Evaluating a Sewershed Urban Storm Water Model for Variability in Parameter Sensitivity and Resolution Effects

Zhaokai Dong,a Daniel Bain,b Murat Akcakayac and Carla Ng \*a,d

a. Department of Civil and Environmental Engineering, University of Pittsburgh, Pittsburgh, PA 15261 USA.

b. Department of Geology and Environmental Science, University of Pittsburgh, Pittsburgh, PA 15261 USA.

c. Department of Electrical and Computer Engineering, University of Pittsburgh, Pittsburgh, PA 15261 USA.

d. Department of Environmental and Occupational Health, University of Pittsburgh, Pittsburgh, PA 15261, USA

Abstract

A high-quality paramer set is essential for reliable stormwater models. Model performance can be improved by optimizing initial parameter estimates. Parameter sensitivity analysis is a robust way to distinguish the influence of parameters on model output and efficiently target the most important parameters to modify. This study evaluates efficient construction of a sewershed model using relatively low-resolution (e.g., 30 meter DEM) data and explores model sensitivity to parameters and regional characteristics using the EPA’s Storm Water Management Model (SWMM). A SWMM model was developed for a sewershed in the City of Pittsburgh, where stormwater management is a critical concern. We assumed uniform or log-normal distributions for parameters and used Monte Carlo simulations to explore and rank the influence of parameters on predicted surface runoff, peak flow, maximum pipe flow and model performance, as measured using the Nash–Sutcliffe efficiency metric. By using the Thiessen polygon approach for sub-catchment delineations, we substantially simplified the parameterization of the areas and hydraulic paramters. Despite this simplification, our approach provided good agreement with monitored pipe flow (Nash–Sutcliffe efficiency: 0.41 – 0.85). Total runoff and peak flow were very sensitive to the model discretization. The size of the polygons (modeled subcatchment areas) and imperviousness had the most influence on both outputs. The imperviousness, infiltration and Manning’s roughness (in the pervious area) contribtuted strongly to the Nash-Sutcliffe efficiency (70%), as did pipe geometric parameters (92%). Parameter rank sets were compared by using kappa statistics between any two model elements to identify generalities. Within our relatively large (9.7 km^2) sewershed, optimizing parameters for the highly impervious (>50%) areas and larger pipes lower in the network contributed most to improving Nash–Sutcliffe efficiency. The geometric parameters influence the water quantity distrubtion and flow conveyance, while imperviousness determines the subcatchment subdivision and influences surface water generation. Application of the Thiessen polygon approach can simplify the construction of large-scale urban storm water models, but the model is senstive to the sewer network configuration and care must be taken in parameterizing areas (polygons) with heterogenous land uses.

Introduction

Urbanization dramatically changes the landscape, particularly by increasing impervious surfaces.1 Higher impervious cover reduces infiltration, increases the velocity and volume of stormwater runoff arriving into steams, and accelerates the transport of nutrients and contaminants.1–4 Further, urban stormwater is one of the leading sources of water pollution in the United States, especially in urban areas.5 Studies evaluating stormwater impact mitigation have pursued innovative stormwater management methods , in particular green infrastructures (GI) which aims to protect, mimic and maintain natural hydrologic conditions by using decentralized retrofit practices.6,7

To evaluate the effectiveness of urban water problem mitigation methods a number of stormwater runoff modelling tools8 have been developed. Among these, the U.S. Environmental Protection Agency (EPA) Storm Water Management Model (SWMM) has been widely applied for urban rainfall-runoff simulations.8,9 SWMM can simulate stormwater quantity and some aspects of water quality (e.g. total phosphorus) both on the land surface and in the sewers, and can also provide the capability to evaluate the effectiveness of GI simulations.10 In recent years, SWMM has been applied to study hydrologic impacts of urbanization11, to model the effects of green and grey (built) infrastructure practices to reduce combined sewer overflows (CSO)12, to investigate the impacts of GI on hydrological responses in urban areas13,14 and to simulate chemical transport15.

Despite these practical applications, improved efficiency in building model networks and enhanced model performance remains critical need, particularly for larger systems. Model output is influenced by many uncertainties, such as data resolution, parameter estimation, model structure, etc.16 SWMM hydrologic and hydraulic components are separately conceptualized as land surface and drainage networks, requiring substantial effort to collect and prepare varied input data, build reliable subcatchment and conduit networks, and estimate and calibrate model parameters. To capture physical processes in the urban system, however, detailed information on sewer drainage networks is often unavailable.17 In the absence of such data, it can be time-intensive to simplify and build simulated networks.18

In addition, discretizing a study area into subcatchments requires the knowledge and consideration of both the surface properties and the sewer connections. High-resolution data (e.g. aerial images) have been used to identify and describe detailed surface information (land use) in SWMM simulations,19–21 such as buildings and roads, which can help to build a model that is closest to reflecting reality. However, delineation of subcatchments based on high-resolution data is time-consuming for larger scale simulations (beyond a few city blocks) and high-resolution data for some area are limited. This motivates the need for more efficient methods using automatic spatial data processing methods for sub-catchment partitioning, such as digital elevation model (DEM) based delineation22 and pipe network-guided division23. However, these strategies introduce additional uncertainties to the flow routing. Unlike in “natural” watersheds, surface flows in urban areas are usually guided by roads or along “built” edges and then are received by pipe networks. Therefore, the feasibility and reliability of these easily-applied approaches for flow routing between subcatchments and in-pipe conveyance needs to be evaluated in urban areas.

Since it is not possible to directly obtain exact values of all parameters for SWMM, common practice is to calibrate and adjust poorly defined parameters. The importance of SWMM parameters on surface water generation and pipe-flow conveyance predictions is established.21,24,25 However, parameters included in model sensitivity analysis can vary for different catchments, when different physical properties of the specific sewershed or pipe network are considered. In addition, the methods and metrics (such as Spearman correlation coefficient, mutual entropy based on Monte Carlo simulation24, and one-at-a-time analysis25) used to indicate parameter impacts can result in different (or conflicting) assessments of parameter influence. Furthermore, parameter sensitivity analyses have mostly focused on a single model performance indicator, such as peak flow or total runoff, which is very useful for single event simulations. The contribution of parameters to the comprehensive model performance, considering both hydrology and hydraulics, has not been investigated comprehensively, especially when considering simulations roughly a month long.

This study is intended to test the ability to efficiently discretize model configurations by drawing Thiessen polygons26 and statistically explore the impacts of the resulting parameterization of model elements on outputs. We built a sewershed model using relatively low-resolution data (>30 meter pixels) in the City of Pittsburgh, which is highly vulnerable to flash floods and combined sewer overflows27. Heterogeneous land cover is divided into smaller, more homogeneous subcatchments by drawing Thiessen polygons. Model performance assessment considered both single outputs and comprehensive indicators. Parameter distributions were constructed based on the frequencies of initial estimates for each parameter across all the model elements (subcatchments or pipe segments). Further, Monte Carlo simulations were used to explore parameter influences and the most important parameters for global model performance. This allows evaluation of the magnitude of a particular parameter’s influence on the magnitude of the model response when a model contains highly variable subcatchments and pipe networks. The work described here consists of four phases: 1) SWMM model discretization and parameterization; 2) parameter sensitivity via Monte Carlo analysis; 3) ranking parameter importance, variability and consistency across model elements; 4) parameter contributions to improve model performance.

Methods

Model setup

The model system was prepared for sewershed M29 in the city of Pittsburgh (Fig. 1) with a drainage area of approximately . Landcover across the sewershed varies from highly urbanized areas to parkland. In addition, the terrain is heterogeneous with numerous steep slopes. These arrangements provide the opportunity to statistically investigate the relationship among various surface physical properties, the sewer networks, and model responses.



Fig. 1 Location and pipe network of sewershed M29 in the city of Pittsburgh. The inset shows the location of sewersheds in Allegheny County28 (the state of Pennsylvania29 in the U.S.); hillshade was derived from 1 m LiDAR DEM30

Table 1 Data sources

|  |  |  |  |
| --- | --- | --- | --- |
| Data | Resolution | Sources | Year |
| Landcover | 30 m×30 m | United States Geological Survey (USGS) | 2011 |
| DEM | 30 m×30 m | 1 Arc-second Shuttle Radar Topography Mission (SRTM) | 2014 |
| Sewer data | N.A. | Allegheny County Sanitary Authority (ALCOSAN) | 2020 |

The model was built using multiple types of spatial data (Table 1). The main input data included: (1) 15-minute rainfall intensity data31 obtained from the University of Pittsburgh rain gauge from March to July in 2008; (2) sewer data digitized on the basis of a system-wide sewer map32 and then corrected and validated by comparing with GIS data provided by the Allegheny County Sanitary Authority (ALCOSAN); (3) 30 m×30 m land cover data33 used to identify land cover types; (4) 30 m×30 m digital elevation model (DEM)34 used to derive slopes (in order to match the resolution of land cover). While the focus of this study was on model sensitivity analysis rather than full model calibration, the model results were compared with monitored pipe flow. Due to the computational cost of the Monte Carlo approach, the simulation was limited to one month (April 2008, provided by ALCOSAN); moreover, this month includes both dry and wet weather conditions.

SWMM parameterization

SWMM can be used for both single-event and long-term rainfall-runoff simulations in urban areas.10 The two main components of SWMM – a hydrologic model and a hydraulic model – receive precipitation and generate runoff; the model then routes this runoff as flow through drainage systems. SWMM surface runoff is mainly governed by precipitation, infiltration and evaporation, modelled as a nonlinear reservoir to derive the governing equation35 eq. (1).

(1)

Where is the change in depth *d* over time *t*; *i* is the rainfall or snowfall rate (Lt-1); *e* is the evaporation rate (Lt-1); *f* is the infiltration rate (Lt-1); *k* is the conversion constant; *W* is the width of subcatchment (L); *A* is the surface area of subcatchment (L2); *n* is the Manning’s roughness coefficient (L-1/3t); *S* is the slope; *ds* is the depression storage (L).

The pipe flow in SWMM is controlled by the St. Venant conservation of mass and momentum equations. Pipe flow36 then can then be derived as eq. (2).

(2)

Where is the change of flow rate per unit of time (L3t-2); *v* is the velocity (Lt-1); *A* is the flow cross-section area (L2); *g* is the gravitational acceleration (Lt-2); *H* is the hydraulic head (L); *Q* is the flow rate (L3t-1); *n* is the Manning’s roughness coefficient (L-1/3t); *R* is the hydraulic radius (L); *k* is the conversion constant.

In this study, we used Thiessen polygons to partition the subcatchments in ArcMap37. This method is widely used in geography to analyze precipitation data from unevenly distributed rain gauges, by finding points within a plane that are closest to the target point.38 In sewershed (M29), the sewer network is quite dense and the pipe inlets (stormwater drains, described as nodes hereafter) are distributed intensively throughout the developed areas but sparsely in the green areas. To simplify the model, we merged the adjacent nodes within 5 meters into one single node. Pipe segments were then merged as well. The simplified sewer network consists of 638 nodes and 636 pipes. The subcatchments were delineated based on sewer data by drawing Thiessen polygons, each plane enclosing only one node. This results in a large number of polygons for the developed areas, which increases homogeneity within each polygon. This method ignores surface topography and assumes that flow is not exchanged between subcatchments (that is, the surface water was only received by the pipe networks). This allowed us to assign flow paths from the surface to sewers, and resulted in 638 subcatchments.

The width of each subcatchment was estimated by dividing the subcatchment area by the longest distance from any point in the subcatchment (polygon) to the node. To estimate imperviousness, land types were extracted from land use and land cover data, followed by the assignment of typical values of imperviousness35,39 for each land type. Then, a weighted value was determined using eq. (3) to calculate the imperviousness, depression storage and Manning’s roughness. Evapotranspiration was not included in this initial model because data were unavailable for the relevant time period.

Target parameter = (3)

In this general equation for weighted target parameters, *i* is the type of land use within a subcatchment, *Pi* is the parameter value for the specific landscape; *Ai* is the corresponding area; and *Atotal*is the total area of the subcatchment.

Slope maps were derived from DEM and the average slope was assigned to each subcatchment. Infiltration-related parameters were based on Horton’s method35,40 as built into SWMM. The parameters that we cannot directly measure or extract from spatial data we estimated based on literature values (e.g. Manning’s roughness41,42 and depression storage41).

Based on the sewer data from ALCOSAN, the pipe diameters have a range from 0.8 ft up to 14 ft. Simplifying sewer network by combining nodes and pipes would result in a larger range of diameters. We assumed circular cross-sections and a constant pipe diameter (8 ft) initially for the 636 pipes and various diameters could be drawn from the subsequent sampling process to change the settings, described in next section. Node depth was assumed 11 ft for the 638 nodes.

Sensitivity analysis

Simplified sewer network and Thiessen polygons introduced uncertainties in determining suitable geometric properties and model element connections, which inevitably influenced the model outputs. Therefore, these physical attributes were considered as variables in the sensitivity analysis. Monte Carlo-based (MC) sensitivity analysis43 was conducted to identify the influence of model inputs on predictions of peak flow, runoff volume and maximum pipe flow. To create the distributions for each parameter, two strategies were used: 1) Parameters extracted and estimated from spatial data and those estimated with different values in subcatchments were assumed as lognormal distributions (based on the parameter frequencies), also preventing negative values when pulling values for MC sampling; 2) Parameters that were constant across subcatchments were assumed to be uniform distributions. The hyperparameter (standard deviation) of the lognormal distributions were then calculated from parameter data sets including initial estimates of parameters from all subcatchments. The initial estimation was used as the mean for each model element. This tested the ability of our method to parameterize the model. The upper and lower limits of the uniform distributions were assumed by considering typical values from literature35,40–42, whereas for the lognormal distributions they were based on different data for each subcatchment. In all, 638 different distributions were generated for each parameter (one for each subcatchment). Then, 1000 Monte Carlo simulations were run using Latin Hypercube Sampling44 (LHS) to evaluate the relationship between the model parameters and predicted outputs. This allowed us to compare the parameter sensitivity in each subcatchment and pipe line and statistically analyze the overall influences on model performance.

To quantify the influence of model parameters on model response, we calculated the Spearman correlation coefficient45 between model outputs — peak flow, total runoff and maximum pipe flow — and each set of parameters generated from the Monte Carlo sampling for each subcatchment or pipe. The distributions used for the main input parameters related to surface runoff and pipe flow are listed in Table 2.

Table 2 Main parameters in SWMM

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | | Distribution | |
| Subcatchment | Area\* | Lognormal | (Area, 0.33) |
| Width\* | Lognormal | (Width, 0.33) |
| Imperviousness\*(Imper) | Lognormal | (Imper, 0.24) |
| Slope\* | Lognormal | (Slope, 0.51) |
| Impervious N\*(N-imp) | Lognormal | (N-imp, 0.559) |
| Pervious N\*(N-perv) | Lognormal | (N-perv, 0.0867) |
| Impervious D\*(D-imper) | Uniform | (0.05, 0.1) |
| Pervious D\*(D-perv) | Lognormal | (D-perv, 0.369) |
| Infiltration (Horton) | Maximum infiltration rate\*(MaxIR) | Uniform | (4, 8) |
| Minimum infiltration rate\*(MinIR) | Uniform | (0.1, 0.3) |
| Decay rate\*(Decay) | Uniform | (2,7) |
| Sewer system | Junction elevation | N.A. | |
| Pipe length\* | Lognormal | (Length, 0.49) |
| Roughness\* | Uniform | (0.011, 0.015) |
| Shape | N.A. | |
| Diameter\* | Lognormal | (8, 0.47) |

Note: parameters with \* are included in the sensitivity analysis; (a, b) for lognormal distribution: a – mean of lognormal distribution (the initial estimation for each model element), b – standard deviation of corresponding transformed normal distribution; (a, b) for uniform distribution: a – lower bound, b – upper bound. As parameter values vary by subcatchment, the mean of the distribution is not reported here, rather the variable name is indicated.

Statistical analysis of parameter influence on global model outputs may overlook detailed parameter ranking information within each model element. If the importance of parameters varies widely among subcatchments (and corresponding pipes), a single set of calibration parameters may result in poor performance for a large model with many subcatchments, such as the model considered here. Therefore, Cohen’s kappa statistic46 was used to investigate the agreement of parameter importance ranks between any two subcatchments or pipes eq. (4).

(4)

= (5)

= (6)

 (eq. (5)) is the observed agreement between parameter ranks.  is the total number of paired observations and is the observed number of parameters with the same order.

 is the hypothetical probability of chance agreement (expected agreement) where *N* is the total number of parameters. In our case, each parameter only has one unique position to be placed per subcatchment. Assuming the parameter ranks were independent among subcatchments, each parameter has a probability of to be assigned to each position. Therefore, the total probability of two parameters sharing the rank position, also termed as , is , which is simplified as (eq. (6)).

We did not include a traditional model calibration here, but we calculated the Nash–Sutcliffe efficiency47 indicating fitness of model outputs to the observation (eq. (7)) after each simulation by comparing modelling results to the monitored pipe flow data. A Nash-Sutcliffe efficiency closer to 1 suggests a better simulation was achieved.

(7)

Here, and are the observed and simulated flow discharge values, respectively, and is the observed mean flow.

The Nash–Sutcliffe efficiency was also included in our sensitivity analyses, in addition to the single indicators mentioned above. Because we varied model parameters simultaneously, we expected to see a relatively wide range of values of Nash–Sutcliffe efficiency. We chose the 5% best and 5% worst simulations (highest efficiency and lowest efficiency, respectively) to further evaluate the parameter settings that most improve model performance.

We calculated the difference (eq. (8)) in parameter values between the best and worst simulations (5%) for each model element (i.e., each subcatchment or pipe). The differences were then normalized by dividing the mean of the parameters to eliminate the influence of different units. The differences in eq. (8) were calculated by both type of parameters and each specific subcatchment or pipe.

= (8)

We then calculated the sum of the absolute differences as the parameter contributions when improving Nash–Sutcliffe efficiency from the worst to best (eq. (9) – eq. (10)). We also normalized the contribution of each subcatchment by dividing by its area, so we could eliminate the influence of the shape of the area and compare it with the contribution per area (eq. (11)).

= (9)

= (10)

= (11)

In these equations, *p* indicates the parameters whereas *c* indicates model elements (i.e., subcatchment or pipe)

Results and discussion

Model performance

The simulation results were compared with observed pipe flow in April 2008. The model performed well based on both Nash–Sutcliffe efficiency metrics and visual comparisons (Fig. 2). The Nash–Sutcliffe efficiency was calculated for 1000 simulations (Fig. 2 (A)). Although we did not implement any model calibration, simulations present good fits of observed pipe flow, with Nash–Sutcliffe efficiency that range from 0.41 to 0.85 for the 1000 simulations. Most of the simulations yield a good number and the most frequent efficiency is 0.72. The monitored flows were within the 95% confidence interval (CI) of model simulations and were close (RMSE of 1.47) to the geometric means derived from them (Fig 2B). For example, results for late April indicate that despite high variability suggested by the 95% CI, our model captures the flow patterns around multiple rain events (Fig. 2C – D). The shapes and magnitudes of flows are well predicted by the geometric mean and the time to peak also corresponds well with most observed data. However, it should be noted that we did not consider evapotranspiration in the current model, which could lead to an over-estimation for water generation. This may be one reason that some flow peaks of the geometric mean were slightly higher than the monitored flow.

Based on this strength of agreement, the Thiessen polygon method was demonstrated appropriately for partitioning our subcatchments. This approach has a major advantage in that it requires relatively minimal effort, especially for larger sewersheds, by making the simplifying assumption of no flow across subcatchments. Our results suggest this strategy is feasible for long-term simulations, in addition to previous use in single event simulatons48. This method assigns the surface water to a sewer inlet by the shortest flow path. It works well in our model, because the simplification of sewer system network nodes was based on 5 meters, preserving most original nodes as possible (94% of simplified nodes were at original locations). This generated more subcatchments, but increased subcatchment homogeneity and reduced potential errors from uncertainties in flow direction (to pipe inlets). However, when re-routing surface flow or having a simplified sewer network with sparse nodes, the model may not work as well as a model developed using high resolution data. Even though the low resolution model had the capability to predict total runoff as high resolution model, the peak flow was affected by model resolution.20 Based on this model performance, the assumed parameter distributions were also reasonable; we next investigated how they reflected the sensitivity of model predictions to the changes in key parameter values.



Fig. 2 Model performance considering 1000 Monte Carlo simulations in comparison to observed pipe flow: A – histogram of Nash-Sutcliffe efficiency based on 1000 simulation B –simulation results compared with observed pipe flow for the month of April, 2008; C – detailed simulation, April 19th to 21st ; D – detailed simulation, April 27th to 29th.

Model parameter sensitivity analyses

Based on 1000 Monte Carlo simulations, Spearman correlation coefficients were calculated for three different model outputs: total runoff, peak flow and maximum pipe flow (Fig. 3). The size of each polygon, that is, the subcatchment area, had the highest correlation coefficient and almost dominated the influence of parameters on total runoff (Fig.3A). SWMM treats subcatchments as reservoirs; the area is therefore the parameter that determines the capacity of the polygons. As the polygons shrunk or enlarged, the entire model configuration was modified and the water generation within each subcatchment was subsequently influenced. As we ignored the surface water exchange, the area was highly important to adjust water quantity distributions across the model system. In our method, each subcatchment includes one node. The water flow schemes can also be reshaped by adjusting the number of polygons, which is determined by the number of nodes (also represented by the radius for node merger). This also suggests the model is sensitive to the simplifications of sewer network. The second most important parameter is the imperviousness, which also has a large positive impact relative to others. The imperviousness plays a role for subcatchment subdivision (impervious and pervious area) and determines the amount of area in subcatchment that directly yield surface flow during rainfall events.



Fig. 3 Spearman correlation coefficients in subcatchments: A – runoff; B – peak flow; C – maximum pipe flow

The remaining parameters all have boxplots (Fig. 3 (A)) with a median of zero that cover positive and negative signs in the range of the Spearman coefficient, indicating that these parameters have no consistent decreasing or increasing effects on total runoff.

The influence of parameters on the predictions of peak flow is much clearer in both their magnitudes and signs (Fig. 3B). The size of polygon (area) is still highly important as it is to controlling total runoff, because the total runoff is derived from the integration of flow. The next most important parameter is debatable, since both imperviousness and Manning’s roughness (in the impervious area) have similar magnitudes, but different signs. While imperviousness and Manning’s roughness may covary in reality, parameters were randomly drawn from independent assumed distributions. The resulting coefficients did not suggest a relationship between these two parameters. The imperviousness determines the percent of area to generate water directly, which leads to the positive sign. Manning’s roughness has a negative influence because this parameter is used to represent the surface friction and will delay surface flow. Width and slope also have comparable influence to one another; for example, steeper slope would speed up the flow rate and the wider width increases the cross-section when the ponded depth (eq. (1)) is solved over a time step, leading to increased flow. The remaining parameters have similar ranges of low values and cross the zero line. Our findings of parameter influence on the predicted surface are similar to previous work that identified the importance of imperviousness49 and area24.

For the pipe parameters, considering the complexity of our model, we only included three parameters, two geometric parameters (pipe diameter and length) and the Manning’s roughness. We simplified sewer networks, so the geometric parameters were not the actual pipe attributes anymore and were similar to the Manning’s roughness to be determined. All Spearman correlation coefficients for these parameters have a zero median, but the pipe diameter and length show substantial fat tails that span a wide range of influence, indicating that the geometric parameters can have higher influence on the routing of flow (Fig. 3C). These results are consistent with the findings of influence of conduit geometry on karst systems by Peterson and Wicks (2006), in which the majority of flow travelled through conduits and were then discharged by an outlet spring.25 In that study, the flow routing in anisotropic aquifer systems was closely linked to the geometry of the conduits, and minor changes in the geometry led to significant changes in the flow.25 We considered the influence of geometry in our sensitivity analysis, so this may have obscured the influence of the pipe Manning’s roughness.24

Parameter importance ranking analysis

In evaluating the Spearman correlation coefficients, a number of parameters had similar low influence on the overall outputs (zero median and within a range between -0.2 and 0.2), making it difficult to compare them. Therefore, after evaluating the overall influence of parameters on these three outputs, we next explored the parameter importance ranking within individual model elements (that is, specific subcatchments and pipes). We ranked parameters by their coefficients for each element to assess the agreement in parameter rankings. The kappa coefficient was calculated by comparing parameter ranks based on the order of the absolute Spearman correlation coefficients for all pairs of subcatchments and pipes. Kappa is scaled from -1 to 1, where values closer to 1 indicate a more perfect match in ranks, whereas negative values mean the agreement is less than chance50. For both total surface runoff and peak flow, the top five parameter ranks are more consistent than when all parameter ranks are considered (Fig. 4). The consistency of parameter ranks for peak flow has more uncertainty compared to total runoff. However, the median of the peak flow ranking was higher than the runoff ranking which means it is more probable for the subcatchments having peak flows to share similar parameter importance ranks. The third quartile of the top five ranks for peak flow reaches a kappa coefficient of 1.0 which means that many of the subcatchments share parameter ranks.



Fig. 4 Kappa coefficients based on comparison of all parameters and top five most important parameters for total runoff, peak flow, and maximum pipe flow

Table 3 Main parameter rank sets and the probability of each set of top five parameter ranks for peak flow

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Probability | No. 1 | No. 2 | No. 3 | No. 4 | No. 5 |
| Set1 | 54% | area | imper | N-imper | width | slope |
| Set2 | 17% | area | N-imper | imper | width | slope |
| Set3 | 16% | area | imper | N-imper | slope | width |
| Set4 | 7% | area | N-imper | imper | slope | width |

We can assemble the parameter ranks of peak flow into a matrix (638 × 11). The rows represented the specific parameter rank sets for subcatchments and columns represented the ranked parameters. We separately calculated the probabilities based on the rank sets of subcatchments (Table 3) and ordered parameter at each position (Fig. 6), representing by its occurred proportion to all subcatchments. Four sets of parameter ranks have 94% probability to explain their rank orders (Table 3). The rank set 1 accounts for 54% of the total subcatchments, much higher than other sets. We can also evaluate the distribution of the parameter rank sets across all 638 subcatchments (Fig. 5). For all four sets, the area is always the most important parameter, but the imperviousness and Manning’s roughness (in the impervious area) switch at the next position. If we combined set 1 and set 3 where imperviousness has the same order, the subcatchments are mostly located far from the center of our sewershed. These areas are highly impervious, so they are more sensitive to the change of percent of imperviousness, which directly determines the area that generates water flow immediately during a rainfall event.

We further focused on the mean values of the associated parameters and found that when rank switching happens the parameter with the higher mean value will show up at the forward position. For example, the mean of imperviousness is larger in the subcatchments in rank set 1 than in set 2, so the imperviousness is placed at the second place in set 2. Thus, a parameter with a higher estimated value would be more important. In other words, the surface properties influence how the physical parameters influence the water generation in the model. For example, the highly pervious areas (set 2 and set 4) were not as sensitive to imperviousness, since the green surfaces are large enough that increasing a little impervious area would not dramatically change the conditions of water reception and infiltration.



Fig. 5 Distribution of top five parameter sets across all the subcatchments (parameter sets shown in Table 3).

The proportion of each parameter occurred at each rank position clearly shows that area is always the first one (Fig. 6), corresponding to the first column in Table 3. Imperviousness has the highest probability (76%) to be ranked as the second most important parameter. As discussed above, the entire sewershed has a large portion of highly impervious area, so the imperviousness was reasonably ranked at the second position with high probability. In general, all the dominant parameters have probabilities larger than 70% in their respective rank positions. When ranking parameters by highest probability at each position, the order is area, imperviousness, imperviousness Manning’s roughness, width, and slope. This is the same rank order as the parameter rank set 1.



Fig. 6 Probability of each parameter in the top five rank positions.

Parameter contribution to model performance

In the Nash-Sutcliffe efficiency, the model outputs were nested together into one expression that presented the fitness of the simulations over each time step relative to the corresponding measurements. This differs from evaluating the influence of parameters on a single output index (e.g. peak flow). We therefore explored improvements in model performance (Nash-Sutcliffe efficiency) by observing the change of parameter values. As the model became more accurate with higher levels of the Nash-Sutcliffe efficiency, the corresponding parameters selected across the LHS based Monte Carlo sample should improve from worse to better, with parameters that underwent the largest change from worst to better having the greatest influence. We used this as an alternative means of evaluating the contribution of parameters to model performance (Fig. 7). These global parameter contributions show similarity to single model output sensitivities. The imperviousness still had substantial contribution, but the area impacts were not as obvious as in the sensitivity analysis. The parameters representing surface properties also have evident influence. Infiltration, impervious depression storage, and Manning’s roughness (in pervious area) together reach nearly 65% of the total contribution. The influence of pipe length and diameter had a contribution as high as 92%, highlighting again the importance of the geometric parameters.



Fig. 7 Parameter contribution to achieve better model performance for subcatchments and pipes.

In addition, we investigated the contribution of each model element (Fig. 8) to model performance. The model elements in the central area in Fig. 8 contribute minimally to performance. This is Schenley Park, park land with low imperviousness. The area surrounding this circle are highly impervious with low Manning’s roughness. The distribution of model element contributions can be linked to the influence of parameters on peak flow. We’ve eliminated the influence of area (eq. (11)), so the water generation is highly related to imperviousness and Manning’s roughness. The peak flow was related negatively to the Manning’s roughness and positively to imperviousness, so the surrounding area would generate relatively high flow rates and the change of parameters on the area contribution more to model performance. Under this condition, the highly impervious area would generate more surface water and have more contribution than the central area. The pipes that received water from the subcatchments with high contribution contribute a lot to model performance, as do corresponding downstream pipes. We assumed all the conduits had the same diameter initially, but the pipe diameters vary based on water conveyance capacity. The diameter from the best simulation ranged from 2.15 ft to 39 ft. The actual upstream pipe diameters are mostly less than 3 ft based on available data. After simplification, even those pipes can have larger diameters due to merging multiple pipes; 8 ft is still a large initial setting for upstream pipes. The pipe segments close to the end-node reach a diameter as large as 14 ft based on the sewer data, but in our system, the diameter could also be much larger due to the merged nodes and reduced pipe segments. For those downstream pipes assuming uniform pipe diameter is likely not appropriate given increased contributions from the upstream sewer network. When initial settings were adjusted to improve model performance, downstream pipes had larger cross-section to allow more water to flow in and out and maintain the stability of the simulation.



Fig. 8 Contribution of model elements (subcatchments or pipes) to Nash-Sutcliffe performance.

Conclusions

In this study we built a sewershed urban stormwater model using low-resolution data. We constructed the model subcatchment configuration by the Thiessen polygon method, assuming no re-routed flow between subcatchments, and a simplified sewer network generated by merging nodes and pipe segments. Despite no traditional calibration procedure, our model, based on a Monte Carlo sample of 1000 simulations, predicted flow patterns well over the month of April 2008. The Nash-Sutcliffe efficiency, ranged from 0.41 to 0.85, with a frequency mode of 0.72 (Fig. 2A). This indicates that the methods we applied to build this sewershed model can work well even with relatively low-resolution data for a multi-event month-long simulation.

We identified the influence of SWMM parameters on three model outputs: surface runoff volume, peak flow and maximum pipe flow. In our changeable model configurations, total runoff predicted by SWMM is highly sensitive to the size of polygons (subcatchment area) and imperviousness. In addition, Manning’s roughness (in the impervious area), width and slope play important roles in determining peak flow. The pipe maximum flow is most sensitive to the pipe geometry (diameter and length).

Parameter ranks for all parameters across all model elements vary and the agreement of parameter importance ranks are not very strong, especially with the consideration of total runoff. However, the first five most important parameters that control the peak flow are consistent. Other studies have previously sought to identify the influence of parameters on model outputs, but usually yielded different emphases24. Based on the consistency of our parameter rank set analysis and the probability for each parameter, it suggests that area, imperviousness, impervious Manning’s roughness, width and slope have the highest influence on peak flow simulation, in order of decreasing importance. These parameter ranks can provide modelers a priority list of parameters for calibrating models to predict peak flow.

The parameter contribution to model performance shows the similarity to the influence of parameters on the three model outputs. The area, Manning’s roughness (in the impervious area), imperviousness and width have 35% of the total parameter contribution. Meanwhile, the filtration parameters (maximum infiltration rate, minimum infiltration rate and decay rate) related to the Horton method and Manning’s roughness (in the pervious area) contribute as well, but we did not observe these influences based on the single parameter analysis (Fig. 3). This indicates the focus of consecutive simulations or some typical indicator would lead to different directions on the emphasis of parameter influence. However, the size of polygon (subcatchment area) and imperviousness were always identified, meaning the two parameters were essential for heterogenous subcatchments based on low resolution data partitioning (Thiessen polygon). In all, the parameter influence was related to the physical surface properties. The parameters in highly impervious area had the most influence on the model performance.

In all, the method we proposed to efficiently build sewershed model networks is sensitive to simplified sewer network and Thiessen polygons. The shape of the polygon and the pipe geometry are critical to achieve good model simulations, influencing flow assignment and conveyance. Polygons can be easily draw based on spatial data and key parameters can be efficiently handed in each polygon and pipe segment by using low resolution data. Even when minimal detail is provided on topography and surface cover, the uncertainties of flow direction and conveyance can be reduced by keeping more of the original nodes and introducing more homogeneous polygons. This suggests there are advantages to using the low-resolution data to build large-scale urban storm water models, for which these important parameters are easy to obtain.

We explored the parameter influence separately on the land surface and in the pipe networks. However, unlike individual subcatchments, the pipe network received water from the surface generation and flow was conveyed by the conduits. Therefore, the surface parameters influenced the pipe flow as well and the influence of the parameters entered into the pipe system as input variables and were transferred downstream. Unfortunately, considering the uncertainty of connections between the pipe network and surface subcatchments, we could not fully analyze this nested influence.

In addition, the results of parameter sensitivity analysis also need to be validated after model calibration to prove its effectiveness to improve model precision. All these issues should be included in future work so as to comprehensively understand parameter impacts on model performance and most efficient calibration strategies for larger-scale implementations of SWMM.

In this paper, we built a sewershed storm water model in SWMM, validated the feasibility of using the Thiessen polygon method to partition subcatchments with minimal effort and analyzed the variation of parameter influence on model outputs and performance for various model elements. With these insights, it is possible to determine optimal model configurations using low resolution data, understand parameter effects on SWMM output and target the most important parameters in model calibration that improve performance both in terms of simple (e.g. single-parameter, single-element, single-output) and more complex (e.g. global) indicators.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. (NSF grant number 1854827). We thank the Allegheny County Sanitary Authority for graciously providing sewer flow data and field data.

Notes and references

1 C. L. Arnold and C. J. Gibbons, Impervious Surface Coverage: The Emergence of a Key Environmental Indicator, *J. Am. Plann. Assoc.*, 1996, **62**, 243–258.

2 J. G. Lee, A. Selvakumar, K. Alvi, J. Riverson, J. X. Zhen, L. Shoemaker and F. Lai, A watershed-scale design optimization model for stormwater best management practices, *Environ. Model. Softw.*, 2012, **37**, 6–18.

3 T. R. Schueler, L. Fraley-McNeal and K. Cappiella, Is Impervious Cover Still Important? Review of Recent Research, *J. Hydrol. Eng.*, 2009, **14**, 309–315.

4 H. Li, L. J. Sharkey, W. F. Hunt and A. P. Davis, Mitigation of Impervious Surface Hydrology Using Bioretention in North Carolina and Maryland, *J. Hydrol. Eng.*, 2009, **14**, 407–415.

5 National Research Council, *Urban Stormwater Management in the United States*, National Research Council, Washington, DC: The National Academies Press, 2009.

6 L. S. Coffman, R. Goo and R. Frederick, in *Low-Impact Development: An Innovative Alternative Approach to Stormwater Management*, 1999, pp. 1–10.

7 L. M. Ahiablame, B. A. Engel and I. Chaubey, Effectiveness of Low Impact Development Practices: Literature Review and Suggestions for Future Research, *Water. Air. Soil Pollut.*, 2012, **223**, 4253–4273.

8 K. Eckart, Z. McPhee and T. Bolisetti, Performance and implementation of low impact development – A review, *Sci. Total Environ.*, 2017, **607–608**, 413–432.

9 L.-Y. Tsai, C.-F. Chen, C.-H. Fan and J.-Y. Lin, Using the HSPF and SWMM Models in a High Pervious Watershed and Estimating Their Parameter Sensitivity, *Water*, 2017, **9**, 780.

10 L. A. Rossman, *Storm Water Management Model User’s Manual Version 5.1*, US EPA Office of Research and Development, Washington, DC, EPA/600/R-14/413 (NTIS EPA/600/R-14/413b), 2015.

11 S. Jang, M. Cho, J. Yoon, Y. Yoon, S. Kim, G. Kim, L. Kim and H. Aksoy, Using SWMM as a tool for hydrologic impact assessment, *Desalination*, 2007, **212**, 344–356.

12 A. Alves, A. Sanchez, Z. Vojinovic, S. Seyoum, M. Babel and D. Brdjanovic, Evolutionary and Holistic Assessment of Green-Grey Infrastructure for CSO Reduction, *Water*, 2016, **8**, 402.

13 A. Palla and I. Gnecco, Hydrologic modeling of Low Impact Development systems at the urban catchment scale, *J. Hydrol.*, 2015, **528**, 361–368.

14 M. P. Abi Aad, M. T. Suidan and W. D. Shuster, Modeling Techniques of Best Management Practices: Rain Barrels and Rain Gardens Using EPA SWMM-5, *J. Hydrol. Eng.*, 2010, **15**, 434–443.

15 A.-K. McCall, R. Palmitessa, F. Blumensaat, E. Morgenroth and C. Ort, Modeling in-sewer transformations at catchment scale – implications on drug consumption estimates in wastewater-based epidemiology, *Water Res.*, 2017, **122**, 655–668.

16 J. Lei and W. Schilling, Parameter Uncertainty Propagation Analysis for Urban Rainfall Runoff Modelling, *Water Sci. Technol.*, 1994, **29**, 145–154.

17 G. Del Giudice and R. Padulano, Sensitivity Analysis and Calibration of a Rainfall-Runoff Model with the Combined Use of EPA-SWMM and Genetic Algorithm, *Acta Geophys.*, 2016, **64**, 1755–1778.

18 M. Zellner, D. Massey, E. Minor and M. Gonzalez-Meler, Exploring the effects of green infrastructure placement on neighborhood-level flooding via spatially explicit simulations, *Comput. Environ. Urban Syst.*, 2016, **59**, 116–128.

19 G. Krebs, T. Kokkonen, M. Valtanen, H. Koivusalo and H. Setälä, A high resolution application of a stormwater management model (SWMM) using genetic parameter optimization, *Urban Water J.*, 2013, **10**, 394–410.

20 G. Krebs, T. Kokkonen, M. Valtanen, H. Setälä and H. Koivusalo, Spatial resolution considerations for urban hydrological modelling, *J. Hydrol.*, 2014, **512**, 482–497.

21 N. Sun, B. Hong and M. Hall, Assessment of the SWMM model uncertainties within the generalized likelihood uncertainty estimation (GLUE) framework for a high-resolution urban sewershed, *Hydrol. Process.*, 2013, **28**, 3018–3034.

22 D. S. Bisht, C. Chatterjee, S. Kalakoti, P. Upadhyay, M. Sahoo and A. Panda, Modeling urban floods and drainage using SWMM and MIKE URBAN: a case study, *Nat. Hazards*, 2016, **84**, 749–776.

23 M. Huang and S. Jin, A methodology for simple 2-D inundation analysis in urban area using SWMM and GIS, *Nat. Hazards*, 2019, **97**, 15–43.

24 C. Li, W. Wang, J. Xiong and P. Chen, Sensitivity Analysis for Urban Drainage Modeling Using Mutual Information, *Entropy*, 2014, **16**, 5738–5752.

25 E. W. Peterson and C. M. Wicks, Assessing the importance of conduit geometry and physical parameters in karst systems using the storm water management model (SWMM), *J. Hydrol.*, 2006, **329**, 294–305.

26 Atsuyuki Okabe, Barry Boots, Kokichi Sugihara and Sung Nok Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, 2nd Edition*, Chichester, UK: John Wiley & Sons., 2000.

27 K. G. Hopkins, D. J. Bain and E. M. Copeland, Reconstruction of a century of landscape modification and hydrologic change in a small urban watershed in Pittsburgh, PA, *Landsc. Ecol.*, 2014, **29**, 413–424.

28 Allegheny County Division of Computer Services Geographic Information Systems Group, *Allegheny County Boundary*, 2016.

29 Pennsylvania Department of Transportation, Bureau of Planning and Research, Cartographic Information Division, *PennDOT - Pennsylvania County Boundaries*, Harrisburg, PA, 2020.

30 Allegheny County Division of Computer Services Geographic Information Systems Group, *Allegheny County Lidar and Terrain Products*, 2017.

31 3 Rivers Wet Weather (3RWW), Rain Gauge and Calibrated Radar Rainfall Data, https://www.3riverswetweather.org/municipalities/calibrated-radar-rainfall-data.

32 Allegheny County Sanitary Authority (ALCOSAN). (2020, March)., Conveyance and Treatment System with Regional Collection Syetem Outfalls, Retrieved from https://www.alcosan.org/docs/default-source/system-mapping/sanitary-sewer\_-combined-sewer\_outfalls\_36x24\_march2020.pdf

33 U.S. Geological Survey, *NLCD 2001 Land Cover (2011 Edition, amended 2014) - National Geospatial Data Asset (NGDA) Land Use Land Cover*, 2014.

34 U.S. Geological Survey, *USGS EROS Archive - Digital Elevation - Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global*, 2014.

35 L. A. Rossman and Wayne C. Huber, *Storm Water Management Model Reference Manual Volume I – Hydrology*, US EPA Office of Research and Development, Washington, DC, EPA/600/R-15/162A, 2016.

36 L. A. Rossman, *Storm Water Management Model Reference Manual Volume II – Hydraulics*, U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-17/111, 2017.

37 Esri Inc, *ArcMap (version 10.5.1)*, Redlands, CA: Esri Inc, 2016.

38 K. E. Brassel and D. Reif, A Procedure to Generate Thiessen Polygons, *Geogr. Anal.*, 2010, **11**, 289–303.

39 U.S. EPA, *Estimating Change in Impervious Area (IA) and Directly Connected Impervious Areas (DCIA) for Massachusetts Small MS4 Permit*, Boston, MA, 2014.

40 R. E. Horton, An Approach Toward a Physical Interpretation of Infiltration-Capacity, *Soil Sci. Soc. Am. J.*, 1941, **5**, 399–417.

41 American Society of Civil Engineers and Water Environment Federation, *Design and Construction of Urban Stormwater Management Systems*, Reston, VA, 2018.

42 P. Bizier, Ed., *Gravity sanitary sewer design and construction*, American Society of Civil Engineers: Environmental & Water Resources Institute; Water Environment Federation, Reston, Va., 2nd ed., 2007.

43 A. Saltelli, S. Tarantola and F. Campolongo, Sensitivity Analysis as an Ingredient of Modeling, *Stat. Sci.*, 2000, **15**, 377–395.

44 M. D. Mckay, R. J. Beckman and W. J. Conover, A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, *Technometrics*, 2000, **42**, 55–61.

45 C. Spearman, ‘General Intelligence,’ Objectively Determined and Measured, *Am. J. Psychol.*, 1904, **15**, 201–292.

46 J. Cohen, A Coefficient of Agreement for Nominal Scales, *Educ. Psychol. Meas.*, 1960, **20**, 37–46.

47 R. H. McCuen, Z. Knight and A. G. Cutter, Evaluation of the Nash–Sutcliffe Efficiency Index, *J. Hydrol. Eng.*, 2006, **11**, 597–602.

48 X. M. Zhu, B. Huang, S. D. Wang, J. L. Zheng, B. Yao and S. Chen, Research for Combined Drainage Networks in Chuanfang River Basin of Kunming City Based on SWMM, *Appl. Mech. Mater.*, 2012, **170–173**, 2380–2385.

49 M. K. Muleta, J. McMillan, G. G. Amenu and S. J. Burian, Bayesian Approach for Uncertainty Analysis of an Urban Storm Water Model and Its Application to a Heavily Urbanized Watershed, *J. Hydrol. Eng.*, 2013, **18**, 1360–1371.

50 A. J. Viera and J. M. Garrett, Understanding Interobserver Agreement: The Kappa Statistic, *Fam. Med.*, 2005, **37**, 360–363.