RESEARCH

Centrality and shortest path length measures for the functional analysis of urban drainage networks

Julian D. Reyes-Silva[1*] , Jonatan Zischg[2], Christopher Klinkhamer[3], P. Suresh C. Rao[3], Robert Sitzenfrei[2] and Peter Krebs[1]

* Correspondence: julian_david.
reyes_silva@tu-dresden.de
[1]Department of Hydrosciences,
Institute for Urban Water
Management, TU Dresden,
Bergstraße 66, 01069 Dresden,
Germany
Full list of author information is
available at the end of the article

Abstract

The objective of this research is to evaluate whether complex dynamics of urban drainage networks (UDNs) can be expressed in terms of their structure, i.e. topological characteristics. The present study focuses on the application of topological measures for describing the transport and collection functions of UDNs, using eight subnetworks of the Dresden sewer network as study cases. All UDNs are considered as weighted directed graphs, where edge weights correspond to structural and hydraulic pipe characteristics which affect flow. Transport functions are evaluated in terms of travel time distributions (TTDs), under the hypothesis that frequency distributions of Single Destination Shortest Paths (SDSP) of nodes to the outlet had similar shapes than TTDs. Assessment of this hypothesis is done based on two-sample Kolmogorov-Smirnov tests and comparisons of statistical moments. Collection analysis, i.e. determination of flow paths, is done based on two approaches: (1) using Edge Betweenness Centrality (EBC), and (2) based on the number of SDSP going through an edge connecting a node to the outlet, referred as Paths. Hydrodynamic simulation results are used to validate the outcomes of graph analysis with actual flow behaviors. Results indicate that given an appropriate edge weighting factor, in this case Residence Time, SDSP has the potential to be used as an indicator for flow transport in UDNs. Moreover, both EBC and Paths values were highly correlated to average flows. The first approach, however, proved to be inadequate for estimating flows near the outlet but appropriate for identifying different paths in meshed systems, while the second approach lead to better results in branched networks. Further studies regarding the influence of UDNs layout are needed.

Keywords: Complex network, Flow, Wastewater, Topological measures, travel time

Introduction

Over the past several years, there has been an increasing use of graph theory for analyzing various natural, human and technological complex real-world systems. Within this approach, concepts and methodologies from graph theory, the study of mathematical structures formed by pairwise relations between objects (Elsner et al. 2009), are applied to understand the structural characteristics, similarities, patterns and scaling laws of different networks (Barabási and Albert 1999). Examples can be found in several fields,

such as social and communications sciences, medicine, epidemiology, meteorology, hydrology, transport and ecology, among others. Cases of network theory analysis and applications in real-world systems can be found for instance in Strogatz (2001) and da Fontoura Costa et al. (2011). Research of urban infrastructure systems have focused mainly on road networks, particularly on their evolution and common topological structures (Porta et al. 2006; Masucci et al. 2014), resilience (Wang 2015; Zhan et al. 2017) and robustness (Masucci and Molinero 2016).

Applications of complex network analysis on urban water networks, i.e. drinking water supply and drainage systems, have been however limited, due to lack of data or difficulty of accessing information. In the case of water supply networks, studies have focused on the usage of metrics derived from complex network analysis for the assessment of topological robustness and vulnerability (Yazdani and Jeffrey 2012; Agathokleous et al. 2017; Hwang and Lansey 2017; Zeng and Li 2017; Krueger et al. 2017; Yang et al. 2017; Nazempour et al. 2018; Ulusoy et al. 2018; Zischg et al. 2018). On the other hand, studies related to graph-theory based analysis of Urban Drainage Networks (UDNs) have focused mainly on the evolution and topological characteristics of both virtual (Ghosh et al. 2006; Möderl et al. 2009; Urich et al. 2010) and real sewer systems (Zischg et al. 2017; Krueger et al. 2017), and on the application of graph-theory based methodologies for determining critical elements in the network (e.g. Meijer et al. 2018). Nevertheless, advantages, limitations and insights on the UDNs' physical characteristics and dynamics that can be derived from the application of complex network analysis are still not fully understood.

Eventhough much research has been done regarding the complex and dynamic processes that take place in UDNs, there is little understanding on the role the structure (i.e. layout) of such networks plays for their function. Since the main objective of these systems is to collect and transport sanitary sewage (i.e. wastewater from households, industries, commercial and institutional buildings), and stormwater runoff produced during rain events to a fixed outlet, e.g. a wastewater treatment plant, two separate functions can be identified, one focused on collection and one on transport.

In this context, the present study aims to characterize these functional properties of UDNs in terms of topological measures. Underlying objective of such analysis is to determine whether graph metrics can be used as a good approximation or as surrogate variables of complex processes in such systems, and hence set a basis for the analysis of UDNs' dynamics in terms of their structure. Because of this, the functional analyses in this study are done under normal operational conditions, i.e. during dry weather conditions (no rain events). Therefore, only the collection and transport functions of sanitary sewage are analyzed. Stormwater is not considered here in order to avoid the potential abnormal behaviors in the system, such as flooding, pressurized flows or overflow events. Based on available information, 8 subnetworks from the sewer system of the city of Dresden, Germany, are analyzed. Since their different physical characteristics are within the typical ranges of urban subcatchments, the results obtained in this study are expected to be representative for UDNs in different geophysical contexts.

Materials and methods
Graph representation of UDNs and their properties
Given that UDNs are generally gravity driven systems that collect and convey wastewater into a specific outlet, they are considered and analyzed as Directed Acyclic Graphs (DAGs),

where the edges correspond to pipe sections, while the nodes are the junctions between them, usually represented as manholes. Furthermore, wastewater flow through the network strongly depends on several characteristics of pipe sections. Based on this, the influences of 3 structural properties, Length, Diameter and Slope; and 3 hydraulic factors, Velocity, Friction Losses and Residence Time, are considered. Those factors are included as edge weights for all network analysis. Since the scope of the study is to analyze the influence of structural properties on the function of UDNs, hydraulic factors are expressed in terms of structural properties of the pipes. As a result, the following weights are used:

- **W1**: Length of the pipe section, in meters [m].
- **W2**: 1/Diameter. Inverse values of this factor are used in order to account for preferential water pathways through pipes with larger diameters. It is measured in $[m^{-1}]$.
- **W3**: 1/Slope. Inverse percentage values are used to account preferential pathways through steeper pipes.
- **W4**: 1/Velocity, measured then in $m^{-1} \times s$. This hydraulic property is expressed in terms of structural characteristics, specifically as:

$$Velocity\ [m/s] \approx Diameter^{\frac{2}{3}} \times Slope^{\frac{1}{2}}$$

This relationship is obtained based on the Manning formula for water flow in open channels and partially filled pipes. Furthermore, inverse values are also used to account for preferential flow paths.

- **W5**: Friction Losses, expressed as the ratio between Length and Diameter. This factor is derived based on the Darcy-Weisbach equation for head loss in pipes.
- **W6**: Residence time, i.e. time water spends to flow through a pipe section, measured in seconds [s]. Since it corresponds to the ratio between pipe length and flow velocity, it is derived based on *W1* and *W4* as follows:

$$Residence\ Time\ [s] \approx \frac{Length}{Diameter^{\frac{2}{3}} \times Slope^{\frac{1}{2}}}$$

- **W7**: Unweighted edges, i.e. all edges have a weight of 1.

The different weighting factors are used for all analyses in order to assess their influence on the efficiency that topological measures may have for describing the transport and collection functions in UDNs. In other words, they are used to evaluate which structural property (Length, Diameter or Slope) or which combination of them (i.e. the hydraulic factors) is more appropriate to use as an edge weight, in order to better represent functional properties of UDNs in terms of topological measures. This allows to determine what is the minimum structural information required to understand the UDNs functions based on network analysis.

Functional analysis
Wastewater transport
Transport in drainage networks, such as rivers and sewers, is traditionally represented with the travel time distribution (TTD), i.e. the frequency distribution of times required

for a water volume to reach the outlet from any point of the network (McDonnell et al. 2010; Kaeseberg et al. 2018). Previous studies such as Seo and Schmidt (2014), have suggested that the frequency distribution of Euclidean distances along the network to its outlet, referred typically as the Width Function, has a similar shape as the TTD. With appropriate scaling, travel times can then be derived by using the Width Functions. In this context, it is hypothesized that geodesic distances, i.e. the minimum length of a shortest path connecting two nodes in a network (Bouttier et al. 2003), could be used in a similar manner. Since UDNs are considered as weighted directed graphs, such lengths correspond to the sum of edge weights. It is then assumed that the frequency distribution of graph distances between any node in the network and the outlet can be interpreted as a Width Function, and hence used to derive the TTDs in urban drainage systems. Geodesic distances are obtained using Dijkstra's algorithm (Dijkstra 1959) for Single Destination Shortest Paths (SDSP).

To test the hypothesis that travel times and SDSP to the network's outlet have similar shapes, i.e. similar distributions, two-sample Kolmogorov-Smirnov tests (Kolmogorv 1933) are applied to compare the cumulative frequencies of SDSP and TTDs for all different weights and for each network. Furthermore, the three first moments (mean, variance and skewness) are also analyzed. All distributions are normalized by their maximum values to avoid any influence of an unknown scaling factor.

Hydrodynamic simulations using the EPA Storm Water Management Model (Rossman 2015) are used to derive the TTDs of each network. This is done following a similar procedure as described in Kaeseberg et al. (2018), based on a 2-min Dirac pulse of a virtual tracer during the daily flow peak (around midday) in all the spatially distributed wastewater inflows. The resulting flow and concentration signals at each network's outlet are combined to obtain pollutographs and then normalized by the total tracer mass to obtain the TTDs.

Wastewater collection

The function of conveying and collecting water in drainage networks is represented by the presence of flow paths. In rivers, such pathways are generally determined based on the Horton-Strahler Order, a numerical measure of the networks branching complexity (Strahler 1957). Several studies have applied this ordering scheme in real and virtual sewer systems (Cantone and Schmidt 2011a; Cantone and Schmidt 2011b; Sitzenfrei et al. 2013; Yang et al. 2017). However, this measure was developed for branching networks (rivers), where only one main flow path exists. Determining the drainge pathways of UDNs based on Horton-Strahler ordering can be challenging and complex since their layout is more heterogeneous, mixing both grid and tree-like configurations.

Two graph-based approaches for determining the wastewater flow pathways in UDNs are tested. In the first case, wastewater flow is considered as the "information" of the network, and therefore it was assumed that edges where more information is conveyed and collected, should correspond to links with higher Edge Betweenness Centrality (EBC), i.e. edges with a higher number of shortest paths connecting any pair of nodes ("all-to-all") in the network (Girvan and Newman 2002). On the other hand, in the second approach it is hypothesized that flow quantities in each edge are directly proportional to the number of SDSP passing through such edge, here simply referred as *Paths*, connecting any node to the outlet ("all-to-one"). In order to test both approaches daily average wastewater flows (Q_{dw}) of each edge are compared with the corresponding EBC values and number of *Paths* for all study cases and

using the different weights. Comparisons are made based on Pearson's correlation coefficients. Note that the Q_{dw} values are calculated based on 7-days hydrodynamic simulations during dry-weather, i.e. in periods with no rain events, and modelled as steady state flow conditions.

Case studies

Availability of structural data of UDNs is often limited due to security reasons or privatization of the information. Moreover, adequate data for the proper development of hydrodynamic models, e.g. flow measurements in the system, is often missing. Based on the available information, 8 subnetworks of the Dresden sewer system are analyzed. All networks are characterized by one destination node, representing the outfall to a central collection system that is connected to the wastewater treatment plant (WWTP) of Dresden. As can be seen in Table 1, not only their size, represented by the number of nodes and the total sewer length, but also their corresponding drainage area and mean daily wastewater flow (Q_{dw}) covers typical ranges of urban subcatchments. Furthermore, their layout is also heterogeneous, as can be seen from Fig. 1 and their Meshness values, i.e. a measure to quantify how close an UDN structure resembles a tree-graph layout (0%) or a grid-like pattern (100%). The values are derived based on the methodology proposed by Reyes-Silva et al. (2019.). For example, the networks N3 and N5 have a layout more similar to a branching system, i.e. low Meshness, while N4 has a more grid-like configuration (higher Meshness degree).

Moreover, calibrated hydrologic-hydraulic models are available for these subnetworks, allowing to validate results from the complex network analysis in terms of actual flow behavior in the drainage networks. Additional information regarding these reference systems can be found in Kaeseberg et al. (2018). Although the hydrodynamic models are only used here to obtain daily average dry-weather flows (Q_{dw}), their availability offers the possibility for future analysis regarding the role of structure on transport and collections functions under more complex and dynamic processes, e.g. influence of daily flow variations or heterogeneous rain events.
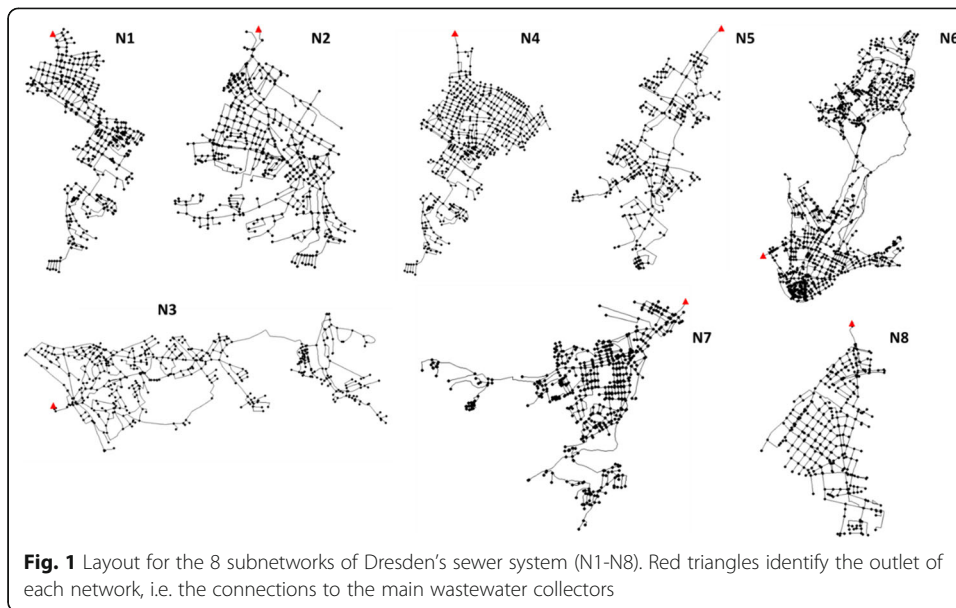
Results and discussion

Wastewater transport

The different cumulative frequencies of the calculated shortest paths and simulated travel times can be seen in Fig. 2. It is found that in the networks N2, N5 and N8, TTDs are significantly similar to SDSP distributions when using Residence Time (W6) as weighting factor (significance level α = 0.01). Similar results are obtained for N2 and N5 when using Friction Losses (W5), and for N2 and N4 when using the inversed pipe diameter (W2) as edge weights. Furthermore, significant results using W1 (pipe length) were
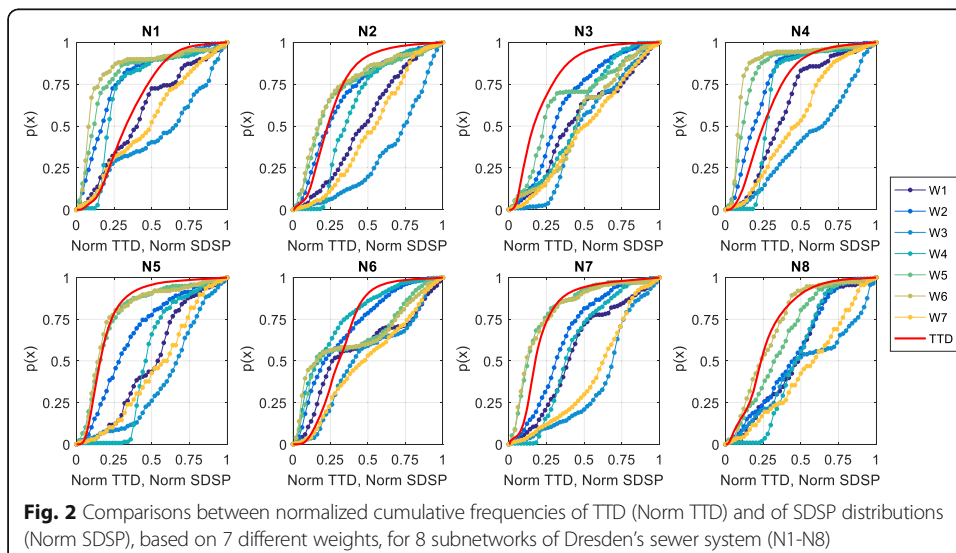
Table 1 Physical characteristics for the 8 subnetworks of Dresden's sewer system (N1-N8)

Network	N1	N2	N3	N4	N5	N6	N7	N8
Nodes	491	563	575	656	357	1235	621	391
Edges	551	629	608	797	384	1352	654	427
Area [km^2]	3.54	5.05	3.73	5.42	3.64	9.21	5.28	2.96
Sewer Length [km]	80.13	114.66	76.46	109.07	61.92	184.82	102.40	59.16
Qdw [L/s]	91.58	87.50	67.52	142.13	58.88	89.33	143.02	45.71
Mean Slope [%]	0.86	1.07	2.11	0.86	1.08	0.71	2.01	1.25
Meshness [%]	31.15	29.03	16.88	45.47	20.99	27.08	17.13	26.55

Fig. 1 Layout for the 8 subnetworks of Dresden's sewer system (N1-N8). Red triangles identify the outlet of each network, i.e. the connections to the main wastewater collectors

reported only for N1. For other weight and network combinations, there is no statistical evidence to suggest that TTD and SDSP have similar probability distributions.

In general, all analyzed networks SDSP distributions show higher mean values than the corresponding TTDs. Nonetheless, when using W5 and W6 as edge weighting factors the mean values of both distributions are relatively similar for the networks N2, N5, N6, N7 and N8. The variances of the SDSP distributions are always higher than those of the TTDs for all networks, regardless of the edge weight used. Regrading skewness, results indicate that for all analyzed networks TTDs are highly positively skewed (i.e. with values greater than 1), except for N1 and N6 which are found to be moderately skewed (0.68 and 0.96 respectively). The SDSP distributions similarly show highly and/or moderate right skewed behavior is observed for almost all edge weights, except for W3 and W7, which yield negatively skewed distributions. Using W6 leads to better

Fig. 2 Comparisons between normalized cumulative frequencies of TTD (Norm TTD) and of SDSP distributions (Norm SDSP), based on 7 different weights, for 8 subnetworks of Dresden's sewer system (N1-N8)
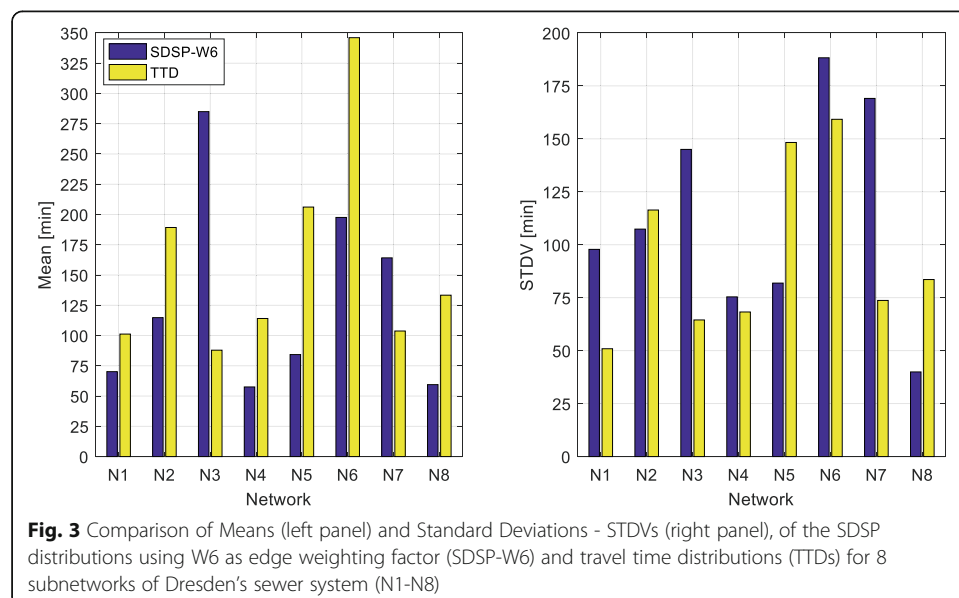
results, i.e. a similar skewness between travel times and SDSP distributions, for networks N2, N5, N7 and N8. For detailed information regarding the results of the statistical moments for all SDSP distributions and TTDs in each network see Additional file 1: Data S1.

Based on the previous results, Residence Time (W6) is identified as the most appropriate edge weighting factor for the description of TTDs in terms of SDSP distributions. This result suggests that the minimum structural information required to understand the UDNs transport function based on network analysis methodologies (i.e. distribution of SDSP to the outlet), is the slope, diameter and length of each pipe section.

Note that these results are obtained based on normalized distributions, therefore the means and standard deviations of the real SDSP distributions using W6 (referred to as SDSP-W6) and TTDs for all networks are compared to better understand the validity of the previous results. As shown in Fig. 3, the mean values of SDSP-W6 are in general lower than in the TTDs, except for N3 and N7, where the graph-based approach results in mean values that are 3.3 and 1.6 times higher than those of the TTDs, respectively. The standard deviations of the SDSP distributions are higher for more than half of the analyzed networks. Direct consequences of these results are that using SDSP-W6 to analyze the transport function will lead to faster and more disperse responses in comparison to the calculated travel times from the hydrodynamic simulations. Lower mean values indicate that water takes less time on average to reach the outlet from any point of the network, and higher standard deviations, thus higher variances, lead to a wider spread of the data. Extreme values, i.e. longer and shorter travel times, are then overestimated.

Nevertheless, SDSP-W6 can be used as a potential indicator for flow transport in UDNs. In fact, fast responses are associated to high flow conditions in the systems, e.g. during precipitation events when high amounts of stormwater enter the system. Based on this, results may indicate that SDSP-W6 could have the potential of representing travel times during rain events, and not during dry-weather conditions, thus explaining the difference between the moments of SDSP-W6 and TTDs in this study. A potential reason for this relies on the approach used to calculate the velocities (W4), and hence

Fig. 3 Comparison of Means (left panel) and Standard Deviations - STDVs (right panel), of the SDSP distributions using W6 as edge weighting factor (SDSP-W6) and travel time distributions (TTDs) for 8 subnetworks of Dresden's sewer system (N1-N8)

the residence times (W6) in the edges. During dry-weather conditions, wastewater level in pipes is considerably lower than the maximum available depth (i.e. diameter), even during daily peak flow conditions. Assuming that the velocities in each edge are proportional to their diameter (see section 2.2) and not a fraction of it, leads to the estimation of discharge rates during full pipe flow conditions, i.e. at the maximum flow. It is then hypothesized that, given the appropriate scaling factor, SDSP-W6 could be used to derive the different travel times during different discharge conditions. For this purpose, further studies regarding, for example, the impacts of diurnal flow variations or different rain intensities are needed. Some rain events might lead to overflow events or pressurized conditions in some pipes which may heavily affect flow velocities and hence travel times. In order to have a robust representation of the transport function it is necessary then to further analyze how these events affect travel times and furthermore how can this be represented from a network perspective.

Although SDSP-W6 shows promising results, the approach for calculating the residence time requires further improvements to have a robust indicator of travel times in UDNs. The fact that using W6 in networks N1, N3 and N5 results in higher frequencies of smaller travel times, while W1 (Length) reported better outcomes (see Fig. 2), suggests that the current approach overestimates the velocities in the systems. In fact, W6 is obtained based on the ratio between W1 and W4, hence smaller travel times are caused by higher velocities. The networks N1, N3 and N5 are particularly sensitive to this, since they are flat networks, i.e. the mean slopes are smaller than 1% (see Table 1) and are therefore associated with lower flow velocities during dry-weather conditions. On the other hand, the current study does not consider the presence of offsets, i.e. when the connection between two or more pipes occur at different elevations. Those can be decisive factors for determining actual flow paths in the network and hence for the estimation of travel times. Including the presence of offsets, either as an additional edge weight or a condition for connectivity, could improve the results obtained using SDSP-W6.

Wastewater collection
EBC approach
A positive correlation between EBC and Q_{dw} is found for all networks, and for all different edge weights. High correlation values, i.e. bigger than 0.7, are reported for almost all networks, the only exception is N5 with correlation coefficients smaller than 0.61. For detailed information see Additional file 1: Data S2. Furthermore, no strong influence of the different edge weighting factors on the results is identified. This result suggests that to understand the UDNs collection function based on network analysis methodologies, in this case EBC, no structural information is required. In other words, edge weights are not needed for this type of analysis.

A closer analysis of the relationship between EBC and Q_{dw} reveals hysteretic behavior between the two variables, regardless of the edge weighting factor used. An example of this can be seen in Fig. 4, which illustrates the relationship between EBC using Residence Time as a weight (EBC-RT) and the normalized daily average dry-weather flows (Qdw_n) for all networks. Flow values are normalized by the maximum average discharge in the system, i.e. at flow at the outlet, in order to facilitate comparisons among the networks. As can be seen in Fig. 4, in all analyzed UDNs flow initially increases proportionally to EBC, however, after a certain threshold higher discharges are related to lower EBC values. The reason for this is that EBC quantifies the number of possible shortest paths passing through an edge connecting any pair of nodes. Hence for links closer to the outlet, i.e.

Fig. 4 Relationship between Qdw_n and EBC-RT, for 8 subnetworks of Dresden's sewer system (N1-N8)
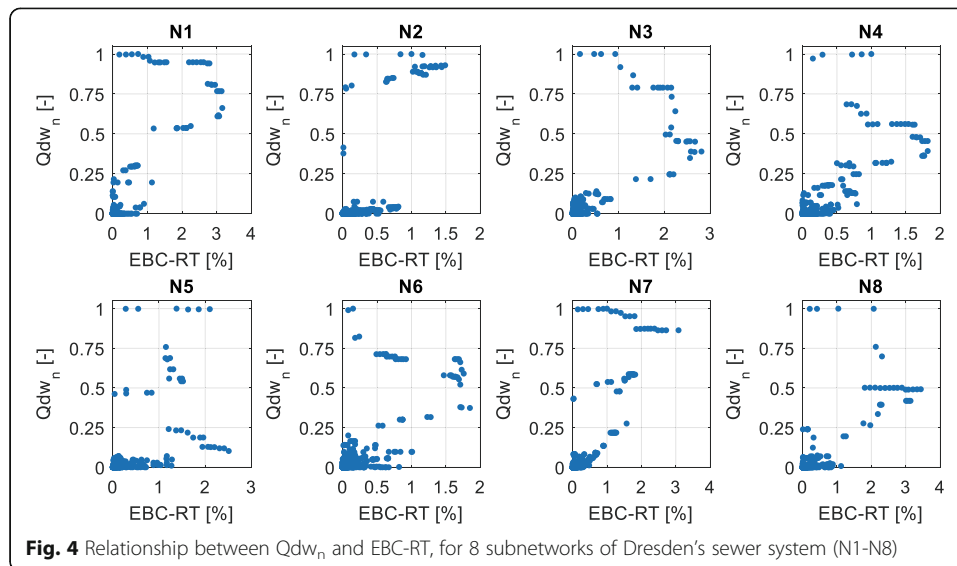
pipes with higher flows, the EBC is lower since the number of possible shortest paths decreases. This is due to the fact that, for historical and topographical reasons, destination nodes such as WWTPs are usually located on the outskirts of cities and not in the topological center, which leads to an inverse relationship between flow and EBC.

The use of EBC for analyzing wastewater flow accumulation in UDNs is therefore not recommended, particularly for pipe sections close to the network's outlet. Nevertheless, it is hypothesized that this topological measure could have other applications in terms of flow accumulation analysis in UDNs. For a given network, edges with higher EBC values can be interpreted as "bridges" between different components of the system (Lu and Zhang 2013). In the case of UDNs, pipe sections with higher EBC values would then correspond to locations where different subnetworks are connected to each other or to the main drainage system. Identifying such critical pipes could, for example, provide the basis for vulnerability and robustness analysis in terms of the disconnection from the WWTP (outlet), and hence the loss of efficiency of wastewater collection and treatment.

Paths approach

Figure 5 and Additional file 1: Data S2 present the results of the paths approach, considering the shortest paths from all nodes to one destination node (all-to-one). In comparison with the previous EBC approach (all-to-all), a more linear tendency and no hysteretic behavior are identified, leading also to higher correlation coefficients, on average approximately 0.91 for all networks and for all different edge weighting factors, which again do not show a strong influence on the results. As for the EBC approach, this result suggests that edge weights are not necessary for this type of analysis.

The influence of network layout, however, becomes evident. In networks were only one possible pathway from each node to the outlet exist, i.e. in branched networks such as N3, a single linear relationship between flow and the number of shortest paths crossing the edge is observed. On the contrary, in more meshed systems, e.g. N2 or N4, the presence of multiple flow paths connecting a node to the outlet leads to more complex relationships between paths and flow. These results suggest that although the current graph-based approach is able to approximate the flow pattern in UDNs quite accurately, further analysis
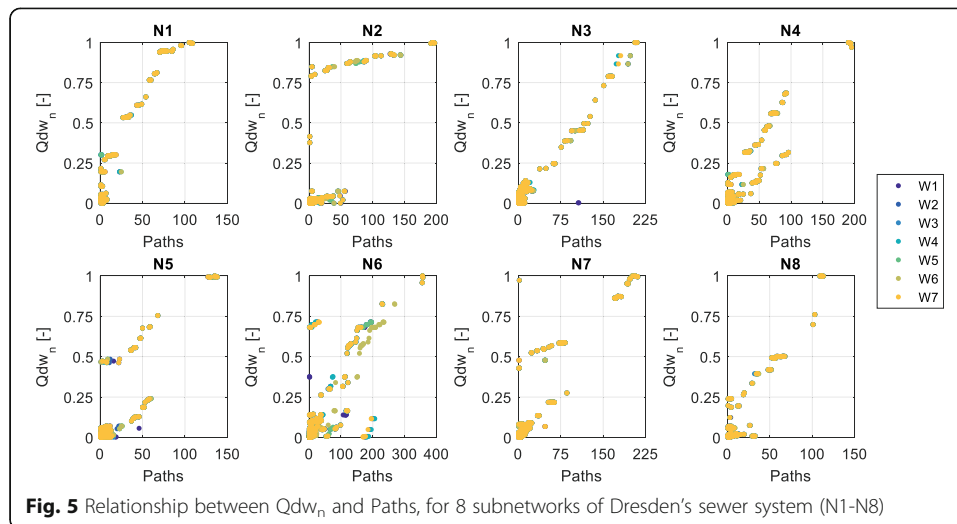
Fig. 5 Relationship between Qdw_n and Paths, for 8 subnetworks of Dresden's sewer system (N1-N8)

of the influence of node connectivity on wastewater quantities are needed. This includes, for example, the consideration of the spatial heterogeneity of dry-weather inflows.

Furthermore, although the current approach focused on the determination of flow paths during dry-weather conditions, the results may also apply, to a certain extent, for the analysis of the flow accumulation function during rain events, i.e. during wet-weather conditions. Sewer systems are generally designed to collect the stormwater produced by recurrent rain events, i.e. events with intensities and durations with a probability of occurrence of typically once every 2 to 5 years (DWA-A118 2006). During rain events with small and medium intensities, the flow regime in the system does not significantly change then in comparison to the dry-weather conditions. In these conditions, the results obtained in the present study can be applied to analyze the collection function during wet-weather conditions. However, there are limitations regarding further extrapolation of these results when analyzing the behavior under intense rainfall events. The large volumes of stormwater entering the system can cause either backflow and/or the activation of combined sewer overflows due to pipe overloading, thus leading to the generation of new flow paths. Although these new water pathways are temporarily, they can have a strong impact on the wastewater quantities in the system. It is then recommended that the analysis of UDNs flow accumulation function during wet-weather conditions using the current graph-based approach should consider the influence of potential new pathways during intense rain events.

Applications

This study focuses on multiple urban drainage subnetworks with contributing areas smaller than $10 \, km^2$. For such small UDNs, local structural differences (like the ones analyzed here) play a key role on the development of complex and dynamic processes that occur during wet-weather conditions, e.g. pipe surcharge, manhole flooding or the activation of combined sewer overflows due to pipe overloading. Such phenomena may have different adverse impacts not only in terms of public health and environmental pollution, but also with regard to structural damages of the drainage system.

In this context, the proposed approaches have the potential to be used for analyzing the influence of subnetwork variability on the occurrence of such events, and hence identify relevant components or strategies for the reduction of adverse effects. Furthermore, the

presented methodologies can also be used for the functional analysis of large scale UDNs, i.e. at city level. It should be noted, however, that with increasing size, local differences may play a smaller role, as flows toward the outlet (e.g. WWTP) are aggregated.

Summary and conclusions

The current study focused on the application of topological measures for describing the main functional properties of urban drainage networks (UDNs), specifically wastewater collection and transport, during normal operational conditions, i.e. during dry-weather conditions. As a study case, 8 subnetworks of the Dresden sewer system were used. The underlying objective of this research was to evaluate whether complex dynamics of UDNs can be expressed in terms of their structure, i.e. topological characteristics. All analyzed UDNs were considered as Directed Acyclic Graphs. Furthermore, 3 structural and 3 hydraulic pipe properties influencing wastewater flow were included as edge weighting factors. This was done to assess which structural information is required in order to analyze functional properties of UDNs from a complex network perspective. Results from 7-days hydrodynamic simulations during dry-weather conditions (no rain events) for each UDN were used to compare the outcomes of the graph analysis with actual behaviors in the systems.

Results indicate that with an appropriate edge weighting factor, in this case Residence Time (W6), the topological measure of Single Destination Shortest Paths (SDSP) to the network's outlet has the potential to be used as an indicator for flow transport in UDNs. For half of the networks, using W6 as edge weighting factor, the normalized SDSP frequency distributions were significantly similar to the normalized travel time distributions, based on two-sample Kolmogorov-Smirnov tests, but with different statistical moments. Consequently, the network responses (i.e. travel times) were found to be faster and more disperse than the ones obtained from hydrodynamic simulations. Based on these results, it is then hypothesized that the current graph-based approach also represents travel times during high flow conditions, e.g. during rain events, and therefore an appropriate scaling of the edge weights is needed to better represent the dry-weather conditions. This, however, has yet to be tested. Furthermore, improvements regarding the estimation of the residence time as the edge weighting factor, such as a more accurate estimation of flow velocities and integrating the presence of offsets, are required to obtain a robust and more realistic representation of travel times.

Regarding the functional analysis of flow accumulation, i.e. the determination of flow paths in the networks, two approaches were tested: one based on Edge Betweenness Centrality (EBC), and one based on the number of SDSP going through an edge connecting a node to the outlet, referred here as *Paths*. Results of the first approach suggest that although relatively high correlations between EBC and wastewater flow values where found, this topological metric is not a good indicator of wastewater flow quantities in UDNs, particularly for pipe sections near the network's outlet. On the contrary, results of the second approach showed a clear linear trend between average flows and the *Paths*, with correlation coefficients higher than 0.8 for all networks. While the correlations were highest for mostly branched network topologies, more than one linear relationship was identified in networks with more grid-like layout, attributed to the presence of multiple flow pathways. In this context, the EBC approach seemed to yield a better fit for identifying the different wastewater paths in meshed systems, since it considers all possible shortest paths passing through an edge. On the other hand, the

second approach leads to better results in branched networks and for pipe section near the system's outlet. Further studies regarding the influence of UDNs layout are needed, including the spatial heterogeneity of dry-weather inflow, changing shortest path preferences during load variation, or the activation of combined sewer overflows due to pipe overloading during heavy rainfall events.

The complex dynamic processes related to transport and accumulation of wastewater flow in UDNs, and the relationship to their network topology is still not fully understood and investigated. In this context, the results obtained in this study are expected to build a basis for future analysis, and to illustrate the potential application of graph-theory based methodologies in the field of urban drainage networks. The main advantage of such approaches is that they rely only on the structural data of the network, i.e. connectivity and pipe structural characteristics, thus reducing the requirement for comprehensive data traditionally needed for functional analysis of UDNs. Results from graph-theory approaches could serve as support for better design, management and operational actions of UDNs, for example, identifying critical locations in the system for the placement of storage units and overflow structures, or improving the layout to reduce the occurrence of unwanted discharges into waterbodies. Furthermore, results could also serve as a basis for the development of a new structural resiliency analysis based mainly on the UDN pattern configuration.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1007/s41109-019-0247-8.

Additional file 1 Data S1. Statistical analysis of all normalized SDSP distributions and TTD for 8 networks. **Data S2.** Correlation analysis between EBC and Avg-Qdw.

Abbreviations
DAG: Directed Acyclic Graph; EBC: Edge Betweenness Centrality; Qdw: Daily Average Dry-Weather Flow; Qdw_n: Normalized Daily Average Dry-Weather Flow; SDSP: Single Destination Shortest Path; SDSP-W6: Single Destination Shortest Path distribution using W6 as edge weight; TTD: Travel Time Distribution; UDN: Urban Drainage Network; WWTP: Wastewater Treatment Plant

Authors' contributions
PK coordinated the research; JDRS, JZ and RS conceived the study; JDRS performed the numerical studies and wrote the initial draft; CK and PSCR contributed to data analysis and discussion. All authors contributed to the writing of the final manuscript and approved the last version.

Availability of data and materials
The datasets generated during and/or analysed during the current study are not publicly available since they contain sensitive infrastructure information but are available from the corresponding author on reasonable request.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Hydrosciences, Institute for Urban Water Management, TU Dresden, Bergstraße 66, 01069 Dresden, Germany. [2]Unit of Environmental Engineering, Department of Infrastructure Engineering, University of Innsbruck, Technikerstraße 13, 6020 Innsbruck, Austria. [3]Lyles School of Civil Engineering & Agronomy Department, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN 47907, United States & KERAMIDA Inc., 401 N College Ave, Indianapolis, IN, USA.

References
Agathokleous A, Christodoulou C, Christodoulou SE (2017) Topological robustness and vulnerability assessment of water distribution networks. Water Resour Manag 31(12):4007–4021
Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286:509–512
Bouttier J, Di Franceso P, Guitter E (2003) Geodesic distance in planar graphs. Nucl Phys B 663(3):535–567
Cantone J, Schmidt A (2011a) Improved understanding and prediction of the hydrologic responseof highly urbanized catchments through developmentof the Illinois urban hydrologic model. Water Resour Res 47:W08538
Cantone J, Schmidt A (2011b) Dispersion mechanisms and the effect of parameter uncertaintyon hydrologic response in urban catchments. Water Resour Res 47:W05503
da Fontoura Costa L, Oliverira ON Jr, Travieso G, Aparecido Rodrigues F, Villas Boas PR, Antiqueira L et al (2011, June) Analyzing and modeling real-world phenomena with complex networks: a survey of applications. Adv Phys 60(3):329–412
Dijkstra EW (1959) A note on two problems in Connexion with graphs. Numer Math 1(1):269–271
DWA-A118 (2006) Arbeitsblatt DWA-A 118 Hydraulische Bemessung und Nachweis Hydraulische Bemessung und Nachweis. DWA Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V, Hennef
Elsner JB, Jagger TH, Fogarty EA (2009) Visibility network pf United States hurrucanes. Geophys Res Lett 36:Ll6702
Ghosh I, Hellweger FL, Fritch TG (2006) Fractal generation of artificial sewer networks for hydrologic simulation. Proc., ESRI international user conference. ESRI, San Diego
Girvan M, Newman MJ (2002) Community structure in social and biological networks. PNAS 99(12):7821–7826
Hwang H, Lansey K (2017) Water distribution system classification using Sysrem characteristics and graph-theory metrics. J Water Resour Plan Manag 134(12):041071071
Kaeseberg T, Kaeseberg M, Zhang J, Jawitz JW, Krebs P, Rao PS (2018) The nexus of inhabitants and impervious surfaces at city scale-wastewater and stormwater travel time distributions and an approach to calibrate diurnal variations. Urban Water J 15(6):576–583
Kolmogorv A (1933) Sulla determinazione empirica di una legge di distribuzione. G Ist Ital Attuari 4:83–91
Krueger E, Klinkhamer C, Urich C, Zhan X, Rao PC (2017) Generic patterns in the evolution of urban water networks: evidence from a large Asian city. Phys Rev E 95:032312
Lu L, Zhang M (2013) Edge Betweenness Centrality. In: Dubitzky W, Wolkenhauer O, Cho K, Yokota H (eds) Encycopledia of Systems Biology. Springer, New York
Masucci AP, Molinero C (2016) Robustness and closeness centrality for self-organized and planned cities. Eur Phys J B 89:53
Masucci AP, Stanilov K, Batty M (2014) Exploring the evolution of London's street network in the information space: A dual approach. h. Phys Rev E 89(1):012805
McDonnell JJ, McGuire K, Aggarwal P, Beven KJ, Biondi D, Destouni G et al (2010) How old is streamwater? Open questions in catchment transittime conceptualization, modelling and analysis. Hydrol Process 24(12):1745–1754
Meijer D, van Bijnen M, Langeveld J, Korving H, Post J, Clemens F (2018) Identifying critical elements in sewer networks using graph theory. Water 10:136
Möderl M, Butler D, Rauch W (2009) A stochastic approach for automatic generation of urban drainage systems. Water Sci Technol 59(6):1137–1143
Nazempour R, Monfraed M, Zio E (2018) A complex network theory approach for optimizing contamination warning sensor location in water distribution networks. Int J Dis Risk Reduction 30(B):225–234
Porta S, Crucitti P, Latora V (2006) The network analysis of urban streets: a dual approach. Phys A 369(2):853–866
Reyes-Silva JD, Helm B, Krebs P (2019) Meshness of sewer networks and its implications for flooding occurrence. In: 9th International Conference on Sewer Processes and Networks (SPN9), Aalborg
Rossman LA (2015) Storm water management model, user's manual, version 5.1. U.S. EPA, Cincinnati
Seo Y, Schmidt AR (2014) Applicat ion of Gibbs' model to urban drainage networks: acase study in southwestern Chicago, USA. Hydrol Process 28:1148–1158
Sitzenfrei R, Urich C, Möderl M, Rauch W (2013) Assessing the efficiency of different CSO positions. Water Sci Technol 67(7): 1574–1580
Strahler AN (1957) Quantitative analysis of watershed geomorphology. Trans Am Geophys Union 38(6):913–920
Strogatz SH (2001) Exploring complex networks. Nature 410:268–276
Ulusoy A-J, Stoianov I, Chazerain A (2018) Hydraulically informed graph theoretic measure of link criticality for the resilience analysis of water distribution networks. Appl Netw Sci 3(1):31
Urich C, Sitzenfrei R, Möderl M, Rauch W (2010) An agent-based approach for generating virtual sewer systems. Water Sci Technol 62(5):1090–1097
Wang J (2015) Resilience of self-organised and top-down planned cities-a case study on London and Beijing street networks. PLoS One 10(12):e0141736
Yang S, Paik K, McGrath GS, Urich C, Krueger E, Kumar P, Rao P (2017) Functional topology of evolving urban Drianage networks. Water Resour Res 53:8966–8979
Yazdani A, Jeffrey P (2012) Applying network theory to quantify the redundancy and structural robustness of water distribution systems. J Water Resour Plan Manag 138(2):153–161
Zeng F, Li X (2017) Modeling complexity in engineered infrastructure system: water distribution network as an example. Chaos 27(2):023105

Zhan X, Ukkusuri S, Rao P (2017) Dynamics of functional failures and recovery in complex road networks. Phys Rev E 96:
    052301
Zischg J, Klinkhamer C, Zhan X, Krueger E, Ukkusuri S, Rao PC et al (2017) Evolution of complex network topologies in urban
    water infrastructure. World Environ Water Resour Congr 2017:648–659
Zischg J, Rauch W, Sitzenfrei R (2018) Morphogenesis of urban water distribution networks: a spatiotemporal planning
    approach for cost-effective and Repliable supply. Entropy 20(9):708

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.