Recurrent Neural Networks and Geostatistics Applied to the Prediction of Severe Rainfall Events and Anomaly Detection

Débora Santos Rodrigues Angélica Nardo Caseri
Sinésio Pesco

Pontifical Catholic University of Rio de Janeiro

Intense rainfall in the city of Rio de Janeiro

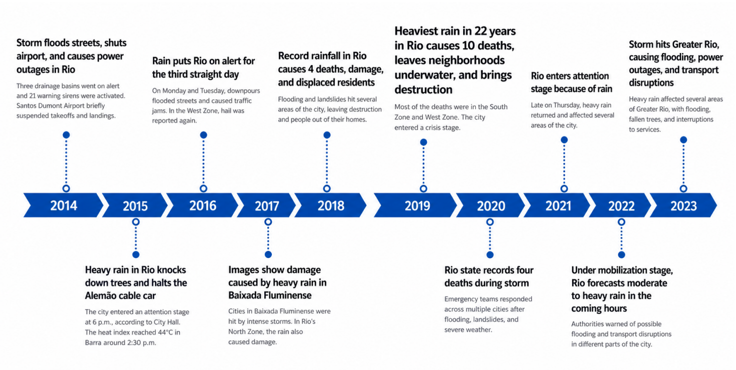


Figura: Timeline of intense rainfall events in the City of Rio de Janeiro from 2014 to 2023, highlighting major news reports related to flooding, landslides, power outages, transport disruptions, and emergency alerts. Source: Adapted from G1, 2014–2023

Rainfall event

Continuous precipitation period lasting at least one hour and with an intensity equal to or greater than 1 millimeter.

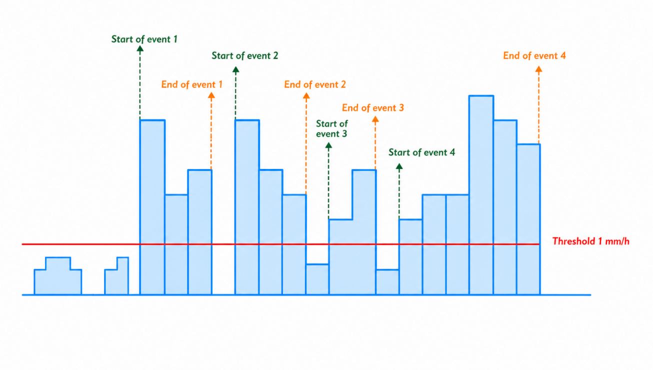


Figura: Hourly precipitation time series, illustrating the identification of rainfall events based on the 1 mm/h threshold.

Set of severe rainfall events

Event	Max (mm/h)	Min (mm/h)	Duration (h)	Accum. (mm)
1	39.4	1.4	11	79.4
2	56.6	1.0	11	134.6
3	66.0	2.0	8	61.6
4	97.4	1.6	9	244.4
5	39.8	1.2	12	97.2
6	57.2	3.4	8	152.6
7	31.6	2.6	22	111.6
8	56.6	1.8	10	125.2
9	97.2	1.8	7	204.0
10	46.6	1.2	13	159.2
11	23.4	1.2	12	68.4
12	64.4	1.0	7	115.8
13	30.6	1.4	7	59.4

Figura: Set of 13 severe to extreme rainfall events selected based on statistical criteria, including the calculation of the third quartile.

Ordinary Kriging

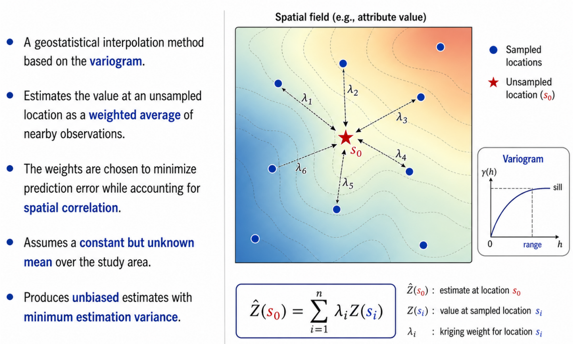
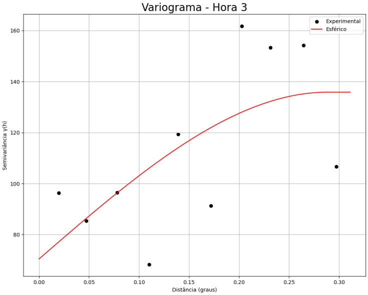
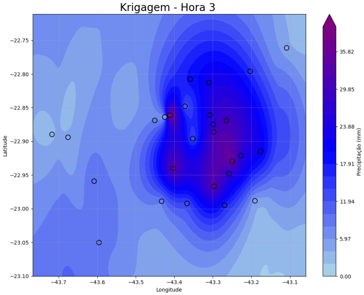
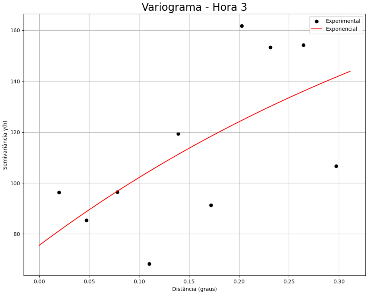


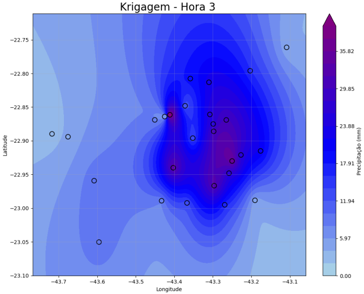
Figura: Z represents the spatial variable of interest, such as rainfall. The variogram $\gamma(h)$ is used to model spatial dependence and calculate the kriging weights λ_i.

Result: Spherical map vs. Exponential map







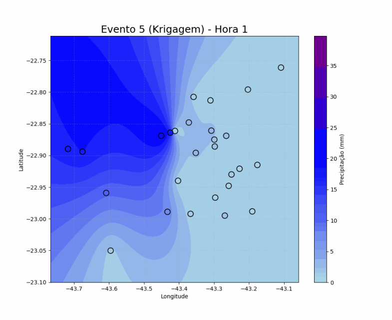


Leave-One-Out cross-validation applied to kriging

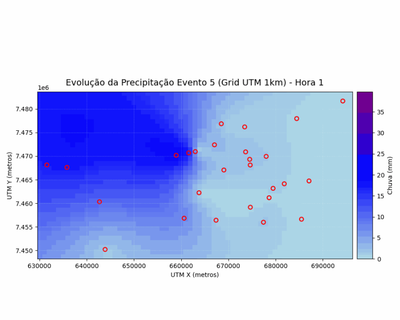
- Leave-One-Out Cross-Validation is a statistical technique used to assess the predictive performance of models, including geostatistical models such as kriging.
- In this approach, each observation is temporarily removed from the dataset, and the kriging model is fitted using the remaining observations.
- The removed observation is then estimated by the model, and the predicted value is compared with the observed value to evaluate the prediction error.

Rainfall events: Kriging vs. UTM projection

Evento 5



Grid UTM 1 km



Temporal sliding windows (lags)

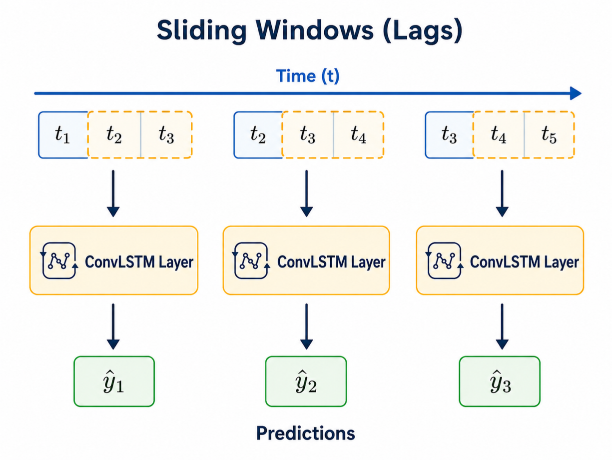


Figura: Sliding windows allow the model to use previous rainfall maps as input. In this study, we tested windows from one to four previous maps to forecast the next rainfall map.

Recurrent Neural Networks (RNNs)

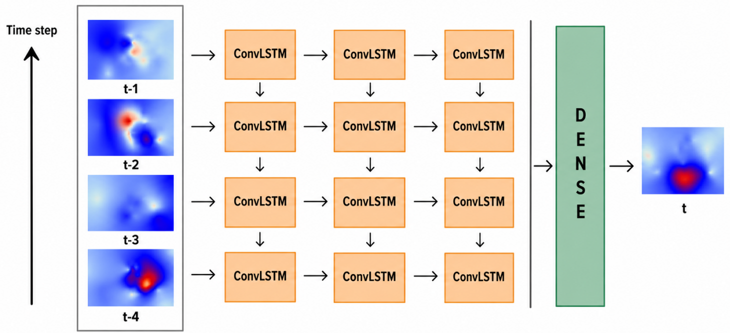


Figura: RNNs are neural networks designed for sequential data. ConvLSTM is a variant that incorporates convolutions to simultaneously capture temporal dependencies and spatial patterns, making it suitable for sequences of images or maps.

Spatiotemporal rainfall forecasting

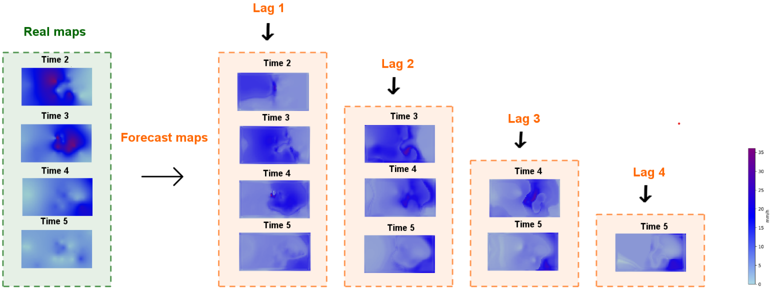


Figura: Comparison between observed precipitation maps and forecast maps for different prediction lags. The observed maps show the rainfall fields from Time 2 to Time 5, while the forecast maps represent the model predictions for lags 1 to 4. The color scale indicates precipitation intensity in mm/h.

Persistence Model

- The persistence model, also known as the naive method, assumes that the current rainfall field will remain unchanged in the next time step.
- In this approach, the forecast is simply given by the most recent observed rainfall map.
- Although simple, this method is commonly used as a baseline to evaluate the performance of more complex forecasting models.

Validation Metrics

- **RMSE** (Root Mean Squared Error) quantifies the average magnitude of the errors between the observed values (y_{true}) and the predicted values (y_{pred}).

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{\text{true}} - y_{\text{pred}})^2}$$

- RMSE is useful when the goal is to penalize large errors, such as in the forecasting of extreme weather events.

Next steps

- Forecasting using different neural network architectures.
- Probabilistic forecasting using ensemble methods.
- Combining satellite and rain gauge data.

Thank you!