Green roof seasonal variation: comparison of the hydrologic behavior of a thick and a thin extensive system in New York City

2016 Environ. Res. Lett. 11 074020

(http://iopscience.iop.org/1748-9326/11/7/074020)

Environmental Research Letters

LETTER

Green roof seasonal variation: comparison of the hydrologic behavior of a thick and a thin extensive system in New York City

R M Elliott[1], R A Gibson[1], T B Carson[1], D E Marasco[1], P J Culligan[1] and W R McGillis[2]

[1] Department of Civil Engineering and Engineering Mechanics, Columbia University, NY, USA
[2] Lamont—Doherty Earth Observatory, Columbia University, NY, USA

E-mail: pjc2104@columbia.edu

Keywords: green roof, urban hydrology, seasonality, antecedent dry weather period, potential evapotranspiration, storm event size, low-impact development

OPEN ACCESS

RECEIVED
21 March 2016

REVISED
31 May 2016

ACCEPTED FOR PUBLICATION
23 June 2016

PUBLISHED
15 July 2016

Abstract

Green roofs have been utilized for urban stormwater management due to their ability to capture rainwater locally. Studies of the most common type, extensive green roofs, have demonstrated that green roofs can retain significant amounts of stormwater, but have also shown variation in seasonal performance. The purpose of this study is to determine how time of year impacts the hydrologic performance of extensive green roofs considering the covariates of antecedent dry weather period (ADWP), potential evapotranspiration (ET_0) and storm event size. To do this, nearly four years of monitoring data from two full-scale extensive green roofs (with differing substrate depths of 100 mm and 31 mm) are analyzed. The annual performance is then modeled using a common empirical relationship between rainfall and green roof runoff, with the addition of Julian day in one approach, ET_0 in another, and both ADWP and ET_0 in a third approach. Together the monitoring and modeling results confirm that stormwater retention is highest in warmer months, the green roofs retain more rainfall with longer ADWPs, and the seasonal variations in behavior are more pronounced for the roof with the thinner media than the roof with the deeper media. Overall, the ability of seasonal accounting to improve stormwater retention modeling is demonstrated; modification of the empirical model to include ADWP, and ET_0 improves the model R^2 from 0.944 to 0.975 for the thinner roof, and from 0.866 to 0.870 for the deeper roof. Furthermore, estimating the runoff with the empirical approach was shown to be more accurate then using a water balance model, with model R^2 of 0.944 and 0.866 compared to 0.975 and 0.866 for the thinner and deeper roof, respectively. This finding is attributed to the difficulty of accurately parameterizing the water balance model.

1. Introduction

Practitioners, policymakers, and researchers have been investigating the use of green infrastructure to reduce the damaging environmental effects of excess stormwater runoff in urban environments (US EPA 2004). One significant opportunity for green infrastructure implementation is on rooftops, which account for approximately 40%–50% of the impermeable urban surface (Stovin *et al* 2012). Many cities, including New York City, which have set benchmarks for local stormwater capture are incentivizing the retrofit of existing buildings with lightweight, extensive green roofs (Carter and Fowler 2008, NYC Department of Buildings 2008). Extensive green roofs require minimal maintenance and consist of several layers, including a vegetation layer (most often in the genus *Sedum*), a 30–150 mm deep substrate layer, and a drainage course.

The ability of an extensive green roof to prevent stormwater runoff depends on the amount of stormwater it can retain during a rain event, which, in turn, depends on its ability to release stored water between rain events. Green roof stormwater retention has been shown to vary with climate, storm size, vegetation type, and season (Mentens *et al* 2006, Lundholm

Table 1. Stormwater retention by season expressed in percent of rainfall retained as reported by Stovin *et al* (2012), Uhl and Schiedt (2008), Kaufmann (1999), Liesecke (2002), and Carson *et al* (2013).

Study		Stovin *et al* (2012)	Uhl & Schiedt (2008)				Kaufmann (1999)	Liesecke (2002)	Carson *et al* (2013)		
Location		Sheffield, UK	Muenster, Germany				Burgdorf, Germany	Hannover, Germany	New York City, U.S.A.		
# of		29	24				48	48	12	12	10
# of Roofs		1	2	4	3	2	8	10	1	1	1
Depth		80	50	80	100	150	100	100	32	100	100
Mean % Retention	Mar	76%	57%	58%	62%	72%	20%	49%	43%	61%	42%
	Apr										
	Ma						48%				
	Jun	52%	73%	71%	76%	83%		70%	70%	43%	72%
	July										
	Aug										
	Sep	35%	57%	58%	62%	72%		49%	22%	40%	78%
	Oct										
	Nov						20%				
	Dec	6%	46%	46%	53%	59%		33%	28%	43%	52%
	Jan										
	Feb										

et al 2010, Voyde *et al* 2010). Green roof hydrological performance is usually assessed as the percent of rainfall captured over a defined period, and can be predicted fairly well using a number of approaches, summarized as: empirical relationships between rainfall and runoff derived from field observations, referred to here as characteristic runoff equations (CREs) (Mentens *et al* 2006, Schroll *et al* 2011) or curve numbers (Carter and Rasmussen 2006, Carter and Jackson 2007), process-based water balance models (Berthier *et al* 2011, Zhang and Guo 2013, Vanuytrecht *et al* 2014), and software such as SWMM (Khader and Montalto 2008, Burszta-Adamiak and Mrowiec 2013) and HYDRUS 1D (Hilten *et al* 2008). Empirical relationships between rainfall and runoff have the advantage of being well-calibrated and simple to use. However, such relationships combine covariates of climate, season, vegetation, and system type making them difficult to generalize. Water balance models and software solutions offer greater predictive capacity, including the ability to account for climatic conditions that lead up to each individual rainfall event. Nonetheless, these models are more complicated to implement and require parameters that are often difficult to measure, such as leaf area index (LAI), plant wilting point, and roof depression storage (Carson *et al* 2015).

Green roof stormwater retention performance has been observed to be consistently higher in the warmer months of the year (Mentens *et al* 2006), largely accounted for by higher potential evapotranspiration (ET_0) (Marasco *et al* 2014). This variability is commonly presented in discrete time periods, often corresponding to seasons. A summary of performance variability reported in prior studies that were undertaken in the northern hemisphere on extensive *Sedum* roofs is provided in table 1, where the greatest percent retention in each study is marked by the darkest shade of red. Despite spanning many different climates within the northern hemisphere, general trends are similar, although the reported magnitude of stormwater retention differs between studies.

While the studies reported in table 1, and others, clearly show a seasonal trend in green roof hydrologic performance, detailed understanding of the primary factors driving the seasonal influence remains lacking. Most investigators include the caveat that the seasonal trends they observe might be masked by the influence of storm event size distribution and the length of antecedent dry weather periods (ADWPs) within each season of their study period (Stovin *et al* 2012, Carson *et al* 2013, Wong and Jim 2014). These confounding variables have been individually examined, with Carson *et al* (2013) finding the seasonal effect to be strongest in storms from 10 to 20 mm in depth, and Poë *et al* (2015) finding the seasonal impact on storage created by ET to be more apparent for events with shorter ADWPs. However, Wong and Jim (2014) were unable to find a significant link between stormwater retention and ADWP.

The purpose of this study is to improve understanding of the factors that contribute to seasonable variability of extensive green roof hydrological performance by (1) providing a context for how ADWP, storm size, and ET_0 relate individually to green roof hydrological performance, and (2) using these variables in adaptive empirical models to determine the significance of their combined effect. To do this, rainfall, runoff, and environmental data were collected over a period of nearly four-years from two full-scale extensive *Sedum* green roofs located in New

Table 2. Summary of storm events in winter (December–February), spring (March–May), summer (June–August), and fall (September–November) considered suitable for analysis.

	Event size (mm)	Spring events (March–May) (#)	Summer events June–August) (#)	Fall Events (September–November) (#)	Winter events (December–February) (#)
W118	0–10	37	56	53	26
	10–20	10	15	4	6
	20–40	10	11	7	5
	40+	5	1	6	4
USPS	0–10	51	30	44	36
	10–20	10	8	8	13
	20–40	9	8	6	9
	40+	3	3	5	4

York City, one 31 mm deep and the other 100 mm deep. The data are then used to identify 503 individual storm events for which event size, rainfall capture, ET_0, and ADWP were determined. The seasonal performance is first evaluated considering the individual factors with a series of exceedance probability (EI) plots. Then a comparative modeling approach is used to determine the significance of *Rf*, ET_0, and ADWP in predicting runoff depth through the development and evaluation of four different models.

2. Methods

2.1. Green roof sites and instrumentation

Monitoring took place in Manhattan, New York, in USDA plant hardiness zone 7B, on two separate extensive green roofs, referred to herein as W118 and USPS. W118 is a Columbia University residence located at 423 West 118th Street (40°48′28″, −73°57′34″) that was outfitted with a Xero Flor America XF301 + 2FL vegetated mat in 2007 (Carson *et al* 2013). The roof is approximately 65 m above mean sea level and has a monitored watershed area of 310 m^2. The growing media is expanded shale, with a water storage capacity of 37%, saturated hydraulic conductivity of 0.021 cm s^{-1}, and relatively thin depth of 32 mm (Marasco *et al* 2014). The vegetated area is 53%, populated with succulent plant species including: *Saxifraga granulata, Sedum acre, Sedum album, Sedum ellacombianum, Sedum hybridum 'Czars Gold', Sedum oregonum, Sedum pulchellum, Sedum reflexum, Sedum sexangulare, Sedum spurium var. coccineum,* and *Sedum stenopetalum*.

The other monitoring site, termed USPS, is atop the US Post Office's Morgan Processing and Distribution Center (40°45′2″, −73°59′55″) where a built-in-place extensive system was installed in 2009. The roof is approximately 45 m above mean sea level and has a monitored watershed area of 390 m^2. The growing media is expanded shale, with a water storage capacity of 35%–65%, saturated hydraulic conductivity of 0.001–0.012 cm s^{-1} and is relatively deep at 100 mm (with a single 2 m × 6 m berm with a depth of 200 mm). The vegetated area is 67% and is primarily populated with the *sedum* species: *Sedum acre, Sedum album 'Coral Carpet', Sedum album murale, Sedum reflexum, Sedum sexangulare, Sedum reflexum 'Blue Spruce', Sedum grisebachii, Sedum kamtschaticum, Sedum 'Matrona', Sedum pluricaule 'Rosenteppich', Sedum spurium 'Roseum',* and *Sedum telephium 'Autumn Joy'*. The 200 mm thick berm is planted with: *Achillea filipendula 'moonshine', Allium schoenoprasum, Coreopsis verticillata 'moonbeam', Silene caroliniana ssp. wherryi, Talinum calycinum,* and *Tradescantia ohiensis* (Marasco *et al* 2014).

Each roof has an Onset Hobo U30 weather station that is instrumented to measure environmental conditions in addition to the water entering the drains and gutters, referred to as runoff. Environmental conditions on each roof are monitored with a THB-M002 2-bit air temperature/relative humidity sensor, an LIB-M003 solar radiation sensor, an S-WCA-M003 wind speed sensor, an S-SMC-M005 EC-5 soil moisture sensor, and an S-RGB-M002 tipping bucket rain gauge. Runoff is measured with custom built and calibrated in-drain V-notch weirs, which continuously measure the flow rate into the drain from watershed areas of 310 m^2 and 390 m^2, for W118 and USPS, respectively (Carson *et al* 2013). A full description of the instrumentation set-up, calibration and monitoring protocols are provided in Culligan *et al* (2014).

2.2. Monitoring data

The continuous runoff monitoring data is processed into discrete storm events using the common criteria that individual storms must be separated by a period of 6 h with no rainfall or, in this case, runoff (Washington State Department of Ecology 2008, Technology Acceptance and Reciprocity Partnership 2001). Once storms are discretized, events that are considered unsuitable for analysis are removed. The following are considered unsuitable events and make a storm unusable for analysis: The peak runoff rate exceeds the limit of the monitoring device (Type 1). Precipitation is in the form of snow or the air temperature is below freezing (Type 2). The

cumulative rainfall is less then the runoff (Type 3). Power to the ultrasonic sensor is interrupted (Type 4). Further details on the instrumentation, parsing, and quality of the monitoring data are available in Carson *et al* (2013).

The data record for W118 runs from 6/29/2011 to 4/7/2015, with 256 usable storms, while the data record for USPS runs from 6/17/2011 to 4/15/2015 with 247 usable storms. A summary of the storms separated by season, with winter (December–February), spring (March–May), summer (June–August), and fall (September–November), is provided in table 2. W118 had 55 unusable storms (20 Type 1, 10 Type 2, 24 Type 3, 1 Type 4) and USPS had 57 unusable storms (0 Type 1, 17 Type 2, 29 Type 3, 11 Type 4). The procedure for maintaining data quality results in a generally lower number of usable winter events, due to snow, freezing air temperature and cold weather related equipment failures.

The storm event size, runoff depth, ADWP and ET_0 are determined for each suitable event. Event size (*Rf*) is the total depth of rain per unit rooftop area and is determined from the tipping bucket data. Runoff (*Ro*) depth is the height of runoff generated per unit rooftop area per storm and is calculated using 5 min flow rate data from the weirs. The ADWP is taken to be the dry period leading up to the rain event as recorded by the tipping bucket, given in days. To capture the variation in the potential of the system to expel water, the potential or reference evapotranspiration (ET_0) is calculated using the well-documented and simple to use Hargreaves and Samani equation (Hargreaves *et al* 1985):

$$ET_0 = 0.0022*RA*(TC - 17.8)*TD^{0.5}, \quad (1)$$

where ET_0 is the reference ET (mm d^{-1}), RA is extraterrestrial radiation (mm equivalent per day) calculated using the day of year and location latitude as described in Allen *et al* (1998), TC is average daily temperature (°C), and TD is the daily temperature range (°C). TC and TD are obtained from Belvedere Castle weather station in Central Park, NYC, which is maintained by the NOAA National Climatic Data Center (ncdc.noaa.gov).

2.3. Comparative modeling approach

The significance of *Rf*, ET_0, and ADWP in predicting runoff depth, *Ro*, is explored using a comparative modeling approach where predicted *Ro* is compared with observed values of *Ro* for four different models. The first (base) model is the empirical CRE, derived from observations between *Rf* and *Ro* for each green roof. CREs have been used by many investigators (Mentens *et al* 2006, Schroll *et al* 2011, Stovin *et al* 2012, Fassman-Beck *et al* 2013, Carson *et al* 2013) to predict green roof performance. Following Carson *et al* 2013, this study uses a quadratic form of the CRE, namely:

$$Ro = C_1Rf^2 + C_2Rf + C_3, \quad (2)$$

where C_1, C_2, and C_3, are empirical fitting coefficients.

For the second model, the CRE is modified to include Julian Day (JD), via:

$$Ro = C_1Rf^2 + C_2Rf + C_3 + f(\mathrm{JD}). \quad (3)$$

Julian Day embodies many of the climatic factors controlling seasonal runoff performance, and is thus a simple means of capturing green roof performance variability for a particular climate region. To determine *f*(JD), a genetic programing (GP) symbolic regression algorithm (Schmidt and Lipson 2009) is used. The GP algorithm generates a population of models and uses stochastic methods to 'evolve' models according to a set of rules (Koza 1992), resulting in equations that best fit the data (as defined by R^2) at several levels of complexity. Similarly, the GP algorithm is used to determine the form of the model functions, *f*, and coefficients in following equations (4)–(6).

For the third model, the CRE is modified to include a function ET_0:

$$Ro = C_1Rf^2 + C_2Rf + C_3 + f(ET_0). \quad (4)$$

Equation (4) allows the modified CRE to reflect overall seasonal change, using ET_0 as a surrogate for season that is more generalizable than JD.

For the fourth model, the CRE is modified to include ET_0, ADWP, *Rf*, and their combinations, as described by:

$$Ro = C_1Rf^2 + C_2Rf + C_3 + f(ET_0) + f(ET_0, Rf) + f(ET_0, \mathrm{ADWP}). \quad (5)$$

Equation (5) allows the modified CRE to reflect overall seasonal change, *f*(ET_0), seasonal change with storm size, *f*(ET_0, *Rf*), and seasonal change with the antecedent dry period, *f*(ET_0, ADWP).

For the fifth model, a simplified reservoir equation (SRE) is evaluated, taking the form:

$$Ro = \max(0, Rf - \max(\mathrm{Kc}*\mathrm{AWDP}*\mathrm{ETo}, S_{\max})), \quad (6)$$

where Kc is the crop coefficient, a unit-less coefficient that accounts for a plants' ability to dispel water and S_{max} is the green roof's maximum available water storage in mm.

In equation (6), *Ro* is predicted as *Rf* minus the available storage in the substrate, unless the available storage is larger then the rainfall—in which case no runoff is generated. The available storage, limited to S_{max}, is estimated using the product of the dry period before the storm (ADWP in days), the potential evapotranspiration (ETo in mm d^{-1}), and the crop coefficient (Kc). For both of the roofs that are the subject of this study, Kc and S_{max} are determined by a best fit of predicted *Ro* to observed *Ro*.

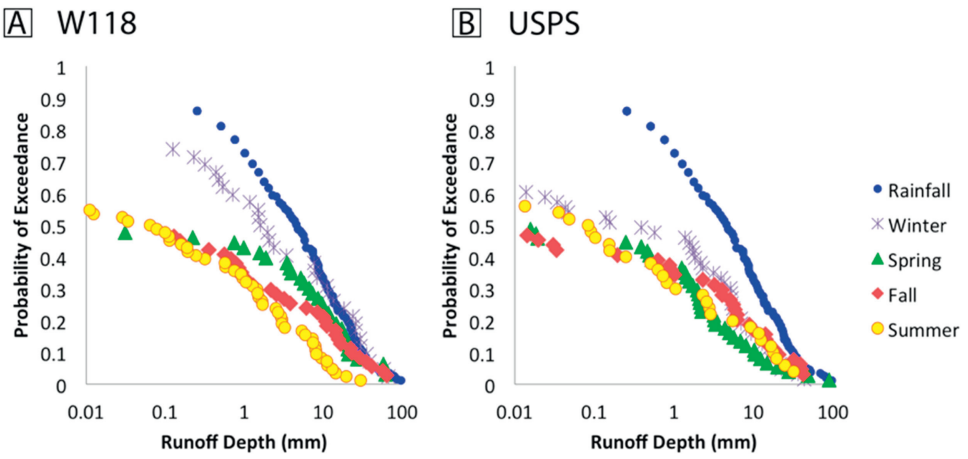

Figure 1. Event exceedance probability for runoff depth separated by winter (December–February), spring (March–May), fall (September–November), and summer (June–August) for W118 (A) and USPS (B).

The accuracy of each model is quantified using recommended statistics for hydrologic model performance; namely, the Nash–Sutcliffe efficiency index (NSE), root mean square error (RMSE) to the standard deviation of measured data (RSR), and percent bias (PBIAS) (Moriasi *et al* 2007, Hakimdavar *et al* 2014). The NSE index ranges between 1 and −1, where 1 represents perfect equivalency and negative values indicate the model is less accurate than the mean value of the observed values (Nash and Sutcliffe 1970). RSR ranges from 0 to a large positive value, where a low value represents a lower RMSE, and better model performance. The magnitude of PBIAS represents how biased the model is, with a positive value representing over-prediction and a negative value representing under-prediction. According to Moriasi *et al* (2007), a watershed model is satisfactory if NSE $>$ 0.5, RSR $\leq$ 0.7 and PBIAS is within $\pm$25%.

3. Results and discussion

3.1. Observed Seasonal behavior

Figures 2–5 present event EP plots, where the EP for each event is calculated as rank/(n + 1), for runoff depth separated by season, ET_0, ADWP and storm size, respectively. The general weather patterns during this study period were similar to a 40 year historical period, with most monthly averages for storm size, ADWP, and ET_0 from the study period falling within the 1st and 3rd quartile of a 40-year historical period. An in depth analysis and comparison of these factors during both the study and a 40 year period spanning from 1971 to 2010 is available in appendix A.

As seen in figure 1, both roofs are more likely to generate higher stormwater runoff values for a given storm size in the winter, which agrees with the findings of Uhl and Schiedt (2008), Kaufmann (1999), Liesecke (2002), and Carson *et al* (2013). Additionally, the thinner W118 extensive green roof demonstrates greater seasonal variability in runoff performance than the deeper USPS roof, with lower runoff depths being much more likely to be exceeded on W118 in the winter than the summer. For example, runoff depths of 1 mm have an EP difference of 0.28 between summer and winter on W118, and an EP difference of 0.16 between summer and winter on USPS. The behavior of each season is statistically unique except for spring and summer for USPS, as demonstrated in the statistical analysis presented in appendix B. In terms of annual performance, the deeper USPS substrate layer is less likely to exceed nearly all runoff depths compared to W118.

The contribution to performance that would vary with season is presumed to be storage created by evapotranspiration, while the ability of the substrate to detain water is expected to be consistent year round (considering days below freezing are removed from the data set). As USPS has a much deeper substrate, the generally higher retention and more consistent performance could be explained by the substrate's detention capacity having a greater contribution to stormwater retention than storage created by ET.

Figure 2 shows the data separated by ET_0 groupings of $0 < ET_0 < 2$, $2 < ET_0 < 4$, and $ET_0 > 4$ mm/day with the behavior of each group being statistically different (see appendix B). While these groupings roughly correspond to the winter, spring/fall, and summer trends, respectively, they more specifically define the environmental conditions surrounding the storm event. As seen, for both roofs, grouping by ET_0 (figure 2) shows a clearer separation in hydrologic performance than grouping by season

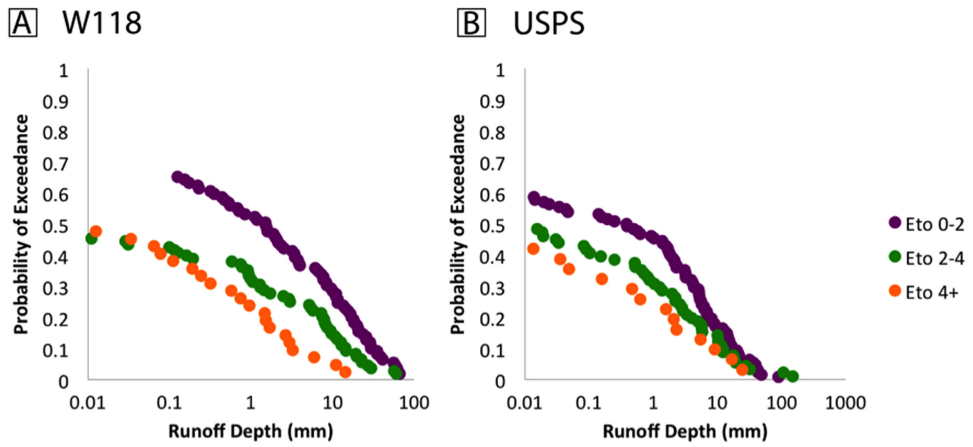

Figure 2. Event exceedance probability for runoff depth separated by ET_0 (0–2, 2–4, 4+ mm d^{-1}).

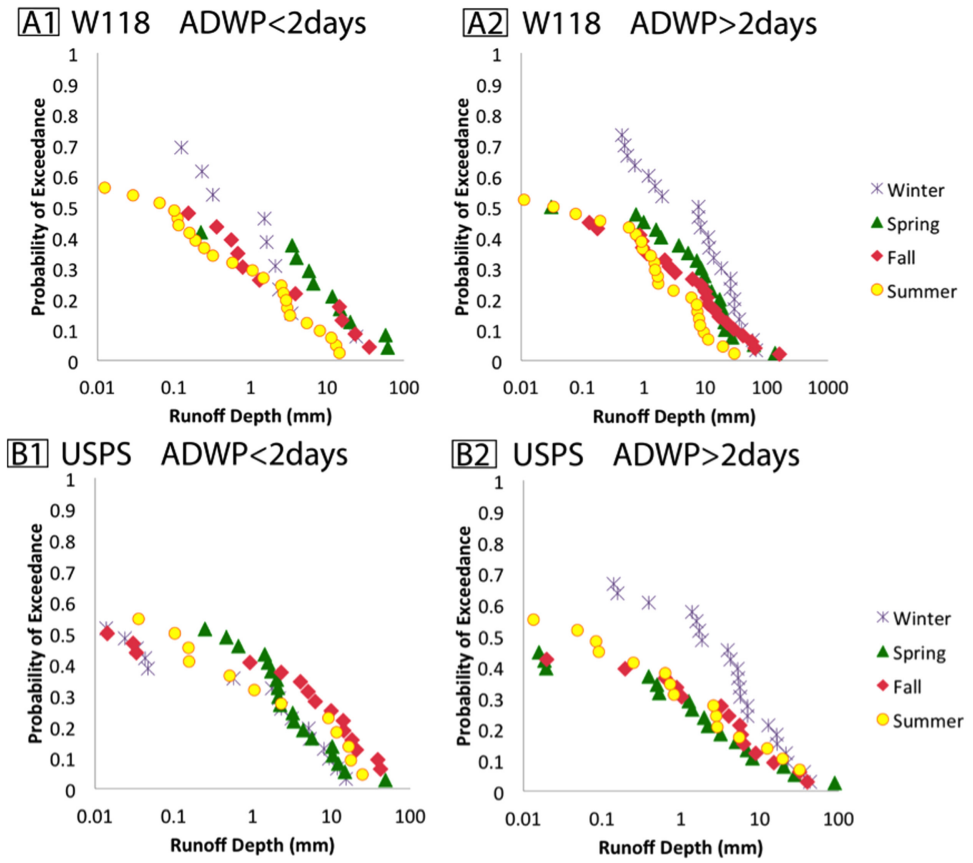

Figure 3. Event exceedance probability for runoff depth in events with different dry periods and separated by season for W118 and USPS events with short (<2 day) (A1) and (B1) and long (>2 day) (A2) and (B2) antecedent dry weather periods (ADWPs).

(figure 1). This is especially evident in USPS, where figure 2(B) shows distinctly separate ET_0 group behavior over the full range of runoff depths, while figure 1(B) showed less distinct seasonal trends, with summer, spring and fall closely tracking one another. This increased distinction can be attributed to ET_0 being more physically defined than season, specifically, the $ET_0 > 4$ category separating only the highest ET_0 events of summer.

For extremely small runoff depths (0.01–0.1 mm), both the medium ($2 < ET_0 < 4$) and high ($ET_0 > 4$) ET_0 conditions exhibit the same probability of exceedance for W118. Extremely small runoff depths on W118 were associated with rainfall events that did not saturate the roof. Thus, it is possible that runoff during such events is generated by flow that occurs along preferential pathways in the thin W118 system, and that such flow is largely independent of antecedent

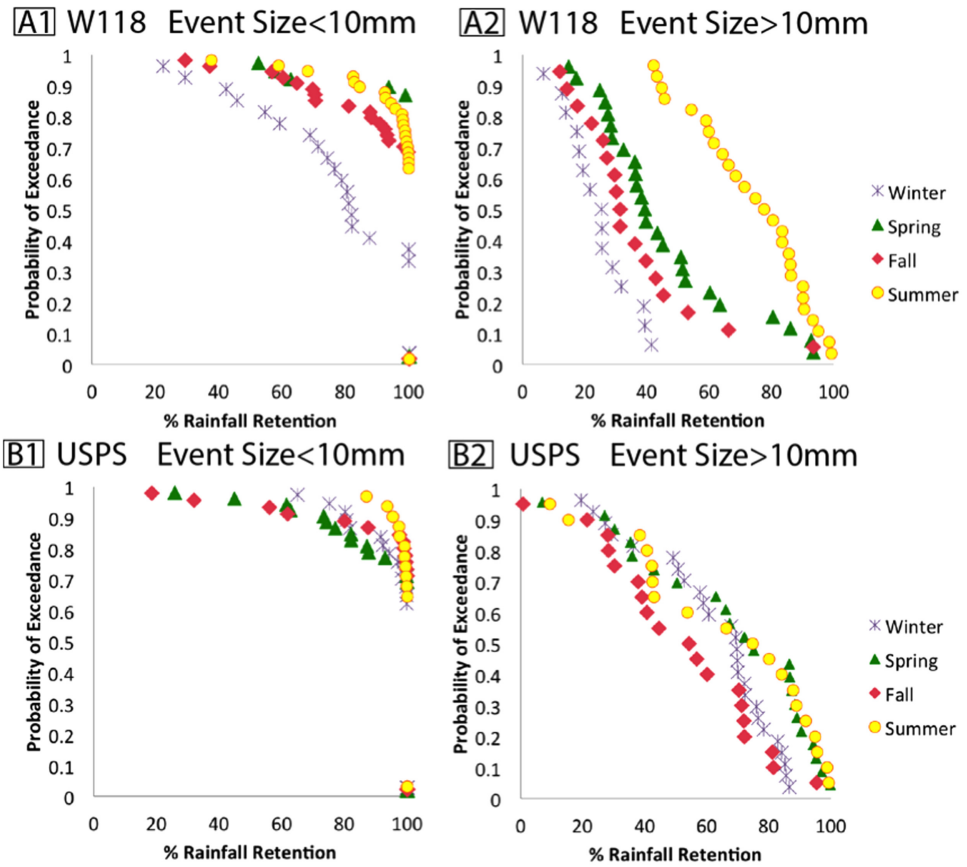

Figure 4. Event exceedance probability for % rainfall retention separated by season for W118 and USPS in small (<10 mm) (A1) and (B1) and large (>10 mm) (A2) and (B2) storm event sizes.

moisture conditions when the roof is relatively dry. Nonetheless, it is important to note there are relatively few runoff depth data points within the range 0.01–0.1 mm for W118, and that the smaller watershed area of W118, in comparison to USPS, make the weir-based measurement system less accurate for extremely small runoff depths on W118. Hence, confirmation of this particular trend really requires further data.

Figure 3 shows the seasonal behavior in runoff separated by short (<2 day) and long (>2 day) ADWPs. For both extensive green roofs, seasonal differences are more apparent with longer ADWPs. For example, the runoff behavior in the winter is statistically different from the other months only in the >2 day plots. Winter also has the pronounced segregation in the >2 day plots, with 1 mm runoff depths having an EP difference of 0.25 and 0.29 between summer and winter for W118 and USPS, respectively, while short ADWPs have corresponding differences of 0.18 and 0.01, respectively.

The greater seasonal variation with longer ADWPs can be explained by the greater storage created by ET given a longer dry period before a storm. As ET_0 is highly variable with season, one would expect events with longer ADWPs to have more pronounced seasonal variation. While Poë *et al* (2015) found the seasonal impact on storage created by ET to be more apparent in events with shorter ADWPs, the results presented here suggest that the seasonal variation in runoff increases with ADWP. Additionally, the finding here differs from Wong and Jim (2014) who found no significant link between runoff and ADWP.

Figure 4 illustrates the seasonal percent rainfall retention for W118 and USPS separated by storm events smaller and larger than 10 mm. Both roofs retain 100% of the rainfall in most smaller storms (<10 mm) and show a range of behavior in larger storms (>10 mm). As in figures 1–3, figure 4 shows a more pronounced seasonal variation in W118, particularly for events larger then 10 mm. This agrees with Carson *et al* (2013), who found the seasonal impact to be strongest in storms from 10 to 20 mm in depth, and can be explained by the ability of the green roof to fully retain most small storms regardless of the seasonal climate. Alternatively, USPS has no apparent difference in seasonal performance with storm size. This can potentially be explained by the greater contribution of substrate detention to stormwater capture, as discussed above. For events larger than 10 mm, all four seasons were statistically different for

Table 3. Various equations used to model runoff (mm) where *Rf* is storm event size (mm) JD is Julian day, ET_0 is potential evapotranspiration (mm d^{-1}), and ADWP is the antecedent dry weather period (days).

	Equation Form	Runoff =	(R^2)	NSE	RSR	PBIAS(%)
W118	CRE (base model) [10.7 < *Rf* < 100 mm]	$1.05 * Rf - 0.002\,28 * Rf^2 - 11.12$	0.945	0.948	0.227	−9.41
	CRE + *f*(JD) [2.7 < *Rf* < 100 mm]	$0.84 * Rf + 5 * \sin(1.47 + 0.0179 * JD) - 6.86$	0.980	0.978	0.147	−2.50
	CRE + *f*(ET_o) [3.0 < *Rf* < 100 mm]	$0.874 * Rf + 0.001\,03 * Rf^2 - 0.854 - 2.59 * ET_o$	0.975	0.975	0.157	−3.39
	CRE + *f*(ET_o,ADWP,*Rf*) [3.5 < *Rf* < 100 mm]	$0.865 * Rf - 7.47 \times 10^{-5} * Rf^2 - 1.28 + ET_o(-0.778 - 0.0567 * Rf - 0.110 * ADWP)$	0.975	0.973	0.164	−2.22
	SRE	$\max(0, Rf - \max(11.35, 0.659 * ADWP * ET_o)$	0.949	0.931	0.263	−6.82
USPS	CRE [6.2 < *Rf* < 100 mm]	$0.444 * Rf + 0.003\,27 * Rf^2 - 2.926$	0.866	0.866	0.366	−1.67
	CRE + *f*(JD) [7.1 < *Rf* < 100 mm]	$0.674 * Rf + 1.3 * \sin(2.46 + 0.0172 * JD) - 6.09$	0.868	0.868	0.362	−3.91
	CRE + *f*(ET_o) [2.2 < *Rf* < 100 mm]	$0.334 * Rf + 0.004\,69 * Rf^2 - 0.363 - 0.572 * ET_o$	0.866	0.866	0.365	−0.84
	CRE + *f*(ET_o,ADWP,*Rf*) [5.3 < *Rf* < 100 mm]	$0.469 * Rf + 0.002\,83 * Rf^2 - 1.995 + ET_o(-0.324 - 0 * Rf - 0.106 * ADWP)$	0.871	0.869	0.361	−2.53
	SRE	$\max(0, Rf - \max(16.71, 0.577 * ADWP * ET_o)$	0.836	0.794	0.453	−16.4

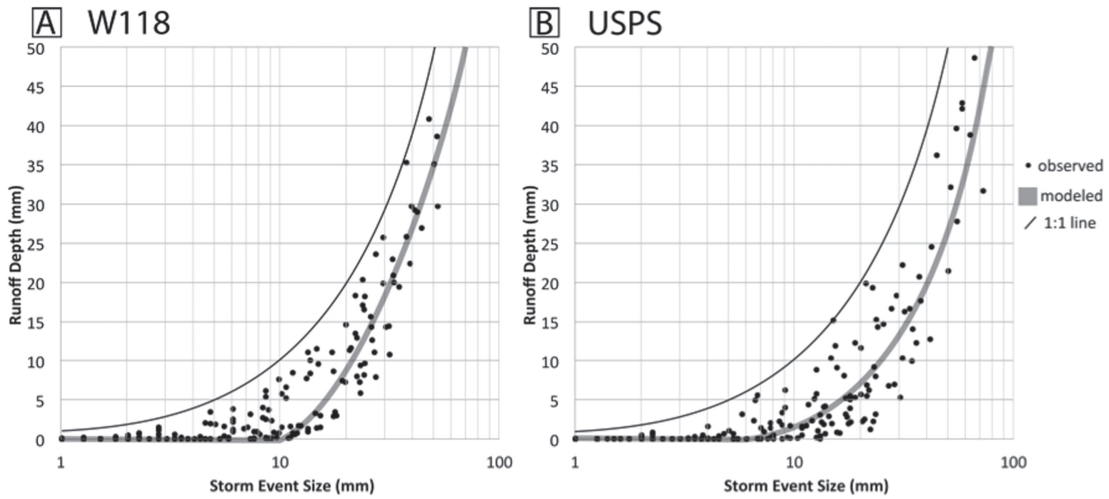

Figure 5. Observed storm event size and runoff depth along with the CRE model for W118 and USPS.

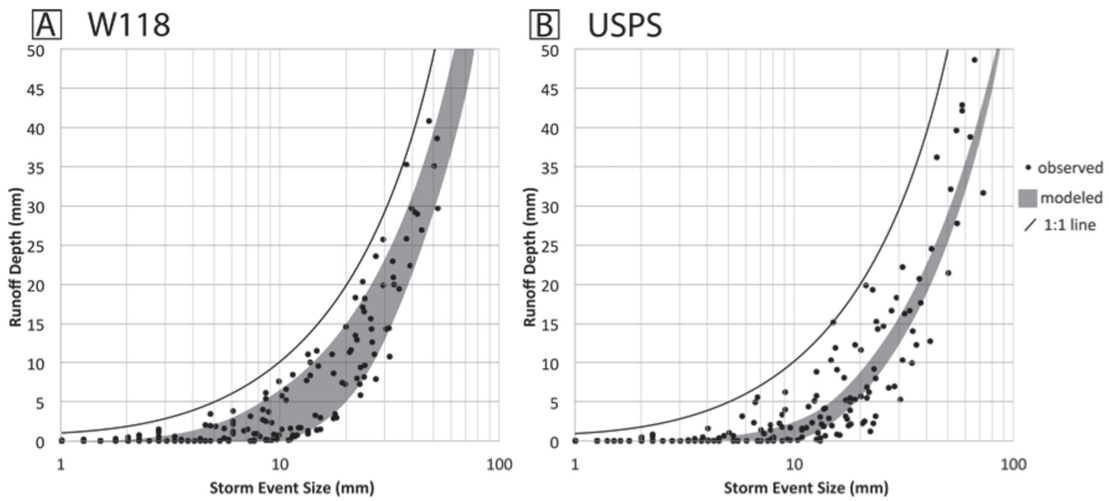

Figure 6. Observed storm event size and runoff depth along with the range of the CRE modified with Julian Day for W118 and USPS.

W118 compared to only fall being statistically different for USPS.

3.2. Modeling results

The fitting results of the genetic programing algorithm are provided in table 3 for the base CRE (2), CRE modified to include Julian Day (3), CRE modified to include ET_0 (4), CRE modified to include ET_0, ADWP, and *Rf* (5) and the SRE (6), along with measures of accuracy.

All of the models' statistics fall well within the satisfactory levels set (NSE $>$ 0.5, RSR $\leqslant$ 0.7 and PBIAS within $\pm$25%) (Moriasi *et al* 2007), however, there are notable differences in their behavior and accuracy. For both roofs, the SREs have the least accurate performance followed the CREs. Both modifications of the CRE increase accuracy, with the CRE modified to include Julian Day most accurate for W118 and the CRE modified to include ET_0 and ADWP most accurate for USPS.

For both roofs, the CRE is able to produce an adequate prediction, shown below in figure 5. As runoff is only estimated using rainfall, the model performance appears as a line.

The CRE modified to incorporate Julian Day (3) has improved model statistics in W118 (R^2 improved from 0.945 to 0.980) and in USPS, (R^2 improved from 0.866 to 0.868). In both cases, the seasonal variance takes the form of a sine curve with a period of approximately 1 year (351 days for W118 and 365.3 days for USPS) and the least runoff is generated around midsummer (July 17th for W118 and June 30th for USPS). The model, illustrated in figure 6, reveals the amplitude of the seasonal variance in W118 to be nearly four times that of USPS (sine coefficient of 5 for W118 and 1.3 for USPS).

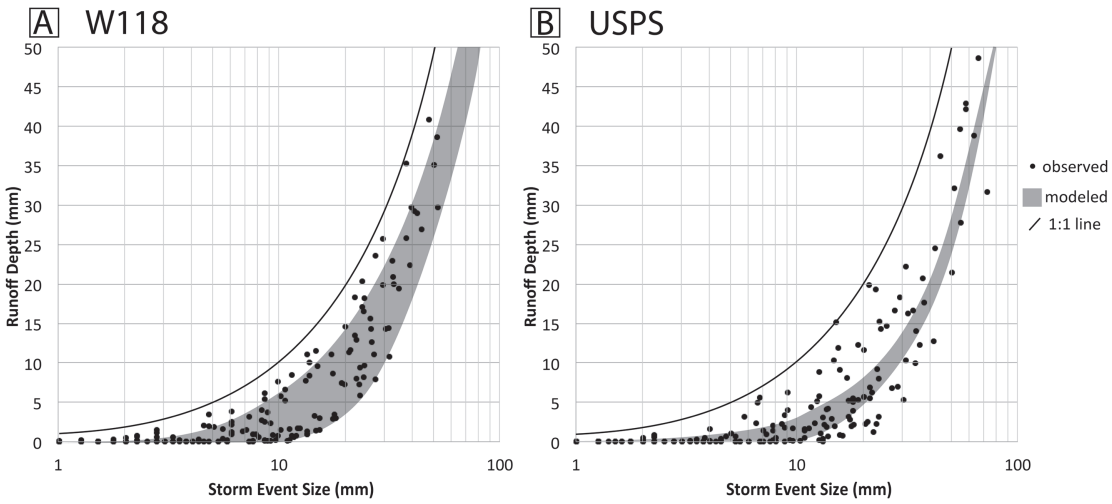

Figure 7. Observed storm event size and runoff depth along with the range of the CRE modified with ET_0 for W118 and USPS.

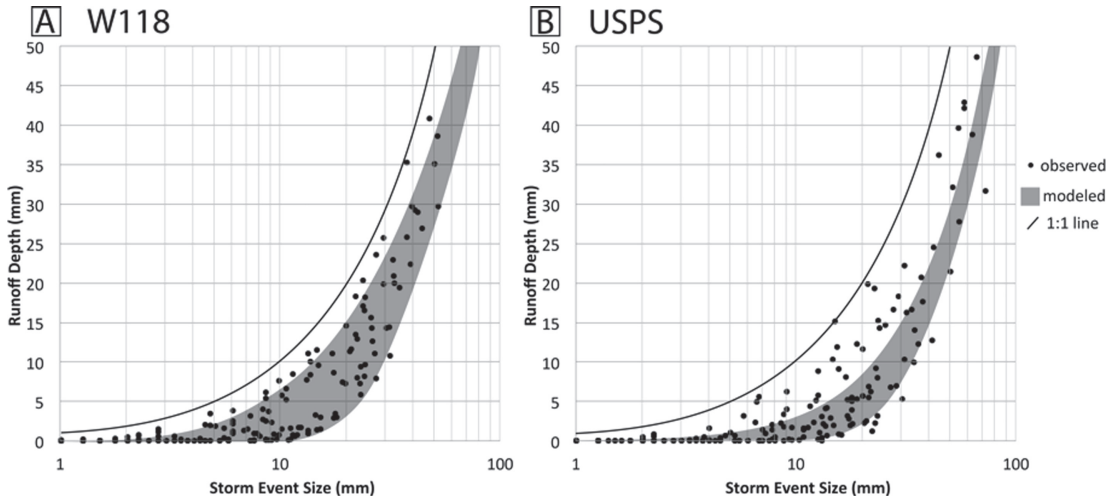

Figure 8. Observed storm event size and runoff depth along with the range of the CRE modified with ET_0 and ADWP for W118 and USPS.

When modified to include ET_0 (4), the model performs with greater accuracy than the base CRE and is close to the Julian Day model. The model behavior is shown below in figure 7.

The CRE modified to include ET_0 and ADWP (5) performs better than using ET_0 alone and is on par with the Julian Day model. Similarly, the seasonal variance and runoff generated is found to be greater in W118, as shown in figure 8.

As a well-fitting and most generally applicable model, the behavior of equation (5) over the course of the year is further illustrated in figure 9.

All coefficients in equation (5) associated with the ET_0 term are found to be negative, resulting in both roofs generating the least runoff when the potential evapotranspiration is highest, as seen in figure 9. Furthermore, the coefficients for all terms including ET_0 are higher for W118, resulting in greater variation throughout the year. W118 is found to have a negative coefficient associated with the ET_0 * *Rf* term, causing the seasonal variation in runoff to decrease with event size. Alternatively, USPS is found to have an insignificant ET_0 * *Rf* term, resulting in the seasonal variation in runoff to not change with event size. This behavior confirms observations in figure 5 and is evident in figure 9; the amplitude of seasonal change is reduced in smaller storms in W118 while remaining consistent with event size in USPS.

Both roofs are found to have a negative coefficient associated with the ADWP * ET_0 term, resulting in less runoff generated in events with longer antecedent dry periods. The impact of ADWP is found to be similar in both roofs, with coefficients of 0.110 and 0.106 for W118 and USPS, respectively. The 0, 3, and 10 day

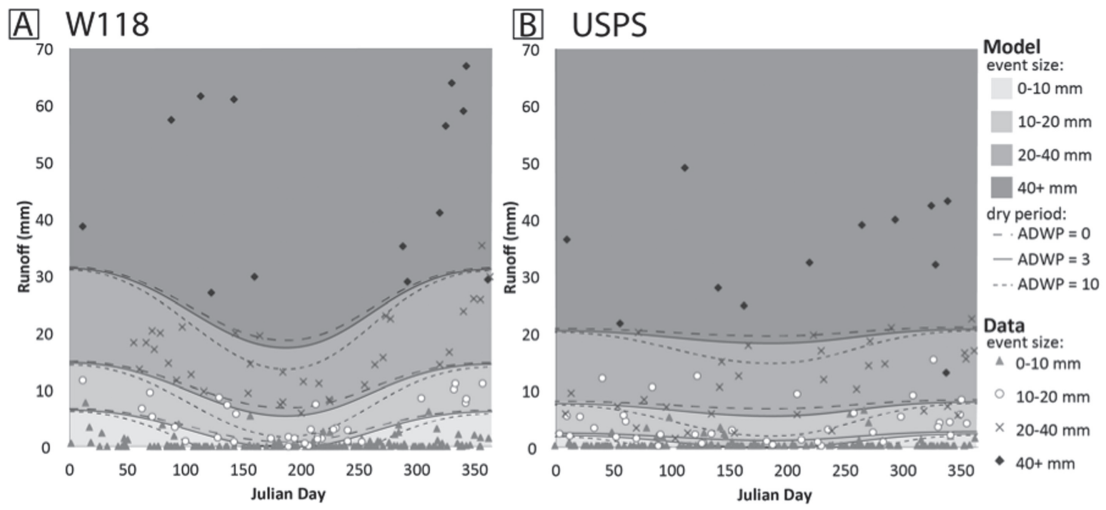

Figure 9. Observed and modeled (CRE modified with ET_0 and ADWP) runoff depth throughout the year for storm event sizes 0–10 mm, 10–20 mm, 20–40 mm, and 40+ mm. Using average historical ET_0 the model shows the behavior during low (0 days), average (3 days), and high (10 days) antecedent dry day conditions.

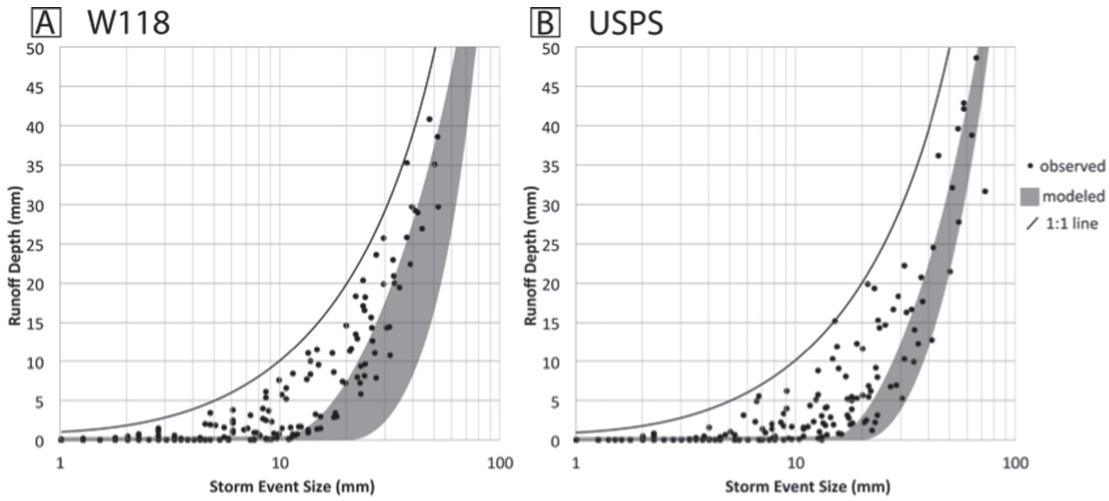

Figure 10. Observed storm event size and runoff depth along with the range of the SRE for W118 and USPS.

ADWP lines in figure 9 illustrate this impact on stormwater retention throughout the year. In colder months ADWP plays a less significant role, with the modeled difference between the 0 and 10 day ADWP smallest in January and February.

The SRE (figure 10), while being the most physically rational, has the worst model performance considering the NSE and RSR for both roofs in addition to R^2 and PBIAS for USPS. The fitted crop coefficients (Kc) are found to be 0.66 for W118 and 0.58 for USPS, which generally agree with previously reported crop coefficients as determined using an energy balance model (Kc of 0.52; Olivieri *et al* 2013), weighing lysimeter (Kc of 0.53; Sherrard Jr and Jacobs 2012), and the FAO-24 method (Kc of 0.85–1.01; Voyde 2011). The fitted maximum storage depth (S_{max}) are found to be 11.4 mm for W118 and 17.1 mm for USPS, which while sensible, are smaller then the S_{max} reported by roofs (17 mm for W118 and 52 mm for USPS). The inability of the SRE to accurately estimate individual storms is attributed to the difficulty of accurately parameterizing the water balance model. Primarily, runoff can occur before the substrate is fully saturated, as discussed in section 3.1, however this behavior is discounted in the simple reservoir equation.

4. Conclusions

Using runoff observations from 503 storms across two extensive green roofs located in New York City, one 31 mm and one 100 mm deep, individual factors of ADWP, storm event size, and potential

evapotranspiration (ET_0) are examined with season. Modeling green roof runoff with these factors confirms several initial observations; stormwater retention was the greatest in warmer months, the deeper roof shows less seasonal variation, roofs retain more rainfall with longer ADWPs, and the thinner green roof system has greater seasonal variation with storm size. Predictive equations are developed for both roofs, with 98% and 87% of the variance in measured runoff accounted for in the thin and deep roof, respectively. This study has shown that full-scale extensive green roofs vary in their ability to retain stormwater throughout the year, and that including seasonal factors can improve runoff model accuracy. While this improvement is clear in the thinner green roof (W118), it is not as significant on the deeper system (USPS).

A limitation of this study is the lack of physical basis for the modified CRE and the inability of the climatic factors to fully represent hydrologic processes. While in reality, the ET_0 of the entire ADWP impacts the amount of storage created through evapotranspiration, only the ET_0 of the event day is used here. Additionally, the plants ability to transpire water (Kc) is expected to vary throughout the year, which is not accounted here. Finally, ADWP does not take into account the extent that the previous storm saturated the substrate, which would impact the amount of available storage and runoff performance accordingly. Further study and modeling work should extend to include these considerations in addition to variables of climate, slope, and plant type.

Acknowledgments

This work is supported, in part, by the National Science Foundation (NSF) grant CMMI-0928604 and the Environmental Protection Agency grant AE-83481601-1. The authors Robert Elliott, Tyler Carson and Daniel Marasco gratefully acknowledge the support of the NSF Integrative Graduate Education and Research Training (IGERT) Fellowship #DGE-0903597. Furthermore, Robert Elliott gratefully acknowledges the support of the NSF Graduate Research Fellowship Program #DGE-11-44155. Any opinions, findings, and conclusions expressed in this paper are those of the author and not meant to represent the views of any supporting institution.

Appendix A. Historical context of monitoring period

In order better understand the context of the climate conditions under which this study occurred, the data from the study period is compared to historical data for storm size, antecedent dry weather (ADWP), and potential evapotranspiration (ET_0). The historic data was obtained from the Belvedere Castle weather station in Central Park, NYC, for a 40 year period spanning from 1971 to 2010. The occurrence of each variable was observed by month, using box plots to compare the median values observed in the study period to the minimum, first quartile, median, third quartile, and maximum values observed in the historical period.

Figure A1 presents storm size, ET_0, and ADWP data, respectively, with boxplots showing the historical range for the years 1971–2010, and points showing the median values for the study period. For all climatic factors, the observed study period medians generally fall within the range of the 1st and 3rd quartile of historical data, which is considered acceptable for the purpose of generalizing some of the results of this study.

As seen in figure A1(A), there is not a strong correlation between storm size and month of the year in NYC, although the median storm size is slightly lower in the spring and summer months than other times of the year. An exceptional deviation in storm size between monitoring period trends and the historical data is noted in February. This deviation is explained by the event suitability criteria used for the monitoring period, which involved ignoring storms with snow or freezing air conditions.

Figure A1(B) reveals the typical antecedent dry weather period to be shorter in the spring and summer months for both the historical data and the study period. The median ADWP for the study period is greater than the historical median for every month except July. This bias is likely due either to the methods used to separate storms; the historical data required a 6 h inter-event period with only no rainfall, while the study data required a 6 h inter-event period with no rainfall or runoff, effectively reducing the amount of very low ADWP events in the study period dataset.

Figure A1(C) shows the potential evapotranspiration to vary greatly with season, with the warmer months having the greatest ET_0. For most months the study period median falls closely to the historical monthly median. The abnormally low ET_0 in the study period for February and July can be explained by the deviations in temperature. While the average temperatures are close for the two data sets, the daily temperature fluctuations are typically larger in the historical period, with average fluctuations of 7.5 °C (historical) and 6.4 °C (study) in January and average fluctuations of 9.0 °C (historical) and 8.3 °C (study) in July.

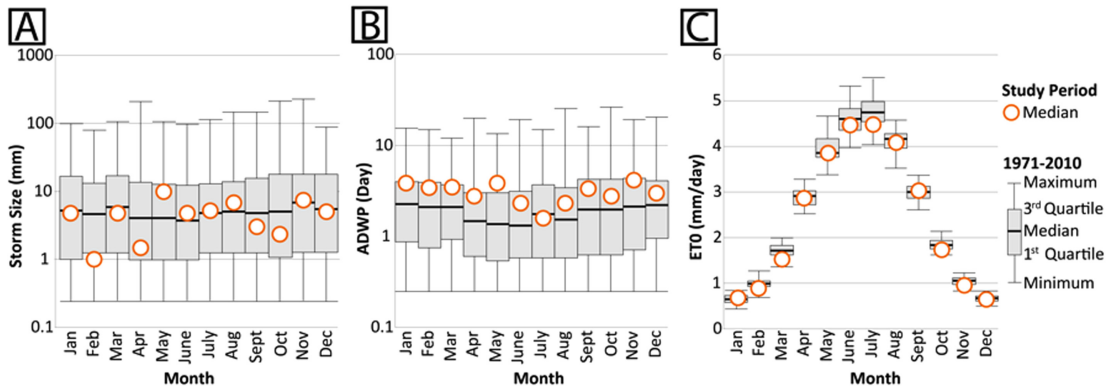

Figure A1. Storm size (A), ADWP (B) and Potential evapotranspiration (C) distribution for the study period compared to historical records.

Table B1. Summary of ANCOVA set-up and results for the event exceedance probability for runoff depth separated by winter (December–February), spring (March–May), summer and fall (September–November), ($\alpha = 0.05$).

	Data set	Truncation		ANCOVA grouping results			
Roof	Corresponding figure	Minimum runoff	Maximum runoff	Winter	Spring	Summer	Fall
W118	Figure 1(A)	0.1 mm	100 mm	A	B	C	D
USPS	Figure 1(B)	0.1 mm	100 mm	A	BC	BC	D

Table B2. Summary of ANCOVA set-up and results for the event exceedance probability for runoff depth separated by ET_0 (0–2, 2–4, 4+ mm d^{-1}) ($\alpha = .05$).

	Data Set	Truncation		ANCOVA grouping results		
Roof	Corresponding figure	Minimum runoff	Maximum runoff	ET_0 0–2	ET_0 2–4	ET_0 4+
W118	Figure 2(A)	0g.1 mm	100 mm	A	B	C
USPS	Figure 2(B)	0.01 mm	100 mm	A	B	C

Appendix B. Statistical analysis

B1. Event EI plots

To test the uniqueness of the trends within each event EP plot, multiple analyses of covariance (ANCOVA) were run. The ANCOVA was set up using the probability of exceedance as the dependent variable, the runoff depth or ETo as the covariate, and the season or ETo grouping, respectively, as the categories being compared.

To meet the assumptions of linearity and normal distribution, the runoff depth data was truncated and log transformed. The ANCOVA was then run setting the level of significance as $p > 0.05$. The corresponding plots, data range, and ANCOVA results are provided in the summary tables below.

Table B1 shows that, when the W118 data grouped by season, the runoff behavior for each category is statistically different. Furthermore, when the same analysis is applied to USPS fall and winter are unique, while spring and summer are not statistically different.

Table B2 shows that when the data are grouped by ET_0 ranges, the runoff behavior for each category is statistically different for both W118 and USPS.

Table B3 shows the difference in seasonal groups when separated into long and short ADWPs. For W118, winter and spring are not statistically different for the shorter dry weather periods (<2 days) while all seasons are statistically different for the longer dry weather period (>2 days). For USPS winter and summer are not statistically different for the shorter dry weather period while only winter is shown to be statistically different for the longer dry weather period.

Table B4 shows the difference in seasonal groups when separated by storm size. For W118 only winter is statistically different for the smaller storms while all seasons are statistically different for the larger storms. Alternatively, for USPS only summer is statistically different for the smaller storms while only fall is statistically different for the larger storms.

B2. Runoff-rainfall plots for W118 and USPS

To determine if the two roofs have distinct runoff behavior, an ANCOVA was performed with each roof as a categorical input and rainfall as the covariate. Both

Table B3. Summary of ANCOVA set-up and results for the event exceedance probability for runoff depth in events with different dry weather periods and separated by season (winter (December–February), spring (March–May), summer (June–August) and fall (September–November)) for W118 and USPS events with short (<2 day) and long (>2 day) antecedent dry weather periods (ADWPs) ($\alpha = .05$).

Data set			Truncation		ANCOVA grouping results			
Roof	ADWP	Corresponding figure	Minimum runoff	Maximum runoff	Winter	Spring	Summer	Fall
W118	<2 Days	Figure 3(A1)	0.1 mm	100 mm	AB	AB	C	D
W118	>2 Days	Figure 3(A2)	1 mm	100 mm	A	B	C	D
USPS	<2 Days	Figure 3(B1)	0.1 mm	100 mm	AC	B	AC	D
USPS	>2 Days	Figure 3(B2)	0.1 mm	100 mm	A	BCD	BCD	BCD

Table B4. Summary of ANCOVA set-up and results for the event exceedance probability for % rainfall retention separated by season (winter (December–February), spring (March–May), summer (June–August) and fall (September–November)) for W118 and USPS in small (<10 mm) (A1 and B1) and large (>10 mm) (A2 and B2) storm event sizes ($\alpha = 0.05$).

Data Set			Truncation		ANCOVA grouping results			
Roof	Event Size	Corresponding figure	Minimum rainfall retention	Maximum rainfall retention	Winter	Spring	Summer	Fall
W118	<10 mm	Figure 4(A1)	0%	70%	A	BCD	BC	BD
W118	>10 mm	Figure 4(A2)	0%	100%	A	B	C	D
USPS	<10 mm	Figure 4(B1)	0%	95%	AD	BD	C	ABD
USPS	>10 mm	Figure 4(B2)	0%	100%	AC	BC	ABC	D

the runoff and rainfall were log-transformed to satisfy the assumption of normal distribution. Additionally, only points with runoff greater than 2 mm were used to meet the assumptions of linearity and similar slopes. The results of ANCOVA revealed that the two roofs have significantly different behavior, with a p-value of 4.95E-05 that falls within the 95% confidence level ($\alpha = .05$).

References

Allen R G, Pereira L S, Raes D, Smith M and Ab W 1998 *Crop evapotranspiration—Guidelines for Computing Crop Water Requirements—FAO Irrigation and Drainage Paper 56* pp 1–15

Berthier E, Ramier D and Gouvello B D 2011 Simulation of green roof hydrological behavior with a reservoir model *2nd Int. Conf. on Urban Drainage (Porto Alegre, Brazil)* pp 10–5

Burszta-Adamiak E and Mrowiec M 2013 Modelling of green roofs' hydrologic performance using EPA's SWMM *Water Sci. Technol.* **68** 36–42

Carson T B 2014 Evaluating green roof stormwater management in New York City: observations, modeling, and design of full-scale systems *Doctoral Dissertation* Columbia University

Carson T B, Keeley M, Marasco D E, McGillis W R and Culligan P J 2015 Assessing methods for predicting green roof rainfall capture: a comparison between full-scale observations and four hydrologic models *Urban Water J.* **9006** 1–15

Carson T B, Marasco D E, Culligan P J and McGillis W R 2013 Hydrological performance of extensive green roofs in New York City: observations and multi-year modeling of three full-scale systems *Environ. Res. Lett.* **8** 024036

Carter T and Fowler L 2008 Establishing green roof infrastructure through environmental policy instruments *Environ. Manage.* **42** 151–64

Carter T and Jackson C R 2007 Vegetated roofs for stormwater management at multiple spatial scales *Landsc. Urban Plan.* **80** 84–94

Carter T L and Rasmussen T C 2006 Hydrologic behavior of vegetated roofs *J. Am. Water Resour. Assoc.* **42** 1261–74

Culligan P, Carson T, Gaffin S, Gibson R, Hakimdavar R, Hsueh D, Hunter N, Marasco D, McGillis W and O'Connor T P 2014 *Evaluation of Green Roof Water Quantity and Quality Performance in an Urban Climate EPA-AE-83484601-0* (EPA/600/R-14/180)

Fassman-Beck E, Voyde E, Simcock R and Hong Y S 2013 4 Living roofs in 3 locations: does configuration affect runoff mitigation? *J. Hydrol.* **490** 11–20

Hakimdavar R, Culligan P J, Finazzi M, Barontini S and Ranzi R 2014 Scale dynamics of extensive green roofs: quantifying the effect of drainage area and rainfall characteristics on observed and modeled green roof hydrologic performance *Ecol. Eng.* **73** 494–508

Hargreaves G L, Hargreaves G H and Riley J P 1985 Agricultural benefits for senegal river basin *J. Irrig. Drain. Eng.* **111** 113–24

Hilten R N, Lawrence T M and Tollner E W 2008 Modeling stormwater runoff from green roofs with HYDRUS-1D *J. Hydrol.* **358** 288–93

Kaufmann P 1999 *Extensiv begrünte Flachdächer - ein Gewinn für die Siedlungsentwässerung* (Burgdorf: Hochschule fur Technik und Architektur)

Khader O and Montalto F A 2008 Development and calibration of a high resolution SWMM model for simulating the effects of LID retrofits on the outflow hydrograph of a dense urban watershed *Low Impact Development for Urban Ecosystem and Habitat Protection* (Reston, VA: ASCE) pp 1–9

Koza J R 1992 *Genetic Programming: On the Programming of Computers by Means of Natural Selection* (Cambridge, MA: MIT Press)

Liesecke H 2002 Weiterentwicklung der einschichtigen Bauweise fur extensive Dachbegrunungen, Teil: III. Wasserrückhaltung und Jahr Abflussbeiwert *Dach+Grün* **11** 10–3

Lundholm J, MacIvor J S, MacDougall Z and Ranalli M 2010 Plant species and functional group combinations affect green roof ecosystem functions *PLoS One* **5** e9677

Marasco D E, Hunter B N, Culligan P J, Gaffin S R and McGillis W R 2014 Quantifying evapotranspiration from urban green roofs: a comparison of chamber measurements with commonly used predictive methods *Environ. Sci. Technol.* **48** 10273–81

Mentens J, Raes D and Hermy M 2006 Green roofs as a tool for solving the rainwater runoff problem in the urbanized 21st century? *Landsc. Urban Plan.* **77** 217–26

Moriasi D N, Arnold J G, Van Liew M W, Bringner R L, Harmel R D and Veith T L 2007 Model evaluation guidelines

for systematic quantification of accuracy in watershed simulations *Trans. ASABE* **50** 885–900

Nash J E and Sutcliffe J V 1970 River flow forecasting through conceptual models: I. A discussion of principles *J. Hydrol.* **10** 282–90

NYC Department of Buildings 2008 1 RCNY §105-01, Green Roof Property Tax Abatement 2–6 (http://nyc.gov/html/dob/downloads/rules/1_RCNY_105-01.pdf)

Olivieri F, Di Perna C, D'Orazio M, Olivieri L and Neila J 2013 Experimental measurements and numerical model for the summer performance assessment of extensive green roofs in a Mediterranean coastal climate *Energy Build.* **63** 1–14

Poë S, Stovin V and Berretta C 2015 Parameters influencing the regeneration of a green roof's retention capacity via evapotranspiration *J. Hydrol.* **523** 356–67

Schmidt M and Lipson H 2009 Distilling free-form natural laws *Science* **324** 81–5

Schroll E, Lambrinos J, Righetti T and Sandrock D 2011 The role of vegetation in regulating stormwater runoff from green roofs in a winter rainfall climate *Ecol. Eng.* **37** 595–600

Sherrard J A Jr and Jacobs J M 2012 Vegetated roof water-balance model : experimental and model results *J. Hydrol. Eng.* **17** 858–68

Stovin V, Vesuviano G and Kasmin H 2012 The hydrological performance of a green roof test bed under UK climatic conditions *J. Hydrol.* **414-415** 148–61

Technology Acceptance and Reciprocity Partnership 2001 The Technology Acceptance Reciprocity Partnership Protocol for Stormwater Best Management Practice Demonstrations (http://mass.gov/dep/water/laws/swprotoc.pdf)

Uhl M and Schiedt L 2008 Green roof storm water retention—monitoring results *11th Int. Conf. on Urban Drainage (Edinburgh, Scotland)* pp 1–10

US EPA 2004 *Report to Congress: Impacts and Control of CSOs and SSOs* (http://cfpub.epa.gov/npdes/cso/cpolicy_report2004.cfm)

Vanuytrecht E, Mechelen C V, Meerbeek K V, Willems P, Hermy M and Raes D 2014 Runoff and vegetation stress of green roofs under different climate change scenarios *Landsc. Urban Plan* **122** 68–77

Voyde E, Fassman E and Simcock R 2010 Hydrology of an extensive living roof under sub-tropical climate conditions in Auckland, New Zealand *J. Hydrol.* **394** 384–95

Voyde E A 2011 Quantifying the complete hydrologic budget for an extensive living roof *Doctoral Dissertation* The University of Auckland

Washington State Department of Ecology 2008 *Guidance for Evaluating Emerging Stormwater Treatment Technologies: Technology Assessment Protocol*

Wong G K L and Jim C Y 2014 Quantitative hydrologic performance of extensive green roof under humid-tropical rainfall regime *Ecol. Eng.* **70** 366–78

Zhang S and Guo Y 2013 Analytical probabilistic model for evaluating the hydrologic performance of green roofs *J. Hydrol. Eng.* **18** 19–28