Project Title: **CBET-EPSRC Efficient Surrogate Modeling for Sustainable Management of Complex Seawater Intrusion-Impacted Aquifers**

Investigators:

Alex S Mayer, University of Texas at El Paso, United States (Project Grant Period: 01/01/2020 - 31/8/2022)

Domenico Bau, University of Sheffield, United Kingdom (Project Grant Period: 01/04/2020 - 5/31/2024)

**Major goals of the project**

The overarching goal of the proposed research is the sustainable management of water resources in coastal regions with diverse geological, hydro-technical and governance settings. Pressures on water resources in coastal regions are already great and are expected to intensify due to increasing populations, standards of living and impacts from climate change and sea level rise (SLR). We focus on coastal areas where aquifer over-drafting has caused seawater intrusion (SWI), thus

deteriorating groundwater quality, and where SLR is expected to further reduce availability of fresh groundwater. Solutions to these problems involve combinations of more efficient pumping schemes, demand reduction, and technological interventions such as desalination. However, determining optimal solutions for these problems poses extreme computational demands. This project will greatly advance the development and application of simulation-optimization (SO) by developing computationally efficient, robust, and accurate surrogate models for coastal groundwater systems.

**Accomplishments under these goals and objectives for 2022**

1. **Modified Henry’s Problem**. As one of the classic seawater intrusion problem set-up, Henry’s problem (Henry,1964), has been widely benchmarked and modified for further analysis (e.g., Javadi et al., 2012; other references). We adopted the original model set up from Henry (1964) in our investigation by enlarging the model domain to a quasi-2D domain with a dimension of 200-m long by 100-m deep by 1-m wide (into the page) and adding a single pumping well located in the central part of the aquifer (Figure 1).

We developed a variable density groundwater model using SEAWAT that couples fluid flow and solute transport to conduct our analysis. Landward recharge is represented by a constant freshwater influx from the left-hand boundary. The coastal boundary is represented by a constant seawater head boundary for groundwater flow and point source (35 g/L) boundary for the solute transport. Top and bottom boundaries are no-flux boundaries for both flow and transport.

We generated 2D heterogeneous hydraulic conductivity fields that are spatially correlated, using a spherical variogram with a fixed vertical correlation length of 10 m. We varied the horizontal correlation length among five different values (10, 40, 70, 100, 200 m) to examine the impacts of aquifer connectivity to the ocean. For the log-normal reference distribution, we kept the mean constant, and varied the variance among 0.5,1, and 2. We conduct a series of 7200 simulations across 12 pumping rates, 3 K-field variances, 5 K-field horizontal correlation scales, with 40 realizations for each of the afore mentioned variables. All realizations have the same geometric mean K equal to the value of K used in the homogeneous simulations.

Please see the Significant results section for results.

|  |
| --- |
| A screenshot of a computer  Description automatically generated with low confidence |
| **Figure 1.** Schematic of the modified Henry’s problem |

The next step in this work will be to develop Gaussian Process (GP) surrogate models (Rasmussen and Williams, 2006) to substitute for the SEAWAT simulations. This work will involve incorporating uncertainty in heterogenous K distributions and analysing trade-offs between the accuracy of the GP models and the number of SEAWAT simulations used to train the GP model (computational effort).

1. **Sensitivity of optimization of groundwater supply in island aquifers to constraint formulation.** This study focuses on a simplified conceptualization of the San Salvador Island aquifer (see Figure 2a), which is used to construct a steady-state variable density flow model relying on USGS’s SEAWAT code (Langevin et al., 2008) (Figure 2b). In this work, we formulate the groundwater management in island aquifers into an optimization problem where the main objective is the minimization of the operating cost associated with the groundwater abstraction and desalination treatment. This objective addresses specifically the increase in costs associated with increasing groundwater demand. Pumping patterns are characterized by the distance of the pumping system from the shoreline (WL), the abstraction screen depth (D) and overall pumping rate (Q), which constitute the decision variables of the optimization problem. We investigate the impact of constraint formulations and pumping strategies on the operation cost, the water table drawdown over the pumping location, and SWI, as quantified by either the reduction in freshwater volume or the salt mass increase in the aquifer.

Please see the Significant results section for results.

|  |  |
| --- | --- |
|  |  |
| **Figure 2.** (a) Location map of San Salvador Island (Moore, 2009). (b) SEAWAT numerical model with boundary conditions and location of pumping system. | |

The next step in this work will be to incorporate the findings regarding sensitivity to constraint formulations into the GP models we've developed for the island-pumping optimization problem.

1. **Application of the Proper Orthogonal Decomposition (POD) method.** In the POD method (Siade et al., 2010), the original model is replaced by a surrogate that is a summation of the R number of time functions multiplied by space orthogonal basis functions as:

y(space, time) = Σar(time) x fr(space) [r=1 to R] (1)

where y is the model output, which depends on space and time, ar is the rth

time function and fr is the rth space function, which depend on time and space, respectively. The more R, the more accurate is the surrogate model. However, there is an optimal value for R, which needs to be found by sensitivity analysis.

As depicted in Figure 3, our application of the POD method is done in the following steps:

1. Create snapshots from MODFLOW, which means producing some results from the original model in different time steps.
2. Use SVD (singular value decomposition) and calculate U matrix elements, representing the basis functions fr.
3. Replace equation (1) in the original differential equation(s) and create a new system reduced order set of equations.
4. Select a value for R.
5. For each time step, solve the new system of differential equations to find the time functions ar.
6. Simple 1D (see Figure 4) and 2D examples are used as proof of concept. The 1D model is solved for 100 days by two different approaches: (a) numerical solving of the original differential equation using MATLAB and (b) applying POD (R=6, using backward-based finite difference method).

The 3D example is a confined cubic aquifer (101 m ×101 m × 50m), and the hydraulic conductivity is 1m/day. The boundary head is constant and equal to 10 m around the domain. There is a pumping well at the model’s center, which pumps 50 m3/day.

Please see the Significant results section for results.

|  |
| --- |
| Diagram  Description automatically generated |
| **Figure 3.** POD Framework Flow Chart |

|  |
| --- |
|  |
| **Figure 4.** 1D aquifer setting adopted for testing the POD scheme. |

The next steps in this work will be to derive the POD system of equations for coupled, variable density flow and transport and to apply the resulting POD model to the Santa Barbara aquifer.

1. **Optimizing Salinity Control Structure locations for the Biscayne Aquifer.** The fully calibrated SEAWAT model of Biscayne Aquifer in the southeast Florida has been developed by our USGS collaborators (Hughes et al., 2016). The active flow model area (denoted by the yellow shades in Figure 5) covers approximately 450 square miles (mi2) in the eastern part of Broward County and the north-eastern part of Miami-Dade County. It is bounded by the Everglades and several other water conservation areas to the west, and by the Atlantic Ocean to the east. The northern and southern boundary of the model are bounded by the Hillsboro canal basin and the C-8 canal basin, respectively. A subset of the domain (190 mi2) simulating active solute transport extends from the Atlantic Ocean to the approximate inland location (4.5 mi) of the seawater/freshwater interface in the Broward County. The prominent management features or decision variables (D.V.) of the Biscayne Aquifer are: (a) stages of the extensive canal network; (b) locations of the salinity control structures; (c) pumping rates of the extraction wells; (d) injection rates of the recharge wells.

We conducted a series of numerical experiments varying the position of the salinity control structure (G54, see Figure 5) at base pumping rate and high projected sea level rise, but we gradually extended the structure further coastward with a series of incremental movement along the canal. The sea level of all scenarios is at the high-level condition (2.03 ft by 2060).

Please see the Significant results section for results.

|  |
| --- |
|  |
| **Figure 5.** Map view of the active model domain for flow and transport in the Biscayne aquifer system. Surface water basins, canals, and salinity control structures are also denoted in the figure. |

Please see the Significant results section for results.

The next steps in this work will be to develop a generic framework for the optimization problem of locating salinity control structures, apply to the Biscayne aquifer, and develop GP surrogate models to replace the SEAWAT based optimization framework.

**Significant Results**

1. **Modified Henry’s Problem**

The results show almost identical relationships between pumping well concentrations (C) and pumping rates (Q\*) for most of the values of variances and horizontal correlation lengths. Figure 6 shows the response curves for C vs. Q\* for the highest K variance and the five values of the horizontal correlation scale (lx) for 10 out of the 40 total realizations. The pumping rates, Q\*, in Figure 6 are normalized to the freshwater recharge from the left-hand boundary. In these cases, well concentration is controlled primarily by the pumping rate relative to freshwater recharge rate and is much less influenced by heterogeneity. As expected, C is close to zero for normalized pumping rates Q\*<1 and increases sharply as the pumping captures more of the freshwater recharge. For Q\*>3, C continues to increase but at a much lower rate, as the pumping has captured almost all the freshwater recharge.

The response curves for the heterogeneous K-fields are nearly identical to curves obtained for homogeneous simulations with the same geometric mean K. However, for the highest horizontal correlation length (lx = 200m), pumping well concentrations are substantially higher for realizations #0 and #3 for pumping rates around Q\* = 1 to2. Figure 6 shows that the total dissolved salt mass (M) is sensitive to the individual realization for all values of the horizontal correlation scale (lx). Most of the M vs. Q\*response curves tend to cluster around a mean and the curves obtained for homogeneous simulations with the same geometric mean K. However, the results in Figure 6 show that for the largest value of the correlation scale (lx = 200m), realizations #0 and #3 again have anomalous behaviour. These results can be explained by looking at the K distributions (see left-hand side of the panel in Figure 7) in the domain and corresponding steady state concentration distributions (see Figure 7) in the domain for four values of Q\* and for the 10 realizations indicated in Figure 6. The K distributions show that for high variances and high horizontal correlation lengths, laterally continuous zones of contrasting hydraulic conductivities appear. The concentration distributions show that for low pumping rates (Q\*=0, 0.5), the vertical position of these continuous zones controls the configuration of the seawater intrusion. For example, for realization #0, the seawater intrudes along a low K (dark blue) layer, even for Q\*=0. This result is explained by the fact that the low K layer prevents the freshwater recharge from the left-hand boundary from “pushing” against the intruding seawater. This phenomenon explains the high values of M for low pumping rates in Figure 6, for higher values of the horizontal scale. However, the intrusion of the seawater responds in an opposite manner as the pumping rate increases (Q\* = 1, 1.5, 2). High K layers (dark red) near the coast (right-hand boundary) aligned with the vertical position of the coast allow the seawater to intrude further towards the well. This phenomenon explains the higher values of C in Figure 6 for the high correlation scale (lx = 200m) and realizations #0 and #3. In general, the anomalies observed in Figure 6 are related to: (a) the distance between the pumping well and the layers; (b) whether the layer has a high or low K, relative to the surrounding area; (c) if the pumping well is directly connected to the saltwater front via the high-K layer; (d) if there is a low-K continuous sealing layer separating the lower part of the saltwater front from the pumping well.

|  |  |  |
| --- | --- | --- |
| Graphical user interface  Description automatically generated | lx=  10 m  lx=  40 m  lx=  70 m  lx=  100 m  lx=  200 m |  |
| **Figure 6.** Response Curves for Well Concentration and Total Salt Mass.. | | |

|  |
| --- |
|  |
| **Figure 7.** K realizations and corresponding concentration maps. |

1. **Sensitivity of optimization of groundwater supply in island aquifers to constraint formulations**

The results of this analysis are summarized in the subpanels of Figure 8. Figure 8a shows that, for larger WLand lower Q, the operation cost fOC tends to decrease with the pumping depth Duntil this remains within 50-100 m. This happens because groundwater abstraction occurs within the original freshwater lens, so that little or no desalination is required, and the predominant cost component is due to groundwater pumping. Instead, for lower WLand larger Q, increases along the depth D, since the dominant cost component is due to the desalination treatment. In general, forD larger than 50-100 m, the total cost increases sharply due to abstraction of groundwater with salt concentration that exceeds the treatment threshold. The profiles in Figure 8a indicate that, for lower WL and larger Q, the cost-optimal pumping strategies are those with the smallest D. For larger WL and lower Q, however, the cost-optimal pumping strategies are found for intermediate values of , between 50 and 100 m.

Figures 8b-c-d show that the drawdown percentage , the percentage of the freshwater volume decrease , and the percentage of salt mass increase share a similar behavior, generally decreasing with Dand increasing with both WL and Q. In Figure 8b, it is interesting to observe that if D is generally over ~150 m, the water table results marginally affected by pumping and most of the groundwater abstracted is resident seawater. In Figures 8c-d, the dependency of and on Q may be explained through simple considerations of aquifer mass balance.

At steady state, the abstraction rate Q is provided in part by the freshwater recharge, and in part by the seawater inflow from the shoreline boundaries, with an overall decrease in the freshwater lens volume. If the pumping depth Dis increased or the distance WL is decreased, both and become progressively less significant, and if Dis larger than ~150 m, they result negative, which implies an overall increase of the freshwater lens volume as pumping removes salt water from underneath, thus promoting the downward flow of freshwater from recharge.

Altogether, the operating cost function and the SWI indicators presented in Figure 8 illustrate the inherent conflicts between the economic cost of groundwater supply and the management SWI. On one hand, for any given demand Q, is minimized by selecting a shallow pumping system situated towards the center of the island center. On another, to limit SWI indicators, such as , and , it is necessary to select deeper pumping systems and closer to the shoreline, which may massively increase the operation cost due to desalination requirements.

|  |  |
| --- | --- |
| (a)  C:\Users\ciq19wy\Desktop\Fig_5_OC.png | (b)  C:\Users\ciq19wy\Desktop\Fig_5_s.png |
| (c)  C:\Users\ciq19wy\Desktop\Fig_5_FV.png | (d)  C:\Users\ciq19wy\Desktop\Fig_5_SM.png |
| **Figure 8.** Profiles of (a) the operation cost and the SWI indicators (b) , (c) and (d) for a number of groundwater abstraction strategies (WL,D,Q). | |

1. **Application of the Proper Orthogonal Decomposition (POD) method**

Figure 9 shows that for the 1D model, the results of the conventional finite difference and POD approaches are almost similar. Similar very good performance is observed for the 3D model (Figure 10). These results indicate that we can correctly derive the POD equations. The results also indicate that we can assess the performance of the POD model with respect to the number of full model snapshots used in the POD simulations, which is the primary method we will rely on to compare the computational effort between conventional and POD based seawater intrusion models.

|  |
| --- |
| Graphical user interface, application  Description automatically generated |
| **Figure 9.** Results of the one-dimensional groundwater flow model. |

|  |
| --- |
|  |
| **Figure 10.** Results of the three-dimensional groundwater flow model, a) Flopy and b) POD |

1. **Optimizing Salinity Control Structure locations for the Biscayne Aquifer**

The concentration contour of the Biscayne Aquifer is shown in Figure 11 and the flow-weighted well field concentration is shown Figure 12. We found that at high sea level rise rate (2.03 ft by 2060) and at current pumping rate (base case pumping), moving G54 has only a local effect and is only beneficial to Dixie well field (see Figure 12b) in terms of reducing the well concentration. Based on the well concentration plot in Figure 12, pushing the control structure, G54, further to the coast than originally proposed new position from the USGS report does not make a difference, given the sea level rise rate and pumping rate. However, the seawater front has indeed been pushed back behind the control structure (Figure 12(e)), which will greatly improve the water quality for Dixie Well field, at a much higher pumping rate and much higher sea level rise rate.

|  |
| --- |
| A picture containing text  Description automatically generated |
| **Figure 11.** Salinity (in fraction of seawater) contours of Layer 10 (base of the Biscayne Aquifer) of the sensitivity analysis with different positions of salinity control structure (G54). The contour with G54 being at the original position is shown in the very left figure. The contours with moving G54 further to the coast along the canal are shown from (a) to (e) with each position being closer to the coast than previous position (i.e., (a) is the one that is farthest from the coast, and (e) is the one that is closest to the coast). The original position of G54 is denoted by the yellow rectangle from (a) to (e). |

|  |
| --- |
|  |
| **Figure 12.** Simulated, flow-weighted, average chloride concentration at each well field for the Scenario of original G54 position (highdsl), Scenario of 1st movement of G54 (G54\_RC1), Scenario of 2nd movement of G54 (G54\_RC2), Scenario of 3rd movement of G54 (G54\_RC3), Scenario of 4th movement of G54 (G54\_RC4), and Scenario of 5th movement of G54 (G54\_RC5). All scenarios have high sea level condition (2.03 ft by 2060). Yellow shading indicates drier than average climatic periods, and blue shading indicates wetter climatic periods. |

**Opportunities for training and professional development**

* The US postdoc, Yipeng Zhang started 9/1/2020. He has learned how to simulate density-dependent subsurface flow and transport, use Gaussian process modelling to develop surrogate optimization models, and to use FlowPy software for setti ngup and processing results from the MODFLOW family of codes. He is a co-author on an upcoming paper on seasonal recharge impacts on seawater intrusion and will be first author on an upcoming paper on using Gaussian process modeling for surrogate models of optimizing Henry's problem with pumping.
* The UK postdoc, Mohammadali Geranmehr started 25/10/2021. He has learned how to simulate density-dependent subsurface flow and transport, as well as use FlowPy software for setting up and processing results from the MODFLOW family of codes. He has been developing reduced-order modelling methods to apply to the develop of surrogate optimization models for the Santa Barbara aquifer and is co-author of a paper in peer review on the impact on SWI constraints on the management of island aquifers.
* A PhD student, Lauren Mancewicz, who works on another NSF seawater intrusion project, has been participating in our surrogate modeling group discussions and has been learning the principles of surrogate modelling. She has also trained the UK PhD student how to use the USGS SEAWAT model and FlowPy and how to set up an island-freshwater lens model. This has given her valuable experience in working on group projects.
* A UK PhD student, Weijang Yu, not directly involved in the project, has been participating in our surrogate modelling group discussions and has been developing machine learning based Gauss process surrogate models of Island aquifers, which has given him valuable experience in working on group projects. He is first author of a paper in peer review on the impact on SWI constraints on the management of island aquifers.

**Results Dissemination**

1. Three presentations on the project were made at the American Geophysical Union's 2022 Frontiers in Hydrology Meeting Conference, San Juan, Puerto Rico, 19-24 June 2022:
   1. Zhang, Y., Mayer, A., Gulley, J., Bedekar, V. and Martin, J.P., "Brackish Water Depletion on Tropical Islands under Seasonal Climate Patterns as Lakes Form and Expand with Rising Sea Level."
   2. Yu, W., Bau, D., Mayer, A., Zhang, Y., Mancewicz, L., and Geranmehr, M., "Comparison of off-line and on-line trained Gaussian process models for island groundwater management."
   3. Mancewicz, L., Mayer, A., Gulley, J., and Martin, J.P., Geological Sciences, Gainesville, United States, "Impact of Sea Level Rise on the Freshwater Lens and Hypersaline Lakes Found on Small Carbonate Islands."
2. Two presentations on the project were made at the Geological Society London’s Ineson Lecture Meeting 2022, London, UK, 8 November 2022:
   1. Geranmehr M., Bau D., Mayer A.S., Mancewicz L. and Yu W, “Developing a Reduced-Order Groundwater Model for Managing Seawater Intrusion in Coastal Aquifers”.
   2. Yu W., Bau D., Mayer A.S., Mancewicz L. and Geranmehr M., “Management of Island Aquifers by Simulation-Optimization”.
3. Two presentations on the project were made at the American Geophysical Union's 2022 Fall Meeting Conference, Chicago, 12-16 December 2022:
   1. Bau D., Yu W., Geranmehr M., Mayer A.S., Zhang Y. and Mancewicz L., "Sustainable Management of Groundwater Resources in Island Aquifers".
   2. Zhang Y., Mayer A.S., Bau D., Yu W., Geranmehr M., and Mancewicz L., "Combined Impacts of Hydraulic Conductivity Heterogeneity and Pumping Intensity on Seawater Intrusion in Coastal Aquifers: A Numerical Investigation Using a Henry Problem with a Pumping Well".
4. A research article has been submitted for peer review to Water Resources Research titled “Investigating the impact of seawater intrusion on the operation cost of groundwater supply in island aquifers”, by Yu W., Bau D, Mayer A.S., Mancewicz L., and Geranmehr M.

**Planned Activities for 2023**

The major activities in the current project year will include the following:

* Modified Henry’s problem: Complete modified Henry’s problem simulations and develop Gaussian process surrogate models for the response surfaces. Submit two papers on this work. Present results at conference.
* Biscayne aquifer: Use the simulation results to develop response surfaces for a range of state variables. Replace response surfaces with Gaussian process surrogate models. Submit paper on this work. Present results at conference.
* POD: Begin simulations of the Santa Barbara aquifer. Develop optimization problems for Santa Barbara aquifer. Develop POD model for the Biscayne aquifer.
* Island aquifer: Complete Gaussian process surrogate models. Complete publication of one paper and submit another paper on this work.

**References**

1. Henry, H. R. (1964). Effects of dispersion on salt encroachment in coastal aquifers, in"Seawater in Coastal Aquifers". US Geological Survey, Water Supply Paper, 1613, C70-C80.
2. Hughes, J. D., Sifuentes, D. F., & White, J. T. (2016). Potential effects of alterations to the hydrologic system on the distribution of salinity in the Biscayne aquifer in Broward County, Florida (No. 2016-5022). US Geological Survey.
3. Javadi, A. A., Abd‐Elhamid, H. F., & Farmani, R. (2012). A simulation‐optimization model to control seawater intrusion in coastal aquifers using abstraction/recharge wells. International Journal for Numerical and Analytical Methods in Geomechanics, 36(16), 1757-1779.
4. Langevin, C.D., D.T. Thorne Jr., A.M. Dausman, M.C. Sukop and Weixing G. (2008). SEAWAT Version 4: A Computer Program for Simulation of Multi-Species Solute and Heat Transport: U.S. Geological Survey Techniques and Methods Book 6, Chapter A22, 39 p.
5. Moore, P. J. (2009). Controls on the generation of secondary porosity ineogenetic karst: examples from San Salvador Island, Bahamas and north-central Florida, USA, (Doctoral dissertation). Gainesville, FL: University of Florida.
6. Siade, A. J., Putti, M., and Yeh, W. W.-G. (2010), Snapshot selection for groundwater model reduction using proper orthogonal decomposition, Water Resour. Res., 46, W08539, doi:10.1029/2009WR008792.
7. Rasmussen C.E. and C.K.I. Williams (2006). Gaussian Processes for Machine Learning, The MIT Press, ISBN 0-262-18253-X.
8. Remy, N., Boucher, A., & Wu, J. (2009). Applied geostatistics with SGeMS: A user's guide. Cambridge University Press.