POLITECNICO DI TORINO
Repository ISTITUZIONALE

Towards Robust Machine Learning Models for the Estimation of Mooring-Line Tension in Floating Wind Turbines under Unknown Sea Conditions

Original
Towards Robust Machine Learning Models for the Estimation of Mooring-Line Tension in Floating Wind Turbines under Unknown Sea Conditions / Kämmerling, M., Nava, V., Mendikoa, I., Ser, J.D.. - In: JOURNAL OF PHYSICS. CONFERENCE SERIES. - ISSN 1742-6588. - 3224:(2026). (The Science of Making Torque from Wind (TORQUE 2026) Bruges (Bel) 3-5 June 2026) [10.1088/1742-6596/3224/6/062030].

Availability:
This version is available at: 11583/3011670 since: 2026-06-04T08:57:10Z

Publisher:
IOP Science

Published
DOI:10.1088/1742-6596/3224/6/062030

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

15 November 2026

Journal of Physics:
Conference Series

PURPOSE-LED
PUBLISHING™

PAPER • OPEN ACCESS

Towards Robust Machine Learning Models for the Estimation of Mooring-Line Tension in Floating Wind Turbines under Unknown Sea Conditions

To cite this article: Max Kämmerling *et al* 2026 *J. Phys.: Conf. Ser.* **3224** 062030

View the article online for updates and enhancements.

You may also like

- AI-enhanced digital twins for predictive maintenance of floating wind structures using surrogate models
Guillermo Rodríguez Llorente, Jose María Moreu Gamazo, Guillermo Silvente Niñirola et al.
- Predicting Mooring Tensions of a Floating Wind Turbine Using Motion Sensors and Neural Networks
Mohammad Arif Payenda, Shuaishuai Wang, Andreas Prinz et al.
- Predictive Maintenance of Floating Offshore Wind Turbine Mooring Lines using Deep Neural Networks
N Gorostidi, V Nava, A Aristondo et al.

This content was downloaded from IP address 79.116.238.137 on 04/06/2026 at 09:44

Towards Robust Machine Learning Models for the Estimation of Mooring-Line Tension in Floating Wind Turbines under Unknown Sea Conditions

Max Kämmerling[1,2], Vincenzo Nava[3], Iñigo Mendikoa[1] and Javier Del Ser[1,2]

[1]TECNALIA, Basque Research & Technology Alliance (BRTA), 48160 Derio, Spain
[2]University of the Basque Country (UPV/EHU), 48940 Leioa, Spain
[3]Politecnico di Torino, 10129 Torino, Italy

E-mail: maximilian.kammerling@tecnalia.com

Abstract. Given the practical limitations of direct measurement in harsh offshore environments, the indirect estimation of mooring-line tension in Floating Offshore Wind Turbines (FOWTs) constitutes a necessary approach for Structural Health Monitoring (SHM) and Remaining Useful Life (RUL) assessment. This study investigates the performance and robustness of several Machine Learning (ML)-based regression models for estimating mooring tensions from floater motion measurements under both known (in-distribution) and unknown (out-of-distribution, OOD) sea states. The analysis is based on wave-tank experiments performed on a scaled model of a typical floating structure with a catenary station-keeping system representative of a FOWT, namely the HarshLab floating laboratory installed at the Biscay Marine Energy Platform (BiMEP). A Multi-Layer Perceptron (MLP) augmented with temporal context outperforms not only traditional tree-based ensemble models but also the more complex sequence-learning architecture evaluated in this study, namely the Long Short-Term Memory network (LSTM). Our results show that all models achieve high accuracy under sea conditions represented in the training data, while their performance degrades as test conditions increasingly deviate from the training distribution, particularly under more severe sea states. In light of this observed performance degradation under OOD conditions, the ensemble variance of a tree-based ensemble regressor is further evaluated as a model degradation indicator (MDI), demonstrating its effectiveness in detecting global estimation-quality deterioration and potentially signaling when model retraining is required to incorporate the effects of previously unseen sea-state dynamics on mooring-line tension.

1. Introduction

From the economic perspective, traditional bottom-fixed offshore wind turbines are only viable in waters shallower than approximately 50–60 meters. This limitation has driven the development of FOWTs, which are suitable for deployment in deeper waters where no continental shelf is present. For FOWTs, mooring systems are essential components of robust station-keeping


Content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

mechanisms. Nevertheless, prevailing economic drivers within the renewable energy industry often promote non-redundant design paradigms, meaning that the failure of a single mooring line may result in the loss of the entire system. This highlights the central role of mooring systems as a critical subsystem within FOWTs. Common sources of mechanical failure in mooring lines are due to excessive loading and fatigue, both of which are influenced by the axial tension at the mooring attachment points (pad eyes) – hereafter referred to as "mooring tension". Monitoring this mooring tension is therefore crucial for estimating the RUL and enabling effective SHM of floating wind systems.

Unfortunately, the direct measurement of mooring tension is technically challenging due to the harsh, submerged offshore environment. These conditions not only make direct methods costly to deploy, but also raise issues with their long-term reliability, especially given the accumulation of marine growth that requires continuous maintenance [1].

A practical alternative is the indirect estimation of mooring tensions. This can be achieved by monitoring the motions of the floating platform, which are easier to capture and generally provide more robust data in different spatial Degrees of Freedom (DoFs) over time. Since mooring tensions and floater motions are strongly correlated with each other, using motion data to estimate mooring tension presents a promising and cost-effective solution. Recent studies including those by [2], [3], [4], and [5] have explored the use of ML techniques to derive models that relate floater motions to mooring tensions.

Unfortunately, a well-known limitation of data-driven models—particularly ML-based regression approaches—is their inability to reliably extrapolate beyond the distribution of the training data, a challenge commonly referred to as OOD generalization [6]. This limitation is especially problematic in offshore environments, where sensor data from extreme sea states are inherently scarce due to their infrequent occurrence. Moreover, during such high-energy conditions, onboard sensing and communication systems often suffer from outages or complete failure, further reducing the availability of reliable measurements.

To address these practical constraints, there remains a strong operational need for models capable of reliably estimating mooring tensions across a broad range of sea conditions, including those for which no sensor measurements were available during the initial training phase. In response to this need, the research presented in this work is structured around two central research questions (RQ):

- **RQ1:** *Which ML regression models are most suitable for estimating mooring tensions from floater motion inputs, and how well do these models generalize to sea states that lie outside the distribution of the training data?*
- **RQ2:** *Can an MDI be developed that quantifies the uncertainty of the regression models' estimates and reliably detects when their estimation quality deteriorates, thereby indicating when model retraining or recalibration is required?*

The remainder of this paper is organized as follows. Section 2 describes the experimental setup and the dataset obtained from the wave-tank campaigns, along with the regression models and evaluation methodology adopted in this study. Section 3 presents and discusses the experimental results, addressing the two research questions posed in this work. Finally, Section 4 summarizes the main findings and outlines directions for future research.

2. Methodology

This section presents the methodology adopted in this work. After describing the wave-tank experimental setup and the dataset obtained from the measurement campaigns, the regression

models and feature construction are introduced. The methodology distinguishes between two separate experimental settings: one dedicated to **RQ1**, focusing on the comparative evaluation of regression models and their out-of-distribution generalization performance, and a second dedicated to **RQ2**, focusing on the formulation and evaluation of MDIs for detecting estimation-quality deterioration.

Dataset. The data used in this research are drawn from measurement campaigns conducted in wave-tank experiments with a Catenary Anchor Leg Mooring (CALM)-type buoy. It is a 1:13.6 Froude-scale model of the *HarshLab*, a floating laboratory owned and operated by TECNALIA [7]. HarshLab is anchored at BiMEP [8], located in the Gulf of Biscay, 2.96 km offshore from the village of Armintza (Bizkaia, Spain) at a water depth of 65 m, as shown in Figure 1a. The purpose of this laboratory is to validate and test materials, components, and equipment in a real offshore environment under controlled conditions. It serves as a testing ground for developments in the offshore industry, enabling accurate assessment of new technologies prior to full-scale deployment.

The 1:13.6 scaled wave-tank model, shown in Figure 1b, consists of three primary structural components: a large-diameter lower cylinder, a narrower upper cylinder, and a boat landing platform. It is held in position by three steel-chain catenary mooring lines, as illustrated in Figure 1c. This configuration provides a controlled experimental environment in which the floating structure's responses arise from wave-induced loads and the dynamics of its mooring system. A detailed description of the wave-tank model and the associated experimental setup can be found in [9]. Due to the lack of publicly accessible measurement data from actual FOWTs, this buoy wave-tank model has been selected as the use case. It serves as a representative proxy for the typical floating structure and catenary station-keeping system of a FOWT.



(a) **(b)** **(c)**

Figure 1. (a) Real scale floating laboratory HarshLab (b) Scaled wave-tank model that provides the data for this study. (c) Chain and load cell of wave-tank model's mooring system.

During the wave-tank experiments, a range of irregular sea states was simulated. Each experiment lasted 15 min at prototype scale and differed in terms of significant wave height H_s, peak period T_p, and the peak enhancement factor γ of the JONSWAP (Joint North Sea Wave Project) spectrum. The experimental setup recorded multiple parameters, including the 6-DoF motions of the floating structure via an optical tracking system, as well as the tension in each mooring line using in-line load cells. Table 1 presents the four sea states considered in this study. They are chosen to represent a wide range of conditions according to their wave energy flux P_{wave}, ranging from very calm conditions (Sea State A) to the most energetic conditions (Sea State D).

Table 1. Characteristics of the sea states considered in the wave-tank experiments that provide the data for this study. Values of H_s, T_p, and P_{wave} are converted and presented at full scale.

ID	Sea State Name	H_s [m]	T_p [s]	γ	P_{wave} [kW/m]	RQ1	RQ2, intra	RQ2, extra
A	Low-Energy	1.0	5.8	1.4	2.84	test	train	test
B	Moderate-Energy	1.9	9.2	1.5	15.87	train	test	train
C	High-Energy	3.0	10.4	1.5	45.88	test	n.c.	n.c.
D	Very High-Energy	4.3	12.3	1.9	111.49	test	train	test

Note: n.c.: *not considered.*

RQ1. To address the first research question, we compare different regressor architectures for the task of estimating mooring tension from floater motions, with a particular focus on their extrapolation capabilities when applied to data outside the range of the training set. The Moderate-Energy Sea State B serves as the training data. To evaluate the OOD generalization capability of the models (i.e., their ability to operate under conditions not represented during training), the regressors are tested on the time series of the Low-Energy Sea State A, whose P_{wave} lies below the training range, and on the High-Energy Sea State C and the Very High-Energy Sea State D, whose P_{wave} exceeds it. Among these, Sea State C is closer to the training conditions, whereas Sea State D constitutes a substantially more extreme scenario characterized by larger values of H_s and T_p.

We consider several regression models for the above comparison. Specifically, two tree-based ensemble models, namely a Random Forest Regressor (RF) and an Extra-Trees Regressor (ET), as well as two neural network architectures, a MLP and an LSTM, are fitted for this purpose. A Bayesian optimization procedure is employed to tune the hyper-parameters of each model, using 50 exploratory and 50 exploitative trials.

The measured motions of the floating structure serve as input features for all models developed in this work. The experimental setup is mechanically symmetric with respect to the central axis of the wave tank, and the wave direction is unidirectional. Because the floating structure itself is axis-symmetric relative to the incoming waves, no motion is theoretically expected in the sway, roll, or yaw DoFs. The sensor data confirm this, as the measurements in these three DoFs remain essentially constant with only minimal variance, likely attributable to sensor noise or to minor inertial, mechanical, or forcing eccentricities. Consequently, the input feature space is reduced to the three DoFs exhibiting meaningful dynamics: surge, heave, and pitch.

To incorporate temporal context into the regression models, the input space is augmented by including, in addition to the current time step, several preceding and succeeding time steps of the three DoFs. The number of past and future time steps are treated as additional hyper-parameters and are optimized during the Bayesian search.

The target variable is the current time step of the mooring tension of mooring line 1, i.e., the port-forward line. All input and output time series are down-sampled to a frequency of 10 Hz.

RQ2. Regarding the second research question, a MDI is developed to identify when the estimation quality of the regression models deteriorates. The proposed approach constructs an ensemble comprising a large number of sub-learners, where the variance among their individual predictions—referred to as ensemble variance—is used as a proxy for epistemic uncertainty. High ensemble variance indicates an elevated likelihood of estimation error and may signal model degradation under distributional shift, thereby suggesting that augmenting the training dataset could be necessary. Conversely, low ensemble variance suggests stable model behavior and reliable predictions, indicating no immediate need for retraining.

Although the experiment of **RQ1** identified the MLP as the best-performing estimator for this use case, tree-based ensemble methods are evaluated as MDIs in this work due to their computational efficiency and interpretability. These models generate predictions based on transparent, human-readable decision rules and enable direct quantification of feature importance without relying on post-hoc attribution techniques.

Among available tree-based ensemble methods, an ET is selected. Compared to the widely used RF, an ET introduces an additional layer of randomness during split selection: although a random subset of features is considered at each split, multiple thresholds are drawn randomly and the best among them is selected. This additional randomness increases heterogeneity among the sub-learners, which is desirable for estimating epistemic uncertainty.

The ET model developed for **RQ2** is based on the hyper-parameter values optimized for the same model used in **RQ1**, except for the proportion of randomly selected input features considered at each node of every tree in the ensemble, which is reduced to 10 %. Lowering the value of this hyper-parameter further increases heterogeneity across sub-learners. This is acceptable, given that the estimation performance on the training dataset remains high, as will be documented in Section 3.

Two OOD generalization experiments are conducted in the context of **RQ2**, differing only in the training data used:

- The first is the *extrapolative* OOD experiment, in which the model is evaluated on sea states whose conditions lie outside the range spanned by the training data. Specifically, the model is trained on the Moderate-Energy Sea State B and evaluated on the Low-Energy Sea State A and the Very High-Energy Sea State D.
- The second is the *interpolative* OOD experiment, where the model is evaluated on a sea state whose conditions lie between those represented in the training data. In this case, the model is trained on the Low-Energy Sea State A and the Very High-Energy Sea State D and tested on the Moderate-Energy Sea State B.

These two experiments are conducted to assess model degradation both when test conditions exceed the envelope of the training data (extrapolative) and when they fall between distinct training domains (interpolative), thereby providing a more complete characterization of the models' robustness under distributional shift. The exact split of training and test data—indicating which sea states are used for training and which for testing across all experiments of **RQ1** and **RQ2**—is provided in Table 1.

3. Results & Discussion

This section presents and discusses the results of the experimental analysis. The performance of the regression models is evaluated under both in-distribution and out-of-distribution sea conditions, and their robustness and limitations are analyzed in relation to the two research questions addressed in this study.

RQ1. The hyper-parameter configurations obtained via Bayesian optimization for all models investigated in the context of the first research question are summarized in Table 2. Model evaluation is performed using the Root Mean Square Error (RMSE) and the coefficient of determination R^2 as performance metrics. The evaluation of these metrics for all regression models is presented in Table 3. Overall, all models perform well on the training time series, but their estimation quality deteriorates as the test conditions deviate further from the training conditions. Among all models evaluated, MLP achieves the highest accuracy for all three unseen sea states,

reaching an R^2 of 0.9704 for Sea State A, 0.9618 for Sea State C and 0.8754 for Sea State D. Interestingly, the LSTM does not outperform the MLP, despite its theoretical advantage in capturing temporal dependencies. This behavior may arise because the dataset exhibits relatively short temporal correlations, limiting the benefits of the LSTM's memory mechanisms relative to a simpler feedforward architecture.

An additional observation arises when examining the model performance on less energetic sea states. Although one might expect a substantial degradation in accuracy in such cases, this is not always observed. This behavior can be explained by the fact that not all waves within a high-energy sea state are extreme. A considerable portion of the time series corresponds to moderate, non-peak conditions whose dynamics resemble those of a lower-energy sea state. Consequently, a model trained on a more energetic sea state is still exposed to a wide spectrum of moderate wave-induced responses, which facilitates accurate prediction when evaluated on the milder Sea State A. This also explains why, for certain models (most notably, the LSTM), the performance on Sea State A can be comparable to, or even slightly better than, the performance on the training Sea State B. Taken together, these results support the interpretation that the less energetic sea state partially overlaps with the regular (non-extreme) dynamical regime present within the higher-energy training sea state.

A closer examination of the estimated time series shown in Figures 2a to 2d offers further insight into the trends observed in the global error metrics. For the training Sea State B, the predicted tensions closely match the measured observations, as shown in Figure 2a. However, across all test sea states, the largest errors occur during the pronounced peaks and deep troughs of the tension signal. The peak values are generally under-predicted, while the troughs tend to be over-predicted. This behaviour appears both when the test conditions are less energetic than the training data (Figure 2b) and when they are more energetic (Figures 2c and 2d).

Consistently with the degradation in global performance reported for Sea States C and D, the discrepancies become increasingly pronounced under more energetic conditions. As sea-state severity increases, both the frequency and the magnitude of these extreme tensions grow substantially. This systematic behaviour aligns with the global metrics and can be plausibly explained by the absence of similarly extreme tension levels in the training Sea State B; as shown in Figure 2a, such high-load peaks and troughs do not occur in the training data. Consequently, the models incur larger errors whenever they encounter previously unseen extreme-load events.

It is also important to note that these peaks and troughs are precisely the segments of the time series in which accurate prediction is most critical, given their relevance for fatigue assessment, structural prognosis, and the overall integrity management of floating offshore systems.

Nevertheless, the MLP distinguishes itself from the other regression models by exhibiting comparatively robust performance across all OOD scenarios. Its predictions remain highly accurate when evaluated on sea states less energetic than those included in the training data, and under high-load conditions its estimates deviate less from the measured values.

RQ2. Before analyzing the results of the MDI, it is important to note that the previously described adaptation of the ET hyper-parameters does not compromise estimation accuracy on the training data. In the extrapolative OOD experiment, the ET achieves a high $R^2 = 0.9772$ on its training Sea State B. Similarly, in the interpolative OOD experiment, the ET attains $R^2 = 0.9809$ on the training Sea State A and $R^2 = 0.9897$ on the training Sea State D.

Table 4 reports both the R^2 of the model estimations and the ensemble variance of the ETs for the two experiments, averaged over the entire time series. In the extrapolative OOD experiment, the averaged ensemble variance for the training Sea State B is 1.2841 N^2. For the unseen and substantially more energetic Sea State D, this value increases markedly to 10.0123 N^2, clearly

Table 2. Hyper-parameter values obtained via Bayesian optimization for all regression models of **RQ1**.

Model	Hyper-parameters
RF	`n_past_timesteps` = 25, `n_future_timesteps` = 7, `min_samples_leaf` = 2, `max_features` = 1.0, `max_samples` = 0.51
ET	`n_past_timesteps` = 20, `n_future_timesteps` = 23, `min_samples_leaf` = 2, `max_features` = 0.81, `max_samples` = 1.0
LSTM	`n_past_timesteps` = 25, `n_future_timesteps` = 12, `bidir`= `False`, `batch`= 128, `stateful` = True, `n_neurons` = 256, `n_layers` = 4, `lr` = 9.17e-4, `dropout` = 0.0
MLP	`n_past_timesteps` = 10, `n_future_timesteps` = 6, `n_neurons` = 512, `n_layers` = 4, `lr` = 1.49e-4, `dropout`= 0.0, `batch` = 512

Table 3. Evaluation of error metrics of regression models included in the experiments addressing **RQ1**.

	Sea State A (Test)		Sea State B (Train)		Sea State C (Test)		Sea State D (Test)	
Model	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
RF	0.8686	1.3034	0.9719	0.5852	0.8791	1.6730	0.6959	4.0418
ET	0.8513	1.3849	0.9866	0.4048	0.8981	1.5368	0.7337	3.7859
LSTM	0.9693	0.6286	0.9494	0.7861	0.9247	1.3214	0.7936	3.3252
MLP	0.9704	0.6188	0.9710	0.5946	0.9618	0.9403	0.8754	2.5989

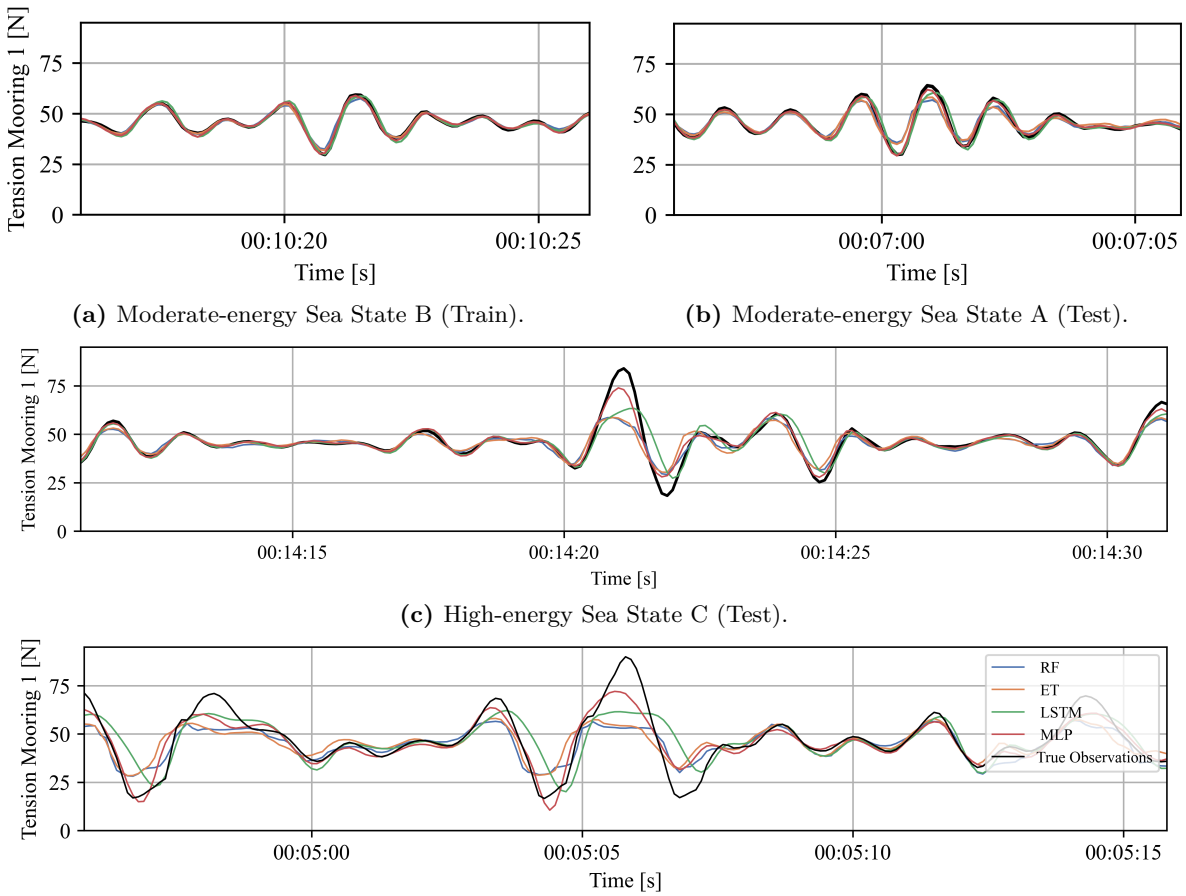
Note: RMSE values are expressed in Newtons (N).

indicating significant model degradation. A minor increase is also observed for the unseen Sea State A, reflecting a comparatively smaller reduction in model performance.

A similar behavior is observed in the interpolative OOD experiment. The averaged ensemble variance for the unseen test Sea State B is 4.0313 N^2, which is approximately 3.0 N^2 higher than in the training Sea State A and 0.4 N^2 higher than in the training Sea State D. This again indicates a degradation in estimation quality when the model is applied to a sea state not included in the training process, even when the unseen sea state lies within the overall range of conditions represented in the training data.

A noteworthy observation arises in the interpolative OOD experiment when examining the averaged ensemble variances reported in Table 4. Although Sea State D is a training sea state in this setting, its ensemble variance (3.6449 N^2) lies closer to that of the unseen test Sea State D (4.0313 N^2) than to that of the other training Sea State A (1.0499 N^2). At first glance, this may suggest that the ensemble variance is behaving less reliably as an MDI in the interpolative case. However, a more detailed examination exposes that this behavior is largely consistent with the relatively small degradation in estimation performance observed for the interpolative test sea state. Indeed, the ratio between the increase in ensemble variance and the corresponding decrease in estimation accuracy remains comparable to that observed in the extrapolative experiment.

For example, in the interpolative experiment, when comparing the Training Sea State D to the Test Sea State B, the R^2 decreases by only 0.15 while the averaged ensemble variance rises by 0.51 N^2. Similarly, in the extrapolative experiment, when comparing the Training Sea State B to the Test Sea State A, the R^2 decreases by 0.14 with a corresponding ensemble-variance increase of 0.39 N^2. These comparable ratios indicate that the ensemble variance still responds proportionally to the degradation in estimation quality, even in the interpolative scenario where



(a) Moderate-energy Sea State B (Train).

(b) Moderate-energy Sea State A (Test).

(c) High-energy Sea State C (Test).

(d) Very high-energy Sea State D (Test).

Figure 2. Comparison of measured and estimated mooring tension time series obtained from the regression models evaluated in **RQ1** for (a) the training sea state and the OOD sea states (b), (c) and (d).

the performance drop is less pronounced. Consequently, although the interpolative case shows less clear separation between training and test ensemble variances, the overall behavior of the ensemble variance remains coherent with the underlying changes in model accuracy, supporting its suitability as a coarse but informative MDI.

Taken together, these findings demonstrate that, when evaluated over entire time series, the ensemble variance constitutes a suitable MDI. It reliably signals when the regression model is applied to OOD data, and its indications are consistent in both the interpolative and extrapolative OOD settings. It should be noted, however, that part of the increase in the averaged ensemble variance may arise from the higher absolute magnitudes of the target variable in more energetic sea states. Across all experiments, higher ensemble variance coincides with more extreme environmental conditions, suggesting that the scale of the target variable itself contributes to the observed level of uncertainty.

For a more detailed evaluation of the suitability of the ensemble variance as an MDI, its local values were analyzed. These were compared with the absolute mean of the residuals computed across all sub-learners of the ET model in the extrapolative OOD experiment. Figure 3 presents the mean predictions of the ET model, together with its residuals and ensemble variance, for the training Sea State B. As expected, the ensemble variance remains low throughout the time series; however, it still displays desirable MDI-like behavior by showing slight increases at time instances where elevated residuals also appear.

Table 4. R^2 and time-series-averaged ensemble variance obtained from the ET model used as an MDI in the two experiments of **RQ2**.

Experiment	Training Sea State	Sea State A		Sea State B		Sea State D	
		R^2	MDI	R^2	MDI	R^2	MDI
Extrapol. OOD	B	0.8311	1.7982	0.9772	1.2841	0.6588	10.0123
Interpol. OOD	A & D	0.9809	1.0499	0.8473	4.0313	0.9897	3.6449

Figure 4 shows the corresponding evaluation for the unseen test Sea State D. In Figure 4a—and for most periods in Figure 4c—very high ensemble variance values coincide with large regressor residuals, and conversely, low ensemble variance aligns with small residuals. Although the ensemble variance does not always peak precisely at the same time instants as the residuals, there is a clear overall correspondence: sections with high prediction errors tend to coincide with sections of high ensemble variance.

Limitations. Unfortunately, the local evaluation also reveals notable weaknesses of ensemble variance as an MDI. As illustrated in Figure 4b, there exist extreme events that lead to large residuals while the ensemble variance remains consistently low. In such cases, the model degradation at these specific time instants would remain undetected.

Another limitation is that the ensemble variance is inherently bounded by the variability among the predictions of the sub-learners. As a result, it cannot "explode" in response to extreme events where the regressor exhibits very large errors. This constraint undermines its ability to function as a reliable local MDI and drives further research into new MDI formulations that overcome these shortcomings.

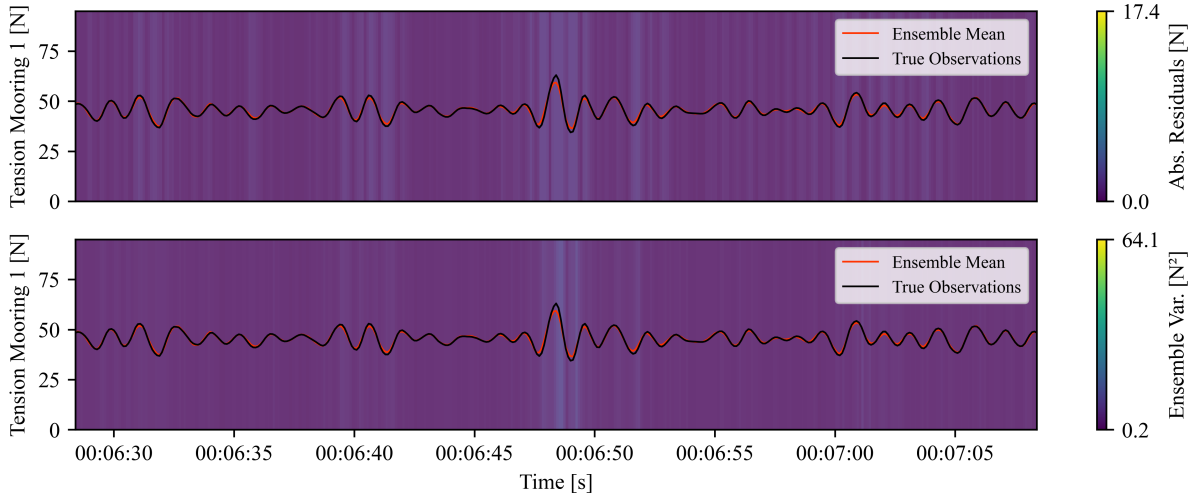


Figure 3. Time-domain evaluation of the ET model in the extrapolative OOD experiment of **RQ2**, shown for a 40 s segment of the time series of training Sea State B. The upper panel shows the absolute mean of the model residuals, visualized using background color; for visualization purposes, the residuals have been capped at the 99th percentile. The lower panel shows the ensemble variance, encoded in the same manner and serving as an MDI.

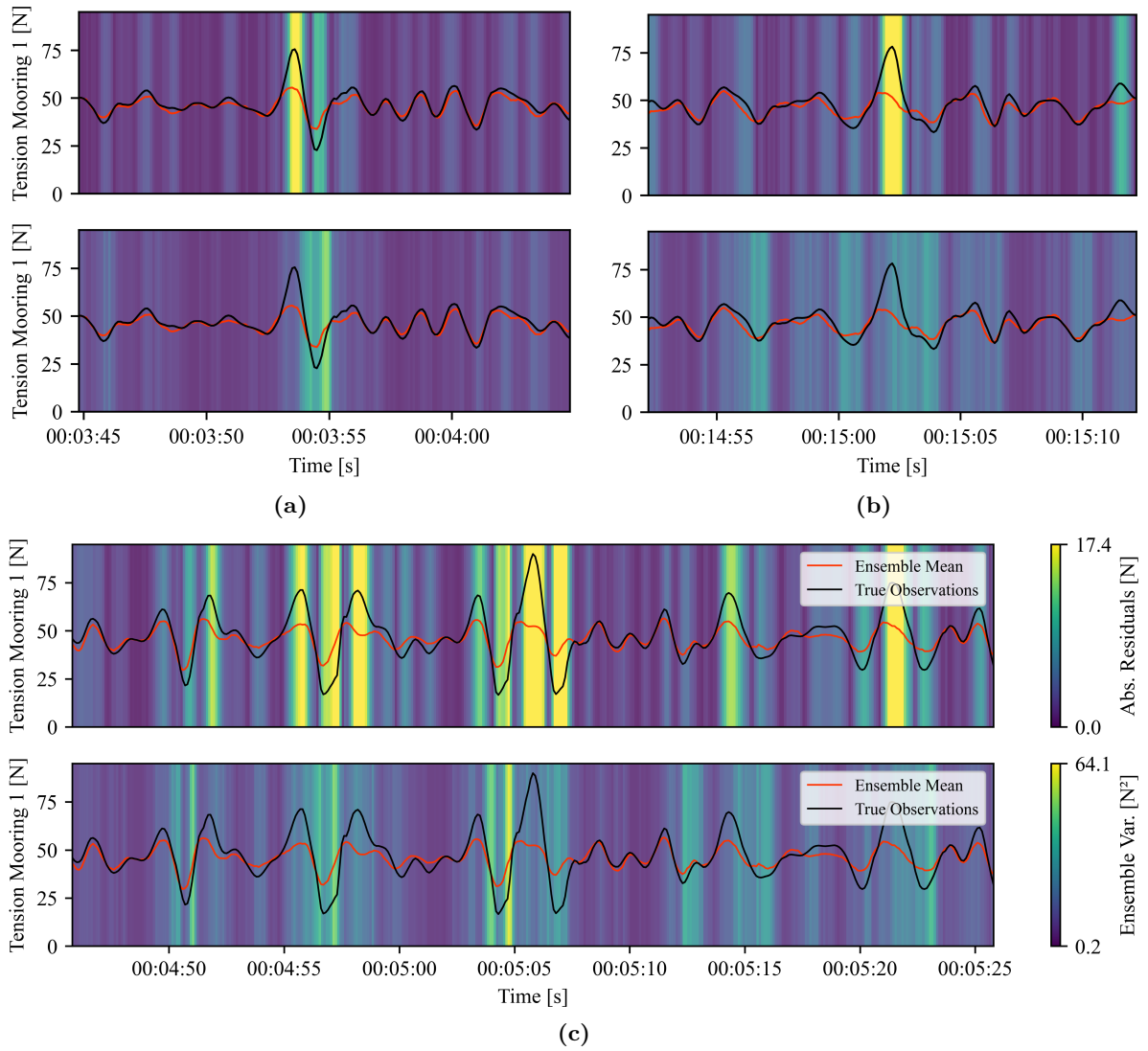


Figure 4. Time-domain evaluation of the ET model in the extrapolative OOD experiment of **RQ2** for three distinct 40 s segments of the unseen test Sea State D.

4. Conclusion & Future Research

This study has addressed the challenge of indirectly estimating mooring tension in FOWTs using ML-based regression models under previously unseen sea conditions. Motivated by the practical limitations of direct tension measurements in harsh offshore environments, the work focused on assessing both the OOD generalization capabilities of different regression architectures (**RQ1**) and the feasibility of uncertainty-based indicators for detecting model degradation (**RQ2**). We have come to the following conclusions:

- In response to **RQ1**, several regression models were evaluated for their ability to infer mooring tension from floater motion data, with particular emphasis on their behavior under OOD conditions. While all models achieved high accuracy within the training domain, their performance degraded progressively as the severity of the sea states exceeded that of the training data. The results highlight that providing explicit temporal context can be as effective as employing specialized time-series architectures. A feedforward MLP augmented with neighbor-

ing time steps consistently outperformed not only traditional tree-based ensemble regressors, but also more complex sequence models (e.g., LSTMs) in extrapolative scenarios.

- Addressing **RQ2**, the study explored the ensemble variance from an ET regressor as an MDI to detect degradation in its estimation quality. When evaluated globally over entire time series, the ensemble variance proved effective in distinguishing between in-distribution and OOD operating conditions, both in interpolative and extrapolative settings. This supports its use as a practical, computationally efficient indicator for identifying when the model is likely operating beyond its domain of validity. However, the local analysis revealed inherent limitations: certain extreme events produced large prediction errors without a corresponding increase in ensemble variance, indicating that degradation may remain undetected at specific time instants.

Future research should primarily focus on improving the extrapolative capabilities of ML regressors for mooring-tension estimation. Potential directions include architectural modifications that explicitly encode physical priors or scaling behavior, as well as hybrid approaches that combine data-driven models with simplified physics-based representations. Such developments may reduce the sensitivity of the regression models to distributional shifts and improve their robustness under unseen sea states.

Acknowledgements

This work was supported by the Basque Government through the BIKAINTEK (grant no. 021-B2/2023) and the ELKARTEK programs (grant no. KK-2024/00086, RUL-ET project).

References

[1] Arramounet V, Chemineau M, Castillo F, Mechinaud L, Mahfouz M Y, Pan Q, Willeke L, Bredmose H, Gözcü O, Pollini N, Borisade F, Trubat Casal P, Molins Borrel C, Guanche R and Somoano M 2023 Design practices and guidelines for mooring, anchoring system design Tech. Rep. D2.4 COREWIND Consortium horizon 2020 project, Ref. Ares(2023)2884487 URL `https://corewind.eu`

[2] Walker J, Coraddu A, Collu M and Oneto L 2022 Digital twins of the mooring line tension for floating offshore wind turbines to improve monitoring, lifespan, and safety *Journal of Ocean Engineering and Marine Energy* **8** 1–16 ISSN 2198-6444, 2198-6452

[3] Ma G, Jin C, Wang H, Li P and Kang H S 2023 Study on dynamic tension estimation for the underwater soft yoke mooring system with LSTM-AM neural network *Ocean Engineering* **267** 113287 ISSN 0029-8018

[4] Yang Y, Peng T and Liao S 2023 Predicting future mooring line tension of floating structure by machine learning *Ocean Engineering* **269** 113470 ISSN 0029-8018

[5] Fathnejat H and Nava V 2026 From augmentation to translation: Data generation by conditional hierarchical variational autoencoder, enhancing monitoring mooring systems in floating offshore wind turbines *Engineering Applications of Artificial Intelligence* **163** 112951

[6] Yang J, Zhou K, Li Y and Liu Z 2024 Generalized out-of-distribution detection: A survey *International Journal of Computer Vision* **132** 5635–5662

[7] Fundación TECNALIA Research & Innovation Harshlab `https://harshlab.eu/en/` accessed on 2025-08-08

[8] BiMEP - Biscay Marine Energy Platform `https://www.bimep.com/en/` accessed: 2025-12-10

[9] Touzon I, Nava V, Gao Z, Mendikoa I and Petuya V 2020 Small scale experimental validation of a numerical model of the HarshLab2.0 floating platform coupled with a non-linear lumped mass catenary mooring system *Ocean Engineering* **200** 107036 ISSN 0029-8018 URL `https://www.sciencedirect.com/science/article/pii/S0029801820301104`