Natural hazard risk of complex systems – the whole is more than the sum of its parts: II. A pilot study in Mexico City

Marcello Arosio[1], Mario L.V. Martina[1], and Rui Figueiredo[1]

[1]Scuola Universitaria Superiore IUSS Pavia, Palazzo del Broletto - Piazza della Vittoria 15, 27100 Pavia (Italy)

5   *Correspondence to:* Marcello Arosio (marcello.arosio@iusspavia.it)

Abstract: Assessing the risk of complex systems to natural hazards is an important and challenging problem. In today's intricate socio-technological world, characterized by strong urbanization and technological trends, the connections, interdependencies and interactions between exposed elements are crucial. These complex relations call for a paradigm shift in collective risk assessments, from a reductionist approach to a holistic one. Most commonly, the risk of a system

10   is estimated through a reductionist approach, based on the sum of the risk of its elements individually. In contrast, a holistic approach considers the whole system as a unique entity of interconnected elements, where those connections are taken into account in order to more thoroughly assess risk. To support this paradigm shift, this paper proposes a new holistic approach to assess the risk in complex systems based on Graph Theory. The paper is organized in two parts: part I describes the proposed approach, and part II presents an application to a pilot study in Mexico City. The choice of

15   Mexico City as the study case allows demonstrating the importance of modelling connections and interdependencies to assess risk in a complex urban environment which, in this case, is characterized by increasing risk of urban flooding due to soil subsidence and the presumable failure of the drainage system. In the application, the complexity of Mexico City is depicted by modelling certain selected typologies of elements (i.e. population, fire stations, hospitals, fuel stations, schools and crossroads) and assuming certain rules of connection between them. The relevant graph properties are

20   computed and interpreted from a natural hazard risk perspective, and a simple hazard scenario is then integrated in the analysis. This study highlights both the potential and the relevance of the graph-based approach for natural hazard risk assessment.

1. Introduction

The increase in exposure in some areas of the globe due to population growth and fast urbanization (Albano et al., 2014),

25   the increase in frequency and intensity of extreme hazards events (e.g. floods, droughts, heat waves) due to climate change (Lankao and Parsons, 2010), and the increase of vulnerability due to a more complex and interconnected society make

disaster risk management a worldwide challenge (Menoni et al., 2002). This prompts scientists, practitioners, policymakers and society to develop and apply methodologies to better assess and manage disaster risk. As we discussed in part I of this paper, present-day urban society is characterized by complex socio-technological networks and

30 increasingly relies on interconnections between critical facilities (Pescaroli and Alexander, 2016). This calls for a more sophisticated and holistic perspective to model risk than the traditional reductionist framework, where collective risk is modelled as a collection of risk elements individually, neglecting the links between them and the broader impacts that they may originate. While various approaches that analyse secondary events (or indirect losses) and cascading effects exist, they are generally limited to the field of critical infrastructure networks (e.g. electric power, Rinaldi et al. 2001),

35 without focusing on collective risk.

This paper is organized in two parts: I. A holistic modelling approach based on Graph Theory, II. A pilot study in Mexico City. In part I, we propose the use of a network to portray the complexity of a risk system, in particular by means of the Graph Theory (i.e. the branch of mathematics that studies the properties of a graph), which provides an alternative angle to assess risk from a holistic and systemic viewpoint. The comparison of the proposed approach with the traditional one

40 exhibits analogies between certain graph properties, relative both to single parts and whole systems, and traditional risk features. The graph properties such as betweenness, closeness, degree distribution, authority and hub not only are favourably connected to the risk factors, but they can be also very useful to understand some important mechanisms of risk (How it is propagated? Where are the weaknesses of the systems?). Moreover, the proposed approach provides information regarding the structure of the whole system (e.g. by the degree distribution of the network) and information

45 about how single parts of the system can have an effect or be affected by the system (e.g. hubs and authority values). The graph and its components are described by different metrics, some considering the neighbourhood of node (e.g. degree and clustering coefficient) and others considering the whole structure of the graph (e.g. closeness and betweenness, hub and authority for directed graphs). Beyond these static metrics, there are properties that explore the dynamic response to a perturbation, such as percolation thresholds and fragmentation modes (Setola, 2010). The integration of all these

50 information in the context of a risk assessment provides a more comprehensive understanding of potential impact and loss, allowing the implementation of more holistic disaster risk reduction strategies. In part I, we also described the two main steps comprising the proposed approach: 1) network conceptualization, and 2) construction of the graph. In the first step, the relevant typologies of exposed elements are identified (e.g. population, schools, power stations) and the connections between typologies are defined (e.g. schools provide education services to students). The construction phase

55 starts with the definition of rules between elements (e.g. students go to the closest school). Based on the typologies, connections and rules, it is then possible to build the complete graph, where all the exposed elements are linked and establish a unique network.

While part I of the paper dealt with the theoretical aspects, part II presents an application of the proposed approach to a pilot study in Mexico City and discusses its practical advantages in more concrete terms. More specifically, the main aims of part II of the paper are:

- to present the feasibility of implementing the new approach through a pilot study in Mexico City;
- to show the benefits of the new approach with respect to the traditional risk assessment for the implementation of decisions related with Disaster Risk Reduction (DRR);
- to suggest future developments and identify the main challenges for wider applications.

2. Pilot study: Mexico City

2.1 Description of the context

Floods, landslides, subsidence, volcanism and earthquakes make Mexico City one of the most hazard-prone cities in the world. Mexico City is situated in a high mountain valley (approximately 2,200 m a.s.l) surrounded by mountains of volcanic origin in the southern part of the Basin of Mexico. Mexico is one of the most seismological active regions on earth (Santos-Reyes et al., 2014), floods and storms are recorded in indigenous documents, and the Popocatépetl volcano has erupted intermittently for at least 500,000 years. At present, people settle in hazardous areas such as scarps, steep slopes, ravines and next to stream channels.

The Mexico City Metropolitan Area (MCMA) is one of the largest urban agglomerations in the world. Located in a closed basin of 9,600 km^2, MCMA spreads over a surface of 4,250 km^2. The MCMA has a metropolitan population estimated at 21.2 million, concentrating 18% of country's population, and generates 35% of Mexico's gross domestic product on a surface equivalent to less than 0.3% of the national territory (Campillo et al., 2011). This pilot study focuses on Mexico City (also called the Federal District - MCFD), where approximately 8.8 million people live, corresponding to 42% of the metropolitan population. The choice of MCFD as a pilot case allows showing the importance of modelling connections and interdependencies in a complex urban environment.

Tellman et al. (2018) show how the risk in Mexico City's history become interconnected and reinforced. In fact, as cities expand spatially and become more interconnected, the risk become endogenous. Urbanization increases the demand for water and land. The urbanized area inhibited aquifer recharge, and the increase in water demand exacerbated subsidence due to an increase of pumping activity out of the aquifer. Subsidence alters the slope of drainage pipes, decreasing the efficiency of built infrastructure and the capacity of the system to both remove water from the basin in floods as well as deliver drinking water to consumers. This exacerbates both water scarcity and flood risk.

2.2 Network conceptualization and construction of the graph

Given the very large scale of the city, certain simplifications and hypotheses had to be assumed for conceptualizing the network. Furthermore, the choice of element typologies, the connections between them and the definition of rules are also done considering the availability of data, which were provided by the Engineer Institute UNAM México D.F.

90   Among the possible exposed elements, we selected 6 typologies that are representative of both emergency management phase (e.g. Fire Stations) and long term impacts (e.g. Schools). The typology of elements considered in this pilot case, which provide and/or receive services reciprocally, are: Fire Stations, Fuel Stations, Hospitals, Schools, Blocks, and Crossroads. The Fire Station represent the node type from which the Recovery service is provided to all the other elements present in the area (except Crossroads). The Fuel Station represents the node type that provides the Power service, the

95   Hospital provides the Healthcare service, and the School provides the Education service; the elements with these three typologies deliver their respective services to all the Blocks. The Block is the node type defined as the proxy for population, which receives services from all the other considered elements. The simulation uses Blocks instead of population, as this enables a reduction in computational demand by lowering the number of nodes from 8 million to few tens of thousands. Finally, the analysis considers 17 Crossroads, which provide the Transportation service to all the other

100   elements. The Crossroads were identified by selecting the major intersections between the main highways present in the road network of MCFD. All the typologies, numbers of elements and the connections between them are presented in the conceptual graph in Figure 1. Figure 2 presents the GIS representation of the providers, and the services that are provided between them.

The link between two elements of two different typologies was set up based on the geographical proximity rule: each

105   specific service is received by the nearest provider (e.g., a Block receives the Education service from the closest School, and the School receives the Recovery service from the closest Fire Station). This simple assumption is due to the lack of data available at this stage; in case of more data, it will be possible to define relation more accurately (e.g. School offer education service to its zoning) but without changing the general validity of the method. The service provided by the road network was modelled considering that each element in the area receives a transportation service from the closest

110   Crossroad between the 17 that were identified. This approach does not model the complete road network system, particularly the paths between nodes or possible alternative paths, but it does allow considering transportation network in the analysis in a simplified manner.

The list of nodes, which contain all the elements of all typologies, together with the list of links between them, both obtained according to the previous considerations, are the inputs to build the mathematical graph. The graph was obtained

115   using the open source igraph package for network analysis of the R environment (http://igraph.org/r/), which provides a

set of data types and functions for the implementation of graph algorithms. The full library of functions adopted are available in Nepusz and Csard (2018).

3. Graph analysis

The proposed approach addresses the whole exposed system and its single parts together. Using a medical analogy, illness
120 in a single part of the body can affect the general health of the person. In order to treat the whole health of the person, one needs to identify the parts that are the sources of the issue and address them. This well accepted concept was already explained in the writing of Plato: "*I dare say that you have heard eminent physicians say to a patient who comes to them with bad eyes, that they cannot cure his eyes by themselves, but that if his eyes are to be cured, his head must be treated; and then again they say that to think of curing the head alone, and not the rest of the body also, is the height of folly. And*
125 *arguing in this way they apply their methods to the whole body, and try to treat and heal the whole and the part together.*" (Plato, 1578).

Analysing the properties of the entire graph allows looking at the studied system as a unique entity that results from the connections and interactions between its parts. These global properties show how the whole system is vulnerable to an external perturbation, as for example a hazardous event that can affect part of it. In parallel, the approach also proposes
130 to assess the properties of the single nodes of the graph in order to check which element, or set of elements (i.e. part), are more critical for the entire system. For instance, as explained in part I, a hub is a node that provides services to many nodes in the graph. In case it is affected by a hazard event, its impacts on the system are relatively much larger than others. The following paragraphs present the results from the graph analysis and show how the properties of whole system and the single elements are assessed, from both provider and receiver perspectives. This analysis provides valuable
135 information that can be synergistically integrated in the traditional reductionist risk assessment approach in order to take more informed decisions for a DRR strategy.

3.1 Whole system: how it is vulnerable to an external hazard?

The analysis in this section shows the structural properties of the whole network (i.e. network topology, arrangement of a network) and investigates how the network, as a unique entity, is vulnerable to a potential external perturbation (e.g.
140 hazardous event).

As anticipated in part I of the paper, there are two types of network, heterogeneous or homogeneous, depending if the degree distribution is respectively heavy tailed or not. Heterogeneous networks have few nodes with a high degree. In other words, they have few hubs that appear as outliers in the degree distribution and make it heavy tailed. This feature can represent a potential weakness of a system, because if one of the hubs is affected by an event, it will propagate the

145   impacts more extensively than other nodes. Note that this is not per se an indication of risk, which is a function of not only the exposed system but also the hazard. However, it may be used to evaluate the vulnerability of the system as a whole, similarly to how single-site vulnerability analyses assess the potential impact of an event regardless of its actual likelihood.

There is an objective way to estimate if the degree distribution is heavy tailed by means of its statistical properties: a
150   distribution is defined heavy tailed if its tail is not bounded by the exponential distribution. In order to verify if the degree distribution of a network is heavy tailed, one can infer the Generalized Pareto Distribution (GPD) on the observation and analyse the shape parameter (Beirlant et al., 1999; Scarrott and Macdonald, 2012)(Scarrott and Macdonald, 2012). The GPD has three parameters, in particular when the shape parameter is equal to zero, the tail of distribution is exponential. Instead, if the shape parameter is greater than zero, the tail of the distribution if fatter than the exponential, and therefore
155   the distribution is heavy tailed. However, in order to fit the GPD to the data, it is first necessary to select a threshold value and consider only the exceeding values. There are different techniques to select the right threshold value (Coles, 2001). Figure 3 shows the values of shape parameter (sp) for the degree-out distribution of the Mexico City network for different values of threshold in terms of data percentile. The shape parameter ε is positive for any value below 0.8; over that value, the degree distribution is meaningless because it is not representing the whole network anymore, but only the extreme
160   values. For this reason, we can conclude that the degree-out distribution is heavy tailed. This confirms that the network built for Mexico City is strongly non-homogeneous, with a few hubs (providers) that are linked to many elements

As mentioned above, the degree distribution has a strong influence on network vulnerability after an external perturbation: heterogeneous graphs are more resistant to random failure, but they are also more vulnerable to intentional attack (Schwarte et al., 2002). In the case of random failure, there is a low probability of removing a hub, but if an intentional
165   attack hits the hub, the consequences for the network could be catastrophic due to the central role of the hubs.

The existence of a few elements that have a central role in the system is analysed in more detail in Figure 4 through the comparison between different provider typologies. Figure 4 presents the empirical Cumulative Distribution Functions (CDFs) of the degree-out, together with the skewness and standard deviation values, for each typology that provides a service. Population is not plotted because it does not provide a service.
170   This figure shows that the Fire Stations have a relatively uniform distribution: the values of degree-out are almost uniformly distributed between zero and one, and the distribution has the lowest skewness value and second highest value of standard deviation. Differently, the other nodes that provide services have a non-uniform distribution with much higher values of skewness and lower values of standard deviation. In particular, more than 90% of the Schools and Fuel Stations nodes have degree-out less than 1000 and few nodes with values above 5000.
175   The uniform distribution in the case of Fire Stations reflects variable values of degree-out, and the high value of the standard deviation confirms their high variability. For this reason, assuming that it is intended to distribute the resources

more homogeneously, it is necessary to intervene in many elements: some elements need to reduce their services and other need to increase. The less uniform distributions with higher skewness of the other providers (in particular Schools and Fuel Stations) show few elements with much higher degree-out compared to the other elements. For these specific

180   typologies, it is prior necessary to intervene in reducing the service in those elements, in fact if an external perturbation (e.g. hazard) will hit the elements with very high degree-out the impact in the network will be much higher.

A decision maker that would like to better plan the distribution of services will look for constant CDFs, i.e. all elements of each typology have degree-out values close to a constant value. In principle, a constant CDF reflects a better strategic localization of the providers that homogeneously distribute the resources on the territory. However, as mentioned above,

185   in order to assess risk and realize a proper DRR strategy, this information needs to be integrated and overlapped with hazard data (e.g. intensity, extension, probability of occurrence).

3.2 Single element vulnerability: which receivers tend to be more isolated in the system?

As described in part I of the paper, the vulnerability of a node in the system is the aptitude to remain isolated from the whole system when the graph is perturbed. The tendency to observe isolated parts is here analysed by the closeness

190   property. Closeness centrality measures the mean distance from a vertex to other vertices, which it is a shortest path through a network between two vertices (Ghoshal, 2009). Figure 5 shows the geographical distribution values of closeness-in values of the Blocks.

In accordance to the model conceptualization, the Blocks increase their distance to the network if the nodes that provide services to them are not connected between each other. As an example, if a School and a Hospital provide services to a

195   Block, the closeness-in of this Block will be higher if the School and the Hospital receive the transportation service from same Crossroad and this one is also serving the Block. In this specific case, where the nodes are more interconnected, the distance between the Block node and the whole network is lower, and by definition its closeness-in is higher.

Figure 5 shows that the region with the majority of Blocks with the highest values of closeness-in is in the southeast part of MCFD. This area is the part of the city that is surrounded by few providers, which are the major hubs as illustrated in

200   the next paragraph in Figure 8. The presence of few providers forces them to exchange services between themselves and to serve all the receivers of the area, meaning that the Blocks have a lower distance to the providers and can therefore be more vulnerable.

3.3 Cascade effects: which providers are more important to the system?

The analysis of the topological structure of the network can assess the nodes from two main points of views: 1) providers, elements that provide services and 2) receivers, elements that receive services. Regarding the providers, it is relevant to

205

explore how some providers compare to others in terms of relevance to the system, according to their connections with the receivers. In particular, we propose a comparison between providers through the analysis of two properties: hub analysis of all nodes that provide service to the population, and betweenness analysis of the Crossroads.

3.3.1 Providers: role of hubs

210 In directed graphs, it is important to explore if some nodes are more important for network function than others. The importance of a node, within the purpose of providers that deliver a service, is closely connected with the concept of topological centrality: the capacity of a node to influence, or be influenced by, other nodes by virtue of its connectivity. In Graph Theory, the influence of a node in a network can be provided by the eigenvector centrality, of which the hub and authority measures are a natural generalization (Koenig and Battiston, 2009). A node with high hub value points to

215 many nodes, while a node with high authority value is linked by many different hubs.

The hub analysis considers all the elements in the graph that provide services; for this reason, Blocks are excluded from this analysis. The first comparison is between the different typologies of providers through the CDF of the hub values for each element typology presented in Figure 6. The plot shows that almost all the School, Fuel Station and Hospital elements have hub values below 0.1, and only a few of them reach high values closer to 1. Differently, only half of the Fire Stations

220 have very low values, while the maximum does not exceed 0.7. The patterns of skewness and standard deviation confirm these differences between typologies: the lowest value of skewness and highest value of standard deviation refer to Fire Stations, while the other elements types have higher values of skewness and lower values of standard deviation.

The CDFs shown in Figure 6 reveal the presence of outliers that are even more evident in Figure 7, which shows the boxplots of each typology. The outliers of hub values are important because they show which are the elements in the

225 graph that, in case of potential failure, can have a large impact on the network, for instance due to their role as major hubs. One Hospital has the highest hub value, which by definition is equal to 1, immediately followed by a Crossroad with value around 0.85, and a few Schools, Fuel Stations and Fire Stations have hub values around 0.5. The ranking of elements according to their hub values can be a very useful information to prioritize intervention actions and maximize the mitigation effects for the whole network. In fact, the hub value allow to assess the single element as a part of the

230 whole, and its ranking shows which element is more important, having more influence on the whole network. If an external perturbation hit an element with very high hub value, the cascading effects on the network is more significant due to its central role in the system. For this reason, implementing a mitigation measure to the element with the higher hub value guarantees to have the higher impact on the whole network.

The hub outliers in Figure 7 are associated to the elements of the network that are geographically located mainly in the

235 southeast part of Mexico City; as shown in Figure 8, the biggest icons are in this part of the city. Based on the available

data, the density of elements that provide services in southeast part is much lower compared to the other areas of the city; as such, the few providers existing in this part become important hubs for the whole system.

This part of the city has few providers that are central hubs of the city and Blocks with very high closeness. Together, these two aspects underline the need for additional providers in this area. This would reduce the respective number of receivers, decreasing the hub values of providers and reducing the number of Blocks depending on each of them.

3.3.2 Crossroads: betweenness analysis

As described in part I of the paper, a network that has nodes with high betweenness values has higher tendency to be fragmented because it has a strong aptitude to generate isolated sub-networks. In this case study, the transportation is the only service that allow to analysis the betweenness values of the nodes. In fact, vehicles (e.g. fire truck or family car) need to pass through Crossroads to go from point A to point B (e.g Fire trucks from Fire Station to an incident place, or family car from house to school). The betweenness analysis presented here shows the number of shortest paths between pairs of nodes that pass through the selected Crossroads. As mentioned previously, the few Crossroads considered in this pilot study are not intended to reproduce the very complex road network of Mexico City, but to present some highlights of the betweenness property.

Figure 9 shows the Crossroads adopted in the analysis, where the dimension of the icons is proportional to the value of betweenness. It can be observed that the Crossroads in the ring road around the city centre have higher values of betweenness, which is due to the fact that they connect the very large suburb areas and the city centre. In particular, the Crossroads in the south have the highest values, because the number of nodes in the south is greater than that in the north of the city. Instead, the Crossroads in the city centre connect mostly the nodes that are inside the ring road, and for this reason they have lower values of betweenness.

The betweenness value shows which Crossroad is more central, or more important and influent in the network, based on shortest paths between the nodes. As an example, in case a Crossroad is flooded, it will reduce or completely interrupt its transportation service. A Crossroad with higher betweenness will influence a higher number of nodes, and as such, if its functionality is affected, this will have a higher impact on the network compared to a Crossroad with lower betweenness.

3.4 Exposure: which receivers have higher centrality in the system?

Regarding the analysis from the receivers' point of view, we explore how the system privileges some receivers compared with others according to their connections with the providers. In particular, we propose a comparison between receivers through the authority analysis.

We first compare the different typologies of elements. The CDFs of the authority values, plotted in Figure 10, show that providers have somewhat similar distributions, while they are very different from the CDF of Blocks, which are the only

node typology that only receives services and does not provide any. In particular, approximately 50% of Schools, Fire Stations, Hospitals and Crossroads have authority values close to zero and the remaining 50% of elements distributed between zero and 0.4. Instead, the population has a distribution closer to the uniform, with authority values that span between zero and one. In this plot, the Crossroads are not reported, as they do not receive any service and therefore have authority values equal to zero. This main difference between Blocks and the rest of the node typologies is confirmed by the skewness coefficients and standard deviations of the authority value distributions: the Blocks distribution has the lowest skewness and highest standard deviation in respect to all the other typologies.

Both Figure 10 and Figure 11 show that the authority of the nodes tend to be clustered around certain values, presenting discontinuities between them. This results from the fact that all Blocks receive exactly five services from five providers (i.e. degree-in=5), and as such have the same values of authority when they receive services from the same provider nodes. Nodes with similar authority values should therefore be geographically located close to one another. This is confirmed in Figure 12, where the Blocks are represented in space and coloured according to their authority values.

Figure 12 shows a clear pattern from low values in the northwest to higher values in the southeast part of MCFD. The Blocks with higher authority values are located in the part of the city that is surrounded by the providers with highest hub values, as illustrated in Figure 8. In contrast, the Blocks in the city centre and in the northwest have the lowest values of authority. In fact, this part of the city has the highest density of providers, which decreases the number of receivers for each provider, and consequently their hub values. Note that this aspect likely results from the assumption of not considering redundancy, meaning that each node can only receive a certain service from its nearest provider. Otherwise, if redundancy was considered, the Blocks in the city centre would receive the same service from many different providers due to the higher density of such nodes.

According to these results, the blocks with higher authority are the ones that depend on the services from the providers with higher values of hubs. For this reason, if a hazardous event hits the Blocks in the southeast part of the city, this will impact the whole system more heavily, because there will be more requests to the same few hubs. Such hubs, which are potentially more overburdened in an ordinary situation due to the high number of services they provide, can put in crisis a considerable part of the network after an external perturbation.

4. Preliminary analysis of the impacts of a flood scenario

The traditional reductionist approach usually analyses and estimates the impacts of a flood scenario only in the area directly affected. Also, the estimation of the risk is the result of the hazard, vulnerability and exposure characteristics of each single element inside the flooded area considered independently. In this section, we propose a preliminary analysis of a flood scenario in the case of Mexico City according to the proposed graph-based approach. The purpose is to show

its potentiality to highlight the impacts on the indirect consequences over the whole system, even outside the flooded area, based on the graph built for this specific study.

The hazard scenario adopted is based on the development of a simplified model that explicitly integrates the drainage system and the surface runoff for the estimation of flood area extension for different return periods, under the condition of possible failure of the pumping system in the drainage system (Arosio et al., 2018). The hydrological and hydraulic simulations are based on EPASWMM (Rossman, 2015) and implemented on the primary deep drainage system (almost 200 km of network, 14 main channels and 108 manholes). For each return period, the flooded areas are computed based on the volume spilled out of each of the main manholes of the drainage system. For each drainage catchment, assumed hydraulically independent from the others, a water depth-area relationship extracted from the DTM is used to compute the flood extension and depth. Figure 13a shows the flood areas for return period of 100 years. The majority of water depth values are between 0 to 1 m (lighter blues), and only a few raster cells (darker blue) have higher values that reach up to 9.83 m in some dip places.

In the flood area there are some provider elements, as shown in Figure 13. These elements provide services to other elements, some which are also inside the flooded areas, but with others located outside. In fact, Figure 13b shows in red all the elements that receive service from the flooded providers. Even if these elements are not directly damaged, they can potentially receive an indirect consequence due to the reduction or interruption of services from the providers that are directly affected.

Using the hub analysis between the providers that are flooded, it is possible to underline the nodes that have more central role and can generate a larger cascade effect for this flood scenario. Figure 14 shows the values of hub values between the 17 providers inside the flood area. The elements with the highest values are in the south part, as they serve more blocks and these have higher authority values. The connections between hubs and authority analysis, as underlined in the previous section, show which part of the city is more critical in case of this flood scenario will occur. The information of the hazard scenario (i.e. flood area for specific a return period), integrated with the hub and authority analysis of the network, allows to qualitatively assess that the red zone of the city has a relatively higher risk compared with the rest of the city. This zone is characterized by few providers that have high values of hubs and many blocks that have high values of authority. This result shows the need for new additional providers in the red zone around the flooded area, in order to reduce the flood impact. As a matter of fact, this would reduce the number of receivers per provider, reducing the hub values of flooded providers. Consequently, the number of affected Blocks outside the flood footprint would be reduced.

This simple example shows how the impacts of potential hazard events materialize not only inside flooded areas, but also outside them. For this reason, the traditional approach based on a GIS representation of the exposure assets, which is used to estimate the direct impact, needs to be coupled with the graph approach that adds information on the connections between these assets. The framework illustrated here can then be used to quantitatively assess indirect impacts, which can

subsequently be integrated into collective risk assessments. This requires that vulnerability functions are associated to such connections. This aspect is outside the scope of the present paper.

330 **5. Final considerations**

The pilot study presented in this part II of the paper implements a new approach for assessing the risk of complex systems based on Graph Theory. This study aims to demonstrate the feasibility and usefulness of the approach through its application to the case of urban flooding in Mexico City.

In this application, the complexity of Mexico City is depicted by modelling certain selected typologies of elements of the

335 urban system and assuming simplified rules of connection between them. However, the flexibility of our approach allows for a graph to be designed with any intended level of detail, depending on the purpose of each specific application and the availability of data. For instance, if a more comprehensive characterization of the road network was required, the graph could be expanded to include additional elements other than the major Crossroads. Another example regards the rules of connections adopted in this study, which do not allow for redundancy, as each node is considered to receive its

340 services from the nearest provider only. A more detailed graph could include, for example, influencing areas for each service, which would allow considering multiple providers for some of them, provided that the required data were available.

We have adopted a simple flood scenario to illustrate how some of the measures of a graph can be used in the context of natural hazard risk assessment. However, within our framework, additional potentially relevant information can be

345 obtained. For example, here we have presented the results of the structural analysis of the graph without looking into functional properties such as the percolation threshold, which characterizes the resilience of a network and can therefore provide valuable information for practical applications. Another possible extension consists in studying how the network evolves with time following external perturbations. More specifically, by modelling how a hazardous event impacts the capacity of affected nodes to provide their services, it would then be possible to assess indirect consequences that may

350 arise in other nodes downstream, and on the network as a whole. Aspects such as these will be explored in further research work.

We believe that this study, despite consisting in a first application where certain simplifications were adopted, unequivocally shows the potential of the graph-based approach for natural hazard risk assessment. We have illustrated how the analysis of certain key measures of the graph can be used to characterize a system beyond its individual elements,

355 by accounting for the interdependencies and connections between them. We then preliminarily show how such information can be used in the context of a hazard scenario, to assess not only direct physical impacts, but also potential indirect consequences. The latter are often more relevant than the former both from the financial and the societal point of

view, and as such, this framework can be used to considerably improve disaster risk reduction and risk management decision-making.

360 **References**

Albano, R., Sole, A., Adamowski, J. and Mancusi, L.: A GIS-based model to estimate flood consequences and the degree of accessibility and operability of strategic emergency response structures in urban areas, Nat. Hazards Earth Syst. Sci., 14(11), 2847–2865, doi:10.5194/nhess-14-2847-2014, 2014.

Arosio, M., Martina, M. L. V, Carboni, E. and Creaco, E.: Simplified pluvial flood risk assessment in a complex urban
365   environment by means of a dynamic coupled hydrological-hydraulic model: case study of Mexico City, in 5th IAHR Europe confress - New Challenges in Hydraulic Research and Engineering, pp. 429–430., 2018.

Beirlant, J., Dierckx, G., Goegebeur, Y. and Matthys, G.: Tail Index Estimation and an Exponential Regression Model, Extremes, 2(2), 177–200, 1999.

Campillo, G., Dickson, E., Leon, C. and Goicoechea, A.: Urban risk assessment Mexico City metropolitan area, mexico,
370   2011.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer-Verlag, London., 2001.

Ghoshal, G.: Structural and dynamical properties of complex networks, University of Michigan., 2009.

Koenig, M. D. and Battiston, S.: From Graph Theory to Models of Economic Networks. A Tutorial, in Networks, Topology and Dynamics, pp. 23–63, Springer-Verlag, Berlin., 2009.

375   Lankao, P. R. and Parsons, J.: Water in Mexico City: what will climate change bring to its history of water-related hazards and vulnerabilities?, Environ. Urban., 22(1), 157–178, doi:10.1177/0956247809362636, 2010.

Menoni, S., Pergalani, F., Boni, M. . and Petrini, V.: Lifelines earthquake vulnerability assessment: a systemic approach, Soil Dyn. Earthq. Eng., 22(9–12), 1199–1208, doi:10.1016/S0267-7261(02)00148-3, 2002.

Nepusz, T. and Csard, G.: Network Analysis and Visualization Author, [online] Available from: https://cran.r-
380   project.org/web/packages/igraph/igraph.pdf, 2018.

Pescaroli, G. and Alexander, D.: Critical infrastructure, panarchies and the vulnerability paths of cascading disasters, Nat. Hazards, 82(1), 175–192, doi:10.1007/s11069-016-2186-3, 2016.

Plato: Platonis opera quae extant omnia, 1578.

Rinaldi, S. M., Peerenboom, J. P. and Kelly, T. K.: Identifying, understanding, and analyzing critical infrastructure
385   interdependencies, IEEE Control Syst. Mag., 21(6), 11–25, doi:10.1109/37.969131, 2001.

Rossman, L. A.: Storm Water Management Model User's Manual, , (EPA/600/R-14/413b), 1–353 [online] Available from: http://www.epa.gov/water-research/storm-water-management-model-swmm, 2015.

Santos-Reyes, J., Gouzeva, T. and Santos-Reyes, G.: Earthquake risk perception and Mexico City's public safety, Procedia Eng., 84, 662–671, doi:10.1016/j.proeng.2014.10.484, 2014.

Scarrott, C. and Macdonald, A.: A review of extreme value thresholds estimation and uncertainty quantification, Stat. J., 10(1), 33–60, 2012.

Schwarte, N., Cohen, R., Ben-Avraham, D., Barabási, A. L. and Havlin, S.: Percolation in directed scale-free networks, Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys., 66(1), 1–4, doi:10.1103/PhysRevE.66.015104, 2002.

Setola, R.: How to measure the degree of interdependencies among critical infrastructures, Int. J. Syst. Syst. Eng., 2(1), 38, doi:10.1504/IJSSE.2010.035380, 2010.

Tellman, B., Bausch, J. C., Eakin, H., Anderies, J. M., Mazari-hiriart, M. and Manuel-navarrete, D.: Adaptive pathways and coupled infrastructure : seven centuries of adaptation to water risk and the production of vulnerability in Mexico City, , 23(1), 2018.

Network conceptualization	Type of nodes	Number of elements	Service Provided	Destination of the service
	Crossroads	17	Transportation	Fire Stations, Schools, Hospitals, Fuel Stations, Blocks
	Fire Stations	11	Recovery	Schools, Hospitals, Fuel Stations, Blocks
	Fuel Stations	103	Power	Blocks
	Hospitals	39	Healthcare	Blocks
	Schools	130	Education	Blocks
	Blocks	64,282	-	-

Figure 1: List of nodes adopted in the network conceptualization.

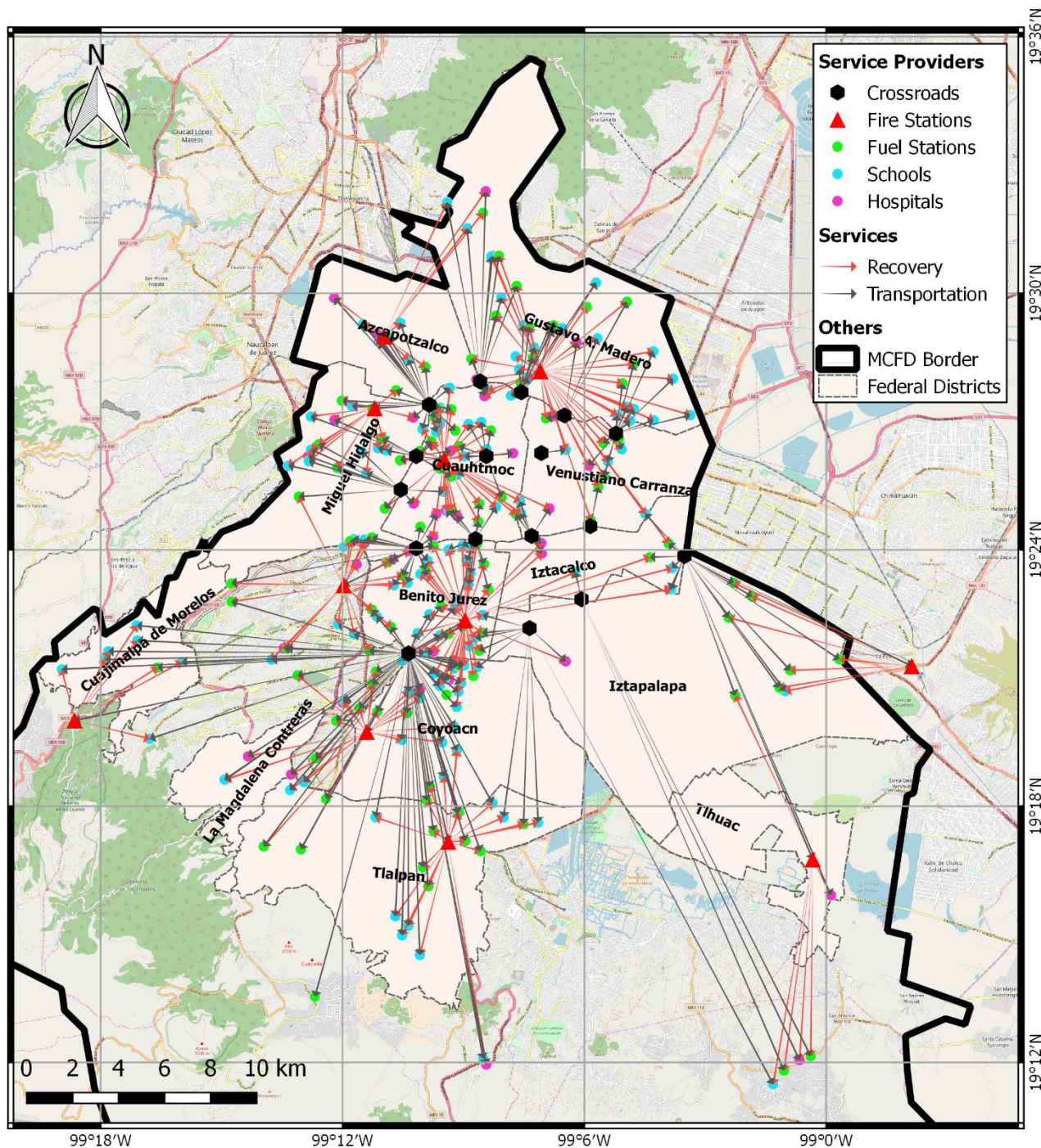



Figure 2: Map of nodes and services provided among them. For readability, Blocks are not included.

405

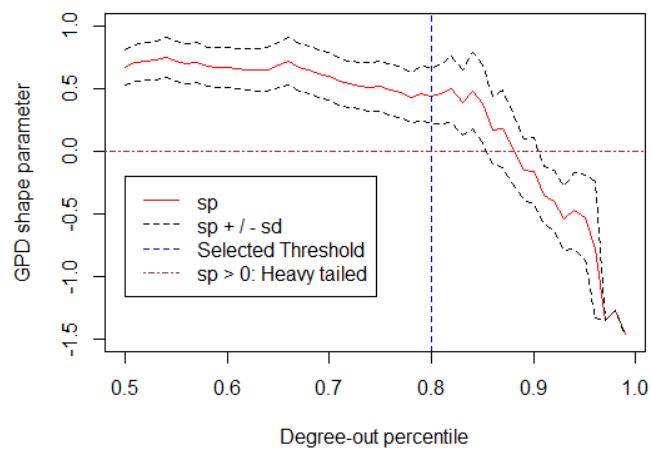
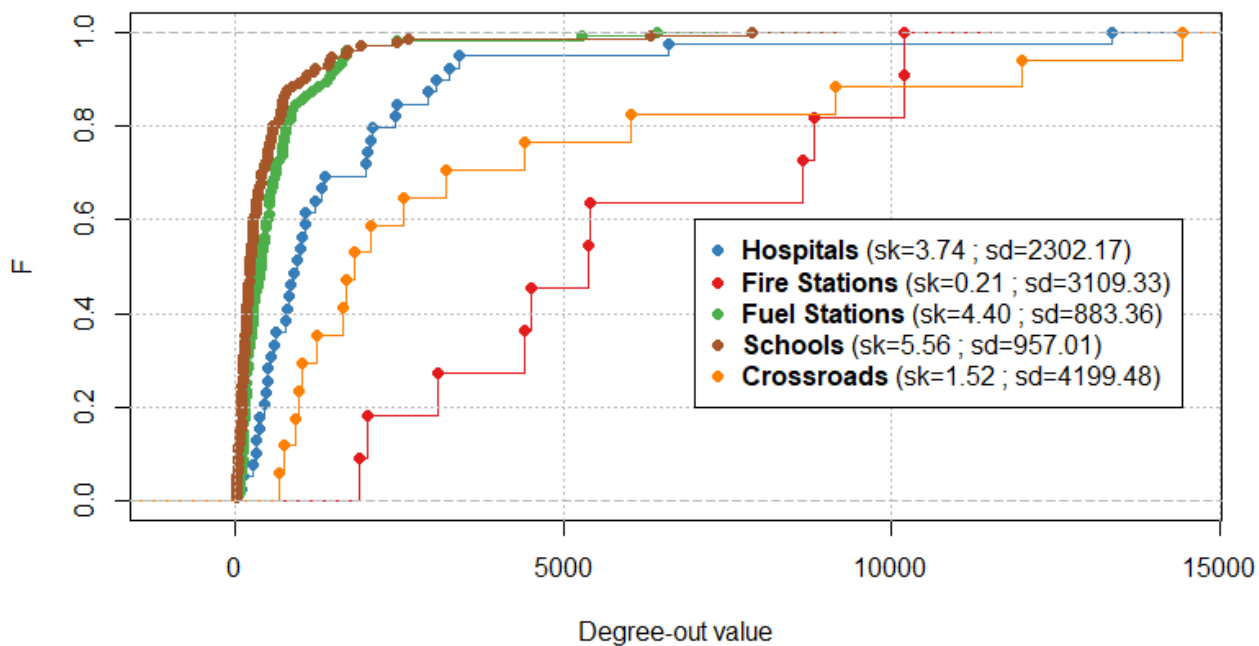



Figure 3: Parameter estimation (sp) against thresholds for degree-out data (sd: standard deviation).

410 **Figure 4: Cumulative Distribution Functions (CDFs) of degree-out for the different typologies of service providers, together with their value of skewness (sk) and standard deviation (sd).**

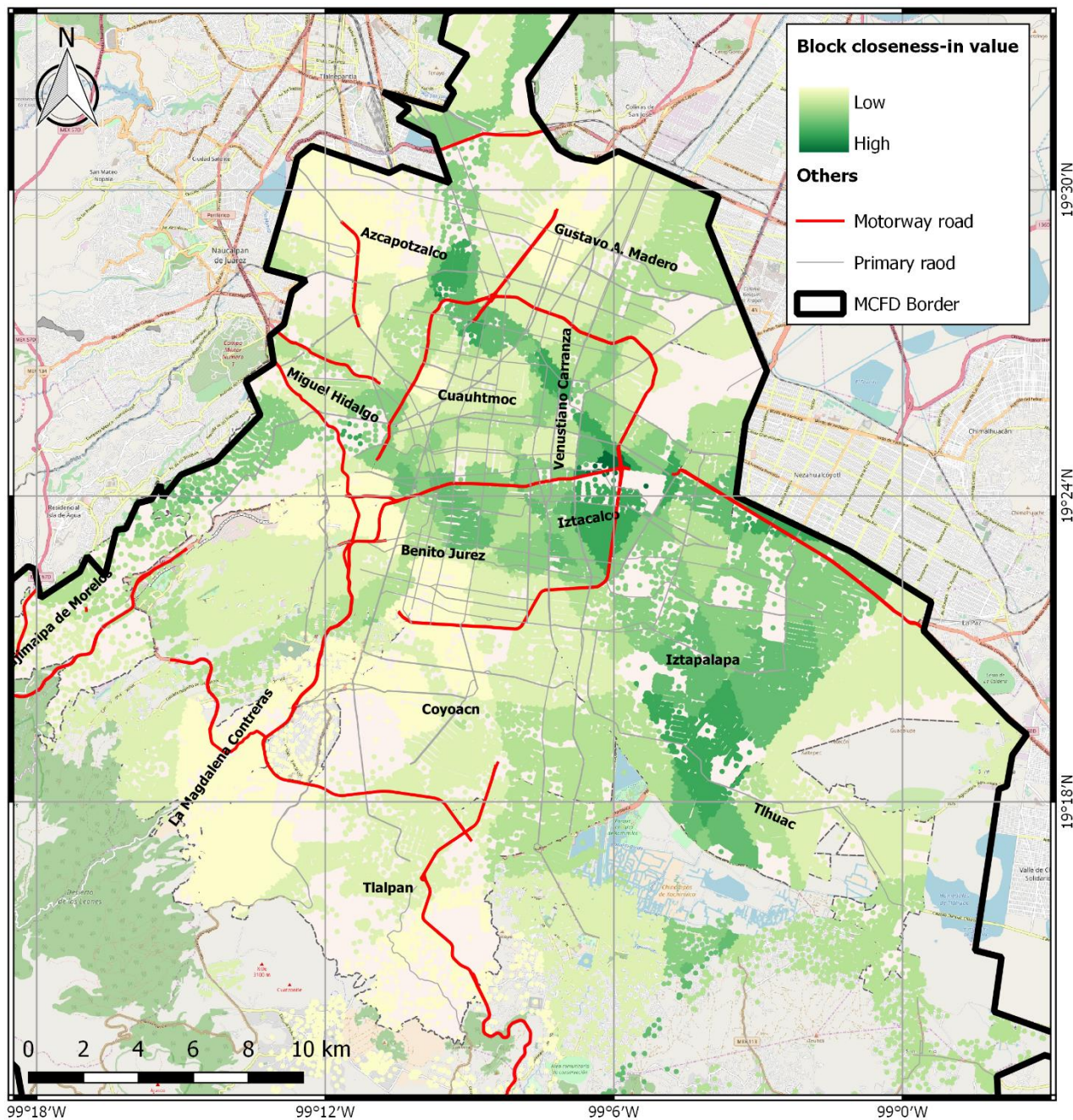



Figure 5: Geographical distribution of the Block closeness-in value.

415

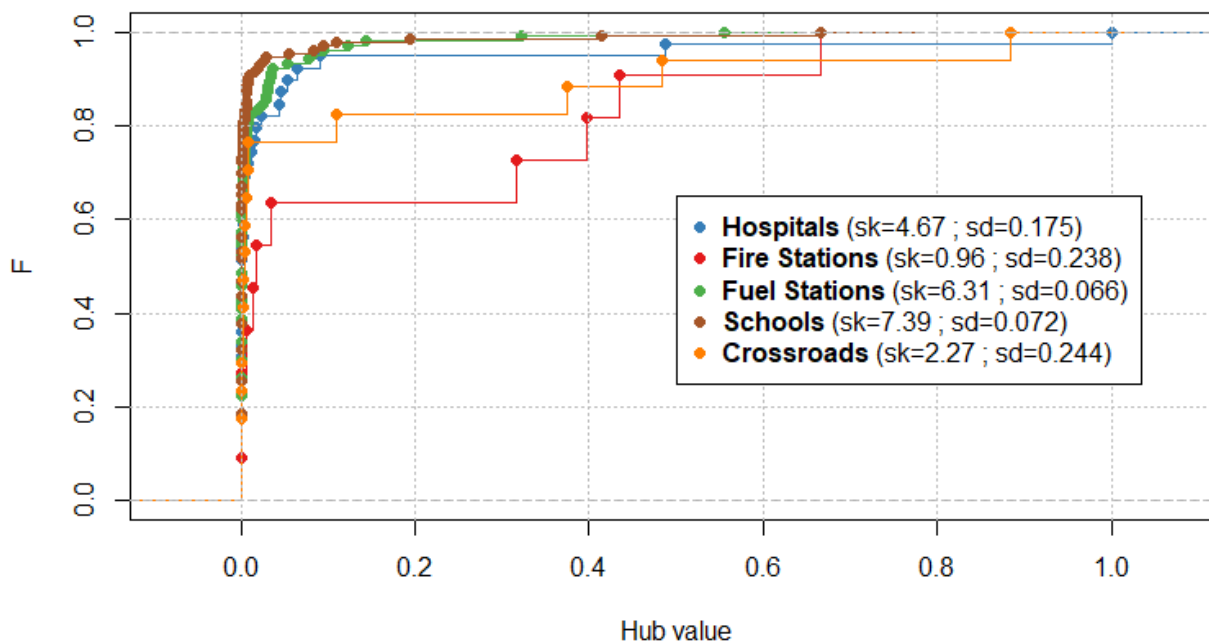



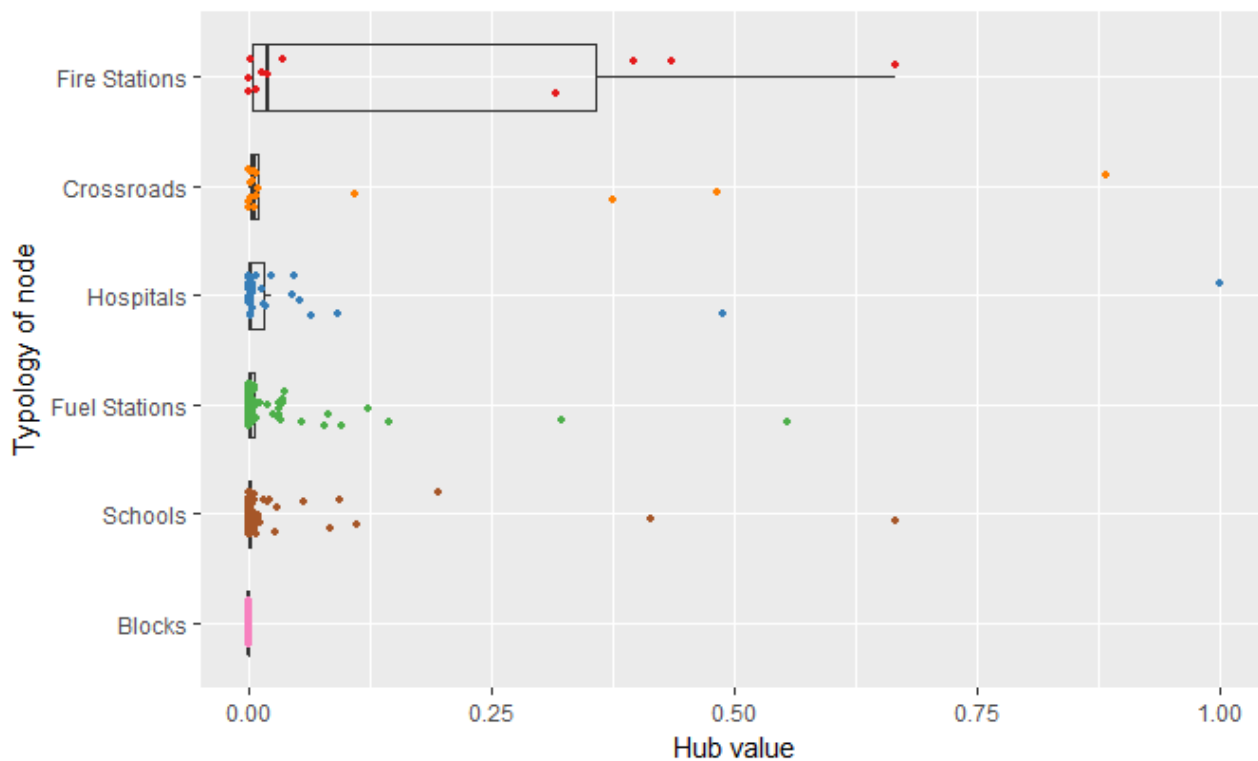
Figure 6: Cumulative Distribution Functions (CDFs) of hub value for the different typologies of service providers, together with their value of skewness (sk) and standard deviation (sd).

420

Figure 7: Boxplots of hub value for different typologies of service providers.

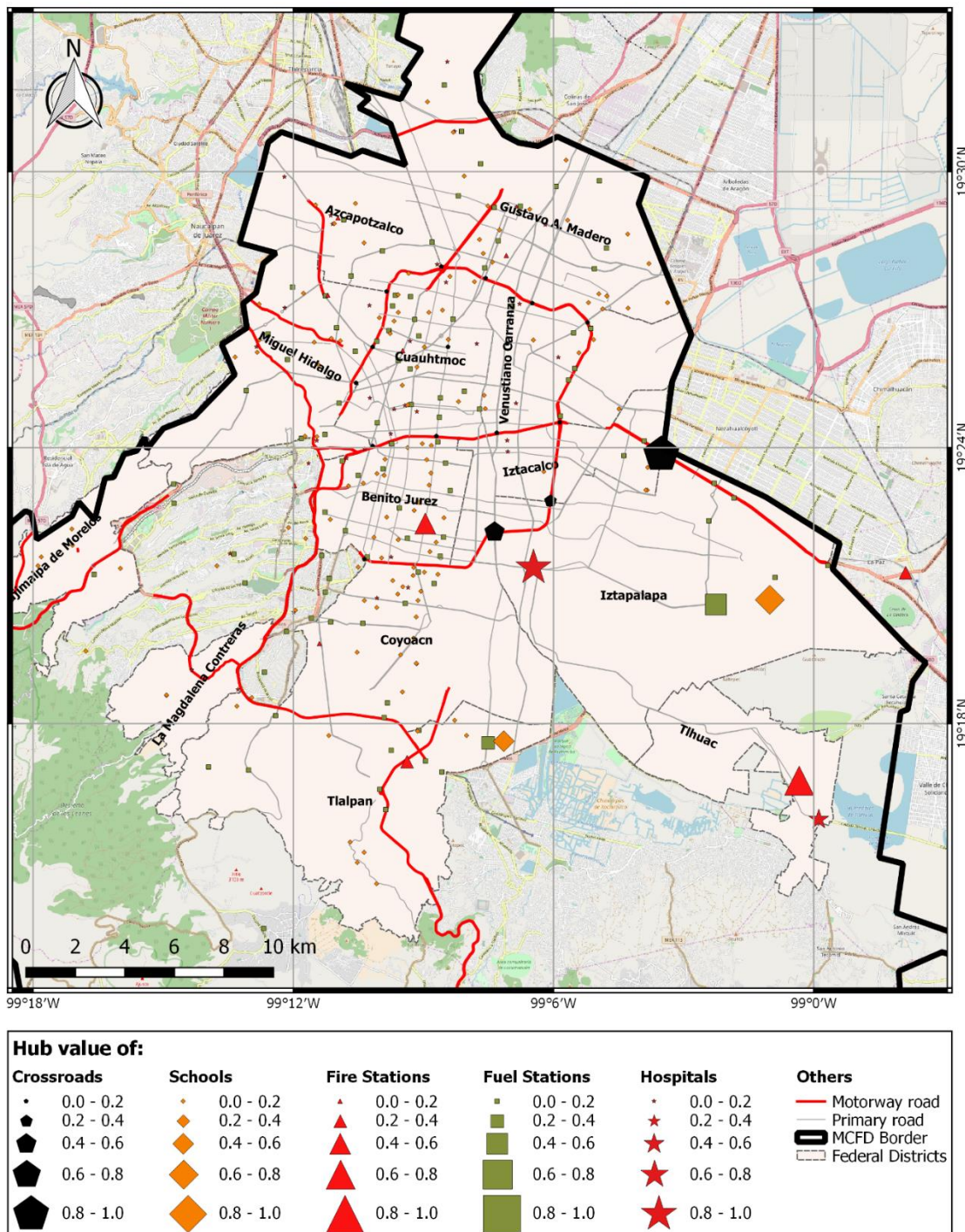



Figure 8: Map of providers. Icon dimensions are proportional to the hub values.

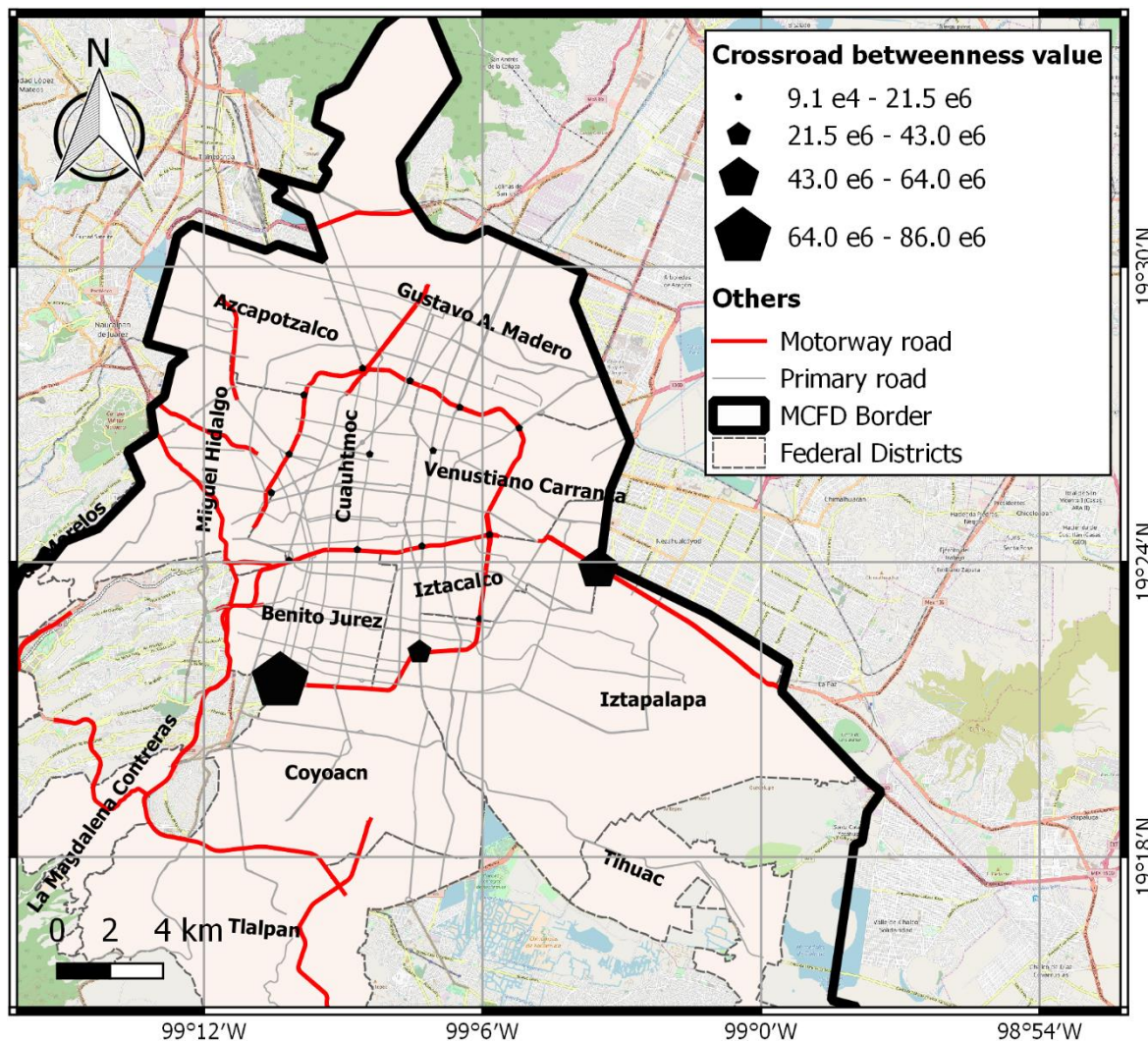



Figure 9: Map of Crossroads. Icon dimensions are proportional to the betweenness values.

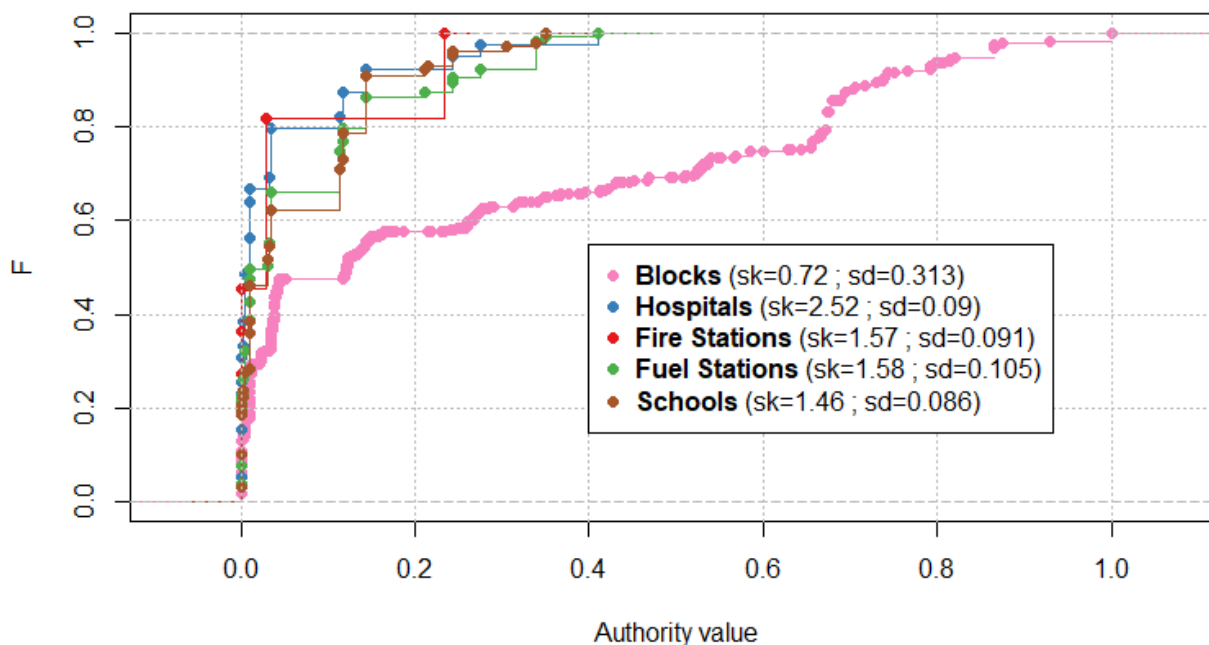



Figure 10: Shape of CDF authority, skewness (sk) and standard deviation (sd) value for different node typologies.

430

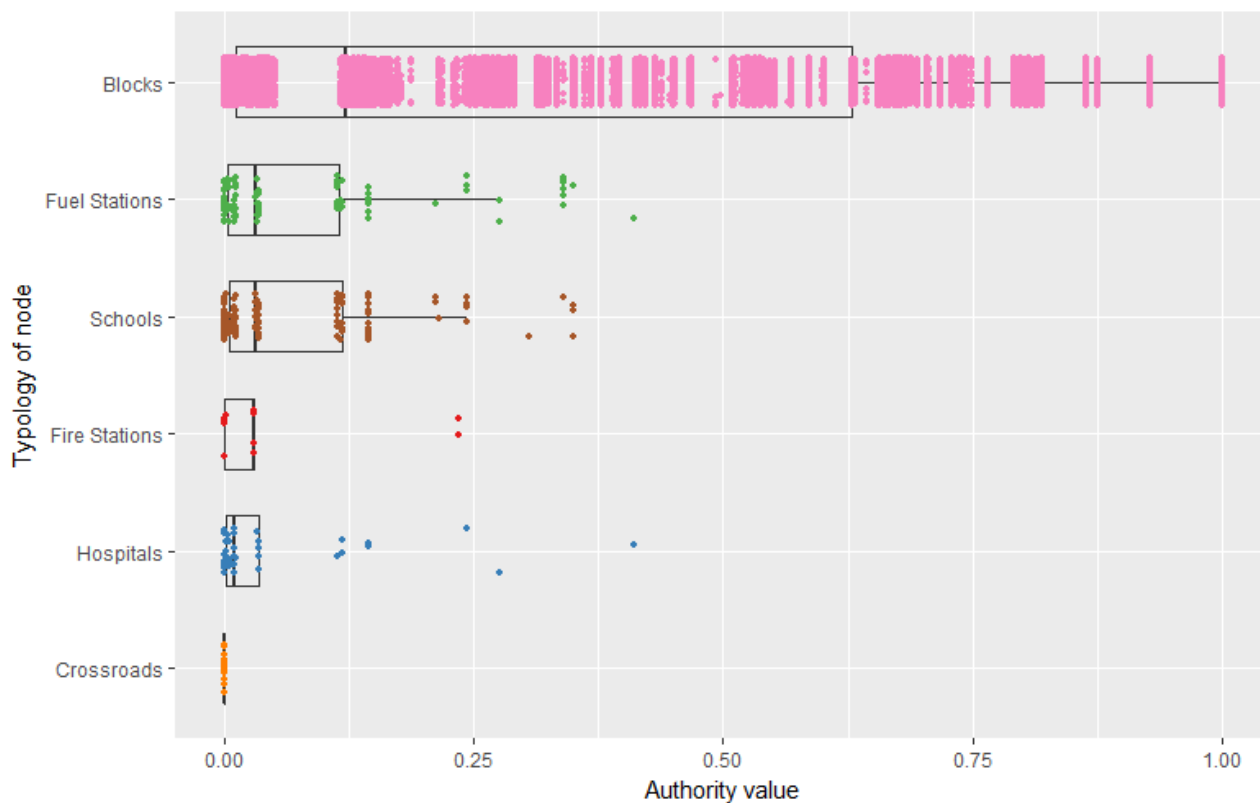



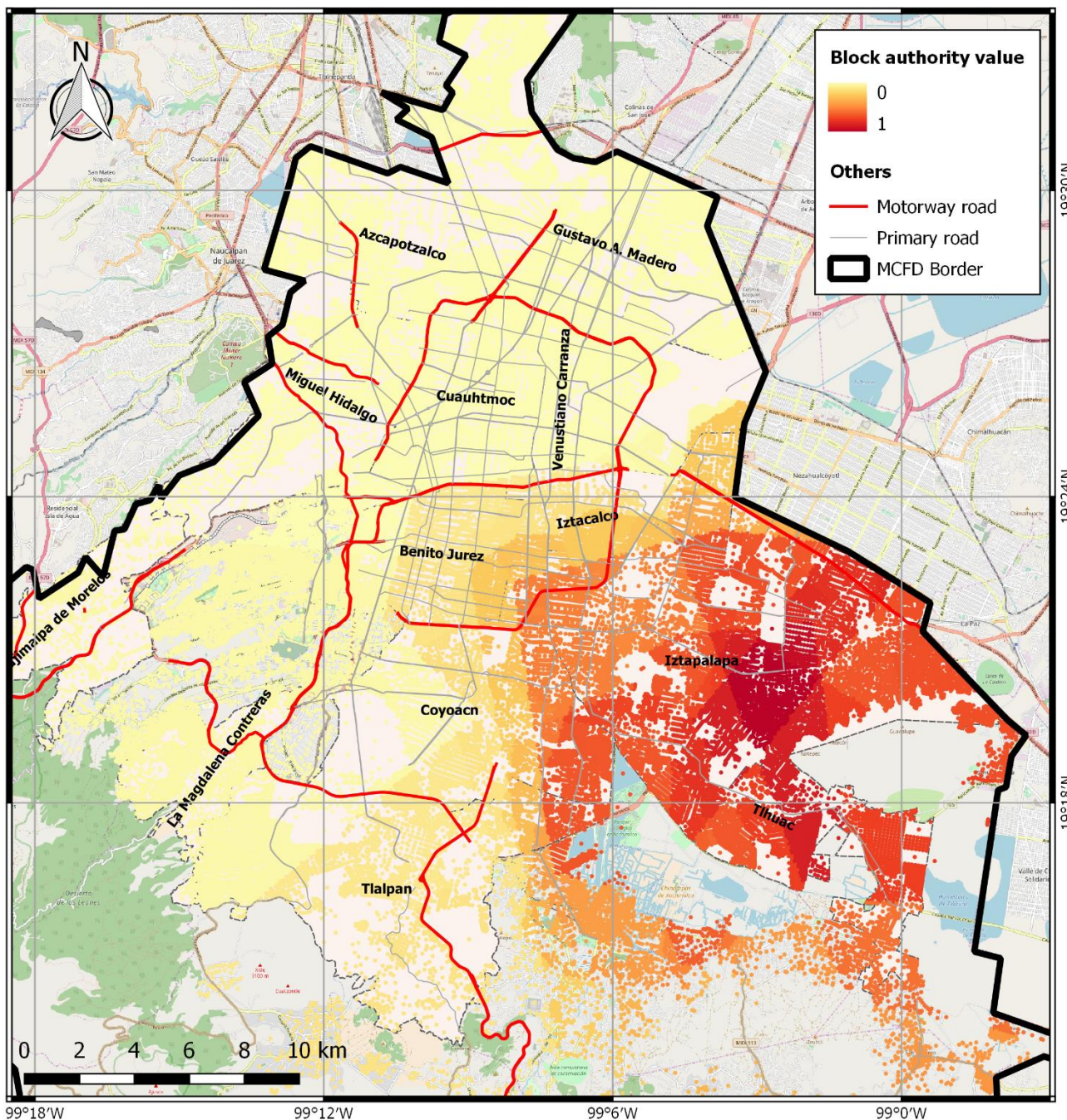
Figure 11: Boxplot of authority values for different provider services.

435

Figure 12: Geographical distribution of the Block authority value.

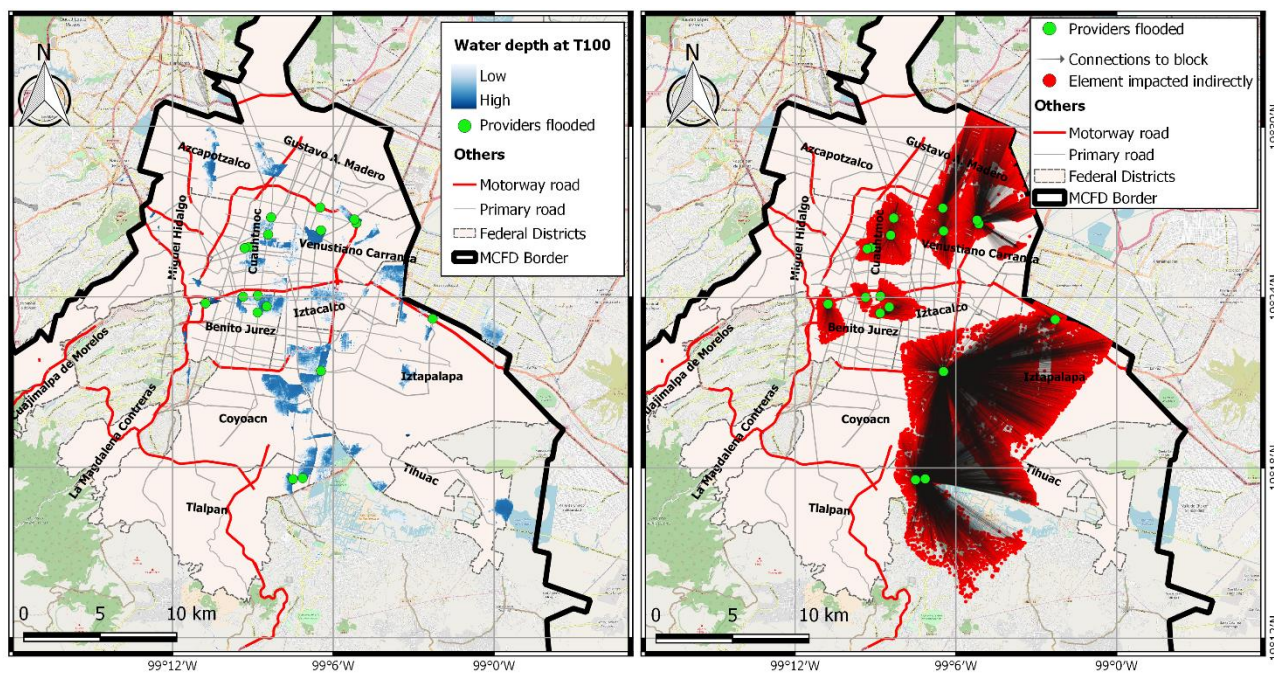



Figure 13: a) Flooded area for T100 and flooded providers; b) blocks connected to the flooded providers.

440

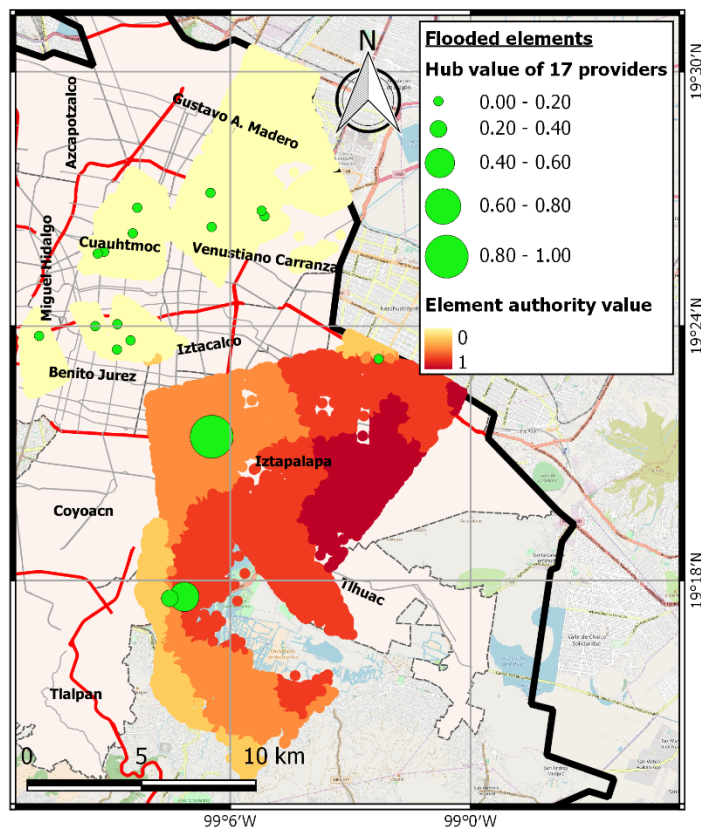



Figure 14: Hub and authority values of flooded nodes.