# Analysis and Visualization of Extreme Weather Events in the City of Rio de Janeiro

**Anderson Silva [1], Thiago Moeda[2], Fabio Porto[1]**

[1]Laboratório Nacional de Computação Científica (LNCC)
Getúlio Vargas, 333 – Petrópolis – RJ – Brazil

{achaves, fabio}@lncc.br

[2]Observatório Nacional, Rio de Janeiro, Brazil

{tsantanna}@on.br

***Abstract.*** *Extreme weather events regularly occur in different locations, causing immense social, environmental and economic impact and damage. Especially in the city of Rio de Janeiro, understanding extreme events related to heavy rains is a fundamental component for the correct prediction of new phenomena, ideally resulting in models capable of predicting when, how and where they will occur. The current work proposes the analysis of rain data collected from rainfall stations positioned in the city of Rio de Janeiro, with the objective of developing a spatial representation that can be used to predict heavy rains from climate models.*

## 1. Introduction

Extreme weather events are studied in many different perspectives, due to their huge potential impact that can lead to catastrophic consequences, and their critical humanitarian, environmental and financial repercussions. These consequences are even more exacerbated by their frequent unpredictability, since even by knowing the governing equations that model an underlying meteorological system, the specific circumstances that precede and are catalyzers to the event still might be difficult to determine. This is because even simple equations can result into very complex chaotic dynamics (Farazmand and Sapsis 2019). Furthermore, extreme events are rare by definition, and in most situations few examples of similar past events are available for study (Ummenhofer and Meehl 2017), making difficult the detection of patterns and indicators for future events.

Rainfall is considered the most influential climate variable to the weather related events that cause the greatest share of worldly impact (Brito et al. 2017; Sobral et al. 2020). This is especially true in the city of Rio de Janeiro (Brazil), where its proximity to the coast and its topography create atmospheric patterns favourable to the development of sudden rainfalls with high precipitation levels that cause frequent urban flooding and landslides, as well as many casualties and socio-economical damage (de Souza et al. 2012a; Brito et al. 2017). The continuous monitoring of climatic conditions as well as the study and understanding of how such phenomena occurs is of great interest to reduce their impacts on the city as a whole.

Climate disaster risk detection and notification systems play a key role in reducing damage caused by extreme events. Such systems make use of advanced numerical prediction models in order to preemptively inform public agencies about potentially dangerous situations. In recent years, machine learning has also been one of the leading strategies extensively applied to a wide variety of problems involving big data, being used in a number of works for the detection of extreme events (Qi and Majda 2020; O'Gorman and Dwyer 2018). However, to better support system operators, visualization techniques must also be applied in order to enable investigation of patterns and trends in real time (de Souza et al. 2022).

Therefore, the current study proposes an analysis of the distribution and visualization of extreme rainfall data collected over 25 years in the city of Rio de Janeiro for the development of techniques capable of indicating in real time (nowcasting) the areas of the city of Rio de Janeiro most subject to the occurrence of extreme events at a given moment. For this, data obtained from 36 pluviometrical stations were used. The remainder of this work is divided as follows: Section 2 presents related works; Section 3 presents in detail the data used, as well as the pre-processing techniques necessary for their preparation and the steps necessary to generate a visualization mesh of such events. Finally, Section 4 presents conclusions and directions for work in progress.

## 2. Related Work

Previous works have already been dedicated to the study and analysis of rainfall patterns in the city of Rio de Janeiro and its surroundings. The Alerta-Rio (Calvello et al. 2015) system is an alert system that aims to identify risk areas and minimize potential damage caused by natural disasters. Currently, the system has information collected by 33 monitoring stations distributed throughout the city, as well as two meteorological radars.

In de Souza et al. (2012b) , data from 30 of these stations is used to detect landslide patterns in the city of Rio de Janeiro. Algorithms based on artificial neural networks (Multilayer Perceptrons) were used to fill in missing rainfall data. The work compares the performance of such networks for forecasting rainfall volume and landslides to that of algorithms based on classification rules.

Recently, Pereira et al. (2022) made use of Ward's method of hierarchical clustering in order to identify homogeneous regions with respect to rainfall distribution and natural disaster risk. The work indicates 4 of these groups for the city of Rio de Janeiro, using analysis of the rain distribution based on the values registered by 14 of the rainfall stations of the Alerta-Rio system.

Studies with the objective of analyzing rainfall patterns including regions adjacent to the city were also carried out. The work by Brito et al. (2017) carried out an analysis of the rainfall distribution similar to that of Pereira et al. (2022) in order to group 100 different stations distributed throughout the state of Rio de Janeiro. In this case, clustering based on Pearson's correlation index and simple linear regression were used to fill in missing values in the time series.

It is important to emphasize that, although some of these works are interested in the analysis of landslide events, none of them carried out the analysis of rainfall patterns with focus on extreme weather events that occurred in the city. The analysis of rainfall data with specific emphasis on extreme rain events is the aim of the current work.

## 3. Pre-Processing of Rainfall Data

The data preparation stage is one of the most important and the one that requires more time and planning in the analysis task. For the current work, data from 36 real-time monitoring stations of the Alerta-Rio system were used, of which 33 are still in operation. The dataset consists of the rainfall volume recorded at 15-minute intervals from the year 1997 to 2021.

The first challenge encountered is related to the rain timestamps provided by each rainfall station. Although the 15-minute collection interval is maintained throughout the period, for certain years, the data collected is not recorded at the same time by the different stations. Thus, standard rainfall recording schedules were defined to be received by each station. Register times were fixed at 15-minute intervals starting at 00:00 in 1997 and extending through all subsequent years. The rainfall value recorded for each station constitutes the value received by that station on the nearest timestamp.

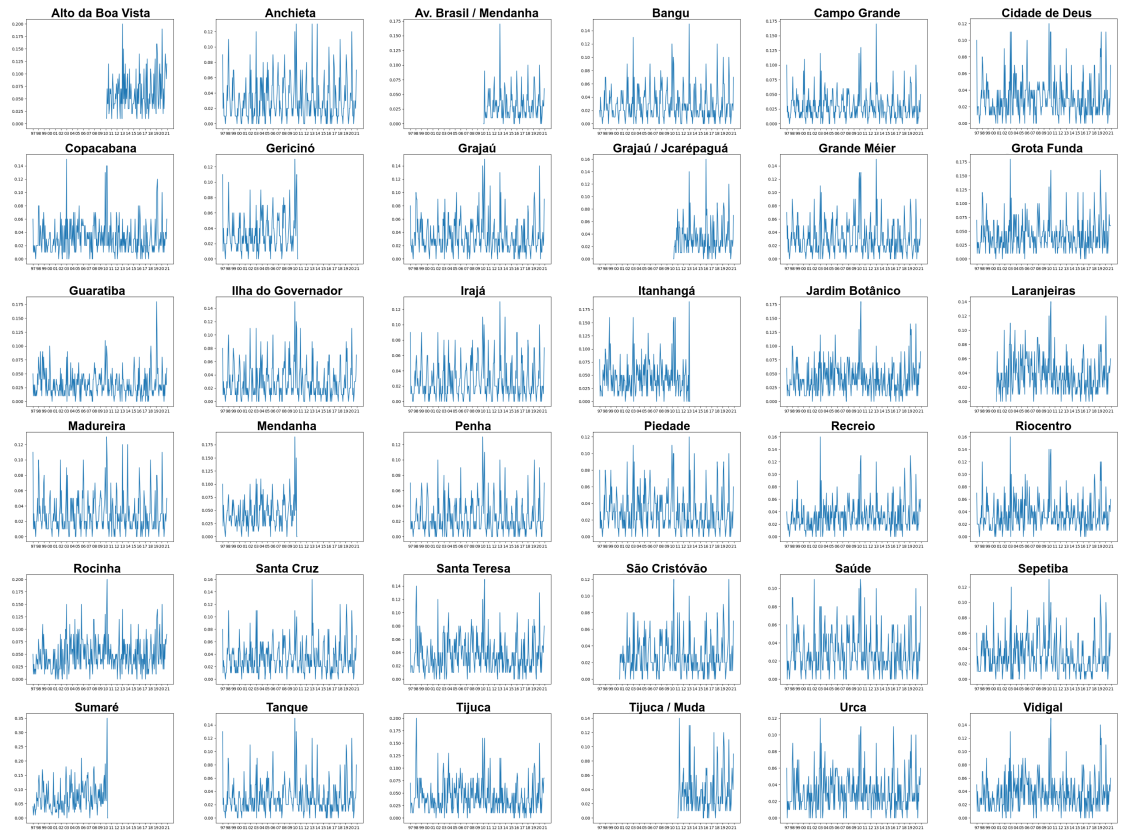


**Figure 1. Average monthly rainfall level over the years 1997 to 2021. Observations recorded based on accumulated precipitation at 15-minute intervals**

The Figure 1 presents the monthly average rainfall recorded by each season (at 15-minute intervals), over the years. It is possible to verify the periodicity in the annual rain peaks, which tend to be registered during summer. It is important to highlight that some of the stations had their operation start or end during the studied period, and for this reason they not shown in the series visualization.

For the analysis of extreme events, the work aimed to generate a three-dimensional mesh (latitude, longitude and time) representing the different rainfall values over the years. For the definition of the mesh, a rectangular region of dimensions 7x7 was established, from the coordinate (lat = –23.1339, lon = –43.8906) to the coordinate (lat = –22.6497, long = –43.0483), at regular intervals of 7.648 km for latitude and 13.397 km for longitude over the city of Rio de Janeiro, as proposed in (Porto et al. 2022). For each region of the grid, the corresponding rainfall value was defined using linear interpolation of the 3 closest stations, based on their locations. The resulting mesh can be seen in Figure 2. Linear interpolation in time was used to treat the time instants where there were gaps in the data provided by the stations.

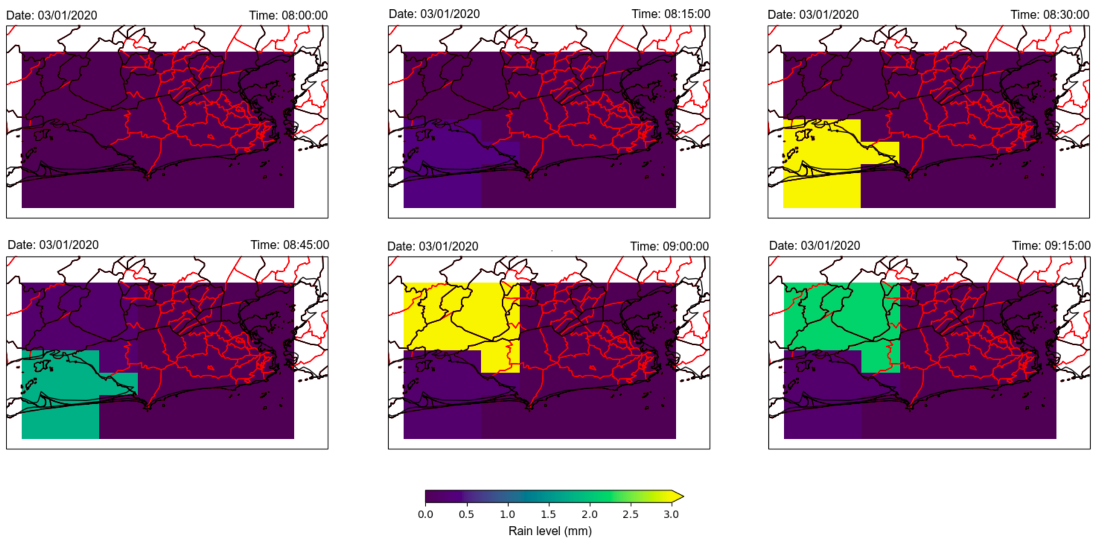


**Figure 2. Mesh generated over the city of Rio de Janeiro, showing the level of rain at different times**

From the generated data mesh, it is possible to visualize the regions most affected by extreme events. The analysis of such events is important to understand the dynamics of their formation, as well as the time and regions affected. Figure 3 shows the visualization of an extreme event that occurred over the city on March 12, 2016. The event registered heavy rain concentrated over the Tijuca Massif, causing the city to enter a state of crisis, leaving several regions flooded.In the figure, the color and circumference of each circle indicate rainfall intensity at each point.

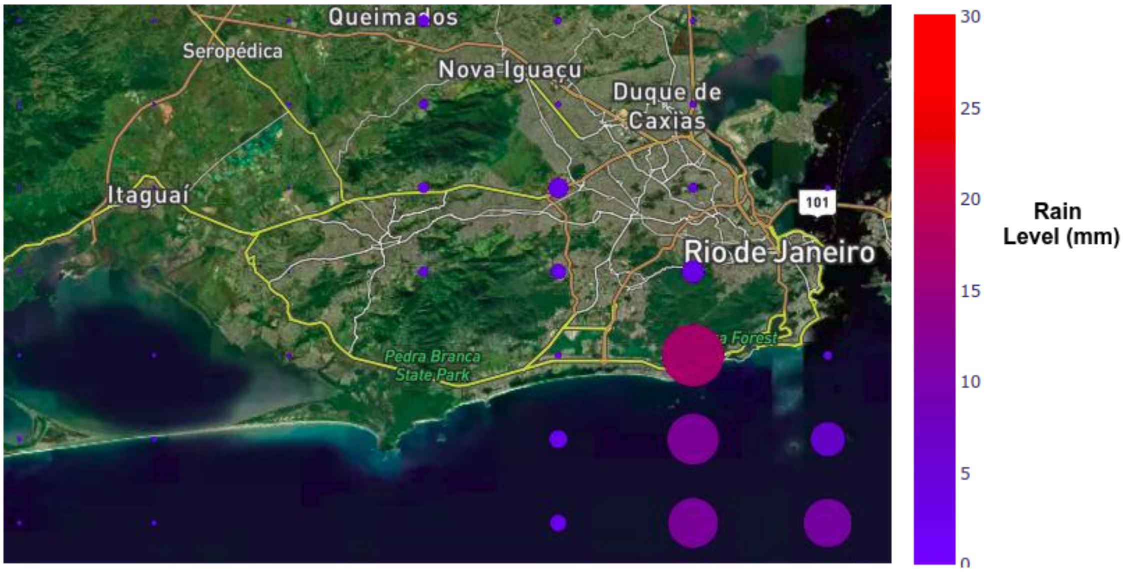


**Figure 3. Dynamic visualization on map of extreme event occurred on March 12, 2016**

## 4. Final Considerations and Research Directions

Extreme events have immense potential for impact, and the correct understanding of the circumstances in which they occur from past event data is essential to reduce huge losses. Thus, monitoring weather conditions is central to predict, detect and react quickly to such events. Given its complexity, the use of visualization techniques, as well as the combination of different modeling strategies are necessary tools to enable more precise and efficient reactions.

As research directions, we intend to develop not only a solution capable of analyzing the phenomena that occurred previously (past), but also a visualization system based on data received in real time (present). Once this objective is achieved, the next step is the development and incorporation of detection models (numerical or based on machine learning) for extreme events prediction (future).

## References

[Brito et al. 2017] Brito, T. T., Oliveira-Júnior, J. F., Lyra, G. B., Gois, G., and Zeri, M. (2017). Multivariate analysis applied to monthly rainfall over rio de janeiro state, brazil. *Meteorology and Atmospheric Physics*, 129(5):469–478.

[Calvello et al. 2015] Calvello, M., d'Orsi, R. N., Piciullo, L., Paes, N., Magalhaes, M., and Lacerda, W. A. (2015). The rio de janeiro early warning system for rainfall-induced landslides: analysis of performance for the years 2010–2013. *International journal of disaster risk reduction*, 12:3–15.

[de Souza et al. 2022] de Souza, C. V. F., Barcellos, P. C. L., Crissaff, L., Cataldi, M., Miranda, F., and Lage, M. (2022). Visualizing simulation ensembles of extreme weather events.

[de Souza et al. 2012a] de Souza, F. T., Ebecken, N., et al. (2012a). A data based model to predict landslide induced by rainfall in rio de janeiro city. *Geotechnical and geological engineering*, 30(1):85–94.

[de Souza et al. 2012b] de Souza, F. T., Ebecken, N., et al. (2012b). A data based model to predict landslide induced by rainfall in rio de janeiro city. *Geotechnical and geological engineering*, 30(1):85–94.

[Farazmand and Sapsis 2019] Farazmand, M. and Sapsis, T. P. (2019). Extreme events: Mechanisms and prediction. *Applied Mechanics Reviews*, 71(5).

[O'Gorman and Dwyer 2018] O'Gorman, P. A. and Dwyer, J. G. (2018). Using machine learning to parameterize moist convection: Potential for modeling of climate, climate change, and extreme events. *Journal of Advances in Modeling Earth Systems*, 10(10):2548–2563.

[Pereira et al. 2022] Pereira, R. M. S., Wanderley, H. S., and Delgado, R. C. (2022). Homogeneous regions for rainfall distribution in the city of rio de janeiro associated with the risk of natural disasters. *Natural Hazards*, 111(1):333–351.

[Porto et al. 2022] Porto, F., Ferro, M., Ogasawara, E., Moeda, T., de Barros, C. D. T., Silva, A. C., Zorrilla, R., Pereira, R. S., Castro, R. N., Silva, J. V., et al. (2022). Machine learning approaches to extreme weather events forecast in urban areas: Challenges and initial results. *Supercomputing Frontiers and Innovations*, 9(1):49–73.

[Qi and Majda 2020] Qi, D. and Majda, A. J. (2020). Using machine learning to predict extreme events in complex systems. *Proceedings of the National Academy of Sciences*, 117(1):52–59.

[Sobral et al. 2020] Sobral, B. S., de Oliveira-Júnior, J. F., Alecrim, F., Gois, G., Muniz-Júnior, J. G., de Bodas Terassi, P. M., Pereira-Júnior, E. R., Lyra, G. B., and Zeri, M. (2020). Persiann-cdr based characterization and trend analysis of annual rainfall in rio de janeiro state, brazil. *Atmospheric Research*, 238:104873.

[Ummenhofer and Meehl 2017] Ummenhofer, C. C. and Meehl, G. A. (2017). Extreme weather and climate events with ecological relevance: a review. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1723):20160135.