Impacts of Development Pattern on Urban Groundwater Flow Regime

Michael L. Barnes[1], Claire Welty[1,2], and Andrew J. Miller[1,3]

[1]Center for Urban Environmental Research and Education, University of Maryland Baltimore County, Baltimore, MD 21250

[2]Department of Chemical, Biochemical, and Environmental Engineering, University of Maryland Baltimore County, Baltimore, MD 21250

[3]Department of Geography and Environmental Systems, University of Maryland Baltimore County, Baltimore, MD 21250

Corresponding author: Claire Welty (weltyc@umbc.edu)

Key Points

- We evaluated subsurface hydrologic response to development pattern across a gradient of urbanization
- Catchments with lowest amounts of impervious cover showed greatest variability in subsurface storage
- Temporal variability in subsurface storage was most responsive to land cover in near-surface layers and to topographic control at depth

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2017WR022146

Abstract

We compare the effects of urban development type and spatial pattern on the hydrology of six small headwater catchments near Baltimore, Maryland utilizing a three-dimensional coupled groundwater-surface water-land atmosphere model (ParFlow.CLM). The catchments range in size from 0.2 - 2 sq km, across a spectrum of older heavily urbanized development to less developed exurban residential areas. The topography and land cover of each model domain is defined using high-resolution LiDAR topography and orthoimagery. Simulations were conducted at an hourly time-step for calendar years 2012-2015 using a 10-m terrain-following horizontal grid with variable dz (0.1 m to 8 m). Differences in development type and pattern across catchments give rise to complex spatial and temporal trends in the water budget. Catchments with the lowest amounts of impervious cover show the greatest variability in total storage response to climatic variation, whereas those with the greatest amount of impervious cover show less variability in response of subsurface storage to annual and seasonal variability in precipitation input. The storage response among catchments tends to be pronounced during prolonged dry and wet periods, with the variability in response being less pronounced over short-term events. A negative correlation is observed across catchments between impervious cover and net subsurface storage variability. Temporal variability in subsurface storage is most responsive to development pattern in near-surface layers, but transitions to topographic control at depth. Spatially, the development footprint controls where recharge and evapotranspiration occur in the unsaturated zone. Infiltration in pervious areas flows laterally beneath impervious surfaces.

1 Introduction

It is a tenet of urban hydrology that percent impervious cover together with efficiency of the storm-drain network is a fundamental driver of urban hydrologic response to rainfall events. The effects of percent impervious cover and density of the storm-drain network on the hydrograph peak and various measures of the time base of the rising limb and other measures of "flashiness" are well-established (Graf, 1977; Leopold, 1968; Smith et al., 2002; Smith et al. 2005). This simple conceptual model has been reiterated frequently in the literature and has survived repeated analyses, albeit with some modification (Schueler et al., 2009). Urban and suburban watershed runoff response to storm events is also affected by complex spatial patterns of development (Brander et al., 2004; Gironas et al., 2009, 2010; Mejia & Moglen, 2009, 2010), rainfall amount and intensity (Mejia & Moglen, 2009, 2010), and soil moisture storage capacity in pervious areas (Smith et al. 2015).

The impact of urban development on hydrologic response of groundwater systems has received less attention; however, it is recognized that groundwater is an integral component of the urban hydrologic cycle (e.g., Bhaskar et al., 2015; Kaushal & Belt, 2012), particularly in terms of base flow generation for urban streams (Bhaskar et al., 2016) that provide numerous ecosystem services. With the recent emphasis by municipalities on implementing infiltration as a stormwater management technique (e.g., Maimone et al., 2011), understanding the groundwater response to this practice is of growing interest.

Subsurface hydrologic response can be quantified in terms of changes in water table elevation or changes in subsurface storage that includes the unsaturated zone. Water table changes can be observed through field measurements, although monitoring well networks in urban areas not dependent on groundwater use for water supply are relatively sparse, and are logistically challenging to install owing to the presence of underground utility lines. Changes in subsurface storage can be calculated on a lumped basis from water balance components for

areas where data are available. However, the spatial pattern of temporal variability is not generally available from field data, again owing to sparse coverage. And data from one well, or aggregate changes in subsurface storage from water-budget calculations, do not provide information needed to evaluate spatial patterns of system behavior.

Bhaskar et al. (2015) used a simulation approach to systematically examine the influence of urban features on subsurface storage. This simulation and scenario-testing approach was carried out at a regional scale across an urban to rural gradient at relatively coarse pixelation (500 m horizontal, 5 m vertical). The investigators found that reduced vegetation and increased impervious surface area are predicted to have negligible effects on subsurface storage compared to leaking infrastructure, at least for the conditions in the case study presented.

The intent here is to build on this previous simulation work by focusing on the effects of development pattern on groundwater dynamics at the small watershed scale, as well as groundwater response to climatic variability, including seasonal and interannual cycles of wetting and drying. Of interest is the spatial variability of watershed response to fine-scale heterogeneity (10 m horizontal, 0.1 m – 8 m vertical) of land cover and land surface processes. By utilizing a highly-resolved distributed modeling approach, we aim to better capture the linkages between the energy balance and shallow subsurface dynamics. In this paper, we focus in particular on the role of varying spatial patterns of land cover and topography, to address the following questions:

(1) How does subsurface storage across a gradient of development respond to seasonal patterns of wetting and drying?

(2) What is the importance of land cover type and distribution versus topography on subsurface storage variability?

Our objective is to use a three-dimensional coupled groundwater/surface water/land-

surface model applied to headwater catchments spanning a range of development types and patterns, using data in the Baltimore metropolitan region as an illustrative example.

2 Study sites

Six headwater catchments were selected for analysis within the Gwynns Falls watershed, located in Baltimore County, Maryland, USA (Figure 1). The Gwynns Falls is a principal observational watershed of the Baltimore Ecosystem Study (BES) Long-Term Ecological Research project (http://beslter.org). The Gwynns Falls empties to Baltimore Harbor, which in turn drains to the Chesapeake Bay. Average annual precipitation in this region is approximately 1100 mm and is distributed evenly year round, with storm events occurring every 7-10 days (National Weather Service, http://www.weather.gov/lwx/bwinme). The study sites were selected in order to evaluate hydrologic response across a range of development types and ages.

The subwatersheds chosen for analysis (Figure 1b) vary with respect to dominant land-cover type (Figure 2), stormwater management infrastructure, topographic configuration, and underlying geology and soils; these characteristics are summarized in Tables 1 and 2. The subwatersheds range in drainage area from 0.21 to 1.63 km^2. Three of the study sites are located within the Red Run watershed, an exurban suburban area in the Baltimore region where new development follows principles of "smart growth" characterized by cluster development,

preserved open space, and extensive stormwater management. Three other sites are located within the Dead Run watershed, an older, densely-developed inner suburban area where there is little to no stormwater management. The catchments all lie in the Piedmont Physiographic Province. Those in Red Run are underlain by Loch Raven Schist; those in Dead Run are underlain by mafic and ultramafic rocks that interfinger with the Potomac Group of the Atlantic Coastal Plain at the watershed southern boundary. Additional details on the study

area are provided in Supplementary Information.

3 Methods

3.1 Hydrologic model

We used ParFlow.CLM, a three-dimensional groundwater flow code (Ashby & Falgout, 1996; Jones & Woodward, 2001) modified to include fully coupled surface flow (Kollet & Maxwell, 2006) as well as energy and plant processes (Kollet & Maxwell, 2008; Maxwell & Miller, 2005). The code is optimized for parallel computing for large-scale, high-resolution simulations.

We selected a horizontal grid resolution for each basin model of 10 m for the catchment study areas that were on the order of 0.2 to 2 km^2. This resolution captured the spatial extent and characteristics of the developed landscape in good detail, while balancing computational considerations. We utilized a terrain-following grid with a variable vertical discretization (Maxwell, 2013) ranging from 0.1 m - 8 m, representing a vertical thickness of 31 m. We implemented no-flow boundary conditions around the edges of each domain. The resulting model domains contained 69,120 to 505,440 finite difference cells (Table 3). We simulated the four-year period from 1 January 2012 to 31 December 2015 using an hourly simulation time step. Non-pressured pipe flows (buried streams, stormwater pipes) were modeled as streams; we simulated surface flow and groundwater pressure heads and compared to relevant available field data. We did not attempt to model other stormwater infrastructure (ponds, infiltration areas, wetlands) explicitly. Some infrastructure features are accounted for in the terrain data, for example, wetlands and some detention basin locations are implicitly included in the model where the water table intersects the land surface.

3.1.1 Model input data

3.1.1.1 Topographic, land cover, and infrastructure data

We utilized a high-resolution land-cover classification for Baltimore County for the year 2007 (0.61-m (2-ft) resolution) developed by the University of Vermont Spatial Analysis Laboratory. The classification uses Ikonos satellite imagery and LiDAR as input to an object-based image analysis (University of Vermont Spatial Analysis Laboratory (UVM SAL), 2007). We resampled the UVM SAL map to 10-m model grid resolution using the mode of the classification, and then mapped the UVM SAL classes to the those in the Common Land Model International Geosphere Biosphere Programme (CLM IGBP) land-cover classes (Table 4). In addition to vegetation type, we used this mapping product to distinguish between pervious and impervious areas, and assigned top layer hydraulic properties accordingly.

A high-resolution LiDAR DEM (1-m horizontal discretization) provided topographic information for the study areas (see hillslope shading in Figure 3). The DEM was derived by UMBC from LiDAR flown for Baltimore County, MD in 2005. We resampled the LiDAR DEM

data at a resolution of 10 m for model input. ParFlow uses a free surface overland flow boundary condition to simulate integrated groundwater-surface water interactions (Kollet & Maxwell 2006). The overland flow component is simulated using the kinematic wave approximation, in which the diffusion terms of the momentum equation are neglected, leaving the friction slope equal to the bed slope. The model user calculates these slope fields using digital elevation model (DEM) data, and provides them as the input for the ParFlow overland flow simulation. We calculated model slopes using a global slope enforcement approach to ensure flow connectivity in the overland flow grid using the method of Barnes et al. (2016).

We utilized mappings of stream and storm drain networks available from Baltimore County, MD Department of Public Works to provide a continuous drainage layer for the purposes of aiding watershed delineation. In cases where a stream was buried in a stormwater pipe, the pipes were burned into the DEM to represent flow channels.

3.1.1.2 Hydrogeology

We utilized hydrogeologic data characterizing the regional geology reported in the literature (Laughlin, 1966; Nutter & Otton, 1969). Individual well data reported in Laughlin (1966) were digitized and imported into a geospatial database. We set the model initial water table depth condition at 5 m below the land surface for each modeling domain, consistent with regional observations of water level from the period of active well drilling. Many of the wells reported in Laughlin (1966) are no longer in service, as the study area has transitioned to a surface-water municipal water supply; however, the legacy data provide a valuable resource for domain characterization.

We represented the subsurface in each catchment model with four hydrogeologic units (Table 5) composed of: soil or impervious surface, saprolite, a transition zone, and fractured

bedrock. The transition zone between saprolite and bedrock consists of weathered bedrock characterized by high hydraulic conductivity. Each hydrogeologic unit was represented by at least one model layer. Fine discretization of the model in the upper layers yielded highly-resolved vertical output, which allowed detailed analysis of shallow subsurface processes.

3.1.1.3 Forcing data

To populate the CLM land surface model component of the model, we utilized forcing data from the North American Land Data Assimilation System (NLDAS2, https://ldas.gsfc.nasa.gov/nldas/NLDAS2forcing.php) project dataset. These data were downloaded and processed, with data extracted for the nearest $1/8^{th}$ degree NLDAS2 pixel.

The CLM model utilized eight fields from NLDAS2: long wave radiation (DLWR) (W/m^2), short wave radiation (DSWR) (W/m^2), precipitation (APCP) (kg/m^2), temperature (TEMP) (K), specific humidity (SPFH) (kg/kg), pressure (PRESS) (Pa), EW wind component (UGRD) (m/s), and NS wind component (VGRD) (m/s). Precipitation data were converted from hourly kg/m^2 to mm/s. Although one 1/8th degree NLDAS2 pixel of forcing was used for the region, spatial variability of the evapotranspiration process was captured through parameterization of land cover at the 10-m pixelation of the model.

3.1.2 Model initialization

We initialized the model using a two stage process. First, we imposed a constant recharge rate of 305 mm/yr (Swain et al., 2004) for a time period of 2 years, with an initial water table set at a depth of 5 m below the land surface. During this initial period, the CLM component of the model was not activated, and the ParFlow spin-up keys (Maxwell et al., 2016, Section 6.1.34) were switched on. Next, we initialized each model with transient meteorological forcing from the years 2008-2011 from NLDAS2 with the overland flow and CLM components of the model active. Each model was thus spun-up for total of 6 years.

3.1.3 Computational resources

To carry out the simulations, we used computational resources from XSEDE (Towns et al., 2014), with the simulations running primarily on the Stampede supercomputer at the Texas Advanced Computing Center (TACC). Simulations used between 64 and 256 processors (Table 3), and required ~200,000 service unit (SU) hours to complete. Utilizing parallel resources significantly reduced the total wall clock time, and enabled efficient completion of the simulations.

3.2 Evaluation of spatial variability of watershed response

We quantitatively documented watershed pattern and spatial response by calculating semivariograms of two inputs and two outputs using the 10-m pixelated model

data. Semivariogram analysis quantifies the distance over which a spatially-varying attribute is autocorrelated; cross-semivariogram analysis quantifies the distance over which two spatially-varying attributes are cross-correlated and the strength of the cross correlation. For inputs, we calculated semivariograms of: (1) leaf area index (LAI) based on land cover type; and (2) the digital elevation model. For outputs, we calculated semivariograms of (1) a snapshot of change in unsaturated zone storage representing the difference between the initial storage and the driest time step of the model (18 September 2012); and (2) the standard deviation of pressure head at the bottom layer of the model representing water table fluctuation, averaged over all time steps. In addition, we calculated cross-semivariograms between (1) LAI and change in unsaturated zone storage; and (2) DEM and pressure response at depth. Semivariograms and cross-semivariograms were computed using GAMV from the GSLIB software library (Deutsch & Journel, 1998).

The semivariograms of LAI utilized information specified as input to CLM for LAI_{max}. One value was assigned to trees, one value was assigned to grass/shrubs, and a value of 0 was assigned to impervious surfaces. Isotropic (omnidirectional) variograms were calculated for LAI and change in storage, since behavior was independent of direction of analysis. Directional (anisotropic) variograms were calculated for DEM and pressure in directions of maximum and minimum correlation. All (auto) semivariogram output was scaled by the variance so that any existing sills had a value of 1.0.

Cross semivariograms were computed for LAI-ΔS and DEM-pressure. For cross-semivariogram analysis, z-scores of the data were calculated for use in the analysis owing to the disparate units of measure of the paired variables. The resulting sills of the cross-semivariogram functions of the z-score data can range from -1 to +1, with +1 indicating perfect correlation and positive values less than 1 indicating the strength of the positive correlation.

To determine the distance over which each variable was autocorrelated or pairs of variables were cross correlated ("effective range" using Deutsch and Journel (1998) terminology), the exponential variogram model was fit to experimental semivariograms of LAI, change in storage, and cross semivariograms of LAI-ΔS. The Gaussian variogram model was fit to experimental cross-semivariograms of DEM-pressure head. For descriptions of variogram models, see Deutsch and Journal (1998, page 25). In addition to the effective range, the model fits were used to determine the best-fit values of the nugget, and in the case of the cross-semivariograms, the sill.

3.3 Hydrologic monitoring

A privately-owned bedrock well in the region no longer in use (location shown in Figure 1b) was outfitted with a continuously-recording pressure transducer (In-Situ Level Troll 500) installed ~ 1 m below the summertime water table elevation. Permission was obtained from the owner to utilize this well in our research. Pressure head was recorded at 1-hour intervals. Pressure head was converted to a continuous record of depth-to-water below land surface. Manual depth-to-water measurements were collected to correct for pressure transducer drift. Data were manually downloaded every 60 days to a field laptop, processed, and subsequently transferred to a database at UMBC. Observed depth-to-water data were used to compare to model output.

A USGS stream gaging station was available at the outlet of one of our study domains (DR5). The site is instrumented with a staff gage and an Accububble pressure measurement device. Stage is recorded at a frequency of 5 minutes. A rating curve has been developed relating stream discharge to stage by making manual streamflow measurements every 8 weeks, under a range of flow conditions. Stage and discharge data are telemetered to USGS hourly and served in near-real time at https://waterdata.usgs.gov/md/nwis/uv?site_no=01589312.

4 Results and Discussion

Model output consisted of pressure head and saturation for every cell in the 6 domains (61,440 to 505,440 cells per domain; see Table 3) for hourly time steps over the four years of simulation time (35,040 time steps total). Derived outputs were calculated to compare to field data: here we calculated subsurface storage scaled by watershed area, which can be compared to the pattern of observed water table fluctuations, and stream discharge, which can be compared to observations for conditions up to bankfull flow. Results are presented in terms of averages over time or space, and spatial snapshots in time.

We do recognize the role of parameter uncertainty in the results. We view our variable land cover and topography across the six domains as representing a range of patterns and density of urban development as well as natural catchment variability. We held the geologic and meteorological forcing data as constant over space so as to be able to examine the influence of the spatially variable land cover and topography on model results. There is uncertainty in subsurface properties that we were unable to quantify given the available data.

4.1 How does subsurface storage across a gradient of development respond to seasonal patterns of wetting and drying?

4.1.1 Subsurface storage

4.1.1.1 Simulation data

Hourly mean-removed subsurface storage scaled by watershed area illustrates event-scale, seasonal, annual and interannual variability over the 6 study domains for the 2012-2015 simulation period (Figure 4). Across all domains, the observed high-frequency temporal variability results from recharge events and from diel fluctuations in evapotranspiration.

Superimposed on the high-frequency fluctuations is an annual pattern of highs and

lows. Subsurface storage reaches a minimum value at mid to late September; the maximum value occurs between mid-May and mid-June, although the pattern of highs is not as well defined interannually compared to the pattern of low values for the four years simulated. Maximum subsurface storage corresponds to late spring hydrologic conditions before full leaf-out; minimum subsurface storage corresponds to the end of the growing season before leaf fall. This pattern illustrates the expected effect of evapotranspiration on annual cycles of subsurface storage, when evapotranspiration is lowest and recharge is highest during leaf-off periods.

The simulation results also illustrate the interannual variability in subsurface storage as a function of climatic conditions. The four-year modeling time period captures extremes in wet and dry conditions. 2012 was the driest dry year overall, as reflected by double the change in minimum subsurface storage for 2012 (-100 mm) compared to the minimum for the wettest year 2014 (-50 mm). In 2012, there is no well-defined seasonal high in subsurface storage. This results from the relatively low precipitation in spring 2012 coupled with the carry-over effects of an extremely dry autumn 2011 period. This persistence of a relatively dry period in most of 2012 was quickly reversed by Hurricane Sandy (at the end of October 2012), characterized by 182 mm of precipitation over a 48-hour period. The subsurface storage response changed from
-70 mm to +40 mm in response to this event, which was nearly as much as the annual variability in other more typical years (e.g., 2015). At the other extreme, the spring 2014 shows the greatest maximum change in subsurface storage across all four years of simulation. This aberration is due to occurrence of an unusually large, long-duration event at the end of April 2014 (135 mm of
precipitation over 4 days) in an already-wet spring. Again, we see that this effect is magnified for the less developed watersheds and muted for the most developed watersheds.

This general pattern of annual highs and lows in subsurface storage is similar over the six study areas, but the magnitude of temporal variability ranges across the domains. Sunnydale and Runnymeade, the two catchments with the lowest amounts of impervious cover, show the greatest variability in total storage response to climatic variation. Kevsway, DR5, Greengage and Redbrook, all with > 42% impervious cover, show less variability in response of subsurface storage to annual and seasonal variability in precipitation input. The variability in storage responses among the catchments tends to be pronounced during prolonged dry periods and prolonged wet periods, with the variability in response being less pronounced over short-term events.

Using percent impervious surface area as a metric, the variability in subsurface storage consistently is reduced as the percent impervious surface area is increased, across the six study domains. A plot of the standard deviation in subsurface storage over all time steps for each domain as a function of percent impervious surface area reveals a strong negative linear correlation between these two parameters (Figure 5). The simulated storage responses across a spectrum of development illustrate the net effect of the competing processes of reduced infiltration (and increased runoff) and lower evapotranspiration (due to reduced vegetative cover) for high-percentage impervious surface area and the opposite effect for low-percentage of impervious area.

In order to simplify the comparison of temporal behavior across watersheds, we calculated the mean daily subsurface storage, averaged these values across Julian day for the four simulation years, and then smoothed the results for each domain (Figure 6a). The NCAR Command Language (NCL) function *smthClmDayTLL* was used to calculate the smoothed values, utilizing a fast Fourier transform technique (The NCAR Command Language, 2017). The resulting plot further highlights predicted annual differences in subsurface storage across the study domains as a function of impervious surface area, with a smoothing out of the

interannual variability caused by wet vs dry years over the simulation period. The variability in mean daily subsurface storage again appears to be correlated to impervious surface area.

Figure 6a also illustrates a slight lag in the annual maxima and minima in subsurface storage across watersheds: the spring peak ranges from Julian day 160 (9 June) for least-developed Sunnydale to Julian day 168 (17 June) for the most-developed Kevsway domain. Conversely, the annual minimum occurs earlier for Kevsway (Julian day 258, 15 September) than for Sunnydale (Julian day 261, 18 September). The dates for maximum and minimum storage for the other four domains are temporally ordered between these endpoints as a function of percent impervious surface area. The net result is a shortening of the time period between the annual minimum and maximum mean daily subsurface storage by 11 days for the most developed watershed compared to the least developed watershed.

A final point to be made regarding the results for mean subsurface storage is that the annual minimum is single-valued and more pronounced than the annual maximum values. The annual maxima are spread out over multiple peaks. This multi-peaked mean behavior is due to significant precipitation events or dry periods during each of the four years of record (Figure 4) that affect the four-year mean values (Figure 6a). It is expected that the multi-peaked nature of the annual maximum values would converge to a well-defined single value over more years of simulation data. The subtle bump in the receding limb of the annual low data around Julian day 300 is due to the influence of Hurricane Sandy from 28 October 2012 to 30 October 2012.

In summary, we have illustrated a pattern of changes in simulated subsurface storage across a gradient of development. The variability in subsurface storage is greatest for the least urbanized catchment. This results from a combination of low impervious surface area and high amount of vegetated land surface, which together provide greater opportunity for infiltration/ recharge as well as evapotranspiration.

4.1.1.2 Comparison to field data

Two sets of field data are available for comparison with the subsurface storage simulation results. One data set (Figure 6b) consists of mean monthly subsurface storage calculated from water balance components over a period of 77 years (Nutter & Otton, 1969) from data collected at the nearby Gunpowder Falls watershed (Figure 1a). As expected, averaging the water balance data set over 77 years smooths out deviations from the long-term mean behavior. This data set clearly shows the annual cycle of highs and lows for the region: in this depiction of mean monthly averages, the annual maximum mean monthly subsurface storage occurs at Julian day 90 (31 March) and the minimum occurs at Julian day 273 (30 September).

Hourly depth-to-water data were collected during the study period at a bedrock well that is located approximately 3 km from the Kevsway site (star in Figure 1a), averaged over Julian day, and smoothed (Figure 6c) over the four-year period of simulation. The annual maximum and minimum depth to water values occur at 91 and 275 days, respectively, which is similar to the long-term pattern exhibited by Figure 6b. Similar to the subsurface storage simulation data for this four-year period (Figure 6a), the annual mean high and low values of depth-to-water are affected by a few extreme events, whereas over the long-term it would be expected that this multi-peaked behavior would converge to one maximum and one minimum value as shown in Figure 6b.

Overall, in comparing the simulation data to these observations, we see that the general pattern of observed annual highs and lows is well captured by the model output. What the field data do not indicate that is captured by the model is the subtle predicted differences in subsurface storage across a gradient of development. Site-specific data are not available to corroborate this behavior.

4.1.2 Stream discharge

Comparisons were made between mean daily stream discharge simulated at the watershed outlet of DR5, and a subset of USGS discharge data available for that location (Figure S1). In general, we found that streamflow was overpredicted by the simulations, but that the timing and peak values of storm discharge were in good agreement between the simulated and observed data. The R^2 correlation coefficient for the period of 2012 evaluated was 0.62, which indicates a reasonable representation of the field data by the model. Streamflow was overpredicted for a number of reasons, some having to do with uncertainty in the data and some having to do with model constraints.

In terms of data uncertainties, according the USGS web site for this station (https://waterdata.usgs.gov/md/nwis/uv?site_no=01589312), up to 25% of the watershed area is non-contributing, meaning that a portion of the delineated watershed does not drain to the stream network. This implies a smaller effective drainage area. If this smaller value were to be used in Figure S1 to scale discharge, the observed runoff values would be increased by 25%. Second, some groundwater may bypass the stream gage owing to the size of the watershed (groundwater divides may not be coincident with surface water divides), and leakage into sanitary sewer pipes crossing beneath the stream (Bhaskar et al., 2012; Bhaskar et al., 2016). This can lead to decreases in observed runoff. Third, the stage-discharge rating curve is adjusted by USGS multiple times per year following field observations of changes in downstream controls; low flows are particularly sensitive to these changes. For 2012, rating shifts at DR5 resulted in streamflow discharge values varying by up to an order of magnitude for low values of stage (0.1 – 0.3 m) and discharge (0.0008 – 0.3 m^3/s).

In terms of the model, there is a disparity in scale between the actual stream geometry and modeled stream geometry. The streams are physically quite small, varying from about 0.5 m to 2 m in width. We chose horizontal model discretization of 10 m x 10 m to balance

computational burden with representation of land cover and topography in adequate enough detail. The fact that the model discretization is larger than the stream width can contribute to difficulty in simulating small values of base flow (< 0.03 m^3/s). More specifically, discharge below a given threshhold will not be simulated with accuracy because the baseflow-generation process begins to occur in the catchment at a scale finer than the model grid. This can lead to overestimation of (erroneously high) base flows as a model result.

A second model issue is that the coupled model may be undersimulating (underestimating) evapotranspiration for watersheds having high percentage of impervious surface area. ET could occur in reality in impervious areas from cracks in parking lots, or in parking lots incorporating trees with roots penetrating beneath impervious surfaces, both of which could contribute to more ET than occurs from the modeled impervious land cover type. If we are undersimulating ET, this water would ultimately exit from the watershed via streamflow, leading to higher modeled streamflow values than would be the case in reality.

The items discussed above contribute to uncertainty (modeled and observed) in the runoff sink term of the water balance for these small urban catchments, making it particularly difficult to accurately estimate this component. The comparison here does indicate a biased model overestimation of runoff on an annual basis, but the magnitude of this overestimation is not known with precision at the scale at which we are working, owing to the complicating factors described. Baseflow overestimation bias could potentially influence the slope of modeled subsurface storage recession, but we did not observe this influence in Figure 6.

4.2 What is the importance of land cover type and distribution versus topography on subsurface storage variability?

Here we examine snapshots or averages of spatial output and evaluate the drivers of the observed spatial patterns.

4.2.1 Spatial variability of subsurface pressure fields

For each domain, the pressure head for each model cell was statistically compared over all time steps to evaluate the long term patterns and correlation to other physical variables. Figure 3 illustrates the standard deviation of the modeled pressure head over all hourly time steps for four years of simulation time at the bottom model layer (23 - 31 m depth) of the simulated domains. The standard deviation in pressure head at the base layer indicates the average fluctuation in water table elevation above. Values for the pressure-head standard deviation range from a maximum of 0.21 m to 0.58 m across the basins. The result is superimposed on the DEM of the watersheds. This figure reveals that DEM and the standard deviation of pressure head at depth exhibit similar spatial patterns. The patterns generally show that the greatest variability in pressure head occurs beneath the ridges and the least amount of variability occurs beneath the valley bottoms. The lower pressure-head variability in the valley bottoms results from drainage by the streams modulating the magnitude of water table changes at these locations. The observed pattern appears to be the case regardless of the spatial distribution of development and the percent impervious surface area of the domain. Generally, catchments with lower overall storage variability (Figure 4) have lower maximum values of standard deviation in pressure head in the hilltops.

In an attempt to quantify this visual observation, we calculated semivariograms of DEM and pressure head variability at depth in directions of maximum and minimum correlation (not shown), and the cross-correlation between these two variables using cross semivariogram analysis (also not shown). It is not surprising that the DEM exhibits maximum, unbounded correlation along the valley axis, and minimum but finite autocorrelation on the cross-valley perpendicular axis. The pressure head variability at depth generally follows this pattern. Cross-semivariogram analysis of z scores shows a similar order-of-magnitude effective range for the cross semivariograms of DEM and at-depth

pressure variability across domains, with no apparent relation to degree of development (percent and pattern of impervious surface area). The strength of the positive cross-correlation, as indicated by the value of the cross-semivariogram sill, ranged from 0.40 (Redbrook and Runnymeade) to 0.73 (Kevsway), also with no relation to the percent or pattern of developed land. This analysis did confirm quantitatively what we observed visually in Figure 3 – that there is a positive cross-correlation between DEM and pressure variability at depth, and that this appears to be the case regardless of the spatial distribution of development and the percent impervious surface area of the domain.

4.2.2 Spatial variability of subsurface storage change

We calculated a metric of maximum decrease in storage across all domains by determining the change in storage between the first simulation time step and the time step at which the greatest negative departure from mean storage occurred. This driest time step (6268) occurred during 18 September 18 2012 (Figure 4). The calculation was carried out for the fifth model layer beneath the land surface (i.e., at a 1-m depth) representing the subsurface zone immediately below the upper one meter of soil or impervious cover. Snapshots of this calculated change in subsurface storage are shown for the Runnymeade domain in Figure 7; the same depiction for all domains is shown in Figure S2.

The subsurface change in storage under dry conditions is due to the daily pattern of evapotranspiration. The pattern of subsurface storage change appears to be strongly related to the vegetation pattern (Figure 2), where the change in storage is greater beneath larger patches of vegetation. It is also related to the pattern of impervious surface coverage. Areas where water cannot infiltrate exhibit less storage change because they do not fully wet up during rain events and evapotranspiration is impeded at these locations.

To quantify the spatial correlation of this observed pattern, we calculated separately the isotropic (omnidirectional) semivariograms of leaf area index and change in subsurface

storage for each model domain, and cross-semivariograms between these two variables for each model domain (Figures S3, S4; Tables S1-S3). Cross-semivariogram analysis between LAI and ΔS

showed a positive, although weak, correlation, with sill values ranging from 0.13 to 0.28. The largest two sill values are associated with two of the most developed watersheds (DR5 and Greengage), whereas the smallest sill value is associated with one of the least developed watersheds (Runnymeade). The distance over which LAI is correlated with ΔS (effective range) for the 6 watershed domains varied from 54 m to 330 m, where the smaller effective ranges are associated with watersheds having the greatest percentage of vegetation, and where the pattern

exhibits fine-scale heterogeneity.

4.2.3 Spatial variability of infiltration patterns

To visualize the effect of development pattern on infiltration into the subsurface, we present an animation of model results for change in storage of the Runnymeade domain during the simulation period for Hurricane Sandy, from model time step 7224 (midnight 28 October 2012) to 7344 (midnight 2 November 2012) (Figure 8). This animation represents change in storage between time step 7223 and the labelled time step.

From this animation, the progression of the three-dimensional nature of the infiltration process, and how the pattern of impervious surface area interacts with this process, can be observed. During precipitation events, the differential pattern of wet-up of pervious vs impervious patches at the land surface is apparent. Water enters the subsurface vertically through pervious/vegetated patches, and then migrates laterally underneath impervious surface patches. The rising water table response in terms of increase in subsurface storage is apparent. Toward the end of the simulation, an increase in storage is also seen toward the bottom of the model, deep in the saturated zone. This is due to increases in storage resulting

from increases in pressure head.

Although not discernable in the visualization, impervious patches impede evapotranspiration of water migrating under them owing to both lack of vegetation to remove water through vegetative pumping and reduced pore space open to the land surface whereby evaporative processes act to remove water from the soil pores. This results in subsurface storage being retained for longer beneath impervious patches.

5 Summary and Implications

We have demonstrated how the pattern of urbanization in six study catchments creates complex spatial and temporal trends in the underlying groundwater resources across a gradient of development type and intensity. Our simulations illustrate that impervious surfaces modulate the location of recharge and the overall magnitude of subsurface storage changes, with the water table aquifer integrating the local groundwater response at the catchment scale. Although the approach was applied to Baltimore suburbs as a case study in order to use real data for the simulation demonstrations, the same method could be applied to any area where fine-scale DEM, land cover, meteorological forcing, and geologic input data are available.

Our findings can be summarized as follows:

1. Using a highly-resolved coupled groundwater-surface water-land surface simulation approach applied to small urban catchments, we were able to capture the effects of spatial variability of development patterns and the temporal effects of climatic variability on subsurface storage across a gradient of urbanization. A novelty of the approach was use of fine-scale land cover and DEM coupled with regional-scale subsurface properties and meteorological forcing.

2. The overall annual cycle of changes in subsurface storage compared well to limited

regional observational data available for comparison to the simulations.

3. The modeling approach revealed spatial variability in temporal subsurface hydrologic response across a range of time scales. Diel and storm-scale responses were well captured.

Mean-removed subsurface storage showed greatest variability in the least developed watersheds and least variability in the most developed watersheds. Negative correlation between hourly subsurface storage and percent impervious surface area was strong (R^2 = .95).

4. Impervious surface area is an important driver in both increasing runoff during storm events and thus diverting water away from recharge areas, and decreasing evapotranspiration (due to less vegetative cover) during seasonal recession periods and during longer-term droughts, with both effects reducing the variability in subsurface storage.

5. At the storm scale, simulations showed that water on upland surfaces infiltrates through pervious zones and then migrates laterally beneath impervious patches. Impervious patches mediate the storage response by reducing evapotranspiration from unsaturated zone patches beneath impervious cover.

6. In the shallow subsurface, land cover patterns had the greatest influence on subsurface hydrologic response, whereas at depth, the topographic signal was the most important control on hydrologic response.

Our findings have implications for future scientific investigations and for policy applications. First, the use of a three-dimensional distributed model provides insights into the spatial variability of subsurface hydrologic response that would not be discernible with a two-dimensional approach. Documentation that the infiltration process must be recognized as being truly three-dimensional and not two-dimensional could have implications for green infrastructure design for stormwater management, where infiltrated stormwater has a horizontal component to subsurface flow. Second, three-dimensional heterogeneous

hydrologic response fields such as those generated by this work provide rich data sets far beyond what is currently measureable in a distributed sense in any watershed-scale landscape. The simulation output could be used for hypothesis testing via deployment of field instrumentation to determine whether the simulated patterns can be observed, or mined to determine where to optimally place observation wells for monitoring purposes.

Finally, the work here should provide further impetus in addition to the work of many before us to collect all hydrogeologic and related data into a national database, so that researcher best efforts are spent on data interpretation rather than data manipulation and assimilation. This kind of effort has been achieved for the atmospheric science community (WRF) and more recently progress has been made for surface water community (National Water Model), but the day has yet to arrive where groundwater models such as ParFlow.CLM can automatically be populated from standardized data sets. Given the increasing emphasis on earth system science and the need to consider the water cycle as a whole, this kind of capability is critical.

Acknowledgments

This work was supported by National Science Foundation grant CBET-1058038 as part of NSF's Water Sustainability and Climate program. This work utilized the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. The data used in this study (NLDAS2) were acquired as part of the mission of NASA's Earth Science Division and archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC). This project was carried out in collaboration with the Baltimore Ecosystem Study Long Term Ecological Research project, supported by NSF grant DEB-1027188. We thank John Kemper and Erin Stapleton for managing instrumentation and data associated with the observing network. Joshua Cole and John Lagrosa provided database management and GIS

support. USGS was funded by the WSC grant to carry out data QA/QC and publication of the DR5 discharge data; we thank Ed Doheny for leading that effort. Model input data are archived at http://doi.org/10.4211/hs.7aef846fbdaf40acb00642768a2e7bde; groundwater observational data can be downloaded from http://userpages.umbc.edu/~weltyc/DRGC_well_data.zip.

References

Ashby, S. F., & Falgout, R. D. (1996). A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations. *Nuclear Science and Engineering*, *124(*1), 145-149. https://doi.org/10.13182/NSE96-A24230

Barnes, M.L., Welty. C., & A.J. Miller. (2016). Global topographic slope enforcement to ensure connectivity and drainage in an urban terrain. *Journal of Hydrologic Engineering*, 21(4). https://doi.org/10.1061/(ASCE)HE.1943-5584.0001306

Bhaskar, A.S., & Welty, C. (2012). Water Balances Along an Urban-to-Rural Gradient of Metropolitan Baltimore, 2001–2009, *Environmental & Engineering Geoscience*, 18(1), 37-50, https://doi:10.2113/gseegeosci.18.1.37

Bhaskar, A.S., Welty, C., Maxwell, R. M., & A.J. Miller. (2015). Untangling the effects of urban development on subsurface storage in Baltimore. *Water Resources Research*, 51, 1158-1181. https://doi.org/10.1002/2014WR016039

Bhaskar A.S., Beesley, L., Burns, M. J., Fletcher, T. D., Hamel, P., Oldham, C.E., & Roy, A.H. (2016). Will it rise or will it fall? Managing the complex effects of urbanization on base flow. *Freshwater Science*, 35(1), 293-310. https://doi.org/10.1086/685084

Brander, K.E., Owen, K. E., & Potter, K. W. (2004). Modeled impacts of development type on runoff volume and infiltration performance. J*ournal of the American Water*

Resources Association, 40(4), 961–969. https://doi.org/10.1111/j.1752-1688.2004.tb01059.x

Deutsch, C. V., & Journel, A. G. (1998), *GSLIB: Geostatistical Software Library and User's Guide, 2nd ed.* New York: Oxford University Press.

Gironas, J., Niemann, J. D., Roesner, L. A., Rodriguez, F., & Andrieu, H. (2009). A morpho-climatic instantaneous unit hydrograph model for urban catchments based on the kinematic wave approximation. *Journal of Hydrology*, *377*(3-4), 317–334. https://doi.org/10.1016/j.jhydrol.2009.08.030

Gironas, J., Niemann, J. D., Roesner, L.A., Rodriguez, F. & Andrieu, H. (2010). Evaluation of methods for representing urban terrain in storm-water modeling. *Journal of Hydrologic Engineering* 15(1). https://doi.org/10.1061/(ASCE)HE.1943-5584.0000142

Graf, W. L. (1977). Network characteristics in suburbanizing streams. *Water Resources Research, 13*(2), 459–463.

Jones, J. E., & Woodward, C. S. (2001). Newton–Krylov-multigrid solvers for large-scale, highly heterogeneous, variably saturated flow problems. *Advances in Water Resources, 24*(7), 763-774. https://doi.org/10.1016/S0309-1708(00)00075-0

Kaushal, S. S., & Belt, K. T. (2012). The urban watershed continuum: evolving spatial and temporal dimensions. *Urban Ecosystems*, *15*(2), 409–435. https://doi.org/10.1007/s11252-012-0226-7

Kollet, S. J., & Maxwell, R. M. (2006). Integrated surface–groundwater flow modeling: free-surface overland flow boundary condition in a parallel groundwater flow model. *Advances in Water Resources, 29*(7), 945-958. https://doi.org/doi:10.1016/j.advwatres.2005.08.006

Kollet, S. J., & Maxwell, R. M. (2008). Capturing the influence of groundwater dynamics on

land surface processes using an integrated, distributed watershed model. *Water Resources Research, 44*(2), W02402. https://doi.org/10.1029/2007WR006004

Laughlin, C.P. (1966). Records of wells and springs in Baltimore County, Maryland (Basic Data Report 1). Baltimore, MD: Maryland Geological Survey.

Leopold, L. B. (1968), Hydrology for urban land planning--A guidebook on the hydrologic effects of urban land use (Geological Survey Circular 554). Washington, D.C.: U.S. Geological Survey.

Maimone, M., O'Rourke, D. E., Knighton, J. O., & Thomas,C. P. (2011). Potential impacts of extensive stormwater infiltration in Philadelphia. *Environmental Engineer: Applied Research and Practice, 14*, 2–12.

Maxwell, R.M. (2013). A terrain-following grid transform and preconditioner for parallel,large-scale, integrated hydrologic modeling. *Advances in Water Resources, (53)*, 109–117. https://doi.org/10.1016/j.advwatres.2012.10.001

Maxwell, R. M., & Miller, N. L. (2005). Development of a coupled land surface and groundwater model, *Journal of Hydrometeorology*, *6*(3), 233-247. https://doi.org/10.1175/JHM422.1

Maxwell, R.M., Kollet, S.J., Smith, S.G., Woodward, C.S., Falgout, R.D., Ferguson, I.M., Engdahl, N., Condon, L.E., Hector, B., Lopez, S.R., Gilbert, J., Bearup, L., Jefferson, J., Collins, C., de Graaf, I., Prubilick, C., Baldwin, C., Bosl, W.J., Hornung, R., & Ashby, S. (2016). ParFlow user's manual (Integrated GroundWater Modeling Center Report GWMI 2016-01). Golden, CO: Integrated GroundWater Modeling Center, Colorado School of Mines.

Mejía, A. I., & Moglen, G. E. (2009). Spatial patterns of urban development from optimization of flood peaks and imperviousness-based measures. *Journal of Hydrologic Engineering*, *14*(4), 416-424. https://doi.org/10.1061/(ASCE)1084-

0699(2009)14:4(416)

Mejía, A. I., & Moglen, G. E. (2010). Spatial distribution of imperviousness and the space-time variability of rainfall, runoff generation, and routing. *Water Resources Research, 46*, W07509. https://doi.org/10.1029/2009WR008568

Nutter, L. J., & Otton, E. G. (1969). Groundwater occurrence in the Maryland Piedmont (Report of Investigations 10). Baltimore, MD: Maryland Geological Survey.

Schueler, T. R., Fraley-McNeal, L., & Cappiella, K. (2009). Is impervious cover still important? Review of recent research. *Journal of Hydrologic Engineering, 1*4(4), 309-315. https://doi.org/10.1061/(ASCE)1084-0699(2009)14:4(309)

Smith, J. A., Baeck, M. L., Morrison, J. E., Sturdevant-Rees, P., Turner-Gillespie, D. F., & Bates P. D. (2002). The regional hydrology of extreme floods in an urbanizing drainage basin. *Journal of Hydrometeorology, 3*(3), 267-282. https://doi.org/10.1175/1525-7541(2002)003<0267:TRHOEF>2.0.CO;2

Smith, J. A., Miller, A. J., Baeck, M. L., Nelson, P. A., Fisher, G. T., & Meierdiercks, K. L. (2005). Extraordinary flood response of a small urban watershed to short-duration convective rainfall. *Journal of Hydrometeorology, 6*(5), 599-617. https://doi.org/10.1175/JHM426.1

Smith, B. K., Smith, J. A., Baeck, M. L., & Miller, A. J. (2015). Exploring storage and runoff generation processes for urban flooding through a physically based watershed model. *Water Resources Research, 51*(3), 1552–1569. https://doi.org/10.1002/2014WR016085

Swain, L.A., Mesko, T. O., & Hollyday, E. F. (2004). Summary of the hydrogeology of the Valley and Ridge, Blue Ridge, and Piedmont Physiographic Provinces in the Eastern United States (Professional Paper 1422-A). Reston, VA: U.S. Geological Survey.

The NCAR Command Language (Version 6.4.0) [Software]. 2017. Boulder, Colorado:

UCAR/NCAR/CISL/TDD. http://dx.doi.org/10.5065/D6WD3XH5

Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D., Peterson, G. D., Roskies, R., Scott, J. R., & Wilkins-Diehr, N. (2014). XSEDE: Accelerating Scientific Discovery. *Computing in Science and Engineering*, *16*(5), 62-74. https://doi.org/10.1109/MCSE.2014.80

University of Vermont Spatial Analysis Laboratory. (2007). Landcover - Baltimore County/metro area, Maryland. http://www.pasda.psu.edu/uci/DataSummary.aspx?dataset=3156

Table 1. Land cover classification of modeled sites

Site Name	Watershed nested within	Basin drainage area (km^2)	Model domain area (km^2)	Model domain % vegetation		Model domain % impervious		
				% Tree canopy	% Grass/ shrubs	% Buildings	% Road/ rail	% Other/ paved
Sunnydale	Red Run	0.23	0.46	86.4		13.8		
				56.7	29.7	5.3	6.3	2.2
Runnymeade	Red Run	0.62	1.23	77.0		22.3		
				40.0	37.0	7.8	9.3	5.2
Redbrook	Red Run	0.21	0.33	57.2		42.0		
				14.0	43.2	6.4	6.8	28.8
DR5	Dead Run	1.63	3.37	52.6		47.4		
				31.0	21.6	12.4	14.2	20.8
Greengage	Dead Run	0.46	0.92	43.0		57.0		
				25.5	17.5	14.0	11.8	31.2
Kevsway	Dead Run	0.24	0.52	29.7		70.4		
				18.8	10.9	15.3	10.3	44.8

* Land cover classes “bare soil” and “water” are not shown; sums exceeding 100% are due to rounding

Table 2. Development characteristics of modeled sites

Site Name	**Median devel-opment date**	**Develop-ment type**	**Storm-water management**	**Water/ waste-water**	**Topographic characterization:** Slope coefficient of variation
Sunny-dale	1967	Residential	None	Wells and septic	0.63
Runny-meade	1996	Residential, high density, new	Detention ponds, wetland	Municipal water and sanitary sewer	0.81
Redbrook	1991	Commercial	Detention treatment train (detention ponds, infiltration basins, wetland)	Municipal water and sanitary sewer	0.99
DR5	1962	Mixed residential/ commercial	None	Municipal water and sanitary sewer	0.76
Green-gage	1962	Mixed residential/ commercial	Extended detention	Municipal water and sanitary sewer	0.72
Kevsway	1987	Mixed residential/ commercial	Infiltration basin, detention pond	Municipal water and sanitary sewer	0.87

Table 3. Cell counts and extents of model domains.

Domain name	nx	ny	nz	dx (m)	dy (m)	dz (m)	Number of model cells	Number of processors utilized
Sunnydale	48	96	15	10	10	0.1 - 8	69120	256
Runnymeade	128	96	15	10	10	0.1 - 8	184320	256
Redbrook	64	64	15	10	10	0.1 - 8	61440	256
DR5	216	156	15	10	10	0.1 - 8	505440	64
Greengage	88	104	15	10	10	0.1 - 8	137280	64
Kevsway	72	72	15	10	10	0.1 - 8	77760	64

Table 4. Mapping of UVM SAL land cover class to CLM IGBP land cover class.

UVM SAL Land Cover Class	CLM IGBP Land Cover Class
Tree Canopy	Deciduous Forest
Grass/Shrubs	Grasslands
Water	Water Bodies
Building	Bare Soil
Road Rail	Bare Soil
Other Paved	Bare Soil
Bare Soil	Bare Soil

Table 5. Hydrogeologic properties assigned to model domain.

Hydrogeologic unit	Depth below land surface (m)	Layer thicknesses comprising unit (m)	Hydraulic conductivity (m/hr)	Porosity (-)
Soil	0 – 1	0.1, 0.2, 0.2, 0.5	0.0227	0.45
Impervious Cover	0 - 1	0.1, 0.2, 0.2, 0.5	0.00212	0.05
Saprolite	1 - 14	0.5, 0.5, 1.5, 1.5, 2.0, 2.0, 2.0, 4.0	0.00556	0.45
Transition Zone	14 – 17	4.0	0.227	0.45
Bedrock	17 – 31	4.0, 8.0	0.00001	0.05

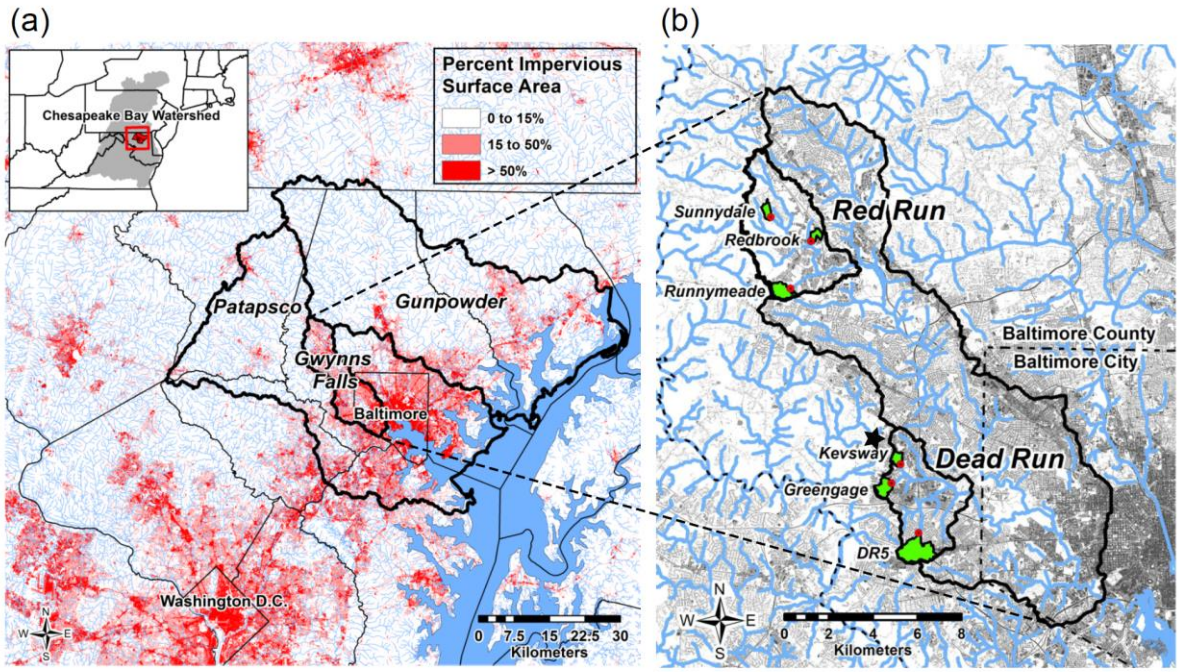

Figure 1. (a) The Baltimore metropolitan region; (b) locations of headwater study sites (shown in green) within Red Run and Dead Run subwatersheds. Red dots indicate stream gage locations. The black star indicates the location of the DRGC bedrock monitoring well.

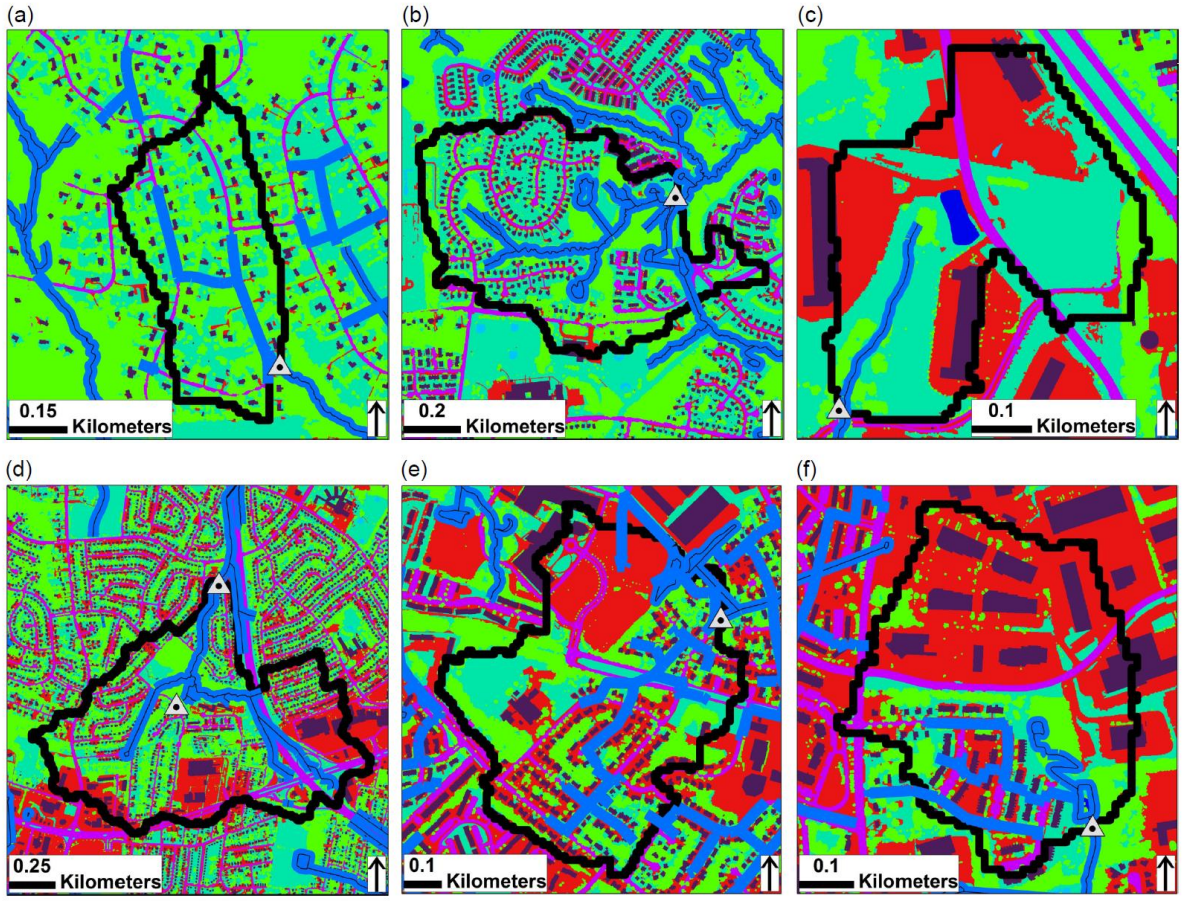

Figure 2. Land cover of six study watersheds: buildings (dark purple); roads (purple); parking lots (red); trees (green); grass (blue-green); water bodies (blue), based on University of Vermont land cover classification. Watershed boundaries are indicated in black; water courses (streams, pipes, and detention basin outlines) are indicated in blue. Watersheds are designated as: (a) Sunnydale, (b) Runnymeade, (c) Redbrook, (d) DR5, (e) Greengage, and (f) Kevsway. The domains are ordered in terms of increasing percent impervious surface area.

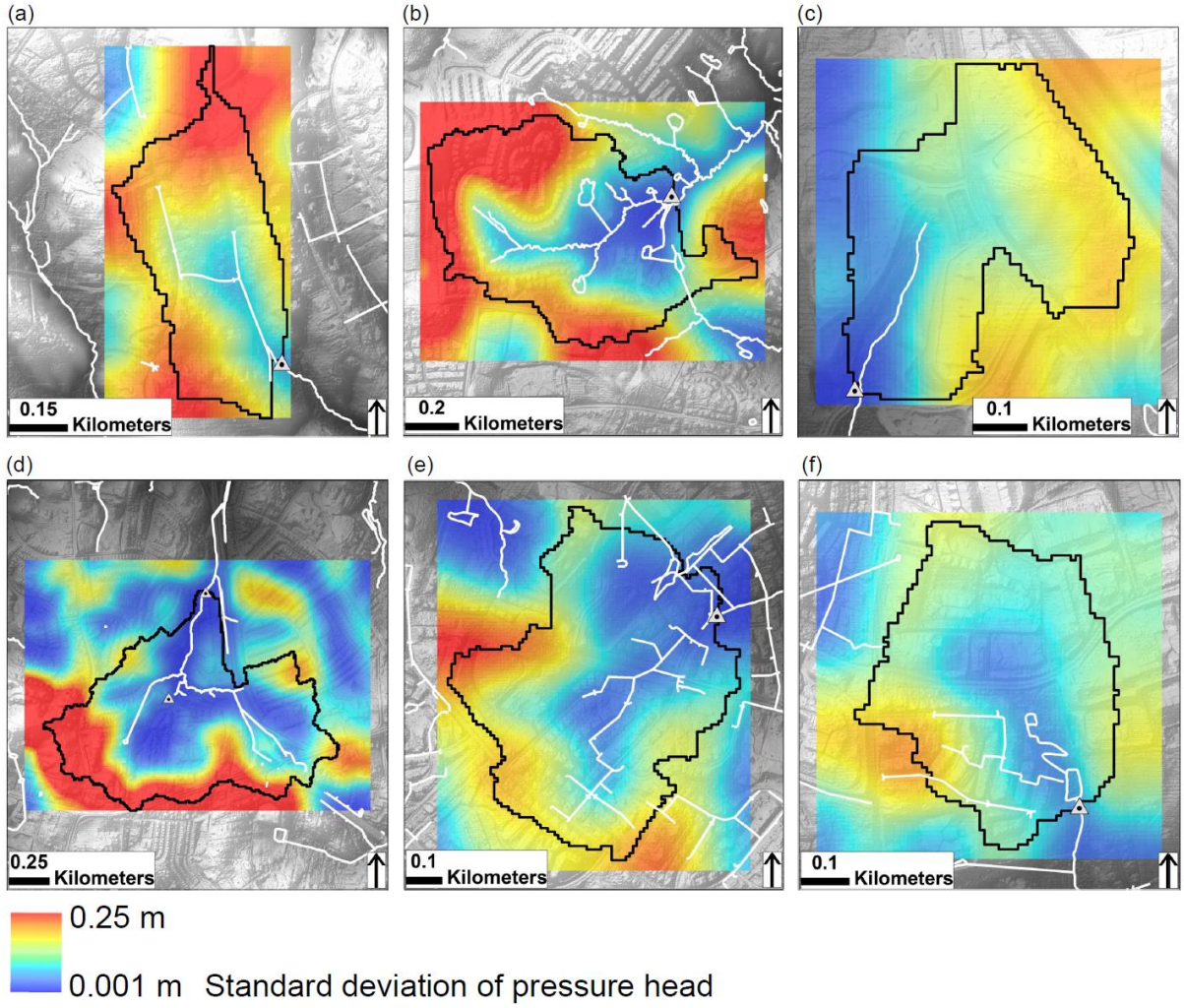

Figure 3. Hillshade digital elevation model (gray scale) derived from LiDAR data overlaid with modeled pressure fields at depth for the six study watersheds: (a) Sunnydale, (b) Runnymeade, (c) Redbrook, (d) DR5, (e) Greengage, and (f) Kevsway. The depicted pressure field is the standard deviation of the modeled pressure head (m) over all hourly timesteps for four years of simulation time at the bottom layer of the model. Watershed boundaries are shown in black; streams and stormwater pipes that contain dry weather flow are shown in white.

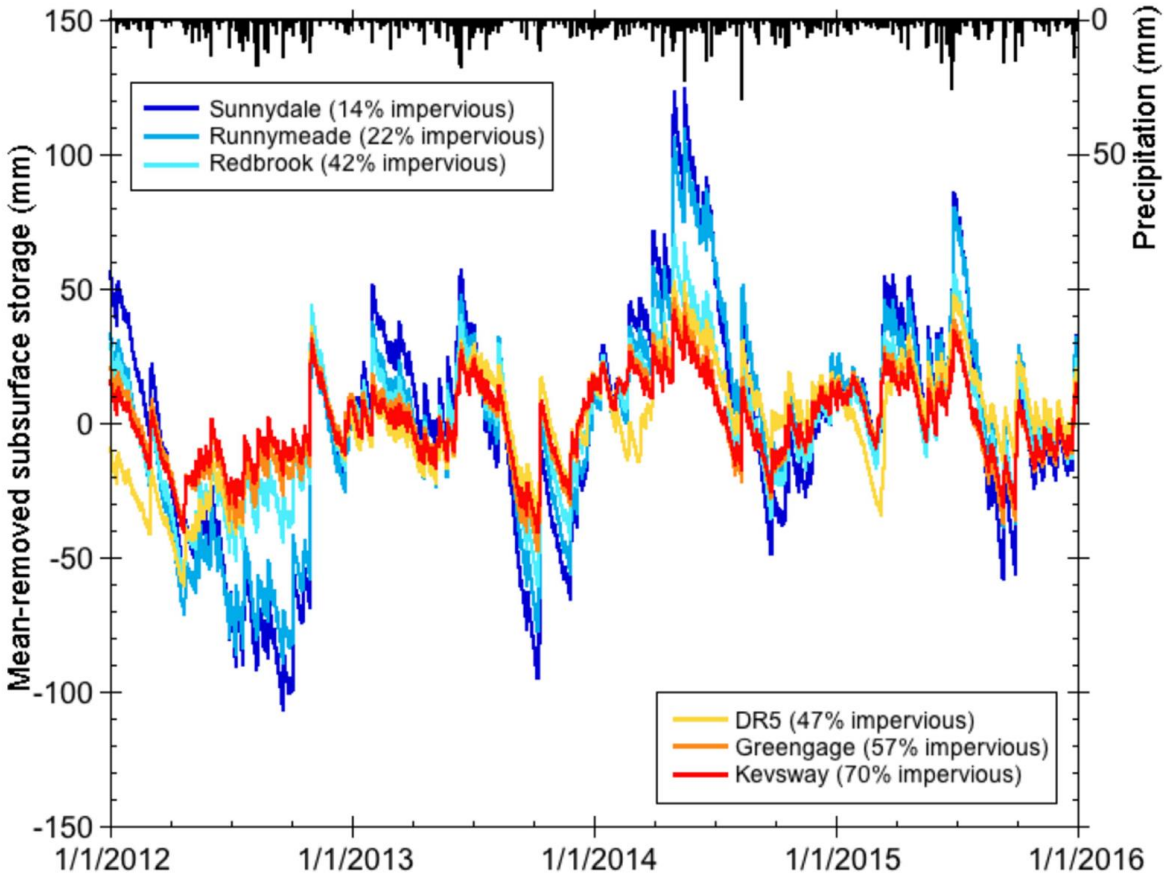

Figure 4. Mean-removed subsurface storage scaled by watershed area for the six study domains, 2012-2015. Precipitation is shown in black. All data are in hourly timesteps.

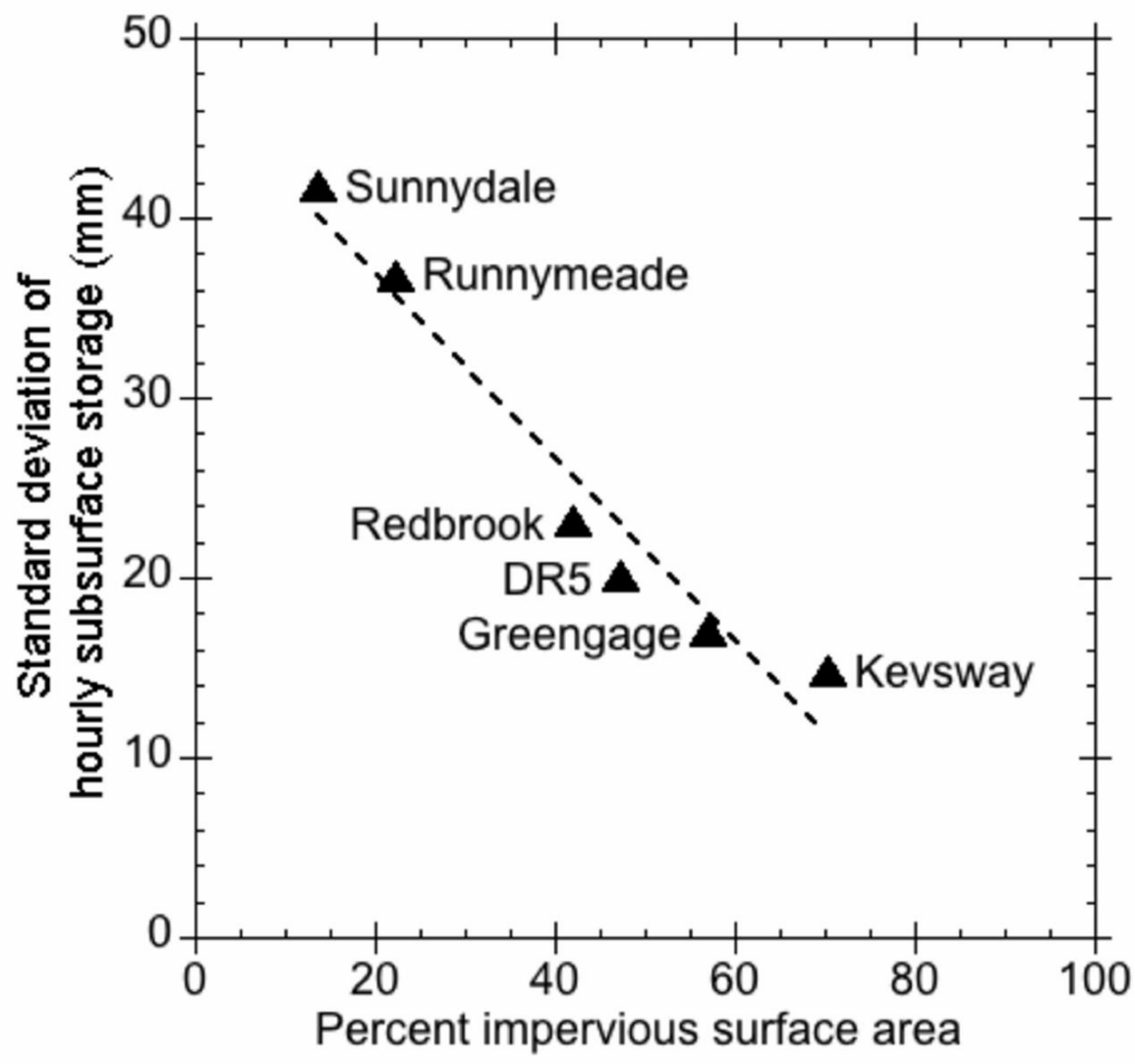

Figure 5. Standard deviation of subsurface storage as a function of percent impervious surface area for the six study watersheds; $R^2 = 0.95$.

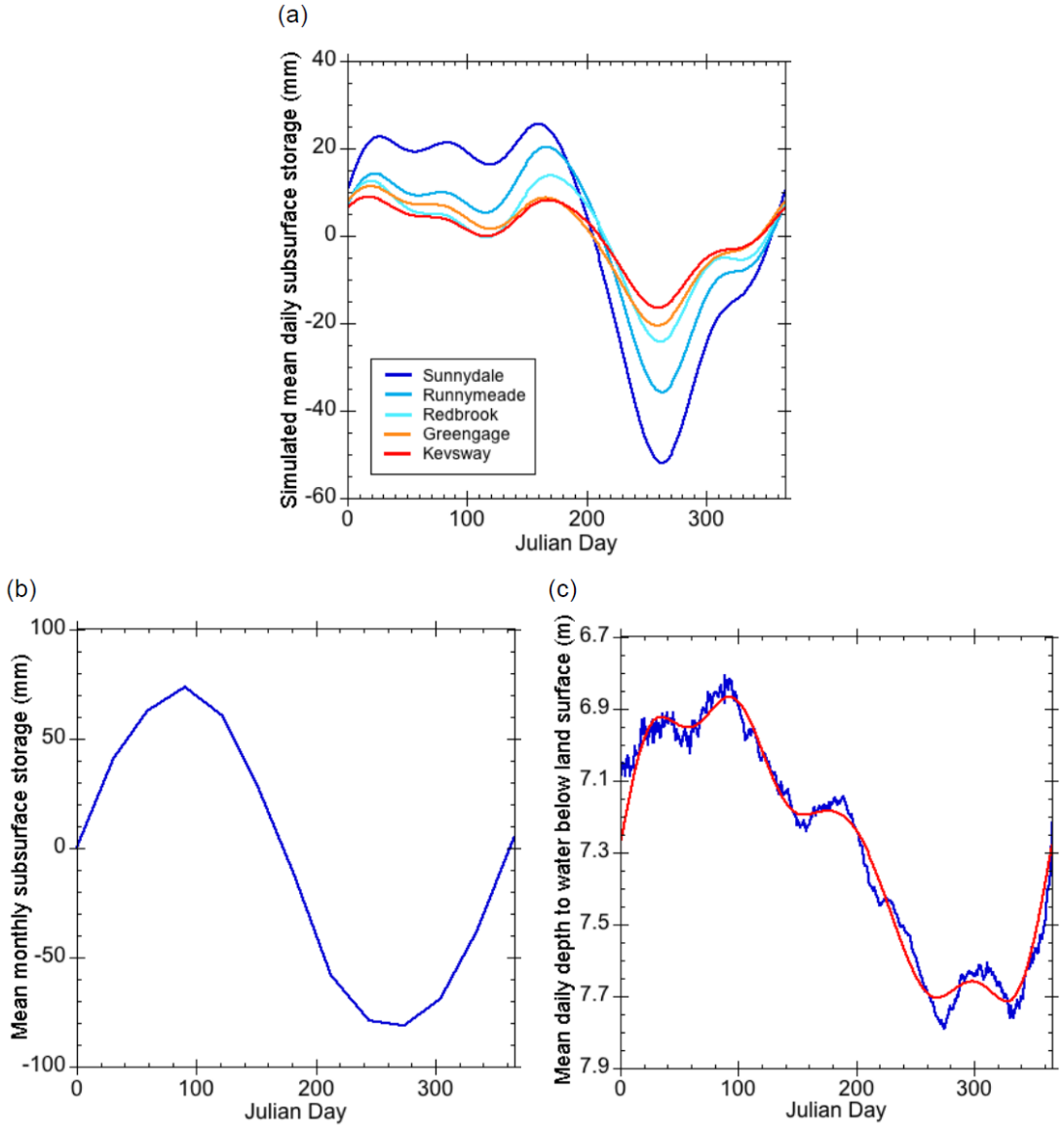

Figure 6. (a) Simulated mean daily subsurface storage as a function of Julian day with temporal smoothing applied. Averages for each Julian day are calculated over the four years (1 January 2012 - 31 December 2015) of model output. (b) Mean monthly change in subsurface storage as a function of Julian day calculated from water budget data for the nearby Gunpowder Falls basin (Figure 1a), 1884 -1960 (Nutter & Otton, 1969). (c) Mean daily depth to water below land surface as a function of Julian day for well DRGC near Kevsway (blue line) with temporal smoothing applied (red line). Data were collected hourly from 1 January 2012 - 31 December 2015; daily averages were calculated and then averaged over the four years of record for each Julian Day.

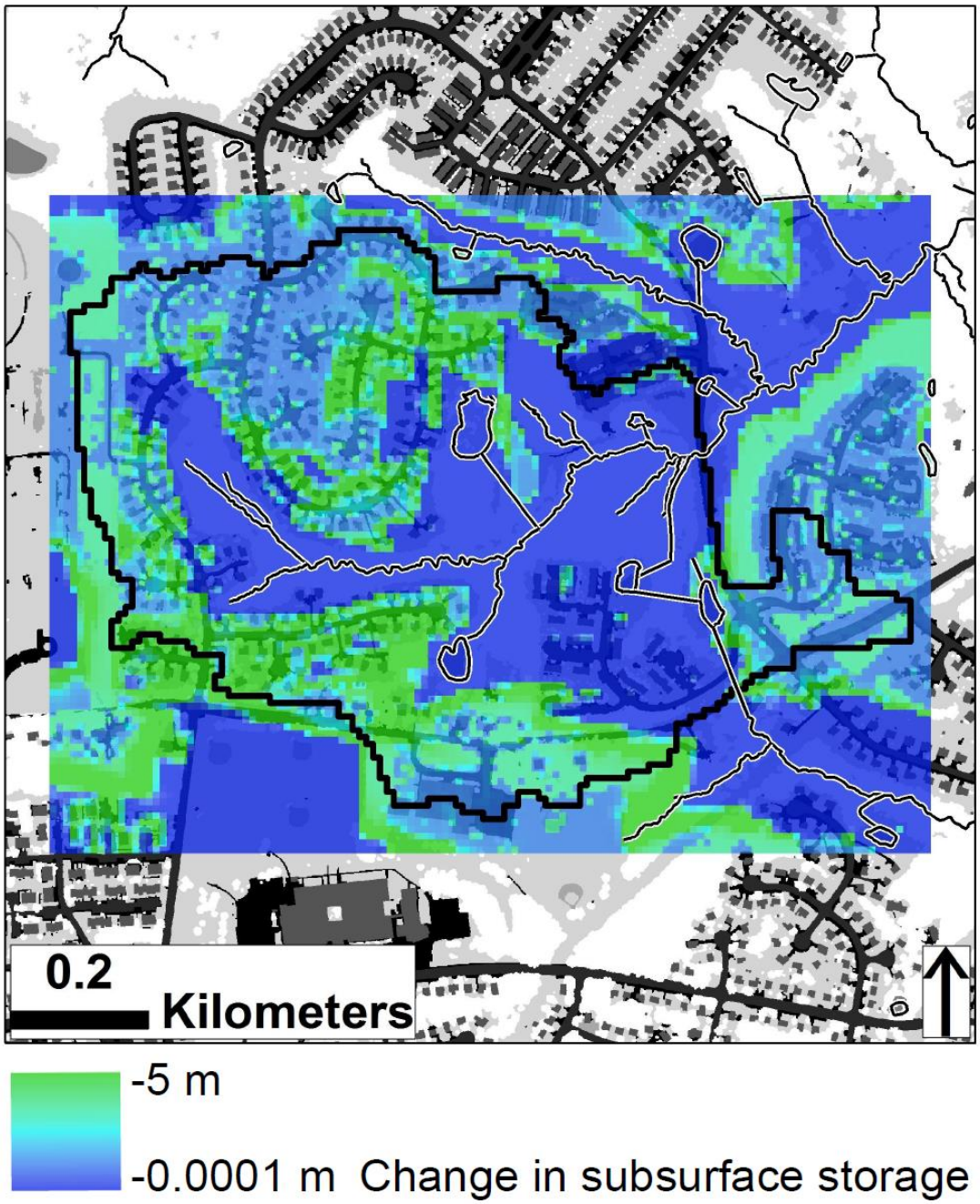

Figure 7. Snapshot of change in subsurface storage in model layer 2 in the unsaturated zone, between the first timestep and the driest timestep of the simulated domain (18 September 2012), draped over a gray-scale version of land cover shown in Figure 2b, for the Runnymeade domain. Blue indicates the smallest change in storage and green indicates largest change in storage. Land cover is depicted with impervious surfaces designated as black and dark gray, grass and shrubs as light gray, and trees as white.

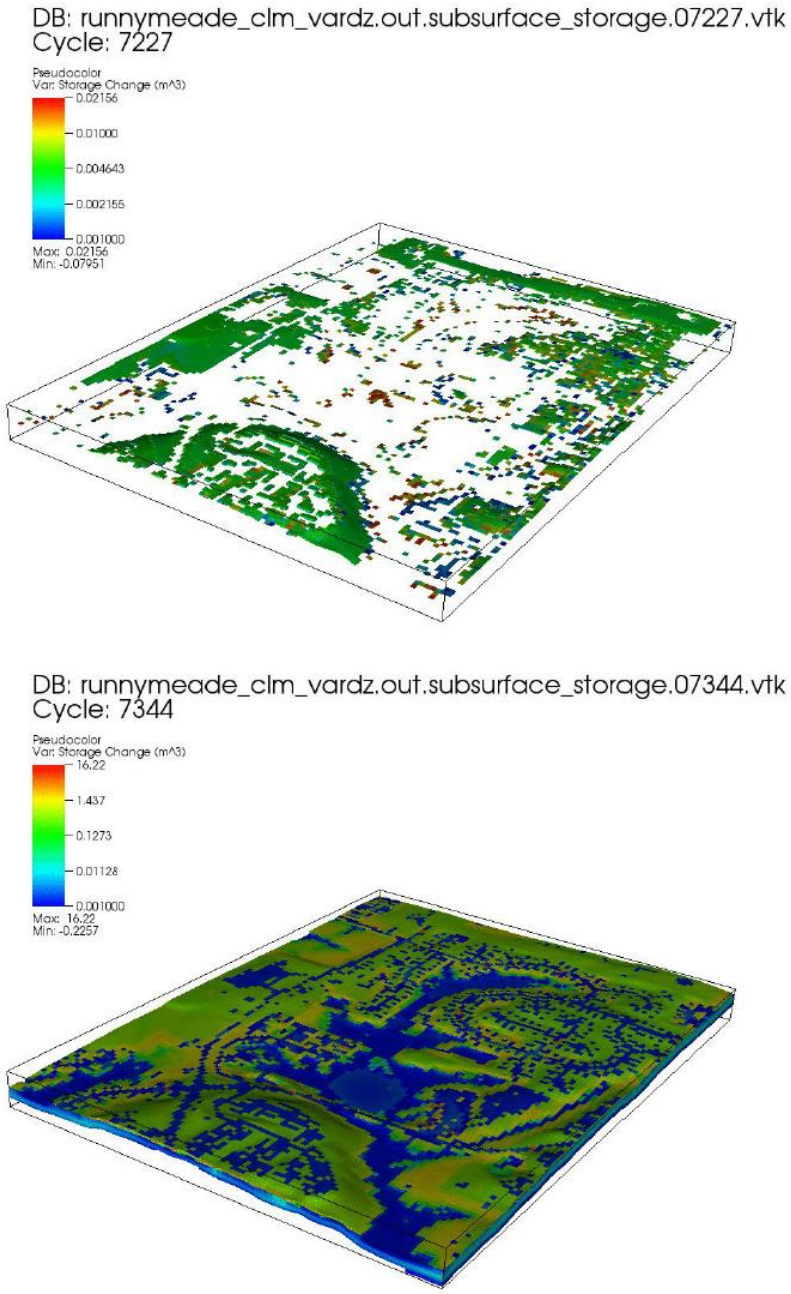

Figure 8. Animation of the infiltration process for Runnymeade for simulation spanning Hurricane Sandy from model time step 7224 (midnight 28 October 2012) to 7344 (midnight 2 November 2012), at an increment of 3 time steps (3 hours). The color scale depicts change in subsurface storage between time step 7223 and the given time step. The view of the lower left side of the domain is from the east as defined by Figures 2b, 3b. Total storm precipitation was 182 mm.