Applications of Statistical Learning Methods in Natural Hazard Assessment

Ziyue Liu[a*], Azin Al Kajbaf[a], Michelle Bensi[a]
[a]University of Maryland, College Park, USA

Abstract: In recent years, there has been a notable advancement in the sophistication of numerical, physical, and other models used to simulate natural hazard events. This progress has led to the development of high-resolution models that better capture the physics underlying these events. However, the improved modeling capabilities have been accompanied by increased computational demands. Simultaneously, statistical and machine learning techniques have been increasingly employed across a wide variety of applications, including the development of hazard prediction models. The widespread availability of user-friendly, open-source software resources has facilitated the adoption of these techniques. Within the context of predicting hazards to which nuclear power plants may be exposed, it is crucial to acknowledge that many general-use machine learning techniques have been focused on applications requiring the building of efficient models that perform well on average. This emphasis is evident in the metrics and methodologies employed for assessing model performance. This paper, along with its associated presentation, explores the evolution of machine/statistical learning techniques that have been implemented in natural hazard assessment in recent years. Furthermore, some general constraints of implementing machine learning in assessing natural hazard events are discussed.

Keywords: statistical learning, machine learning, natural hazard

1. INTRODUCTION

Physical, numerical, and other process models are used to estimate characteristics of external hazard events to support applications ranging from deterministic and probabilistic hazard assessments to real-time event forecasting (e.g., predicting the depth of water levels and spatial extent of flooding during coastal surge events). As the degree of sophistication of these models has increased, the computational demands have also grown. Practical resource constraints often restrict the number of high-fidelity simulations (e.g., [1, 2]). This problem may be particularly acute when a high degree of spatial resolution is needed to understand the impacts of hazard events on infrastructure [3].

A range of strategies have been used to address these practical considerations. For example, in the context of real-time tropical storm forecasting applications, techniques have included software-based execution strategies [4] as well as multiple downscaling approaches that have been developed to derive model results at a refined spatial scale [3, 5]. Other strategies include relying on lower-fidelity or empirically-derived regression models in lieu of high-fidelity models (e.g., [6, 7, 8]). Such approaches are commonly employed to assess seismic hazards in both risk assessment and real-time contexts [6, 9, 10].

To support probabilistic coastal hazard assessments, techniques have been proposed that use response surface approaches based on physically-informed interpolation approaches (e.g., [11-13]) and Kriging methods (e.g., [11,14]) that seek to interpolate values of response quantities (e.g., storm surge) based on the results of a limited number of model runs. Optimal sampling techniques have also been employed that aim to sufficiently "cover" the storm parameter space through the selection and weighting of a small number of parameter values in a manner that minimizes the number of storms required to produce a hazard curve [15]. Other response-based sampling selection strategies have also been used for efficient hazard analysis [16, 17].

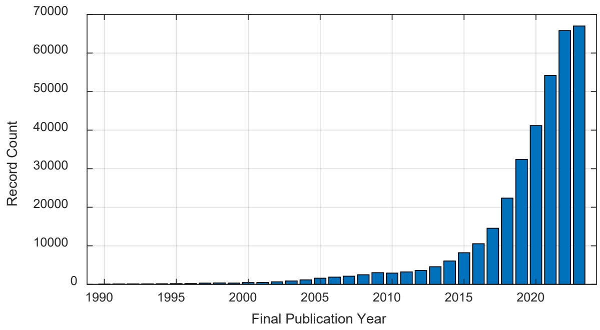
Recently, machine (ML) or statistical learning methods have been increasingly used to emulate sophisticated physical, numerical, and other process models. ML has been used to build models using recorded data such as from strong-motion databases [18-20] or meteorological records [21]. However, surrogate models are increasingly being developed using synthetic input/output from high-fidelity models for training, validation, and testing [22-25].

2. GROWTH IN THE USE OF STATISTICAL AND MACHINE LEARNING APPROACHES

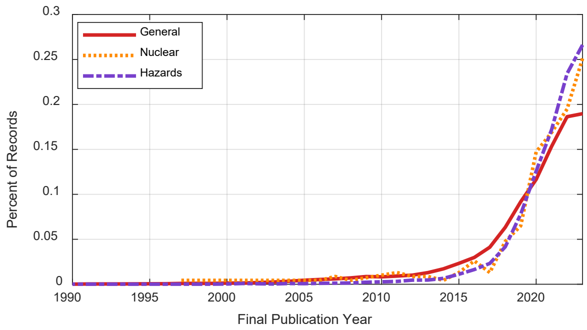
ML models provide an opportunity to balance efficiency and accuracy by using parametric or nonparametric models to emulate more resource-intensive computational models. Figure 1 (a) shows the rapid growth in literature from 1990 to 2023 related to the general application of ML methods. This data was obtained from a search of the Web of Science[1] using the topical search[2] criteria outlined in Table 1 and filtered to include only articles, proceedings, and review articles with a known final publication date between 1990 and 2023.

Figure 1 (b) shows how the growth of application-specific ML-focused literature compares with this overall trend using data obtained from a similar topical search of the Web of Science using the criteria shown in Table 1. Specifically, the figure shows the growth trend of literature published between 1990 and 2023 focused on nuclear power and natural hazards applications compared with the growth of the general application of ML methods. This comparison of growth trends is performed by identifying the percentage of the total number of articles[3] published under each topical domain each year between 1990 and 2023.

While the absolute number of articles published under each domain differs, the growth of hazards-focused and nuclear power-focused literature is actually outpacing the overall trend in recent years. (though it is noted that the total number of published literature in the nuclear power domain remains small). Highlighting potential research needs, an attempt to perform a topical search intersecting the nuclear power- and hazards-focused domain yielded less than ten articles, with only a fraction appearing to be directly relevant.



(a)



(b)

Figure 1. (a) Growth in literature related to machine and statistical learning methods; (b) Percentage of total published literature (by year) for selected topical areas

[1] Certain data included herein are derived from Clarivate™ (Web of Science™). © Clarivate 2024. All rights reserved. [Search performed May 29, 2024]
[2] Topical searches within the Web of Science consider an article title, abstract, author keywords, and Keywords Plus®.
[3] 353,305, 231, and 10,695 articles were identified for the general, nuclear power-focused, and hazard-focused area, respectively, using the specified date range and topical search criteria shown in Table 1.

Table 1. Web of Science Search Criteria

Topic of Interest	Web of Science Topical Search (TS) Criteria[4]
Use of ML in general	TS=("statistical learning" OR "machine learning")
Use of ML in nuclear power-focused applications	TS=(("statistical learning" OR "machine learning") AND ("nuclear power"))
Use of ML in hazard-focused applications	TS=(("statistical learning" OR "machine learning") AND ("external hazard*" OR "natural hazard*" OR earthquake* OR seismic OR "ground shaking" OR "ground motion" OR hurricane* OR "tropical storm*" OR "tropical cyclone*" OR typhoon* OR "cyclonic storm*" OR cyclone* OR "tropical depression*" OR rain* or precipitation OR "wildfire*" OR drought* OR heatwave* OR landslide* OR tsunami* OR seiche* OR "storm surge*" OR "storm tide*" OR hail* OR avalanche* OR blizzard* OR snow* OR volcano* OR volcanic OR tornado* OR mudslide*))

3. COMMENTARY

The rapid growth in ML applications has been enabled by advancements in software resources, which have substantially lowered the barriers to implementing ML approaches. Many ML methods and software resources have been developed with a focus on “big data” applications and ensuring models perform well “on average.” This scale and objective may not reflect the context of natural hazards risk assessment for nuclear power plants, which seeks to capture the contributions of severe hazards to risk. This issue may be particularly acute given the range of return periods of relevance. The subsections below highlight several potential considerations of relevance to applying ML within the context of external hazard risk assessment for nuclear power plants. This commentary is augmented by presentation content and is intended to elicit discussion among researchers and practitioners.

3.1. Variability in Predicted Quantities

From the perspective of statistics, a regression model prediction (whether the regression uses conventional or ML-based approaches) returns the best estimate of the response, often represented as the conditional expectation of the response quantity given the predictors. However, this “expected prediction” derived from an algorithm that targets to minimize loss function, tends to reduce variability when compared to the actual response. To preserve this variability, error terms can be leveraged. In probabilistic hazard assessment such as probabilistic seismic hazard analysis (e.g., [26]), the error in the predictive regression model, which is typically developed through statistical analysis of observational data, is integrated into the hazard curve as a reflection of variability. In other applications (e.g., [2]), the error is used to develop confidence or other error bounds. In recent applications, ML approaches have been used in conjunction with synthetic training data derived from numerical models (e.g., [27]). In this case, there are errors associated with both the numerical model and the surrogate ML model, and these errors may or may not linearly superimpose.

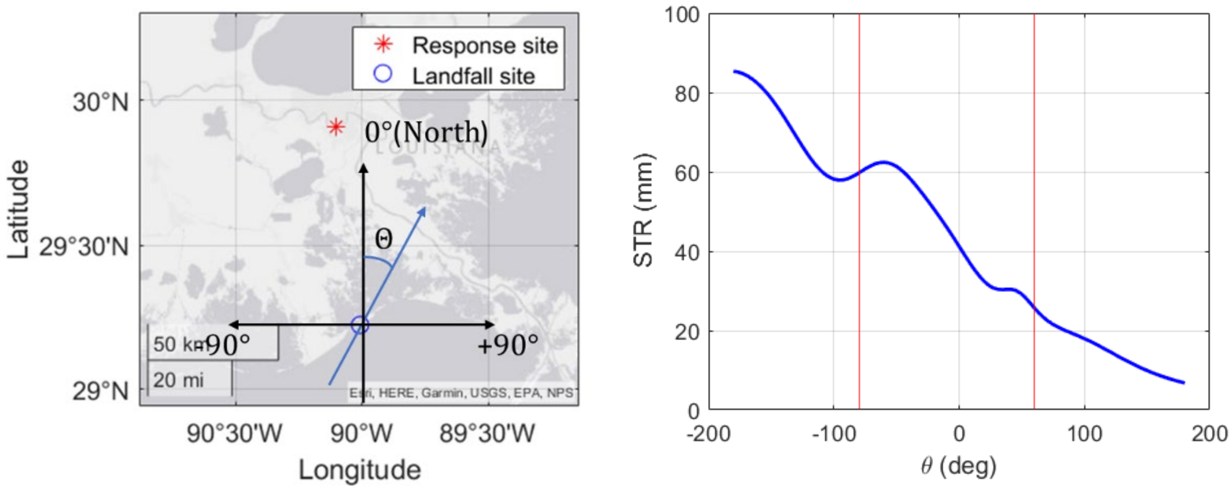
ML-derived regression models have also been developed for applications beyond the prediction of response quantities. For example, the challenge of limited and incomplete historical storm datasets has been addressed by leveraging ML techniques to impute missing data [1, 2, 28]. As demonstrated by [28], this can reduce the variability within the imputed data. Moreover, if the imputation model is built using multiple variables within a dataset, the use of an imputation model can enhance the correlation between imputed variables and other input variables in the dataset. This can impact subsequent statistical assessments performed using the imputed data, especially in joint probability analysis [29].

3.2. Interpolation and Extrapolation

[4] All searches are filtered to include only articles, proceedings, and review articles.

Unreliable (and physically unfeasible) predictions may arise when seeking to predict responses for input (predictor) values that fall outside the range of the training data (extrapolation). Given the limited data generally available related to natural hazards, similar issues may arise when seeking to perform interpolation if the training data has gaps within the parameter domain of interest. These issues may be exacerbated when using synthetic data that is not randomly sampled.

There are several strategies that can be used to assess the impacts of extrapolation outside the range of training data. One helpful strategy is to visualize response functions, which show the predicted value of a quantity given variation in the input quantities.[5] For example, Figure 2 shows an example of a response function for an ML-derived storm total rainfall (STR) model considering variation in the predictor quantity storm heading direction (θ) and fixing all other predictor quantities.[6] An artificial neural network (ANN) STR model based on one developed by [30] is used here. It is emphasized that this model is intentionally used outside of its intended purpose for illustrative purposes. It can be observed that when the input θ falls below -90 degrees (the minimum value of θ in the training dataset, as denoted by the red vertical line in Figure 2(b)), the predicted STR tends to increase disproportionately. Such behavior appears unreliable when considering the coastal geometry (as depicted in Figure 2 (a)).



(a) (b)

Figure 2. (a) Geometry and coordinate used for θ; (b) ML STR model response function, the red vertical line indicates the range of training data

In addition to unreliable predictions that may occur as a result of inputs falling outside the range of the training data, ML models may have difficulty in predicting unseen parameters within the training range. A useful means of addressing this issue is to perform strategic sampling in selecting training data so that all possible values of predictors expected for a study region or a specific application are covered. In the case of applying ML models for predicting storm surges, [31] evaluated the performance of three ML models (ANN, Gaussian process regression (GPR), and support vector regression (SVR)) for predicting storm surges associated with unseen values of the parameters Δp and θ. Consistent with many recent applications of ML in coastal hazard assessment (e.g., [14, 32, 33]), the data used for ML model development in [31] was synthetic and derived from a series of numerical model runs that did not involve random sampling of the input parameter space. Certain parameters within the input parameter space were systematically discretized, leading to a limited number of parameter combinations in the training/testing data. As a result, the application of conventional validation approaches (i.e., leave one out cross-validation or k-fold cross-validation) would lead to training and testing sets that may contain similar events and not provide a reflection of the model's out-of-sample capabilities as would be likely for randomly sampled training/testing data (e.g., as is more likely when working with observational data). To address this issue, [31] used strategic sampling in which all storms corresponding

[5] Partial dependence plots can also be used to show marginal effects of one or two predictors on a response quantity.
[6] When generating the response, other input parameters are set as "fixed" values that are representative of the mean of historical storm data

to a specified value of Δp and θ are eliminated from the database and used as a testing set while the rest of the data is used as training (as opposed to randomly holding out a set of storms). Figure 3 (adapted from [31]) illustrates the error of prediction in terms of root mean squared error (RMSE) for the three ML models. For all models, use of prediction in an entirely unseen quantity may cause significant increases in error of prediction.

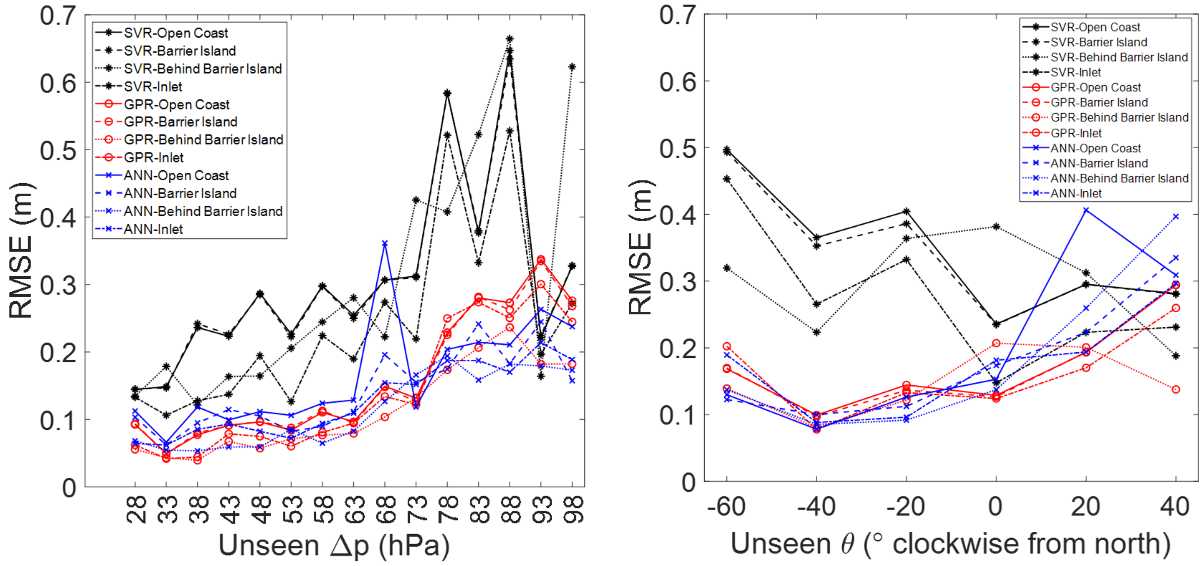


Figure 3. RMSE of ML models at different unseen Δp and θ (adapted from [31])

3.3. Imbalanced Sample

The distribution of target data is also important as it can affect the credibility of the ML model's performance accuracy. Having imbalanced target variables (absence or having too few target variables in a specific range) can negatively influence the reliability of ML models' predictions in the range that target training data is scarce. Furthermore, the presence of many near-zero target values can be problematic when evaluating model performance. As noted above, [31] comparatively explored the performance of three ML models of ANN, GPR, and SVR in prediction of storm surge using the United States Army Corps of Engineers North Atlantic Coast Comprehensive Study (NACCS) synthetic dataset [27]. The majority of surge heights in the NACCS dataset are small in magnitude (i.e., surge heights are clustered near zero) because the synthetic storm suite is intended to generate surges along the entire North Atlantic coast. In the assessment of risks from coastal hazards, larger surge events may be of particular interest. The accuracy of ML models is usually evaluated using aggregated error metrics (e.g., RMSE). The inclusion of many small magnitude values in the computation of those metrics will decrease the overall averages, and thus, those metrics may not reflect the performance of the model under the larger events of interest.

As shown in [31], a possible solution would be to truncate the training/testing set so that the error metric is focused on risk-important events. Figure 4 (adapted from [31]) shows the RMSE and correlation coefficient (R) values for ANN, GPR, and SVR models through 10-fold cross-validation testing. Upper and lower bars show the max and min values in 10 folds, and the middle symbol shows the mean. While the models are trained on data associated with the full range of surge values, the models are tested on a truncated dataset where all surges below the lower bound truncation threshold (as indicated in the x-axis) are removed so that error measures focused on model performance over the range of surge heights of interest. For all three models, the error increases (R decreases) with the exclusion of smaller surge values.

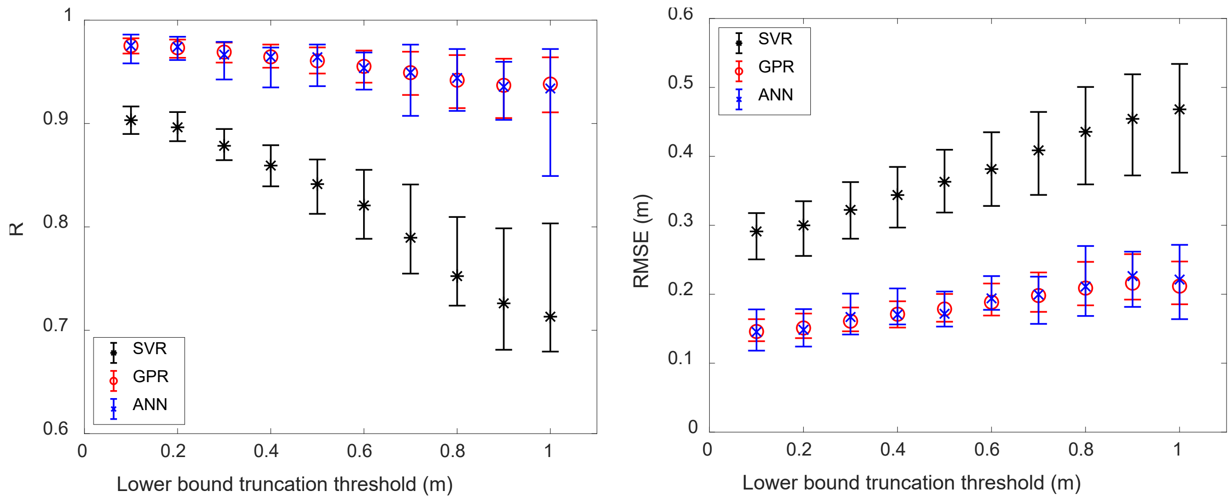


Figure 4. Max, min and mean of R and RMSE across truncated testing sets. For each testing set, all surges below the lower bound truncation threshold are removed (adapted from [31])

3.4. Physical Constraints

While ML-prediction models are generally highly efficient, they lack inherent physical constraints, unlike well-defined numerical models. As a result, it is important to understand the mechanisms of ML models and avoid using them as "black boxes" (which can be enabled by increasingly easy-to-use software tools). Response functions can be a bridge connecting the physical phenomena and the ML models and help assess their physical reasonableness or rationality. Response functions provide a means to graphically investigate the relationship between input and output parameters [21]. To develop a two-dimensional response function, all variables are locked at a specific value except for two variables, which are varied over a defined range.[7] An ML model calculates the target response corresponding to each pair of variables (which is shown with the colors on the graph). In the case of the application of ML models to temporally downscale precipitation, [21] created response functions to explore the different mechanisms of the ML models of ANN and boosted trees (BT) in predicting target precipitation. Figure 5 (adapted from [21]) shows the response functions of ANN and BT models for the pair of variables, daily precipitation and 2-hour precipitation, that are used as predictors to downscale 2-hour precipitation to 1-hour precipitation. As shown in Figure 5, ANN has a smooth transition between different values, while BT shows sudden changes evident by the plaid or striped patterns. The abrupt changes seen in the behavior of BT model are due to the mechanism of its learner algorithm, an ensemble of decision trees that consists of splitting paths upon specific values. It is noted that response functions often highlight physical considerations that are not captured by conventional error metrics. For example, it may be possible for a physically impractical result to lead to smaller error metrics, which may be caused by overfitting or the limited range of applicability of the model. For example, as shown in [21], the BT model shows a good performance in terms of aggregated error metrics. However, downscaled precipitation predictions occasionally fall into unrealistic physical domains (i.e., predictions are negative).

[7] Two-dimensional partial dependence plots can also be used within this context.

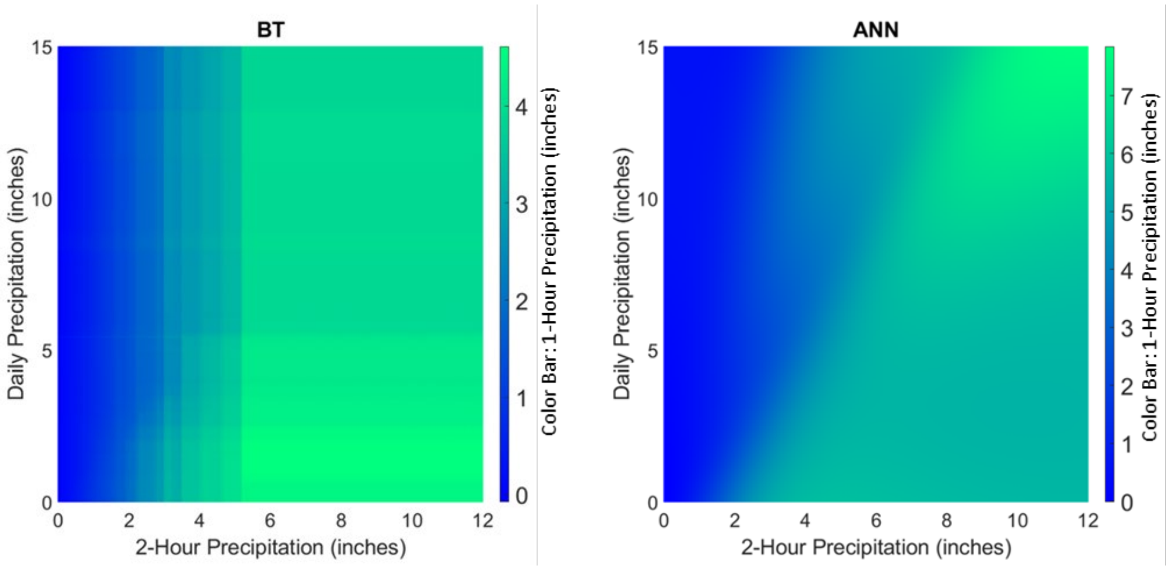


Figure 5. Response functions for different target precipitation of ANN and BT models (adapted from [21])

4. CONCLUSION

Over the last decade, there has been a substantial increase in the application of statistical and machine learning techniques for a wide range of applications, including external hazard assessment. The wide availability of low-barrier, open-source general software resources has enabled applications of these techniques.

Despite the proliferation of ML approaches/software, evaluations of model performance often overlook the distinctive characteristics of predictive modeling for severe hazard events. Both the nature of ML algorithms, aimed at predicting "expected" values, and the constraints of "big data" utilization—where ML models are trained on assigned data rather than direct physical attributes of hazard events—can potentially impact the quality of predictions generated. This paper explores the growth in ML approaches in the modeling of hazard events, highlighting key themes/trends. It further provides commentary on assessing the performance of ML models for severe natural hazard events.

Acknowledgments

The commentary provided in this paper benefitted from discussions with co-authors as part of multiple projects documented in cited references. The authors gratefully acknowledge these collaborations and particularly acknowledge the value of discussions with Dr. Kaye Brubaker, Dr. Meredith Carr, Dr. Norberto Nadal-Caraballo and Dr. Jennifer Irish.

Disclaimer

Any opinions, findings, and conclusions expressed in this paper/presentation are those of the authors and do not necessarily reflect the views of any organization.

References

[1] N. C. Nadal-Caraballo, M. O. Campbell, V. M. Gonzalez, M. J. Torres, J. A. Melby, and A. A. Taflanidis, "Coastal Hazards System: A Probabilistic Coastal Hazard Analysis Framework," *Journal of Coastal Research*, vol. 95, no. sp1, p. 1211, May 2020, doi: 10.2112/SI95-235.1.

[2] N. C. Nadal-Caraballo *et al.*, "Coastal Hazards System–Louisiana (CHS-LA)," Vicksburg, MS: US Army Engineer Research and Development Center, ERDC/CHL TR-22-16, Aug. 2022. Accessed: Sep. 12, 2022. [Online]. Available: http://dx.doi.org/10.21079/11681/45286.

[3] C. A. Rucker *et al.*, "Downscaling of real-time coastal flooding predictions for decision support," *Natural Hazards*, vol. 107, no. 2, pp. 1341–1369, Jun. 2021, doi: 10.1007/s11069-021-04634-8.

[4] J. G. Fleming, C. W. Fulcher, R. A. Luettich, B. D. Estrade, G. D. Allen, and H. S. Winer, "A Real Time Storm Surge Forecasting System Using ADCIRC," in *Proceedings of the 10th International Conference on Estuarine and Coastal Modeling*, Newport, Rhode Island, Nov. 2007. Accessed: Jan. 03, 2024. [Online]. Available: https://ascelibrary.org/doi/10.1061/40990%28324%2948.

[5] C. A. Rucker, "Improving the Accuracy of a Real-Time ADCIRC Storm Surge Downscaling Model | Coastal & Computational Hydraulics Team," NorthCarolinaStateUniversity, Raleigh,NorthCarolina, 2020. Accessed: Jan. 03, 2024. [Online]. Available: https://ccht.ccee.ncsu.edu/ms-2020-carter/.

[6] J. Douglas, "Ground motion prediction equations 1964-2021." Mar. 2022. [Online]. Available: http://www.gmpe.org.uk/gmpereport2014.pdf.

[7] B. Glahn, T. Arthur, K. Nicole, and S. Wilson, "THE ROLE OF THE SLOSH MODEL IN NATIONAL WEATHER SERVICE STORM SURGE FORECASTING," *National Weather Digest*, vol. 33, no. 1, pp.3–14,2009,[Online].Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=4fe41d7103868cd88e0608a3308b9c1659a6df2d.

[8] J. Kaplan and M. DeMaria, "A Simple Empirical Model for Predicting the Decay of Tropical Cyclone Winds after Landfall," *Journal of Applied Meteorology and Climatology*, vol. 34, no. 11, pp. 2499–2512, Nov. 1995, doi: 10.1175/1520-0450(1995)034<2499:ASEMFP>2.0.CO;2.

[9] USGS, "2.3. Data Processing," ShakeMap Documentation (USGS GitHub). Accessed: Jan. 03, 2024. [Online]. Available: https://usgs.github.io/shakemap/manual4_0/tg_processing.html.

[10] USGS, "Hazards | U.S. Geological Survey." Accessed: Jan. 03, 2024. [Online]. Available: https://www.usgs.gov/programs/earthquake-hazards/hazards.

[11] G. Jia and A. A. Taflanidis, "Kriging metamodeling for approximation of high-dimensional wave and surge responses in real-time storm/hurricane risk assessment," *Computer Methods in Applied Mechanics and Engineering*, vol. 261, pp. 24–38, 2013.

[12] D. T. Resio, "White paper on estimating hurricane inundation probabilities," *This Digital Resource was created from scans of the Print Resource.*, May 2007, Accessed: Jul. 21, 2022. [Online]. Available: https://erdc-library.erdc.dren.mil/jspui/handle/11681/22643.

[13] A. A. Taflanidis, G. Jia, A. B. Kennedy, and J. M. Smith, "Implementation/optimization of moving least squares response surfaces for approximation of hurricane/storm surge and wave responses," *Natural Hazards*, vol. 66, no. 2, pp. 955–983, Mar. 2013, doi: 10.1007/s11069-012-0520-y.

[14] G. Jia, A. A. Taflanidis, N. C. Nadal-Caraballo, J. A. Melby, A. B. Kennedy, and J. M. Smith, "Surrogate modeling for peak or time-dependent storm surge prediction over an extended coastal region using an existing database of synthetic storms," *Natural Hazards*, vol. 81, no. 2, pp. 909–938, 2016.

[15] G. R. Toro, D. T. Resio, D. Divoky, A. Wm. Niedoroda, and C. Reed, "Efficient joint-probability methods for hurricane surge frequency analysis," *Ocean Engineering*, vol. 37, no. 1, pp. 125–134, Jan. 2010, doi: 10.1016/j.oceaneng.2009.09.004.

[16] F. L. Santiago-Collazo, M. V. Bilskie, P. Bacopoulos, and S. C. Hagen, "An Examination of Compound Flood Hazard Zones for Past, Present, and Future Low-Gradient Coastal Land-Margins," *Front. Clim.*, vol. 3, p. 684035, Jul. 2021, doi: 10.3389/fclim.2021.684035.

[17] M. Sohrabi, H. Moftakhari, and H. Moradkhani, "Efficient Tropical Cyclone Scenario Selection Based on Cumulative Likelihood of Potential Impacts," *Earth's Future*, vol. 11, no. 10, p. e2023EF003731, Oct. 2023, doi: 10.1029/2023EF003731.

[18] A. Alimoradi and J. L. Beck, "Machine-Learning Methods for Earthquake Ground Motion Analysis and Simulation," *Journal of Engineering Mechanics*, vol. 141, no. 4, p. 04014147, Apr. 2015, doi: 10.1061/(ASCE)EM.1943-7889.0000869.

[19] A. Joshi, B. Raman, C. K. Mohan, and L. R. Cenkeramaddi, "Application of a new machine learning model to improve earthquake ground motion predictions," *Natural Hazards,* Oct. 2023, doi: 10.1007/s11069-023-06230-4.

[20] F. Khosravikia and P. Clayton, "Machine learning in ground motion prediction," *Computers & Geosciences*, vol. 148, p. 104700, Mar. 2021, doi: 10.1016/j.cageo.2021.104700.

[21] A. Al Kajbaf, M. Bensi, and K. L. Brubaker, "Temporal downscaling of precipitation from climate model projections using machine learning," *Stoch Environ Res Risk Assess*, vol. 36, no. 8, pp. 2173–2194, Aug. 2022, doi: 10.1007/s00477-022-02259-2.

[22] E. Adeli, L. Sun, J. Wang, and A. A. Taflanidis, "An advanced spatio-temporal convolutional recurrent neural network for storm surge predictions," *Neural Comput & Applic*, vol. 35, no. 26, pp. 18971–18987, Sep. 2023, doi: 10.1007/s00521-023-08719-2.

[23] J.-W. Lee, J. L. Irish, M. T. Bensi, and D. C. Marcy, "Rapid prediction of peak storm surge from tropical cyclone track time series using machine learning," *Coastal Engineering*, vol. 170, p. 104024, Dec. 2021, doi: 10.1016/j.coastaleng.2021.104024.

[24] N. C. Nadal-Caraballo, M. O. Campbell, V. M. Gonzalez, M. J. Torres, J. A. Melby, and A. A. Taflanidis, "Coastal Hazards System: A Probabilistic Coastal Hazard Analysis Framework," *Journal of Coastal Research*, vol. 95, no. SI, pp. 1211–1216, Jun. 2020, doi: 10.2112/SI95-235.1.

[25] Y. Qin, C. Su, D. Chu, J. Zhang, and J. Song, "A Review of Application of Machine Learning in Storm Surge Problems," *Journal of Marine Science and Engineering*, vol. 11, no. 9, Art. no. 9, Sep. 2023, doi: 10.3390/jmse11091729.

[26] J. W. Baker, B. A. Bradley, and P. J. Stafford, *Seismic Hazard and Risk Analysis*. Cambridge, England: Cambridge University Press, 2021. Accessed: Dec. 09, 2022. [Online]. Available: https://www.cambridge.org/core/books/seismic-hazard-and-risk-analysis/177B4BA01FC2600AA53A22B27598A06C.

[27] N. C. Nadal-Caraballo, J. A. Melby, V. M. Gonzalez, and A. T. Cox, "North Atlantic Coast Comprehensive Study–Coastal Storm Hazards from Virginia to Maine," Vicksburg, MS: US Army Engineer Research and Development Center, ERDC/CHL TR-15-5, 2015. [Online]. Available: http://hdl.handle.net/11681/7715.

[28] Z. Liu, M. L. Carr, N. C. Nadal-Caraballo, M. C. Yawn, A. A. Taflanidis, and M. Bensi, "Machine Learning Motivated Data Imputation of Storm Data Used in Coastal Hazard Assessments," *Coastal Engineering*, vol. 190, 2024, doi: https://doi.org/10.1016/j.coastaleng.2024.104505.

[29] Z. Liu, M. L. Carr, N. C. Nadal-Caraballo, L. A. Aucoin, M. C. Yawn, and M. Bensi, "Comparative Analysis of Joint Distribution Models for Tropical Cyclone Atmospheric Parameters in Probabilistic Coastal Hazard Analysis," *Stochastic Environmental Research and Risk Assessment*, vol. 38, pp. 1741–1767, 2024, doi: https://doi.org/10.1007/s00477-023-02652-5.

[30] Z. Liu, M. L. Carr, N. C. Nadal-Caraballo, M. C. Yawn, and M. Bensi, "A Bayesian Network Method for Deaggregation of Coastal Hazards," *Natural Hazards (under review)*, 2024.

[31] A. Al Kajbaf and M. Bensi, "Application of surrogate models in estimation of storm surge:A comparative assessment," *Applied Soft Computing*, vol. 91, p. 106184, Jun. 2020, doi: 10.1016/j.asoc.2020.106184.

[32] A. P. Kyprioti, A. A. Taflanidis, N. C. Nadal-Caraballo, and M. O. Campbell, "Incorporation of sea level rise in storm surge surrogate modeling," *Natural Hazards,* vol. 105, no. 1, pp. 531–563, Jan. 2021, doi: 10.1007/s11069-020-04322-z.

[33] J. Zhang, A. A. Taflanidis, N. C. Nadal-Caraballo, J. A. Melby, and F. Diop, "Advances in surrogate modeling for storm surge prediction: storm selection and addressing characteristics related to climate change," *Natural Hazards*, vol. 94, no. 3, pp. 1225–1253, Dec. 2018, doi: 10.1007/s11069-018-3470-1.