International Journal of Naval Architecture and Ocean Engineering 14 (2022) 100421

Contents lists available at ScienceDirect

International Journal of Naval Architecture and Ocean Engineering

journal homepage: http://www.journals.elsevier.com/international-journal-of-naval-architecture-and-ocean-engineering/

# Standardization in building an ANN-based mooring line top tension prediction system

Daesoo Lee [a], Seungjae Lee [b, *], Jaeyong Lee [c]

[a] Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway
[b] Division of Naval Architecture and Ocean Systems Engineering, Korea Maritime and Ocean University, Busan, Republic of Korea
[c] Department of Naval Architecture and Ocean Engineering, Dong-eui University, Busan, Republic of Korea

ARTICLE INFO

*Article history:*
Received 23 April 2021
Received in revised form
15 October 2021
Accepted 16 November 2021
Available online 1 December 2021

*Keywords:*
Mooring line top tension prediction
Artificial Neural Network (ANN)
Wave state selection

ABSTRACT

ANN-based mooring line top tension prediction systems are trained on ship motion and mooring line top tension time histories from multiple wave states with a certain simulation length. In the previous studies, selection of the wave states and the simulation length differs between the studies and they are not standardized. Also, a plain neural network is mostly used. In this paper, tension prediction performances with respect to a distribution shape of the wave states, a number of the wave states, and the simulation length are first studied. Then, the prediction performances with respect to Batch Normalization (BN) and Learning Rate Decay (LRD) are studied, in which BN and LRD are very common components in modern neural network models. Lastly, a guideline for selecting the wave states and the simulation length is proposed, and BN and LRD are proven to be advisable to use to improve the prediction performance.


## 1. Introduction

Typically, dynamic simulation for a moored floating structure and fatigue analysis on mooring lines takes a long computational time because it requires a huge amount of numerical calculation to represent system's dynamic behavior under various sea states. To speed up the dynamic simulation and fatigue analysis, several studies have been conducted. Simões et al. (2002) proposed a neural network architecture to predict the top tension on mooring lines and a hawser in a system where a turret-FPSO is connected with a shuttle tanker by the hawser to perform a faster simulation by replacing a huge amount of numerical calculation to represent an actual system's dynamic behavior. Guarize et al. (2007) proposed a hybrid Artificial Neural Network (ANN)-Finite Element Method (FEM) approach where the ANN is trained with short motion and top tension time histories generated by a FEM-based time-domain simulation and the trained ANN calculates the rest of the top tension history with an input of a prescribed motion time history for a faster simulation by partially replacing the FEM with a neural network model. Christiansen et al. (2013) proposed an ANN-based mooring line fatigue analysis approach for the fast fatigue analysis by predicting tension time histories on unseen environmental conditions and a procedure for the fatigue analysis based on the proposed approach. Sidarta et al. (2017) proposed an ANN-based mooring line top tension prediction system that receives previous motion time histories only as an input unlike Christiansen et al. (2013) where the ANN receives previous motion time histories as well as previous predicted top tension histories as the input. Yetkin et al. (2017) and Yetkin and Kim (2019) proposed the mooring top tension prediction systems based on a NARX (Nonlinear AutoRegressive eXogenous) neural network for a faster simulation by predicting tension time histories on unseen environmental conditions. The NARX is often used for time series prediction (Diaconescu, 2008; Xie et al., 2009). Despite all the previous developments, there are some common limitations in the previous studies: 'non-standardized selection for the wave states and the simulation length' a plain neural network architecture'; To train a neural network model for the tension prediction, a training dataset is required, in which the training dataset usually consist of ship motion and top tension time histories. The training dataset is commonly obtained by running dynamic simulations on multiple wave states with a certain simulation length. However, selection of the wave states and the simulation length differs between the previous studies and they are not standardized. Also, a plain neural network architecture is mostly used in the previous studies. In this paper, the tension prediction performances of a neural network

* Corresponding author.
*E-mail address:* slee@kmou.ac.kr (S. Lee).
Peer review under responsibility of The Society of Naval Architects of Korea.

https://doi.org/10.1016/j.ijnaoe.2021.11.004

model with respect to a distribution shape of the wave states, a number of the wave states, and the simulation length are first studied by running the neural network training. Then, the prediction performances of the trained neural network models are explained by proposed approximate measures that are based on a statistical variance and a statistical theorem, central limit theorem (Wheelan, 2013). The proposed approximate measures can be utilized to approximate the prediction performance even before training the neural network models. Also, effects of Batch Normalization (BN) proposed by Ioffe and Szegedy (2015) and Learning Rate Decay (LRD) are studied to improve the prediction performance further. The BN normalizes outputs from a hidden layer, which prevents a gradient vanishing problem and allows faster learning. The LRD decreases the learning rate over training iterations, which allows the training process to converge and avoid oscillation in the gradient descent process. The BN and the LRD are very common components in modern neural network models and known to improve neural network's performances fairly well (Ioffe and Szegedy, 2015; You et al., 2019). Lastly, a guideline for selecting the wave states and the simulation length is proposed along with the approximate measures for the prediction performances with respect to a number of the wave states and the simulation length, respectively. Also, it is proven that the use of the BN and LRD in the ANN-based mooring line top tension prediction system improves the tension prediction performance.

## 2. Motivation

The mooring line top tension prediction performances of the neural network model are studied with respect to the following four perspectives: 1) a distribution shape of the wave states, 2) a number of the wave states, 3) the simulation length, 4) the use of the BN and the LRD.

### 2.1. With respect to the distribution shape of the wave states

It is a well-known fact that data diversity in a training dataset leads to robust performance of a neural network model. Then, in order to have the robust tension prediction performance of the neural network model across various wave states within a wave scatter diagram, it is crucial to create the training dataset from various wave states across the wave scatter diagram. Then, it is natural to think of selecting the wave states uniformly across a wave scatter diagram as in Fig. 1(a) where *Hs* and *Tz* denote a significant wave height and a zero-crossing period, respectively. However, if each wave state is represented by a representative vector which consists of some statistical properties of the ship motion and tension time histories from the corresponding wave state, then it is obvious that the difference between the representative vectors within the low-*Hs* wave states of 1m and 3m would be much lower than the high-*Hs* wave states of 11m and 13m. This concept is illustrated in Fig. 2. Therefore, it might be more beneficial to select the high-*Hs* wave states more densely than the low-*Hs* wave states to ensure the data diversity in the training dataset as in Fig. 1(b). The first idea of the uniform wave state selection is utilized in the previous study such as (Sidarta et al., 2017), and the second idea of the high-*Hs*-focused wave state selection is utilized in (Christiansen et al., 2013). In this paper, the study of the tension prediction performance with respect to the distribution shape of the wave states is inspired to find out which distribution shape of the wave states results in the better prediction performance.

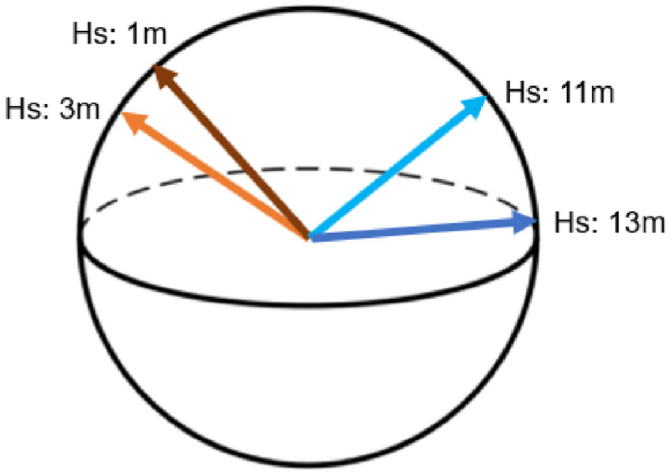
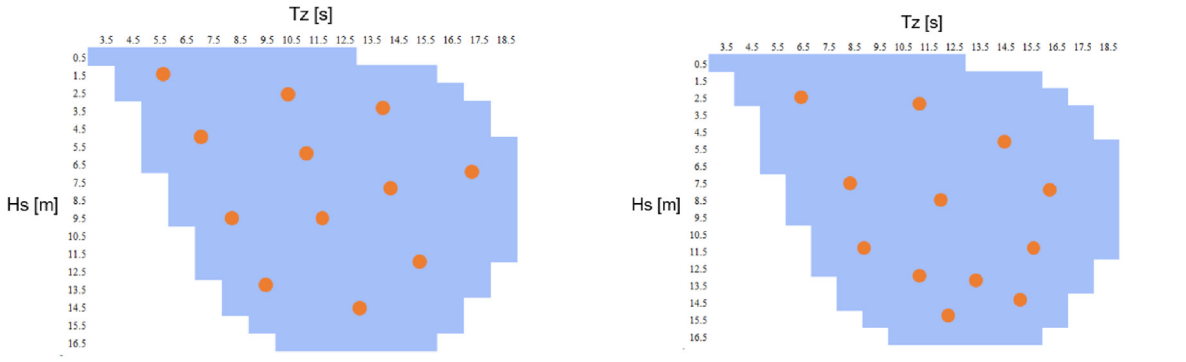



**Fig. 2.** Illustration of the representative vectors that consist of some statistical properties of the motion and tension time histories from the corresponding wave states.

### 2.2. With respect to a number of the wave states

A large amount of training data is crucial for performance of a neural network model in general. Therefore, a higher number of the




(a) Uniform distribution of the wave states (uniform wave state selection)

(b) High-*Hs*-focused distribution of the wave states (high-*Hs*-focused wave state selection)

**Fig. 1.** Two distribution shapes of the wave states across a wave scatter diagram.

wave states is very likely to result in the better prediction performance assuming that the simulation length for each wave state remains the same. Examples of a low number of the wave states and a high number of the wave states are shown in Fig. 3 where the black lines denote boundary lines of valid cells in a wave scatter diagram. However, the number of the wave states cannot be increased indefinitely due to the computational cost for the dynamic simulations. Since the prediction performance would converge as a number of the wave states increases, it is important to select a number of the wave states that is just large enough to maximize the prediction performance.

### 2.3. With respect to the simulation length

Aside from the distribution shape of the wave states and a number of the wave states, the simulation length needs to be determined. The simulation length denotes the dynamic simulation's length on each wave state to obtain the motion and tension time histories. The longer the simulation length is, the larger training dataset the neural network model is trained on, which leads to the better tension prediction performance. In a nutshell, the simulation length should be long enough to capture a global statistical distribution of the motion and tension time histories. For instance, 3h dynamic simulations are commonly used for mooring design analysis since it is considered to be statistically reliable enough to capture the maximum top tension (Kang et al., 2021). However, the simulation length cannot just be as long as possible since the dynamic simulations are computationally expensive. The tension prediction performance would converge as the simulation length increases, therefore, it is important to find the simulation length that is just long enough to maximize the tension prediction performance.

### 2.4. With respect to batch normalization and learning rate decay

The BN and LRD are very common components in modern neural network models and are known to improve neural network's performance fairly well. The BN is treated as a layer and used between a hidden layer and an activation layer. It normalizes outputs from a hidden layer, which prevents a gradient vanishing problem and allows the faster learning. The BN is illustrated in Fig. 4(b). It can be observed that many outputs without the BN fall into a zero-gradient zone, which causes the gradient vanishing problem while the BN prevents the outputs fall into the zero-gradient zone.

With the LRD, an initial learning rate gradually decreases over training iterations, which allows to converge and avoid oscillation in the gradient descent process. Also, You et al. (2019) found that the LRD helps neural networks learn complex patterns better. One of the most common LRD is an exponential LRD. Its equation is shown in Eq. (1) where $\eta$ denotes the learning rate, $i$ denotes a step (training iteration), and $D_{rate}$ and $D_{step}$ are parameters where $D_{rate} \in (0, 1)$ and $D_{step} > 1$. If $\eta_0$ (initial learning rate), $D_{rate}$, and $D_{step}$ are set to 0.001, 0.96, and 1,000, respectively, then the decayed learning rate looks as Fig. 5.

$$\eta_{i+1} = \eta_i D_{rate}^{i/D_{step}} \tag{1}$$

In the previous studies, neither the BN nor the LRD is utilized to the best of our knowledge. Given the fact that these two provide neural network models with robust performance improvement in various domains, it is likely that the BN and LRD would provide improvement in the tension prediction performance. Additionally, it should be noted that the BN and LRD can easily be implemented in open-source deep learning libraries such as Tensorflow and PyTorch.

## 3. Target ship and mooring configuration

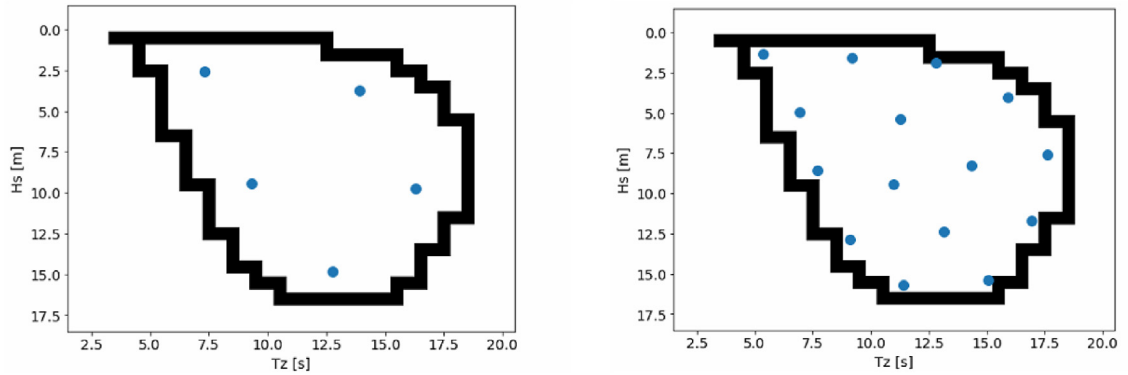
A target ship used in the dynamic simulations is a Floating Production Unit (FPU) and its principal dimension is shown in Table 1. The LBP, GM, XCG, YCG, VCG, $k_{xx}$, $k_{yy}$, and $k_{zz}$ denote length between perpendiculars, metacentric height, surge-directional center of gravity, sway-directional center of gravity, heave-directional center of gravity, and radius of gyrations along the surge-directional axis, the sway-directional axis, and the heave-directional axis, respectively.

A mooring configuration of the target ship is shown in Fig. 6 where the OBA (Outer Bundle Angle) and IBA (Inner Bundle Angle) are set to 50° and 4°, respectively. The small numbers at the ends of the mooring lines indicate the mooring line index numbers which start from an upper left side and count counterclockwise. Properties of the mooring lines are shown in Table 2. An attack angle of environmental direction is 20° to the target ship. This paper is a study using only waves, and additional environmental loads such as wind and current will be considered in the future.

## 4. Design of the neural network model and metric for its tension prediction

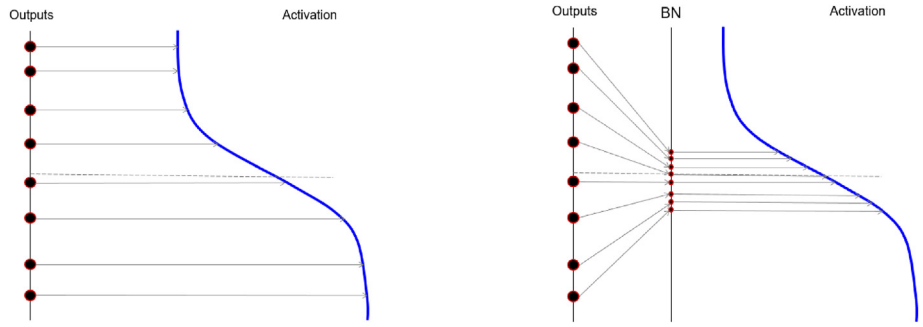
### 4.1. Design of the neural network model

Fig. 7(a) presents the plain Neural Network (NN) architecture that is used in most of the previous studies. In this paper, the NN




(a) Low number of the wave states (5 wave states) (b) High number of the wave states (15 wave states)

**Fig. 3.** Examples of low and high numbers of the wave states.




(a) Without the BN (b) With the BN

**Fig. 4.** Illustration of the batch normalization.

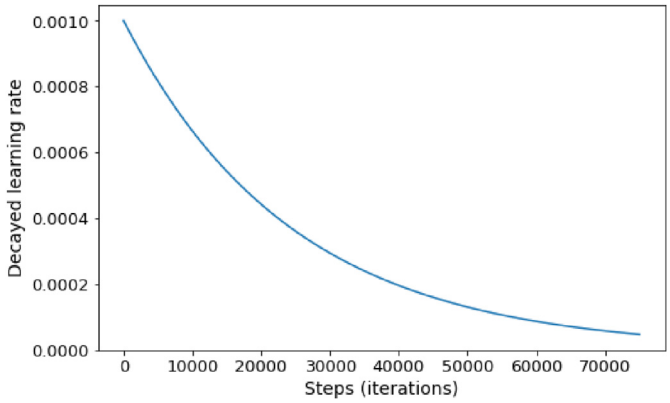



**Fig. 5.** Example of the decayed learning rate.

architecture with the BN as shown in Fig. 7(b) is used for all neural network training unless specified differently. Configurations of the input $\boldsymbol{X}$ and the output $\boldsymbol{Y}$ used in this paper are illustrated in Fig. 8 and presented in Eqs. (2) and (3) in detail.

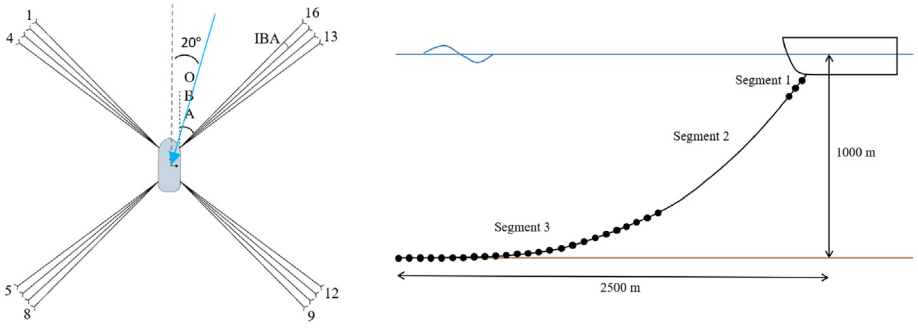
$$\boldsymbol{X}(t,L,\Delta t)=\begin{Bmatrix} x(t-L),x(t-(L-\Delta t)),x(t-(L-2\Delta t)),\cdots,x(t),\\ y(t-L),y(t-(L-\Delta t)),y(t-(L-2\Delta t)),\cdots,y(t),\\ z(t-L),z(t-(L-\Delta t)),z(t-(L-2\Delta t)),\cdots,z(t),\\ R_x(t-L),R_x(t-(L-\Delta t)),R_x(t-(L-2\Delta t)),\cdots,R_x(t),\\ R_y(t-L),R_y(t-(L-\Delta t)),R_y(t-(L-2\Delta t)),\cdots,R_y(t),\\ R_z(t-L),R_z(t-(L-\Delta t)),R_z(t-(L-2\Delta t)),\cdots,R_z(t), \end{Bmatrix} \tag{2}$$

$$\boldsymbol{Y}(t)=\{T_1(t),T_2(t),\cdots,T_M(t)\} \tag{3}$$

where $x$, $y$, $z$, $R_x$, $R_y$, and $R_z$ denote the surge, sway, heave, roll, pitch, and yaw ship motions, respectively. $\Delta t$ and $L$ denote a timestep interval and memory length of $\boldsymbol{X}$, respectively. $T$ denotes the tension with its subscript $M$ denoting a number of mooring lines. In this paper, $L$ and $\Delta t$ are set to 10s and 0.2s based on the previous studies (Guarize et al., 2007; Simões et al., 2002; Yetkin et al., 2017). The training dataset consists of the input-output pairs and it is obtained from the dynamic simulations on the selected wave states within a wave scatter diagram. For the dynamic simulation software, OrcaFlex developed by Orcina is used in this paper. For the wave scatter diagram, a wave scatter diagram from DNV-GL (2018) is used. The used wave scatter diagram is shown in Fig. 9.

**Table 1**
Principal dimension of the target ship.

| LBP | Breath | Draft | Volume | GM | XCG | YCG | VCG | $k_{xx}$ | $k_{yy}$ | $k_{zz}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 244m | 50m | 18.6m | 169,614 $m^3$ | 1.99m | 117.7m | 0m | 19.5m | 17.4m | 59.3m | 60m |

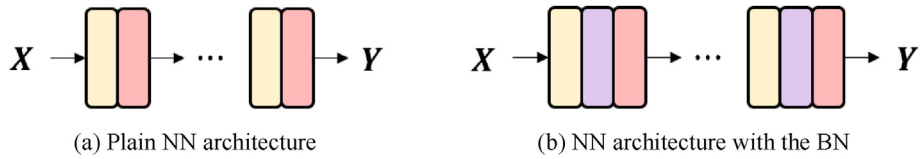



(a) Top (b) Profile

**Fig. 6.** Mooring configuration of the target ship.

**Table 2**
Properties of the mooring lines.

| | Type | Length [m] | Diameter [mm] | Weight in air [kg/m] | Weight in water [kg/m] | MBL [kN] |
|---|---|---|---|---|---|---|
| Segment 1 | R4 studless chain | 62 | 130 | 336 | 292 | 15,559 |
| Segment 2 | Spiral strand wire | 900 | 120 | 57 | 50 | 9120 |
| Segment 3 | R4 studless chain | 1800 | 130 | 336 | 292 | 15,559 |




(a) Plain NN architecture (b) NN architecture with the BN

**Fig. 7.** Architecture of the neural network model (Yellow: a hidden layer, Red: an activation layer, Purple: a batch normalization layer).

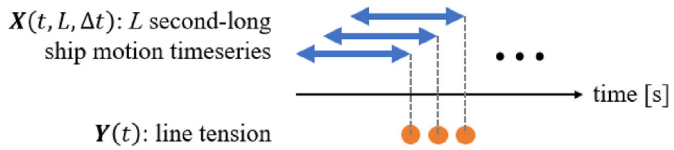



**Fig. 8.** Arrangement of the input-output pairs to be used to train the NN model.

OrcaFlex performs global static and dynamic analysis of a wide range of offshore systems, typically including boundary conditions such as vessels, buoys, etc., as well as finite element (FE) modeling of line structures. The procedure of the simulation consists of two steps: 1) static analysis, 2) dynamic analysis. In the static analysis, OrcaFlex is to find positions and orientations for each element in the model such that all forces and moments are in equilibrium. The equilibrium is used as an initial point for the dynamic analysis. The time-domain dynamic analysis is fully nonlinear, and mass, damping, stiffness, loading, etc. are evaluated at each time step, considering the instantaneous, time-varying geometry. In the time domain dynamic analysis, the following equation of motion is solved:

$$M(p,a)+C(p,v)+K(p)=F(p,v,t) \tag{4}$$

where $M(p,a)$ is the system inertial load, $C(p,v)$ is the system damping load, $K(p)$ is the system stiffness load, $F(p,v,t)$ is the external load, $p, v, a$, and $t$ are the position, velocity, acceleration vectors, and simulation time, respectively. As for the line modeling, OrcaFlex uses a Finite Element (FE) model for a line as shown in Fig. 10. The line is divided into a series of line segments which are then modeled by a straight massless model with a node at each end. The model segments only model the axial and torsional properties of the line. The other properties (mass, weight, buoyancy, etc.) are all lumped to the nodes, as indicated by the arrows in the figure.

The hyperparameters for the neural network model used in this paper are shown in Table 3. Throughout the paper, this hyperparameter setting is used unless specified differently.

The neural network model was implemented using the deep learning open-source library, Tensorflow, and its training was conducted on a single GPU (Geforce GTX 1070).

### 4.2. Metric for the tension prediction

After training the neural network model, the trained model needs to be evaluated on a test dataset. The test dataset is obtained

| Tz | 3.5 | 4.5 | 5.5 | 6.5 | 7.5 | 8.5 | 9.5 | 10.5 | 11.5 | 12.5 | 13.5 | 14.5 | 15.5 | 16.5 | 17.5 | 18.5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Hs | | | | | | | | | | | | | | | | |
| 0.5 | 1.3 | 134 | 866 | 1186 | 634 | 186 | 36.9 | 5.6 | 0.7 | 0.1 | | | | | | |
| 1.5 | | 29.3 | 986 | 4976 | 7738 | 5570 | 2376 | 704 | 161 | 30.5 | 5.1 | 0.8 | 0.1 | | | |
| 2.5 | | 2.2 | 198 | 2159 | 6230 | 7450 | 4860 | 2066 | 645 | 160 | 33.7 | 6.3 | 1.1 | 0.2 | | |
| 3.5 | | | 34.9 | 696 | 3227 | 5675 | 5099 | 2838 | 1114 | 338 | 84.3 | 18.2 | 3.5 | 0.6 | 0.1 | |
| 4.5 | | | 6 | 196 | 1354 | 3289 | 3858 | 2686 | 1275 | 455 | 131 | 31.9 | 6.9 | 1.3 | 0.2 | |
| 5.5 | | | 1 | 51 | 498 | 1603 | 2373 | 2008 | 1126 | 464 | 151 | 41 | 9.7 | 2.1 | 0.4 | 0.1 |
| 6.5 | | | 0.2 | 12.6 | 167 | 90 | 1258 | 1269 | 826 | 387 | 141 | 42.2 | 10.9 | 2.5 | 0.5 | 0.1 |
| 7.5 | | | | 3 | 52.1 | 270 | 594 | 703 | 525 | 277 | 112 | 36.7 | 10.2 | 2.5 | 0.6 | 0.1 |
| 8.5 | | | | 0.7 | 15.4 | 97.9 | 256 | 351 | 297 | 175 | 77.6 | 27.7 | 8.4 | 2.2 | 0.5 | 0.1 |
| 9.5 | | | | 0.2 | 4.3 | 33.2 | 102 | 160 | 152 | 99.2 | 48.3 | 18.7 | 6.1 | 1.7 | 0.4 | 0.1 |
| 10.5 | | | | | 1.2 | 10.7 | 37.9 | 67.5 | 71.7 | 51.5 | 27.3 | 11.4 | 4 | 1.2 | 0.3 | 0.1 |
| 11.5 | | | | | 0.3 | 3.3 | 13.3 | 26.6 | 31.4 | 24.7 | 14.2 | 6.4 | 2.4 | 0.7 | 0.2 | 0.1 |
| 12.5 | | | | | 0.1 | 1 | 4.4 | 9.9 | 12.8 | 11 | 6.8 | 3.3 | 1.3 | 0.4 | 0.1 | |
| 13.5 | | | | | | 0.3 | 1.4 | 3.5 | 5 | 4.6 | 3.1 | 1.6 | 0.7 | 0.2 | 0.1 | |
| 14.5 | | | | | | 0.1 | 0.4 | 1.2 | 1.8 | 1.8 | 1.3 | 0.7 | 0.3 | 0.1 | | |
| 15.5 | | | | | | | 0.1 | 0.4 | 0.6 | 0.7 | 0.5 | 0.3 | 0.1 | 0.1 | | |
| 16.5 | | | | | | | | 0.1 | 0.2 | 0.2 | 0.2 | 0.1 | 0.1 | | | |

**Fig. 9.** Wave scatter diagram used in this paper.

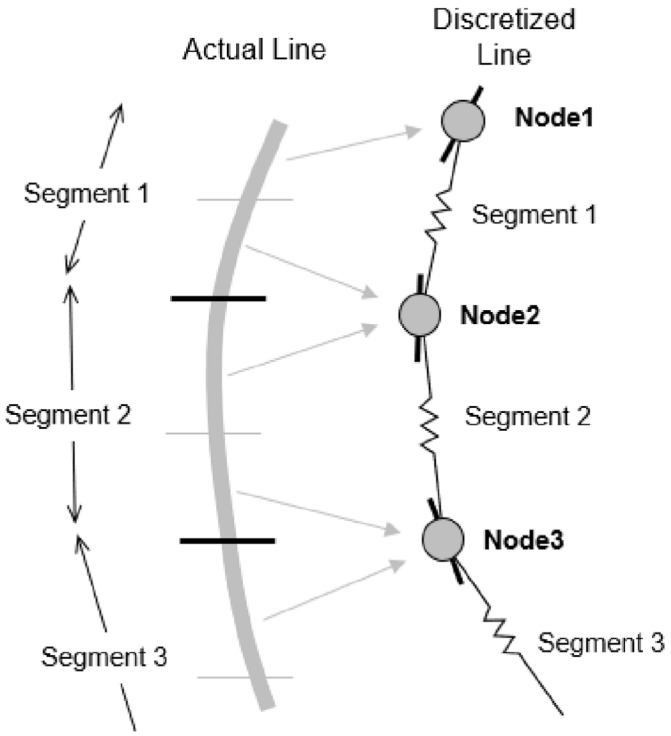



**Fig. 10.** Line modelling in OrcaFlex.

**Table 3**
Hyperparameters for the neural network model.

| Number of hidden layers | Hidden layer size | Use of the BN layer | Optimizer, Initial learning rate | Batch size | Epochs | Learning Rate Decay | |
|---|---|---|---|---|---|---|---|
| | | | | | | $D_{rate}$ | $D_{step}$ |
| 3 | 128 | Yes | Adam, 0.001 | $2^{14}$ | 15,000 | 0.96 | 1000 |

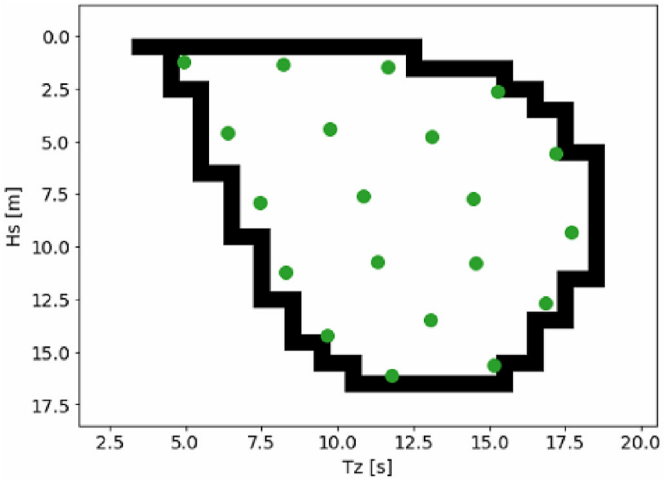



**Fig. 11.** Test wave states used to obtain the test dataset (20 wave states).

from the dynamic simulations on various test wave states across the wave scatter diagram with a different random seed from that of the training dataset. The test wave states used to obtain the test dataset are shown in Fig. 11.

Then, given that the output $\boldsymbol{Y}(t)$ consists of the top tensions on the 16 mooring lines at $t$ and there are 20 wave states for the test, the metrics for the tension prediction can be established as Eq. (5) and Eq. (6), where $M$ denotes a number of mooring lines, $\widehat{\boldsymbol{Y}}$ denotes predicted tension, $S$ denotes a number of the wave states, and $t$ here denotes a timestep. Two metrics are defined to compute the metric on two different cases: 1) metric on a single wave state, 2) metric on multiple wave states. $\mathbb{M}_s$ suits the former purpose and $\mathbb{M}_S$ suits the latter purpose. The higher metrics indicate the better prediction performance.

$$\mathbb{M}_s = -\frac{1}{M}\sum_{m}^{M}\frac{\sum_t^T|\boldsymbol{Y}_m(t)-\widehat{\boldsymbol{Y}}_m(t)|}{T} \tag{5}$$

$$\mathbb{M}_S = -\frac{1}{S}\frac{1}{M}\sum_{s}^{S}\sum_{m}^{M}\frac{\sum_t^T|\boldsymbol{Y}_{s,m}-\widehat{\boldsymbol{Y}}_{s,m}|}{T} \tag{6}$$

Depending on a random seed for the dynamic simulation, the resulting motion and tension time histories differ. Hence, a choice of the random seed may affect the performance of a trained neural network model since the neural network model is trained on those motion and tension time histories, then eventually, the metrics may differ depending on the random seed. Yet, if the training dataset of the motion and tension time histories is statistically consistent with respect to the random seed, the trained neural network model would result in the consistent performance, which results in the consistent metrics. To verify that the results presented in this paper are consistent with respect to the random seed, the following figure is presented:

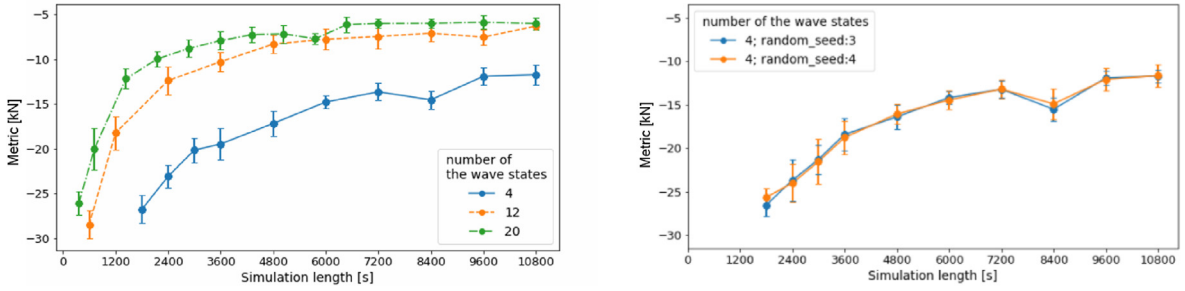
Fig. 12 shows the metrics of Eq. (6) with respect to a number of the wave states over the simulation length. The experiment was conducted focused on a number of the wave states of 4. It is apparent that the metrics remain statistically consistent with respect to the random seed for the dynamic simulation. A detailed result analysis regarding the prediction performance with respect to a number of the wave states is presented in Chapter 5.2.

## 5. Analysis

The mooring line top tension prediction performances of the neural network model are studied with respect to the following four perspectives: 1) a distribution shape of the wave states, 2) a number of the wave states, 3) the simulation length, 4) the use of the BN and the LRD.

### 5.1. With respect to distribution shape of the wave states

As stated in Chapter 2.1, there exist two distribution shapes of the wave states: a) uniform distribution of the wave states by the uniform wave state selection, b) high-*Hs*-focused distribution of the wave states by the high-*Hs*-focused wave state selection. In the previous studies, although one of the two distribution shapes is used, the wave state selections are conducted manually. In this

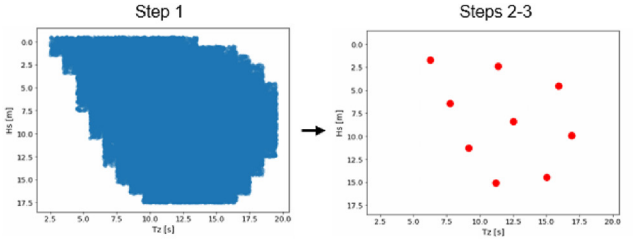



(a) random seed of 2 (b) random seed of 3 and 4

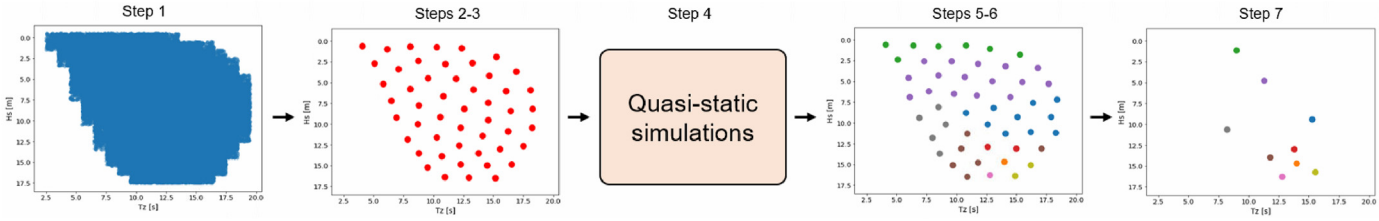
**Fig. 12.** Verification of consistent metrics with respect to a random seed.

**Table 4**
Standardized procedures for the wave state selection methods.

| (a) Uniform wave state selection (quasi-static) | (b) High-*Hs*-focused wave state selection (quasi-static) |
|---|---|
| Define a number of the wave states $S$ | Define a number of the wave states $S$ |
| 1. Randomly generate a large number of wave states (e.g. 50,000) within a wave scatter diagram.<br>2. Run a K-means algorithm and obtain $S$ clusters.<br>3. Obtain $S$ cluster centers which is the *selected wave states.* | 1. Randomly generate a large number (e.g. 50,000) of wave states within a wave scatter diagram.<br>2. Run the K-means algorithm and obtain multiple (e.g. 50) clusters.<br>3. Obtain the 50 cluster centers which is termed *pseudo-selected wave states.*<br>4. Run quasi-static simulations on the pseudo-selected wave states, and obtain motion and tension time histories for each pseudo-selected wave state.<br>5. Create a *representative vector* for each pseudo-selected wave state.<br>6. Run the K-means algorithm on the representative vectors, and obtain $S$ clusters.<br>7. Find a center point within each cluster on the wave scatter diagram domain, which results in the *selected wave states.* |




(a) Distribution shape of the wave states by the uniform wave state selection



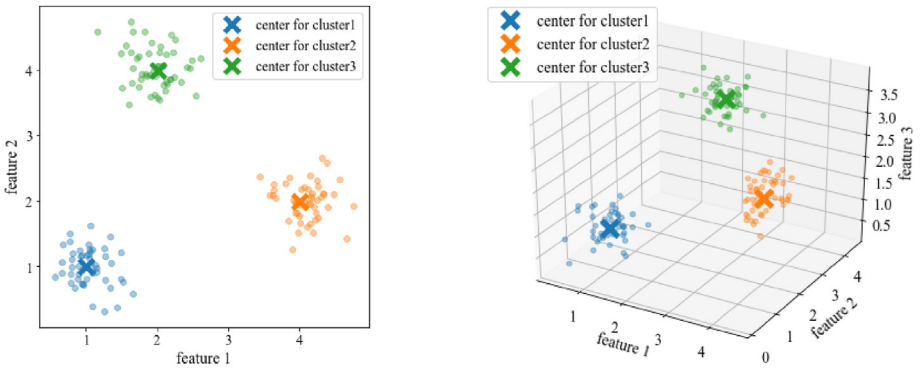

(b) Distribution shape of the wave states by the high-*Hs*-focused wave state selection

**Fig. 13.** Illustrations of the standardized procedures for the wave state selection methods.

paper, standardized procedures for both selection methods are proposed to replace the manual selection. The procedures of the standardized wave selection methods are presented in Table 4 and illustrated in Fig. 13.

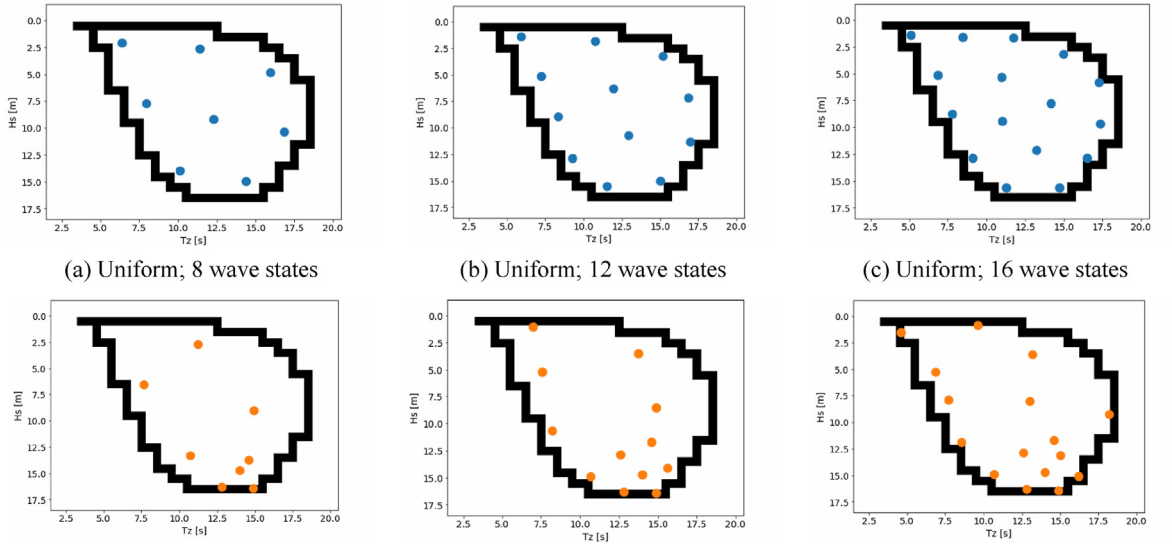
In Table 4, the K-means algorithm introduced by Macqueen (1967) is one of the most popular data clustering algorithms in machine learning. Examples of the data clustering by the K-means are shown in Fig. 14 where scattered data is clustered into three groups and a cluster center in each cluster is presented as a cross mark. It should be noted that input features are usually normalized before running the K-means algorithm to standardize distributions of each feature space. The representative vector used in Table 4 is under the same idea as the representative vector introduced in Chapter 2.1, and it is expressed as Eq. (7) to capture statistical properties of the motion and tension time histories. $X', Y', Z', R'_X, R'_Y$, and $R'_Z$ denote normalized six Degrees-of-Freedom (DOF) ship motion time histories, and $T'$ denotes a normalized mooring line top tension time history. $\mu$ and $\sigma$ denote mean and standard deviation, respectively. $\mathrm{FFT}_{\mathrm{CG}x}$ refers to a $x$-axial center of gravity of a Fast Fourier Transform (FFT) graph.

$$\text{Representative vector} = \left\{ \begin{array}{l} \mu(X'), \mu(Y'), \mu(Z'), \mu(R'_X), \mu(R'_Y), \mu(R'_Z), \\ \sigma(X'), \sigma(Y'), \sigma(Z'), \sigma(R'_X), \sigma(R'_Y), \sigma(R'_Z), \\ FFT_{\mathrm{CG}x}(X'), FFT_{\mathrm{CG}x}(Y'), FFT_{\mathrm{CG}x}(Z'), FFT_{\mathrm{CG}x}(R'_X), FFT_{\mathrm{CG}x}(R'_Y), FFT_{\mathrm{CG}x}(R'_Z), \\ \mu(T'_1), \mu(T'_2), \cdots, \mu(T'_M), \\ \sigma(T'_1), \sigma(T'_2), \cdots, \sigma(T'_M), \\ FFT_{\mathrm{CG}x}(T'_1), FFT_{\mathrm{CG}x}(T'_2), \cdots, FFT_{\mathrm{CG}x}(T'_M) \end{array} \right\} \quad (7)$$




(a) 2D (b) 3D

**Fig. 14.** Examples of data clustering by the K-means.




(a) Uniform; 8 wave states (b) Uniform; 12 wave states (c) Uniform; 16 wave states

(d) High-*Hs*-focused; 8 wave states (e) High-*Hs*-focused; 12 wave states (f) High-*Hs*-focused; 16 wave states

**Fig. 15.** Wave states selected by the two wave state selection methods.

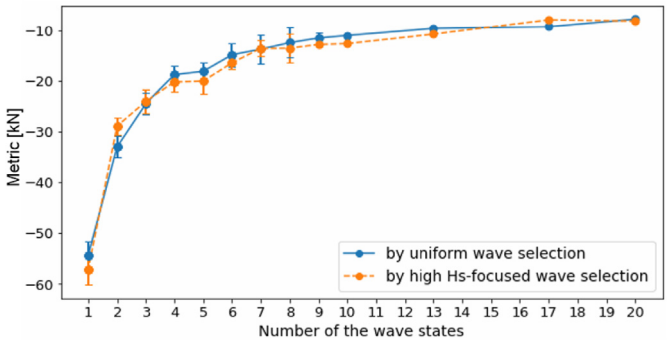



**Fig. 16.** Tension prediction performance with respect to the distribution shape of the wave states.

The proposed standardized procedures are used to select the wave states to run the dynamic simulations on, and the training dataset can be obtained from motion and tension time history results of the dynamic simulations. The wave states selected by the two wave state selection methods are presented in Fig. 15 where 'Uniform' refers to the uniform wave state selection and 'High-*Hs*-focused' refers to the high-*Hs*-focused wave state selection.

To study the prediction performance with respect to the distribution shape of the wave states, the neural network model is trained on the training dataset created from different numbers of

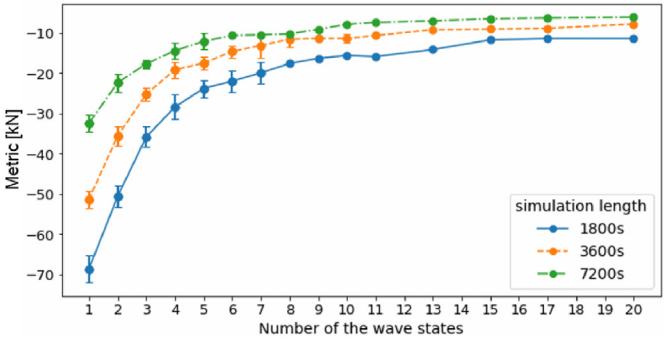



**Fig. 17.** Tension prediction performance with respect to a number of the wave states.

the wave states generated by each wave selection method. Here, the simulation length is set to 3600s. The trained neural network is tested on the test dataset created from the test wave states as Fig. 11 with the metric of Eq. (6), $\mathbb{M}_S$. The result is presented in Fig. 16. The result shows that the two wave selection methods result in almost the same tension prediction performances. Since the procedure of the uniform wave state selection is much simpler, it should be preferred over the other.

### 5.2. With respect to a number of the wave states

To study the prediction performance with respect to a number of the wave states, the neural network model is trained on the training dataset created from different numbers of the wave states by the uniform wave selection method. To verify the generic trend of the prediction performances with respect to a number of the wave states over different simulation lengths, three different simulation lengths are considered: 1800s, 3600s, 7200s. The trained neural network is tested on the test dataset created from the test wave states as Fig. 11 with the metric of Eq. (6), $\mathbb{M}_S$. The result is presented in Fig. 17. The result shows the general trend that the prediction performance increases significantly in the beginning and gradually converges as a number of the wave states increases.

#### 5.2.1. Approximate measure for the prediction performance w.r.t a number of the wave states

The tension prediction performance of the neural network model crucially depends on the amount of the training dataset and the data diversity within the training dataset. In other words, the prediction performance is positively affected by the following three components: 1) a number of the wave states, 2) dispersion of the wave states across the wave scatter diagram, 3) the simulation length. If the uniform wave state selection method is used, since a number of the wave states is proportional to the dispersion of the wave states, the three components can be reduced to two components: 1) a number of the wave states, 2) the simulation length. Then, from the inverse perspective, the tension prediction performance should be able to be explained by these two components. This is the hypothesis, $\mathscr{H}$, established in this paper to approximate the tension prediction performance as Eq. (8) where $\alpha$ denotes a scalar, $S$ denotes a number of the wave states, $L$ denotes the simulation length, and the prediction performance is represented by $\mathbb{M}_S$. Then, $f(S,L)$ represents a relative location of the prediction performance scaled by $\alpha$. In this paper, $f(S,L)$ is termed the approximate measure.

$$\mathscr{H}: \ \mathbb{M}_S \approx \alpha f(S,L) \tag{8}$$

In Chapter 5.2, the prediction performance is studied with respect to a number of the wave states. If the simulation length is set to a certain constant, $l$, then a trend of the prediction performance should be able to be approximated by $f(S,L=l)$ according to the hypothesis. The equation to compute $f(S,L=l)$ is shown in Eqs. 9–11 where $s$ denotes a certain wave state among the $S$ wave states, $\mathrm{FFT}_{\mathrm{CG}_x}(.)$ computes a $x$-axial center of gravity of a FFT graph, and $T_m(s,l)$ denotes a tension time history of an $m$-th mooring line with the simulation length of $l$. It should be noted that $f(S,L=l)$ is in a standard deviation form.

$$f(S,L=l) = \sqrt{\frac{\sum_s^S x_{s,l}^2}{S} - \mu_S^2} \tag{9}$$

$$\mu_S = \frac{1}{S}\sum_s^S x_{s,l} \tag{10}$$

$$x_{s,l} = \frac{1}{M}\sum_m^M \mathrm{FFT}_{\mathrm{CG}_x}(T_m(s,l)) \tag{11}$$

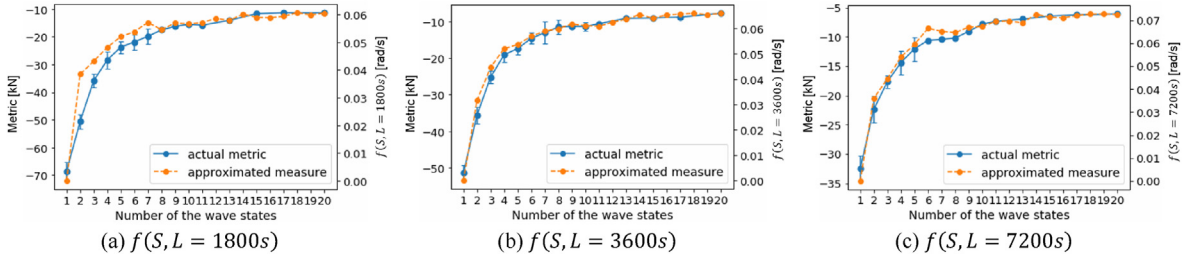
The approximate measure, $f(S,L=l)$, is computed and compared with the metric of the tension prediction performance in Fig. 18, where $S \in [1,20]$ and $L \in \{1800s,\ 3600s,\ 7200s\}$. In Fig. 18, the metric and the approximate measure are presented in the 1st and 2nd $y$-axes, respectively. It can be observed that the proposed approximate measures can approximate the trend of the prediction performances (metrics) fairly well, which verifies the hypothesis subject to a number of the wave states.

One useful application of this approximate measure is to find an ideal number of the wave states even before training the neural network model so that computational cost can be reduced. If the neural network model is trained on a low number of the wave states, its training would be finished quickly but its prediction performance would not reach its maximum. On the other hand, if the model is trained on an unnecessarily high number of the wave states, its training would take unnecessarily long for a mere improvement in the prediction. Hence, using an ideal number of the wave states is important. Searching it by training the neural network model for every number of the wave states would take too much computational time. This proposed approximate measure helps to find it very quickly, which saves computational time and resources.

The approximate measure, $f(S,L=l)$, has a form of a standard deviation. Therefore, its magnitude is proportional to a dispersion of $x_{s,l}$ from Eq. (11). The dispersion of $x_{s,l}$ with respect to the number of the wave states is presented in Fig. 19, which explains the result in Fig. 18.

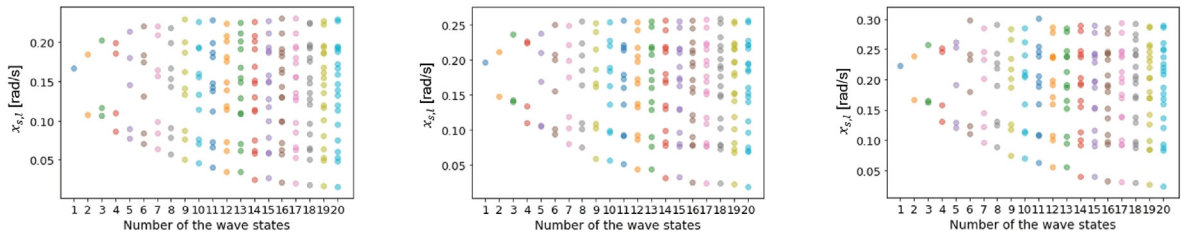
### 5.3. With respect to simulation length

To study the prediction performance with respect to the simulation length, the neural network model is trained on the training




(a) $f(S,L=1800s)$ (b) $f(S,L=3600s)$ (c) $f(S,L=7200s)$

**Fig. 18.** Approximate measure for the tension prediction performance with respect to a number of the wave states.




(a) $f(S, L = 1800s)$ (b) $f(S, L = 3600s)$ (c) $f(S, L = 7200s)$

**Fig. 19.** Dispersion of $x_{s,l}$ with respect to the number of the wave states.

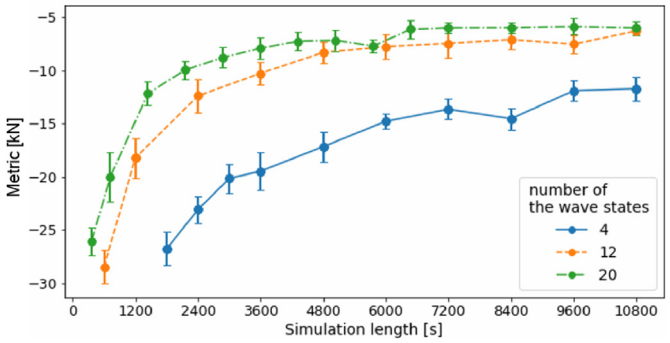



**Fig. 20.** Tension prediction performance with respect to the simulation length.

dataset created from a certain number of the wave states by the uniform wave selection method with various simulation lengths. To verify the generic trend of the prediction performance with respect to the simulation length over different numbers of the wave states, three different numbers of the wave states are considered: 4, 12, and 20. The trained neural network is tested on the test dataset created from the test wave states as Fig. 11 with the metric of Eq. (6), $\mathbb{M}_S$. The result in presented in Fig. 20. The result presents the general trend that the prediction performance increases significantly in the beginning and gradually converges as the simulation length increases.

### 5.3.1. Approximate measure for the prediction performance w.r.t simulation length

Based on the hypothesis, $\mathscr{H}$, the prediction performance should be explainable by the approximate measure, $f(S,L)$. In Chapter 5.3, the prediction performance is studied with respect to the simulation length. Therefore, if a number of the wave states is set to a certain constant, $S'$, then a trend of the prediction performance should be able to be approximated by $f(S=S',L)$ according to the hypothesis. The equation to compute $f(S=S',L)$ is shown in Eqs. (12)-(16). $\mathscr{M}$ consists of the six-DOF ship motions, $U(0,b)$ denotes a uniform distribution with a range of 0 to $b$, $N$ denotes sample size for $x^2_{s,L,t}$, and $\tilde{\mathscr{M}}$ denotes a subsequence of $\mathscr{M}$ with a time-range of $[t, t+L]$. In these equations, $L$ denotes the simulation length of $\tilde{\mathscr{M}}$ and $b$ is determined by the entire simulation length minus $L$. For the entire simulation length, 10800s is used. Intuitively, $\sum_{t\sim U(0,\,b)}^{N} x_{s,L,t}$ in Eqs. 13 and 14 can be explained as follows: A random subsequence of a motion history is sampled from an entire motion history and reduced to a single value, $x_{s,L,t}$, by $\mathrm{FFT}_{\mathrm{CG}_x}$, and this is repeated $N$ times. This process follows the central limit theorem in a sense that the weighted means of the $N$ sampled subsequences of the motion history form a sampling distribution and it would follow a normal distribution as the definition of the central limit theorem, in which the weighted mean is computed by $\mathrm{FFT}_{\mathrm{CG}_x}$. In other words, $x_{s,L,t}$ is randomly sampled $N$ times and it would form the normal distribution. Then, $\sigma_{\mathscr{M}}$ has the same concept as a standard error of a sampling distribution. Finally, $\sigma^2_{\mathscr{M}}$ is summed over the six-DOF motion histories and the selected wave states.

$$f(S=S',L) = -\sum_{s}^{S'} \sum_{\mathscr{M}\in\left\{X,Y,Z,R_X,R_Y,R_Z\right\}} \sigma^2_{\mathscr{M}} \tag{12}$$

$$\sigma^2_{\mathscr{M}} = \frac{\sum_{t\sim U(0,\,b)}^{N} x^2_{s,L,t}}{N} - \mu^2 \tag{13}$$

$$\mu = \frac{\sum_{t\sim U(0,\,b)}^{N} x_{s,L,t}}{N} \tag{14}$$

$$x_{s,L,t} = \mathrm{FFT}_{\mathrm{CG}_x}\left(\tilde{\mathscr{M}}_{s,L,t}\right) \tag{15}$$

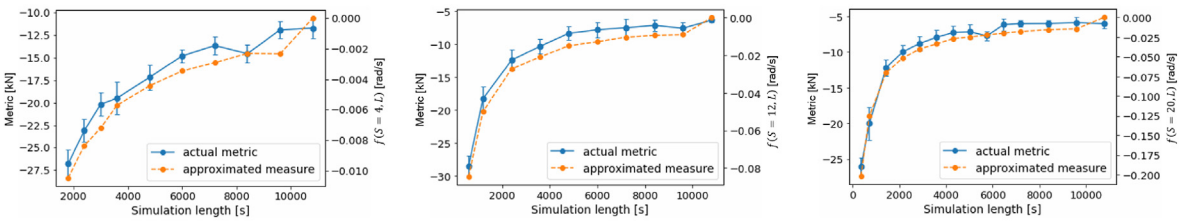
$$\tilde{\mathscr{M}}_{s,L,t} = \mathscr{M}_{t\in[t,t+L]} \tag{16}$$

The approximate measure, $f(S=S',L)$, is computed and compared with the metric of the tension prediction performance in Fig. 21, where $S\in\{4,12,20\}$. In Fig. 21, the metric and the approximate measure are presented in the 1st and 2nd $y$-axes, respectively. It can be observed that the proposed approximate measure, $f(S=S',L)$, can approximate the trend of the prediction performances (metrics) fairly well, which verifies the hypothesis subject to the simulation length. Similar to $f(S,L=l)$ introduced in Chapter 5.2.1, $f(S=S',L)$ can be used to find the optimal simulation length to reduce the computational cost.

The magnitude of $f(S=S',L)$ depends on $\sigma^2_{\mathscr{M}}$ which is the variance of the sampling distribution of $x_{s,L,t}$. In the central limit theorem, if the sample size is small, a variance of a corresponding sampling distribution is large, and vice and versa. In the same concept, if the simulation length, $L$, is small, the variance $\sigma^2_{\mathscr{M}}$ of the sampling distribution of $x_{s,L,t}$ should be large, and vice and versa. Examples of the sampling distributions of $x_{s,L,t}$ by histogram are presented in Fig. 22 where the wave state has $Hs$ of 16.8m and $Tz$ of 11.8s. It can be observed that the shorter simulation length results in the larger variance and the longer simulation length results in the smaller variance.

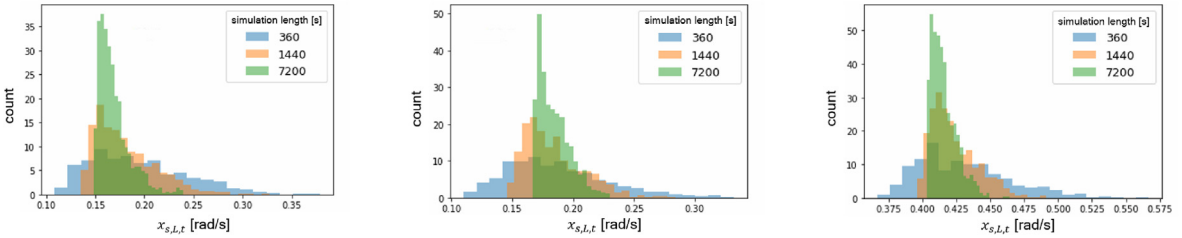
### 5.4. Trade-off between a number of wave states and simulation length

In the previous chapters, the prediction performances are



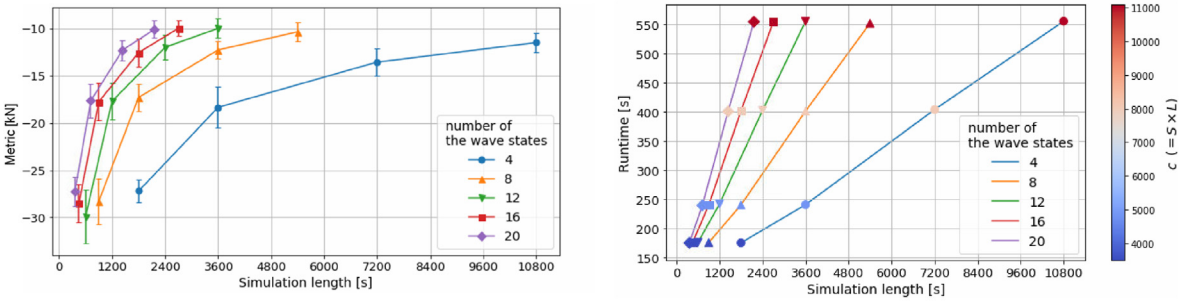

(a) $f(S=4,L)$ (b) $f(S=12,L)$ (c) $f(S=20,L)$

**Fig. 21.** Approximate measure for the prediction performance with respect to the simulation length.




(a) Surge; $\mathcal{M}=X$ (e) Sway; $\mathcal{M}=Y$ (f) Heave; $\mathcal{M}=Z$

**Fig. 22.** Examples of the sampling distributions by histogram.




(a) Tension prediction performances (b) Computational costs

**Fig. 23.** Tension prediction performance with respect to a number of the wave states and the simulation length.

studied with respect to a number of the wave states and the simulation length, separately. It was discovered that the higher number of the wave states and the higher simulation length result in the better prediction performances in general. However, increases in a number of the wave states and the simulation length come with the computational cost. In this chapter, the tension prediction performance is studied with respect to a number of the wave states and the simulation length while the computational cost is kept constant by keeping the following condition: {number of the wave states $\times$ simulation length $=$ constant $c$} as Eq. (17). This way, it is possible to find out which one is better between the following options: 1) higher number of the wave states & shorter simulation length, 2) lower number of the wave states & longer length, 3) somewhere between the first and second options.

$$S \times L = c \tag{17}$$

The neural network model is trained on the training dataset created from different numbers of the wave states and different simulation lengths that follow the condition of Eq. (17). The trained neural network is tested on the test dataset created from the test wave states as Fig. 11 with the metric of Eq. (6), $\mathbb{M}_S$. The result is presented in Fig. 23(a) and its computational cost as runtime is shown in Fig. 23(b). Fig. 23(a) shows that the better prediction performance is achieved by following 'higher number of the wave states & shorter simulation length' at the same computational cost. However, it should be noted that the performance improvement converges over the number of the wave states of 12. Fig. 23(b) verifies that the same $c$ from Eq. (17) results in the same computational cost.

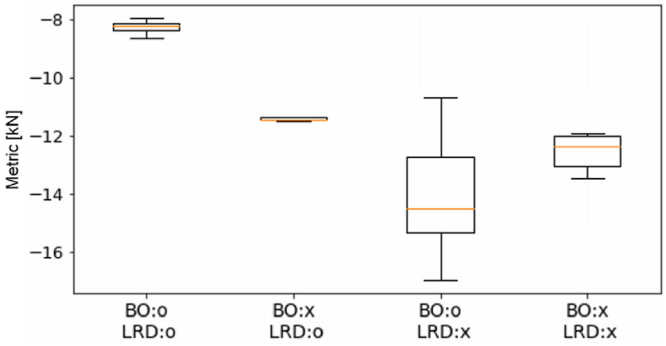



**Fig. 24.** Tension prediction performance with respect to the BN and LRD.

### 5.5. With respect to batch normalization and learning rate decay

The tension prediction performance with respect to the BN and LRD is studied. The neural network models with and without the BN and LRD are trained on the training dataset created from the 8 wave states by the uniform wave selection method with the simulation length of 3600s. The trained neural network is tested on the test dataset created from the test wave states as Fig. 11 with the metric of Eq. (6), $\mathbb{M}_S$. The result is shown in Fig. 24 where 'o' and 'x' denote 'used' and 'not used', respectively, therefore, four cases are considered in total. The result shows that the use of both the BN and the LRD provides an apparent improvement in the prediction performance with the low variance in the prediction performance.

Despite the fact that the Adam optimizer carries a decaying learning rate internally, an explicit learning rate decay provides the performance improvement. It can be explained by observing the parameter update equation by the Adam optimizer which is shown in Eqs. 18–22.

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)\nabla_\theta J(\theta_t) \tag{18}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2)(\nabla_\theta J(\theta_t))^2 \tag{19}$$

$$\widehat{m}_t = m_t/(1-\beta_1^t) \tag{20}$$

$$\widehat{v}_t = v_t/(1-\beta_2^t) \tag{21}$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\widehat{v}_t + \varepsilon}} \cdot \widehat{m}_t \tag{22}$$

where $m_t$ and $v_t$ are called a biased first moment estimate and a biased second moment estimate, respectively, and $\beta_1$ and $\beta_2$ are the exponential decay rates for $m_t$ and $v_t$, respectively. For settings of the parameters, $m_t$ and $v_t$ are initialized to zero, and $\beta_1$ and $\beta_2$ are usually set to 0.9 and 0.999, respectively. It can be observed that the learning rate decay occurs by an increase of $\widehat{v}_t$ in Eq. (6). Since $\widehat{v}_t$ increases as the training step goes on, it may be questionable to use a learning rate decay additionally. Our speculation is that although the Adam provides the effect of learning rate decay, if a parameter is already nearby an optimal point when the Adam's $\widehat{v}_t$ has not increased enough (*i.e.,* parameter update is still conducted by a large extent), it would experience oscillation in its weight update process. Additionally, it may take a long time (epochs) until $\widehat{v}_t$ increases large enough to make the parameter update converge. Instead, by providing an explicit learning rate decay, the parameters can be encouraged to converge faster.

## 6. Conclusion

The mooring line top tension prediction performances of the neural network model are studied with respect to the following four perspectives: 1) a distribution shape of the wave states, 2) a number of the wave states, 3) the simulation length, 4) the use of the BN and the LRD. Based on the results for each perspective, a guideline for the wave state selection is as follows: a) Use the uniform distribution of the wave states rather than the high-*Hs*-focused distribution since they result in the same performances and the uniform distribution is easier to implement, b) Select a high number of the wave states (i.e. at least 8) and choose the longest simulation length that is affordable given the available computational power, c) If the computational cost needs to be reduced, use the proposed approximate measure to quickly find the optimal simulation length, in which the performance is almost maximized, then the computational cost can be reduced by not using the bigger training dataset for the mere improvement. d) Use the batch normalization and the learning rate decay since they result in the apparent performance improvement.

Our study is the first and initial study to the best of our knowledge that analyzes the tension prediction performances of the neural network model with respect to a distribution shape of the wave states, a number of the wave states, and the simulation length in an effort to standardize them. As the relative previous works considered one main direction only, our work considers one main direction only to verify our hypothesis in a relatively smaller scale. Our work can be up-scaled and extended by considering various environmental directions in addition to various environmental conditions, which is to be the future direction of this work.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

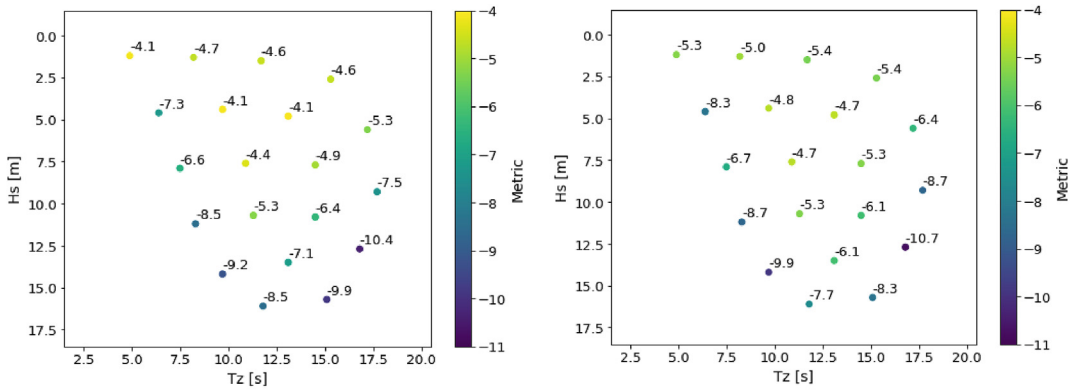
## Acknowledgements

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning(KETEP) grant funded by the Korea government(MOTIE) (Project Number:20213030020200, Project Name: Development of fully-coupled aero-hydro-servo-elastic-soil analysis program for offshore wind turbine system).

## A. Detailed Prediction Performances

In the main chapters, the prediction performances are evaluated by the metric of Eq. (6), $\mathbb{M}_S$, which calculates the metric on multiple wave states. But one might be curious about the metric on each wave state in the wave scatter diagram which can be computed using Eq. (5), $\mathbb{M}_s$, and the predicted tension time histories compared to the actual tension time histories from the dynamic simulation. In this supplementary chapter, the metric by Eq. (5), $\mathbb{M}_s$, and the predicted tension time histories are presented for various results in the main chapters.

In Chapter 5.1, the tension prediction performances with respect to the distribution shape of the wave states are studied, and there are two distribution shapes: 1) uniform distribution, 2) high-*Hs*-focused distribution. It was shown that the two distribution shapes result in almost the same performances. In the following figures, the metrics of the prediction performances by $\mathbb{M}_s$ are presented in Fig. A.1 and the predicted tension time histories compared to the actual tension time histories are presented in Fig A.2, in which a number of the wave states is set to 12 and the simulation length is set to 10800s for creating the training dataset. For Fig A.2, the tension time histories are presented on the test wave state with the lowest metric (*Hs*: 12.7m, *Tz*: 16.8s), the mooring line number of the presented tension time histories is 16 since this line is most affected by the wave load given its direction, and the 'actual' and 'predicted' refer to the tension time histories from the dynamic simulation and the predicted tension time histories by the neural network model, respectively. Since both distribution shapes result in the similar metrics and almost the same on the test wave state of {*Hs*: 12.7m, *Tz*: 16.8s}, the predicted tension time history of the neural network model with respect to the uniform distribution is presented only.




(a) With respect to the uniform distribution (b) With respect to the high-*Hs*-focused distribution

**Fig. A.1.** Detailed tension prediction performance with respect to the distribution shape of the wave states on the test wave states

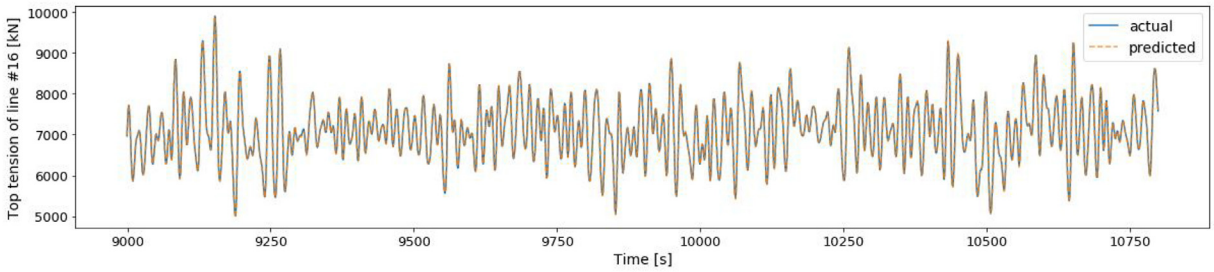
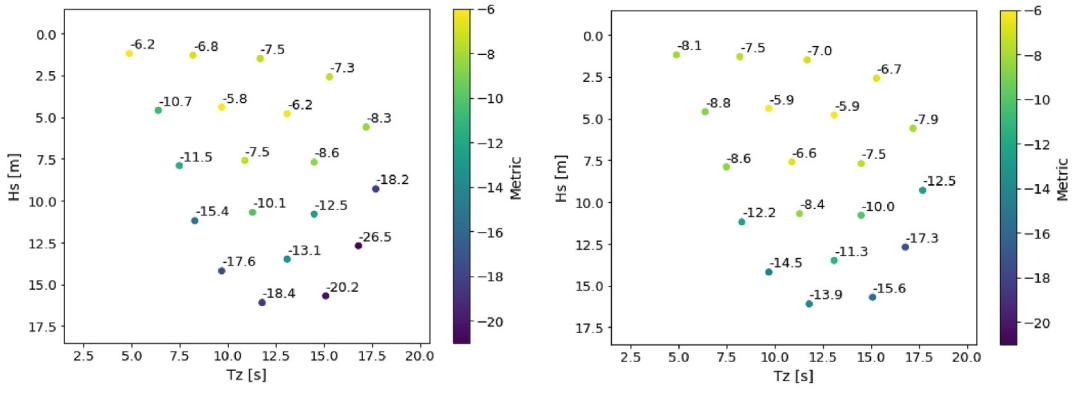


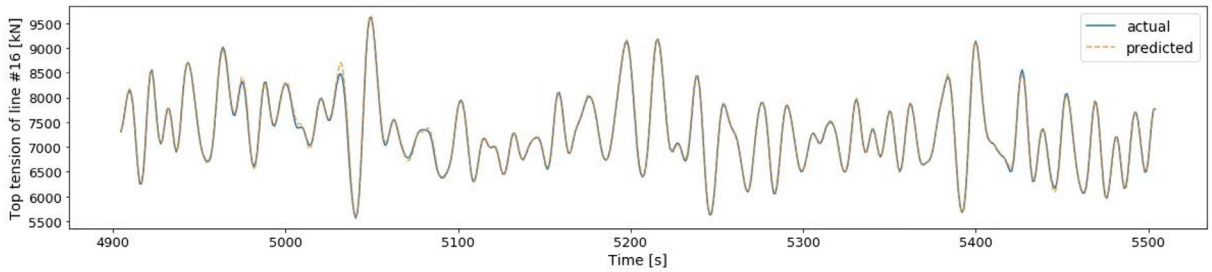

**Fig. A.2.** Actual and predicted tension time histories on the test wave state of {*Hs*: 12.7m, *Tz*: 16.8s}. This tension time history is based on a test random seed which is different from a training random seed. The time history would look different if a different random seed is used, but the prediction performance would remain the same as shown in Section 4.2.

In Chapter 5.4, the tension prediction performances are studied with respect to a number of the wave states and the simulation length altogether. It was concluded that 'higher number of the wave states & shorter simulation length' results in the better prediction performances than 'lower number of the wave states & longer simulation length' for the same computational cost. In the following figures, two cases are compared: a) number of the wave states: 4 & simulation length: 10800s, b) number of the wave states: 20 & simulation length: 2160s. The detailed tension prediction performances are presented in Fig. A.3. It is shown that the prediction performances are better in case of {*S*:20, *L*:2160s} than {*S*:4, *L*:10800s} on most of the test wave states. The actual and predicted tension time histories of both cases are presented in Fig. A.4, in which the test wave state is {*Hs*: 12.7m, *Tz*: 16.8s}.

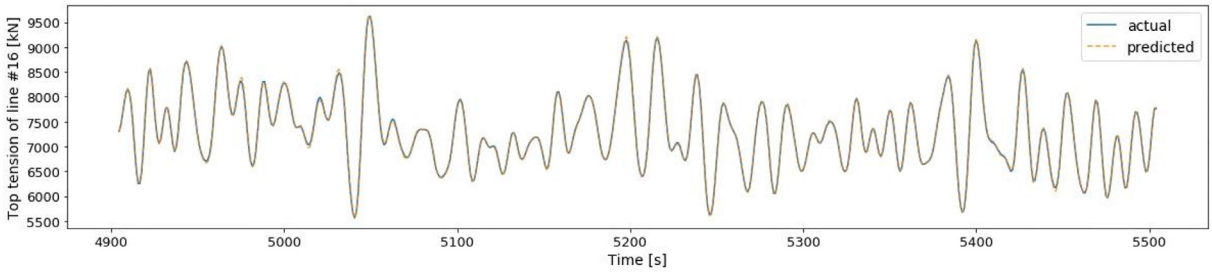



(a) *S*:4, *L*:10800s (b) *S*:20, *L*:2160s

**Fig. A.3.** Detailed tension prediction performance with respect to a number of the wave states and the simulation length on the test wave states



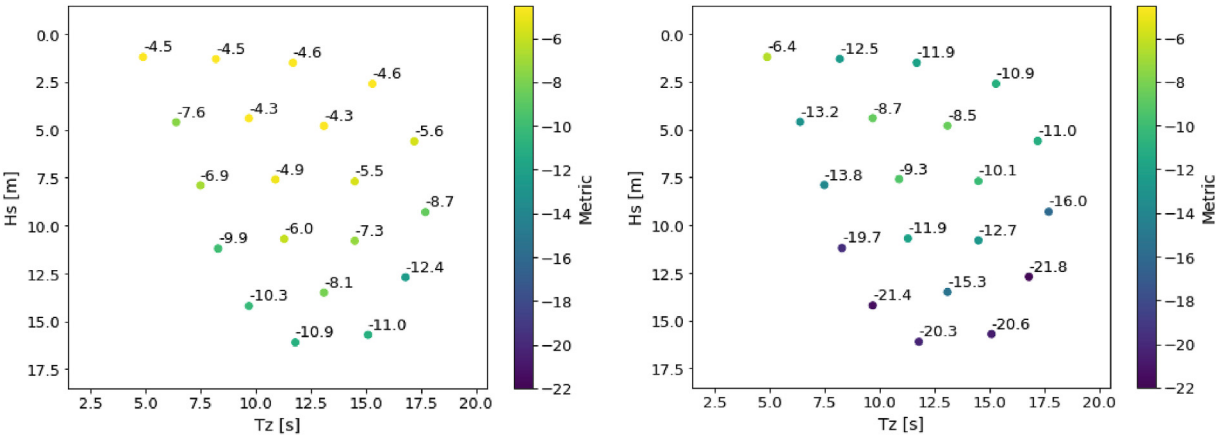

(a) $S$:4, $L$:10800s




(b) $S$:20, $L$:2160s

**Fig. A.4.** Actual and predicted tension time histories on the test wave state of {*Hs*: 12.7m, *Tz*: 16.8s}

In Chapter 5.5, the tension prediction performances are studied with respect to the batch normalization and learning rate decay. It was concluded that the use of the BN and LRD provided an apparent improvement in the prediction performance. The detailed prediction performances with respect to the BN and LRD are presented in Fig. A.5. It is shown that the use of the BN and LRD improves the overall prediction performance across various test wave states.




(a) BO: used; LRD: used (b) BO: not used; LRD: not used

**Fig. A.5.** Detailed tension prediction performance with respect to the BO and LRD on the test wave states

## References


Christiansen, N.H., Voie, P.E.T., Høgsberg, J., Sødahl, N., 2013. Efficient mooring line fatigue analysis using a hybrid method time domain simulation scheme. In: Proceedings of the AMSE 2013 32nd International Conference on Ocean, Offshore and Arctic Engineering.

Diaconescu, E., 2008. The use of NARX neural networks to predict chaotic time series. WSEAS Trans. Comput. Res. 3, 182–191.

DNV-GL, 2018. Class Guideline (DNVGL-CG-0130): Wave Loads.

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. Appl. Ocean Res. 29, 191–198. https://doi.org/10.1016/j.apor.2008.01.002.

Ioffe, S., Szegedy, C., 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: 32nd International Conference on Machine Learning, ICML 2015. International Machine Learning Society (IMLS), pp. 448–456.

Kang, H.H., Lee, D.S., Lim, J.S., Lee, S.J., Jang, J., Jung, K.H., Lee, J., 2021. A study on the effectiveness of the heading control on the mooring line tension and position offset for an arctic floating structure under complex environmental loads. J. Mar. Sci. Eng. 9, 1–17. https://doi.org/10.3390/jmse9020102.

Macqueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability.

Sidarta, D.E., Kyoung, J., O'Sullivan, J., Lambrakos, K.F., 2017. Prediction of offshore platform mooring line tensions using artificial neural network. In: Proceedings of the 27th International Ocean and Polar Engineering Conference, pp. 1–11.

Simões, M.G., Tiquilloca, J.L.M., Morishita, H.M., 2002. Neural-Network-based prediction of mooring forces in floating production storage and offloading systems. IEEE Trans. Ind. Appl. 38, 457–466. https://doi.org/10.1109/28.993167.

Wheelan, C., 2013. Naked Statistics: Stripping the Dread from Data. W. W. Norton & Company. Informa UK Limited. https://doi.org/10.1080/09332480.2014.890874.

Xie, H., Tang, H., Liao, Y.H., 2009. Time series prediction based on NARX neural networks: an advanced approach. In: Proc. 2009 Int. Conf. Mach. Learn. Cybern., vol. 3, pp. 1275–1279. https://doi.org/10.1109/ICMLC.2009.5212326.

Yetkin, M., Kim, M.S., Kim, Y., 2017. Mooring line top-tension prediction using NARX. In: Proceedings of the AMSE 2017 36th International Conference on Ocean, Offshore and Arctic Engineering, pp. 196–202.

Yetkin, M., Kim, Y., 2019. Time series prediction of mooring line top tension by the NARX and Volterra model. Appl. Ocean Res. 88, 170–186. https://doi.org/10.1016/j.apor.2019.04.013.

You, K., Long, M., Wang, J., Jordan, M.I., 2019. How Does Learning Rate Decay Help Modern Neural Networks? (arXiv).