An urban flood database for Disaster Risk Reduction studies in São Paulo, Brazil

Lívia Rodrigues Tomás[1], Leonardo B. L. Santos[1]

[1] National Center for Monitoring and Early Warning of Natural Disasters (Cemaden) – Estrada Dr. Altino Bondensan, 500 - 12247-016 - São José dos Campos – SP - Brazil

liviatomas@gmail.com, santoslbl@gmail.com

Abstract. *This paper describes the urban flood database of an area in the São Paulo municipality, Brazil. The database is organized into input, processed, output and auxiliary data, and it contains vector files (shapefiles) and alphanumeric files. The input data have their original spatial topology. The processed data presents the spatialization of alphanumeric data as well as generated by-products. The output data presents the variables aggregated in a regular grid. It is available on GitHub online repository. The geography coordinate system is EPSG 4326 – SIRGAS2000.*

Background and summary

Flooding is one of the most frequent and damaging natural disaster affecting millions of people around the globe (Jongman et al., 2015a). It can be defined as hydrological phenomena caused by an excess capacity of surface runoff and urban drainage systems forming accumulations of water in impermeable areas (CENAD, 2014). In Brazil, the hydrometereological disasters accounted for 84% of the unsheltered from 2000 and 2019 (Perez et al., 2020).

For flood detection, data collected by gauge systems in rivers, remote sensing (Revilla-Romero et al., 2014), on-site surveys or other sources, such as media reports and government agencies (Jongman et al., 2015b), are normally used (de Brujin et al., 2019).

The Flood Alert System of São Paulo (SAISP) generates reports at two-hour intervals for each rain event, informing about its progress, power, and duration. The municipality also counts on the Climate Emergency Management Center (CGE) to monitor flood areas in the city of São Paulo and mitigate the effects of the occurrences. CGE uses radar, rain gauge measurements and flood reports provided by the Municipal Civil Defense Coordination (COMDEC) and the Traffic Engineering Company (CET) to monitor and to emit flood alerts (López and Rodriguez, 2020).

The CGE keeps the flood point records on its website (CGE, 2022) and anyone can access the data. It is necessary to search by date to access the history of floods. The search result is shown on the screen and it contains information on the location of the flood, the start and end time of the road interdiction and a classification of the point (whether it is active or inactive and whether the road is passable or impassable). It is only possible to consult one date at a time and the system does not export the result.

The occurrence of flooding with the prohibition of an important road has a great impact on people's lives, and may hinder or preclude to return to residence. In this sense, accurate and timely information on the geographic location and impacts of ongoing floods is essential for disaster response (Jongman et al., 2015b).

The potential of a hazard to become a disaster depends on the degree of exposure of the population, vulnerability, probability of occurrence, amongst others (Dickson et al., 2012). Thus, these variables must be included in the disaster risk reduction management.

Study area and datasets

The study area (Figure 1) belongs to Tamanduateí river basin in São Paulo municipality. It is a built-up area with impervious surfaces with 115.3 km², and 1,637,265 inhabitants (IBGE, 2016). It is an area with great attraction for trips, where 4,760,743 people work or study there (METRO, 2019).

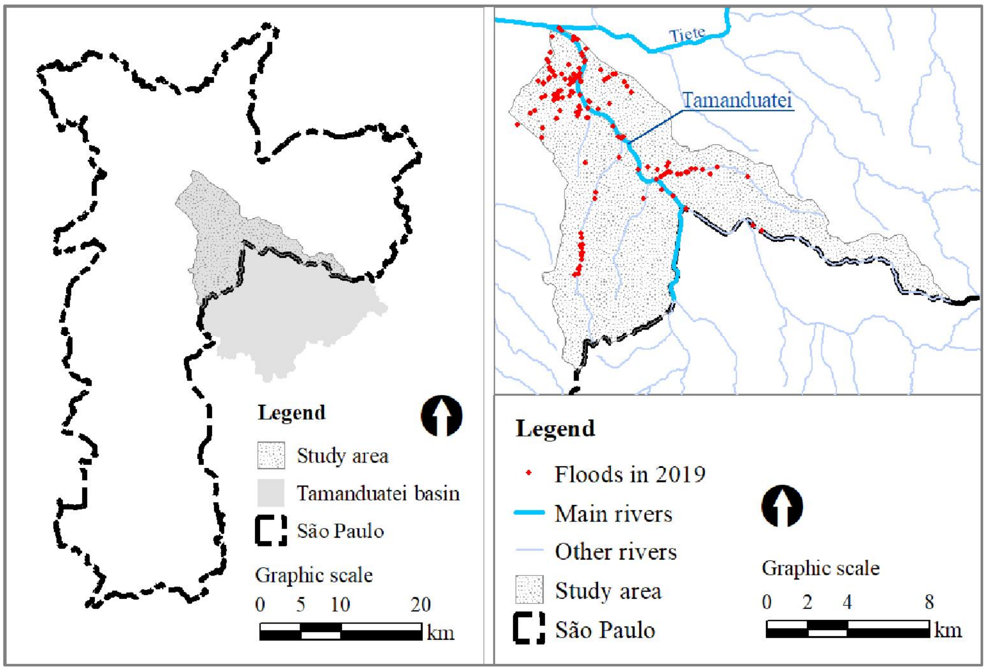

Figure 1. Study area location

To estimate the risk of flooding, and the variables associated with it, we created a geographic database considering socioeconomic data, occurrence of flooding, urban mobility, soil relief and characteristics of the road system. These input data have different sources and spatial topologies, and they are available as follows:

(a) Regular grid from 2010 Brazilian census (Fundação Instituto Brasileiro de Geografia e Estatística - IBGE, 2016);
(b) Inhabitants by cell grid (Fundação Instituto Brasileiro de Geografia e Estatística - IBGE, 2016);
(c) Geographic coordinates of workplaces and the number of people who work at each location from the Origin Destination Research Database (OD Database) (METRO, Companhia do Metropolitano de São Paulo, 2019);

(d) Geographic coordinates of educational institutions and the number of people who study at each location from the OD Database (METRO, Companhia do Metropolitano de São Paulo, 2019);
(e) Flood registries in 2019 obtained from the Emergency Management Center of the city of São Paulo - CGE website (Climate Emergency Management Center - CGE, 2021);
(f) Road system vector data from the Center for Metropolitan Studies - CEM (2021);
(g) SRTM Digital Elevation Model - DEM (Farr et al., 2007).

We generate new layers, after processing the input data. The alphanumeric data of items (c), (d), and (e) were used to generate vector files (shapefiles) with points representing the location of workplaces, education institutions and floods occurrence respectively. The SRTM DEM model (item g) was used to obtain the Height Above the Nearest Drainage (HAND) raster. Roads vector was used to calculate the roads vulnerability index.

The processed data was then aggregated in the IBGE regular grid considering their quantitative value. Each grid cell received these attributes, considering their spatial location and these approaches:

- Sum of commuters and resident population, resulting in the exposed population;
- Sum of road closure time, resulting in local vulnerability;
- Maximum value of the node vulnerability of the intersected lines, resulting in network vulnerability;
- Minimum HAND value among those that intersect it.

These variables aggregated in the grid cell were then classified and combined in three classes: moderate, high and very high, resulting in the flood risk map and its related components. The detailed methodology and all steps are described in Tomás et al., 2022. The output is the regular grid with five attributes: exposure, vulnerability, flood susceptibility, potential impact and flood risk.

The database is available at *Github* repository (Tomás, 2022), organized into the following folders: input, processed, output, and auxiliary.

Final Remarks

For most cities in developing countries, collecting reliable and timely data is a challenging task, especially from a multidisciplinary point of view. Integrating knowledge from multiple fields better reflects the geographic environment in which various natural processes, human activities, and information interactions exist. In this sense, methodologies comprised of hydrology, mobility, and Geographic Information Systems (GIS) have proved effective in many case studies to model complex urban systems. Furthermore, GIS are of great help to analyze data and produce meaningful information for urban management and resilience planning.

This database combines elements from natural geographies, such as hydrological indexes, and elements from human geographies, such as urban mobility data, in both cases using different sources and geographic units. It was originally created to map the risk of flooding in the city of São Paulo, but allows its use for other areas of knowledge, since it has variables of different natures. We hope that it can contribute to sustainable urban planning and can be added to other perspectives on the city.

References

Center for Metropolitan Studies - CEM. (2021). "Digital Geo-referenced Cartographic Base of Roads System in the Metropolitan Region of São Paulo – RMSP - 2020 Edition". https://centrodametropole.fflch.usp.br/pt-br/download-de-dados

CGE - Climate Emergency Management Center. (2022). "Flood records in São Paulo, Brazil". https://www.cgesp.org

de Bruijn, J. A., de Moel, H., Jongman, B., de Ruiter, M. C., Wagemaker, J., & Aerts, J. C. J. H. (2019). A global database of historic and real-time flood events based on social media. In *Sci Data*, 6, 311.

Dickson, E., Baker, J. L., Hoornweg, D., & Tiwari, A. (2012). Urban risk assessments: An approach for understanding disaster

Farr, T. G., et al. (2007). The shuttle radar topography mission. In *Reviews of Geophysics*, 45(2), RG2004. https://doi.org/10.1029/2005R G000183

Fundação Instituto Brasileiro de Geografia e Estatística - IBGE. (2016). "Grade Estatística Censo 2010". https://www.ibge.gov.br/geociencias/downloads-geociencias.html

Jongman, B. et al. (2015a). Declining vulnerability to river floods and the global benefits of adaptation. In *Proc. Natl. Acad. Sci.*, 112, E2271-E2280.

Jongman, B., Wagemaker, J., Revilla-Romero, B., & de Perez, E. C. (2015b). Early Flood Detection for Rapid Humanitarian Response: Harnessing Near Real-Time Satellite and Twitter Signals. In *ISPRS Int. J. Geo-Information*, 4, 2246–2266.

López, A. S. V. and Rodriguez, C. A. M. (2020). Flash Flood Forecasting in São Paulo Using a Binary Logistic Regression Model. In *Atmosphere*, 11, 473.

METRO, Companhia do Metropolitano de São Paulo. (2019). "Pesquisa Origem-Destino 2017". http://www.metro.sp.gov.br/pesquisa-od/

Ministério de Integração Nacional, Secretaria Nacional de Defesa Civil & Centro Nacional de Gerenciamento de Riscos e Desastres – CENAD. (2014). "Anuário Brasileiro de Desastres Naturais".

Perez, L. P., Rodrigues-Filho, S., Marengo, J. A., Santos, D. V., Mikosz, L. (2020). Climate change and disasters: analysis of the Brazilian regional inequality. In *Sustainability in Debate*, 11 (3), 260-277

Revilla-Romero, B., Thielen, J., Salamon, P., De Groeve, T., & Brakenridge, G. R. (2014). Evaluation of the satellite-based global flood Detection System for measuring river discharge: Influence of local factors. In *Hydrol. Earth Syst. Sci.*, 18, 4467-4484.

Tomás, L. R. (2022) "Urban flood database of an area of São Paulo municipality", https://github.com/liviatomas/floods_saopaulo.

Tomás, L. R., Soares, G. G., Jorge, A. A. S., Mendes, J. F., Freitas, V. L. S. & Santos, L. B. L. (2022a). Flood risk map from hydrological and mobility data: A case study in São Paulo (Brazil). In *Transactions in GIS*, 26, 2341– 2365. https://doi.org/10.1111/tgis.12962