Supporting Information for "DU Pathways$_{ERAS}$: Cooperative Water Supply Investments that are Equitable, Robust, Adaptive and Stable"

David F. Gold[1], Patrick M. Reed[1], David E. Gorelick[2,3], Gregory W. Characklis[2,3]

[1]Department of Civil and Environmental Engineering, Cornell University, Ithaca, NY, USA
[2]Center on Financial Risk in Environmental Systems, Gillings School of Global Public Health and UNC Institute for the Environment, University of North Carolina at Chapel Hill, North Carolina, USA
[3]Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Contents

S1 Synthetic streamflow generation

Synthetic streamflow generation by the (Kirsch et al., 2013) generator begins by log transforming and whitening the record of historical weekly inflows, $\boldsymbol{Q_k} \in \boldsymbol{R}^{(80\times52)}$ to create a matrix $\boldsymbol{Z_k} \in \boldsymbol{R}^{(80\times52)}$ for each gage k. Next, a matrix of integer indices $\boldsymbol{M} \in \boldsymbol{R}^{(1000\times52)}$ is generated by sampling with replacement from $(1, 2, \ldots, 80)$. $M_{i,j}$ represents the historical year that will be used to create the streamflow value for synthetic year i in week j. $\boldsymbol{M}$ is used to make a matrix of uncorrelated synthetic flows, $\boldsymbol{C_k}$ with entries $C_{k_{i,j}} = Z_{k_{M_{(i,j)},j}}$. The same matrix $\boldsymbol{M}$ is used to for all sites to preserve spatial correlation for synthetic records. Next, a matrix of historical autocorrelation, $\boldsymbol{p_{H_k}} = corr(\boldsymbol{Z_k})$ is created for each gage and a Cholesky decomposition is used to find an upper triangular matrix $\boldsymbol{U_k} \in \boldsymbol{R}^{(52\times52)}$ such that $\boldsymbol{p_{H_k}} = \boldsymbol{U_k}\boldsymbol{U_k^T}$. Upper triangular matrix $\boldsymbol{U_k}$ is then used to impose the historical autocorrelation structure on matrix $\boldsymbol{C_k}$ to make a new synthetic record $\boldsymbol{S_k} = \boldsymbol{C_k}\cdot\boldsymbol{U_k}$. Finally, $\boldsymbol{S_k}$ is transformed back into real space to gen-

Corresponding author: David F. Gold, `dfg42@cornell.edu`

erate a record of reservoir inflows that preserve the spatial and temporal correlation structures of the historical record.

To improve the inter-annual correlations of synthetic streamflows, this process is repeated using a shifted version of historical inflows, $\boldsymbol{Q_{k'}}$ beginning at week 27 of each year and ending at week 26 of the following year. Matrices $\boldsymbol{Z_{k'}}$, and $\boldsymbol{U_{k'}}$ and created based off this shifted record and $\boldsymbol{C_{k'}}$ is created separately shifting matrix $\boldsymbol{C_k}$. A new matrix of synthetic inflows, $\boldsymbol{S_{k'}}$ is created using the operation $\boldsymbol{S_{k'}} = \boldsymbol{C_{k'}} \cdot \boldsymbol{U_{k'}}$ and transforming the product back to real space. The final set of synthetic streamflows is comprised of columns 27-52 of $\boldsymbol{S_k}$ and columns 1-26 of $\boldsymbol{S_{k'}}$. For more details on the synthetic generation process, refer to Kirsch et al. (2013) and Herman et al. (2016).

The number of streamflow samples used in this paper were chosen based on empirical assessment. (Trindade et al., 2017) empirically assessed the number of the number of realizations needed to estimate the objective functions for the Research Triangle test case by examining sample sizes varying from 100 to 5000 realizations. Results of the empirical assessments showed that 1000 evaluates per modeling run is sufficient to approximate the mean and variances of the Monte Carlo distributions used to determine candidate solutions' objectives. The approach used by (Trindade et al., 2017) is derived from early studies of metaheuristic search dynamics given noisy objective functions (e.g. (Miller & Goldberg, 1996; Smalley et al., 2000)) which show that relatively small Monte Carlo samples per function evaluations can provide good approximations when verified with much larger samples after search has been completed.

S2 Runtime Diagnostics

Multiple instances of MOEA search are run ensure the algorithm has overcome any biases in search generated by the initial population (Salazar et al., 2017). In this experiment, a total of 10 random seeds were run, using the multi-master configuration of the Borg MOEA with two seeds per master. The true Pareto set for this problem is not known, so to assess the convergence convergence we measure relative hypervolume (Zitzler et al., 2003), which compares performance of the approximate Pareto sets discovered at set checkpoints within search to the final "reference set", which contains non-dominated solutions across all seeds. If the relative hypervolume is found to plateau, we conclude that the algorithm has converged to a satisfactory approximation of the true Pareto set.

Runtime diagnostics for all seeds optimizations are shown in Figure S1. There was very little variance across seeds, and the hypervolume of all defection optimizations plateaued after around 50,000 function evaluations.

S3 Distribution of Unit Cost objective for the DSFR compromise

Figure S2 shows the distribution of the unit cost of expansion objective for Durham across the 2,000 SOWs used for DU reevaluation for the DFSR compromise. Of the 2,000 DU SOWs, over 1,900 return unit costs near zero. However, the extreme tail of the unit cost of expansion increases to over $1,000/kgal. This extreme tail explains the high regional value of the unit cost objective shown in Figure 4a - because DU optimization calculates values in expectation across all sampled futures, extreme values in the tails have a large impact on the objective value. Future work may reduce the impact of these extreme SOWs by using other summary statistics such as the median or 90th% unit cost.

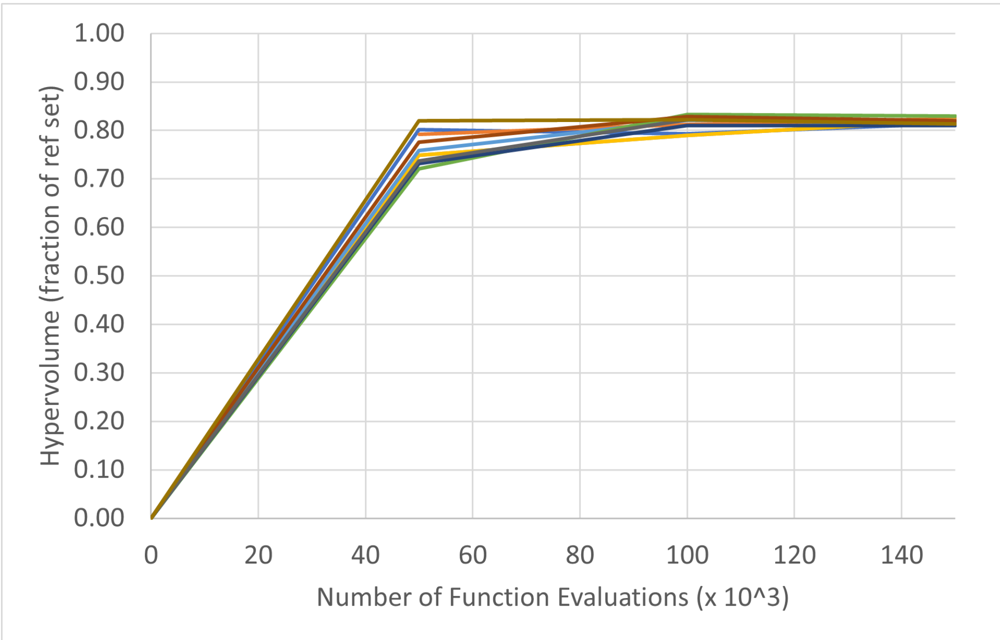

Figure S1. Runtime diagnostics for 10 random seeds. The plateau of hypervolume across all seeds for all formulations indicates that number of function evaluations (NFE) were enough to achieve maximum attainable convergence.

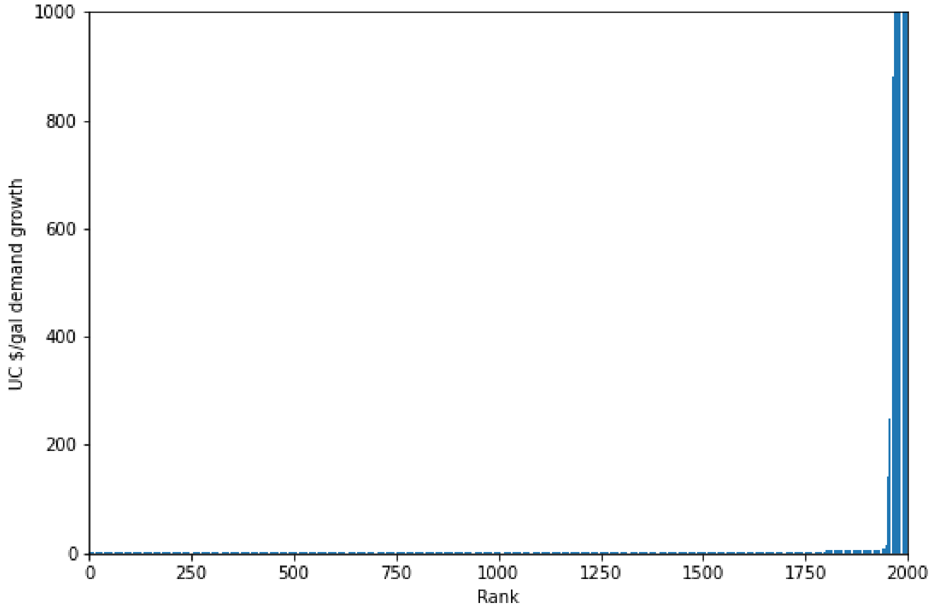

Figure S2. Distribution of Unit Cost for Durham across 2,000 DU SOWs

References

Herman, J. D., Zeff, H. B., Lamontagne, J. R., Reed, P. M., & Characklis, G. W. (2016, November). Synthetic Drought Scenario Generation to Support Bottom-Up Water Supply Vulnerability Assessments. *Journal of Water Resources Planning and Management*, *142*(11), 04016050. Retrieved 2019-03-17, from `https://ascelibrary.org/doi/full/10.1061/(ASCE)WR.1943-5452.0000701` doi: 10.1061/(ASCE)WR.1943-5452.0000701

Kirsch, B. R., Characklis, G. W., & Zeff, H. B. (2013, July). Evaluating the Impact of Alternative Hydro-Climate Scenarios on Transfer Agreements: Practical Improvement for Generating Synthetic Streamflows. *Journal of Water Resources Planning and Management*, *139*(4), 396–406. Retrieved 2019-03-17, from `https://ascelibrary.org/doi/10.1061/%28ASCE%29WR.1943-5452.0000287` doi: 10.1061/(ASCE)WR.1943-5452.0000287

Miller, B. L., & Goldberg, D. E. (1996). Optimal sampling for genetic algorithms. In *Proceedings of the artificial neural networks in engineering (annie'96) conference* (Vol. 6, pp. 291–297).

Salazar, J. Z., Reed, P. M., Quinn, J. D., Giuliani, M., & Castelletti, A. (2017). Balancing exploration, uncertainty and computational demands in many objective reservoir optimization. *Advances in water resources*, *109*, 196–210.

Smalley, J. B., Minsker, B. S., & Goldberg, D. E. (2000). Risk-based in situ bioremediation design using a noisy genetic algorithm. *Water Resources Research*, *36*(10), 3043–3052.

Trindade, B., Reed, P., Herman, J., Zeff, H., & Characklis, G. (2017). Reducing regional drought vulnerabilities and multi-city robustness conflicts using many-objective optimization under deep uncertainty. *Advances in Water Resources*, *104*, 195–209.

Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., & Da Fonseca, V. G. (2003). Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on evolutionary computation*, *7*(2), 117–132.