Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira,
Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria
 Published by Research Publishing, Singapore.
doi: 10.3850/ESREL2026061419_esrel26-p26304-cd

A Surrogate Model for Quantifying Hydrodynamic Forces on Coastal Structures under Uncertain Compound Flooding Scenarios

Thirusankaranand Mathavan

Laboratory of Engineering Sciences for the Environment (LaSIE), University of La Rochelle, France.
E-mail: thirusankaranand.mathavan@univ-lr.fr

Erwan Liberge

Laboratory of Engineering Sciences for the Environment (LaSIE), University of La Rochelle, France.
E-mail: erwan.liberge@univ-lr.fr

Emilio Bastidas-Arteaga

Laboratory of Engineering Sciences for the Environment (LaSIE), University of La Rochelle, France.
E-mail: ebastida@univ-lr.fr

Extreme storm events can lead to flooding in coastal cities, introducing significant uncertainties in floodwater depths and wind-driven storm surges and waves. Flooding from these events exerts static loads as well as wave-generated cyclic hydrodynamic forces on coastal structures. Typically, flood impact assessments focus solely on static water levels, overlooking dynamic load contributions. In addition, most numerical methods are computationally expensive when accounting for dynamic wave action due to the inherent complexity of accurately modeling such interactions, while empirical formulas tend to under- or over-estimate these dynamic load patterns. This research proposes a physics-based surrogate model to estimate hydrodynamic loading on residential buildings subjected to compound flooding from the combined effects of storm surge and waves. The surrogate model, based on Polynomial Chaos Expansion (PCE), aims to estimate fluid forces on these buildings under uncertain water levels and wave loads. The proposed methodology lays the foundation for scalable assessment of wave-induced loads in compound flooding scenarios by significantly reducing computational effort and has the potential to support future applications in coastal risk analysis, building fragility assessment, and resilience-oriented design.

Keywords: Computational Fluid Dynamics (CFD), Polynomial Chaos Expansion (PCE), Uncertainty Quantification, Residential Buildings, Compound Flooding.

1. Introduction

Due to climate change and rising sea levels, extreme storm events are becoming more frequent and intense, leading to flooding in coastal cities and introducing significant uncertainties in floodwater depths and wind-driven storm surges and waves. According to the European Commission, Directorate-General for Maritime Affairs and Fisheries (2023), coastal flooding currently causes economic damages of approximately €1.2 billion per year in the European Union. Under projected climate change scenarios, this figure could rise dramatically, reaching €137 billion annually under moderate-emission pathways and up to €814 billion under high-emission scenarios, depending on the severity of sea-level rise and storm impacts. During storm conditions, storm surge produces a quasi-static water level (ponding depth) in front of buildings, upon which wind-generated waves are superimposed. This phenomenon imposes cyclic hydrodynamic loads on buildings in addition to the hydrostatic component. However, most flood studies consider only static flood depth and fail to include wave action. As shown by Tomičzek et al. (2014), buildings subjected to lower flood depths but higher wave conditions were found to experience more severe damage than buildings exposed to higher flood levels with smaller waves. This highlights the importance of considering wave action in coastal flood-related studies.

Most commonly used empirical formulations

suffer from significant uncertainties inherent to their assumptions and parameterizations, often leading to unreliable predictions. While Computational Fluid Dynamics (CFD) modeling provides a more robust and physics-based alternative (Do et al. (2020); Dang et al. (2024)), its widespread use is constrained by prohibitive computational demands. Consequently, Polynomial Chaos Expansion (PCE)–based meta-modeling emerges as an efficient approximation strategy for representing high-fidelity numerical models with significantly reduced computational cost.

In this study, the main objective is to develop a surrogate model based on the Polynomial Chaos Expansion method to estimate fluid forces on buildings subjected to compound flooding from the combined effects of storm surge and waves. To ensure that the surrogate model captures realistic coastal flood conditions, it is trained using a real study site: *Île de Noirmoutier* on the French Atlantic coast. Building geometry parameters are derived from empirical datasets describing the local building stock, while environmental parameters are sampled from physically admissible flood scenarios consistent with the regional hydrodynamic climate. The sampled combinations of building and environmental conditions form the basis of a metamodeling framework, which is demonstrated in this study using a simplified analytical model. Future work will replace this "toy model" with high-fidelity Computational Fluid Dynamics (CFD) simulations to train and validate the PCE surrogate for realistic coastal building configurations.

2. Study area

Île de Noirmoutier is a low-lying island on the French Atlantic coast (Figure 1), with an area of about 48 km^2 and a permanent population of around 9,400 inhabitants. It is one of the most popular summer tourist destinations in the region, and the building stock is largely associated with resort and tourism activities, with a dominant share of secondary residences (65.5% according to CNES (2021)). Because the island's economy is primarily recreational, the built environment is predominantly residential, with a relatively small proportion of permanent dwellings and a large number of holiday homes.

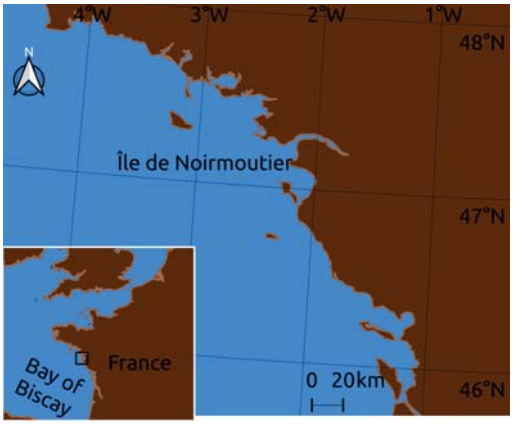

Fig. 1. Study area: Noirmoutier island in France

Several parts of the island lie at or below 0 m NGF–IGN69, and many areas are close to or even below mean sea level. Ground elevation is a key variable in coastal flood risk Creach et al. (2022) and given the island's generally low relief, combined with located in southern portion of the Armorican Shelf, the wide and shallow continental shelf along the Atlantic coast—which enhances wave transformation and storm-surge penetration. Therefore, the physical setting makes Île de Noirmoutier particularly vulnerable to coastal flood hazards.

Approximately 60% of residential buildings are exposed to some degree of flooding, and overall 51% are exposed to water depths exceeding 1 m (Creach et al. (2016)), a threshold that can potentially lead to structural component failure and life-threatening situations, as observed during storm Xynthia in the French department of Charente-Maritime (Vinet et al. (2011)).

The island is currently protected by an extensive system of seawalls. Breaches of these defenses during the 1937 and 1978 floods caused severe inundation, after which substantial efforts were undertaken to strengthen and redesign coastal protection systems. Although no events of comparable magnitude have occurred

since, future climate change and increasing hydrometeorological variability are expected to exacerbate coastal flood hazards. Île de Noirmoutier is therefore an ideal case study for developing a metamodel to estimate fluid forces on residential buildings under compound flooding scenarios, which can support flood risk assessment and the design of mitigation and adaptation strategies.

3. Metamodeling Framework

In this study, a metamodeling approach based on Polynomial Chaos Expansion–Gaussian Kriging (PCE–GK) (Schöbi et al. (2015)), also referred to as PC–Kriging (PCK), is employed to construct surrogate models for uncertainties in wave-induced forces on coastal building façades. The surrogate models are developed using the UQLabpy framework (Lataniotis et al. (2021)).

In the PCE–GK formulation, the model response $Y(\mathbf{X})$ is expressed as the sum of a polynomial expansion and a Gaussian component:

$$Y(\mathbf{X}) = \sum_{i=0}^{P} a_i \, \Psi_i(\mathbf{X}) + Z(\mathbf{X}) \qquad (1)$$

where $\mathbf{X} = (X_1, X_2, \ldots, X_d)$ denotes the vector of uncertain input variables, $\Psi_i(\mathbf{X})$ are multivariate orthogonal polynomial basis functions, a_i are the corresponding PCE coefficients, and $Z(\mathbf{X})$ is a zero-mean Gaussian process accounting for residual variations not captured by the polynomial expansion.

In this hybrid approach, the PCE component captures the global behavior of the system response, while the Gaussian component accounts for local deviations and nonlinearities not represented by the polynomial expansion. To evaluate the influence of different surrogate modeling strategies, standalone Kriging, standalone PCE, sequential PC–Kriging , and optimal PC–Kriging models are compared. Finally, Model performance is assessed using several error metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2), generalization error (GenErr) evaluated on an independent validation dataset, as well as the leave-one-out (LOO) cross-validation error computed from the training samples.

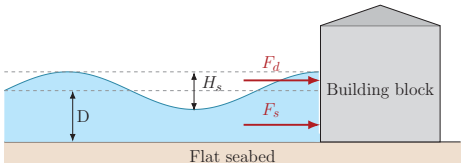

Fig. 2. Schematic of wave H_s impact on a building façade under flood depth D, showing hydrostatic force F_s and dynamic wave force F_d.

3.1. *Mathematical Model*

Prior to uncertainty modeling and surrogate construction, a simplified analytical framework is adopted to estimate wave-induced forces on building façades under compound flooding conditions. This preliminary model serves as a physically consistent link between sampled environmental parameters and subsequent high-fidelity CFD simulations.

Wave-induced loading is estimated using a simplified analytical formulation (hereafter referred to as the "toy model") based on the pulsating pressure approach of Goda (2000). The model is employed to provide order-of-magnitude force estimates and to generate training data for surrogate-model development, rather than to replace CFD simulations.

The building façade is idealized as a rigid vertical wall subjected to head-on, non-breaking wave attack over a flat seabed, with partial submergence due to elevated water levels during flooding (Figure 2). Under these assumptions, the total horizontal force acting on a façade of width B is expressed as

$$F = B\,(F_{\text{stat}} + F_{\text{dyn}}) \qquad (2)$$

$$F_{\text{stat}} = \frac{1}{2}\,\rho g D^2 \qquad (3)$$

$$F_{\text{dyn}} = \frac{1}{2}(p_H + p_L)D + \frac{1}{2}p_H\,h_c^* \qquad (4)$$

where, D denotes the still-water depth at the wall, h_c^* is the effective vertical extent of wave action above the still water level, ρ is the density of seawater, and g is the gravitational acceleration. p_H denotes the wave-induced (excess) pressure acting

on the wall at the still water level corresponding to the incident wave crest. p_L denotes the wave-induced (excess) pressure acting on the wall at the seabed elevation, accounting for the vertical decay of dynamic wave pressure with depth. p_4 denotes the wave-induced (excess) pressure acting at the upper limit of the wetted wall above the still water level.

3.2. *Modelling Uncertainties*

The hydrodynamic forces acting on buildings are governed by six uncertain input parameters, grouped into (i) environmental variables and (ii) building variables. Environmental variables describe the hydrodynamic forcing and include the still-water level and wave characteristics (significant wave height and peak period), while building variables describe the geometry of the exposed structures (length, width, and height). These parameters are derived from real datasets and sampled to generate combinations of environmental and structural conditions for surrogate-model construction.

3.2.1. *Data Collection and Distribution Fitting*

Building geometry data were obtained from the BD TOPO® September 2025 dataset provided by the *Institut national de l'information géographique et forestière* (IGN, 2025). Residential buildings were retained by filtering out structures with a floor area below 9 m^2 or a ceiling height below 2.2 m, following French housing standards (République française, 2002). All retained buildings were idealised as simple rectangular prisms for the CFD simulations to reduce computational cost while preserving the dominant geometric controls on hydrodynamic loading.

Due to the absence of long-term, high-resolution observations of compound flooding in the Noirmoutier, environmental variables were derived from the best available offshore and coastal dataset, which could represent the local storm driven hydrodynamic characteristics.

Local water-level observations at the Noirmoutier island are limited to short records and are therefore insufficient for characterising long-term sea-level variability, as at least 30 years of continuous data are generally used for climate-related analyses (World Meteorological Organization, 2017). Water-level data were therefore taken from the nearest long-record station, the Saint-Nazaire tide gauge (longitude: $-2.20155°$, latitude: $47.266862°$). Located approximately 40 km away, this station provides a continuous 40-year record (1986–2025) and exhibits tidal characteristics and storm-driven variability very similar to those at L'Herbaudière, with minimal phase lag along this section of the French Atlantic coast. It therefore represents an appropriate long-term proxy for estimating the distribution of regional water levels. Flood depth was then approximated from tide-gauge water levels relative to the minimum island elevation (0 m NGF–IGN69), and only water levels exceeding the 95th percentile were retained to focus on extreme, storm-driven but plausible flood conditions.

Wave data were obtained from the Copernicus Marine Service global wave reanalysis product, as no operational Copernicus in-situ buoy or other direct wave measurements exist in the immediate vicinity of Île de Noirmoutier. A virtual buoy location was therefore defined near the island to extract the reanalysis wave dataset for the period 1985–2025 using the methods proposed by Law-Chune et al. (2021); Rovere (2025).

Once all data were collected, probability distributions were fitted to each input variable using the available datasets. the selected distributions are summarised in Table 1.

Table 1. Fitted probability distributions for building and environmental parameters.

Var	Distr	Shape	loc	Scale
B	GenExtreme	0.036	8.455	2.915
L	GenExtreme	-0.031	13.276	4.515
H	GenExtreme	-0.037	3.376	0.729
D	Weibull	2.441	-0.171	1.733
H_s	Lognormal	0.667	0.077	1.101
T_p	Logistic	–	10.872	1.685

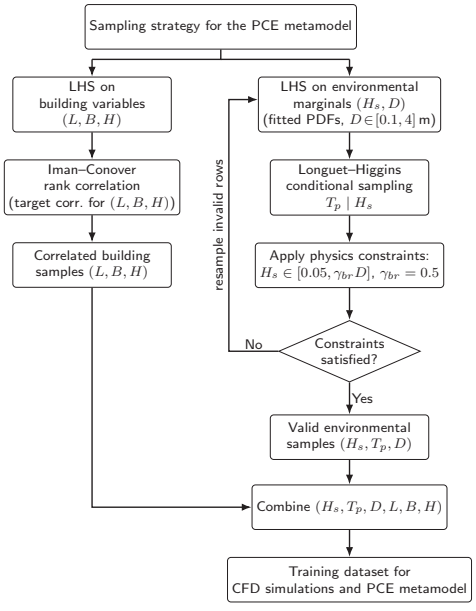

Fig. 3. Sampling strategy

3.2.2. *Sampling*

To construct the PCE surrogate model, the training samples were generated using a carefully designed multi-stage sampling strategy that preserves the natural patterns of the key variables (Figure 3). First, Latin Hypercube Sampling (LHS) was applied to generate building geometries (length, width, and height) according to their fitted probability distributions, ensuring good coverage of the variability observed in the real building stock. To preserve the correlation structure among the building variables (L, B, H), the Iman–Conover rank-correlation method (Iman and Conover, 1982) was then applied to the LHS samples, producing a synthetic building dataset that retains both the marginal distributions and the dependence structure of the real buildings.

The environmental parameters were sampled using a multistage procedure designed to preserve both the natural statistical behaviour of the variables and the underlying physical constraints of wave dynamics. First, LHS Sampling was employed to generate values of the inundation depth D and the significant wave height H_s, each drawn from their fitted probability density functions. The inundation depth D followed a Weibull distribution with physically based bounds, where the upper limit $D_{\max} \approx 4\,\mathrm{m}$ corresponds to the ~ 1000-year return-period inundation depth at the lowest elevation point in the study area, and a minimum value $D_{\min} = 0.1\,\mathrm{m}$ was imposed to exclude unrealistically dry conditions.

The physical dependence between the significant wave height H_s and the peak wave period T_p was incorporated using a Longuet–Higgins–type formulation Longuet-Higgins (1952). For a given sampled H_s, the associated wave period was generated conditionally according to

$$\ln(T_p) = a + b\,\ln(H_s) + \varepsilon \tag{5}$$

where $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. The coefficients a, b, and the variance σ^2 were calibrated using long-term local wave observations representative of the offshore wave climate near Île de Noirmoutier.

Following the initial sampling step, physics-based constraints were applied to eliminate unrealistic wave combinations. To avoid unrealistically small waves, a minimum significant wave height of $0.05\,\mathrm{m}$ was imposed. In addition, the maximum significant wave height was limited to a fraction of the local inundation depth to exclude breaking-wave conditions, as the present study focuses on non-breaking waves. Specifically, the significant wave height was restricted to

$$H_{s,\min} = 0.05\,\mathrm{m}, \qquad H_{s,\max} = \gamma_{\mathrm{br}} D,$$

where $\gamma_{\mathrm{br}} = 0.5$ is the breaker index and D denotes the sampled inundation depth. Any sample failing to satisfy these criteria was automatically discarded and resampled until all environmental conditions met the imposed physical limits.

A total of 500 valid samples were generated. For each sampled combination of building geometry and environmental conditions, the corresponding hydrodynamic force was computed using the simplified analytical “toy model” described in Section 3.1. The resulting force responses form the training dataset used to construct the Polynomial Chaos Expansion surrogate model.

4. Results

This section presents the performance and convergence characteristics of the proposed surrogate modeling framework for estimating hydrodynamic forces on building façades.

4.1. *Surrogate Model Comparison: Kriging, PCE, and PCE–Kriging (PCK)*

To construct an efficient surrogate for the façade force response, four metamodel families were evaluated on a common validation dataset: Kriging , PCE, and two hybrid PCK approaches based on sequential and optimal basis-selection strategies.

Table 2 summarizes the validation performance of the considered surrogate models. While Kriging captures the overall response trends well, as reflected by a high coefficient of determination, but results in large absolute prediction errors than the polynomial-based approaches. PCE further improves accuracy, indicating that the dominant trends in the response are well captured by a polynomial representation. The hybrid PCK models achieve the lowest prediction errors, with sequential and optimal formulations showing nearly identical performance. Based on these results, the sequential PCK model is selected for subsequent analyses, as it attains comparable accuracy to the optimal formulation while offering improved computational efficiency.

Table 2. Validation performance of Kriging, PCE, and PCE–Kriging (PCK) surrogates using a common validation set.

Model	RMSE	MAE	R^2	GenErr
Kriging	2201	971	0.9997	3.39×10^{-4}
PCE	1274	665	0.9999	1.13×10^{-4}
Sequential PCK	1107	489	0.9999	8.56×10^{-5}
Optimal PCK	1107	489	0.9999	8.56×10^{-5}

Model performance was further examined as a function of polynomial degree and the error metrics obtained are summarized in Table 3. Increasing the polynomial degree or employing more complex basis-selection strategies did not result in noticeable improvements in predictive accuracy, while incurring higher computational cost. Consequently, a second-order sequential PCE model was identified as an effective compromise between accuracy and efficiency, providing performance comparable to higher-order representations at reduced computational expense.

Table 3. Surrogate-model performance as a function of polynomial degree.

Polynomial degree	R^2	LOO error
2	0.9999	7.22×10^{-5}
3	0.9999	7.22×10^{-5}
4	0.9999	7.22×10^{-5}
5	0.9999	7.22×10^{-5}
6	0.9999	7.22×10^{-5}
7	0.9999	7.22×10^{-5}
8	0.9999	7.22×10^{-5}
9	0.9999	7.22×10^{-5}
10	0.9999	7.22×10^{-5}

Figure 4 illustrates the influence of training sample size on the accuracy of the PCE–Kriging surrogate. As the number of training samples increases, the prediction error, quantified by the RMSE, exhibits an overall decreasing trend, indicating improved predictive accuracy with increasing training data.

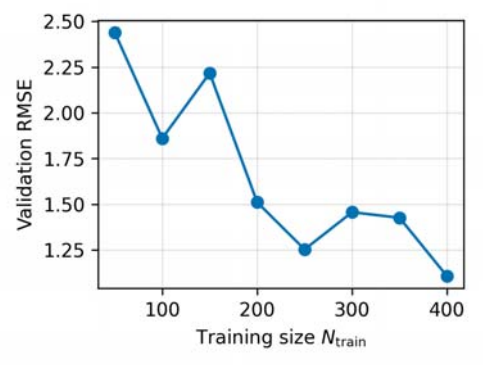

Fig. 4. Learning curve of the PCE–Kriging (PCK) surrogate showing the evolution of prediction error as a function of training sample size.

Although minor non-monotonic variations in

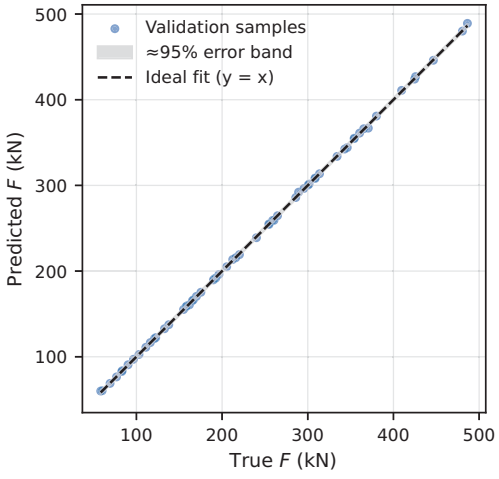

Fig. 5. Parity plot comparing surrogate-predicted and reference force values for the sequential PCE–Kriging model on the validation dataset.

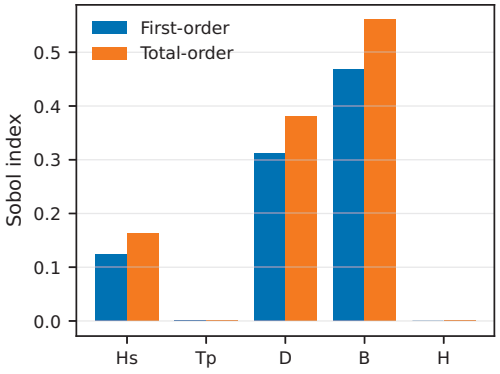

Fig. 6. First-order and total-order Sobol sensitivity indices for the input parameters of the PCE–Kriging surrogate model.

RMSE are observed for intermediate sample sizes, a marked improvement is achieved beyond approximately $N_s = 250$. While the lowest RMSE is obtained for $N_s = 400$ in the present analysis, further increases in training size yield diminishing returns relative to the associated computational and memory requirements. Consequently, a training size on the order of $N_s \approx 250$ is selected as a practical compromise between surrogate accuracy and computational efficiency.

The final results of the surrogate model, based on a sequential second-degree PCE–Kriging formulation trained on 250 samples, is presented in Figure 5. The parity plot demonstrates good agreement between reference force values and surrogate predictions, indicating satisfactory predictive performance within the sampled parameter space.

Sobol sensitivity analysis indicates that the system response is primarily governed by the geometric parameter B and the water depth D, followed by the significant wave height H_s. Together, these parameters account for the majority of the output variance, with first-order Sobol indices summing to approximately 0.91. The differences between total- and first-order indices for B, D, and H_s indicate the presence of moderate interaction and nonlinear effects.

The peak wave period T_p exhibits a negligible influence on the response within the investigated parameter ranges, while the building height H shows no measurable sensitivity. This behavior is attributed to the fact that the flood depth D remains below the building height for all considered cases, such that the effective loaded height in the Goda-type force formulation is capped by the flood depth and therefore independent of H.

In the present formulation, T_p enters only through the estimation of the wavelength L. Since the Goda-type force model is dominated by the vertical water column height, variations in T_p (and thus L) have a negligible influence on the total force.

5. Conclusions and future work

This study presented a physics-informed surrogate modeling framework for estimating hydrodynamic forces acting on coastal buildings subjected to compound flooding from combined storm surge and wave action. By coupling a simplified force formulation with a Polynomial Chaos Expansion–Kriging (PCE–Kriging) metamodel, the proposed approach enables efficient prediction of building-scale loads while preserving key physical dependencies on wave characteristics, water levels, and geometric parameters. The surrogate model demonstrates strong predictive performance and provides physically interpretable sensitivity information, offering a practical alternative to purely depth-based damage representations.

The central objective of the proposed framework is to move beyond depth-only damage metrics toward physically based, load-driven damage assessment. By developing the surrogate model using force estimates derived from high-fidelity CFD simulations, hydrodynamic loads acting on all building façades, global overturning moments, and forces on individual structural components can be explicitly quantified.

At the community scale, the proposed metamodel provides a practical pathway for translating compound flood hazards into physically meaningful structural demands. By retaining information on load magnitude and spatial distribution, the framework enables assessment of multiple damage mechanisms, including component-level failure and global overturning. This capability is essential for realistic estimation of coastal flood damage and for evaluating the effectiveness of mitigation strategies under compound flooding scenarios.

Acknowledgement

This research was jointly funded by l'Agence Nationale de la Recherche (ANR), project ANR-23-CE22-0011 and the Research Grants Council (RGC) Joint Research Scheme (A-PolyU502/23)). For the purpose of open access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

References

CNES (2021). Pays de loire – l'île de noirmoutier : entre convoitises, nécessité de développement et gestion des risques.

Creach, A., E. Bastidas-Arteaga, S. Pardo, and D. Mercier (2022). Method to identify the likelihood of death in residential buildings during coastal flooding. *Buildings 12*(2), 125.

Creach, A., E. Chevillot-Miot, D. Mercier, and L. Pourinet (2016). Vulnerability to coastal flood hazard of residential buildings on noirmoutier island (france). *Journal of Maps 12*(2), 371–381.

Dang, H. V., H. Park, S. Shin, T. Ha, and D. T. Cox (2024). Numerical modeling and assessment of flood mitigation structures in idealized coastal communities: OpenFOAM simulations for hydrodynamics and pressures on the buildings. *Ocean Engineering 307*, 118147.

Do, T. Q., J. W. van de Lindt, and D. T. Cox (2020). Hurricane surge-wave building fragility methodology for use in damage, loss, and resilience analysis. *Journal of Structural Engineering 146*(1), 04019177.

European Commission, Directorate-General for Maritime Affairs and Fisheries (2023). The eu blue economy report 2023. Technical report, Publications Office of the European Union, Luxembourg. Accessed: 2025-08-29.

Goda, Y. (2000). *Random Seas and Design of Maritime Structures* (2nd ed.). Singapore: World Scientific.

IGN (2025). Bd topo® – édition septembre 2025. Base de données vectorielle, IGN Géoservices.

Iman, R. L. and W. J. Conover (1982). A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics – Simulation and Computation 11*(3), 311–334.

Lataniotis, C., S. Marelli, and B. Sudret (2021). Uncertainty quantification in the cloud with UQCloud. In *Proceedings of the 4th International Conference on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP 2021)*.

Law-Chune, S., B. Levier, B. Romain, and E. Greiner (2021). The copernicus marine environment monitoring service global wave reanalysis: Multi-year assessment and product improvements. *Ocean Dynamics*.

Longuet-Higgins, M. S. (1952). On the statistical distribution of the heights of sea waves. *Journal of Marine Research 11*(3), 245–266.

République française (2002). Décret n^{o} 2002-120 du 30 janvier 2002 relatif aux caractéristiques du logement décent. Code de la construction et de l'habitation, France. Pièce principale : 9 m^2 et 2,20 m de hauteur ou 20 m^3 de volume minimal.

Rovere, A. (2025). Github repository: Coastalhydrodynamics. `https://github.com/alerovere/CoastalHydrodynamics`.

Schöbi, R., B. Sudret, and J. Wiart (2015). Polynomial-chaos-based kriging.

Tomičzek, T., A. Kennedy, and S. Rogers (2014). Collapse limit state fragilities of wood-framed residences from storm surge and waves during hurricane ike. *Journal of Waterway, Port, Coastal, and Ocean Engineering 140*(1), 43–55.

Vinet, F., L. Boissier, and S. Defossez (2011). La mortalité comme expression de la vulnérabilité humaine face aux catastrophes naturelles : deux inondations récentes en france (xynthia, var, 2010). *VertigO – la revue électronique en sciences de l'environnement 11*(2).

World Meteorological Organization (2017). Wmo guidelines on the calculation of climate normals.