Supporting Information for "Learning and flexibility for water supply infrastructure planning under groundwater resource uncertainty"

Sarah Fletcher $^{1,2}$, Kenneth Strzepek $^{2,3}$, Adnan Alsaati $^4$, Olivier de Weck $^5$

$^1$Civil and Environmental Engineering, Massachusetts Institute of Technology
$^2$Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology
$^3$Hydrology, Water Resources, & Environmental Fluid Mechanics, University of Colorado Boulder
$^4$Center for Complex Engineering Systems, King Abdulaziz City for Science and Technology
$^5$Institute for Data, Systems, and Society, Massachusetts Institute of Technology

Contents

1. Groundwater model
2. Sensitivity on Prior Choice
3. SDP Planning Model
4. Results
5. References

Corresponding author: Sarah Fletcher, sfletch@mit.edu
Groundwater model

We model the Minjur aquifer in the Riyadh area using the numerical groundwater modeling software MODFLOW (Harbaugh, 2005). The Minjur is a Triassic aquifer composed primarily of sandstone and shale, extending over 800 km across the central Arabian peninsula. There is an outcrop area approximately 100 km west of Riyadh. The aquifer is estimated to be 315 m thick in the Riyadh area. Recharge in the outcrop area is small, estimated between 3 and 25 mm per year (Al-Saleh, 1992; Williams & Al-Sagaby, 1982). A hydrostratigraphic cross section is shown in Figure 1.

The MODFLOW model is based on that of (Williams & Al-Sagaby, 1982), which is a study by the USGS that was used in the Saudi government’s 1984 Water Atlas (Ministry of Agriculture and Water, 1984). The USGS report is the most recent publicly available study of the Minjur, and it provides a model appropriate for demonstrating our planning framework. Following the USGS study, we use a 423 km by 288 km two-dimensional rectangular grid with 315 m thickness. The lower left corner of the grid is positioned at latitude 22.5° and longitude 45.79°, putting Riyadh and the major pumping well fields in the center of the study area. The Western side of the grid is bounded by an irregular no-flow boundary representing the outcrop of the aquifer. The rest of the grid is also bounded by no-flow boundaries; these boundaries are far enough away that they have negligible impact on the main study area in and surrounding Riyadh. The grid cells are 1 km x 1 km in the main study area and gradually increases to as large as 15 km. While the large grid spacing may underestimate drawdown in individual wells, it allows us to keep the computational cost of the model small and is appropriate for assessing long-term regional impacts on head (Anderson et al., 2002). We model the impact of 120 pumping wells in the Riyadh area. The locations and historical pumping rates of these wells are provided in (Williams & Al-Sagaby, 1982). We have 2010 data on 60 of these wells which account for over 80% of the withdrawals. The starting head is between 200 and 250 m.b.l.s.; the range reflects a substantial cone of depression around Riyadh. Recharge is assumed to be 5 mm/y in the outcrop area. During the model development, we tested recharge rates ranging from 3 to 25 mm/y and found this variation to have negligible impacts on the results. Therefore, variability and climatic uncertainty in recharge were not considered in the study. We use a transient model with a 30-year simulation horizon and weekly time steps. The study area, including groundwater pumping and desalination infrastructure, and a sample simulation of the groundwater model are shown in...
Figure 1: Hydrostratigraphic cross section of study area including Minjur aquifer. Reproduced and edited with permission from (Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) and Dornier Consulting International (DCI), 2016)

Figure 2. Note that because historical time series of head are not available, we have not validated the groundwater model on historical data; the use of the previously validated USGS model is intended to address this limitation.

Because our approach requires numerical calculation of many posterior distributions using the groundwater model, we develop a surrogate model that captures the key dynamics of the MODFLOW model with greater computational efficiency. Our application uses an ANN, trained on output data from the MODFLOW model with a variety of parameter inputs. This provides an instantaneous mapping from any realization of $\theta$ to the corresponding value of $h(t)$. This surrogate approach, rather than an analytical alternative like a response matrix, enables generalization to unconfined nonlinear groundwater models in future applications.

To train the ANN, latin hyper-cube sampling is used on the prior parameter distributions to generate 400 K and S samples as input to the MODFLOW model, yielding 400 output time series of hydraulic head; the set of these parameter combinations span the full possible range. Pre- and post-processing was completed using the Python
Figure 2: Schematic of MODFLOW grid with hydraulic head contours from one sample simulation. Pumping wells shown with red dots. Outcrop area colored black on left side of grid.
library FloPy (Bakker et al., 2016). One representative well, chosen from the largest well field in the center of Riyadh, is used to represent head decline in the planning model. We therefore train the ANN to predict the 30-year time series in that well using K and S as inputs; this prediction does incorporate the impact of all 120 pumping wells in the MODFLOW model. The MODFLOW simulations are run with 100 time steps per year and recorded at a single grid cell representing the Buwayb well field. This yields a total of 4,000 data points for hydraulic head with varying inputs for $K$, $S$, and $t$. This dataset is randomly split into a 70-15-15 train-validation-test partition. A feedforward ANN is trained using MATLAB’s scaled conjugate gradient backpropagation algorithm. Many different network architectures varying the transfer function, number of hidden layers, and number of neurons are tested. We select the network with the lowest root mean square error (RMSE) on the test partition. This architecture has 2 hidden layers, with 6 neurons each, and a sigmoid transfer function. Because of the SDP formulation which imposes a drawdown-limit of 50 meters, only observations above the drawdown limit are included in calculating the RMSE; this allows us to choose the model that performs best in the range it will be used in the SDP. The RMSE calculated using this approach is 1.02 m, indicating excellent performance for our long-term regional planning application.
Figure 3: Error histogram comparing drawdown estimates from MODFLOW to corresponding estimates from the ANN.
Sensitivity on Prior Choice

Figure 4: Alternate prior distribution for K: We assess the impact of a reduced-variance prior in K with the same median as the original but half the variance. These distributions are shown here.
Figure 5: Sensitivity to K prior: The impact of the alternate reduced variance prior is shown on the posterior for log K (left column) and the posterior for drawdown (right column). Each row illustrates a different posterior using a unique head observation. The impact of the reduced variance prior is negligible, demonstrating that the posterior is more sensitive to the data given the wide range of uncertainty and therefore uninformative prior.
SDP planning model

Details of the formulation of the SDP Bellman equation from equation (3) in the main text are provided below:

\[
S = \{h_t, x_t\} \\
A = \{p_t, e_t(x_t)\} \\
C_t = P(h) * p_t + E * e_t + O * x_t * e_t + S * \text{max}(0, D - x_t)
\] (1)

where

- \( t \) = time step in the model in years between 1 and 30
- \( h_t \) = hydraulic head (meters) in representative well at time \( t \)
- \( x_t \) = desalination capacity (MCM/y) at time \( t \). \( x_t \) starts at 0 and changes to 108 two years after the expansion option is exercised if at all.
- \( p_t \in\{0, 1\} \) where 0 and 1 indicate respectively that pumping is off or on at time \( t \)
- \( e_t(x_t) \in\{0, 1\} \) where 0 and 1 indicate respectively that infrastructure option \( i \) is not expanded or expanded. If \( x_t = 108 \), indicating that the desalination plant has already been added, then \( e_t \) is constrained to equal 0.
- \( P(h) \) = Marginal cost of pumped groundwater at hydraulic head \( h \) (USD/MCM)
- \( E \) = Infrastructure expansion cost (USD)
- \( O \) = Desalination marginal cost (USD/MCM)
- \( S \) = Shortage cost (USD/MCM) reflecting damages that are incurred if the depth limit is reached before adding desalination infrastructure, and
- \( D \) = Demand (MCM/y), assumed to be 108 equivalent to estimates for current withdrawal rates from the Minjur.

The transition probabilities for the individual state variables are assumed to be independent. The transition probabilities for hydraulic head are determined by the groundwater model and Bayes’ theorem as described in equations (1) and (2). The transition for the capacity state is deterministic, determined by the expansion action \( e_t \). We assume the desalination capacity is available two years after the decision to expand. This reflects the flexible planning process described previously in which upfront planning enables timely capacity additions.
Figure 6: Cost assumptions for groundwater SDP formulation. Pumping and brackish treatment costs are incurred for water supplied from groundwater; pipeline pumping and desalination opex are incurred for water supplied from desalination. Desalination capex is incurred when a new desalination plant is brought online.
Figure 7 illustrates the cost assumptions used in the formulation. The marginal cost of pumped groundwater $P$ is the sum of two components: pumping costs and pretreatment costs because of the brackish quality of the water. Pumping costs were estimated using the cost of energy needed to raise water the height of drawdown in the well plus head losses due to friction estimated using the Darcy-Weisbach equation (Hwang & Houghtalen, 1996). Pumping cost range between $0.40/m^3$ at the assumed starting depth of 337 m.a.s.l and $0.47/m^3$ when the 50 m depth limit is reached. Brackish treatment costs were assumed to vary between $0.3/m^3$ and $0.35/m^3$ for the starting depth and maximum depth respectively. The marginal cost of desalinated water is estimated as the sum of pumping costs through an existing pipeline from the desalination plant on the Arabian Gulf to Riyadh plus desalination opex. Pipeline pumping costs are assumed to be $1.35/m^3$ and were estimated as the cost of energy needed to raise water the elevation difference between the desalination plant at sea level and Riyadh at 612 m.a.s.l. plus head losses due to friction again estimated using the Darcy-Weisbach equation (Hwang & Houghtalen, 1996); changes in elevation over the pipeline’s path were not considered. RO desalination opex and capex were estimated to be $0.48/m^3$ and M$304 based on a 108 MCM/y capacity plant using the Cost Estimator tool from Global Water Intelligence’s ”Desal Data” database (Global Water Intelligence, n.d.). This size was chosen to be equivalent to estimated withdrawals from the Minjur aquifer.
Results

Figure 7: CDFs of a) capital costs b) total operating costs over 30 years and c) total shortage costs over 30-years by infrastructure alternative under the base case. Operating costs include desalination production and groundwater pumping costs. Top row: base case; lower rows: sensitivity analysis. Horizontal axis truncated for visual clarity.
References

Al-Saleh, M. A. (1992). Declining groundwater level of the Minjur Aquifer, Tebrak area, Saudi Arabia. *The Geographical Journal, 158*(2), 215–222.

Anderson, M. P., Woessner, W. W., & Hunt, R. J. (2002). *Applied Groundwater Modeling*. London: Academic Press.

Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) and Dornier Consulting International (DCI). (2016). *Detailed Water Resources Studies of Khuff Jil Jh Minjur Dhurma and Overlying Aquifers* (Tech. Rep.). Ministry of Environment Water & Agriculture.

Global Water Intelligence. (n.d.). *Desal Data Cost Estimator*. https://www.desaldata.com/cost estimator.

Harbaugh, A. W. (2005). *MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model — the Ground-Water Flow Process*. Reston, Virginia. Retrieved from https://pubs.usgs.gov/tm/2005/tm6A16/PDF/TM6A16.pdf

Hwang, N. H. C., & Houghtalen, R. J. (1996). *Fundamentals of Hydraulic Engineering Systems* (Prentice-Hall, Ed.). Upper Saddle River, New Jersey.

Ministry of Agriculture and Water. (1984). *Water Atlas of Saudi Arabia* (Tech. Rep.). Riyadh, Saudi Arabia: Ministry of Agriculture and Water in cooperation with the Saudi Arabian-United States Joint Commission on Economic Cooperation.

Williams, J. F., & Al-Sagaby, I. (1982). *Simulated changes in water level in the minjur aquifer, Riyadh area, Saudi Arabia* (Tech. Rep.). Saudi Arabian Ministry of Agriculture and Water; United States Geological Survey. Retrieved from https://pubs.er.usgs.gov/publication/ofr82371