Towards a Definition for Extreme Weather Events in Rio de Janeiro City

Mariza Ferro[1,3] Eduardo Bezerra[2], Eduardo Ogasawara[2], Nilton Moraes[4] , Fabio Porto[1]

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

[2]Centro Federal de Educação Tecnológica Celso Suckow da Fonseca
Rio de Janeiro, Brazil.

[3]Instituto de Computação
Universidade Federal Fluminense – Niterói, RJ – Brazil.

[4]Sistema Alerta Rio da Prefeitura do Rio de Janeiro
Rio de Janeiro – RJ – Brazil.

{mariza,fporto}@lncc.br, eogsawara@ieee.org,ebezerra@cefet-rj.br

Abstract. *Extreme weather events is a new area of research that has recently attracted the attention of researchers from different disciplines. In this paper, we investigate the phenomena from a data-driven forecast standpoint for severe rainfall occurring in the city of Rio de Janeiro. We aim at a formal definition for the phenomena that can clarify its concept and help drive the research. Our initial result is a characterization of the problem inspired by a reformulation of a numerical model based forecast framework.*

1. Introduction

Extreme weather is responsible for many natural disasters that cause material damage and loss of life. Precipitation nowcasting, a few hours ahead forecast for extreme rainfall events, is an essential component of early warning systems and consecutive actions within crisis management and risk prevention. As the effects of climate change become more severe, more frequent episodes of extreme events are likely to be observed [Gates, 2021]. Severe events recorded across the Rio de Janeiro nearby town demonstrate the need to improve the ability to monitor, predict and issue landslide alerts in advance to minimize related damage and protect human lives and properties. However, despite the efforts dispensed to improve strong events forecast accuracy, results of the Alerta Rio [1], a system of the department of the municipality responsible for throwing alerts of extreme events needs improvements [2]. In an internal study, [daSilva, 2019] analyzed 168 rain alerts, from February 2019 to May 2019. On the total forecasts emitted by Alerta Rio classified as *strong rain*, only 12% did materialize as such. More interestingly, 35% of the these alerts corresponded to actual *no rain* observed. Conversely, considering the total forecast events for *moderate rain*, 49% had *no rain* and 2% faced *heavy rain*. Thus, there is a clear need to improve on extreme weather forecasts in urban areas, and in particular for the Rio de Janeiro city. In this regard, our ongoing research project [Porto et al., 2022] aims at

[1]http://alertario.rio.rj.gov.br/
[2]It is worth mentioning that the current monitoring is not done through predictive models with ML.

building Machine Learning (ML) predictive models for extreme precipitation forecast in urban areas, in particular in the Rio de Janeiro City. For this type of modeling with ML, it was felt the need to clearly define what an extreme event is to improve the forecasting modeling task.

Severe weather watches with a few hours lead time, if they exceed a certain threshold [Wang et al., 2017]. However, these thresholds are usually determined by local and regional factors. These are challenging in Rio since, in association with the socio-spatial processes, the interaction of physiographic aspects of the city, notably the relief (Tijuca, Pedra Branca, and Medanha massifs), the proximity to the Atlantic Ocean and two Basins (Guanabara and Sepetiba), impose risk factors concerning heavy precipitation and natural disasters. Thus a skilled local forecaster is necessary for a correct interpretation. Defining an extreme/severe weather event after its occurrence is easier. They usually involve human and material losses with high levels of precipitation, as in the examples mentioned. Analyzes of rainfall levels and measurements of observed atmospheric data help to indicate the factors that caused heavy rain.

The objective of this work is to make a study that helps in the definition of what is an extreme precipitation event in the city of Rio de Janeiro. Furthermore, two possible definitions are analyzed:

- What defines an extreme rainfall event to be considered in a training database of ML algorithms for nowcasting extreme precipitation. For this case, we present a literature review and historical databases which could support this definition;
- What defines an extreme event for the classification and regression model with ML, allowing the nowcaster to identify with a certain probability where and when an intense rain with the potential for floods and/or landslides will occur.

2. Extreme Weather Events in Rio de Janeiro

Extreme rainfall events have always occurred in the city of Rio de Janeiro, especially in the spring and summer months, causing great inconvenience to the local population. This section presents a review of the extreme weather events that hit the city of Rio de Janeiro. The objective of this study is to gather elements for a definition of what an extreme event is for the city of Rio, enabling ML approaches for a future nowcasting decision-making process, under development. It is worth mentioning that the events studied here delimit the problem for precipitation events with rainfall. Other events such as winds and coastal events will not be considered, not even droughts, which are also related to precipitation.

Luiz-Silva [2022] provide an overview of the main climatological characteristics of precipitation extremes throughout the State of Rio de Janeiro. Daily precipitation data are derived from 56 rainfall stations during the second half of the twentieth century and the 2000s. Also, the data for their extreme events study must include the occurrence of flood and landslide events, weighing at least one dead. They indicate and examine eight indices related to extreme precipitation, similar to those recommended by the Expert Team on Climate Change Detection and Indices of the Commission for Climatology of the World Meteorological Organization (WMO). Indices are mainly based on daily and annual data of maximum and minimum precipitation to identify trends. Specifically for the city, the climatological values of rainfall above 30mm, which exceed 20 days in some parts of the Guanabara Bay basin, particularly around the Tijuca massif in the city of Rio de Janeiro, are highlighted. The topography and significant total precipitation in short periods contribute considerably to landslides.

Dereczynski et al. [2017] present a survey of extreme rainfall events and their effects on the population of the city of Rio de Janeiro, carried out from 1881 to 1996. The objective is to recover the main events of extreme rainfall from the 19th century by assessing heavy rainfall reports and evaluating the usefulness of meteorological information to alert the population against disasters caused by heavy rains. First, they pre-selected cases with daily rainfall of 100 mm or greater. Of the total of 100 pre-selected, Eighty-two cases were considered extreme events due to the significant damages caused to the population, including deaths, injuries, material damages, landslides, and floods. Throughout the analyzed period, the weather forecasts in the newspapers were not sufficient to alert the population in cases of heavy rains.

d'Orsi et al. [2017] survey the 50 accidents of a geological-geotechnical nature, most impacting the city of Rio de Janeiro between 1966 and 2016. This work considers the GeoRio database, news published in the media, magazines, articles, and scientific papers. The fifty accidents consider information and characteristics of the occurrences, such as the number of fatal victims, buildings totally and/or partially destroyed, precipitation volume, disturbances generated, obstruction of roads, etc. From there, they developed a methodology to score each "Accident Category" according to its magnitude. Based on this total score, accidents were grouped into three classes which indicate the severity and potential of accidents to generate inconvenience and impacts for the city. Class I represents the least expressive accidents, and Class III is the most expressive. For example, among the 16 class-III events registered in the event on April 5, 2010, in *Morro dos Prazeres* had landslides, 30 deaths, obstruction of roads and houses destroyed.

Aires et al. [2017] show that the Brazilian Regional Atmospheric Modeling System (BRAMS) is a useful tool for forecasting heavy rains. Simulations were performed with BRAMS to understand the extreme event that occurred on April 7, 2010, in Rio and Niterói (a neighborhood of Rio de Janeiro, 14km away and located in the Guanabara Bay basin, known as Morro do Bumba). Intense rainfalls with more than 300mm in 24h caused the landslip inflicting 50 deaths in Niteroi and 66 in Rio, with more than 11439 people displaced in addition to several impacts on the cities infrastructure. The simulations use synoptic analyses and satellite images provided by CPTEC. The work also presents a study of the literature highlighting factors, from an observational point of view, of cases of heavy rains that occurred in the State and the city of Rio de Janeiro. For the city, observations use images from radar - Pico do Couto, GOES-16 satellite, and a radiosonde in Galeao airport. Among the atmospheric systems responsible for the analyzed events, the following can be highlighted: dynamic and thermodynamic mechanisms associated with these systems, which prove to be useful for understanding and predicting extreme events.

Climatological characterization of the precipitation extremes from these works is mainly based on observed meteorological indices and association with hydrometeorological impacts, like landslides and deaths. These works bring us elements to characterize and define extreme events for the database definition. However, for the learning model, what really matters seems to be able to learn the conditions that precede the extreme event and have models that can advance in time with the precipitation variable. As presented in [Aires et al., 2017], dynamic and thermodynamic mechanisms associated with these systems from observational data from radar and satellite data seem to be the most important aspect. Next sections aims to study approaches and frameworks focused in this characterization for the learning model.

3. An Extreme Events study framework

From the history of extreme weather events reported in the city of Rio de Janeiro, as well as those occurring in other urban areas, it becomes clear that instruments that could mitigate the effects of such events are paramount, particularly in a climate change scenario we have been observing, with an increase in the frequency of occurrences of very strong rainfall. In an attempt to specify a framework within which extreme events can be reasoned about, [Farazmand and Sapsis, 2019] have structured the problem into inter-correlated components, as depicted in Figure 1.

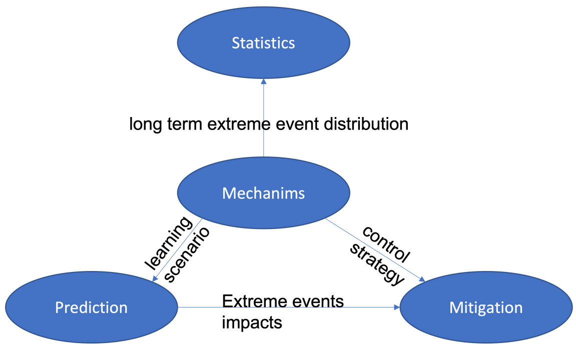

Figura 1. Extreme events problem components inter-correlations, adapted from [Farazmand and Sapsis, 2019].

In Figure 1, extreme event phenomena can be studied from four different perspectives: statistics, mechanisms, real-time prediction and mitigation. The *statistics* and *real-time prediction* components refer to approaches to infer the occurrence of an extreme-event, while the *mechanisms* one explores techniques for modeling the ensemble of phenomena that gives rise to extreme-events; and *mitigation* is about actions and strategy to manage their effects over the cities and their population. In this work we are mostly interested in data-driven techniques that support extreme-events predictions, which drives us to discuss the *prediction* and *statistics* components. However, understanding the conditions (i.e. state transitions) that precede an event is essential for the design of a learning strategy and, ounce the predictions have been made, authorities can plan actions to mitigate their effects. Thus, the proposed framework sets an interesting panorama for the study of extreme events.

3.1. Exploring the extreme events framework

Considering the three framework components, presented in Figure 1 and of relevance to this work, in this section, we discuss them in light of our interest for the data-driven prediction of extreme events.

Statistical approaches infer the frequency and probability of an extreme event occurrence on a large distribution of events. The *Extreme Event Theory* is one such approach that has a large body of work, for instance [Haans and Ferreira, 2006, Watson, 1954, Lucarini et al., 2016]. The theory aims at computing $M_n = \max(X_1, X_2, ..., X_n)$, the set of maximum values occurring in a distribution, where $X_1, X_2, ...$ is a sequence of random variables, and $n \rightarrow \infty$. Although developed considering independent and identically distributed (i.i.d) values, the approach has been extended to cope with time series data [Mudelsee, 2010]. Another important statistics approach in this context is the *large deviation theory*, [Touchette, 2011]. Due to the large number of extreme events' occurrence needed to support their prediction, the statistics approach can be used as the basis for governmental policies dealing with extreme rainfall in the city, but it is not a candidate for nowcasting of rainfall events [Buch-larsen et al., 2005] as the latter requires a precise forecast about their future occurrences.

The *mechanism* component considers the context that triggers the occurrence of an extreme event associated with dynamic and thermodynamic effects. Through the study of such mechanisms, one should be able to understand the conditions that may entail a transformation from an observable state with typical values into an extreme event (EE) state. Examples of known mechanisms for generating extreme events include:(i) multiscale systems,(ii) homoclinic and (iii) heteroclinic bursting and (iv) noise-induced transitions, [Farazmand and Sapsis, 2019]. A main aspect of the study of mechanisms underlying the development of extreme events is the discovery of indicators. The latter are observables of type $f : \mathcal{U} \rightarrow \mathbb{R}$, representing some quantity in a domain space $\mathcal{U}$, whose observation along a trajectory $u(t)$ signals the occurrence of a set of extreme values $f(u_t) \geq f_e$, for time instants $t_0 \leq t \leq t_n$. Thus, f_e defines a threshold for the boundary between typical and extreme-values space for f. In the case of weather forecast, f could model the precipitation observations. An extreme event threshold value could be $f_e = 50$ mm/h, referring to the value of precipitation in one hour. One may observe that the discovery of indicator functions concept is in line with the selection of features that shall be learned by a machine learning forecasting model [Xiang et al., 2020]. In this sense, a feature (or set of therein) determines the behavior to be learned in predicting a target distribution (i.e. a random variable). It is interesting to note that for learning tasks, being able to understand the complex scenario preceding an extreme event is the basis for an accurate prediction. In a data-driven prediction scenario, this entails learning the signals that indicate the conditions favorable for the rise of an extreme event. In this context, the definition of an exact extreme event threshold, such as f_e above, may not be as relevant as learning the last minute signals before the burst of the extreme event.

Finally, by analyzing the predictions for the extreme rainfall, indicating the volume of rain, the affected spatial region and the expected timing, authorities can more accurately plan for actions to mitigate their effects.

4. Final Considerations

Extreme rainfall events are becoming more frequent and producing extended damage to the population in urban areas. In this paper, we focus on extreme weather events occurred in the city of Rio de Janeiro. We investigated the current literature describing these phenomena and the one proposing methods for nowcasting extreme rainfall events. In an attempt to find a formal definition for the phenomena, we characterize it according to four main approaches that reason about extreme weather events prediction and their effects. Given our focus on forecasting, it seems that the *mechanisms* and *prediction* components form the basis of a framework for tackling the problem. Moreover, fixing a threshold for values delimiting extreme events is not necessary for solving the prediction problem.

Referências

M. Aires, J. L. F. de Oliveira, J. M. de Castro Junior, M. M. F. de Oliveira, and N. F. F. Ebecken. Numerical simulation of the atmosphere related to landslides triggered by heavy rainfalls. *Engevista*, 19(1):37–64, 2017.

T. Buch-larsen, J. P. Nielsen, M. Guillén, and C. Bolancé. Kernel density estimation for heavy-tailed distributions using the champernowne transformation. *Statistics*, 39(6): 503–516, 2005. doi: 10.1080/02331880500439782.

F. daSilva. Projeto pesquisa operacional, 2019. Internal Report, in PT.

C. Dereczynski, R. Calado, and A. Barros. Extreme rainfall in the city of rio de janeiro: History from the 19th century. *Anuário do Instituto de Geociências - UFRJ*, 40:17–30, 09 2017. doi: 10.11137/2017_2_17_30.

R. N. d'Orsi, N. M. Paes, M. A. Magalhães, R. da Silva Coelho, L. R. da Silva Junior, and L. R. S. Valente. Os 50 maiores acidentes geolÓgico-geotÉcnicos na cidade do rio de janeiro entre 1966 e 2016. `http://alertario.rio.rj.gov.br/wp-content/uploads/2017/01/PDF_ESTRUTURA-DO-LIVRETO_50-MAIORES-ACIDENTES-_A5_.pdf`, 2017. Accessed: 2022-06-24.

M. Farazmand and T. Sapsis. Extreme events: Mechanisms and prediction. *Applied Mechanics reviews*, 71(5), Sept. 2019. doi: https://doi.org/10.1115/1.4042065.

B. Gates. *How to avoid a Climate Disaster: The Solutions We Have and the Breakthroughs We Need*. Random House Large Print Publishing, 2021.

L. Haans and A. Ferreira. *Extreme Value Theory An Introduction*. Springer, 2006. ISBN 978-0-387-34471-3.

V. Lucarini, D. Faranda, A. de Freitas, J. de Freitas, J. Holland, T. Kuna, M. Nicol, M. Todd, and S. Vaienti. *Extremes and Recurrence in Dynamical Systems*. Wiley, 2016. ISBN 978-1-118-63219-2.

A. Luiz-Silva, W.and Oscar-Júnior. Climate extremes related with rainfall in the state of rio de janeiro, brazil: a review of climatological characteristics and recorded trends. *Nat Hazards*, (10), 2022. doi: https://doi.org/10.1007/s11069-022-05409-5.

M. Mudelsee. *Extreme Value Time Series*. Springer, 2010.

F. Porto, M. Ferro, E. Ogasawara, T. Moeda, C. D. Tenorio de Barros, A. Chaves Silva, R. Zorrilla, R. Silva Pereira, R. Nascimento Castro, J. V. Silva, R. Salles, A. J. Fonseca, J. Hermsdorff, M. Magalhães, V. Sá, A. A. Simões, C. Cardoso, and E. Bezerra. Machine learning approaches to extreme weather events forecast in urban areas: Challenges and initial results. *Supercomputing Frontiers and Innovations*, 9(1):49–73, May 2022. doi: 10.14529/jsfi220104.

H. Touchette. A basic introduction to large deviations: Theory, applications, simulations. 2011. doi: 10.48550/ARXIV.1106.4146.

Y. Wang, E. Coning, A. Harou, W. Jacobs, P. Joe, L. Nikitina, R. Roberts, J. Wang, J. Wilson, A. Atencia, B. Bica, B. Brown, S. Goodmann, A. Kann, P.-w. Li, I. Monterio, F. Schmid, A. Seed, and J. Sun. *Guidelines for Nowcasting Techniques*. 11 2017. ISBN 978-92-63-11198-2.

G. S. Watson. Extreme values in samples from m-dependent stationary stochastic processes. *The Annals of Mathematical Statistics*, 25(4):798–800, 1954.

Y. Xiang, J. Ma, and X. Wu. A precipitation nowcasting mechanism for real-world data based on machine learning. *Mathematical Problems in Engineering*, 2020:1–11, 11 2020. doi: 10.1155/2020/8408931.