An alternative flood model calibration strategy for urban watersheds: the case study of Riohacha, Colombia

Antonio Krishnamurti Beleño de Oliveira, Osvaldo Moura Rezende, Matheus Martins de Sousa, Andrea Nardini and Marcelo Gomes Miguez

ABSTRACT

The city of Riohacha (Colombia) has a complex urban setting that, under the pressure of recurring intense rains, experiences increasing flood damage. With the aim of identifying a systemic solution to flood problems, a hydrodynamic mathematical modelling exercise was conducted. Within the modelling process, calibration and validation are two fundamental actions that must precede the use of the model. However, most of the river basins around the world lack hydrometeorological information, which is indispensable for the calibration process. This paper presents an original approach to collecting such information for the calibration process, based on interviewing inhabitants. The results of this effort were surprisingly good, when considering the kind of approximations involved in using people’s answers as hard data. This encouraged us to promote it as a working solution for many other similar cases, which all suffer from lack of suitable data.

Key words | calibration, Colombia, data scarcity, hydrodynamic modelling, urban flood simulation

INTRODUCTION

Urban flooding is a growing problem in many urban settlements (Freni et al. 2010). Disorderly urban growth and lack of suitable urban water infrastructure aggravate flood disaster statistics (Balica et al. 2009). This problem is particularly significant in Latin America/Caribbean countries where, according to the United Nations (United Nations Department of Economics and Social Affairs Population Division 2015), the population will increase by 25% by 2061, reaching a peak of 793 million people.

This paper assumes that mathematical modelling plays a key role in such a context, as a tool to explore possible solutions for flood control, to help in planning urban development and to make prognoses about different scenarios (Viero et al. 2015). In particular, model calibration difficulties in the reality of developing countries (Hagen & Lu 2011) are discussed and a proposition is made to face this challenge, overcoming shortcomings in terms of data records and topographic information. A specific study was developed for the town of Riohacha, located in the northernmost Caribbean Region of Colombia (in La Guajira) as an exploratory case to test the practical approach proposed. The town of Riohacha has a great historical importance and a high environmental potential; however, it has a history of disordered growth and fast population increase followed by flooding problems and environmental degradation. According to the Demographic Census of 2005 (DANE 2005), the municipality has seen a population growth of more than 100% in 11 years. This rapid urbanization has raised the ground’s imperviousness and consequently worsened floods in the city. These effects are representative of most developing cities (Suriya & Mudgal 2012; Chen et al. 2015).

Additionally, the local topography and hydrography contribute to this situation, as Riohacha has several natural terrain depressions and lagoons. These natural ponds are used to store storm waters during the rainy season. After
reaching a given water level, they start to connect with each other, creating a superficial drainage net. However, this natural drainage system has been deeply disrupted by urbanization and the road infrastructure implemented during the city’s growth. Now, the expanding water surface reaches the streets and uses them as channels (Pérez et al. 2018). In addition, the whole town lies only slightly above mean sea level, hindering rainwater flow drainage, as already shown in Balica et al. (2012). The depressions in the urban area of Riohacha, shown in a topographic representation, together with an image of one of the biggest lagoons (La Esperanza), can be seen in Figure 1.

This topographic setup is a challenge for modelling and makes the whole situation interesting to discuss, reinforcing the use of this city as an exploratory case. However, Riohacha did not have an available and detailed topographic data set at the time of this study. This data gap limits the use of physically based modelling tools, which need extensive and precise topographic data. In fact, flow measurements were also lacking, and this is something that limits the use of almost every model, since calibration and validation play a core role in the confidence level of the modelled results.

Owing to the difficulty of data acquisition, the aim of this article is to develop and present an original strategy for model calibration and validation, based on interviews and field inspections to retrieve historical flooding information and to gain knowledge about the real physical situation. Flood levels as well as flow paths and their characteristics obtained from interviews were taken as the basic information to recreate a flooding surface for the calibration and validation processes. The data generated a Georeferenced Database that allowed the reproduction of ponding effects and a network of flow paths.

MODCEL was chosen as a model particularly suited to representing urban environments lacking detailed topographic data. This hydrodynamic model simulates the physical situation through a set of flow-cells, which are compartments capable of representing average characteristics of the urban surface, store water and communicate with each other through hydraulic links, composing a flow network of ponds and flood paths (Miguez et al. 2012; Miguez et al. 2017). The flood surface retrieved from the field strategy was used to compare the water levels (ponds) and water depths (flood paths) obtained by MODCEL simulation and hence support calibration.

**LITERATURE REVIEW**

In order to approach the watershed as a whole, while looking for a systems approach, mathematical models stand out for supporting flood diagnosis and prognosis in the solution design process. However, mathematical modelling always represents a certain degree of simplification and models need to be physically representative. The traditional hydrodynamic modelling approach tends to consider fluvial systems as one-dimensional (1D) courses. However, numerical flood simulations through 1D models can lead to distortions in mass balance when flooding spreads over vast areas, thus restricting the use of this type of simulation (Barnard et al. 2007). When simulating complex terrain features, like urban land, this situation can be even more difficult. By contrast, 2D, quasi-2D and 3D models are
better for simulating hydraulic specificities, but they have to be supported by higher computational efforts and higher topographic data requirements. With the computational processing capacity development and the improvement in the process of obtaining topographic data, there has been a move towards models with dimensions higher than 1D for the solution of hydrodynamic problems (Kuiry et al. 2010). Although defining the dimensions of the model is most often fundamental for the strategies used in the modelling processes, the choices made by the modeller for representing geometry, topography and bathymetry, the scale representation and the co-related information can be even more important for the accuracy and representativeness of the hydrodynamic results obtained (Eleutério & Mosé 2011).

Thus, key questions in mathematical modelling refer to understanding the physical reality, identifying data needs and gaps, defining basic hypotheses and simplifications, and defining dimension and model scale. The definition of model scale is one of the most important aspects of modelling studies (Singh & Kumar 2017) and it directly affects the data needs.

Therefore, considering that the quality and availability of data are relevant aspects to the modelling process, the choices and assumptions made during this process may threaten the representativeness of the results obtained (Hurford et al. 2010).

Within the modelling process, calibration and validation are two steps that must precede any simulation for planning or design purposes (Tejaswini & Sathian 2018). These steps are responsible for validating the topographic and hydraulic representation, giving reliability to the physical interpretation and consequent representation. They are needed to improve the consistency of the model and to reduce parameter uncertainty (Wang et al. 2017).

The complexity of cities and the usually large flooded areas tend to point to 2D models. Modelling 2D urban flows, however, can bring additional challenges when compared to rural floods. If a detailed terrain model is available, with a refined grid resolution of a few metres, for example, but representing just the local topography while ignoring the buildings, the correct flow patterns may be not adequately represented. In this case, when representing flood flows passing through the city, buildings will need to be defined by the insertion of solid blocks or/and by artificially increasing the roughness coefficient in the location of the buildings (Bellos & Tsakiris 2015). In this way, when calibration data are available, the Manning coefficients are frequently adjusted with values greater than the ones corresponding to the actual terrain (Fewtrell et al. 2008). If no calibration data are available, the modeller may not account for this effect and the obtained results may be not valid. Furthermore, in engineering practice, non-calibration of models may result in under- or overestimation of design values, resulting in (significant) wasted resources (Tscheikner-Gratl et al. 2016) or unexpected disasters.

Calibration and validation of a model is highly influenced by data availability or quality issues (Freni et al. 2009; Kleidorfer et al. 2009). Most of the river basins around the world (especially small urban catchments) lack hydrometeorological information (López et al. 2017), which is an indispensable input for the calibration and validation processes. This is frequently the case in Latin American countries, where the lack of (reliable) hydrometeorological information is a central problem in the use of hydrodynamic models to support urban drainage and flood control projects (Santos Júnior & Santos 2014).

Scarcity and uncertainty of the available data are the main limiting factors for a comprehensive assessment of urban waters, but this can be solved through model adaptations to compensate for the incompleteness of the data (Espinosa & Otterpohl 2014). However, model adaptations or model simplifications should be supported by a consistent physical interpretation. The lack of data cannot itself justify the simplification of a model. If this is the case, the modeller can make multiple mistakes, distorting the mathematical equations that should represent physical reality and using a non-representative data set.

**MATERIALS AND METHODS**

Van Dijk et al. (2014) discuss the difficulty of choosing the appropriate modelling technique for certain situations. The choice of model affects simulation results and decision making related to extreme rains. Choosing the most appropriate modelling technique means taking into account the accuracy, time spent on the computing step and data availability. The selection process should also take into account the fact that these models can hardly ever be calibrated for extreme rainfall situations since these events are very rare and descriptions of flooding reaching very high levels are often not accurate enough.

Besides, when discussing urban floods, not only can the calibration introduce uncertainties (especially because data measurements are not so frequent in small rivers or storm drains) but also the complex setup of cities brings a modelling challenge. Flooded urban areas may comprise different water levels, separated by walls and buildings or
other urban structures. Abily et al. (2013) discuss the possibilities, performances and limits of standard modelling tools for high-resolution runoff simulations of a dense area with complex structures that affect (and alter) drainage paths. These authors (ibid.) argue that modelling complex sites deserves special attention to rapid changes in flow regime, small water depths and high gradient properties, and local effects introduced by some of the aboveground structures. Physical obstacles, like walls or earthworks, may disconnect surface flow levels, or link separate flooded areas by weirs or gaps, for example.

Considering the peculiarities of the Riohacha case study, including the terrain characteristics and the lack of data, MODCEL was chosen due to its capability for dealing with topographical interference over flood flows. MODCEL is a quasi-2D hydrodynamic model in which the watershed is discretised into flow-cells; such cells are defined as homogeneous compartments that represent a portion of the terrain (natural or urban) and act in an integrated way, linking each other in a flow mesh using only one-dimensional hydraulic equations. The resulting grid is irregular, in the sense that each cell may assume the proper form to represent a certain portion of the terrain. Different scales may co-exist in the representation of the cells’ grid. The cell limits can be defined by topographic characteristics, using watershed reasoning, but they can also be defined by urban features, such as building contours. The representation of a schematic urban cell can be found in Figure 2.

Different hydraulic relationships may define the connection between cells. The most commonly used is the Saint-Venant dynamic equation, but local structures, like weirs, gaps gates and pumps may also be represented.

MODCEL is suitable for simulation of the particular ‘flattened egg box’ topography of Riohacha. A similar exercise was performed by Sousa (2018), who carried out the tests proposed by the UK Environment Agency through the comparison of MODCEL with several mathematical models (Néelz & Pender 2013). One of the tests consisted of modelling a square region of 2,000 m by 2,000 m with a ‘flattened egg box’ shape. MODCEL showed satisfactory results for this test when compared

---

**Figure 2** | Schematic representation of an urban cell and how it works in a model mesh.
to the other models. Water levels and flow velocities were reproduced consistently (Figure 3), using a much lower number of elements when compared with 2D hydrodynamic models using digital terrain models (DTM), proving that the lack of detailed topographic data does not limit the use of this model.

A usual calibration database for Riohacha is not available. To address this gap, and to create a valid calibration database, a survey that included a significant sample of households was carried out to map the effects of two intense rainfall events: September 2011 (used for calibration purposes) and November 2011 (for validation purposes). In particular, a key feature was the necessity to relate the flood depth observed inside houses to the local flood elevation referred to streets, since these field surveys were only able to gather relative flooding information from house owners, but not the absolute elevation values. Understanding the uncertainties that such a procedure could raise is necessary to critically assess the possible impacts on the results. However, it is worth remembering that uncertainties are inherent to the modelling process, and may come from several sources, ranging from the model formulation itself to the data availability for the calibration process (Deletic et al. 2012).

Geographic information system (GIS) tools allowed us to merge the information about the territory with the information collected during workshops, field surveys and local inquiries conducted within a broader participatory process.
In particular, the ability to geo-graphically locate the flooded houses and the co-related flood depths and, consequently, flood elevations, was very useful. This spatial information was used to produce a flood surface that helped in understanding how flow paths develop and where water is stored in ponds. When relating these characteristics to the cells and the resulting flow mesh, the ponds will refer to the storage capacity of the cells, while the flow paths are drawn between each pair of cells in the mesh or through a set of cells that compose a continuous line of flow.

As anticipated, field survey teams were not able to directly measure the elevation of the water in the flooded households; in fact, only height differences could be measured, as shown in Figure 4. According to this figure, the flooded water depth $h$ of the lowest sector of the house was measured, as indicated by residents. Then, the relative height difference $ΔE$ between the street in front of the house and the house entrance was also measured (some houses have a kind of flood gate), as well as $ΔI$ between that threshold and the lowest floor, both taken with their algebraic sign according to figure convention. Once the elevation $Y_{road}$ of the road in front of the house was known (from a DTM raster built on the basis of the available topographic map), it was possible to translate such relative information into absolute flood water elevations.

After the detailed analysis and comparison of each house, it was observed that some elevations or depths of water would be physically impossible to occur. Some responses were inaccurate due to memory weakness or overstatements (for example, sometimes one house was flanked by two or three others that consistently reported a certain water level, while the first house reported a much higher inundation). Then, a reliability analysis was performed for the calibration and validation process, spatially crossing the information obtained. Data were discarded when at least one of the following conditions occurred:

(a) the reported value represented a high water depth at a point of high slope or small contribution basin;
(b) the reported value had a water depth that was very different from most points nearby;
(c) water was reported only inside the house, without water records on the street.

Surveyed data (after such filtering) were then compared with the simulation results. However, a complementary distinction was made bearing in mind the natural characteristics of the watershed: cells were classified either as ‘storage ponds’ or as ‘flow paths’. ‘Storage ponds’ occur in flat areas with depressions; there, the main characteristic of the cell representing such a portion of the terrain refers to its storage capacity with virtually no water in motion (at least, not as a predominant character). Such cells may present irregular terrain surface; thus, the appropriate variable to consider when calibrating the model in these types of cells is the water elevation (which will be mostly horizontal over the cell area), and not the water depth, since several different measured points falling within such a cell should display a very similar absolute elevation value. ‘Flow path’ cells, instead, characterize zones where local slope and urbanization patterns indicate the presence of predominantly running water, without significant storage. Hence, the water depth is the appropriate variable to be used in these types of cells, in the surveyed-modelled comparison, since these depths will be directly related to the conveyance of the flow cross sections established by the running waters. Therefore, the water elevation values will be different over the cell surface, requiring interpolations between cells to find the proper values in the modelled urban space. Figure 5

![Figure 4](image_url) Scheme adopted to reconstruct the elevation of the water table in flooded areas – the flooded water elevation can be calculated as $Y_{flood} = Y_{road} + ΔE - ΔI + h$.

![Figure 5](image_url) Explanatory section of the difference between ‘flow paths’ and ‘storage ponds’.
presents an explanatory schematic cross section showing the difference between the two approaches mentioned. A flooding surface was generated by this process, with the aim of providing a reference for model calibration. The flooding surface was obtained as follows: the raster surface obtained by interpolating water elevations (obtained from the survey) was subtracted from the raster of the ground surface and only the positive portions were selected.

Figure 6 | (a) Flow directions calibration. Calibration (b) and validation (c) results for ‘storage ponds’ and ‘flow paths’.
RESULTS AND DISCUSSION

As a preliminary consideration, from Figure 6(a), it can be observed that, in general terms, the model adequately represented the flow paths between cells. The flow directions obtained from interviews and those simulated by MODCEL were consistently represented. Note, however, that the model only provides this information along its established connections, which do not overlap with all the sites for which people provided inputs. It can also be seen that sometimes a given street can present flows in two opposite directions, due to the ‘flattened egg-box’ topography. Discharge values reported in Figure 6(a) correspond to the peak flows in each link during the same event.

In the calibration process, surveyed-simulated comparisons were performed in 16 ‘storage ponds’ and 21 ‘flow paths’. Water depths in some ponds may reach a few metres. A point was considered calibrated if an absolute difference of no more than 0.15 m appeared between the surveyed point and the modelled result in the same point, if it is a storage pond. Conversely, if the point referred to a flow path, it was considered calibrated if the difference was less than 10%. The calibration results can be seen in Figure 6(b). The average difference between the modelled and surveyed value for ‘storage ponds’ was 0.069 m and for ‘flow paths’ was 9.4%.

Considering the validation process, 13 points were compared to the ‘storage pond’ criterion and 43 points to the ‘flow path’ criterion. The validation results can also be seen in Figure 6(c).

Therefore, good results were observed for the calibration and validation events. The flood calibration map obtained can be seen in Figure 7.

Figure 7 | Flood map obtained in the calibration result (equivalent to a rainfall event with a return period of 84 years).
Calibration results tend to be (at least) slightly better than validation results. Taking into account that the calibration event has used a very high rainfall intensity and was the one that had recently caused the most damage, it was expected (and apparently it was confirmed) that residents would have the most recent flood levels fresh in their memories, presenting the best records for comparison.

In order to facilitate the interpretation of the results, Figures 6 and 7 will be considered together. Note, for example, that for cell 402, which is a natural lagoon (called Boca Grande), the pond comparison gave similar values around the elevation of 13 m, with more than 1 m of water depth. On the other hand, cells 404, 406 and 405, connected by streets (thus, linked by the Saint-Venant dynamic equation), were considered to be flow paths, with water depths of about 0.30 m. The link between the neighbouring cells 402 and 404 is a weir, which justifies the difference in the water depths shown in Figure 7. This situation is the same for the other lagoons (304 – Esperanza; 102 – Mano de Dios; 603 – Salada). It is important to highlight that a cell is not a precise and detailed representation of a portion of terrain. In fact, it is a modelled element that interprets and averages the information of an area representing a portion of terrain. In this way, cell 402 is a lagoon, represented by a depth × volume curve with weirs placed at the street levels that appear inside (in the border) of this cell. In this way, when looking at the flooded area in Figure 7, the water level is referred to the cell centre (inside the lagoon) and not over the streets (representing the weirs). Mass balance is correctly accounted for in the calculation, but its visual representation is simplified. Note that in the vicinity of the Boca Grande Lagoon there are cells that are representative of typical urban areas (like cells 403 and 408, for example), with elevations similar to that of the streets that worked as weirs in the lagoon representation. In these cells no significant flooding occurred, which means that overflows were not important or even did not occur to these areas. In contrast, in the vicinity of cell 603, overflows produced flooded cells around this Salada Lagoon. This discussion is illustrated in Figure 8, which shows a partial and schematic representation of the hydraulic interaction between a lagoon and its vicinity.

The same type of analyses may be done for cells 102, 208, 304 and 603, which relate to four other lagoons in depressed ground, namely: Mano de Dios, El Patron, Esperanza and Salada.
CONCLUSIONS

Despite the lack of proper and formal information for model calibration, a field campaign based on a structured questionnaire and supported by a local complementary topographical survey and the search for watermarks on urban structures, allowed the reconstruction of the flood level of past events, recovering information that was previously unavailable.

With this information and using a quasi-2D model that could represent the local topography in an interpretative way, a calibration process was successfully held and the differences between the calculated events and the recovered information were acceptable. Validation was also considered successful. This avoided the need for a detailed DTM, which is a model with a fine topographic representation and a small grid area, which includes interference by buildings. Thus, it is possible to conclude that the model choice and the calibration procedures offer a feasible alternative to setting up a planning tool; they were sufficiently reliable even in situations with very scarce data, opening the way to finding efficient solutions to flood problems. Thus, the result of this work and the proposed methods can be used to support the calibration process of hydrodynamic models in places where there is little or no information available.

REFERENCES


Balica, S. F., Douben, N. & Wright, N. G. 2009 Flood vulnerability indices at varying spatial scales. *Water Science and Technology* 60 (10), 2571–2580.


First received 17 January 2019; accepted in revised form 7 June 2019. Available online 20 June 2019