the left hand side of Table 4.4 and those in Table 4.3, with negative values indicating higher preference among the full set of spatial units shown in red and positive values indicating a higher preference among the lowest quartile set shown in in blue.

The difference is most pronounced in policy P1, which has a higher preference among units with the lowest average education and lowest access to water within the dwelling unit at all scales, and with the lowest access to healthcare at the census block scale and to a lesser degree at the cluster scale. This is because the main portion of the basin where P1 is not among the nondominated set is also the least marginalized among the socioeconomic indicators. The largest discrepancy is in the P4 policy at the municipal scale, which has a much larger preference basinwide than among the marginalized municipalities. This is likely because the southwestern municipalities that experience the greatest decrease in groundwater availability are among the most vulnerable in terms

of water access and the northeastern municipalities that do not experience a significant increase in groundwater table levels are similarly disadvantaged with respect to water access. These findings highlight the importance of considering social indicators when calculating policy preference as the majority of the households within more marginalized spatial units would be less likely than the general population to be able to mitigate the negative impacts from a given alternative that has poor performance in an objective such as the urban flooding, water quality, or groundwater availability objectives.

4.4.4 Limitations and future work

The social and hydrogeologic heterogeneity within the region create many layers of model complexity that can be explored in future work. Specific components of the methods proposed can be further evaluated to confirm the robustness of the method under endogenous uncertainties that are outside of the scope of this study. For example, only the municipal wells within the model area were considered for the pumping policies because there is an assumed difficulty in regulating private pumping based on the current lack of resources at the basin level to monitor, let alone restrict, existing concessions. However, future work could evaluate the effects of different levels of restriction and compliance of reduction of private pumping in addition to municipal pumping.

Additionally, the model used for this study could be expanded to include more complex flow and transport relationships. For example, instead of risk indicators for urban flooding and water quality, physics and chemical process based modeling could allow for the direct calculation of surface hydrology and groundwater quality. Similarly, the resolution of the groundwater model could be improved to allow for the use of the census block data directly in place of the smoothing methods implemented.

Another valuable improvement would be further customization of the spectral clustering method used to create the intermediate-scale units to better represent communities within the region. The clustering method used in this study gives equal weight to the x, y, and z coordinates and the four social indicators used to cluster the grid cells for the intermediate scale. This results in a

few clusters that are characterized by disconnected grid cells spread over a relatively large area, based on the social indicators selected. Similarly, there is no restriction based on population or total area, meaning that individual units evaluated may represent much larger or smaller portions of the population or region. A method that ensures spatial continuity within clusters would allow for a more accurate and physically meaningful calculation of the planning objectives, in addition to a higher probability of shared interests within the population represented, and an equal area or population across clusters would allow for better comparison between clusters.

Finally, the framework presented here could be leveraged for use in multi-objective optimization to find an optimal pumping scenario for the planning objectives evaluated. Specifically, given the use of Pareto optimality in the final analysis presented, this method could be adapted for use in an evolutionary algorithm that treats the pumping volume for each pumping well or groups of pumping wells as a decision variable, with a constraint on the maximum pumping volume for each well. Depending on the dimension of the problem, it may benefit from the use of a surrogate model to more efficiently represent the relationship between pumping decisions and the steady-state performance objectives.

4.5 Conclusion

In this study, we explore the effects of scale and spatial heterogeneity on aquifer management alternative performance. Reductions in pumping across four subgroups of municipal groundwater wells are evaluated under distinct groundwater planning objectives and compared against spatially distributed social indicators to understand how pumping policies interact with existing inequalities in health, education, income, and water access. Three scales of spatial units are evaluated to determine the robustness of the analysis to the method used to create the spatial units. Overall, this study results in three major findings.

First, alternatives selected based on management objectives at the regional scale can worsen the challenges faced by marginalized populations at the subregional scale. For example, pumping

policies show better performance in the water quality objective for some spatial units with higher rates of health insurance, and poorer performance under the urban flooding objective for some units with higher rates of poverty at both the municipal and cluster scales. Therefore, the comparison of system performance at different scales and across spatial units of varying degrees of marginalization shows the potential for adverse impacts to be concentrated in the communities least able to manage them.

Secondly, both large and small unit scales can cause difficulties in accurately representing the effects of management alternatives on marginalized populations. The relationships found between marginalization and pumping policy performance are more pronounced at the cluster scale than at the municipal scale, showing a benefit to higher-resolution spatial units. However, this effect is obscured at the even higher resolution of the census block scale because of the limits of the grid cell resolution of the numerical groundwater model. Thus, the evaluation of pumping policies can be highlighted or obscured depending on the average size of the spatial units used to calculate the planning objectives. The intermediate-scale resolution obtained by clustering according to social indicators is designed to address this challenge, and creates avenues for future work.

Finally, subregional aquifer management policy preference can diverge from regional preference based on the aggregation of management objectives, a method common in multi-objective groundwater decision-making analysis. While at the regional scale all policies are situated along the Pareto optimal front, at the subregional scale, policy P2 is shown to be the only nondominated solution in more than 10% of spatial units and multiple, but not all, policies are nondominated in over 30% of spatial units at all three scales. Similarly, preferences for policies differ according to the scale evaluated and the marginalization of the spatial units, as in the case of the P1 policy in the lowest quartile group of the education and water access indicators and the AllP policy in the education indicator. Thus, subregional evaluation of pumping policy performance under spatially determined planning objectives should be considered at multiple scales and among at-risk communities when making regional groundwater decisions.

This research underscores the importance of considering social indicators and appropriate

subregional scales when making groundwater management decisions. By incorporating social indicators into the analysis of policy performance and preference, socially informed spatial analysis can better serve the needs of decision-makers who must consider the equity implications of aquifer recharge planning. Future work should focus on further testing the robustness of the findings under variations of the pumping policies, planning objective calculations, and grid-cell clustering methods, and leveraging the findings for use in optimization studies. In conclusion, the findings from this study have meaningful implications for environmentally just and hydrogeologically effective regional aquifer management that can be extended to other complex urban groundwater systems.

4.6 Appendix

Figures 4.12, 4.13, and 4.14 are additional figures similar to Figure 4.3 in the main text, but with the health, education, and poverty indicators represented, respectively. Note that indicator data in Figure 4.14 is only available at the municipal level.

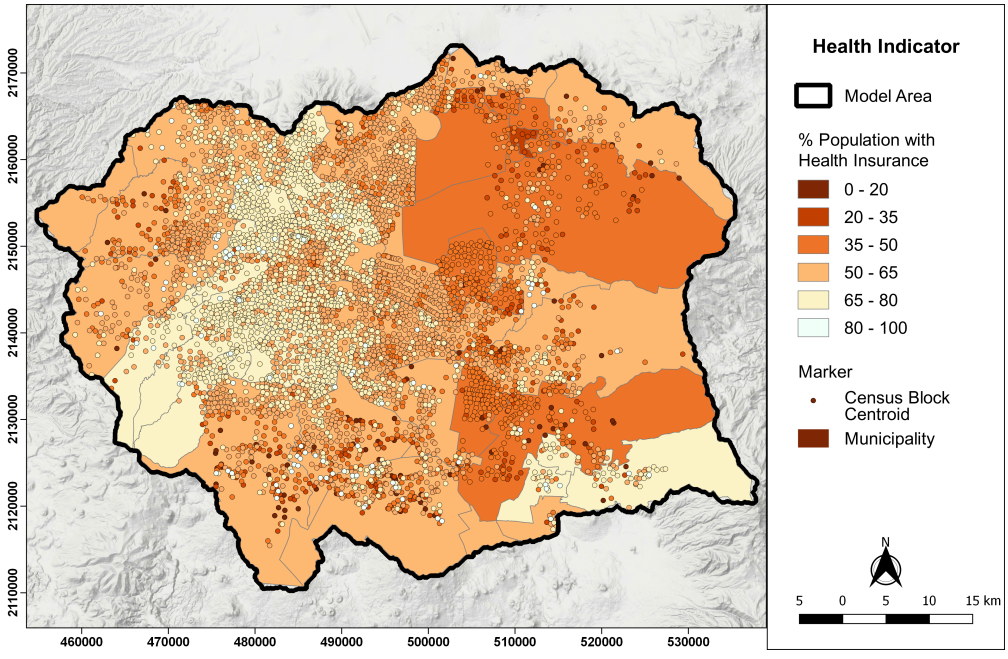

Figure 4.12: Health indicator for each of the municipality units (polygons) and centroids of the census block units (points). Darker indicates a smaller fraction of the population with private or public health insurance.

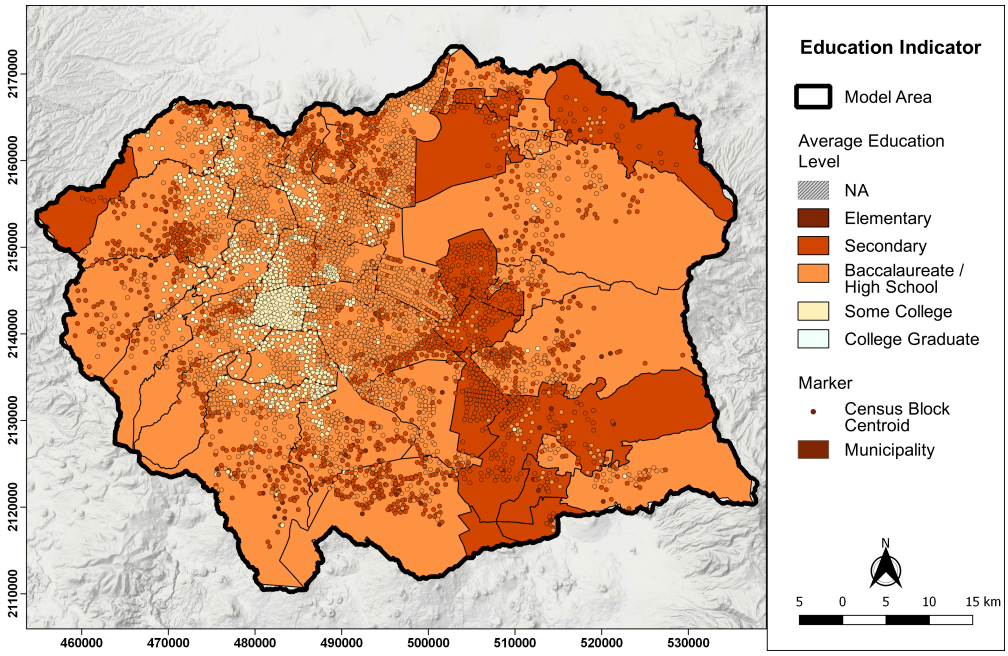

Figure 4.13: Education indicator for each of the municipality units (polygons) and centroids of the census block units (points). Darker indicates lower average education level among the adult population.

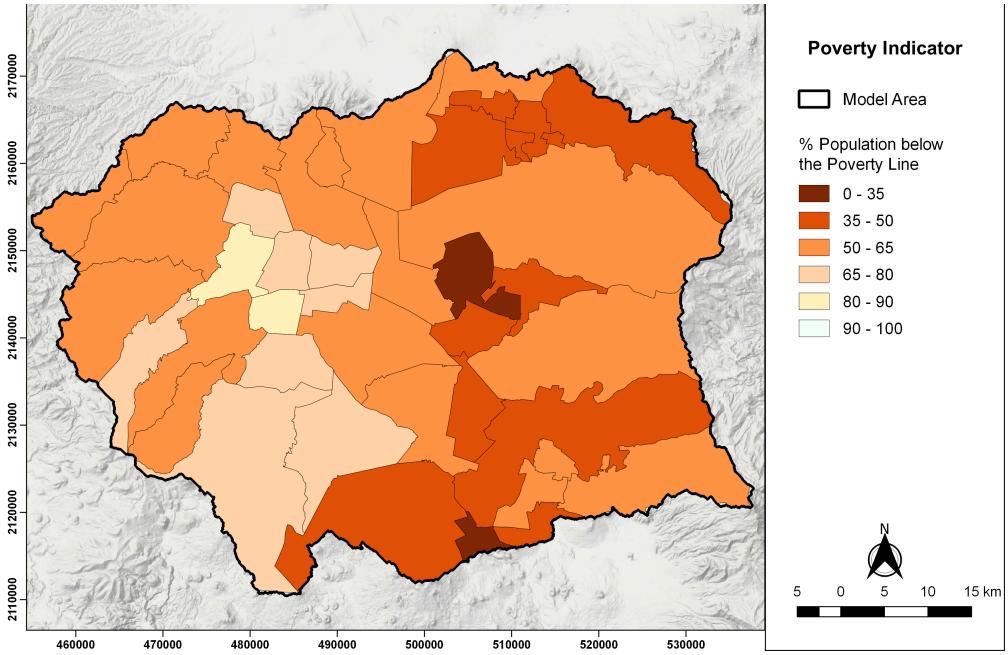

Figure 4.14: Poverty indicator for each of the municipality units. Darker indicates a lower fraction of the population above the threshold for basic needs used to indicate poverty in Mexico.

Figures 4.15 and 4.16 are additional figures similar to Figure 4.11 in the main text, but at the municipal and census block scales, respectively.

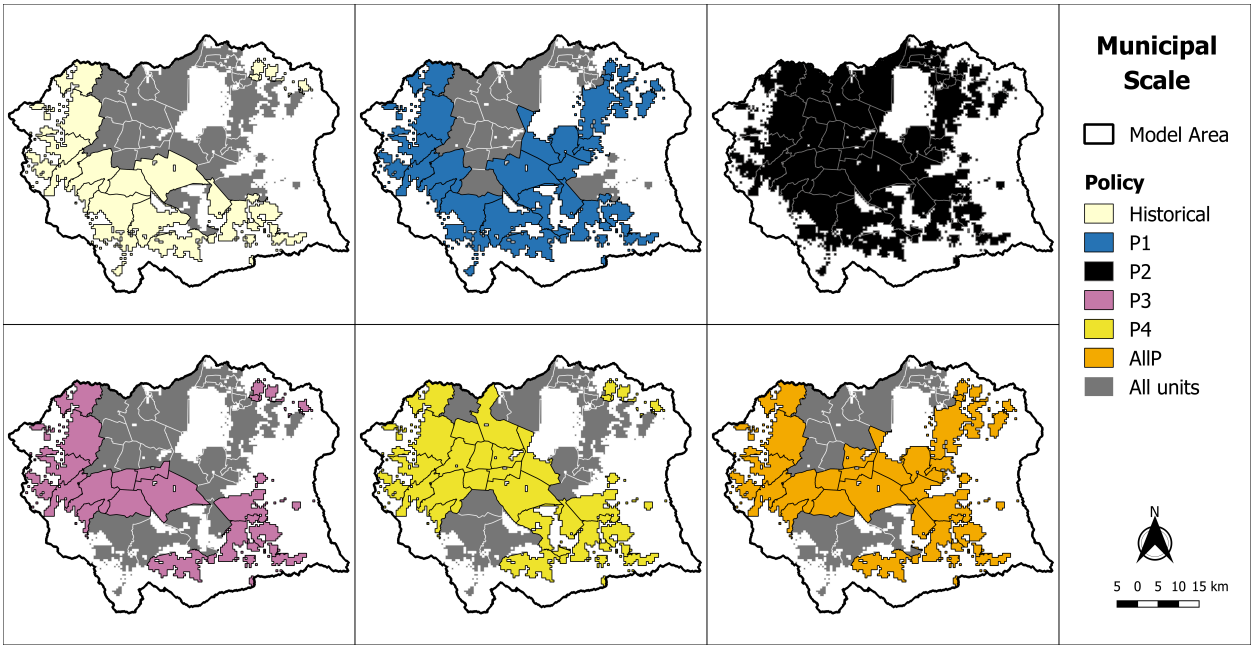

Figure 4.15: Spatial distribution of the spatial units at the census block scale that contain the indicated policy within the nondominated solutions. A layer with all units is shown below the spatial extent of the preference of each of the policies for easier comparison. The P2 policy (black) is nondominated across all spatial units, leaving no gray visible.

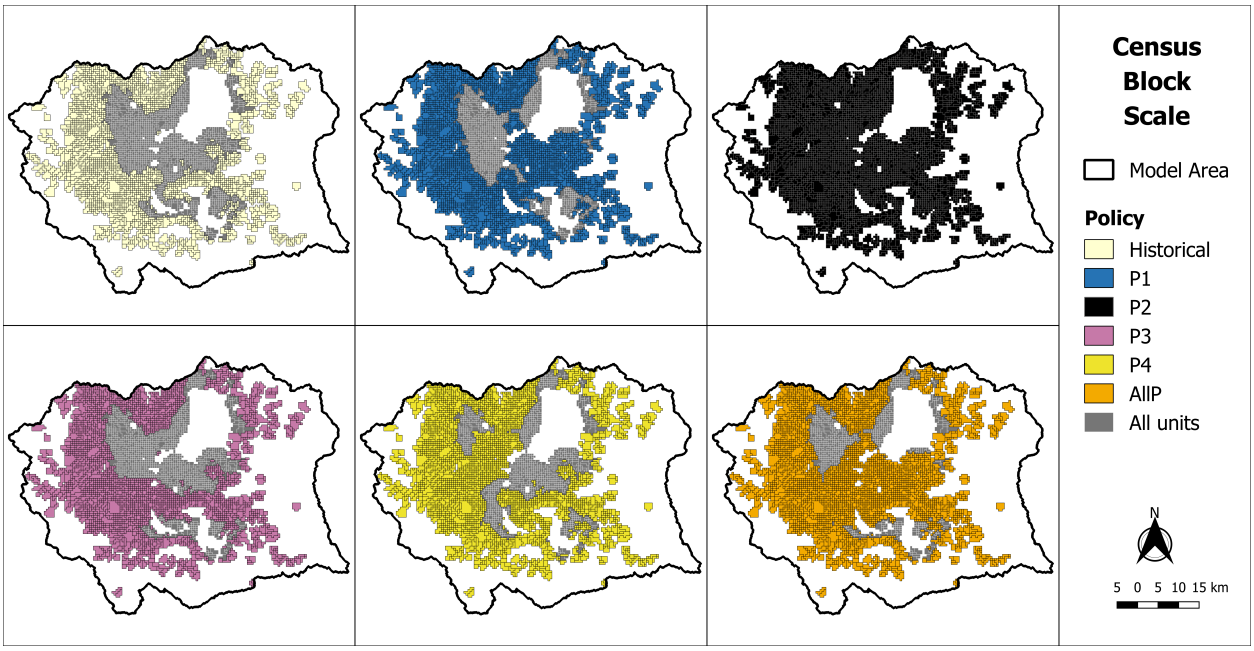

Figure 4.16: Spatial distribution of the spatial units at the municipal scale that contain the indicated policy within the nondominated solutions. A layer with all units is shown below the spatial extent of the preference of each of the policies for easier comparison. The P2 policy (black) is nondominated across all spatial units, leaving no gray visible.

4.7 Data Availability

The model with input datasets, observations, results, and postprocessing scripts are available in a GitHub repository at https://github.com/mrlmautner/Spatial-GW-Planning.

4.8 Acknowledgements

This work has been supported in part by the Ford Foundation Predoctoral Fellowship Program of the National Academies of Science, Engineering, and Medicine. Research trips to Mexico City were funded in part by the University of California Davis Henry A. Jastro Graduate Research Award. We thank the Organismo de Cuencas: Aguas del Valle de México (OCAVM) of the National Water Commission (CONAGUA) of Mexico and the Instituto de Geofísica of the Universidad Nacional Autónoma de México (UNAM) for their input on initial model development and pumping data.

Bibliography

Bach, P. M., Rauch, W., Mikkelsen, P. S., McCarthy, D. T. and Deletic, A. (2014). A critical review of integrated urban water modelling – Urban drainage and beyond, *Environmental Modelling & Software* **54**: 88–107.

Bakker, M., Post, V., Langevin, C. D., Hughes, J. D., White, J. T., Starn, J. J. and Fienen, M. N. (2016). Scripting MODFLOW Model Development Using Python and FloPy, *Groundwater* **54**(5): 733–739.

Barthel, R., Foster, S. and Villholth, K. G. (2017). Interdisciplinary and participatory approaches: the key to effective groundwater management, *Hydrogeology Journal* **25**(7): 1923–1926.

Cabello, V., Hernández-Mora, N., Serrat-Capdevila, A., del Moral, L. and Curley, E. F. (2012). Implications of spatially neutral groundwater management: Water use and sustainability in the Tucson basin, *in* F. Poupeau, H. Gupta, A. Serrat-Capdevila, M. A. Sans-Fuentes, S. Harris and L. G. Hayde (eds), *Water Bankruptcy in the Land of Plenty: Steps towards a transatlantic and transdisciplinary assessment of water scarcity in Southern Arizona*, CRC Press, Taylor & Francis Group.

Chahim, D. (2022). Governing beyond capacity, *American Ethnologist* **49**(1): 20–34.

Ciullo, A., Kwakkel, J. H., De Bruijn, K. M., Doorn, N. and Klijn, F. (2020). Efficient or Fair? Operationalizing Ethical Principles in Flood Risk Management: A Case Study on the Dutch-German Rhine, *Risk Analysis* **40**(9): 1844–1862.

de Graaf, I. E. M., Gleeson, T., (Rens) van Beek, L. P. H., Sutanudjaja, E. H. and Bierkens, M. F. P. (2019). Environmental flow limits to global groundwater pumping, *Nature* **574**(7776): 90–94.

Foster, S. (2020). Global Policy Overview of Groundwater in Urban Development—A Tale of 10 Cities!, *Water* **12**(2): 456.

Gain, A. K., Hossain, S., Benson, D., Di Baldassarre, G., Giupponi, C. and Huq, N. (2021). Social-ecological system approaches for water resources management, *International Journal of Sustainable Development & World Ecology* **28**(2): 109–124.

Galán-Breth, R. I. (2018). *Modelación matemática de nitratos en el agua subterránea en la región Sur de la Ciudad de México*, PhD thesis, Universidad Nacional Autónoma de México (UNAM).

Hadjimichael, A., Quinn, J., Wilson, E., Reed, P., Basdekas, L., Yates, D. and Garrison, M. (2020). Defining Robustness, Vulnerabilities, and Consequential Scenarios for Diverse Stakeholder Interests in Institutionally Complex River Basins, *Earth's Future* **8**(7): e2020EF001503.

Herrera-Zamarrón, G., Cardona-Benavides, A., González-Hita, L., Gutiérrez-Ojeda, C., Hernández-Calero, R., Hernández-García, G., Hernández-Laloth, N., López-Hernández, R. I., Martínez-Morales, M., Pita de la Paz, C., Sánchez-Díaz, L. F., Báez-Durán, J. A., Cruickshank-Villanueva, C. and Herrera-Revilla, I. (2005). Estudio para obtener la disponibilidad del acuífero de la Zona Metropolitana de la Ciudad de México, *Technical Report Contract No. 06-CD-03-1O-0272-1-06*, Secretaría del Medio Ambiente del Gobierno del Distrito Federal, Sistema de

Aguas de la Ciudad de México (SACM), and Instituto Mexicano de Tecnología del Agua (IMTA), Mexico City.

INEGI (2016). Principales resultados por AGEB y manzana urbana, *Technical report*, Instituto Nacional de Estadística y Geografía.
URL: https://www.inegi.org.mx/app/scitel/Default?ev=10

Jacobson, C. R. (2011). Identification and quantification of the hydrological impacts of imperviousness in urban catchments: A review, *Journal of Environmental Management* **92**(6): 1438–1448.

Jafino, B. A., Kwakkel, J. H., Klijn, F., Dung, N. V., van Delden, H., Haasnoot, M. and Sutanudjaja, E. H. (2021). Accounting for Multisectoral Dynamics in Supporting Equitable Adaptation Planning: A Case Study on the Rice Agriculture in the Vietnam Mekong Delta, *Earth's Future* **9**(5): 1–20.

Jafino, B. A., Kwakkel, J. H. and Taebi, B. (2021). Enabling assessment of distributive justice through models for climate change planning: A review of recent advances and a research agenda, *Wiley Interdisciplinary Reviews: Climate Change* **12**(4): 1–23.

Kind, J., Wouter Botzen, W. J. and Aerts, J. C. (2017). Accounting for risk aversion, income distribution and social welfare in cost-benefit analysis for flood risk management, *Wiley Interdisciplinary Reviews: Climate Change* **8**(2): 1–20.

La Rosa, D. and Pappalardo, V. (2020). Planning for spatial equity - A performance based approach for sustainable urban drainage systems, *Sustainable Cities and Society* **53**(October 2019): 101885.

Lerner, D. N. (2002). Identifying and quantifying urban recharge: a review, *Hydrogeology Journal* **10**(1): 143–152.

Lopez-Alvis, J. (2014). *Calibración de un modelo de flujo del Acuífero de la Zona Metropolitana de la Ciudad de México (AZMCM)*, PhD thesis, Universidad Nacional Autónoma de México (UNAM).

Mautner, M. R., Foglia, L., Herrera, G. S., Galán, R. and Herman, J. D. (2020). Urban growth and groundwater sustainability: Evaluating spatially distributed recharge alternatives in the Mexico City Metropolitan Area, *Journal of Hydrology* **586**(April): 124909.

Mautner, M. R. L., Foglia, L. and Herman, J. D. (2022). Coupled effects of observation and parameter uncertainty on urban groundwater infrastructure decisions, *Hydrology and Earth System Sciences* **26**(5): 1319–1340.

Megdal, S. B., Gerlak, A. K., Varady, R. G. and Huang, L. Y. (2015). Groundwater Governance in the United States: Common Priorities and Challenges, *Groundwater* **53**(5): 677–684.

Neal, M. J., Greco, F., Connell, D. and Conrad, J. (2016). The Social-Environmental Justice of Groundwater Governance, *in* A. J. Jakeman, O. Barreteau, R. J. Hunt, J.-D. Rinaudo and A. Ross (eds), *Integrated Groundwater Management*, Springer International Publishing, Cham, chapter 10, pp. 253–272.

OCAVM (2014). Programa Hídrico Regional 2014-2018: Region Administrativo Hidrológico XIII, Aguas del Valle de México, *Technical report*, Comisión Nacional del Agua, Tlalpan, Mexico, D.F.

Quinn, J. D., Reed, P. M., Giuliani, M. and Castelletti, A. (2017). Rival framings: A framework for discovering how problem formulation uncertainties shape risk management trade-offs in water resources systems, *Water Resources Research* **53**(8): 7208–7233.

Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Lo Piano, S., Iwanaga, T., Becker, W., Tarantola, S., Guillaume, J. H., Jakeman, J., Gupta, H., Melillo, N., Rabitti, G., Chabridon, V., Duan, Q., Sun, X., Smith, S., Sheikholeslami, R., Hosseini, N., Asadzadeh, M., Puy, A., Kucherenko, S. and Maier, H. R. (2021). The Future of Sensitivity Analysis: An essential discipline for systems modeling and policy support, *Environmental Modelling and Software* **137**(December 2020).

Re, V., Manzione, R., Abiye, T., Mukherji, A. and MacDonald, A. (2022). Groundwater for Sustainable Livelihoods and Equitable Growth, *Groundwater for Sustainable Livelihoods and Equitable Growth* .

Reed, P. M., Hadka, D. M., Herman, J. D., Kasprzyk, J. R. and Kollat, J. B. (2013). Evolutionary multiobjective optimization in water resources: The past, present, and future, *Advances in Water Resources* **51**: 438–456.

Rittel, H. W. and Webber, M. M. (1973). Dilemmas in a general theory of planning, *Policy Sciences* **4**(2): 155–169.

Schirmer, M., Leschik, S. and Musolff, A. (2013). Current research in urban hydrogeology – A review, *Advances in Water Resources* **51**: 280–291.

Tellman, B., Bausch, J. C., Eakin, H., Anderies, J. M., Mazari-Hiriart, M., Manuel-Navarrete, D. and Redman, C. L. (2018). Adaptive pathways and coupled infrastructure: Seven centuries of adaptation to water risk and the production of vulnerability in Mexico city, *Ecology and Society* **23**(1): art1.

Wada, Y., Bierkens, M. F. P., de Roo, A., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., Konar, M., Liu, J., Müller Schmied, H., Oki, T., Pokhrel, Y., Sivapalan, M., Troy, T. J., van Dijk, A. I. J. M., van Emmerik, T., Van Huijgevoort, M. H. J., Van Lanen, H. A. J., Vörösmarty, C. J., Wanders, N. and Wheater, H. (2017). Human–water interface in hydrological modelling: current status and future directions, *Hydrology and Earth System Sciences* **21**(8): 4169–4193.

Walker, W., Harremoës, P., Rotmans, J., van der Sluijs, J., van Asselt, M., Janssen, P. and Krayer von Krauss, M. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support, *Integrated Assessment* **4**(1): 5–17.

Wigle, J. (2020). Fast-track Redevelopment and Slow-track Regularization: The Uneven Geographies of Spatial Regulation in Mexico City, *Latin American Perspectives* **47**(6): 56–76.

Yang, G., Giuliani, M. and Castelletti, A. (2022). Operationalizing equity in multipurpose water systems, *Hydrology and Earth System Sciences* **Preprint**.

Zagonari, F. (2010). Sustainable, Just, Equal, and Optimal Groundwater Management Strategies to Cope with Climate Change: Insights from Brazil, *Water Resources Management* **24**(13): 3731–3756.

Zhou, Y. (2009). A critical review of groundwater budget myth, safe yield and sustainability, *Journal of Hydrology* **370**(1-4): 207–213.

Chapter 5

Conclusion

Spatially distributed groundwater modeling has opened the door to testing innumerable aquifer management options under a variety of uncertainties. However, the opportunity to simulate viable policy choices comes with the responsibility to try to understand the limitations of such modeling processes along with the social and environmental implications of the output from such efforts. This dissertation has proposed a suite of management objectives, sensitivity analyses, and disaggregation techniques that will allow decision-makers and researchers alike to evaluate and visualize groundwater supply planning policies in an equitable and transparent framework. The methods described herein add to the hydrogeologic modeling literature through the definition of regional spatially explicit urban groundwater objectives for use in multi-objective analysis, while also furthering the literature on robust multi-objective decision-making through the novel implementation of global sensitivity analysis through to the alternative ranking process, and of rival spatial framings.

Chapter 2 formulates a spatially distributed multi-objective aquifer management problem that defines novel, basin-relevant management objectives revealing conflicting regional needs. The development of planning objectives with distinct spatial and temporal extents to properly understand the regional effects of groundwater supply planning schemes has broad applicability in other rapidly urbanizing, groundwater dependent regions. Using the multi-objective problem developed in Chapter 2, Chapter 3 develops planning-driven uncertainty analysis that describes the sensitivity of decision alternative performance to plausible ranges of hydrogeologic model parameters and subsets of groundwater well observations. An important finding is that metrics that are generally

used to determine predictive ability of a groundwater model, such as the sum of squared weighted residuals, are not necessarily aligned with the decision-making applications for which models are applied. Additionally, results show that observational uncertainty may play a much larger role in the sensitivity of the objectives than the management alternatives themselves. This approach exemplifies how the propagation of multiple endogenous uncertainties throughout the modeling process can ultimately affect the outcomes of regional groundwater supply planning. Ultimately, Chapter 4 contributes to the literature in evaluating rival spatial framings of a groundwater management problem in the context of social marginalization according to multiple socioeconomic indicators. Methods evaluating the performance of spatially distributed urban groundwater pumping policies at multiple subregional scales yield a better understanding of how decision-making is affected by the spatial aggregation of system performance in the problem framing. The novel clustering method proposed to aggregate social indicators at a scale relevant for the calculation of groundwater management objectives provides the opportunity to more critically evaluate aquifer management policies against the possibility of exacerbating existing socioeconomic inequities. Additionally, subregional aquifer management policy preference is found to diverge from regional preferences based on the aggregation of management objectives and preferences for the policies differ according to the marginalization of the spatial units. Thus, subregional evaluation of pumping policy performance under spatially determined planning objectives should be considered at multiple scales and among at-risk communities when making regional groundwater decisions.

In conclusion, these three studies provide a diverse set of methods to assess much needed policy and infrastructural interventions in urban aquifers experiencing negative effects from excessive groundwater abstraction and modified hydrologic regimes. The modeling practices developed and implemented provide a clear guide from model conceptualization through sensitivity analysis of inputs to novel incorporation of social context in the modeling process. As a whole, this dissertation creates a robust modeling framework for the evaluation of the implementation of managed aquifer recharge alternatives that have the potential to improve the sustainability of future groundwater supply planning in the Mexico City Metropolitan Area and beyond.