Progress report for “Scaling of hydrologic and land-surface responses: Are the right processes represented at the right scale?”

N.B. Engdahl and A.S. Richey

1. Overview

Groundwater is declining across much of Washington State [*Burns et al*., 2012; *Vaccaro et al.*, 2015] and there is a general consensus that climate change will place significantly more stress on existing resources [*Pitz, 2016*]. The Washington State Department of Ecology (Ecology) has called for establishing formal mechanisms to monitor and assess current and future groundwater depletion in the state’s aquifers [*Pitz*, 2016], which will require an exhaustive inventory of existing resources as well as dynamic models to infer future changes. This is problematic because there is a discrepancy between the scale of hydrologic measurements, often scattered points, and the scale of management decisions across the entirety of a region. Long-term planning and management typically involves combining observations with modeling, so the disparity of the information sources could lead to large discrepancies between planned usage and actual usage, making the problem very real. As such, the central theme of this project is the transfer of information across scales and whether or not a model constructed at one scale (resolution) is equivalent to another model constructed at a different scale over the same area.

Data interpolation and numerical modeling efforts are often combined to coarsen local measurements for regional applications but doing so confidently requires an understanding of how hydrologic processes interact across scales. The average response of a fine resolution model for total water content, for example, may not match the result of a coarse resolution model even when they are calibrated to the same data [*Hill and Tiedeman*, 2007]. The reason for this phenomenon is twofold, with part of it being the mathematical issue of non-uniqueness and the other being the nonlinear response of the complex processes. The former is unavoidable, but the latter occurs because the numerical solution (integration) of the governing equations changes when solved at different scales. The main goal of the project is understanding how these kinds of disparities affect the results of groundwater models. Specifically, we ask, how do process interactions and the hydrological response at small scales translate to larger scales? The reason this question is so critical is that, overwhelmingly, the only comparison metrics for hydrologic simulations are point observations (heads, streamflow, volumetric water content, etc…), and if two models of the same site can be fit to the same data equally well, how can one say which is correct? The approach for investigating these questions is to use multi-scale numerical models to quantify the magnitude and spatial trends in the differences seen at the different scales for the same sites. We have made significant progress toward understanding how the nonlinear response of these hydrologic systems differ and are nearing completion of the project.

1.1 Study sites

The project originally intended to use the Cook Agronomy Farm (CAF) located near Pullman, WA which is managed by Washington State University (WSU) in collaboration with Pullman USDA/ARS scientists (http://css.wsu.edu/cook/). The CAF (Figure 1) is part of the Long-Term Agroecological Reserve (LTAR) network established by the U.S. Department of Agriculture and the site covers an area of roughly 0.57km^2, spanning an area several hundred meters across. The CAF is home to a large number of multi-disciplinary research projects focusing on agricultural efficiency and process-oriented applied research. Based on conversations with colleagues who have worked at the site, we were led to believe that the abundant data (both characterization and observation) was readily available. However, this was found not to be the case. Apparently, despite numerous persons believing otherwise, no central data repository for the site exists. The limited data sets we could obtain were not detailed enough and everyone involved with this data “thought” someone else had more of it. Clearly, sorting out these discrepancies is an issue for the CAF team, not ours, and accordingly we sought alternatives to circumvent the lack of data availability.

Our alternative was to use two sites, one synthetic to eliminate uncertainty and the other based on a real watershed that does have observation data. The synthetic domain is a common geometry used in testing hydrologic models referred to as a 3-D Tilted-V watershed. This simple domain has an analytically defined geometry (surface slopes) and uniform hydrologic parameters within three analytically defined regions, meaning that “exactly equivalent” versions of the problem can be constructed at any grid resolution. By exactly equivalent, we mean that no resampling or averaging of parameters is needed, so a model with 10m by 10m cells should have precisely the same output as one constructed with twice as many 5m by 5m cells if there are no grid effects or nonlinear scaling effects. The second site we selected is the Dry Creek Experimental Watershed (DCEW) North of Boise, ID. This location has a climate similar to the Palouse and the monitoring data for the site is openly maintained on a public website. The data includes hydrologic (streamflow and soil moisture) and meteorological data (precipitation, temperature). The site is roughly 36km^2 and has a combination of grassy and forested slopes, with variable slope angles, and good characterization of the soils in the upper 2m of the watershed. Overall, the data at DCEW has the level of support and confidence that typically leads to an accurate integrated model of a site. Conceptually, the synthetic domain is similar to CAF but DCEW is fundamentally different from CAF because it is larger, steeper, and forested instead of farmed. However, the interaction of slopes and vegetation with profoundly different water demands (grasses versus trees) is a more challenging, and broadly transferable, problem with which to test scaling laws and the synthetic domain retains many similarities to CAF.

2. Progress and Results

The PhD student supported by this project has made excellent progress in modeling the systems and analyzing the results. The simulations and analysis of the 3-D Tilted-V case are complete, and these mainly involved running the same benchmark problem for different computational grids and comparing their outputs. The simulations varied the spatial discretization laterally and vertically. The base-case scenario, which comprises the synthetic “true” result, used 1m by 1m cells laterally, and 0.1m think cells vertically. The lateral resolution was then changed 2m, 5m and 10m, respectively, and the simulations re-run. Vertical resolutions included 0.1m, 0.25m, 0.5m and 1m, but all models occupy the same domain volume. Each run used the same boundary conditions and the simulation represented a 12-hour rainstorm followed by a long period for it to drain off. We found that increasing the grid resolution increased the streamflow and also had significant impacts on other portions of the water budget. The largest compensatory effect was a decrease in the volume of saturated groundwater in the system, which was accompanied by small shifts in variably saturated soil moisture. We also observed significant changes in the spatial patterns of overland flow and soil moisture, where larger areas were inundated with surface water in the coarse grid simulations after the storm passed. Presently, the project student is proposing and developing scaling laws to describe these trends (*i.e.* an exponential relationship between streamflow and grid size) and preparing the results for inclusion in a publication. These results were presented at an international conference in December and were well-received by the scientific community.

The second portion of the project is nearing completion. The study at DCEW involves a larger, more complex domain, which requires longer model runtimes, but it also has observation data to consider. A similar approach to exploring scaling behaviors is being used but we limit this to three lateral grid resolutions because of computational limitations. There are several tradeoffs to consider because the model must be calibrated to ensure reasonable reproduction of the data. Our approach is to adopt the finest resolution (20m laterally) as the "truth" and calibrate the model parameters to this scale. However, translating this information to the coarser scales (40m and 60m lateral cell resolution) can be done in several ways. One option is to average the values from the small model to the larger models and another is to independently calibrate each resolution to reproduce the data (the outflow hydrograph) as closely as possible. Since it is unclear which of these gives the fairest comparison, both are being evaluated. The parameter upscaling (averaging) is already complete and we have found that this produces large differences in the magnitude of streamflow, but a similar trend to the 3-D Tilted-V was observed where larger grids gave more surface flow. This result is promising because it suggests that some scaling behaviors related to grid selection may have general trends that describe them, even when the nature of the flow systems is drastically different. Presently, we are completing the multi-scale calibrations and once this is completed we will have all of the simulated data needed to complete the scaling analysis.

3. Remaining tasks and anticipated timeline

The only tasks remaining are to complete the multi-scale simulations at DCEW and to complete the scaling analysis with those results. These simulations are taking longer than expected due to: 1) longer-than expected runtimes of the 20m resolution integrated model, and 2) the difficult, and often unpredictable, transient calibration process. We chose to use a real, specific storm in late October of 2012 for the DCEW study and physically-based hydrologic models have a large number of parameters that interact in complex ways, so these kind of delays during calibration are not unusual, but they also cannot be reliably estimated ahead of time. Once these simulations are completed, the student will work up an analysis of scaling behaviors at DCEW similar to what she has already done for the 3-D Tilted-V, and then a correlation analysis of the trends observed at both sites. She will also propose scaling laws for upscaling or downscaling results from one grid resolution to others and quantify the anticipated variability of the simulations. As these results become available, she will continue to make progress on her manuscript describing these results, for which we anticipate an August 2018 submission date to the Journal of Hydrology. The simulations are currently running, almost around the clock, and we expect them to be completed within the next 3-weeks. We originally intended to directly simulate land-surface process in these simulations, but it became clear that doing so would add too much complexity too soon. These simulations will still be done as part of the PhD students dissertation work but there was insufficient time for them to be considered in this project. Regardless, with further independent testing by other researchers, we expect that the scaling laws we are proposing may be able to provide the most reliable method for describing the range of variability one should expect from a calibrated model run at one scale relative to other scales. These relationships are already showing great promise in their resilience across domains that vary drastically in complexity, so we expect that similar scaling laws will be discovered when land-surface processes are included in the future.