Decision-making support for reeled pipeline installation using a machine learning based method

Øystein Døskeland[1, 2], Zhen Gao[1], Yuri Novoseltsev[2], Nicolas Holguin[2]
[1] Department of Marine Technology, NTNU, Trondheim, Norway
[2] Subsea 7 Norway AS, Stavanger, Norway

ABSTRACT

Installation of subsea pipelines used for transportation of hydrocarbons, water, or CO2, is carried out by ship-type installation vessels, which are highly sensitive to wave conditions. The prediction of installation loads in the pipeline is an essential input to the decision-making process for safe operation during offshore execution. Predictions may be required up to five days into the future. They can be produced from a physics-based simulation model with nonlinear calculations in the time domain and probabilistic representations of the response parameters based on multiple simulations for the forecasted wave spectra. Such calculations are computationally costly and, therefore, normally produced in advance by considering a set of generic parameter-based wave spectra. This paper describes how a machine learning model can be established, verified, and used to support decision-makers during a reeled pipeline installation operation. Compared to a physics-based simulation model, this method enables computationally efficient calculation of pipeline responses from forecasted wave spectra during offshore execution to provide more accurate input to decision-makers.

KEY WORDS: Machine learning; pipeline installation; modelling.

INTRODUCTION

Subsea pipelines are an essential part of the infrastructure required for offshore energy production and transportation. Some examples include transportation of oil and gas from a subsea production well to a production and storage unit, injection of water into a reservoir to control the well pressure, flowlines delivering oil and gas from an offshore field to an onshore terminal, transportation of hydrogen produced from offshore wind power to an onshore storage facility and transportation of CO2 from an onshore terminal for storage below the seabed. There is a significant need for such infrastructure also in the future.

The installation can only be done in operable weather conditions, meaning that the environmental conditions are sufficiently benign to avoid damage to the pipe or the installation equipment. Specialized monohull offshore construction vessels provide an efficient method for both transportation and installation of subsea pipelines, but monohull vessels are also very susceptible to wave conditions (especially wave period and direction) due to possible resonant motions. The process of classifying a condition as operable or non-operable typically includes dynamic structural analysis of the pipeline installation. This analysis is time-consuming, especially if a considerable number of environmental conditions and several time-domain simulations for each environmental condition need to be evaluated.

Offshore decision making for a pipeline installation project may be based on pre-determined environmental limits, often in terms of a limiting significant wave height (Hs) for a selected range of spectral wave periods (such as the spectral peak period, Tp) and mean wave direction. The calculated limits are then based on generic and parameterized wave spectra, such as JONSWAP (Hasselmann et al., 1973), which does not consider simultaneously occurring wave systems. In reality, both swell and wind sea systems are often present. The Torsethaugen wave spectrum takes into account both swell and wind sea systems but is limited to generic representations with coinciding directionality. In fact, the wave systems can vary significantly from the average/generic representations, and normally come from different directions.

Ship type installation vessels have natural heave, roll and pitch periods close to the main wave periods, leading to large resonant motions. A parameterized wave spectrum will often not capture the wave energy close to these periods well, leading to significant errors in the predicted vessel motion.

An alternative, and more accurate approach, is to base the decision-making on forecasted two-dimensional numerical wave spectra that are available closer to the time of execution. Whereas this approach is more accurate, it also requires a vast amount of analysis to be performed after the wave-forecast issuance to generate input for the decision-makers. Moreover, one needs to run multiple time-domain simulations with random waves for the same environmental condition to reduce the statistical uncertainty of the response parameters. Analysis that is purely based on FEM model simulations may not be feasible, because the wall clock time required for the simulations might not be tolerable.

Guarize et al. (2007) proposed a hybrid method for reducing the dynamic analysis simulation time for slender marine structures, introducing machine learning in the analysis process. Specifically, an artificial neural network (ANN) was trained from FEM based simulations, making it possible to generate long timeseries with relatively little computational effort. The scheme was applied to two cases: mooring line tension and SLWR (Steel Lazy Wave Riser) tension in deep waters. Other work discussed later in this article also applies the hybrid method for design of in-place systems, typically considering the long-term weather statistics.

The current work presents a novel application of the hybrid method for decision-making during offshore installation of pipelines based on two-dimensional forecasted wave spectra. Two different implementation strategies for the hybrid method are applied to a case study example. A detailed account is provided of the model input and output, as well as the neural network model architecture and parameter selection. The performance of the hybrid method is presented and compared to a pure physics-based approach.

RELATED WORK ON THE HYBRID METHOD FOR SLENDER STRUCTURES SUSPENDED FROM A FLOATING VESSEL

Machine learning methods that replace a physics-based model are useful if they can produce sufficiently accurate and reliable results. The input parameter selection and structure must be carefully considered when constructing the ANN model. Wu (2021) introduces the term Physics-Based Machine Learning (PBML) for a modelling approach that, similarly to the hybrid method, combines a physics-based model and machine learning method for predictions of future wave conditions. In this case, the spectral wave parameters Hs and Tp are predicted based on wind forcing more efficiently than pure physics-based modelling.

As described in the introduction, Guarize et al. (2007) applied the hybrid method for fatigue calculation in mooring lines and risers. They used a ANN with a single hidden layer and a single output node. The input parameters included the floater heave, surge, and sway motions at the current time step in addition to previous time steps. The size of the hidden layer and the length of the delay for previous time steps were selected as tuning parameters. The model was trained for each sea state.

Christiansen et al. (2013) proposed a hybrid method for the prediction of fatigue due to top tension in mooring lines connected to an FPSO. The ANN architecture included a single hidden layer. The input consisted of the 6 DOF floater motion, including previous time steps, and also previous time steps of the mooring line tension (target). The model was trained with fifteen different wave conditions, selected to represent the most significant sea states in the long term scatter diagram for the area. The resulting model was applied throughout the fatigue calculations.

Chaves et al. (2015) applied the hybrid method for fatigue in flexible pipes by predicting tension and curvatures in the bend stiffener. The network was trained on each sea state using a NARX (non-linear auto regressive network with exogenous input) architecture, which is similar to the structure used by Christiansen et al (2013).

Srikonda et al. (2018) applied the hybrid method to predict wellhead bending moment during operation, using a RNN (Recurrent Neural Network) trained from simulated data. The input to the machine learning model was inclination and acceleration at two positions close to the bottom of the riser: the lower riser joint (LRJ) and the blow out preventor (BOP), and tension at the LRJ. The model was trained using FEM simulations, while predictions were made using recorded MRU data during the operation.

Da Silva & de Arujo (2020) used a convolutional neural network combined with a NARX architecture (NARX-CNN) to predict tension and curvature on a flexible riser. Further work by da Silva et al. (2022) applied a RNN structure to predict responses at multiple positions on a catenary and lazy wave flexible riser.

PHYSICS-BASED MODEL IN PIPELINE INSTALLATION PROJECTS

Objective of dynamic finite element analysis in operational planning

Extensive dynamic finite element analysis is required during the detailed planning phase of pipeline installation operations. The analysis typically has multiple objectives: to increase the understanding of the operation, to determine the environmental limits, and to determine the amount of fatigue damage that might be accumulated during the installation process.

Each pipeline installation operation has three main phases: initiation, normal lay, and laydown. The catenary profile is different for each step of the operation. Additionally, more complex installations may involve riser sections with buoyancy attachments, strakes and similar ancillaries, or a pipeline with inline structures. The modelling supports operational planning by developing step-by-step static configurations. Dynamic sensitivity studies are applied to identify steps that are potentially critical in a dynamic situation. At each step, multiple quasistatic conditions of current level, current direction and vessel offset is possible. This is an iterative job where the findings are used to improve the operational procedures, provides detailed step-by-step installation tables, and highlight critical phases to the offshore crew where special caution and attention is needed.

The stochastic nature of waves means that the limiting environmental conditions must be based on a probabilistic criterion. For pipeline installations, the criterion is often a limit imposed on the most probable extreme (MPE) value for a representative duration. The MPE for a specific sea state can be found by fitting the extreme value distribution of the loads or, as a practical approach, conservatively approximated as the average extreme value from multiple realizations. Time series of the loads are obtained by introducing a random wave train generated from the target sea state and resolving the finite element model in time domain through time stepping.

In this paper, the finite element model is termed a physics-based model. It stands in contrast to machine learning models, which are data driven and have no prior knowledge of the laws of physics. The hybrid method can not replace the physics-based model. The physics-based model is a prerequisite for applying the hybrid method and a new objective is also introduced: to train the machine learning model. In turn, the machine learning model can be used for efficient evaluation of new wave conditions.

Case study model

OrcaFlex is a widely used tool for dynamic simulation of slender structures using a finite element modelling approach. The modelling and simulations for this paper are done using OrcaFlex v. 11.2.

In this article, a single pipe lay configuration is studied. The configuration is a pipeline in a catenary shape with two inline structures. The water depth is approximately 1200 m. The focus of the study is the prediction of maximum von Mises stresses at the top of the catenary, i.e., at the pipeline hang-off position on the stern of the vessel. Even though the von Mises stress is normally not directly used as an acceptance criterion, it represents a value closely related to the pipeline utilization, and is known to be difficult to predict for the top hang-off location using linear regression. An overview of the model is shown in figure 1.

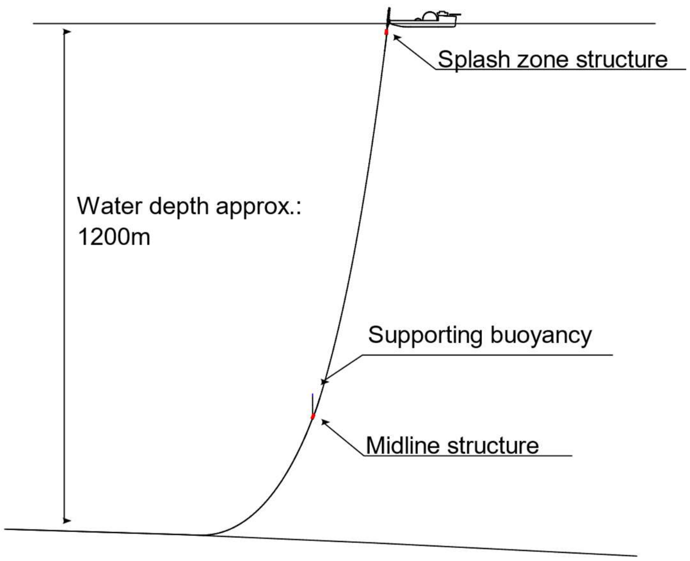

Figure 1: Pipeline installation example in 1200 m water depth with two inline structures as modelled with OrcaFlex v 11.2

The pipeline is 8" in diameter and is installed with two structures in the catenary. This represents a complex but realistic model of a pipeline installation operation that may represent one of the critical phases. The purpose of structures, such as those included in the model, may be for branching of the pipelines (in-line tee) or to enable intervention (such as valve operations/isolation or pigging). The weight of the structures is 59 metric tonne each.

The pipe catenary is modelled from the vessel exit to the seabed using 568 segments in total, meaning that there are 567 nodes. Each node represents three degrees of freedom (x, y and z global position), and each segment represents five degrees of freedom (bend angle at each side of the segment, spring offset at the middle of the segment, and torsion angle at the middle of the segment). The structures each have 6 degrees of freedom. In total the system has 1847 degrees of freedom.

A modal analysis shows that the first mode of the system has an eigenperiod of 160 seconds. This represents the longest memory effect of the system.

Simulation is done using an implicit time stepping scheme, where equilibrium is found at each time step through iterations. An adaptive time step is applied in this model, but with a maximum value of 0.05 seconds. Results are logged every 0.1 seconds. Typically, a simulation is run for 3 hours after a short buildup period, which is 20 seconds in this study.

A common assumption for pipeline installation projects is that the system can be adequately modelled without hydrodynamic coupling between the vessel and the pipeline, and that the presence of the pipeline has negligible influence on the 1st order wave frequency motion of the vessel. However, both vessel and pipeline are subjected to the same wave field. Moreover, the pipeline responses are strongly influenced by the motions of the top-end at the vessel. Vessel drift within the limits of the dynamic positioning capability and the effect of current is normally modelled as quasi-static processes. These assumptions are also made in the current case study.

MACHINE LEARNIG MODEL

For good performance of a machine learning model, a causal relationship between the input and output parameters is required. The selected machine learning model must have the ability to replicate the input-output relationship. The input should be sufficient to lead to the output, and the dynamic memory effect of the system must be captured. In the context of pipeline installation this input includes the wave field, and in extension the vessel motions.

Model architecture

The machine learning (ML) model for the case study is developed with a specific objective: to replicate the physics-based model's ability to predict the average extreme von Mises stress at the top of the catenary. The model produces the desired output (target) based on one or more inputs (predictors), and the relationship between the target and the predictors is learned by reviewing a sufficient number of samples.

The problem is approached by fitting an artificial neural network on time domain simulation results generated by the physics-based model. Neural networks are attractive because they inherently have the ability to learn nonlinear relationships with interaction between the predictors without any prior knowledge. A general reference on neural networks can be found in e.g. Goodfellow et al. (2016), but a brief description is provided herein in order to give some background to the terms and concepts used in this study.

The perceptron was introduced by Rosenblatt (1958) as a model that could mimic the biological brains behavior in storing and handling information. The perceptron consists of only two layers, an input layer and an output layer, which limits its application to solve linear relationships. However, the multi-layer perceptron (MLP) is a development that, by adding at least one hidden layer, can be shown to have universal approximation properties given that it has enough hidden units (Hornik et al., 1989). Visually, it can be seen as a network of nodes in multiple layers interconnected through weights. An example of such a representation is shown in figure 2. Each layer in the MLP is fully connected, such that each node in the layer is a function of all nodes in the previous layer. The model can be made more flexible and able to handle higher order relationships by increasing the width (i.e., the number of nodes in each layer) and/or the depth (i.e., the number of layers).

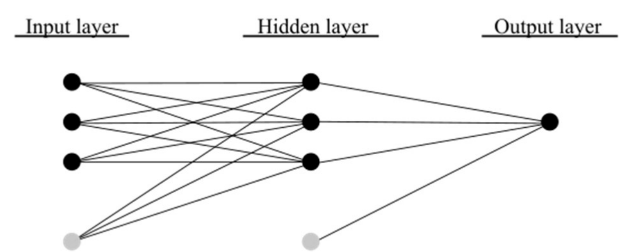

Figure 2: Visual representation of a MLP with one hidden layer. The black circles represent nodes, the grey circles represent bias, and the lines represent weights which are determined through the training

The MLP can be categorized as a feed forward network, because the information propagates only in a single direction. Although it is suitable for handling interactions and non-linear relationships, special considerations are needed to handle the dynamic nature of a pipeline installation model.

Christiansen et al. (2013) and Chaves et al. (2015) use a NARX (Non-linear Auto Regressive with Exogenous input) type architecture, e.g. (Chaves et al., 2015), to capture the dynamics of the system, where the target from the previous time steps is included as an input to the model, together with previous time steps from each predictor variable. The NARX model is trained as a feed forward network, using the true time series of the target during training, but it is recursive for predictions, as it uses the predicted target from the previous time steps for predicting the current time step.

Recurrent Neural Networks (RNNs) are a development of neural networks that is especially adapted to handle predictions of sequential data, by moving recurrently through an ordered sequence (such as a time series) and including the inner state of each layer at the previous step as an input to the model. This process is more complex, but it increases the model's ability to handle the dynamic behavior compared to the autoregressive architecture, since the inner state of the network contains substantially more information than the single output node. Plain RNNs have been shown to be difficult to train, but the problem is reduced by introducing special units/cells such as the LSTM (Long Short Term Memory) (Hochreiter & Schmidhuber, 1997). The GRU (Gated Recurrent Unit) cell structure (Cho et al., 2014) is a somewhat simpler alternative to LSTM which addresses the same issues. Using these architectures, RNNs have shown great performance on a number of problems. As earlier mentioned, Srikonda et al. (2018) has applied this architecture for prediction of loads in a drilling riser with good results. In this study, a recurrent neural network is applied using the GRU cell structure.

The LengthNet model proposed by da Silva et al. (2022) applies a recurrent network structure both through time and space (the length of the catenary). This is an interesting extension to the RNN structure that increases the prediction speed when multiple points need to be assessed but is only found to give slight improvements in the prediction accuracy. The approach is not adopted in the current study. Rather, the neural network models developed in this study are trained to output results for a single location on the catenary. Even though several locations may be of interest, these often correspond to different quasi-static conditions, such as current direction and vessel drift. Critical stresses might, for example, occur both at the top of the catenary and at the sag-bend, where the top location is critical if the catenary is stretched, and the sag-bend location is critical if it is compressed. A single location per model is therefore often sufficient.

Output and input selection

As previously described, the target output of the neural network is the time series prediction of von Mises stress at the top of the catenary. This means that the model is used to solve a regression problem. Alternatively, it could be treated as a classification problem. For a classification problem, the model is trained to find the boundary, which in this case separates acceptable and unacceptable loading. The model is then fine-tuned to this boundary, and less affected by the possibly different physics that govern responses far away from the boundary. It is also less sensitive to changes in the acceptance criteria, e.g., between a criterion on the bending moment, the von Mises stress, or a combined loading criterion.

The input-output selection determines what type of problem the ML model needs to solve. At the highest level, the two-dimensional wave surface alone determines the load in the pipeline. Information about the wave field should therefore be sufficient to make predictions since it determines both the vessel motions and direct wave loads on the structure. Current and other conditions are treated as quasi static. Under those conditions, the ML model needs to infer the vessel RAOs as well as the dynamic behavior of the catenary to make accurate predictions for an arbitrary point on the pipeline.

At the lowest level, the load on a specific pipeline segment at a specific point in time is determined by an equilibrium of loads transferred from adjacent nodes, external loads on the segment, linear and quadratic damping, and dynamic motion. In this scenario, the scope of the ML model is limited to learning the explicit equilibrium equation. However, this is not practical for real-world application, since it would require simulation in a physics-based software up to the point in time of the prediction to resolve the input.

A practical approach is to reduce the complexity of the problem as much as possible by providing the lowest level input to the neural network that can be resolved with little computational effort. Since the physics-based model assumes that 1^{st} order wave frequency motion of the vessel is unaffected by the pipeline, both wave elevation and vessel motions are resolved explicitly and can be generated with negligible computational effort for long time series. In this paper, vessel responses are provided at the hang-off point of the pipeline, which means that the neural network does not need to infer any of the vessel properties, and the model can in principle be used across different vessels. The wave field is represented through water particle velocity at the surface and points close to the surface, specifically, 10m, 20m and 30m depth along the pipeline catenary. An overview of the predictors is provided in table 1.

Table 1: Overview of the predictors provided to the ML model

Parameter	# Predictors
Vessel surge, sway, and heave motion	3
Vessel roll, pitch, and yaw motion	3
Water particle x-velocity, evaluated at surface, 10m, 20m and 30m depth	4
Water particle y-velocity, evaluated at surface, 10m, 20m and 30m depth	4
Water particle z-velocity, evaluated at surface, 10m, 20m and 30m depth	4
Total	**18**

Note: The vessel surge, sway and heave motion are evaluated at the position where the pipeline exits the tensioner system on the vessel (hang-off point). The water particle velocities are evaluated in a global coordinate system that is aligned with the vessel, where x is positive in positive surge direction and z points vertically upwards.

The predictors are provided to the model as time series, and the total number of data points included in each sample therefore depends on the choice of time step (sample interval) and the window used for each sample (sample window length). An example calculation is provided in table 2.

Table 2: Example calculation of the total number of predictor values provided in each sample

# Predictors	Sample interval	Sample window length	Total data points
18	0.5s	10s	$18 \cdot (10/0.5 + 1)$ $= 378$

The desired output to be used in decision making is a probabilistic description of installation criteria, however, the output from the ML model is a single deterministic value shown in table 3, representing the value of the target variable at the end of the sample window.

Table 3: Overview of the target output of the ML model

Parameter	# Targets
Max von Mises stress at the top of the pipeline	1

For practical use it is required to evaluate the model sequentially to generate time series of the target that can be used in a probabilistic assessment. In this study, the representative load is calculated as the average of the extreme maxima from five 3-hour realizations. The number of evaluations required on the ML model to generate this output is shown in table 4.

Table 4: Number of evaluations required on the machine learning model to generate a probabilistic output

# Seeds	Simulation length	Time step	Total evaluations
5	10800s (3-hours)	0.1s	$5 \cdot 10800/0.1$ $= 540'000$

PRACTICAL BUILDING AND TRAINING OF THE ML MODEL

Datasets for training, validation, and testing

In accordance with the hybrid method, training data is generated by running multiple simulations with the physics-based model for different sea states. Each sea state represents a unique forecasted two-dimensional numeric wave spectrum for a location offshore West Africa. Two data sets are generated: The first dataset consists of 300 simulations, each 15 minutes long from a unique sea state between 2.0m Hs and 3.5m Hs. In total 4500 minutes. The second dataset consists of simulations from 10 unique sea states between 2.0m and 2.5m, each with five different random realizations and three hours long. In total 50 simulations and 9000 minutes. The sea states included in the two datasets are non-overlapping. The sample interval is 0.1 seconds for both.

The simulation time is approximately 1:1, meaning that it takes about three hours on the clock to produce a simulation of three hours length on a single processor.

Two different model approaches are presented in this study and listed below.

1) The unseen sea state model: The model is trained for prediction on unseen wave spectra, meaning that the time series used for testing are generated from a wave spectrum which is not included in the training data set.

2) The single sea state model: The model is trained for prediction on a single known wave spectrum, meaning that the time series used for both training and testing are generated from the same single wave spectrum

The unseen sea state model has a very high prediction efficiency during operation, since it is trained prior to the operation, and, in principle, does not require any further evaluation of the physics-based model. This approach is adopted by e.g. Christiansen et al. (2013). On the other hand, the single sea state model has a much smaller scope and requires less training data, but it must be trained for every wave spectrum that needs to be evaluated. In sum, the required training data may be larger than for the unseen sea state model, and it needs to be processed during operation. An advantage of the single sea state model is that it will not encounter un-expected wave spectra, such as an unseen wave direction that was not part of the training set, or poorly represented. This approach is applied by e.g. Guarize et al. (2007).

The unseen sea state model is trained from the first data set with 15-minute simulations. The dataset is split into a training set and a validation set during training. The training set contains 270 simulations, and the validation set includes 30 simulations. To ensure maximum independence between the test set and validation set, the simulations are sorted chronologically by forecast date before splitting, and the validation set is taken from the end of the list.

The single sea state model is trained from the second dataset with 3-hour simulations. Simulation length between 5 minutes and 60 minutes is used for the training, taken from the beginning of the time series, and the validation data is taken as 15 minutes from the end of the data set.

The performance of both models is tested on the second data set with 3-hour simulations. Even though part of the test data for the known sea state model is not independent of the training and validation data, this represents a realistic evaluation of the performance on actual operations since the output from the physics-based model is also available to decision makers with this method.

Model parameter selection

The building and training of a recurrent neural network is not straight forward and involves multiple parameter choices. Some parameters can be selected from experience, but there is often no obvious choice, and selection must be done by repeated training and testing to identify the best possible configuration. The parameters selected for tuning the model are called hyper-parameters. For this study they are separated into three different types: convergence parameters, architecture parameters and physical parameters.

The neural network is trained by finding the weights that minimize the prediction error (loss) for a set of training samples. This process is done using a gradient based optimizing algorithm, and in the case study, the Adam optimization algorithm is applied. The convergence parameters include learning rate, batch size, and dropout. The learning rate determines the step size for each gradient update and should not be selected too large; otherwise, the training will be unstable. A too low value will require an excessive number of iterations to converge.

During optimization, gradients can be updated on each sample (stochastic gradient decent), after reviewing all samples (batch gradient decent) or using an intermediary setting, which is more common and applied in this case study (mini-batch gradient decent). The number of samples included for each update represents the batch-size.

Dropout is a method for regularizing neural networks, where a random selection of the weights is assigned a value of zero between two layers, for the case study, dropout is included after each of the GRU layers. The dropout is re-selected between each sample. The dropout rate specifies the ratio of the weights that are nulled.

Table 5: Convergence parameter selection

Parameter	Value
Learning rate	5e-4
Batch size	32
Dropout	0.1

The number of RNN layers and the number of units within each layer determines the network architecture. As the number of units and layers increase, it allows the network to handle problems of greater complexity. However, the issues related to overfitting and need for regularization of the network also increases. For simplicity, the same number of units are used in each layer.

Table 6: Architecture parameter selection

Parameter	Value
Number of layers	4
Number of units each layer	32

RNNs are trained on samples of limited length time series. These samples are found by splitting the training data files into smaller time series sections and resampling them at specified time steps. The sample windows may overlap but should be separated sufficiently so the target output value can be assumed independent. A separation of 1 minute is adopted, which should be more than sufficient.

In principle, the maximum length time series and minimum choice of time step increase the amount of information provided to the network. This choice maximize performance up to the model's highest modal period, which is approximately 150 seconds. Even so, both long sample durations and short time steps increase the sequence length. For practical implementation this increases the need to regularize the network and, therefore, may not lead to increased performance.

Table 7: Physical parameter selection

Parameter	Value
Sample window length	10 seconds
Sample interval	0.3 / 0.6 seconds

The parameters applied in the case study are provided in tables 5, 6 and 7. They are identified based on extensive and iterative tuning of the hyper parameters. In this tuning process, it is essential to separate the validation data and the training data, otherwise, there is a risk of overfitting against the tuning parameters. While separate training and validation sets are used in this study for both approaches, a cross-validation approach is also possible, and could be particularly useful for the single-sea state model creation.

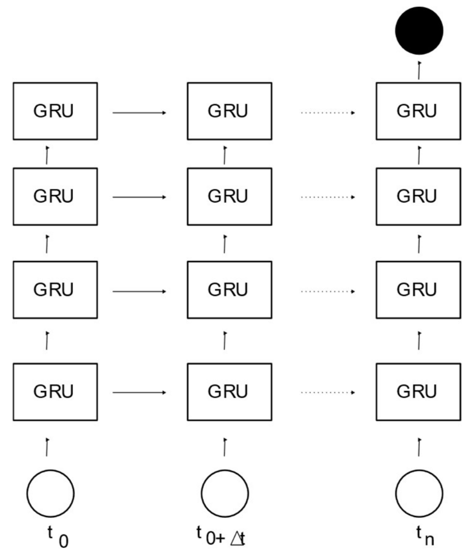

Figure 3: Outline of RNN model architecture

Performance

The performance of the model is measured by the predicted average extreme value over the five seeds in each of the 10 sea states represented in the 3-hour simulation dataset. Results for the unseen sea state model shows that the prediction error of the von Mises stress at the top of the catenary vary between 1 and 12 MPa.

Table 8: Performance of the unseen sea state model in terms of the average extreme von Mises stress over 5 seeds

Sea state number	True value [MPa]	Predicted value [MPa]	Difference [MPa]
#1	394	383	12
#2	396	385	10
#3	385	378	6
#4	394	387	7
#5	398	390	8
#6	404	402	1
#7	409	406	4
#8	408	406	2
#9	414	408	6
#10	413	407	7

Even though there is significant under estimation for some sea states, the predicted and true time series reveal that the model is very proficient at predicting the location of the peaks. The time series around the extreme peak for one of the seeds is shown in figure 4.

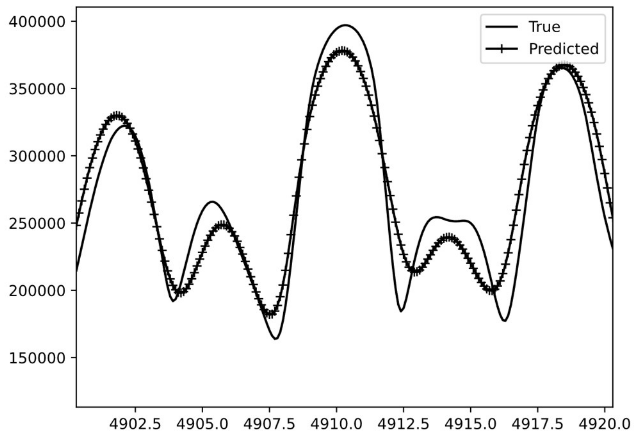

Figure 4: Extreme peak in sea state #1, first seed, plotted against predicted values

If the time of occurrence of the extreme response is known, it is straight forward to apply the physics-based model for a limited duration around the peak value. Typically, such a simulation should start well in advance of the peak to allow the system to settle in a stationary condition, and can end a few seconds after the occurrence of the peak. A duration of approximately 60 seconds is adequate for most models. Figure 5 shows the sensitivity of the case-study model to the lead time when re-running peaks. All the simulations converge to the same value, indicating that a 15 second lead time is sufficient. Many of the largest predicted peaks may be identified for such a rerun to improve predictions. The results including a three-peak re-run strategy are shown in table 9, where the process is approximated by extracting the extreme value from the true time series for a duration +/- 10 seconds from the predicted peak locations.

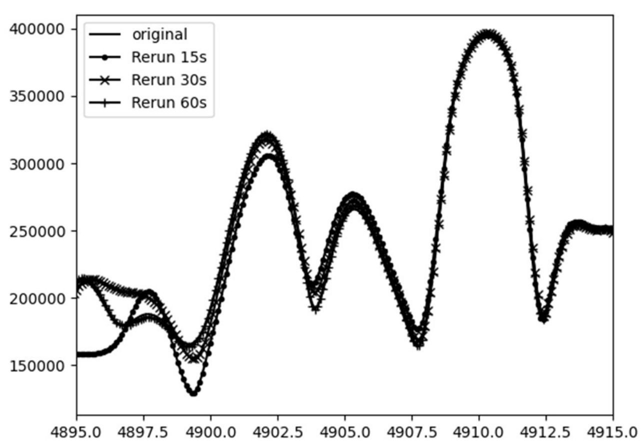

Figure 5: Results obtained by short simulations around the peak of the original simulation using the physics-based model

Table 9: Performance of the unseen sea state model in terms of the average extreme von Mises stress over 5 seeds, allowing re-run of time series for the 3 highest predicted peak locations

Sea state number	True value [MPa]	Predicted value [MPa]	Difference [MPa]
#1	394	394	0.2
#2	396	395	0.7
#3	385	385	0.0
#4	394	394	0.0
#5	398	397	1.1
#6	404	402	1.5
#7	409	409	0.0
#8	408	407	1.5
#9	414	414	0.0
#10	413	411	2.6

The tables show the average extreme value, the distribution of the values for the five seeds are exemplified by sea state #1, which is plotted in figure 6 for the unseen sea state model.

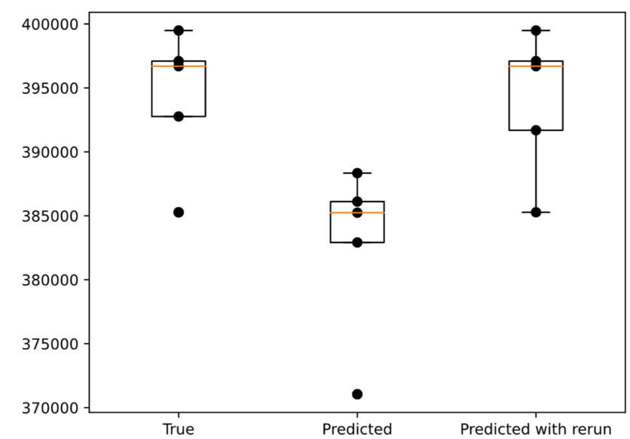

Figure 6: Results for the five seeds in sea state #1 for the true value, the value predicted purely by the unseen sea state model, and the value found from rerun of three highest peaks predicted by the model

The single sea state model can perform well with significantly reduced training time. While the unseen sea state model is trained on 300 samples of 15 minutes each, 4500 minutes in total, the single sea state models in this study are trained on time series between 15 minutes and 60 minutes.

Table 10: Performance of the 15minute single-sea state model in terms of the average extreme von Mises stress over 5 seeds

Sea state number	True value [MPa]	Predicted value [MPa]	Difference [MPa]
#1	394	386	8
#2	396	361	35
#3	385	361	23
#4	394	385	9
#5	398	396	2
#6	404	372	31
#7	409	399	10
#8	408	405	4
#9	414	399	15
#10	413	415	-2

The performance of the 15-minute single-sea state model is worse than that of the unseen sea state model. However, similar improvements are seen if the peak-values are extracted from short windows around predicted occurrence on the true time series.

Table 11: Performance of the 15-minute single-sea state model in terms of the average extreme von Mises stress over five seeds, allowing re-run of time series for the three highest predicted peak locations

Sea state number	True value [MPa]	Predicted value [MPa]	Difference [MPa]
#1	394	394	0.2
#2	396	393	2.2
#3	385	385	0.0
#4	394	393	1.1
#5	398	398	0.3
#6	404	397	6.7
#7	409	409	0.0
#8	408	406	2.5
#9	414	413	1.2
#10	413	413	0.4

Results are also presented for a single sea state model trained on 60-minute simulation data. By increasing the training data, the difference between predicted and true values are reduced, which is the expected effect, but the results are still worse in comparison with the unseen sea state model.

Table 12: Performance of the 60minute single-sea state model in terms of the average extreme von Mises stress over 5 seeds

Sea state number	True value [MPa]	Predicted value [MPa]	Difference [MPa]
#1	394	395	-1
#2	396	378	17
#3	385	365	19
#4	394	390	3
#5	398	393	5
#6	404	396	8
#7	409	408	1
#8	408	405	4
#9	414	416	-3
#10	413	408	5

Table 13: Performance of the 60minute single-sea state model in terms of the average extreme von Mises stress over 5 seeds, allowing re-run of time series for the 3 highest predicted peak locations

Sea state number	True value [MPa]	Predicted value [MPa]	Difference [MPa]
#1	394	394	0.0
#2	396	395	0.7
#3	385	385	0.0
#4	394	393	1.1
#5	398	398	0.3
#6	404	403	0.7
#7	409	409	0.0
#8	408	408	0.0
#9	414	414	0.0
#10	413	413	0.4

SUMMARY

The process of building a machine learning model that can make efficient predictions during operations is presented in this paper, including the detailed steps and configuration for a realistic and challenging case study.

Two main approaches are used for training the machine learning model, a single-sea state approach, where the scope of the model is only to predict new time series generated from the same sea state on which it is trained, and a unseen sea state model, where the scope of the model is to predict time series for completely new sea states.

Even though the single-sea state model requires much less training time, the unseen sea state model is much more efficient for operations, because all the training can be performed before project execution. The unseen sea state model also performs better than the single sea state model, even when the single sea state models training duration is increased up to 60 minutes.

The unseen sea state model should, however, be used with caution. All relevant sea states must be sufficiently represented in the training data set. If the model is exposed to sea states that have unseen features, such as a different directionality compared to the training data set, the predictions cannot be trusted.

For both these methods, significant improvement is seen if the physics-based model is used to generate target responses around the occurrence of extreme peaks. If three peaks from each sea state are selected for rerun, the error is found to be negligible for all the 10 sea states included in the test set, independent of the model applied. This approach increases the simulation time somewhat, but it is still reduced by a factor of 10 compared to running a complete set of 5 three-hour simulations (assuming a single sea state model trained on a 60-minute time series). The data generated from peak-rerun can also be used as a validation set, to safeguard against wrongful application on unseen and unexpected wave conditions.

ACKNOWLEDGEMENTS

The authors would like to thank the Research Council of Norway for their economic support of this project, and Subsea7 for providing realistic vessel and pipeline models, as well as recorded forecast data for the case study presented herein.

REFERENCES

Chaves, V., Sagrilo, L. V. S., da Silva, V. R. M., & Vignoles, M. A. (2015). Artificial Neural Networks Applied to Flexible Pipes Fatigue Calculations. *Volume 5B: Pipeline and Riser Technology*, V05BT04A022. https://doi.org/10.1115/OMAE2015-41650

Cho, K., van Merrienboer, B., Bahdanau, D., & Bengio, Y. (2014). *On the Properties of Neural Machine Translation: Encoder-Decoder Approaches* (arXiv:1409.1259). arXiv. https://doi.org/10.48550/arXiv.1409.1259

Christiansen, N. H., Torbergsen Voie, P. E., Høgsberg, J., & Sødahl, N. (2013). Efficient Mooring Line Fatigue Analysis Using a Hybrid Method Time Domain Simulation Scheme. *Volume 1: Offshore Technology*, V001T01A035. https://doi.org/10.1115/OMAE2013-10682

da Silva, V. R. M., & de Araujo, B. S. (2020, December 18). *Convolutional Neural Networks Applied to Flexible Pipes for Fatigue Calculations*. ASME 2020 39th International Conference on Ocean, Offshore and Arctic Engineering. https://doi.org/10.1115/OMAE2020-18212

da Silva, V. R. M., Sagrilo, L. V. S., & de Araujo, B. S. (2022). Applied Deep Learning for Slender Marine Structure Dynamic Analysis. *Journal of Offshore Mechanics and Arctic Engineering*, *144*(2), 021701. https://doi.org/10.1115/1.4052243

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. The MIT Press.

Guarize, R., Matos, N. A. F., Sagrilo, L. V. S., & Lima, E. C. P. (2007). Neural networks in the dynamic response analysis of slender marine structures. *Applied Ocean Research*, *29*(4), 191–198. https://doi.org/10.1016/j.apor.2008.01.002

Hasselmann, K., Barnett, T. P., Bouws, E., Carlson, H., Cartwright, D. E., Enke, K., Ewing, J., Gienapp, H., Hasselmann, D., Kruseman, P., Meerburg, A., Muller, P., Olbers, D., Richter, K., Sell, W., & Walden, H. (1973). Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP). *Deut. Hydrogr. Z.*, *8*, 1–95.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*, 386–408. https://doi.org/10.1037/h0042519

Srikonda, R., Haakonsen, R., Russo, M., & Periyasamy, P. (2018). Real-Time Wellhead Bending Moment Measurement Using Motion Reference Unit (MRU) Sensors and Machine Learning. *Volume 3: Structures, Safety, and Reliability*, V003T02A095. https://doi.org/10.1115/OMAE2018-78301

Wu, M. (2021). *Uncertainty of Machine Learning-Based Methods for Wave Forecast and its Effect on Installation of Offshore Wind Turbines* [Doctoral thesis, NTNU]. https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2788401