**Modeling the hydrologic effects of watershed-scale green roof implementation**

**Abstract**

Green roofs are among the most popular type of green infrastructure implemented in highly urbanized watersheds due to their low cost and efficient utilization of unused or under-used space. However, few studies have evaluated the performance of large-scale adoption of green roofs across entire watersheds. Therefore, in this study, we examined the effectiveness of green roofs to attenuate stormwater runoff across a large metropolitan area. We utilized a spatially explicit ecohydrological watershed model called Visualizing Ecosystem Land Management Assessments (VELMA) to simulate the resulting stormwater hydrology of implementing green roofs over 25%, 50%, 75%, and 100% of existing buildings within four urban watersheds in Seattle, Washington, United States. We simulated the effects of two types of green roofs: extensive green roofs, which are characterized by shallow soil profiles and short vegetative cover, and intensive green roofs, which are characterized by deeper soil profiles and can support larger vegetation. While buildings only comprise approximately 10% of the total area within each of these watersheds, our simulations showed that 100% implementation of green roofs on these buildings can achieve approximately 10-15% and 20-25% mean annual runoff reductions for extensive and intensive green roofs, respectively, over a 28-year simulation. These results provide an upper limit for volume reductions achievable by green roofs in these urban watersheds. We also showed that stormwater runoff reductions are proportionately smaller during higher flow regimes caused by increased precipitation, likely due to the limited storage capacity of saturated green roofs. In general, green roofs can be effective at reducing stormwater runoff, and their effectiveness is limited by both their areal extent and storage capacity. Our results showed that green roof implementation can be an effective stormwater management tool in highly urban areas, and we demonstrated that our modeling approach can be used to assess the watershed-scale hydrologic impacts of the widespread adoption of green roofs across large metropolitan areas.

Keywords: Green roofs, urban watershed modeling, VELMA, stormwater

**1. Introduction**

In cities throughout the United States (US) (Hoghooghi et al., 2018; Sarkar et al., 2018) and the world (Tzoulas et al., 2007), green roofs are a type of green infrastructure (GI) that is considered a best management practice (BMP) for managing storm runoff as well as other ecosystem services including heat reduction, habitat de-fragmentation, nutrient management, green space, and recreation (Berardi et al., 2014; Golden and Hoghooghi, 2018; Passeport et al., 2013; Woznicki et al., 2018). Green roofs are a popular GI type implemented in highly urbanized watersheds due to their low cost and efficient utilization of unused or under-used space (Carter and Jackson, 2007).

Green roofs, also known as vegetated roofs, refer to plants in a growing medium installed on the top of a building. In terms of managing storm water, green roofs store water within the soil column until the maximum retentive capacity is reached, and capacity is dependent on the volume and composition of the media. After the media becomes saturated, the subsequent flow of water is determined by the hydraulic conductivity of the media (Bollman et al., 2019). The amount and type of vegetation grown is also important for water retention, since some water is intercepted by the canopy, and water is also removed through evapotranspiration (Cascone et al., 2019).

Numerous studies have experimentally investigated the ability of green roofs to delay and retain storm water. For example, Li and Babcock Jr (2014) reviewed results from 19 laboratory and roof-scale experiments and found that green roofs can reduce stormwater runoff volume by 30-86% and delay peak flow by 0 to 30 minutes. Akther et al. (2018) reviewed 60 studies that evaluated the hydrologic performance of green roofs within different climatic regimes and found that median runoff volume reductions ranged from 57-78%, with standard deviations that ranged from 25-35%. While these laboratory and roof-scale findings are important for understanding the possible effectiveness of green roofs, their inherent dependence on local conditions and resulting variability makes it difficult to reliably translate local results to watershed scales.

In addition to experimental studies on the performance of green roofs, numerous studies have modeled the performance of individual green roofs at small scales – for example, using EPA’s stormwater management model (SWMM) (Burszta-Adamiak and Mrowiec, 2013; Cipolla et al., 2016), HYDRUS-1D (Hilten et al., 2008) and HYDRUS-2D (Li and Babcock Jr, 2015), or comparing multiple methods (Carson et al., 2017). Both semi-distributed and spatially explicit watershed models have also been used to simulate the hydrologic impacts of green roofs. Her et al. (2017) proposed a framework within the Soil and Water Assessment Tool (SWAT), a semi-distributed watershed model, to simulate low-impact design including green roofs. Sarkar et al. (2018) used a spatially explicit (i.e., gridded) model called the Regional Hydro-Ecological Simulation System (RHESSys) to simulate various GI practices including green roofs at the block level under various climatic conditions across the US.

Numerous studies from Europe and Asia have also investigated the large-scale implementation of green roofs at watershed scales – for example, in France (Versini et al., 2016; Versini et al., 2015), Italy (Ercolani et al., 2018; Masseroni and Cislaghi, 2016), and Singapore (Schmitter et al., 2016), all of which assessed various green roof configurations and rainfall events and their impacts of green roof performance. Generally, these studies found reductions in flow peaks and volumes at the network outlet that were proportional to the conversion percentage of green roofs throughout the catchment. Green roof performance varied by climate and precipitation patterns, antecedent soil moisture conditions, land use and watershed configuration, and green roof characteristics, where runoff volume reductions for 100% implementations of green roofs ranged from 2% in the tropical climate of Singapore to 25-35% in Italy and France (Masseroni and Cislaghi, 2016; Versini et al., 2016; Versini et al., 2015) and up to 70%, also in Italy (Ercolani et al., 2018). Peak discharge reductions also widely varied, ranging from 10-80% among the aforementioned studies.

In this paper, we use a spatially explicit watershed model called Visualizing Ecosystem Land Management Assessments (VELMA) to evaluate the performance of widespread adoption of green roofs at watershed scales in four watersheds in the Pacific Northwestern US. We simulate the resulting hydrologic impacts of implementing green roofs over 25%, 50%, 75%, and 100% of existing buildings within four urban watersheds in Seattle, Washington, US. We simulated the impacts of implementing extensive green roofs, which are characterized by shallow soil profiles and low-level vegetative cover. We also simulated intensive green roofs, which are characterized by deeper soil profiles and can support larger vegetation. Our study makes two major contributions. First, by simulating the impacts of 100% green roof implementation, we provide an upper limit on the possible runoff volume reductions that can be expected for a large urban watershed such as these in Seattle. Therefore, these results can inform decision makers when crafting programs to support the adoption of urban GI including green roofs and identify the maximum benefit of green roofs for stormwater management. Second, we employ VELMA, a heretofore unutilized watershed model for simulating green roofs. VELMA has been used to model natural and engineered green infrastructure for water quality protection in other systems to compare the effects of GI and climate scenarios on water quality and associated co-benefits and trade-offs for other ecosystem services (Abdelnour et al., 2013; Abdelnour et al., 2011; Golden et al., 2012; Golden et al., 2014; Knightes et al., 2014). Our parameterizations and model improvements of VELMA provide another application of this model that can be useful for future studies to simulate the impacts of green roofs and to compare other green and traditional stormwater infrastructure. Finally, our approach is intended to be useful to assess effectiveness of green roof implementation in other large metropolitan areas.

Green roofs are especially attractive for urban planning policy makers, since they can be implemented with relatively low costs to provide a spatially disaggregated array of ecosystem services scattered throughout watersheds and metropolitan areas, including using unutilized space in public or commercial buildings (GSA, 2011).

**2. Materials and Methods**

**2.1. Study Areas**

We focused on four distinct urban watersheds within the greater Seattle, Washington, United States (US) metropolitan area: Taylor Creek, Thornton Creek, Longfellow Creek, and Pipers Creek (Figure 1). Pipers and Longfellow watersheds drain west into Puget Sound, and Thornton and Taylor watersheds drain into Lake Washington.

**<Insert Figure 1 Here>**

The percentage distribution of land use for each of the four watersheds was derived using 1-m land use/ land cover data obtained from the University of Washington’s Remote Sensing & Geospatial Analysis Laboratory (Styers et al., 2014) (Table 1). Note that the 1-m data were resampled to 10 m to match the digital elevation data (see Input Data section).

**<Insert Table 1 Here>**

The four watersheds vary in size from approximately 3 km2 to 31 km2, yet the land use classification characteristics are remarkably similar (Table 1). For example, the percentages of buildings were 10%, 10%, 10%, and 11% in Taylor, Thornton, Longfellow, and Pipers watersheds, respectively.

Longfellow Creek is located in the southwestern corner of Seattle, Washington and is the most urbanized watershed among the four based on its percentage of buildings and impervious surfaces (e.g., roads, parking lots, and sidewalks). The High Point neighborhood, accounting for approximately 10% of the Longfellow Creek watershed, has worked with Seattle Public Utilities since the 1980s to adopt green infrastructure practices such as grass and vegetated swales, porous pavement, and a large storm-water pond to slow runoff and filter contaminants before reaching the creek and ultimately heading to the Puget Sound (Seattle Public Utilities, 2018).

Thornton Creek, the largest of the four watersheds (31 km2), is located in northeastern Seattle, Washington. The watershed is heavily urbanized and is intersected by Interstate 5, which cuts through the western portion of the watershed. Numerous GI and low-impact design studies have been implemented in Thornton Creek, including the Thornton Creek Water Quality Channel (US EPA, 2016).

Pipers Creek, located on the western side of Seattle, Washington, flows directly into Puget Sound. The watershed holds the highest percentage of forests (46%) of all the watersheds included in this study, and approximately 11% of the watershed area is covered by buildings.

Taylor Creek, located in the southeastern region of Seattle, Washington, flows into Lake Washington. Taylor is the smallest watershed in our sample (3 km2), and the total areal percentage of buildings within the watershed is 10%. Numerous restoration efforts led by the Seattle Public Utilities have been conducted throughout the watershed since 1971, yet, as with the other watersheds included in this study, the large-scale potential of green roof implementations has not been investigated.

**2.2. Watershed Model**

**2.2.1. *Model Overview***

To simulate the effects of green roof implementation scenarios on hydrologic discharge, we used the Visualizing Ecosystem and Land Management Assessments (VELMA v2019-07-22) model (Abdelnour et al., 2011). VELMA is a spatially explicit (i.e., gridded) watershed model that integrates hydrologic and biogeochemical (C and N) sub-models to simulate numerous environmental attributes, including watershed-scale discharge and plant growth. Spatially explicit maps for elevation, land use, and soil type are used to characterize the baseline configuration and attributes of the watershed, and either discrete climate monitoring stations or climate interpolation maps are used as climatic forcers (i.e., temperature and precipitation) that drive the model at a daily time step. Each cell has a surface layer and four soil layers, in which water and dissolved nutrients can flow or be stored within the soil column, and quantities can also flow laterally between cells depending on local elevation changes and saturation levels. Evapotranspiration, combining evaporation and plant transpiration, as well as plant uptake, growth, and decomposition are all mechanistically modeled using various sub-models. In addition, VELMA incorporates spatially explicit disturbances (e.g., atmospheric deposition, harvests, burns, fertilizer application, irrigation) to simulate the effects of external drivers and land management activities, which in turn affect the storage and flow of water and nutrients and the growth of vegetation throughout the landscape.

VELMA has been tested in a variety of ecosystem types, including grassland prairie ecosystems (Barnhart et al., 2015), forests in the Pacific Northwest (Abdelnour et al., 2013; Abdelnour et al., 2011; McKane et al., 2014a), and urbanized mixed-use ecosystems (Hoghooghi et al., 2018). Within these studies, it has been validated to reproduce historical stream flow and carbon stocks on daily, annual, and decadal time scales. A complete description of the model and its sub-components can be found in Abdelnour et al. (2011), Abdelnour et al. (2013), and in the VELMA user manual (McKane et al., 2014b).

**2.2.2. *Model Improvements***

VELMA has recently been used to model semi-urbanized watersheds for implementation of GI (Hoghooghi et al., 2018) and has not yet been used in fully urbanized watersheds or to explicitly model green roofs. Figure 2 depicts a single VELMA voxel that describes how VELMA models vertical flows and lateral flows within the soil subsurface. The left panel designates a traditional VELMA voxel that includes an optional impermeable layer, as implementable in VELMA 2.0. This optional impermeable layer limits the percentage of water that can infiltrate from the surface to the first soil layer and allows VELMA to simulate increased surface runoff and less infiltration caused by the increased impermeability of urbanized surfaces (e.g., buildings, roads, parking lots, sidewalks).

In addition to utilizing the optional permeable layer to better represent urbanized surfaces such as roads and parking lots, we manually parameterized a new soil type to represent green roofs. For this study, the traditional VELMA voxel representation (Figure 2, left panel) was altered to accommodate green roofs (Figure 2, right panel). The first layer of the green roof soil type is characterized by the soil properties of the green roof, whereas the remaining three soil layers are characterized by the soil properties of soil under the building. Lateral flow is allowed both in and out of the first soil layer (i.e., the green roof) and in and out of the lower soil layers, but vertical flow is limited between soil layers 1 and 2 by manually setting the first-layer value of *setSoilLayerKsLateralValues* to a small but non-zero quantity. This essentially limits flow between soil layers 1 and 2 to a negligible quantity while also preventing model crashes due to divisions by 0. We note that the allowance of lateral flow into the green roof is a model simplification that is not reflected in the real world. However, the digital elevation model did not include buildings, and preventing lateral flow between cells via parameterization caused unrealistic flow patterns and model crashes. Therefore, we chose to adopt this model simplification to reflect an approximate mechanistic representation of green roofs in urban environments (Figure 2).

**<Insert Figure 2 Here>**

VELMA disturbances were also used to simulate irrigation (0.3629 cm day-1 including precipitation) and fertilizer (0.005 kgN m-2yr-1) application on grass voxels that were not green roofs, following recommendations by Sarkar et al. (2018), who referenced Milesi et al. (2005) and Carey et al. (2012).

***2.2.3. Input Data***

Standard, spatially distributed inputs are required to construct watershed models including VELMA including a digital elevation model, soil and land use/land cover maps, a stream network, and weather drivers including daily temperature and precipitation (Table 2).

**<Insert Table 2 Here>**

VELMA cells were set to be 10 m x 10 m as a compromise between simulation accuracy and required time. Therefore, each raster input was chosen to match this spatial resolution or was modified, as described in this section, to match the 10-m resolution.

A 10-m digital elevation model (DEM) was acquired from the USGS (Table 2). This product was chosen over lidar-based digital terrain models, which provide higher spatial resolution, because higher resolutions require more voxels to be simulated within VELMA, which in turn increase the total simulation time. Therefore, we chose a 10-m DEM, and more generally 10-m voxels throughout the watershed, as a compromise between accuracy and simulation time. The 10-m DEM was flat-processed using the JPDEM-Dredge processing tool (McKane et al., 2014b; Pan et al., 2012). A hand-digitized stream network obtained from the City of Seattle was used to aid the JPDEM-Dredge processing tool that enforces pre-determined flow routing within the DEM. Downspouts and municipal sewers and pipes were not incorporated into the flow routing used by the model due to a lack of data, yet each of the four watersheds are heavily urbanized and contain downspouts, sewers, and municipal pipes, so this is a known simplification and possible source of error.

A single soil type, sandy loam, was initially used for all four watersheds. An additional soil type was then created for cells that implemented green roofs and were characterized by intensive and extensive green roof media characteristics, as described in a subsequent section.

Land use data were acquired from the University of Washington’s Remote Sensing & Geospatial Analysis Laboratory (Styers et al., 2014) and consisted of 1-m land use land cover data across the Seattle metropolitan area (Tables 1-2). Four land use categories were included: grass, trees, buildings, and other impervious surfaces (e.g., roads, sidewalks, parking lots). Therefore, existing GI (e.g., established green roofs, ponds, impervious pavement) were not included in the baseline model simulations. The land use data were resampled to 10-m cells via majority rule, meaning that the most prevalent land use replaced the others when scaling down from 1-m to 10-m cells. This resulted in an average increase of 0.58% in building area for the four watersheds, which was considered inconsequential. For these and other geospatial and statistical techniques used in this analysis, scripts were written using R 3.1.2 statistical software (R Core Team, 2013) and Python 2.7.12 (Python Software Foundation, 2016) programming language. Visualizations, sampling location analysis, and basic map editing were made with ArcGIS 10.3 (ESRI, 2014).

Three NOAA-referenced weather stations (Sand Point, Portage Bay, and Boeing Field) and Daymet modeled data were used to compile daily mean temperature and precipitation estimates for the duration of our model runs(NOAA, 2016; Thornton et al., 2017). All three stations were within the municipal boundaries of the City of Seattle and were located between 2-21.5 km of either Thornton or Pipers creeks. The Sand Point weather station had 10,076 recorded daily weather observations between 1986 and 2015, including 526 missing daily observations, 13 precipitation NA’s, and 1 average temperature NA observations. Between 1986-1-1 and 1998-4-30, Sand Point had 151 missing daily observations, 4 precipitation NA’s, and 1 average temperature NA, which were gap filled with Portage Bay recorded weather. Boeing Field weather observations were used to gap fill 153 days of missing Sand Point daily data between 1998-12-5 and 2015-12-31 as well as being used to replace the 9 remaining precipitation NA’s. From 1998-5-1 through 1998-12-4, there were no recorded weather observations at Sand Point, Portage Bay, or Boeing Field, so these days were completely gap filled with Daymet modeled data (Thornton et al., 2017). Daymet model output data were acquired for the 1-km cell at the Sand Point station latitude and longitude. R 3.1.2 statistical software (R Core Team, 2013) was used for gap filling observed NOAA weather station data with Daymet daily gridded modeled weather parameters, using the “daymetr” package (Hufkens et al., 2018) for single cell sampling.

**2.4. Baseline Calibration and Validation**

A semi-automatic calibration procedure was implemented to tune VELMA’s calibration parameters in order to match daily simulations of watershed discharge with daily observed streamflow for a baseline model of Taylor Creek including buildings with no green roofs. First, the nondominated sorting genetic algorithm II (NSGA-II; (Deb et al., 2002)) was implemented using the MOEA Framework (Hadka, 2014) to test various sets of VELMA input parameters and to maximize the Nash Sutcliffe efficiency (NSE; Nash and Sutcliffe [1970]) criterion (Equation 1):

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

where O is the observed value, S is the simulated value, and is the mean of the observed values. NSE values range from -∞ to 1.0 where one represents a perfect fit with the observed data.

VELMA was calibrated for Taylor Creek only, and the resulting calibrated model parameters were then transferred to the remaining three watersheds. Taylor Creek was chosen because it is the smallest watershed and therefore produced the fastest computational runtime, and it also had the longest period of observed hydrologic data available. The close proximity of the watersheds (~6-10 km) justified the use of the same calibration parameters across all watersheds, since it was considered unlikely for the mechanisms to substantially vary. Observed daily discharge data from the mainstem of Taylor Creek near the outlet was obtained from Seattle Public Utilities (station 401) (Seattle Public Utilities, 2016), and calibration was performed with daily data from 2004-2009 (Table 3).

**<Insert Table 3 Here>**

The *petparam1* parameters for both conifer and grass cover types determine the allowable potential evapotranspiration using the Hamon equation (Hamon, 1960), and the parameter *be* adjusts the fraction of actual evapotranspiration compared with the potential values. Meanwhile, the parameter *surfaceKs* is the surface saturated hydraulic conductivity (mm-day-1), and the parameters *ksLat* and *ksVert* are unitless multipliers that determine the respective rates of decrease in lateral and vertical flow with depth.

The implementation of NSGA-II with VELMA using the MOEA Framework tested approximately 3,500 parameter sets, and the NSE values between daily observed and simulated discharge were maximized. In total, the best 848 parameter sets produced simulations 0.60 > NSE > 0.62 and were retained. From these, a single parameter set was manually chosen and applied to each of the four watersheds to serve as baseline models. While aboveground biomass values were not explicitly incorporated into the automatic portion of the calibration procedure, final VELMA input parameter sets were chosen such that the simulated maximum annual aboveground biomass values reached approximately 240 and 1000 gC m-2 yr-1 in order to match data from experiments conducted by Getter et al. (2009). Evapotranspiration values were also verified to

The calibrated baseline models were then compared with the green roof scenarios (Section 2.5). While all simulations with 0.60 > NSE > 0.62 could feasibly be run to estimate the relative uncertainty of model outputs, the required simulation times, which were approximately 6.5, 13, 18.5, and 57 hours for Taylor, Pipers, Longfellow, and Thornton watersheds, respectively, were prohibitive and prevented this type of analysis.

**2.5. Green Roof Scenarios**

Green roofs are generally categorized as either intensive or extensive (McIntosh, 2010). Intensive green roofs (IGRs) are characterized by thick soil columns (e.g., >15 cm) and can include landscaped gardens, mixtures of trees, bushes and grass. They require substantial structural support and are typically installed on large, commercial buildings that may allow pedestrian access. Extensive green roofs (EGRs) are characterized by shallow soil depths (e.g., 5-15 cm) and short vegetation that typically covers a large proportion of the roof. EGRs can be implemented on buildings with less structural support than IGRs, and typically do not require maintenance such as irrigation or fertilization (McIntosh, 2010).

To model green roofs in VELMA, both the cover and soil characteristics of green roof voxels were changed to match those of intensive or extensive green roofs (Table 4).

**<Insert Table 4 Here>**

The values for the general soil type were chosen to match a sandy loam soil type and were taken from McKane et al. (2014b). The intensive and extensive green roof soil characteristics were taken from the technical specifications of a proprietary source of green roof media (Rooflite Extensive 600 Media and Rooflite Intensive 700 Media; (Rooflite, 2020) that was designated as an approved media source to obtain stormwater reduction credit by the City of Seattle (Magnusson Klemencic Associates and Seattle Public Utilities, 2008)

In addition to the soil characteristics, we also changed the land cover from a traditional building (i.e., no vegetation biomass) to cover characteristics of grass that match either extensive or intensive green roofs. As noted previously, the simulated maximum annual aboveground biomass values reached approximately 240 and 1000 gC m-2 yr-1, which match data from experiments conducted by Getter et al. (2009).

No additional routing of green roof runoff, via downspouts or other drain connections, were included in the model. In addition, sewer networks were not explicitly incorporated into the flow routing of the model; therefore, flow routing merely followed elevation changes according to the 10-m digital elevation model used. Additions of downspouts and sewer networks may improve the realism of this simplified model and will be left for future research.

Green roof scenarios were simulated for intensive and extensive green roofs separately, and four different scenarios were created for each that randomly converted different proportions of existing buildings to green roofs (i.e., 25%, 50%, 75%, and 100%) (Figures 3-6). Note that the method of random placement applied to all 10-m cells, and therefore some green roof implementations only covered a portion of the total roof area of a given building.

Simulations were run for 29 years (1987-2015), and the first year was designated as a spin-up year and not included in the results.

**<Insert Figure 3 Here>**

**<Insert Figure 4 Here>**

**<Insert Figure 5 Here >**

**<Insert Figure 6 Here>**

**3. Results and Discussion**

The automatic calibration procedure conducted with Taylor Creek resulted in 848 out of 3,500 parameter sets with 0.60 < NSE < 0.62 (see Table 3) between the daily observed and simulated discharge for a location near the outlet of Taylor Creek watershed. These solutions were represented by a wide range of parameter values, indicating that different combinations of parameters (e.g., a high *be* value combined with a low *Petparam1* value compared with a low *be* value combined with a high *Petparam1* value) produced similar hydrographs. Fully addressing parameter redundancy is beyond the scope of this investigation and will be left for future research. Therefore, for the present investigation, a single parameter set was chosen, as shown in Table 3, and transferred to the other four watersheds. Further improvements to the calibration procedure could be made, for example, by including additional objectives related to soil moisture and evapotranspiration or calibrating base and surface flows separately.

The results of the four model scenarios (25%, 50%, 75%, and 100% of buildings converted to green roofs) in addition to the baseline scenario (0%) for both extensive and intensive green roof types are shown in Figure 7. As the total percentage of buildings converted to green roofs increases, the total runoff reductions increase among the scenarios, although the response is not linear. The intensive green roof scenarios have higher storage capacity and are therefore able to reduce total annual runoff values more effectively that the extensive green roofs, producing maximum simulated runoff reductions of 20-25% and 10-15% for intensive and extensive green roofs, respectively. These results appear to be low compared to Li and Babcock Jr (2014), who reviewed results from 19 laboratory and roof-scale experiments and found runoff volume reductions ranged from 30-86%, as well as compared to Akther et al. (2018), who reviewed 60 green roof studies amidst different climatic regimes and found that median runoff volume reductions ranged from 57-78% (SD: 25-35%). The difference in size scales between these studies and our watershed assessments may account for these discrepancies.

To compare specifically with other watershed models, Sarkar et al. (2018) used a similar spatially explicit model called RHESSys and found that green roofs provided a 33% (median) reduction in annual water yields, which represents slightly greater reductions than our simulations for intensive green roofs**.** However, the RHESSys simulations were applied to a small set of archetypal urban subunits (AUSs) that only comprise a single city block and adjacent road leading to a pour point rather than an entire watershed. In our study, by leveraging land use data to determine the actual locations of buildings within each of our watersheds, we provide a realistic upper bound on the runoff reductions that would be feasible within these four watersheds by employing green roofs alone as a stormwater management tool.

Another way to contextualize our runoff results is to consider that the total watershed area converted to green roofs in each of the models was only approximately 10%. Also, the land use percentages were remarkably similar among the four watersheds, even though the watersheds varied in size from 3 to 31 km2 and were located in four distinct regions of the greater Seattle metropolitan area. Therefore, the 10-15% and 20-25% runoff reductions were achieved by converting only 10% of the watershed area to extensive and intensive green roofs, respectively. For comparison, Hoghooghi et al. (2018) used VELMA to simulate various GI, and while they did not explicitly model green roofs, they showed that implementing rain gardens over 9% of the total watershed area produced a 22% reduction in surface runoff, which is in line with our current estimates for intensive green roofs. This may not be surprising since the same model was used and since rain gardens embody similar characteristics to intensive green roofs.

While our general hypothesis regarding green roofs reducing runoff was confirmed, some simulation results are counterintuitive. For example, the 75% and 100% extensive green roof simulations for Pipers Creek appear to be anomalous outliers. These simulations produce runoff reductions that are less than the 50% scenarios, which is counterintuitive, and the runoff simulations for the 100% green roof scenarios even stretch above the baseline (0%) simulations. One reason for these anomalous results may be due to limitations associated with the transferability of flow-related parameters from Taylor to Pipers Creek. Alternatively, the presence and function of storm drains and sewer networks were not included within the model, nor in any of the models used in this study, and this lack of realism may have contributed to these anomalous results. Therefore, these limitations need to be considered and amended in future model developments.

For 100% implementations of green roofs across the four watersheds, both the extensive and intensive green roof scenarios show that runoff reductions vary with annual precipitation (Figure 8). These trends are slightly more pronounced in the extensive green roof scenarios compared with the intensive green roofs, but overall, they indicate that wetter years decrease the annual effectiveness of green roofs to reduce runoff. The green roofs simply become saturated more often throughout the year and cannot retain water beyond their capacity. These results agree with the roof-scale experiments conducted by Speak et al. (2013), who found that retention rates of green roofs significantly decreased during high rainfall events, and are also supported by Hoghooghi et al. (2018), who used VELMA to simulate rain gardens at the watershed scale which also become saturated at some level of precipitation.

**4. Conclusions**

We examined the hydrologic impacts of large-scale green roof implementations in four heavily urbanized watersheds in Seattle, Washington, US. We found that 20-25% and 10-15% median annual runoff reductions were achievable when only 10% of the total watershed areas were converted to green roofs when using intensive and extensive green roof types, respectively. This result may help guide city planners who seek to mitigate excessive stormwater runoff in highly urbanized watersheds using green infrastructure approaches.

One of the advantages of using VELMA, a spatially explicit (i.e., gridded) watershed model, is the ability to assess the hydrologic impacts of spatially precise implementations of GI and management. Because converting all roof areas to green roofs may not be feasible in most metropolitan areas, spatially explicit approaches for placing green roofs non-randomly in urban watersheds can optimize GI effectiveness (Martin-Mikle et al., 2015). Future research should therefore investigate the impacts of different spatial configurations of green roofs across watersheds to determine whether prioritizing particular areas can increase their effectiveness.

We also showed that the runoff reductions exhibited increasing trends with annual precipitation, indicating that green roofs become saturated more often throughout wetter years and cannot retain water beyond their capacity, which decreases their effectiveness. Therefore, future methods to increase the storage capacity of green roofs will be important in addition to considering their areal extent and spatial configuration. As one example, Bollman et al. (2019) created various green roof soil media mixtures that optimized water storage capacity and wet weight, amongst other parameters. These types of studies, in addition to their incorporation into watershed models, will contribute to improving the effectiveness of green roofs across all scales.

While mechanistic watershed models, in general, and spatially explicit versions, specifically, have been proven capable for simulating GI across watershed scales, one drawback is their long computational runtimes. In our study, runtimes per scenario were 6.5, 13, 18.5, and 57 hours for Taylor, Pipers, Longfellow, and Thornton watersheds, respectively. Future research on statistical meta-models (e.g., Semiromi et al. (2018)) should therefore be considered in order to characterize the relationships between drivers and responses specifically embodied in mechanistic models but in a more scalable and computationally efficient manner.

It remains unknown how different model simplifications may have affected the resulting model prediction errors. For example, a single soil type, 3 climate monitoring stations, and 10-m rasters were used as inputs for the spatially explicit watershed models. While we showed negligible differences in characterizing land use at 10-m compared with 1-m – there was 0.58% change in average building area across all four watersheds – it is possible that spatial heterogeneity of soils, climate, or other variables beyond those captured by the model may be present and contribute to each model’s predictive error. Only flat roofs were considered, neglecting slopes, which may overestimate the extent to which green roofs (especially intensive green roofs) can be feasibly implemented. Downspouts and the municipal sewer system were not explicitly incorporated into the model, which would impact flow routing and perhaps achievable runoff. Finally, the same calibration parameters were used for all watersheds, assuming homogeneity among the mechanistic model formulations across the watersheds. These, among other model simplifications leveraged by this study, should be considered and potentially addressed in future work.

In general, green roofs can be effective at reducing stormwater runoff, and their effectiveness is limited by both their areal extent and storage capacity. While these results are applicable to city planners working in urban watersheds near Puget Sound, our results showed that green roof implementation can be an effective stormwater management tool in highly urban areas, and we demonstrated that our modeling approach can be used to assess the watershed-scale hydrologic impacts of the widespread adoption of green roofs across large in other metropolitan areas. In particular, urban planning policies to promote the adoption of green roofs in urban watersheds and metropolitan areas should consider similar modeling exercises to determine the maximal achievable runoff reductions to more accurately inform cost and benefit analyses. Green roofs also offer a suite of other ecosystem services beyond runoff delay and retention (GSA, 2011), and therefore future work should attempt quantify these affects using basin-wide models and analyses.

**Acknowledgements**

Heather Golden made numerous constructive comments on earlier drafts of the manuscript. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency. Any mention of trade names, products, or services does not imply an endorsement by the U.S. Government or the U.S. Environmental Protection Agency. The EPA does not endorse any commercial products, services, or enterprises."

**References**

Abdelnour, A., B. McKane, R., Stieglitz, M., Pan, F., Cheng, Y., 2013. Effects of harvest on carbon and nitrogen dynamics in a Pacific Northwest forest catchment. Water Resources Research 49, 1292-1313.

Abdelnour, A., Stieglitz, M., Pan, F., McKane, R., 2011. Catchment hydrological responses to forest harvest amount and spatial pattern. Water Resources Research 47, W09521.

Akther, M., He, J., Chu, A., Huang, J., Van Duin, B., 2018. A review of green roof applications for managing urban stormwater in different climatic zones. Sustainability 10, 2864.

Barnhart, B.L., Mckane, R., Brookes, A., Schumaker, N., Papenfus, M., Pettus, P., Halama, J., Powers, B., Djang, K., Groskinsky, B., 2015. Integrated Modeling to Assess the Ecological and Air Quality Trade-offs of Agricultural Burning in the Flint Hills of Eastern Kansas, AGU Fall Meeting Abstracts.

Berardi, U., GhaffarianHoseini, A., GhaffarianHoseini, A., 2014. State-of-the-art analysis of the environmental benefits of green roofs. Applied Energy 115, 411-428.

Bollman, M.A., DeSantis, G.E., DuChanois, R.M., Etten-Bohm, M., Olszyk, D.M., Lambrinos, J.G., Mayer, P.M., 2019. A framework for optimizing hydrologic performance of green roof media. Ecological Engineering 140, 105589.

Burszta-Adamiak, E., Mrowiec, M., 2013. Modelling of green roofs' hydrologic performance using EPA's SWMM. Water Science and Technology 68, 36-42.

Carey, R.O., Hochmuth, G.J., Martinez, C.J., Boyer, T.H., Nair, V.D., Dukes, M.D., Toor, G.S., Shober, A.L., Cisar, J.L., Trenholm, L.E., 2012. A review of turfgrass fertilizer management practices: Implications for urban water quality. HortTechnology 22, 280-291.

Carson, T., Keeley, M., Marasco, D.E., McGillis, W., Culligan, P., 2017. Assessing methods for predicting green roof rainfall capture: A comparison between full-scale observations and four hydrologic models. Urban Water Journal 14, 589-603.

Carter, T., Jackson, C.R., 2007. Vegetated roofs for stormwater management at multiple spatial scales. Landscape and urban planning 80, 84-94.

Cascone, S., Coma, J., Gagliano, A., Perez, G., 2019. The evapotranspiration process in green roofs: A review. Building and Environment 147, 337-355.

Cipolla, S.S., Maglionico, M., Stojkov, I., 2016. A long-term hydrological modelling of an extensive green roof by means of SWMM. Ecological engineering 95, 876-887.

Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE transactions on evolutionary computation 6, 182-197.

Ercolani, G., Chiaradia, E.A., Gandolfi, C., Castelli, F., Masseroni, D., 2018. Evaluating performances of green roofs for stormwater runoff mitigation in a high flood risk urban catchment. Journal of Hydrology 566, 830-845.

ESRI, 2014. ArcGIS Desktop: Release 10.3. Environmental Systems Research Institute, Redlands, CA.

Getter, K.L., Rowe, D.B., Robertson, G.P., Cregg, B.M., Andresen, J.A., 2009. Carbon sequestration potential of extensive green roofs. Environmental science & technology 43, 7564-7570.

Golden, H., Knightes, C., Conrads, P., Davis, G., Feaster, T., Journey, C., Benedict, S., Brigham, M., Bradley, P., 2012. Characterizing mercury concentrations and fluxes in a Coastal Plain watershed: Insights from dynamic modeling and data. Journal of Geophysical Research: Biogeosciences 117.

Golden, H.E., Hoghooghi, N., 2018. Green infrastructure and its catchment‐scale effects: an emerging science. Wiley Interdisciplinary Reviews: Water 5, e1254.

Golden, H.E., Lane, C.R., Amatya, D.M., Bandilla, K.W., Kiperwas, H.R., Knightes, C.D., Ssegane, H., 2014. Hydrologic connectivity between geographically isolated wetlands and surface water systems: A review of select modeling methods. Environmental Modelling & Software 53, 190-206.

GSA, U., 2011. The Benefits and Challenges of Green Roofs on Public and Commercial Buildings. A Report of the United States General Service Administration.

Hadka, D., 2014. MOEA framework user guide. <http://moeaframework.org/>.

Hamon, W.R., 1960. Estimating potential evapotranspiration. Massachusetts Institute of Technology.

Her, Y., Jeong, J., Arnold, J., Gosselink, L., Glick, R., Jaber, F., 2017. A new framework for modeling decentralized low impact developments using Soil and Water Assessment Tool. Environmental modelling & software 96, 305-322.

Hilten, R.N., Lawrence, T.M., Tollner, E.W., 2008. Modeling stormwater runoff from green roofs with HYDRUS-1D. Journal of hydrology 358, 288-293.

Hoghooghi, N., Golden, H., Bledsoe, B., Barnhart, B., Brookes, A., Djang, K., Halama, J., McKane, R., Nietch, C., Pettus, P., 2018. Cumulative Effects of Low Impact Development on Watershed Hydrology in a Mixed Land-Cover System. Water 10, 991.

Hufkens, K., Basler, D., Milliman, T., Melaas, E.K., Richardson, A.D., 2018. An integrated phenology modelling framework in R. Methods in Ecology and Evolution 9, 1276-1285.

Knightes, C.D., Golden, H.E., Journey, C.A., Davis, G.M., Conrads, P.A., Marvin-DiPasquale, M., Brigham, M.E., Bradley, P.M., 2014. Mercury and methylmercury stream concentrations in a Coastal Plain watershed: A multi-scale simulation analysis. Environmental pollution 187, 182-192.

Li, Y., Babcock Jr, R.W., 2014. Green roof hydrologic performance and modeling: a review. Water science and technology 69, 727-738.

Li, Y., Babcock Jr, R.W., 2015. Modeling hydrologic performance of a green roof system with HYDRUS-2D. Journal of Environmental Engineering 141, 04015036.

Magnusson Klemencic Associates, Seattle Public Utilities, 2008. Memo Draft - Green Roof Media Recommended Specifications. <http://www.seattle.gov/dpd/cs/groups/pan/@pan/documents/web_informational/p2371388.pdf>.

Martin-Mikle, C.J., de Beurs, K.M., Julian, J.P., Mayer, P.M., 2015. Identifying priority sites for low impact development (LID) in a mixed-use watershed. Landscape and urban planning 140, 29-41.

Masseroni, D., Cislaghi, A., 2016. Green roof benefits for reducing flood risk at the catchment scale. Environmental Earth Sciences 75, 579.

McIntosh, A., 2010. Green roofs in Seattle: A survey of vegetated roofs and rooftop gardens. <https://www.seattle.gov/Documents/Departments/OSE/Green-Roofs-In-Seattle.pdf>, 1-8.

McKane, R., Brookes, A., Djang, K., Halama, J., Pettus, P.B., Papenfus, M., Phillips, D., Dewitt, T., Brown, C.A., Stecher, H., 2014a. Quantifying ecosystem service tradeoffs in response to alternative land use and climate scenarios: Pacific Northwest applications of the VELMA ecohydrological model. Presented at Salish Sea Ecosystem Conference, Seattle, WA, April 30-May 2, 2014.

McKane, R., Brookes, A., Djang, K., Stieglitz, M., Abdelnour, A., Pan, F., Halama, J., Pettus, P., Phillips, D., 2014b. Velma Version 2.0: User Manual and Technical Documentation. Environmental Protection Agency Office of Research and Development National Health and Environmental Effects Research Laboratory, Corvallis, OR, USA.

Milesi, C., Running, S.W., Elvidge, C.D., Dietz, J.B., Tuttle, B.T., Nemani, R.R., 2005. Mapping and modeling the biogeochemical cycling of turf grasses in the United States. Environmental management 36, 426-438.

NOAA, 2016. Global Historical Climatology Network. <https://www.ncdc.noaa.gov/data-access>.

Pan, F., Stieglitz, M., McKane, R.B., 2012. An algorithm for treating flat areas and depressions in digital elevation models using linear interpolation. Water Resources Research 48, W00L10, doi:10.1029/2011WR010735.

Passeport, E., Vidon, P., Forshay, K.J., Harris, L., Kaushal, S.S., Kellogg, D.Q., Lazar, J., Mayer, P., Stander, E.K., 2013. Ecological engineering practices for the reduction of excess nitrogen in human-influenced landscapes: A guide for watershed managers. Environmental management 51, 392-413.

Python Software Foundation, 2016. Python Language Reference. <https://www.python.org/>.

R Core Team, 2013. R: A language and environment for statistical computing. <https://www.r-project.org/>.

Rooflite, 2020. Certified Green Roof Media, <https://www.rooflitesoil.com>.

Sarkar, S., Butcher, J.B., Johnson, T.E., Clark, C.M., 2018. Simulated Sensitivity of Urban Green Infrastructure Practices to Climate Change. Earth Interactions 22 (2018), 13:11-37.

Schmitter, P., Goedbloed, A., Galelli, S., Babovic, V., 2016. Effect of catchment-scale green roof deployment on stormwater generation and reuse in a tropical city. Journal of Water Resources Planning and Management 142, 05016002.

Seattle Public Utilities, 2016. <https://www.seattle.gov/utilities>.

Semiromi, M.T., Omidvar, S., Kamali, B., 2018. Reducing Computational Costs of Automatic Calibration of Rainfall-Runoff Models: Meta-Models or High-Performance Computers? Water 10, 1440.

Speak, A., Rothwell, J., Lindley, S., Smith, C., 2013. Rainwater runoff retention on an aged intensive green roof. Science of the Total Environment 461, 28-38.

Styers, D.M., Moskal, L.M., Richardson, J.J., Halabisky, M.A., 2014. Evaluation of the contribution of LiDAR data and postclassification procedures to object-based classification accuracy. Journal of Applied Remote Sensing 8, 083529.

Thornton, M., Thornton, P., Wei, Y., Vose, R., Boyer, A., 2017. Daymet: Station-Level Inputs and Model Predicted Values for North America, Version 3. ORNL DAAC, Oak Ridge, Tennessee, USA.

Tzoulas, K., Korpela, K., Venn, S., Yli-Pelkonen, V., Kaźmierczak, A., Niemela, J., James, P., 2007. Promoting ecosystem and human health in urban areas using Green Infrastructure: A literature review. Landscape and urban planning 81, 167-178.

US EPA, 2016. City green: Innovative green infrastructure solutions for downtowns and infill locations. EPA230R16001 <https://www.epa.gov/sites/production/files/2016-06/documents/city_green_0.pdf>.

Versini, P.-A., Gires, A., Tchinguirinskaia, I., Schertzer, D., 2016. Toward an operational tool to simulate green roof hydrological impact at the basin scale: a new version of the distributed rainfall–runoff model Multi-Hydro. Water Science and Technology 74, 1845-1854.

Versini, P.-A., Ramier, D., Berthier, E., De Gouvello, B., 2015. Assessment of the hydrological impacts of green roof: From building scale to basin scale. Journal of Hydrology 524, 562-575.

Woznicki, S.A., Hondula, K.L., Jarnagin, S.T., 2018. Effectiveness of landscape‐based green infrastructure for stormwater management in suburban catchments. Hydrological processes 32, 2346-2361.