

## Abstract Template

*Rohmer Jeremy*  
[j.rohmer@brgm.fr](mailto:j.rohmer@brgm.fr)

### **Early-warning system for cyclone-induced wave overtopping aided by a suite of random forest approaches**

Authors : Rohmer Jeremy<sup>1</sup>, Lecacheux Sophie<sup>1</sup>, Pedreros Rodrigo<sup>1</sup>, Idier Deborah<sup>1</sup> & Bonnardot François<sup>2</sup>

<sup>1</sup> BRGM, 3 av. C. Guillemin - 45060 Orléans Cedex 2 – France

<sup>2</sup>Direction Régionale de Météo-France pour l'Océan Indien, 50 boulevard du Chaudron, BP 4, 97491 Sainte-Clotilde cedex

**Proposed topic:** Maths & hydroinformatics

**Key words:** Tropical cyclones; Uncertainty; Flooding; Forecast; Machine Learning.

**Speaker:** (Rohmer/ [j.rohmer@brgm.fr](mailto:j.rohmer@brgm.fr))

**Preference for an Oral communication**

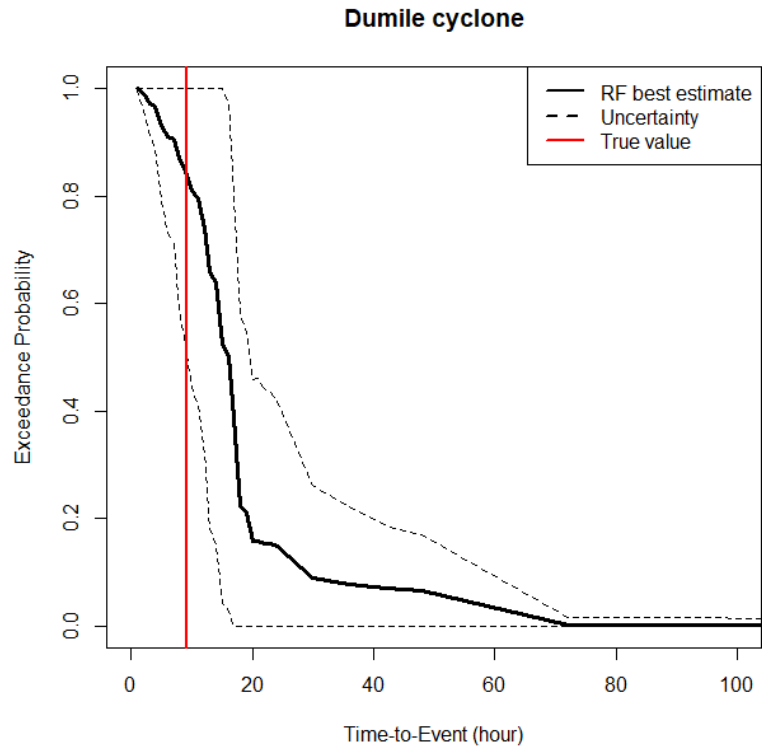
Full process-based hydrodynamic models enable to reproduce with high fidelity the complexity of cyclone-induced wave dynamics as well as their potential impact in terms of overtopping-induced inundation. Yet, their major drawback is the high computational time cost (typically of several hours), which prevent them from a direct integration in a real-time early-warning system to support crisis / emergency management.

A possible alternative relies on the statistical exploitation of pre-calculated simulation results to build a fast (low computation time cost) predictor given the offshore conditions (wave and sea level). In the present study, we present a suite of random forest (RF) techniques for fast prediction of key indicators to support forecasting of cyclone-induced wave overtopping, namely: 1) the occurrence likelihood of the overtopping event using a RF-based classification method; 2) the maximum cumulative volume overtopping using a RF-based regression method; 3) the starting time of the overtopping event (referred to as time-to-event) using a multi-output RF-based classification method. We apply the technique at Sainte-Suzanne city at Reunion Island (Indian Ocean basin) using a database of simulation results, which relate offshore conditions (wave and sea level) induced by ~500 synthetic cyclones and their consequences in terms of wave overtopping, namely the time evolution of the cumulative overtopping volume in the vicinity of a key spatial location for emergency management, namely the fire station.

Through an extensive cross-validation exercise, we show the high performance of the RF models with respect to the prediction of the three indicators. More specifically, the accuracy and the area under the ROC curve (AUC) of the RF-based classifier reaches values above 95%; the R-squared of the RF-based regression model reaches values above 80%; the time-averaged accuracy of the multi-output RF-based classifier reaches values above 80%. As a complementary analysis, the comparison to simulated historical cases (Dumile in 2013, and Dina in 2002) shows error less than 10% on indicators 2 and 3. Finally, we take advantage of the probabilistic

information provided by RF models to evaluate some measure of confidence associated to the prediction result.

As an illustration, the figure depicts the probabilistic RF-based prediction, which corresponds to the exceedance probability for the starting time of the overtopping event. We see that the “true” time-to-event for Dumile cyclone (at ~9 hours) falls into the tight confidence interval at 90% provided by the RF model.



**Figure.** Exceedance probability (black) estimated by the RF model to predict the starting time of the overtopping event (Time-to-Event). The vertical red line corresponds to the true value.