Water Management in the U.S. Southwest: A Systems View

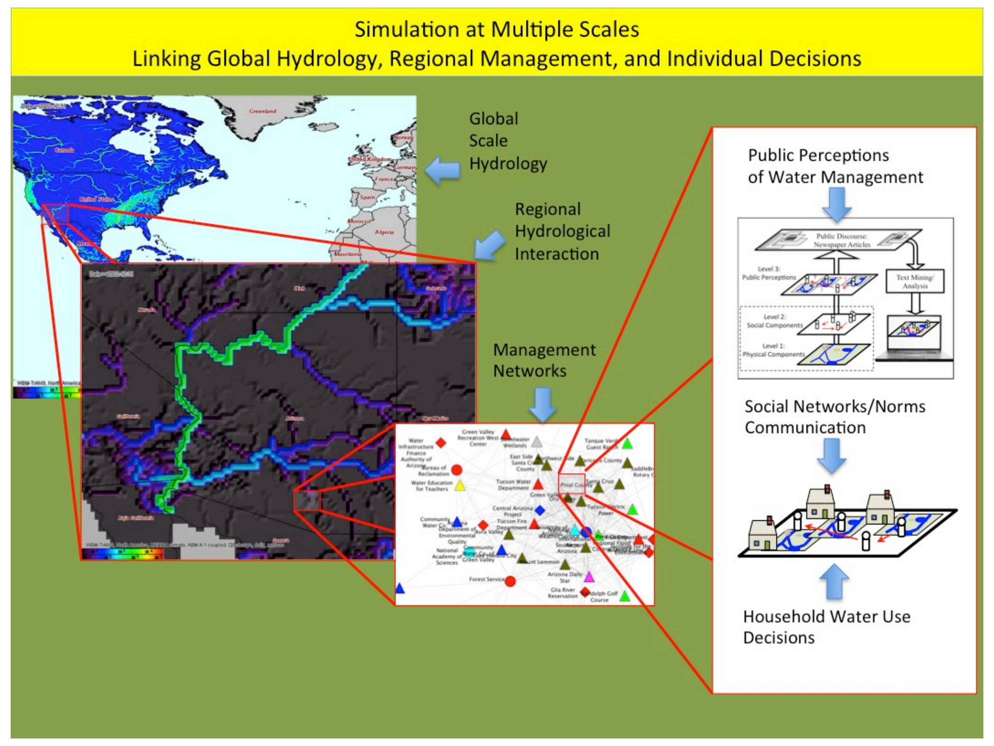
John T. Murphy[1,2], Jonathan Ozik[1,2], Nick Collier[1,2], Mark Altaweel[1,3], Richard Lammers[4], Alex Prusevich[4], Andrew Kliskey[5], Lilian Alessa[5], Chun Ho Chow[1]

[1]University of Chicago; [2]Argonne National Lab; [3]University College London; [4]University of New Hampshire; [5]University of Idaho Moscow

Overview

The management of water is critical in the U.S. and around the world. Our project conceptualizes water management as occurring at several distinct but related levels:

- Global hydrology is driven by large-scale climate and weather patterns
- Regional water management moves water across or between watersheds
- Local water management occurs in a context of multiple peer institutions that form a network
- Individual or household water use is driven by human decisions that are shaped by individual characteristics and perceptions about the resource being used



Our project uses data mining and simulation modeling to understand how these levels interact. Specifically, we examine:

- How individual perceptions shape household consumption
- How networks of water management, and the legal framework within which institutions must act, shape municipal water demand
- How municipal demand leads to impacts on regional hydrology

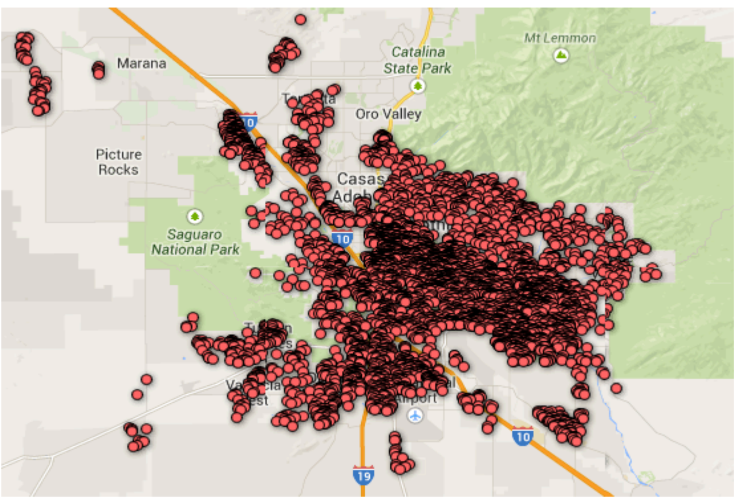
Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. BCS-1114851.

The authors gratefully acknowledge Tom Arnold and the Tucson Water Department for providing data for the Tucson household usage study. We additionally thank Michael Wilde and Justin Wozniak (Argonne Mathematics and Computing Sciences) for their kind support with the Swift/K and Swift/T software. Stanley Glidden (UNH) provided comparative data for the WBM/ABM coupled runs.

Residential Water Use

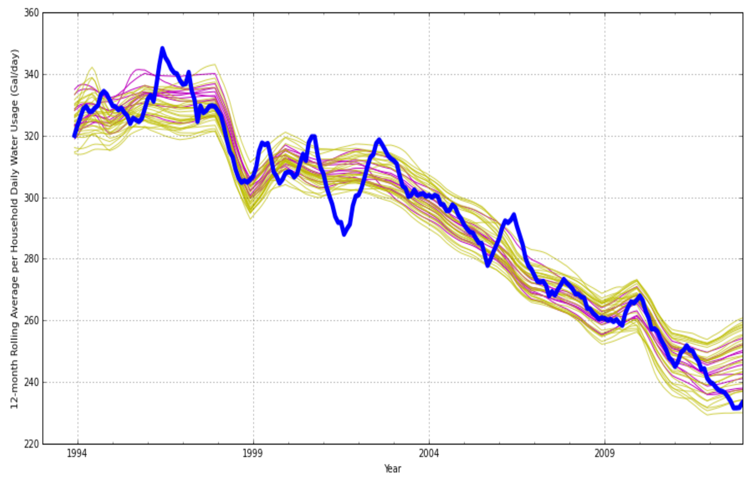
Residential water use is of critical interest to planners and water managers throughout the U.S. and the world. However, predicting fluctuations in water consumption, and understanding the causes of those fluctuations, is challenging. To address this, we created an **agent-based model** of household water use, and calibrated it to water use in **Tucson, Arizona**. Each agent represented a single household, and was endowed with:



- A unique baseline value for water use
- A value for sensitivity to price increases or 'shocks'
- A value for the weight given to price as a concern for shaping usage
- A value representing conservation attitudes
- A value for the weight given to conservation as a concern for shaping usage
- Connections in two distinct social networks through which conservation messages could spread

More than 200,000 households could be simulated, representing all of Tucson; alternatively, a subset of these could be sampled. Real-world pricing data was used to capture price increases and potential price shocks during the period studied. Calculated demand was also adjusted for climate based on historic weather data. In some test runs, conservation messages could be delivered to randomly selected households in the social networks; the assumed 'effectiveness' of these messages could be varied, with the effect of a message being a change in the conservation attitudes held by the household.

Approximately **27,000 runs** were performed using the Swift/T test harness. The figure below (left) gives a comparison with actual usage data for Tucson from **1994-2012** and the top 40 runs as assessed by closeness-of-fit to the empirical data (using least-squares difference method). The empirically attested data is in blue; magenta lines indicate the usage of the top 10 runs, yellow the next top 30 runs. At right is a figure showing the combinations of price and conservation weighting and their closeness to the empirical data (blue is closer, red is more distant).



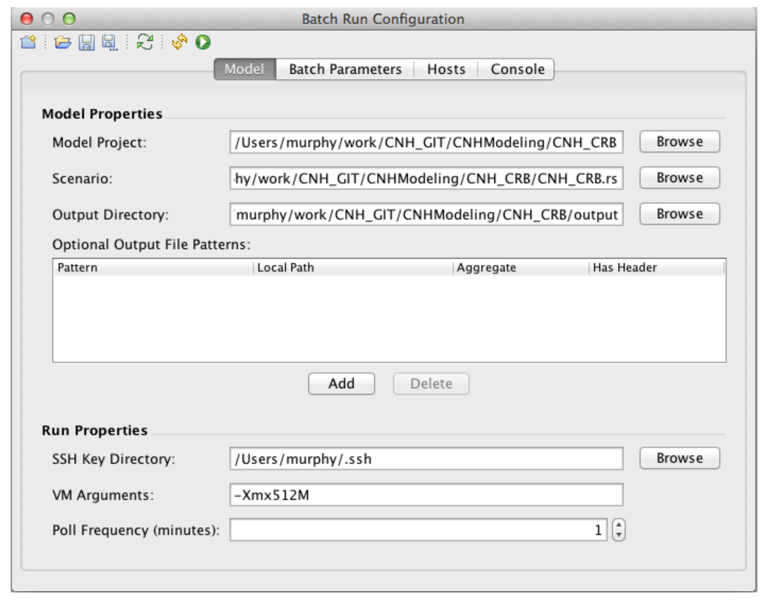
messagingBeginYear (All)

Average of DIFF — Social Factor Weight (rows) × Price Weight (columns)

Row Labels	0	0.025	0.05	0.075	0.1	0.125	0.15	0.175	0.2	0.225	0.25
0	57.9	53.9	49.5	44.7	40.7	36.2	31.8	27.1	22.1	18.1	17.5
0.025	53.8	49.6	45.2	40.8	35.6	31.9	27.2	22.9	17.4	15.3	12.9
0.05	49.7	45.6	41.1	36.7	32.3	27.3	21.8	18.2	15.7	12.3	11.6
0.075	45.7	41.6	36.9	32.0	27.3	23.0	18.8	14.8	13.5	12.1	13.1
0.1	41.7	37.4	33.1	28.9	23.9	20.5	16.5	12.5	11.1	13.1	17.1
0.125	37.9	33.9	29.3	24.7	20.9	16.6	14.1	12.5	11.1	15.7	18.9
0.15	34.2	30.2	25.9	21.6	17.4	15.4	12.2	13.2	15.1	20.4	24.7
0.175	30.8	26.8	22.9	19.2	15.6	14.1	13.8	16.1	18.5	24.0	27.2
0.2	27.6	23.7	20.1	17.2	15.3	14.8	15.8	18.1	22.8	26.0	32.2
0.225	24.8	21.5	18.7	16.9	15.9	18.2	19.7	24.2	28.6	31.2	34.9
0.25	22.7	19.9	17.9	17.6	17.4	20.8	22.8	27.8	31.2	37.4	41.2
0.275	21.3	19.2	18.7	19.2	21.9	24.2	25.8	34.0	37.3	39.4	45.7
0.3	20.7	19.8	19.7	21.8	24.7	27.6	28.8	37.0	42.6	47.1	46.4
0.325	21.1	21.3	22.5	24.9	28.7	31.9	35.1	39.6	43.8	48.8	56.2
0.35	22.5	23.6	25.5	27.9	31.2	35.6	40.3	44.6	50.2	54.1	58.1
0.375	24.5	26.2	28.7	32.9	36.3	38.8	43.5	48.6	53.9	61.3	64.2
0.4	27.1	29.3	32.7	36.0	40.5	44.4	46.9	52.9	59.8	62.7	66.1

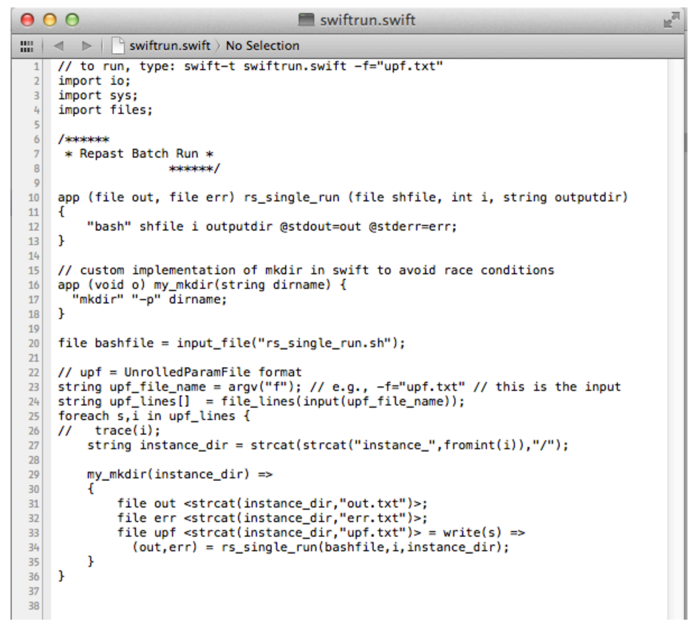
Adaptive Simulation

The complexity of the Agent-Based models that we are using to explore these issues commonly leads to a profusion of parameters, and, therefore, to an extremely large parameter space. The computing facilities at the RCC permit large numbers of runs to be performed. However, the logistical challenges of organizing these large ensemble runs are not insignificant.



To overcome these challenges, we use the **Swift** workflow management tool, developed by Argonne National Lab. This tool allows the specification of a graph of tasks and dependencies (as in the script below), and then automates the invocation of the executables needed to create the component nodes in the graph.

We use this in conjunction with our Agent-Based Modeling toolkit, **Repast Simphony**. Repast Simphony includes robust tools for preparing large ensembles that can be run on a variety of platforms. The user interface (right) allows the creation of a file that specified the combinations of parameters to be used. Repast can then launch ensembles on a variety of platforms, or, as in our case, can be manually launched with Swift.



Using Swift also allows **adaptive simulation**. This implies that given a large parameter space, the strategy to explore that space can be calculated and/or adjusted as simulations are in progress. We have developed simulation harnesses in Swift that allow evolutionary approaches (genetic algorithms), various types of sampling and filtering.

Swift exists in two versions: Swift/K is the original tool, and Swift/T is a newer implementation that uses MPI and the Asynchronous Dynamic Load Balancing library (ADLB) to extend Swift's power to the largest HPC platforms. We have employed both Swift/K and Swift/T for our projects here.

Swift can be used both in the generation of results over large ensemble runs and in the post-run analyses. Because the ensemble runs often generate very large datasets yielding tens of thousands of individual result files, Swift is employed to process these files as well. This processing can be included in the original run or performed as a separate job, depending on whether the specifics of the analyses needed are known in advance of the simulation run or if they are decided upon completion and inspection.

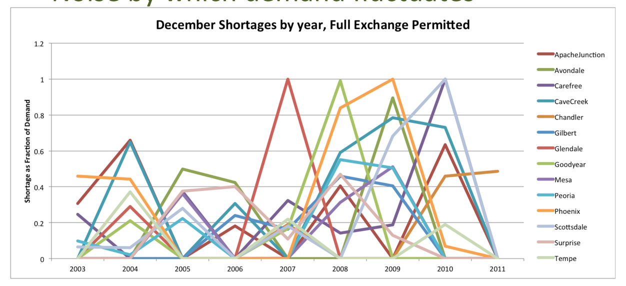
Water Management

Water is delivered to Phoenix via the CAP Canal. The CAP system serves municipal water management agencies; these provide water to their cities according to the city's population and gallons per capita per day (GPCD) demands.

The CAP system requires each manager to submit a schedule in October for deliveries per month for the following year. This is a projection, and may not be accurate. During the year, shortages and surpluses can be exchanged. Each manager must decide when to return water, or when to try to obtain more water, and (in either case) how much.

We simulate this by assuming that each manager will schedule a given year based on the previous year's usage. During the year each manager has a parameter for:

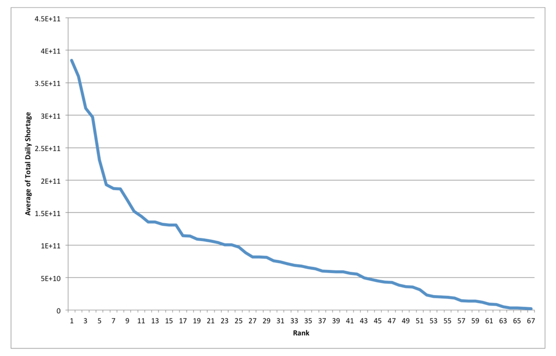
- Percentage **over prediction** at which water is returned
- Percentage **under prediction** at which water is sought
- Fraction of surplus returned
- Fraction of shortage requested
- Noise by which demand fluctuates



In single runs (left), our interest is in which cities are more or less likely to have end-of-year shortages as their predictions are less accurate and demand exceeds supply. For these runs, all CAP customers have the same values for their parameters, and we can compare results with exchanges permitted vs. those without.

Alternatively, and more realistically, we can assume that each customer has different thresholds, fractions, and noise levels, and ask whether the values for specific customers (e.g. very large cities), are determinant, and, if so, which kinds of values (e.g. overage vs. shortage thresholds) are more significant.

The figure to the right ranks a collection of **29,464 runs** performed using the Swift/T parallel scripting language. The runs are a strategic sample of the very large parameter space possible with 14 customers x 5 parameters/customer. For each parameter several values are explored, and the slope of the regression through the average total shortage for all runs with that value is calculated. The list at far right shows these results by parameter. Note that 'shortage threshold' seems much more significant than other factors, especially 'noisiness'.



Variable	Slope
Goodyear_shortageThreshold	3.84528E+11
Gilbert_overageThreshold	3.5977E+11
Tempe_shortageThreshold	3.10557E+11
Mesa_shortageThreshold	2.96931E+11
Carefree_shortageThreshold	2.30866E+11
CaveCreek_shortageThreshold	1.92848E+11
Avondale_overageThreshold	1.8707E+11
Mesa_amtToObtain	1.86322E+11
CaveCreek_overageThreshold	1.69394E+11
Scottsdale_shortageThreshold	1.52278E+11
Peoria_amtToObtain	1.4442E+11
Glendale_shortageThreshold	1.35741E+11
Surprise_shortageThreshold	1.35435E+11
Carefree_overageThreshold	1.31908E+11
Chandler_overageThreshold	1.30983E+11
ApacheJunction_shortageThreshold	1.30883E+11
Avondale_shortageThreshold	1.14664E+11
Chandler_amtToObtain	1.13893E+11
Scottsdale_overageThreshold	1.09022E+11
Peoria_amtToReturn	1.08173E+11
Peoria_shortageThreshold	1.06208E+11
Avondale_amtToObtain	1.04129E+11
Avondale_amtToReturn	1.00697E+11
Surprise_amtToReturn	1.00223E+11
ApacheJunction_amtToReturn	96839780882
Phoenix_amtToReturn	88165057443
Phoenix_amtToObtain	81758772832
CaveCreek_amtToReturn	81624519331
CaveCreek_amtToObtain	80950457912
Surprise_overageThreshold	75951618464
Phoenix_shortageThreshold	74202296805
Surprise_amtToObtain	71199156982
Carefree_amtToReturn	68861403787
Chandler_shortageThreshold	67819395073
CaveCreek_noisiness	65427751129
Gilbert_noisiness	63518633465
Gilbert_shortageThreshold	59885098610
Gilbert_amtToObtain	59370103196
Gilbert_amtToReturn	58883591134
Carefree_noisiness	58798669126
Surprise_noisiness	56767606891
Glendale_overageThreshold	55573215839
Goodyear_amtToReturn	49433327015
Tempe_noisiness	46919401969
Goodyear_overageThreshold	44638739806
Scottsdale_amtToReturn	42815756127
Tempe_overageThreshold	42735510989
Glendale_noisiness	38569125669
Chandler_noisiness	36006131297
Chandler_amtToReturn	35541926190
ApacheJunction_amtToObtain	31372826026
Scottsdale_amtToObtain	23364558604
Phoenix_noisiness	21047871817
Phoenix_overageThreshold	20329696368
Glendale_amtToReturn	19429015421
Tempe_amtToReturn	18434685639
Tempe_amtToObtain	14375938976
Mesa_overageThreshold	14051196891
Mesa_noisiness	13598772102
Glendale_amtToObtain	12163018820
Avondale_noisiness	9341928717
Mesa_amtToReturn	8571740034
ApacheJunction_overageThreshold	5265487411
Goodyear_amtToObtain	3289237465
Goodyear_noisiness	3116341667
ApacheJunction_noisiness	2535395911
Scottsdale_noisiness	1859606376

Regional Hydrology

In previous studies, we have coupled our **Agent-Based Model (ABM)** with a regional- to global-scale hydrology model called the **Water Balance Model(WBM)**. The WBM operates on grid cells with resolution as fine as 6" lat/lon. It provides an accounting of surface flow and groundwater for each cell. Surface flow is modeled through an abstract network that connects grid cells into representations of water sheds.

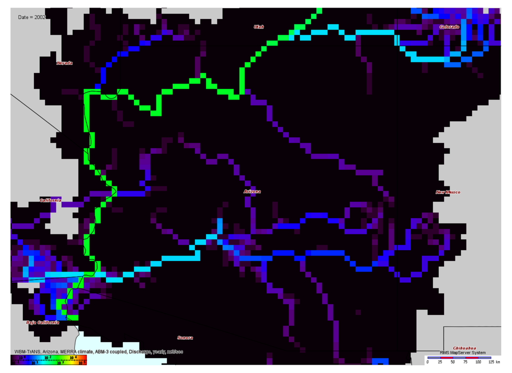
The motivation for linking the WBM and ABM was two-fold:

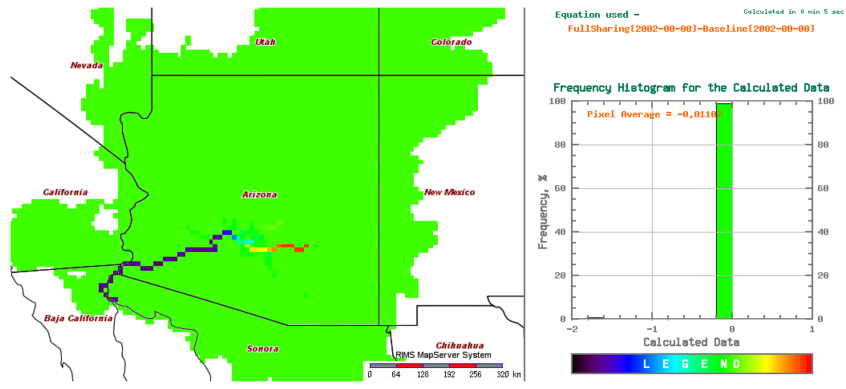
- **To allow the ABM to model accurate hydrology**: The impacts of human decisions on the physical hydrology of the region could be captured
- **To improve the WBM's representation of human decisions**: The WBM implemented generic rules for irrigation or the operation of dams, but these were abstract and not constrained by the actual situation found in a specific context

Our procedure for coupling the ABM and the WBM overcame several conceptual and practical difficulties:

- To avoid duplication, the WBM's human system had to be 'turned off' at targeted points
- WBM and ABM resolution had to be resolved: because the ABM operates at a finer resolution, a strategy by which ABM results and interactions could be aggregated to match WBM grid cells was implemented.
- Because the WBM model run separately from the ABM, the **ZeroMQ** protocol was used to pass messages from the WBM to the ABM and back
- Because the WBM natively includes irrigation, 'Farmer' agents had to be included in the model. These received information from the WBM on the amount of water to draw, but did not follow the WBM's generic rules for finding sources of the water, and instead drew water from the CAP canal system.

Our objective was to ensure that a coupled run could duplicate the original runs using the de-coupled WBM. An example of the baseline run is given at left, below. The figure on the right is a subtraction of a coupled ABM/WBM run from the baseline run. While we are still working to interpret these results, the close correspondence suggests that our mechanisms are functioning properly and will provide useful results.





A current objective is to install the WBM on the Midway cluster, and to allow large-ensemble runs to be done with the coupled WBM/ABM system.