ACCUMULATED RAINFALL AND POST-RAINFALL RECOVERY FOR NOWCASTING UNSUITABLE BATHING CONDITIONS

Tainah Frota Carvalho[1]; Iran Eduardo Lima Neto[2]; Newton Célio Becker de Moura[3];Francisca Dalila Menezes Vasconcelos[4]; Matheus Castelo Branco de Araújo[5]

key-words – Bathing water quality; Accumulated rainfall; Nowcasting.

INTRODUCTION

Waterfront recreational spaces play an important role in city dynamics considering their leisure function, which consequently also configures an economic function, especially in touristic areas. The fate and transport of Fecal Indicator Bacteria (FIB), such as *Escherichia coli* (*E. coli*) and *Enterococcus* spp., in these environments are influenced by local and regional hydrometeorology, including rainfall, tidal forces, sediment re-suspension, wave activity and river flow (Zimmer-Faust *et al*., 2021). In tropical countries such as Brazil, peak contamination levels are higher than those reported in studies from high-income countries with wastewater treatment plants (Lloyd *et al*., 2026). In addition, in such urban centers, these recreational activities often conflict with incomplete sewage coverage, resulting in water pollution during rainy periods (Ribeiro *et al*., 2024).

FIB are commonly used as indicators of water quality (Verhougstraete *et al*., 2020). However, this method is costly and slow, typically requiring 18–24 h to yield results, which delays necessary public health advisories (Li and Zhu, 2026; Guo and Lee, 2021). In Fortaleza, Brazil, bathing water quality classification is carried out by SEMACE[6] based on the criteria established by CONAMA Resolution No. 274/2000[7], using a retrospective series of microbiological samples. Although appropriate for regulatory purposes, these bulletins do not necessarily reflect current water conditions, especially after intense rainfall events.

Given this context, this study asks whether routine bathing water quality and rainfall data can be transformed into a simple, interpretable, and operational nowcasting system to estimate the daily risk of unsuitable bathing conditions at Fortaleza's urban beaches. The study aims to develop

[1]) Afiliação: Universidade Federal do Ceará, tainahcarvalhoarq@gmail.com
[2]) Afiliação: Universidade Federal do Ceará, iran@deha.ufc.br
[3]) Afiliação: Universidade Federal do Ceará, newtonbecker@ufc.br
[4]) Afiliação: menezes.dalila@gmail.com
[5]) Afiliação: branco.castelo@outlook.com
[6]) SEMACE: Superintendência Estadual do Meio Ambiente do Ceará, environmental agency responsible for monitoring and publishing bathing water quality bulletins in the state of Ceará.
[7]) CONAMA: Conselho Nacional do Meio Ambiente, Brazilian national environmental council responsible for establishing environmental standards and regulations, including Resolution No. 274/2000, which defines bathing water quality criteria.

and evaluate a predictive methodology based on historical bathing water quality data from SEMACE and rainfall data from FUNCEME, with particular attention to accumulated rainfall and post-rainfall recovery. The proposed approach seeks to support the creation of a new predictive layer in Fortaleza's Climate Risk Observatory, contributing to the timely identification of sanitary risks following rainfall events.

METHODOLOGY

The analytical dataset was structured in a longitudinal format, with each of the 33 bathing water monitoring points distributed along Fortaleza's approximately 34 km coastline observed in each monitoring week as the unit of analysis. The response variable corresponded to the bathing water quality classification published by SEMACE, coded in binary format for modeling purposes. This structure was associated with rainfall indicators derived from FUNCEME[8], calculated across four antecedent windows prior to sampling: 24 h, 48 h, 72 h, and 7 days. The precipitation variables were intended to represent the effect of recent and accumulated rainfall events on water quality. This strategy follows the logic of synchronizing water quality data with antecedent hydrometeorological variables, similar to that adopted in studies on post-rainfall recovery and bathing water nowcasting.

An Extreme Gradient Boosting (XGBoost) classifier was used to estimate the probability of unsuitable bathing conditions. This algorithm was selected because it can capture non-linear relationships and interactions between accumulated rainfall, critical rainfall events, post-rainfall recovery time, and the spatial identity of each monitoring point. The model was trained and evaluated using a temporal split, ensuring that predictions were tested on observations from periods subsequent to those used for training. To support a preventive public health application, the probability threshold was selected using the F2-score criterion, which assigns greater importance to the recall of unsuitable conditions and, therefore, to reducing false-negative predictions.

RESULTS

The analysis used 15,706 records with valid bathing water quality classifications from a total dataset of 20,658 observations. To ensure comparability across the rainfall windows, each rainfall variable was divided into the same four ordinal groups (Q1 to Q4), representing progressively lower to higher antecedent rainfall conditions. Each quartile included approximately 3,926 to 3,927 observations. The association between rainfall and unsuitable bathing conditions increased as the temporal window expanded: Pearson correlation coefficients rose from 0.09 at 24 h to 0.12 at 48 h, 0.16 at 72 h, and 0.23 at 7 days, while Spearman correlations increased from 0.19 to 0.28. Although the association remained modest when rainfall was considered in isolation (Spearman's $\rho = 0.28$ for the 7-day window), it was consistent across all temporal windows and should not be interpreted as a standalone measure of predictive performance. Rather, it indicates that accumulated rainfall provides a relevant hydrometeorological signal for identifying periods in which sanitary risk increases after rainfall events. This signal was subsequently integrated with the spatial identity of monitoring points to support operational prediction.

[8]) FUNCEME: Fundação Cearense de Meteorologia e Recursos Hídricos, a state institution responsible for meteorological and hydrological monitoring in Ceará, including rainfall data used for environmental and climate-related analyses.

Post-rainfall recovery was assessed considering a critical event defined as accumulated rainfall of at least 10 mm over seven days. The proportion of unsuitable bathing conditions was highest on the day of the event (61.9%), decreasing to 42.3% within 1–3 days and to 29.3% after 4–7 days. In the following periods, unsuitable conditions remained close to the background level, ranging from 29.9% to 35.2%. This pattern indicates that the sanitary risk associated with rainfall is not limited to the immediate runoff period and supports the inclusion of post-event recovery variables in the nowcasting framework.

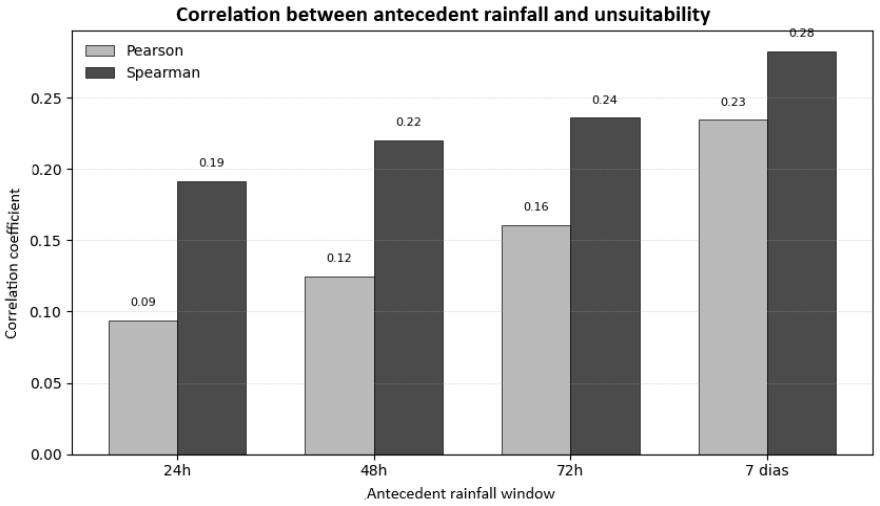


Figure 1. correlation between antecedent rainfall and unsuitability. Source: authors.

The quartile analysis reinforced this pattern. Across all rainfall windows, bathing water suitability declined progressively from Q1, representing the lowest antecedent rainfall conditions, to Q4, representing the highest. Suitability decreased from 63.3% to 37.8% in the 24 h window, from 62.7% to 36.5% in the 48 h window, from 65.2% to 34.8% in the 72 h window, and from 67.1% to 34.0% in the 7-day window. The discrimination metrics further supported the selection of 7-day accumulated rainfall as the main predictor: ROC-AUC[9] increased from 0.60 for the 24 h window to 0.66 for the 7-day window, while PR-AUC[10] rose from 0.54 to 0.61. Although the largest absolute difference between the highest and lowest rainfall quartiles occurred in the 48 h window, the 7-day window provided the most consistent overall evidence, combining the highest correlation coefficients with the best discrimination metrics. Therefore, 7-day accumulated rainfall was adopted as the central variable for the analysis of critical rainfall events, post-rainfall recovery, and predictive modeling of unsuitable bathing conditions.

While accumulated rainfall alone explained only part of the observed variation in bathing water conditions, its integration with sector and monitoring point generated robust operational performance. The final XGBoost model combined 7-day accumulated rainfall, sector, and monitoring point, using a temporal test set (n = 1,089). The model achieved an accuracy of 0.713, balanced accuracy of 0.718, recall of 0.951 for unsuitable conditions, ROC-AUC of 0.857, and

[9] ROC-AUC is the area under the ROC curve. This metric evaluates the model's ability to distinguish between two classes (in this case, appropriate and inappropriate) across various decision thresholds. The higher the ROC-AUC, the better the model separates risky situations from non-risky ones.

[10] PR-AUC is the area under the Precision-Recall curve. It evaluates the relationship between the model's ability to identify positive cases and the precision of those alerts. It is particularly important when the focus is on detecting risk events.

PR-AUC of 0.846. Under the preventive threshold selected using the F2-score, the model correctly identified 507 unsuitable observations and produced only 26 false-negative predictions. These results show that the purpose is not to attribute unsuitable bathing conditions exclusively to rainfall, but to use rainfall as a trigger for an event-driven sanitary risk process whose local expression varies across the coastline. This constitutes an initial step toward tools capable of anticipating cascading sanitary impacts following intense rainfall events and, in future developments, supporting the interpretation and prediction of territorial vulnerability in Fortaleza's Climate Risk Observatory.

CONCLUSIONS

- Seven-day accumulated rainfall showed the most consistent association with unsuitable bathing conditions, outperforming the 24 h, 48 h, and 72 h antecedent rainfall windows in correlation and discrimination metrics.
- Following critical rainfall events (≥10 mm accumulated over seven days), unsuitable bathing conditions decreased from 61.9% on the event day to 29.3% after 4–7 days, indicating a measurable post-rainfall recovery pattern.
- The results suggest that rainfall effects on coastal water quality extend beyond immediate runoff, supporting the inclusion of post-rainfall recovery patterns in operational monitoring strategies.
- The proposed approach uses routine bathing water quality and rainfall data and may support a predictive layer in Fortaleza's Climate Risk Observatory, strengthening preventive public health alerts for beach users.
- The proposed framework should be understood as an initial operational step for anticipating sanitary risks triggered by rainfall events. Future work should test the model under extreme rainfall conditions, incorporate tidal, hydrodynamic, and territorial vulnerability variables, and evaluate its capacity to support risk interpretation and prevention at the municipal scale.

REFERENCES

LLOYD, S. D.; POPOVIC, G.; OSMOND, P.; ROSER, D. J.; KHAN, S. J. (2026). "*Swimming in urban estuaries: Understanding stormwater contamination events and recovery from historical data*". Water Research, v.297, 125686.

RIBEIRO, H. M. C.; FREITAS, R. C.; RODRIGUES, C. V. A.; BEZERRA, F. C. M.; MARÇAL, C. S.; ANAISSE, C. C. R. (2024). "*Water Quality and Environmental Challenges: A Study at Chapéu Virado and Farol Beaches, Mosqueiro Island, Pará-Belém*". Revista de Gestão Social e Ambiental, v.18, n.3, p.1–11, e04436.

VERHOUGSTRAETE, M. P.; POGREBA-BROWN, K.; REYNOLDS, K. A.; LAMPARELLI, C. C.; SATO, M. I. Z.; WADE, T. J.; EISENBERG, J. N. S. (2020). "*A critical analysis of recreational water guidelines developed from temperate climate data and applied to the tropics*". Water Research, v.170, 115294.

LI, P.; ZHU, Z. (2026). "*STGCN-WQ: A new Spatio-Temporal Graph Convolutional Network for predicting beach Water Quality*". Environmental Modelling and Software, v.195, 106731.

GUO, J.; LEE, J. H. W. (2021). "*Development of Predictive Models for 'Very Poor' Beach Water Quality Gradings Using Class-Imbalance Learning*". Environmental Science & Technology, v.55, p.14990–15000.

ZIMMER-FAUST, A. G.; STEELE, J. A.; XIONG, X.; STALEY, C.; GRIFFITH, M.; SADOWSKY, M. J.; DIAZ, M.; GRIFFITH, J. F. (2021). "*A Combined Digital PCR and Next Generation DNA-Sequencing Based Approach for Tracking Nearshore Pollutant Dynamics Along the Southwest United States/Mexico Border*". Frontiers in Microbiology, v. 12, 674214. https://doi.org/10.3389/fmicb.2021.674214.