

Water Resources Research®



RESEARCH ARTICLE

10.1029/2024WR038394

Key Points:

- A comprehensive framework is proposed to quantitatively evaluate the skeletonization impacts on urban drainage network model predictions
- A new metric is developed to simultaneously account for flood volume and spatial range when assessing the model skeletonization errors
- Proposed framework is applied to assess two existing compensation methods for the mitigation of simulation errors induced by skeletonization

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

F. Zheng,
feifeizheng@zju.edu.cn

Citation:

Ji, Y., Zheng, F., Yang, Y., Shuai, J., Huang, Y., Kapelan, Z., & Savic, D. (2025). A comprehensive framework for evaluation of skeletonization impacts on urban drainage network model simulations. *Water Resources Research*, 61, e2024WR038394. <https://doi.org/10.1029/2024WR038394>

Received 10 JUL 2024

Accepted 20 JAN 2025

Author Contributions:

Data curation: Yongfei Yang, Jia Shuai

Methodology: Feifei Zheng

Writing – original draft: Yiran Ji

Writing – review & editing:

Feifei Zheng, Yuan Huang,
Zoran Kapelan, Dragan Savic

A Comprehensive Framework for Evaluation of Skeletonization Impacts on Urban Drainage Network Model Simulations

Yiran Ji¹ , Feifei Zheng¹ , Yongfei Yang², Jia Shuai², Yuan Huang³ , Zoran Kapelan⁴ , and Dragan Savic^{5,6} 

¹College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, China, ²China Energy Construction Group Guangdong Electric Power Design and Research Institute, Guangzhou, China, ³College of Water Conservancy & Hydropower Engineering, Hohai University, Nanjing, China, ⁴Department of Water Management, Delft University of Technology, Delft, The Netherlands, ⁵KWR Water Research Institute, Nieuwegein, The Netherlands, ⁶Centre for Water Systems, University of Exeter, Exeter, UK

Abstract Urban drainage network models (UDNMs) have been widely used to facilitate flood management. Typically, a UDNM is developed using data from Geographic Information Systems (GIS), and hence it consists of many short pipes and connection nodes or manholes. To improve modeling efficiency, a GIS-based model is generally skeletonized by removing many elements. However, there has been surprisingly a lack of knowledge on to what extent such skeletonization can affect the model's simulation accuracy, resulting in uncertainty in flood risk estimation. This paper makes the first attempt to quantitatively evaluate multidimensional impacts of different skeletonization levels on hydraulic properties of UDNMs. This goal is achieved by a new evaluation framework comprising of eight existing and new metrics that make use of hydrographs, main pipe hydraulics and flood distribution properties. A real-life UDNM is used to illustrate the new framework under various rainfall conditions and different skeletonization levels. The new framework is also used to compare the performance of two compensation methods in mitigating impacts caused by model skeletonization. Results obtained show that: (a) model skeletonization can significantly affect the magnitude of peak flow at the outfall, with a maximum overestimation of up to 20%, (b) hydraulics in main pipes can also be affected by model skeletonization with the maximum flow increasing up to 35%, and (c) model skeletonization may significantly alter the flood distribution properties which has been largely ignored in past studies. These findings provide guidance for UDNM skeletonization, where their associated impacts should be aware in engineering practice.

1. Introduction

Urban floods have been one of the most damaging hazards to cities in recent years, leading to huge economic losses and casualties (Zheng et al., 2015). Many actions have been taken to reduce the frequency of urban floods or mitigate their associated impacts, including urban drainage expansion (Lin et al., 2020), sponge city construction (Yin et al., 2021) and flood warning system development (Oh & Bartos, 2023). Implementations of these actions often rely on urban drainage network models (UDNMs) that are developed to simulate the hydraulic processes of the urban runoff (Radinja et al., 2021) on the ground and within the drainage pipes. For example, an effective flood warning is typically derived from the simulation results of the UDNM conditioned on a given rainfall event. Therefore, the development of the UDNMs is vital to enable the effective flood management and mitigation (Huang et al., 2022).

Generally, a UDNM is developed by collecting data from a Geographic Information System (GIS), which are often available in many cities (Niemi et al., 2019; Zhang & Pan, 2014). A GIS often consists of the typology of the drainage system, pipes, manholes and other hydraulic structures, representing a comprehensive database that can greatly facilitate the UDNM development. However, one challenge associated with the GIS-based model is that it often consists of many unnecessary details and elements such as a large number of short pipes, connection nodes or manholes. While being able to simulate the hydraulic process for each single pipe, such a model is often computationally demanding. This computational overhead may not be significant when only one or several simulations are conducted, but it can substantially go beyond the computational resources that are typically available when uncertainty analysis or optimization is needed (Maier et al., 2023). In addition, developing such detailed models often has high costs due to the extensive data collection and processing. Furthermore, a model

© 2025. The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

with a large number of elements would also affect the management (e.g., real-time control) and analysis efficiency in engineering practice (Huang et al., 2019).

To improve the modeling efficiency, surrogate models (also known as metamodels or emulators) or parallel computing technology can be used (Garzón et al., 2022; Nkwunonwo et al., 2020). Specifically, the former either uses simplified governing equations or data-driven models to directly capture the input-output relationships (Farina et al., 2023; Thrysøe et al., 2019). While being efficient, the surrogate models cannot ensure their performance when future conditions being considered (configurational changes) go beyond the data range used for model development (Farina et al., 2023; Kim & Han, 2020). While the parallel computing technology is effective in speeding up simulations for many different models (Kotyra, 2023; Xia et al., 2018), its implementation for a UDNM is not straightforward due to the underlying strong hydraulic connections between different elements within the UDNM (Burger et al., 2014). In addition, parallel computing resource is not easily accessible in engineering practice, hindering its wide application (Sadler et al., 2019). Another approach to improve the UDNM's modeling efficiency is the development of a conceptual model that is built by replacing pipes/structures/sub-networks with equivalent artificial tanks/reservoirs (Farina et al., 2023; Fischer et al., 2009; Mahmoodian et al., 2018). However, this type of approach, typically based on the preservation of mass balances only, ignores the hydraulics within the subnetworks that have been replaced by equivalent reservoirs, and hence is only of limited use in certain applications such as real-time control (e.g., van der Werf et al., 2023). In practice, where is the line draw between simplification and full catchment scale is application specific as it depends on the modeling scope and the level of detail information required (Achleitner et al., 2007). For instance, some model applications require looking at a catchment as a whole with more or less detail whilst in other cases it is sufficient to simulate only the urban drainage network, either with more detail (e.g., system optimization) or less detail (e.g., optimal operation of pumps or real time control).

In contrast to the above approaches, model skeletonization can be an effective approach to enable efficient model management and simulation (Huang et al., 2019). Skeletonization refers to a process that removes (or replaces with equivalent) some hydraulic elements that are believed to have rather limited impacts on the hydraulics of the original UDNM, such as pruning short pipes at terminal nodes or merging some short pipes into single equivalent pipes (Davidsen et al., 2017). The resulting skeletonized model is expected to represent the main hydraulic properties of the original full UDNM, while simplifying the complexity of model structures (Cantone & Schmidt, 2009). The main merit of the model skeletonization is its improvement in modeling efficiency whilst preserving prediction accuracy (Davidsen et al., 2017).

In engineering practice, skeletonization of UDNMs is considered necessary when (a) a skeletonized model can significantly improve simulation efficiency for a real, large UDNM under a long time rainfall event (e.g., 1 year, Huang et al., 2022); (b) a skeletonized model can be useful for testing many rainfall scenarios or simulations required during the optimization design, optimal operation or uncertainty analysis (Huang et al., 2019), (c) the management efficiency (e.g., real-time control) can be substantially enhanced by a skeletonized model, which can greatly facilitate its practical implementation, and (d) model skeletonization is able to maintain the main physical structure or process of the UDNMs, but this cannot be achieved by black-box or simplified methods such as machine-learning surrogate models (Garzón et al., 2022) or conceptual models (Fischer et al., 2009; Mahmoodian et al., 2018). In other words, the skeletonized models are more interpretable and transparent relative to simplified models, as the main hydraulic dynamics and relationships between elements within the network are preserved.

It is acknowledged that model skeletonization can induce simulation errors and additional uncertainty compared to the original full model (Hellbach et al., 2012). Such uncertainty combined with other sources of uncertainties can be one of the main limitations for model's practical implementation (Pedersen et al., 2022). However, there have been surprisingly few studies investigating the potential impacts associated with model skeletonization. In other words, while model skeletonization has been widely used in engineering practice, its induced simulation errors and uncertainties have been largely ignored or have not been systematically investigated. Limited references include Davidsen et al. (2017) who showed that the peak flow of the outfall of a skeletonized model can deviate from the full model by about 15%. Cantone and Schmidt (2009) demonstrated that there is a possibility that the user might not correctly predict the magnitude, timing, and shape of the outfall hydrograph when using simplified models. Therefore, there is still a lack of knowledge of the extent such skeletonization can affect the model's simulation accuracy which, in turn, leads to uncertainty in flood risk estimation.

Table 1
Summary, Purposes, and Importance Statements of Evaluation Framework Metrics

Metrics	Purpose	Importance
Magnitude of peak flow at the outfall	Assess the changes in the outfall peak flow	Critical for model calibration and flooding prediction
Timing of peak flow at the outfall	Focuses on the time that peak flow occurs	Important for model calibration and flooding analysis
Hydrograph similarity at the outfall	Assess the hydrograph similarity at the outfall	Important for model calibration and runoff analysis
Hydrograph similarity at monitoring locations	Evaluate the hydrograph similarity at the intermediate locations within the UDNM	Important for flooding location analysis
Overall water depth in main pipes	Understand the overall water depth changes in main pipes induced by model skeletonization	Import for overflow prediction and management
Overall velocity in main pipes	Reveal the overall pipe velocity variations induced by model skeletonization	Important for model calibration and sediment analysis
Overall flow in main pipes	Show the overall pipe flow changes induced by model skeletonization	Important for model calibration and system operation
Flood property similarity	Evaluate the deviations of flood properties induced by model skeletonization	Critical for flooding analysis and management

To this end, this paper makes the first attempt to develop a comprehensive framework that quantitatively evaluates the skeletonization impacts on UDNM predictions. The framework considers different aspects of the UDNM simulations, including the magnitude and timing of peak flow at the outfall, the hydrograph similarity of the outfall flow, the hydrograph similarity at monitoring locations, the overall water depth, velocity and flow in main pipes as well as the flood volume (depth) and flood ranges. The main contributions/novelties of this study include: (a) the proposal of a new framework that enables a comprehensive and systematic evaluation of UDNM's skeletonization on its predictions, (b) the development of a new flood property similarity (FPS) metric that is able to simultaneously account for flood volume (depth) and spatial range when assessing the model skeletonization errors, (c) assessment of the proposed framework's contribution to two existing compensation methods for the mitigation of simulation errors induced by model skeletonization, and (d) the theoretical exploration between model skeletonization errors, rainfall properties (e.g., peak intensity (PI) and return periods (RPs)) and model outputs (predictions).

2. Methodology

2.1. Evaluation Framework

The framework developed for evaluating the impacts of model skeletonization on UDNMs is introduced here. As outlined in Table 1, the framework consists of eight key metrics designed to provide a thorough assessment, covering the main hydraulic processes of a UDNM including the outfall, the monitoring locations, the main pipes, and the flood outcomes. As shown in Table 1, the hydraulics and the hydrograph at the outfall are typically important in engineering practice as they represent direct indicators for model performance evaluation (Davidsen et al., 2017). In addition to the outfall, this study considers the hydrographs at the monitoring locations within the UDNM as it can offer an overall hydraulic assessment of the UDNM. The rationale behind this is that a skeletonized model may produce small hydraulic errors at the outfall but it may still generate larger simulation deviations at intermediate urban drainage network locations. For a UDNM, it is critical to ensure predictions for the main pipes (pipes with larger diameters, e.g. diameters greater than 800 mm) match the real values as these pipes can significantly affect the system delivery ability. Therefore, the hydraulics in the main pipes, including water depth, velocity and flow are considered in the proposed framework for assessing the performance of a skeletonized model. Typically, the main purpose of a UDNM is the prediction of urban flood for a given rainfall event, and hence the associated flood simulation errors produced by model skeletonization are of great importance too. In recognizing this, a new metric called FPS is proposed in this study.

It is noted that pumps and pressure mains are not directly considered in the metrics, but they can be easily accounted for in applications. More specifically, pump flows can be assessed using the metric of hydrograph similarity at monitoring locations (i.e., the pumping flows are often monitored). The pressure mains can be

assessed by the metrics of overall velocity and flow in main pipes. The summary, purposes and importance statements of evaluation framework metrics are given in Table 1.

2.1.1. Magnitude and Timing of Peak Flow at the Outfall

The magnitude and timing of peak flow at the outfall are critical aspects in assessing a UDNM's simulation ability during rainfall events, and thus they are included in the proposed framework. This is because accurate prediction of peak flow is vital for effective flood warning, mitigation and emergency response planning. In this study, two metrics called the peak value deviation percentage (PDP) and the peak time change (PTC) are utilized to evaluate changes of the outfall peak flow induced by model skeletonization. These two metrics are defined as follows:

$$\text{PDP} = \frac{P_s - P_o}{P_o} \times 100\% \quad (1)$$

$$\text{PTC} = PT_s - PT_o \quad (2)$$

where P_s and P_o are the peak value of a hydrograph of the skeletonized model and the original model respectively, PT_s and PT_o are the peak time of a hydrograph of the skeletonized model and the original model respectively.

2.1.2. Hydrograph Similarity at the Outfall and Monitoring Locations

Hydrographs similarity at the outfall and locations with sensors can be important metrics to evaluate the impacts of the skeletonization on the model's performance. A number of metrics are available to calculate the hydrograph similarity, such as the Kling-Gupta Efficiency (KGE, Gupta et al., 2009), the Nash-Sutcliffe Efficiency and root mean square error. In this study, the KGE is utilized, as it is commonly used for assessing the similarity between two hydrographs (Zheng et al., 2023), but the use of other metrics does not affect the application of the proposed framework. The KGE is defined as follows:

$$\text{KGE} = 1 - \sqrt{\left(\frac{\text{cov}(Q_o, Q_s)}{\sigma_{Q_o} \sigma_{Q_s}} - 1\right)^2 + \left(\frac{\sigma_{Q_s}}{\sigma_{Q_o}} - 1\right)^2 + \left(\frac{\mu_{Q_s}}{\mu_{Q_o}} - 1\right)^2} \quad (3)$$

where Q_s and Q_o are the hydrograph of the skeletonized model and the original model respectively. $\text{cov}(Q_o, Q_s)$ is the covariance of these two hydrographs. σ_{Q_s} and σ_{Q_o} are the standard deviation of the hydrograph of the skeletonized model and the original model respectively. μ_{Q_s} and μ_{Q_o} are the mean of a hydrograph of the skeletonized model and the original model respectively. A higher KGE value indicates a greater similarity between two hydrographs, with $\text{KGE} = 1$ indicating identical hydrographs.

2.1.3. Overall Hydraulic Properties in Main Pipes

In addition to the hydrograph similarity at the outfall and locations with sensors, overall hydraulic properties in main pipes can be important to evaluate the simulation impacts caused by model skeletonization. The water depth, velocity and flow in the main pipes represent the predominant response characteristics of a UDNM to a rainfall event. Focusing solely on main pipes (say pipe with diameter larger than 800 mm) is because these pipes are more likely to play a more important role than smaller-diameter pipes due to their associated large flow. Therefore, main pipes are typically retained during model skeletonization. However, the impacts of model skeletonization on these hydraulic properties has been largely neglected in the literature. To this end, this paper considers the following metric to assess the hydraulic changes in the main pipes caused by model skeletonization:

$$\text{RMM(H)} = \frac{\frac{1}{N} \sum_i^N \max(H_{s,i}) - \frac{1}{N} \sum_j^N \max(H_{o,j})}{\frac{1}{N} \sum_j^N \max(H_{o,j})} \times 100\% \quad (4)$$

where H is a state variable of main pipes, such as water depth, velocity and flow, and RMM(H) is the relative mean error of that state variable between the skeletonized model and the original model. $H_{s,i}$ and $H_{o,j}$ are the time series of a particular state variable at the i th main pipe in the skeletonized model and the original model

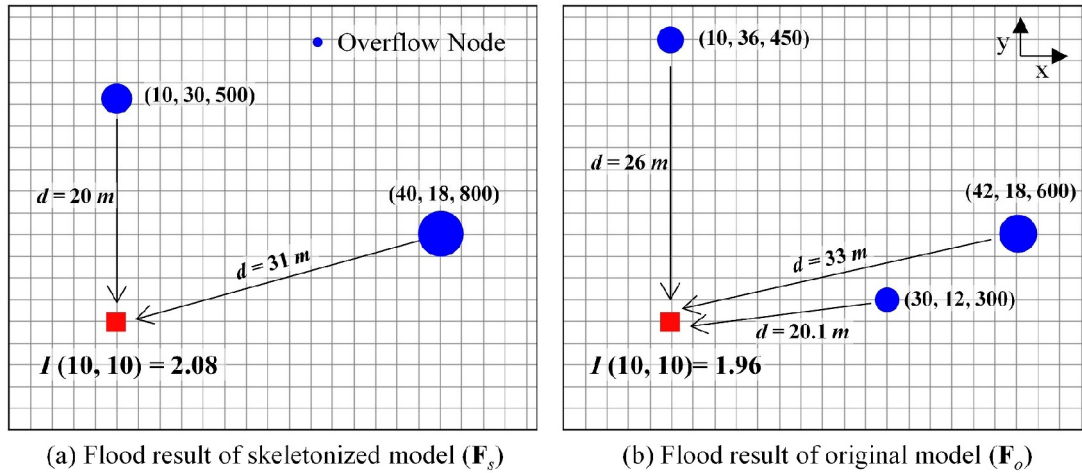


Figure 1. Spatial distribution of flood results F_s (a) and F_o (b) shown on a quadratic grid, where the triplet (a, b, V) represents x and y -axes of the overflow node and V the overflow volume.

respectively. N is the total number of main pipes, which is the same in both models. $\frac{1}{N} \sum_i^N \max(H_{s,i})$ and $\frac{1}{N} \sum_j^N \max(H_{o,j})$ are the averages of maximum values of a state variable H of all main pipes in the skeletonization model and the original model, respectively. In this study, the pipes with diameters greater than 800 mm are considered as the main pipes, but a different choice of the main pipes would not affect the application of the proposed evaluation metric.

2.1.4. Flood Property Similarity

As previously stated, one purpose of a UDNM is to simulate the flood process driven by a given rainfall event, and hence the flood simulation accuracy is crucial for a UDNM. The flood property can be mainly assessed by the flood volume and the flood locations, where the former aspect has been reported in literature (Davidsen et al., 2017) but the latter has not been explicitly considered as an evaluation metric. The number and/or the spatial locations of the overflowing manholes is important in addition to the flood volume or depth to enable effective flood mitigation decision-making. However, it is not straightforward for a single metric to assess the FPS that can simultaneously accounts for flood volume and locations.

This study proposes a new metric, termed “flood property similarity” (FPS), to simultaneously measure the flood volume (depth) and overflow manhole location differences due to model skeletonization. Consider that F_s and F_o , represent flood results of the skeletonized and the original model respectively. Each input F is a set of three-dimensional arrays, that is, $F = \{(a_1, b_1, V_1), (a_2, b_2, V_2), \dots, (a_n, b_n, V_n)\}$, where n is the number of flooded manholes, referred to as overflow nodes. Each (a_i, b_i, V_i) represents the coordinates (a, b) in meters of the i th overflow node and its associated total overflow volume V in cubic meters.

Figure 1 illustrates the calculation of the FPS. In Figure 1a, the skeletonized model (F_s) has two overflow nodes, $(a = 10, b = 30, V = 500)$ and $(40, 18, 800)$, whereas the original model (F_o) in Figure 1b has three overflow nodes, $(10, 36, 450)$, $(30, 12, 300)$ and $(42, 18, 600)$. When comparing the difference between these two flood maps, it is not straightforward to differentiate them using a single quantitative value. This is because the flood volumes at each overflow node, the locations of the overflow nodes and the number of the overflow nodes are all different between the skeletonized and original models. To address this issue, the FPS proposed in this study measures a difference to all other grid nodes within the flood area using the impact values induced by the overflow nodes.

More specifically, the FPS metric quantifies the similarity between F_s and F_o in two steps: (a) by generating impact value maps for both models using Equations 5 and 6, and (b) by comparing these maps to derive the final FPS value. As shown in Equation 5, the impact value $f(x, y, a, b, V)$ is a function of the overflow node (a, b, V) , where (x, y) is the grid cell being affected by this overflow node. Let us assume that impact of an overflow

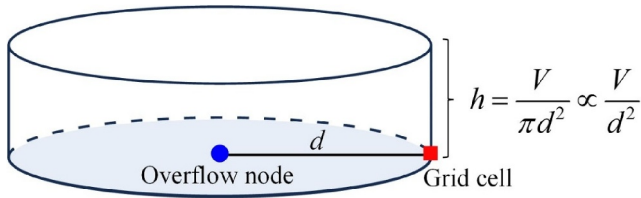


Figure 2. Illustration of the physical representation of the impact value associated with each overflow node on its surrounding grid cell.

network node on its surrounding grid cells exhibits a cylinder shape with an influence radius d (as shown in Figure 2), where the water depth h at each grid node is used to represent the impact value. This h can be approximately estimated using $\frac{V}{d^2}$, as shown in Figure 2 and Equation 5.

In engineering practice, for the grid cells that are very close to the overflow node, the d can be very small, leading to a very large value of $\frac{V}{d^2}$. To solve this problem, a $d_{\min}$ is defined where all grid cells within the $d_{\min}$ distance to the overflow node have identical impact value.

Following above method, for the grid cell at (10, 10) in Figure 1a, the impact values produced by the overflow nodes (10, 30, 500) and (40, 18, 800) are 1.25 and 0.83 respectively. Consequently, the total impact value produced by the two overflow nodes on the cell (10,10) is $I = 2.08$, as calculated using Equation 6. In a similar way, the impact value produced by the three overflow nodes on the cell (10,10) (see Figure 1b) is $I = 1.96$.

$$f(x, y, a, b, V) = \frac{V}{d^2}$$

$$d = \begin{cases} d_{\min}, & \text{if } (x - a)^2 + (y - b)^2 \leq d_{\min}^2 \\ (x - a)^2 + (y - b)^2, & \text{if } (x - a)^2 + (y - b)^2 > d_{\min}^2 \end{cases} \quad (5)$$

$$I(x, y) = \sum_{i=1}^n f(x, y, a_i, b_i, V_i) \quad (6)$$

As shown in Equations 5 and 6, the impact values across the area are influenced by the flood volume and coordinates of all overflow nodes. Changes in either would alter the impact value distributions. Therefore, the FPS metric can be calculated using the impact value distributions as shown Equation 7.

$$\text{FPS} = 1 - \frac{\sum_{j=1}^G |I(x_j, y_j, \mathbf{F}_s) - I(x_j, y_j, \mathbf{F}_o)|}{\sum_{j=1}^G I(x_j, y_j, \mathbf{F}_o)} \quad (7)$$

where $I(x_j, y_j, \mathbf{F}_s)$ and $I(x_j, y_j, \mathbf{F}_o)$ are impact value of the j th grid node, and the total number of the grid cell is G .

The numerator in Equation 7 is the sum of the absolute differences of impact values at each grid cell before and after model skeletonization, while the denominator is the sum of impact values for all grid cells in the original model. The value of FPS ranges from $-\infty$ to 1 (similar to KGE), where a larger FPS indicates a greater similarity in the flood property between the skeletonized model ($\mathbf{F}_s$) and the original model ($\mathbf{F}_o$), where this value can simultaneously consider the variation in flood volume and distribution. FPS = 1 represents that the flood simulation results from the $\mathbf{F}_s$ are identical to those from $\mathbf{F}_o$. It is noted that the FPS depends on the direct surface distance. The advantage of this metric over the use of the network graph is that it can directly account for the spatial distribution of floods. This is because the network graph indicates the flow directions of the runoff in the pipes, which can be significantly different from the flooding locations on the ground surface which is mainly dependent on the surface elevation.

2.2. Skeletonization Method

Various techniques for model skeletonization are available, including trimming, pruning, merging of branches (Leitão et al., 2010) and removal of pipes based on their importance (Meijer et al., 2018). Model trimming is used to remove some parts of the UDNM that are considered to be not important, which is often applied to deal with case studies with many areas composed of small pipes. Merging of branches is often employed to connect pipes by removing their intermediate nodes, and this method can be adopted when it is aimed at reducing the number of nodes is the purpose. The method of removal of pipes based on their importance is often carried out by deleting some pipes that have rather limited impacts on the UDNM's overall hydraulic performance, which is conditioned on the hydraulic simulation results of the UDNMs. The present study does not focus on the application of different

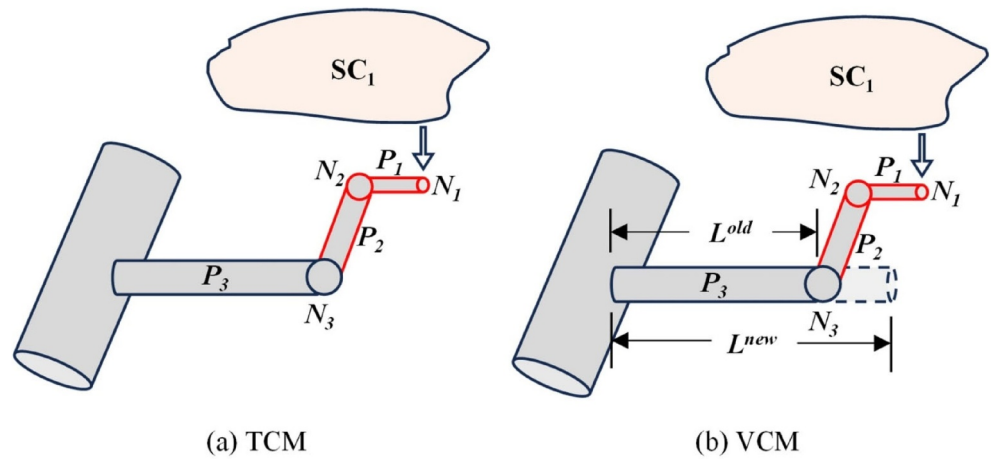


Figure 3. Illustration of time-compensation method and volume-compensation method.

skeletonization methods, and hence a straightforward and widely accepted pruning technique is used here. More specifically, this skeletonization method removes the pipe when the following two specific conditions are simultaneously met: the pipe does not have any upstream pipes, and its length is shorter than a user-defined threshold, L_t . When the pipe is removed, its corresponding sub-catchments are reconnected to its downstream node. Such a process is carried out from the downstream ends of the UDNM to its upstream pipes until no pipe meets the specified conditions. Notably, a larger L_t indicates a larger level of skeletonization. While the use of different skeletonization methods does not affect the application of the proposed framework, future studies should explore the performance of other skeletonization methods such as the removal of pipes based on their importance (Meijer et al., 2018).

2.3. Compensation Methods

In this study, two compensation methods proposed by Davidsen et al. (2017) are compared in their performance to mitigate the simulation errors introduced by model skeletonization. This is done with the aid of the proposed framework. These two methods are the time-compensation method (TCM) and the volume-compensation method (VCM).

2.3.1. TCM

When a pipe is removed during the skeletonization process its upstream sub-catchments are connected to its downstream node. This indicates that the runoff from these sub-catchments would reach the downstream node earlier compared to the original model due to ignoring the traveling time in the pipe that is removed. The TCM accounts for the effect by adding extra time to each sub-catchment's rainfall time series.

Figure 3a is used to illustrate the TCM, where a sub-catchment (SC₁) is associated with outlet N₁. Assuming that pipes P₁ and P₂ have been removed due to their short length, and hence SC₁ is now connected to node N₃. The travel time for each removed pipe j can be estimated using Equation 8, and the compensation time for SC₁ is the sum of the travel time in all removed pipes.

$$T_i = \sum_j \frac{n_j \cdot L_j}{R_{h,j}^{2/3} \cdot S_j^{1/2}} \quad (8)$$

where T_i is the compensation time associated with sub-catchment i ; L is the pipe length [m], n is Manning's coefficient [s/m^{1/3}], R_h is the hydraulic radius [m], S is the pipe slope, and subscript j denotes j th pipe that has been removed. Different sub-catchments have different compensation times and this value is used to delay the start time of the rainfall event whose value is T_i for this sub-catchment.

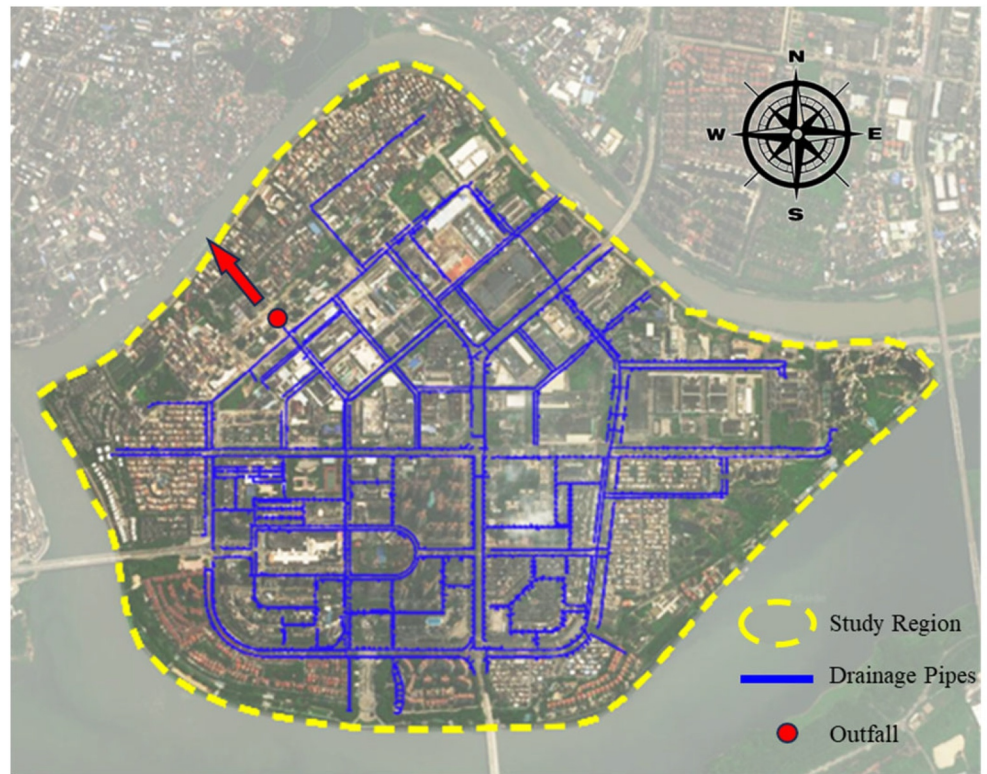


Figure 4. The urban drainage network model for the SL city located in Guangdong Province, China.

2.3.2. VCM

Whilst the model skeletonization simplifies the topology of the UDNM it inevitably reduces the total volume (i.e., storage capacity of) of the network. The VCM aims to preserve the original volume characteristics of the network as much as possible. As shown in Figure 3b, the total volume of all pipes that have been removed during the skeletonization process is added to the pipe directly connecting to these removed pipes, with added volume determined as follows:

$$L_i^{\text{new}} - L_i^{\text{old}} = \frac{\sum_j L_j \cdot A_j}{A_i} \quad (9)$$

where L_i^{new} is the new pipe length of pipe P_i associated with SC_i , where its upstream pipes are removed as shown in Figure 3b. L_i^{old} is the original pipe length. A_i and A_j are the pipe cross-section area of P_i and P_j , respectively.

3. Case Study

3.1. Description

The urban drainage network tested in this paper is located in SL, a small town in Guangdong Province, China. As shown in Figure 4, it covers an area of 4 km² and serves approximately a population of 100,000 people. The network was built in the 1990s and was designed to collect, convey, and discharge rainwater only. To construct the UDNM for this case study area, the pipe data, the digital elevation model data (5 m × 5 m resolution) and the land use data were all collected from the GIS database provided by the local government drainage department. The analyzed network was modeled using the Storm Water Management Model (SWMM), which has been widely used in engineering practice (Rossman, 2015). The D8 method (Warsta et al., 2017) was applied to identify the sub-catchments for each manhole of the UDNM. The full model consists of 5,104 pipes with a total length of 58 km, and has one outlet located in the northeast of the study area as shown in Figure 4.

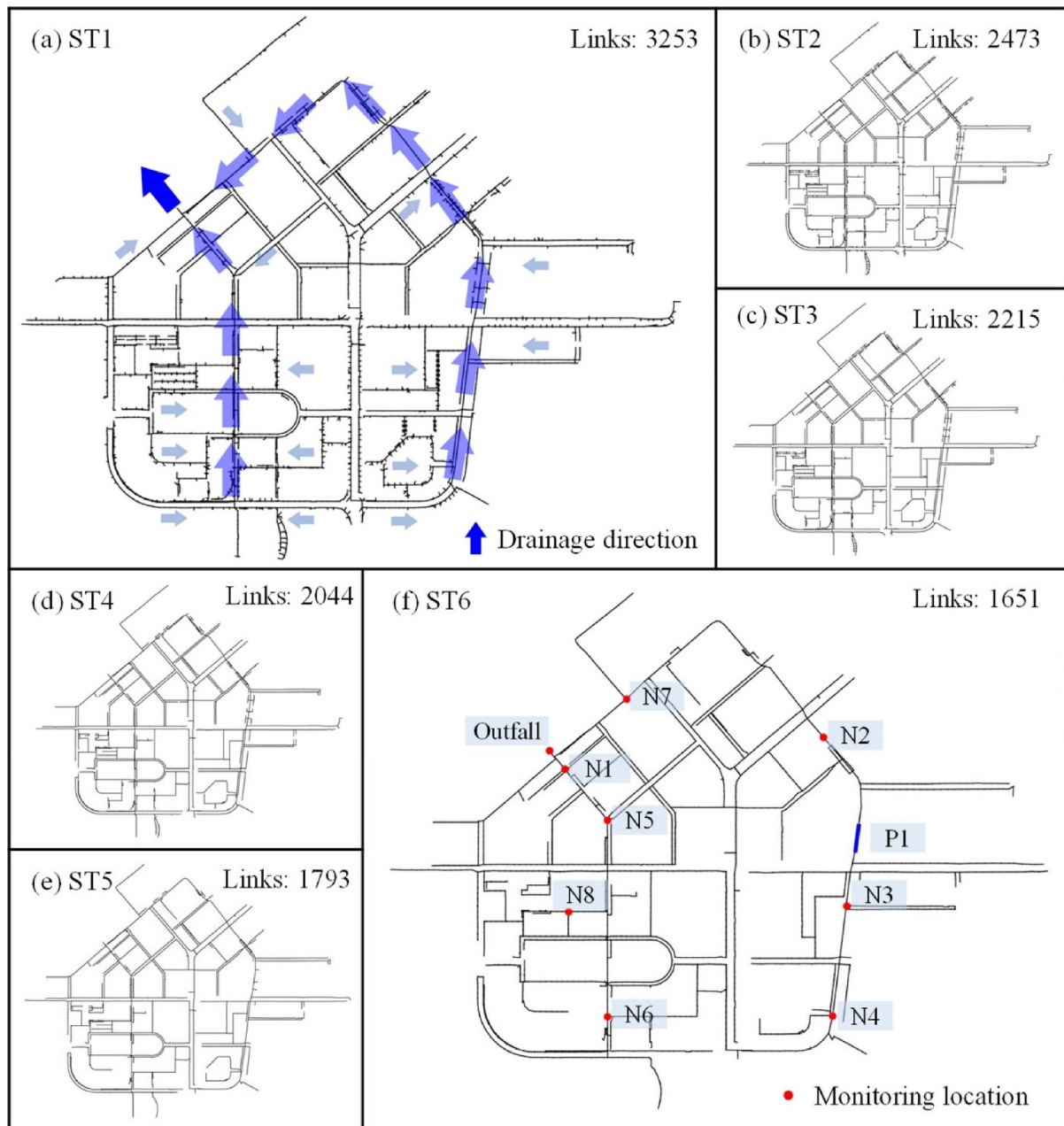


Figure 5. Comparative diagrams of the urban drainage network model with six levels of skeletonization (ST1-ST6).

The original full model was skeletonized by the method proposed in Section 2.2 at six different levels by using length thresholds of $L_i = 5, 10, 15, 20, 25,$ and 30 m respectively. The six skeletonized models were labeled as ST1, ST2, ST3, ST4, ST5, and ST6, respectively. ST1 represents the lowest degree of skeletonization corresponding to a $L_i = 5$ m and ST6 represents the highest degree of skeletonization corresponding to a $L_i = 30$ m. Detailed information about skeletonized networks is displayed in Figure 5, where Figure 5a also shows the drainage flowrate direction and Figure 5f shows the monitoring locations of this UDNM.

3.2. Rainfall Events Used in This Study

To understand how skeletonization affects the model's performance, a total of 64 rainfall events were used using combinations of eight rainfall intensities and eight rainfall patterns. The eight rainfall intensities considered represent different RPs of 1, 2, 5, 10, 20, 30, 50, and 100 years, respectively. The eight rainfall patterns were taken

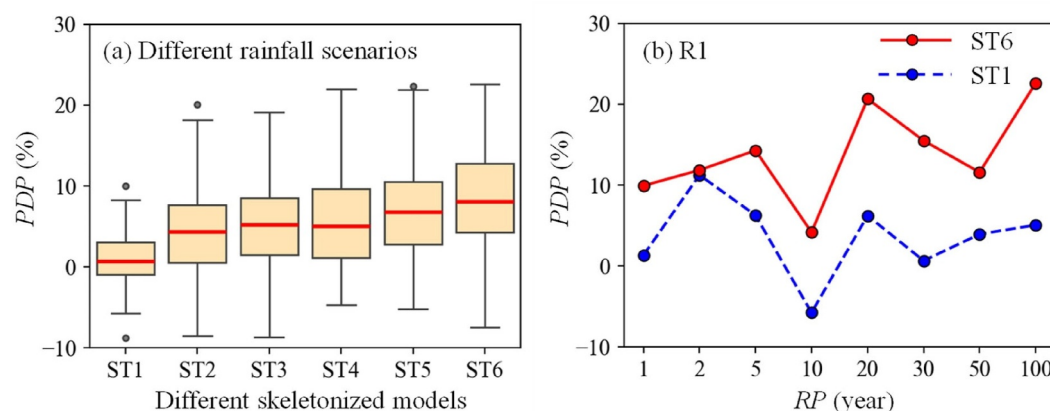


Figure 6. (a) Peak flow deviation percentage (PDP) at the outfall across 64 rainfall scenarios and (b) PDP values of ST1 and ST6 models under R1 pattern.

from seven recorded events (Figures S1a–S1c and S1e–S1h in Supporting Information S1) and a single typical design event of Chicago-type (Lin et al., 2020) with a peak ratio of 0.5 and a duration of 2 hr (Figure S1d in Supporting Information S1). The 64 rainfall events generated this way were provided as inputs into both original and skeletonized models of the analyzed urban drainage network. At the end, TCM and the VCM introduced in Section 2.3 were implemented for different skeletonization levels.

4. Results and Discussions

4.1. Performance Analysis of Skeletonized Models

4.1.1. Magnitude and Timing of Peak Flow at the Outfall

Figure 6a illustrates the distribution of the peak value deviation percentage (PDP) at the outfall for different skeletonized models for 64 analyzed rainfall events. As shown in this figure, the PDP values at the outfall are increasing for higher levels of model skeletonization with a mean PDP value for ST6 being about 8.6%. It can also be observed that the maximum PDP value obtained at the outfall that the peak flow of the skeletonized model can be as much as 20% larger than the corresponding original model value. This highlights that model skeletonization can result in a significant overestimation of flood risk, which should be given sufficient attention in engineering practice.

Figure 6b shows that while higher model skeletonization can overall induce a larger peak flow simulation value compared to the original model, large variations can be also observed. Interestingly, the skeletonized model can underestimate the outfall flowrate under some particular rainfall patterns and intensities as negative PDP values are observed. For example, for RP of 10 years, the ST1 model produces an outfall flow with PDP about -8% , but for RP of 5 years the corresponding PDP is about $+5\%$. Similar observations can be found for the ST6 model, where PDP is about 4% , which is significantly lower than the values from $RP = 5$ or 20 years as shown in Figure 6b.

We analyze further the variations of the PDP values of skeletonized models caused by increasing rainfall intensity (RPs) and the presence of negative PDP values in Figure 6b. This is mainly because the peak flow at the outfall is determined by the superposition of the peak flow of all its upstream pipes, and the latter is a function of the rainfall intensity and patterns. In other words, the peak flows at the pipes of the UDNM can exhibit a significant spatiotemporal variation, which can be affected by both the rainfall properties and pipe structure (e.g., length, slope). Therefore, for the skeletonized models, a low or even negative value of PDP is possible for a large rainfall event. This suggests an improved understanding of the impacts of the model skeletonization is very important. The performance of skeletonized UDNMs should be also thoroughly evaluated under a wide range of rainfall scenarios before engineering application.

In addition to examining the magnitude of peak flow deviation, PTC has been also investigated in this study. The results (not shown in the paper) indicate that the timing of the peak flow is not substantially affected by model

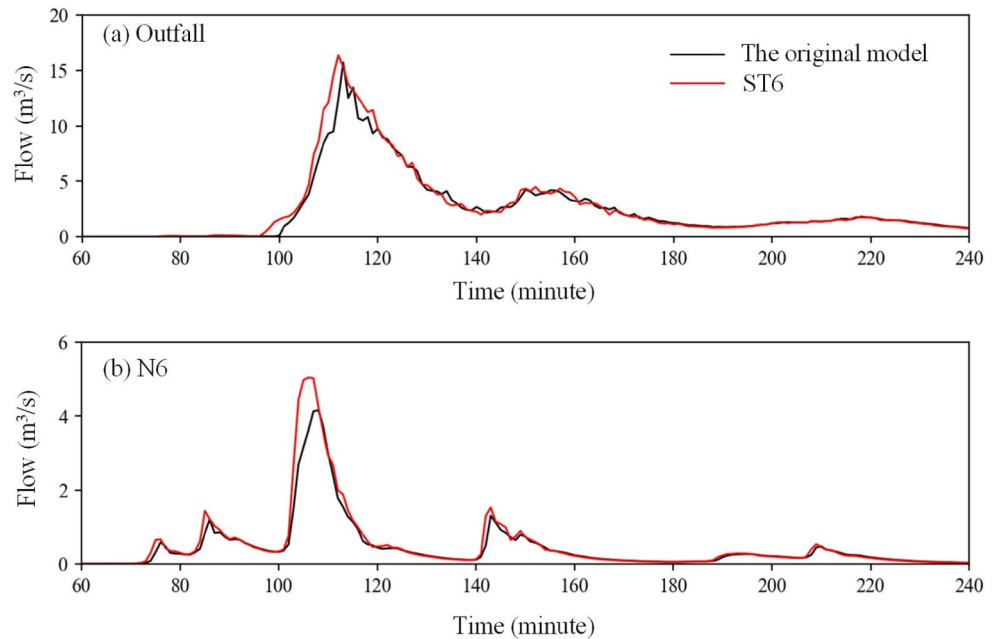


Figure 7. Hydrography of the original model and ST6 under R1 and RP = 10.

skeletonization, matching well with the findings reported by Cantone and Schmidt (2009). More specifically, for this case study, the PTC values are within ± 4 min for all different skeletonized models under different rainfall events. It is anticipated that such a deviation would not significantly affect the engineering applications of the skeletonized models.

4.1.2. Hydrograph Similarity at the Outfall and Monitoring Locations

Figure 7 shows the flow hydrograph of the ST6 and the original model at the outfall and N6 monitoring location (locations shown in Figure 5). As it can be seen from this figure the two model hydrographs are similar in both cases with monitoring location N6 exhibiting a larger deviation in peak flow than the outfall. This shows that the model skeletonization would not significantly influence the overall flow evolving pattern of the UDNMs. This can also be seen from the obtained KGE values shown in Table 2 (the values shown are averages across 64 rainfall events).

Table 2
Average Kling-Gupta Efficiency Values Across 64 Rainfall Scenarios

Node	Models					
	ST1	ST2	ST3	ST4	ST5	ST6
N1	0.97	0.93	0.92	0.92	0.90	0.89
N2	0.99	0.98	0.98	0.98	0.97	0.96
N3	0.90	0.89	0.88	0.89	0.88	0.87
N4	0.93	0.93	0.92	0.92	0.90	0.87
N5	0.97	0.93	0.92	0.92	0.91	0.90
N6	0.99	0.84	0.84	0.83	0.83	0.82
N7	0.99	0.98	0.98	0.97	0.96	0.95
N8	0.98	0.98	0.96	0.95	0.95	0.92
Outfall	0.99	0.96	0.96	0.95	0.94	0.93
Mean	0.97	0.94	0.93	0.93	0.92	0.90

As it can be seen from Table 2, the KGE values generally decrease with the increasing model skeletonization level but the deterioration of the model performance is rather minor. This is reflected by the fact that the mean of the KGE for ST6 is 0.90, which is also acceptable for engineering application. Note also that whilst the KGE value of the outfalls are consistently high some of the upstream nodes have relatively low KGE values. For example, the KGE of the outfall is 0.93 for ST6, but this value is only 0.82 for node N6 that is located upstream. This implies that a good hydrograph similarity between the original and skeletonized model at the outfall does not necessarily indicate the same high performance for the nodes at the intermediate locations within the UDNM. This shows that a comprehensive evaluation in addition to the outfall is necessary to gain an improved understanding on how the model skeletonization affects its performance.

4.1.3. Water Depth, Velocity, and Flow in Pipes

The hydraulics in the main pipes (defined in this case study as pipes with diameter greater than 800 mm) are analyzed here in terms of water depth,

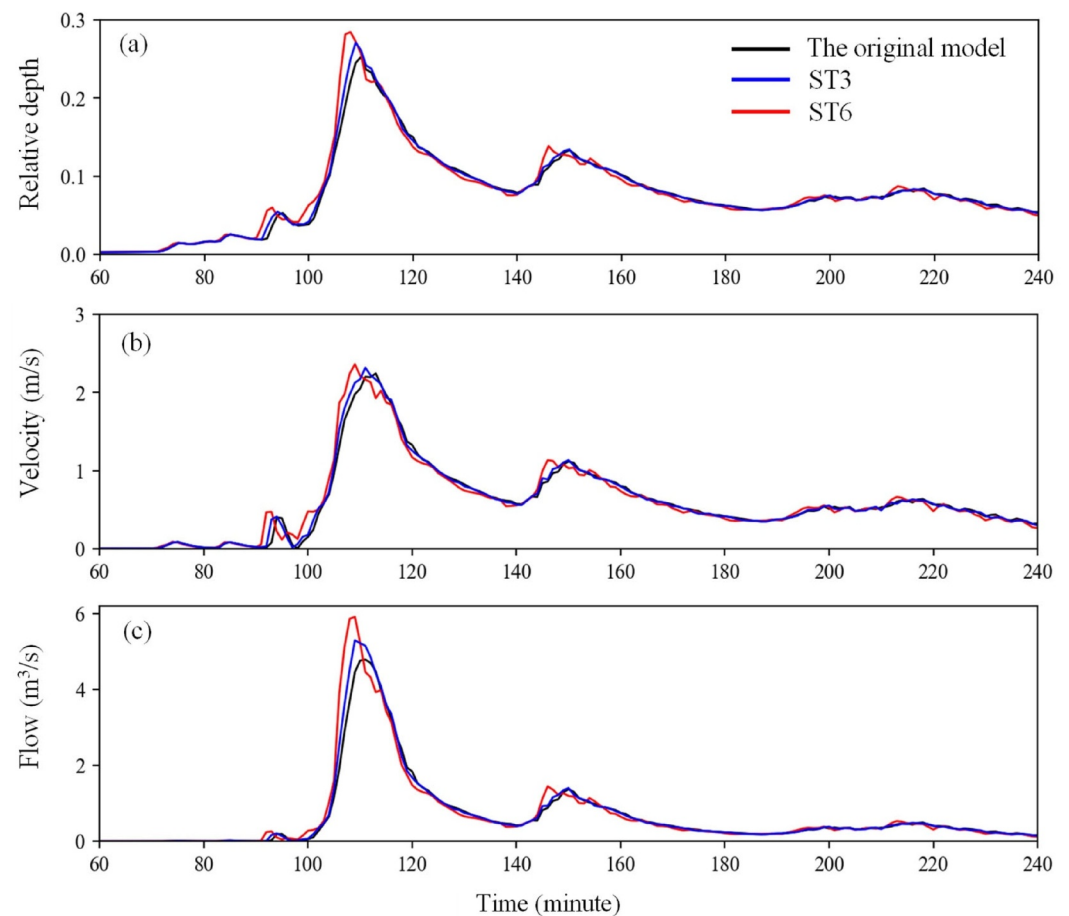


Figure 8. The relative depth, velocity and flow of a typical pipe (P1 in Figure 5) for the original model, ST3, and ST6 respectively.

velocity and flow. The water depth is represented by the relative depth which is defined as the ratio of water depth and pipe diameter. Figure 8 shows the relative depth, velocity and flow for a typical pipe (P1 shown in Figure 5f, whose diameter is 2,300 mm) between the original, ST3 and ST6 skeletonized models. As it can be seen from this figure, the model with a high skeletonization level (ST6) tends to have a larger peak flow and a slightly earlier occurrence of peak flow time when compared to the original model. Figure 9 (for R1, R2, R3 and R4) and Figure S2 (for R5, R6, R7, and R8, in Supporting Information S1) shows the relative mean error of the maximum values (RMM) for the relative depth, velocity and flow between the original model and the model with different skeletonization levels for different rainfall RPs.

As it can be seen from Figure 9 and Figure S2 in Supporting Information S1, model skeletonization consistently introduces an overall larger maximum value of the water depth, velocity and flow values in the main pipes compared to the original model, where a higher skeletonization level is more likely to produce a larger value of these hydraulic variables. In addition, such impacts can also be affected by the rainfall patterns. For instance, when compared to the original model, the maximum water level and the maximum flow increase by an average of 15% and 30% respectively in the main pipes when the R1 rainfall pattern is used for ST6, but these two values are reduced to about 10% and 20% if R4 is used. When R5-R8 rainfall events are considered, RMM values for the relative depth, velocity and flow between the original model and the model with different skeletonization levels (and for different rainfall RPs) become small. This is because the peak intensities of the R5-R8 rainfall events are overall lower than those of R1-R4 events.

To explore the impacts of PI of different rainfall patterns on RMM values, Figure 10 presents the average RMM values for relative depth, velocity and flow over different RPs for skeletonized models ST3 and ST6 under eight

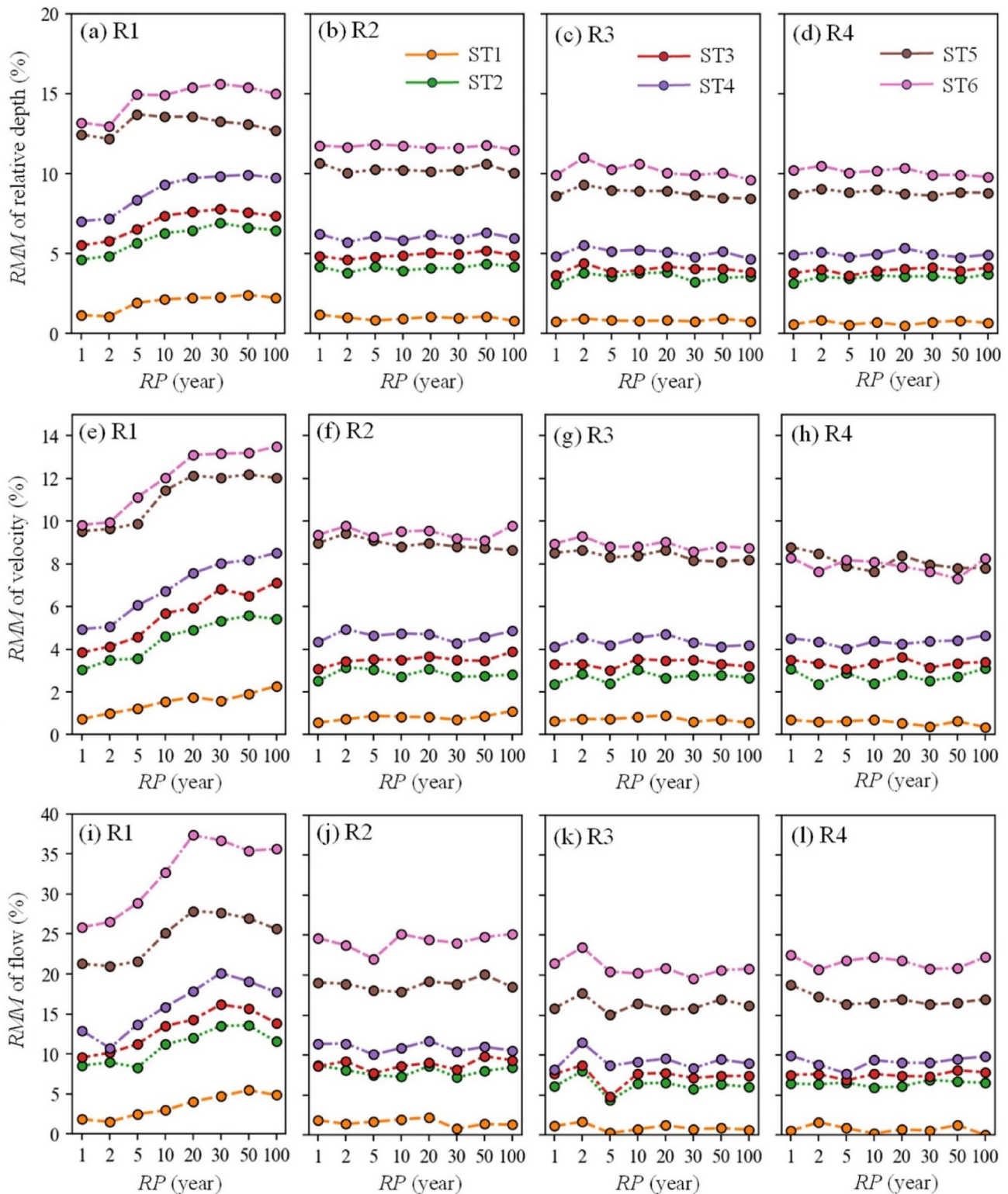


Figure 9. RMM values of relative depth, velocity, and flow of skeletonized models across R1, R2, R3, and R4 rainfall events.

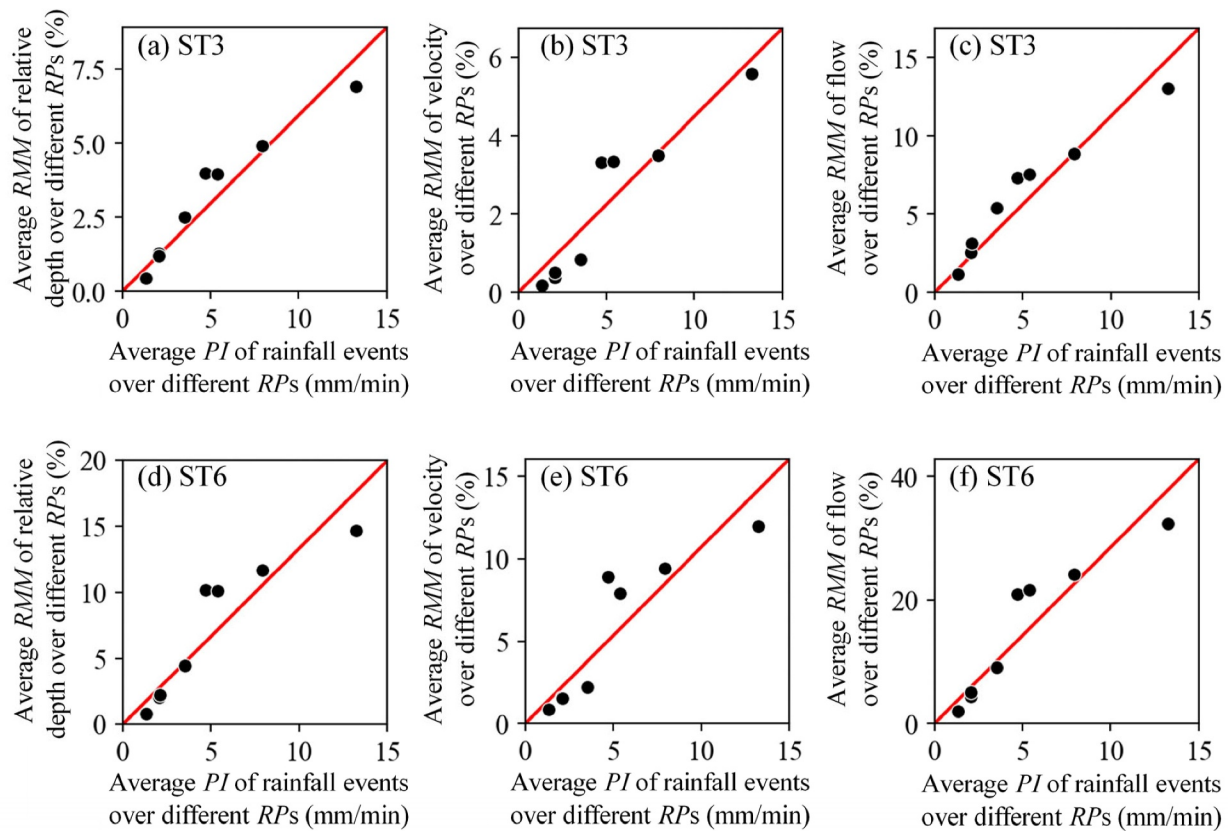


Figure 10. Average RMM values versus average peak intensity of rainfall events over different return periods.

rainfall events. The x-axis in Figure 10 shows the average PI of each of the eight rainfall events over eight RPs. For instance, R1 has the highest average PI = 13.2 mm/min, while R6 has the lowest average PI = 1.3 mm/min. The results in Figure 10 reveal that there is a strong linear relationship between the average PI of rainfall events and the average RMM values for relative depth, velocity, and flow. In other words, for a skeletonized model, a rainfall event with a larger PI is likely to lead to larger simulation errors for main pipes.

4.1.4. FPS

Figure 11 illustrates the difference in flood locations between the original model and ST6 under rainfall pattern R1 with RP = 100 years. In this figure, the black dots represent flood nodes in both the original model and skeletonized model, blue dots indicate flood nodes in the original model that have been removed by skeletonization, and red dots denote new flood nodes induced by model skeletonization. As shown in Figure 11, the flood has significantly changed spatially due to skeletonization (as a number of blue and red dots exist).

More specifically, many areas exist with blue dots but no red dots, indicating that many flood areas in the original model are not recognized in the skeletonized model. More importantly, many red dots appear as a result of model skeletonization, implying an overestimation of flood risk for these locations.

Figure 12 displays FPS values for different skeletonized models across 64 rainfall events, where FPS value of 1 representing an identical flood and a lower FPS values indicating a difference in flood predictions between the original and skeletonized models. As shown in Figure 12a, despite some variations, a higher level of model skeletonization is likely to induce a lower value of FPS, and hence a larger variation in flood property. For example, the mean of the FPS is 0.6 for ST1, but it becomes about 0.1 for ST6. Figure 12b shows how FPS varies as a function of changing rainfall intensities for ST6. While exhibiting moderate variations, a rainfall event with a larger intensity is more likely to lead to a higher FPS for the skeletonized model, indicating a low difference in flood property. This can be attributed to the fact that the flood property is dominated by rainfall intensity rather

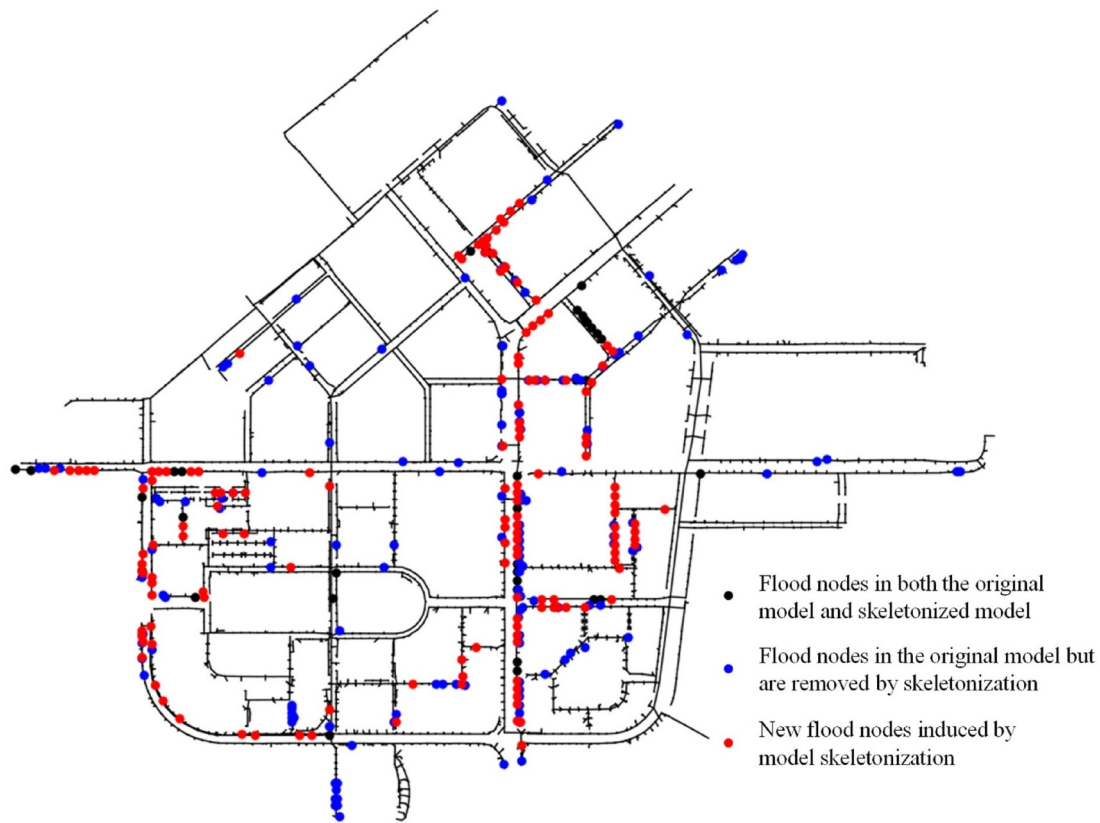


Figure 11. Flood distribution before and after model skeletonization for ST6 under rainfall R1 with $RP = 100$.

than the model skeletonization when an extreme rainfall event occurs. Similar observations can be made for other skeletonized models and hence they are not presented in the paper.

Results in this subsection suggest that model skeletonization can substantially affect the flood predictions of the UDNM, which has not been sufficiently recognized before. This is because previous studies focus only on the variation of hydrograph at the outfall caused by model skeletonization. Given that the main purpose of a UDNM is to predict the flood for a given rainfall event, the knowledge of model skeletonization on flood property of the UDNM is important. Therefore, the proposed quantitative metric, the FPS, is an important contribution of the present study.

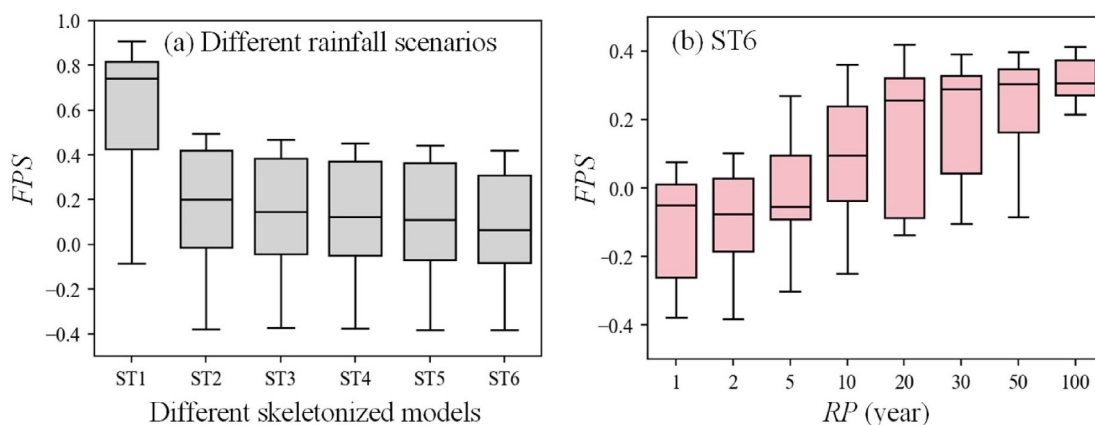


Figure 12. Flood property similarity distribution: (a) across different skeletonized models (b) across different return periods for model ST6.

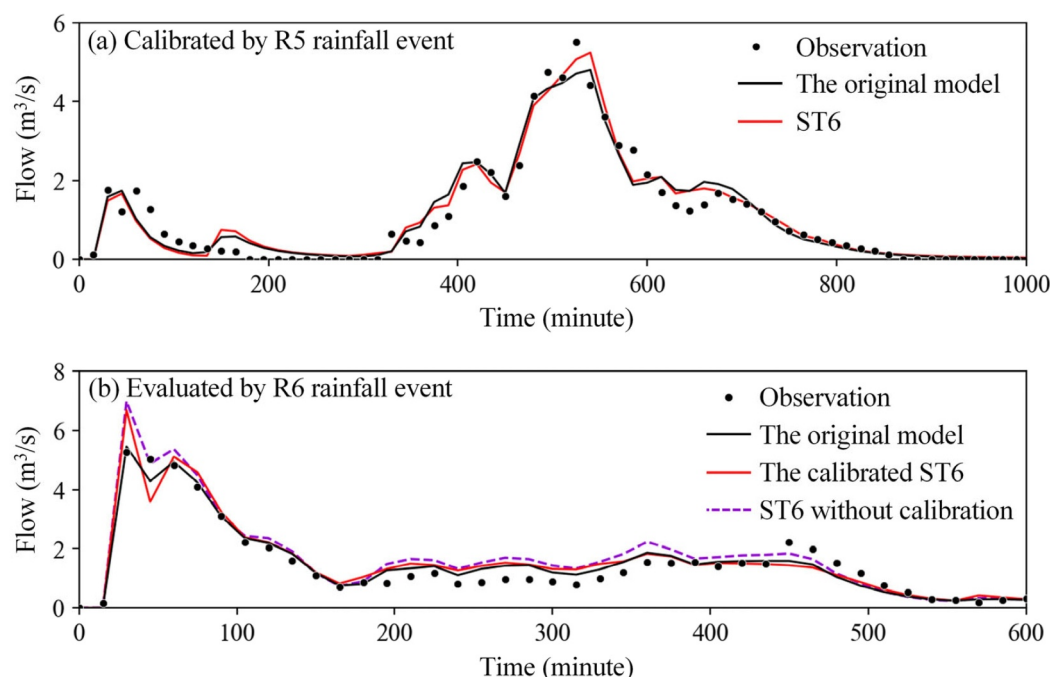


Figure 13. (a) Calibration and (b) evaluation results for flow at the outfall monitoring site based on the original model and the skeletonized model ST6, where ST6 without calibration indicates that it uses the same parameterization as the original model.

4.2. The Comparison of Model Performance Using Observations

The performance analysis of skeletonized models in the previous section considers the original model as a reference, based on the assumption that it can reliably replicate real-world conditions. In addition, in the present study, only the original (i.e., full/detailed) model is calibrated using the observations, where the skeletonized models are not re-calibrated. This rationale behind the above is that (a) the use of the observations and the re-calibration of skeletonized models can result in additional uncertainty to the analysis, as the performance impact can be simultaneously affected by model skeletonization, observation uncertainty and model calibration process (Huang et al., 2022); (b) the impact of model skeletonization on the simulation performance is typically ignored in engineering practice, or at least such impact is believed to be negligible, and hence the re-calibration of the skeletonized models is not often carried out in practice, and (c) a compensation method (Davidsen et al., 2017) is often used to mitigate the potential impacts caused by model skeletonization, instead of using the re-calibration on the skeletonized models, where the performance of different compensation methods have analyzed in this study.

Nonetheless, to further demonstrate the potential impacts caused by model skeletonization, the flow observations at the outfall are used to enable the model comparison. In addition, some of the observations are used to calibrate the original model and the skeletonized ST6 model respectively. More specifically, the real rainfall event R5 and its corresponding observation data (flow data at the outfall with a 15-min time resolution) are used for model calibration for the original model and the ST6, with results given in Figure 13a. It can be seen from this figure that both the original model and the skeletonized model ST6 are well calibrated with a KGE of 0.94. When using the test data for model evaluation (the R6 event, see Figure 13b), the original model is able to produce a high KGE of 0.95, but the KGE is reduced to 0.88 and 0.78 for the calibrated ST6 and the ST6 without calibration (use the same parameterization as the original model), respectively. Regarding the peak flow deviation percentages (PDP), the original model achieves a value of 3.5%, which is significantly lower than the calibrated ST6 (PDP = 27%) and the ST6 without calibration (PDP = 33%). Similar observations can be made for other rainfall events and skeletonized models. These results indicate that the skeletonization can result in large model errors based on both the full model simulations and the observations. Additionally, it is found that while re-calibration of skeletonized

Table 3
Computational Time (Seconds) and Time Savings Relative to the Original Model Based on a PC With Intel CPU i7-12700H

Model	Original	ST1	ST2	ST3	ST4	ST5	ST6
Time	1,253	660	502	476	411	367	347
Saving	/	-47.3%	-59.9%	-62.0%	-67.1%	-70.7%	-72.3%

models can improve model performance on the calibration event, its performance significantly deteriorates when using a new evaluation rainfall event. It is noted that the observations are only available at the outfall for the case study, and many other metrics in Table 1 (e.g., hydrograph similarity at monitoring locations or overall water depth in main pipes) cannot be assessed using observations. However, Figure 13 shows that the findings derived by using detailed model as the reference are overall similar to those using the outfall observations as the reference. Consequently, it is believed that the use of the detailed model output as the reference is effective for analysis. In addition, given that model calibration can result in additional uncertainty and the focus of the present study is to explore the impacts of model skeletonization, the skeletonized models are not re-calibrated in this study in order to isolate its potential impacts. However, results in Figure 13 also can show that the re-calibration of the skeletonized models does not significantly affect the findings made in Section 4.1.

4.3. Computational Time Savings

Time saving is one of the primary advantages of model skeletonization, as it enables more efficient repetitive model applications such as real-time prediction and uncertainty analysis. Table 3 shows the total computational times for various models (original and skeletonized) under different RPs across eight different rainfall patterns based on a PC with Intel CPU i7-12700H. As shown in this table, the skeletonized models achieved significant model run time reductions, ranging from 47.3% to 72.3% compared to the original model.

4.4. Performance Comparison of Different Compensation Methods

In this subsection, the TCM and the VCM are evaluated for their performance using the proposed evaluation framework, with results given in Figure 14. As shown in this figure, the VCM consistently produce a superior performance in mitigating the errors caused by model skeletonization. For example, the VCM is able to reduce the PDP value at the outfall from the mean of 13% (ST6) to about 8%, but the TCM even shows a slight increase in PDP value. Similarly, the VCM is capable of reducing the RMM values of the water depth, velocity and flow in main pipes substantially as shown in Figure 14, but the TCM shows no improvement.

Table 4 displays average KGE values across 64 rainfall scenarios at both the outfall and monitoring locations for ST6, the TCM and VCM. The VCM is consistently able to produce a larger KGE value relative to TCM. In addition, the VCM also shows a better ability to improve the flood simulation accuracy as measured by the metric FPS. Similar observations can be made for other skeletonized models and rainfall patterns. This indicates that the VCM can be a better option to mitigate the impacts caused by model skeletonization, which can be used for engineering practice.

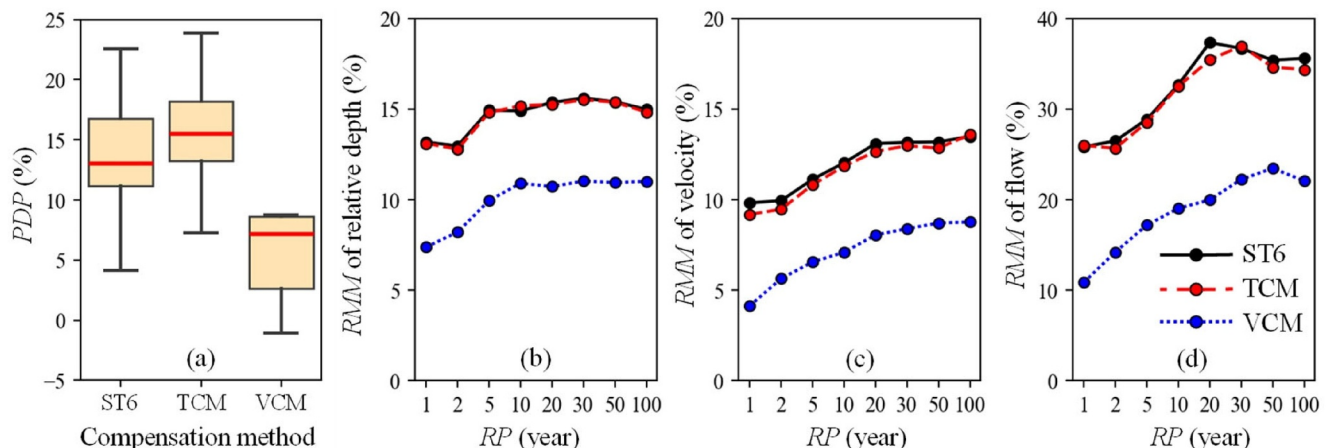


Figure 14. Impact of compensation methods on PDP, RMM of the relative depth, velocity, and flow in main pipes for ST6 under R1 across different return periods. TCM and VCM are the time-compensation method and the volume-compensation method respectively.

Table 4

Average Kling-Gupta Efficiency Values Across 64 Rainfall Scenarios for ST6

Node	Compensation methods		
	ST6	TCM	VCM
N1	0.89	0.89	0.93
N2	0.96	0.96	0.97
N3	0.87	0.88	0.88
N4	0.87	0.89	0.89
N5	0.90	0.90	0.93
N6	0.82	0.82	0.85
N7	0.95	0.95	0.97
N8	0.92	0.92	0.96
Outfall	0.93	0.93	0.97
Mean	0.90	0.90	0.93

5. Conclusions

UDNMs have been widely used in engineering practice for a range of different applications such as the flood prediction, system design and management. These models are often built using the detailed data stored in the corresponding GIS. Once this is done UDNMs are usually skeletonized to reduce the complexity of the initially built model with the aim to improve simulation efficiency when conducting the analysis based on many rainfall and other scenarios, and enhance the management efficiency during the daily work. However, there has been surprisingly few efforts made to understand how model skeletonization affects the model's behavior and accuracy, leading to an uncertainty in how these models perform for different applications including flood risk analysis.

This paper proposes a new comprehensive framework to systematically evaluate the impacts of model skeletonization on prediction accuracy. The metrics used in the framework include the magnitude and timing of peak flow at the outfall, the hydrograph similarity of the outfall flow, the hydrograph similarity at monitoring locations, the overall water depth, velocity and flow

in main pipes as well as the flood volume and flood ranges. In addition, two compensation methods are compared for their performance in mitigating the simulation errors induced by model skeletonization with the aid of the proposed framework. A real UDNM is used in this study to assess the utility of the proposed framework, where a total of six skeletonized models and 64 rainfall events are considered. The main findings and implications of this study are outlined as follows:

1. Model skeletonization can significantly affect the magnitude of peak flow at the outfall, with the maximum overestimation up to about 20%, but its impact is negligible on the timing of the peak flow of the outfall is minor. Model skeletonization may also underestimate the peak flow magnitude at the outfall for a few rainfall scenarios, which is likely caused by changes in the peak flow timing of the upstream pipes. Therefore, the impacts of skeletonization on model's performance can be complex as it can be influenced by the model structure and rainfall patterns.
2. The system hydraulics (water depth, velocity and flow) in the main pipes can also be affected by the model skeletonization. Generally, skeletonization can introduce an overall higher water depth, velocity and flow in main pipes, where the maximum flow increase can be up to 35%. This implies that the traditional studies that focus only on hydraulics or hydrograph at the system outfall location could be misleading as model skeletonization can also influence the hydraulics of pipes at the intermediate locations of the UDN.
3. Model skeletonization may significantly alter the flood volume and extent properties derived from the UDNM simulations and this aspect has been largely ignored in the past. Specifically, model skeletonization may produce a relatively large number of new flood locations which do not exist when using the original model and vice versa. Special attention should be paid to this when using the skeletonized model for flood risk analysis.
4. It is found that when flooding analysis for extreme rainfall events (e.g., RP = 100 years) is conducted, the flood property is dominated by rainfall intensity rather than the model skeletonization. In addition, simulation errors caused by model skeletonization are highly correlated to the PI of the rainfall events, where a rainfall event with a larger PI is more likely to result in larger simulation errors in the main pipes.
5. By using the new evaluation framework it was found that the VCM consistently shows improved performance when compared to the TCM for mitigating the errors induced by model skeletonization. However, the errors remaining after the use of TCM are still moderate, posing a danger for possible underestimation or overestimation of flood risk.

The above findings are important for engineering practice as they provide insights into how skeletonization affects the model's simulation performance. The above findings are conditioned on the real case study used in this work but the evaluation framework is transferable to other case studies where it can be easily applied.

Given that model skeletonization can cause large simulation errors as demonstrated in this study, skeletonization should be used when the potential benefits outweigh the downsides. Based on the findings of this study, if the model is used to simulate floods for higher RP events (e.g., RP = 100 years) or long-time rainfall events with relatively low PI, the use of skeletonized model is unlikely to introduce large errors (RMM is around 8%). For

other cases, the volume-compensation method (VCM, Davidsen et al., 2017) needs to be applied to the skeletonized models in order to mitigate the simulation errors, or improved skeletonization methods needs to be developed in future.

Data Availability Statement

The data are available at Ji (2024).

Acknowledgments

This work was supported by National Natural Science Foundation of China (NSFC, 52350130, 52261160379, 52179080, and 52000156), and the National Key R&D Program of China (Grants 2023YFC3208901, 2023YFC3208900).

References

- Achleitner, S., Möderl, M., & Rauch, W. (2007). CITY DRAIN © – An open source approach for simulation of integrated urban drainage systems. *Environmental Modelling & Software, Bayesian networks in water resource modelling and management*, 22(8), 1184–1195. <https://doi.org/10.1016/j.envsoft.2006.06.013>
- Burger, G., Sitzenfrie, R., Kleidorfer, M., & Rauch, W. (2014). Parallel flow routing in SWMM 5. *Environmental Modelling & Software*, 53, 27–34. <https://doi.org/10.1016/j.envsoft.2013.11.002>
- Cantone, J. P., & Schmidt, A. R. (2009). Potential dangers of simplifying combined sewer hydrologic/hydraulic models. *Journal of Hydrologic Engineering*, 14(6), 596–605. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000023](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000023)
- Davidsen, S., Löwe, R., Thrysoe, C., & Arnbjerg-Nielsen, K. (2017). Simplification of one-dimensional hydraulic networks by automated processes evaluated on 1D/2D deterministic flood models. *Journal of Hydroinformatics*, 19(5), 686–700. <https://doi.org/10.2166/hydro.2017.152>
- Farina, A., Di Nardo, A., Gargano, R., van der Werf, J. A., & Greco, R. (2023). A simplified approach for the hydrological simulation of urban drainage systems with SWMM. *Journal of Hydrology*, 623, 129757. <https://doi.org/10.1016/j.jhydrol.2023.129757>
- Fischer, A., Rouault, P., Kroll, S., Van Assel, J., & Pawlowsky-Reusing, E. (2009). Possibilities of sewer model simplifications. *Urban Water Journal*, 6, 457–470. <https://doi.org/10.1080/15730620903038453>
- Garzón, A., Kapelan, Z., Langeveld, J., & Taormina, R. (2022). Machine learning-based surrogate modeling for urban water networks: Review and future research directions. *Water Resources Research*, 58(5), e2021WR031808. <https://doi.org/10.1029/2021WR031808>
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Hellbach, C., Möderl, M., Sitzenfrie, R., & Rauch, W. (2012). Influence of network properties and model purpose on the level of skeletonization. In *World Environmental and Water Resources Congress 2011: Bearing Knowledge for Sustainability* (pp. 137–145). [https://doi.org/10.1061/41173\(414\)15](https://doi.org/10.1061/41173(414)15)
- Huang, Y., Zhang, J., Zheng, F., Jia, Y., Kapelan, Z., & Savic, D. (2022). Exploring the performance of ensemble smoothers to calibrate urban drainage models. *Water Resources Research*, 58(10), e2022WR032440. <https://doi.org/10.1029/2022WR032440>
- Huang, Y., Zheng, F., Duan, H.-F., Zhang, T., Guo, X., & Zhang, Q. (2019). Skeletonizing pipes in series within urban water distribution systems using a transient-based method. *Journal of Hydraulic Engineering*, 145(2), 04018084. [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0001560](https://doi.org/10.1061/(ASCE)HY.1943-7900.0001560)
- Ji, Y. (2024). Urban drainage network model: SL case [Dataset]. *Zenodo*. <https://doi.org/10.5281/zenodo.13920128>
- Kim, H. I., & Han, K. Y. (2020). Data-driven approach for the rapid simulation of urban flood prediction. *KSCSE Journal of Civil Engineering*, 24(6), 1932–1943. <https://doi.org/10.1007/s12205-020-1304-7>
- Kotrya, B. (2023). High-performance watershed delineation algorithm for GPU using CUDA and OpenMP. *Environmental Modelling & Software*, 160, 105613. <https://doi.org/10.1016/j.envsoft.2022.105613>
- Leitão, J. P., Simões, N. E., Maksimović, Č., Ferreira, F., Prodanović, D., Matos, J. S., & Sá Marques, A. (2010). Real-time forecasting urban drainage models: Full or simplified networks? *Water Science and Technology*, 62(9), 2106–2114. <https://doi.org/10.2166/wst.2010.382>
- Lin, R., Zheng, F., Savic, D., Zhang, Q., & Fang, X. (2020). Improving the effectiveness of multiobjective optimization design of urban drainage systems. *Water Resources Research*, 56(7), e2019WR026656. <https://doi.org/10.1029/2019WR026656>
- Mahmoodian, M., Carbajal, J. P., Bellos, V., Leopold, U., Schutz, G., & Clemens, F. (2018). A hybrid surrogate modelling strategy for simplification of detailed urban drainage simulators. *Water Resources Management*, 32(15), 5241–5256. <https://doi.org/10.1007/s11269-018-2157-4>
- Maier, H. R., Zheng, F., Gupta, H., Chen, J., Mai, J., Savic, D., et al. (2023). On how data are partitioned in model development and evaluation: Confronting the elephant in the room to enhance model generalization. *Environmental Modelling & Software*, 167, 105779. <https://doi.org/10.1016/j.envsoft.2023.105779>
- Meijer, D., Van Bijnen, M., Langeveld, J., Korving, H., Post, J., & Clemens, F. (2018). Identifying critical elements in sewer networks using graph-theory. *Water*, 10(2), 136. <https://doi.org/10.3390/w10020136>
- Niemi, T. J., Kokkonen, T., Sillanpää, N., Setälä, H., & Koivusalo, H. (2019). Automated urban rainfall–runoff model generation with detailed land cover and flow routing. *Journal of Hydrologic Engineering*, 24(5), 04019011. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001784](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001784)
- Nkwunonwo, U. C., Whitworth, M., & Baily, B. (2020). A review of the current status of flood modelling for urban flood risk management in the developing countries. *Scientific African*, 7, e00269. <https://doi.org/10.1016/j.sciaf.2020.e00269>
- Oh, J., & Bartos, M. (2023). Model predictive control of stormwater basins coupled with real-time data assimilation enhances flood and pollution control under uncertainty. *Water Research*, 235, 119825. <https://doi.org/10.1016/j.watres.2023.119825>
- Pedersen, A. N., Brink-Kjær, A., & Mikkelsen, P. S. (2022). All models are wrong, but are they useful? Assessing reliability across multiple sites to build trust in urban drainage modelling. *Hydrology and Earth System Sciences*, 26(22), 5879–5898. <https://doi.org/10.5194/hess-26-5879-2022>
- Radinja, M., Škerjanec, M., Šraj, M., Džeroski, S., Todorovski, L., & Atanasova, N. (2021). Automated modelling of urban runoff based on domain knowledge and equation discovery. *Journal of Hydrology*, 603, 127077. <https://doi.org/10.1016/j.jhydrol.2021.127077>
- Rossman, L. A. (2015). Storm water management model user's manual version 5.1.
- Sadler, J. M., Goodall, J. L., Behl, M., Morsy, M. M., Culver, T. B., & Bowes, B. D. (2019). Leveraging open source software and parallel computing for model predictive control of urban drainage systems using EPA-SWMM5. *Environmental Modelling & Software*, 120, 104484. <https://doi.org/10.1016/j.envsoft.2019.07.009>
- Thrysoe, C., Arnbjerg-Nielsen, K., & Borup, M. (2019). Identifying fit-for-purpose lumped surrogate models for large urban drainage systems using GLUE. *Journal of Hydrology*, 568, 517–533. <https://doi.org/10.1016/j.jhydrol.2018.11.005>

- van der Werf, J. A., Kapelan, Z., & Langeveld, J. G. (2023). The impact of blue-green infrastructure and urban area densification on the performance of real-time control of sewer networks. *Water Resources Research*, 59(6), e2022WR033591. <https://doi.org/10.1029/2022wr033591>
- Warsta, L., Niemi, T. J., Taka, M., Krebs, G., Haahti, K., Koivusalo, H., & Kokkonen, T. (2017). Development and application of an automated subcatchment generator for SWMM using open data. *Urban Water Journal*, 14(9), 954–963. <https://doi.org/10.1080/1573062X.2017.1325496>
- Xia, C., Wang, H., Zhang, A., & Zhang, W. (2018). A high-performance cellular automata model for urban simulation based on vectorization and parallel computing technology. *International Journal of Geographical Information Science*, 32(2), 399–424. <https://doi.org/10.1080/13658816.2017.1390118>
- Yin, D., Chen, Y., Jia, H., Wang, Q., Chen, Z., Xu, C., et al. (2021). Sponge city practice in China: A review of construction, assessment, operational and maintenance. *Journal of Cleaner Production*, 280, 124963. <https://doi.org/10.1016/j.jclepro.2020.124963>
- Zhang, S., & Pan, B. (2014). An urban storm-inundation simulation method based on GIS. *Journal of Hydrology*, 517, 260–268. <https://doi.org/10.1016/j.jhydrol.2014.05.044>
- Zheng, F., Chen, J., Ma, Y., Chen, Q., Maier, H. R., & Gupta, H. (2023). A Robust strategy to account for data sampling variability in the development of hydrological models. *Water Resources Research*, 59(3), e2022WR033703. <https://doi.org/10.1029/2022WR033703>
- Zheng, F., Westra, S., & Leonard, M. (2015). Opposing local precipitation extremes. *Nature Climate Change*, 5, 389–390. <https://doi.org/10.1038/nclimate2579>