Modeling of Coupling Effects in Neural Networks for Ship Motion Prediction

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

10-07-2021 / 25-01-2022

CITATION

Zhang, Chong; Hu, Peida; Zhang, Rong; Wu, Qiuping; Yang, Anlan (2021): Modeling of Coupling Effects in Neural Networks for Ship Motion Prediction. TechRxiv. Preprint.
<https://doi.org/10.36227/techrxiv.14945544.v2>

DOI

[10.36227/techrxiv.14945544.v2](https://doi.org/10.36227/techrxiv.14945544.v2)

Modeling of Coupling Effects in Neural Networks for Ship Motion Prediction

Chong Zhang, Peida Hu, Rong Zhang, Qiuping Wu, Anlan Yang

This work has been submitted to the IEEE for possible publication. Copyright may be transferred without notice, after which this version may no longer be accessible.

Abstract—Neural networks (NNs) have been applied to predict complex ship motions for stable ship maneuvering and operations at sea, but the coupling effects of different DoFs are still not effectively modeled in black-box NNs. To this end, we propose a novel deep NN framework that explicitly models the coupling effects through factorization machines, and significantly improves the prediction accuracy without introducing much extra complexity. The proposed framework is also lightweight and frequency-aware, which serves real-time and high-frequency prediction with high accuracy. We experimentally demonstrated that the framework is not only friendly to real-time prediction and online learning, but also interpretable and with great potential to be complemented by common machine learning tricks and go deeper.

Index Terms—Coupling effects, interpretable machine learning, neural network architecture, ship motion prediction.

I. INTRODUCTION

SHIP motion prediction is vital for ship maneuvering [1] and control of shipborne systems, such as launch of air vehicles [2], shipboard landing [3] [4], robotic manipulation [5], and control of offshore cranes [6]. Briefly speaking, a good control system at sea can never go without the accurate prediction of ship motions.

However, accurate prediction of ship motions is difficult. Ship motions are non-linear [7], chaotic [8], time-varying [9], and susceptible to environmental factors [10], such as waves [11], currents [12], and winds [13]. The different DoFs are highly coupled [14] [15], which adds to the difficulty of modeling ship dynamics. Moreover, the requirements of real-time and high-frequency prediction limit the complexity from going high. Therefore, it becomes an open challenge to trade off the prediction accuracy and the model complexity.

With a rising tide of machine learning lifting all the boats, neural networks (NNs) that are structurally complex but computationally elegant have been widely applied to ship motion prediction [16]. The emerging concept of learning-based control also asks for learning of dynamics [17], where deeper NNs typically have better expressive power but higher complexity [18]. Hence, an investigation into effective and efficient modeling of ship dynamics in NNs is of interest.

The authors are with the Department of Precision Instrument, Tsinghua University, Beijing 100084, China. (Corresponding author: Peida Hu.) (e-mail: chong-zh18@mails.tsinghua.edu.cn, hpd211@mail.tsinghua.edu.cn, rongzh@mail.tsinghua.edu.cn, wuqiuping@mail.tsinghua.edu.cn, yal20@mails.tsinghua.edu.cn)

This work was supported by China's National Training Program of Innovation and Entrepreneurship for Undergraduates.

Code implementations are provided on request of emails.

Fig. 1. Ship motions with 6 DoFs.

Specifically, in this paper, we study the explicit modeling of coupling effects in NNs which can significantly improve the prediction accuracy. Based on this, we further propose a novel framework called multi-resolution recurrent neural factorization machines (mRNFM). Besides the modeling of coupling effects, the framework is frequency-aware and interpretable, but of low complexity and friendly to real-time prediction and online learning. As machine learning techniques evolve, NNs are expected to go deeper, and our proposed framework can serve as a generic framework for ship motion prediction.

A. Related Work

1) Non-NN Prediction Models:

The prediction of ship motions, usually expressed by the 6 DoFs in Fig. 1, has long been investigated, and non-NN methods are well developed in this context. Typically, there are two kinds of non-NN methods: one kind based on parameterized models, such as Kalman filters [19] and model identification [20], and the other kind based on linear models, such as Autoregressive method (AR) [21]. These methods are based on specific assumptions, e.g., non-linear terms can be omitted for linearization [22], and a priori knowledge, such as simplified models of ship dynamics [23].

For the methods mentioned above, more detailed assumptions and knowledge can usually bring better accuracy, but there are always cases that these assumptions and knowledge cannot work, e.g., on rough sea [24] [25]. Methods based on nonlinear regression methods such as support vector regression [26] [27] [28] can be regarded as extensions of linear regression, but they are more complex and will be discussed later with NN-related methods.

The non-NN methods with a priori knowledge can be combined with NN models for better representation [29]

[30]. However, such representation requires not only detailed knowledge of specific vessels dynamics but also access to data of developed sensor systems. For shipborne robotic manipulators and air vehicles, ship motion data can be the only accessible data. Modeling of ship dynamics also requires specific knowledge of ship hull and mass distribution that can be changeable, which limits the use of such hybrid models. We don't mean to reject such models, but this paper will be only focused on the NN part for more generic development.

2) NN-related Prediction Models:

For NN-related methods, a typical way is to predict with shallow neural networks, such as RBF-NN [31] and FGWN [32], and parameters are updated based on certain criteria. Despite recent success of such methods in online prediction, the complexity is high because frequent parameter updates can lead to massive re-fitting of models. Also, for each new update, only a short sliding window is utilized for fitting, which limits the NNs from going deeper and learning from big data. Consequently, to achieve real-time prediction and online learning, the expressive power of NNs is limited, and so is the prediction accuracy. Similar things happen to methods based on support vector regression, where parameters are updated through combinatorial optimization [26] [27] [28].

Models based on recurrent neural networks (RNNs) [33] are preferred in this paper due to their capacities of time-series modeling and the structure friendly to real-time prediction. Here the RNNs include the vanilla RNN and its variants, such as LSTM [34] and GRU [35]. Contrary to the models mentioned in the above paragraph, RNN models can learn from the whole sequence instead of a sliding window. There can be only one-step-forward calculation for each new prediction, and parameters are updated through gradient descent. Therefore, RNN models can go wider and deeper while maintaining real-time prediction, and get complemented by novel NN techniques, such as attention [36], and other algorithms, such as Gaussian Process [37].

Other NN models, such as CNNs [38] and echo state networks [39], have shown advantages in neither prediction accuracy nor computational complexity when compared with RNNs. Hence, we do not consider them in this paper.

3) Wavelet Transform:

Since the attitude of ships varies quasi-periodically [26], frequency information can help improve the representation in NNs. A typical tool is wavelet transform (WT), a multi-resolution analysis method with time-frequency localization [40]. There are three ways to combine WT with NN: 1) to use wavelet functions as activation functions [41] [42] [32], 2) to apply WT directly for feature engineering [36] [43] [44], and 3) to embed wavelet decomposition into NNs [45].

Among the three ways mentioned above, the first way hardly utilizes the frequency since it only assumes that the target can be represented by the combination of wavelet kernel functions. The second way successfully utilizes the frequency information, but some adaptations are necessary for further applications in RNNs. The third way is not friendly to real-time prediction, because it allocates independent parallel neural networks to each sub-series and then fully connects them, leading to considerable extra complexity.

Fig. 2. A diagram of the preferred RNN structure. The model can sequentially learn from the entire dataset and supports gradient-based online learning. A new prediction does not require the recalculation for the previous inputs, and the past information is encoded in the hidden states.

Besides WT, there are other methods of signal decomposition, such as EMD [46], but they have shown neither better interpretability nor better performance than WT, so we don't consider them in this paper.

4) Coupling Effects:

Numerical studies have been conducted on coupling effects [47] [48], but such implementations are not realistic for real-time prediction models. In addition, since we cannot guarantee the long-term accuracy of the numerical simulation results, our models are directly trained on real ship data.

In non-NN models, typical modeling of coupling effects is based on the second-order terms, i.e., the product of 2 DoFs or the square of 1 DoF, and lower-order terms, while terms of higher orders are omitted because they are of small values and technically hard to achieve [49] [50]. Such approximation can greatly improve the modeling of ship dynamics. In NN-based models, different DoFs are typically implicitly coupled by black-box NN models, while no explicit representation of the coupling effects has been achieved to our knowledge.

B. Motivation and Our Contribution

Machine learning is moving toward big data and big models, but for ship motions, real-time prediction requires low-complexity models, and online learning is necessary for time-varying ship dynamics. RNN models can outperform other methods in terms of computational complexity and utilization of big data because: 1) gradient-based optimization allows RNNs to learn from the entire dataset and quickly adapt to new data, and 2) past information is encoded in hidden states so that previous data are not reprocessed for prediction. A diagram of such an RNN scheme is shown in Fig. 2.

Regrettably, existing works [36] [37] [38] typically used shallow RNNs (≤ 3 hidden layers), a small number of data ($< 0.1M$ data points), and the sliding window, which can hardly show the advantages of RNNs. In this paper, we show that RNN models can go much deeper with our framework. Also, for RNN models with WT [36], the feature of low computational complexity is underutilized because past inputs are repeatedly used for different new predictions with sliding windows, which calls for an adaptation of WT (see Sec. II-A).

Modeling of the coupling effects through second-order terms makes sense for non-NN methods [49] [50], but it

is not easy to directly apply it in NNs, especially when we are applying WT. To utilize frequency information, we apply WT as many works did [36] [43] [44], and 5 levels of WT are empirically necessary for common 10-Hz data [36], decomposing each signal dimension to at least 6 dimensions. Consider when we use the 6 DoFs in Fig. 1, there can be 36 sub-series after WT, thereby generating $C_{36}^2 = 630$ second-order terms, i.e., hundreds of dimensions to represent the coupling effects. Such a high dimensionality can lead to the curse of dimensionality [51] for a real-time prediction model.

Hence, we cannot directly apply the second-order terms after WT as features. Instead, we use a factorized expression (see Sec. II-B) that is of low complexity but capable of modeling interactions between different features. In this way, not only the coupling DoFs but also the coupling high-frequency and low-frequency parts can be modeled.

A generic framework is necessary as well, because there can be different inputs that different controllers can access, and the input dimensions can be different from the output dimensions. Therefore, instead of achieving a state-of-the-art (SOTA) model, we are exploring the possibilities of RNNs going deeper and getting improved via explicit modeling of coupling effects.

In this paper, our contributions include:

- 1) an adaptation of WT for the scheme of real-time ship motion prediction and online learning,
- 2) a real-time frequency-aware deep RNN framework that explicitly models the coupling effects in ship motions,
- 3) implementations of techniques and visualization that can be complementary to the proposed framework.

II. METHODOLOGY

We present our methodology in this section, and the used symbols and notations are listed in Table I for easy lookup.

A. Adapted Wavelet Transform for Real-Time Prediction

In our proposed RNN scheme, the previous inputs should not be reused for a new prediction, i.e., the applied WT should be translation-invariant. Otherwise, the previous input can be changed when a new time step comes, leading to the change of the previous output that should not happen. Therefore, the commonly used discrete WT [52] cannot be applied here. Instead, we use the time-invariant stationary WT [53].

Mathematically, the stationary WT is determined by a high-pass filter HP and a low-pass filter LP :

$$\begin{aligned} HP &= [HP_1, HP_2, \dots, HP_p], \\ LP &= [LP_1, LP_2, \dots, LP_p], \end{aligned} \quad (1)$$

where p is the length of the filters. The standard multilevel stationary WT can be defined with dilated convolution [54]:

$$\begin{cases} A^{[j]} &= X *_1 LP, & j = 1 \\ D^{[j]} &= X *_1 HP, & j = 1 \\ A^{[j]} &= A^{[j-1]} *_2 LP, & j = 2, 3, \dots, J \\ D^{[j]} &= A^{[j-1]} *_2 HP, & j = 2, 3, \dots, J \end{cases} \quad (2)$$

where X is the input sequence of length l for stationary WT, $A^{[j]}$ and $D^{[j]}$ represents the approximate component and the

TABLE I
SYMBOLS AND NOTATIONS

Symbols	Meaning
x, y, ζ	the surge, sway, and heave DoFs in Fig. 1
ϕ, θ, ψ	the roll, pitch, and yaw DoFs in Fig. 1
D	the set of used DoFs and other sequences as inputs
J	the number of levels of WT
LP	the low-pass filter coefficients for WT
HP	the high-pass filter coefficients for WT
p	the length of the filters LP and HP
$A^{[j]}$	the approximate component after the j -th level of WT
$D^{[j]}$	the detail component after the j -th level of WT
l	the length of the input sequence for WT
X	the input sequence for WT
S	the output sequences for WT
s	the used sequences in S after WT for FMs
z	the outputs of the multiple-FM layer
FM_i	the output of the i -th FM
m	the number of the time distributed FMs
w_i	the linear weights for the i -th FM
B_i	the factorization matrix for the i -th FM
b	the length of the row vectors in B_i
σ_z	the activation function of the multiple-FM layer
h^k	the hidden states of the k -th hidden layer in RNN
K	the number of hidden layers in RNN
O_t	the outputs at the t -th time step
I_t	the inputs at the t -th time step

detail component respectively after the j -th level of WT, J is the maximum level of WT, and $*_d$ represents the d -dilated convolution for any positive integer d . Such a process is illustrated in Fig. 3. In standard stationary WT, the padding is used to keep the length of each component the same as l after convolution.

However, such padding, especially the padding at the future end of a sequence, is problematic for real-time prediction with one-step calculation because it hurts the translation invariance. During the convolution, the padding at the future end is fulfilled by assumed values that are different from the real values to come. With a new time step forward, some of the assumed values are replaced by the real values, which can lead to changes of the previous outputs. Specifically, the values that are polluted by the future-end padding are displayed in Fig. 4, while the past-end padding can be replaced by the past real values.

To achieve our scheme of one-step calculation, we should only take the valid values after convolution. Therefore, only the causal convolution [56] can be applied without the polluted values. In the standard stationary WT formulation, the valid outputs S of WT at time step t should be

$$S_t = [X_t, A_{t-\frac{p}{2}}^{[1]}, D_{t-\frac{p}{2}}^{[1]}, \dots, A_{t-\frac{(2^J-1)P}{2}}^{[J]}, D_{t-\frac{(2^J-1)P}{2}}^{[J]}]^T. \quad (3)$$

Since all of the components after WT have time delay that may affect the short-term prediction accuracy, we include the input sequence X here. Not all elements in the S_t are necessarily used, and (3) is an enumeration of the full set.

Fig. 3. An illustration of the standard WT process. The padding is used to keep the length of each component the same as the length of the input sequence X .

Fig. 4. Polluted values in each component of WT due to the future-end padding. For standard stationary WT, the filter length p is even, and $\frac{p}{2}$ and $\frac{p-2}{2}$ are respectively the padding length at the future end and the past end. The pollution is also called the boundary effects [55].

We do not recommend wavelet filters of large lengths, because a large p can bring great delay in high-level components of WT according to (3) and Fig. 4. Instead, we would recommend Haar wavelet [55] with $p = 2$,

$$LP = [\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}], HP = [\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}]. \quad (4)$$

B. Factorization of Coupling Effects

As is mentioned in Sec. I-A4 and I-B, an introduction of high-order terms, especially the second-order terms, can help the representation of coupling effects in the model and improve the performance, but it is not easy to achieve this in NNs when combined with WT.

To this end, we propose to factorize the coupling effects by factorization machines (FMs) [57]. Specifically, with the coupled sequences s , we use m FMs to achieve an informative time distributed representation, and the output of the i -th FM working on the t -th time step of s is

$$FM_i(s_t) = w_{i,0} + \sum_{u=1}^{\dim s_t} w_{i,u} s_{t,u} + \sum_{u=1}^{\dim s_t} \sum_{v=u+1}^{\dim s_t} \langle B_{i,u}, B_{i,v} \rangle \cdot s_{t,u} s_{t,v}, \quad (5)$$

where $w_i \in \mathbf{R}^{1+\dim s_t}$, $B_i \in \mathbf{R}^{\dim s_t \times b}$ are model parameters to be estimated through learning, and b is a positive integer to be determined by human. In this way, there is no increase in the input dimensions, and $m \times \dim s_t \times b$ additional learnable parameters are used to represent the second-order terms.

According to [57], both the computation and the update through gradient descent are in $O(b \cdot \dim s_t)$ for one FM, which can be much smaller than the number of second-order terms $O(\dim s_t^2)$. Therefore, FM is of quite low complexity. Since there are many FMs used, we consider $b = 2$ as enough for representation. Because of NNs' strong capabilities of non-linear fitting, the factorization to the second-order is sufficient when combined with NNs [58], although FMs can be of higher orders [59].

The m FMs can make a multiple-FM layer of which the outputs at the t -th time step are

$$z_t = \sigma_z([FM_1(s_t), \dots, FM_m(s_t)]^T), \quad (6)$$

where $\sigma_z(\cdot)$ can be any activation function used in NNs. In our framework, z_t will be used as the inputs of follow-up NNs.

C. Multi-resolution Recurrent Neural Factorization Machines

Based on the adapted WT in Sec. II-A and the factorization of coupling effects in Sec. II-B, a frequency-aware deep RNN framework that explicitly models the coupling effects in ship motions is proposed, with the architecture and the scheme illustrated in Fig. 5. We call it multi-resolution recurrent neural factorization machines (mRNFM).

In mRNFM, the input sequences are the used DoFs D which can include the 6 DoFs in Fig. 1 and their derivatives of any order with respect to time. Other sequences can also be included in D if available. Such a setting can facilitate the generalization of the proposed framework to systems with different accessible data. The sequences in D are respectively decomposed by WT to extract frequency information.

After WT, we get the output sequences S , some of which (denoted by s) are used as the inputs of the time distributed multiple-FM layer. The multiple-FM layer models the coupling effects in different sub-series, and its outputs serve as the inputs of the follow-up RNN which then generates the predictions as the outputs. The RNN can be of arbitrary structure, as long as it can perform the scheme in Fig. 2. For a new prediction, a new time step of s is obtained through WT, and

Fig. 5. The proposed framework called multi-resolution recurrent neural factorization machines (mRNFM). Black arrows for prediction, and green arrows for optimization.

TABLE II
STATISTICAL PARAMETERS OF USED DATA

DoF	Surge Rate	Sway Rate	Heave	Roll	Pitch	Yaw Rate
Feature	$\dot{x}$	$\dot{y}$	ζ	ϕ	θ	$\dot{\psi}$
Unit	$m \cdot s^{-1}$	$m \cdot s^{-1}$	m	deg	deg	$deg \cdot s^{-1}$
Mean	5.72	-0.230	0	0	0	≈ 0
SD	0.578	0.414	0.470	1.26	0.596	0.302
Maximum	7.03	1.76	2.08	8.13	2.51	1.23
Minimum	4.31	-1.83	-2.09	-7.13	-2.74	-1.30

then the multiple-FM layer and the RNN execute the one-step calculation.

III. EXPERIMENTAL RESULTS

A. Data and Task

We validated our method directly on the data collected from a ship maneuvering test, because deep NNs must learn from massive data and we cannot guarantee the long-term accuracy of the numerical simulation results, especially when the coupling effects are involved. The used data were collected from a shipborne inertial navigation system (INS) which consists of two dual-axis gyroscopes with low drift rates and three pendulous accelerometers, as we used in [60]. The accelerometer bias and gyro drift are respectively $2 \times 10^{-5}g$ and $1 \times 10^{-3} deg/h$ [61].

The dataset was sampled at 10 Hz in East China Sea and spans a period of 18 hours, i.e., $\sim 0.6M$ data points, with the sea state varying between 3 $\sim$ 5. The test includes multiple sections of straight sailing and changes of direction. The used DoFs are surge rate $\dot{x}$, sway rate $\dot{y}$, heave ζ , roll ϕ , pitch θ , and yaw rate $\dot{\psi}$. The statistical parameters are shown in Table II, with the mean values of roll ϕ , pitch θ , and heave ζ set as zeros.

The task is to forecast the roll motions of 8 s ahead, i.e., 80 output dimensions. Such a prediction horizon can be long because of frequent and unscheduled maneuvering. The used sub-series after WT are $X, A^{[5]}, D^{[1]}, D^{[2]}, D^{[3]}, D^{[4]}, D^{[5]}$ for all of the 6 DoFs, i.e., 42 sequences in total in s , and the Haar wavelet in (4) is used. Figure 6(a) shows the roll values of the vessel when the amplitude reaches 8.13 deg, together with other used sub-series for roll.

The first 9 hours of the data are used for training, while the rest are for prediction and online learning. The online

Fig. 6. Data of ship roll motions. (a) shows the sub-series of roll motions after WT when the roll attitude reaches 8.13 deg. The divisions in the legend are to make the signals contain the proper amount of energy. (b) shows the roll attitudes for the entire dataset. The offline training part and the prediction part have different data distributions, which necessitates online learning.

optimization is executed per 1 s, i.e., 10 predictions. The roll attitude shown in Fig. 6(b) indicates that online learning is necessary.

For better learning, all of the 6 used DoFs are standardized to have a mean of 0 and a standard deviation of 1, and the sub-series take the divisions shown in the legend of Fig. 6(a) to contain the proper amount of energy. These are specified for the task as part of manual feature engineering.

B. Comparative Tests

We tried different models for the prediction task with same training hyperparameters, some for ablation studies and some for explorations of advanced combinations with common machine learning techniques, as is shown in Fig. 7. In the ablation studies, effects of the multiple-FM layer, the WT layer, and the online learning are studied. To show that the proposed mRNFM can develop with new deep learning techniques, two trials are made: to use deeper RNNs, and to combine the framework with a simple multiplicative attention mechanism on the sub-series.

We also introduced the AR method with Akaike's Information Criterion (AIC) [62] as a baseline, the performance of which can generally reflect the difficulty of prediction. We do not add other methods for comparative tests because:

- 1) it is hard to achieve fair comparisons with way too many other methods that can be tricky to implement and sensitive to configurations;
- 2) the advantages over the AR model can be used for implicit comparisons with other models;

Fig. 7. Model configurations for comparative tests. The middle is the standard mRNFM model for comparison. The left lists the ablation studies to verify the efficacy of our proposed method. The right presents the advanced trials to show that mRNFM has the potential to evolve together with deep learning techniques.

TABLE III
EXPERIMENTAL RESULTS

Model	RMSE (deg)	# Param.	Time cost per * (s)	
			Prediction	Optimization
AR	0.647	127	0.002	0.071
<i>Ablation Studies</i>				
Offline	0.545	323k	0.007	-
w/o WT w/o FM	0.500	307k	0.005	0.012
w/o FM	0.476	312k	0.006	0.012
w/o WT	0.497	309k	0.005	0.012
Standard mRNFM [†]	0.458	323k	0.007	0.012
<i>Advanced Trials</i>				
Deeper [†]	0.453	421k	0.009	0.020
w/ Attn. [†]	0.450	326k	0.008	0.013

[†] Proposed methods.

- 3) too many factors, including how data are collected, can affect the results and are hard to control;
- 4) we are not trying to achieve a SOTA model, but to explore the possibilities of deep learning for ship motion prediction;
- 5) and most importantly, our aim is to validate the efficacy of our proposed technique in such an NN framework.

The offline training stage for NNs is to execute RMSprop optimizer with MSE loss for slices in batch, and the online learning stage uses the SGD optimizer with hypergradient descent [63] for the accumulated gradient at 1-s intervals. For AR method, the best order is 126 according to AIC on the offline training set, and online learning is achieved by refitting the model on the latest 1260 points per 1 s.

The results are shown in Table III, with the metrics calculated over the entire time horizon, i.e., $\sim 9 \times 3600 \text{ s} \times 10 \text{ step/s} \times 80 \text{ points/step}$ in total. Since the sea state varies, we also calculated the RMSE for each hour in the prediction stage and show them in Fig. 8. The time cost is tested on $8 \times 1.80 \text{ GHz}$ CPU with Python 3.7 + Pytorch 1.9, indicating that the proposed models can run on FPGA [64] in real time and spare enough time for follow-up controllers.

It can be seen from the results that online learning is necessary. The mere introduction of FM on the 6 DoFs does not make much sense, but the combination of WT and FM can make a great improvement. The FM does not significantly increase the model complexity, and fast prediction and online learning are achieved by the one-step calculation and the

Fig. 8. Prediction RMSE of the tested models, separately calculated for each hour in the prediction stage. "Amp" is the maximum roll attitude during each hour indicating the sea state, and it is multiplied by 0.1. Solid lines with squares are our proposed models. Dashed lines with circles are for ablation and comparison.

Fig. 9. Visualization of the learned second-order coupling effects in the multiple-FM layer. The upper right triangular part is for the "Standard mRNFM" model, and the lower left triangular part is for the "Deeper" model. For each DOF-related block, the 7 sub-blocks are respectively for the sub-series $D^{[1]}, X, D^{[2]}, D^{[3]}, D^{[4]}, D^{[5]}, A^{[5]}$, left to right or top to bottom, in an order from low-frequency part to high-frequency part.

gradient-based optimization. Both a deeper NN and a better architecture can bring better accuracy, but the latter one is more efficient in complexity.

The learned second-order coupling effects in the multiple-FM layer are visualized in Fig. 9. The "Standard mRNFM" model and the "Deeper" model are used for visualization, and the similar two patterns indicate that the learned patterns are correct and interpretable. Specifically, the values for visualization are calculated by

$$\text{Vis}[u, v] = \sqrt{\frac{1}{128} \sum_{i=1}^{128} (\langle B_{i,u}, B_{i,v} \rangle)^2}, \quad (7)$$

where $1 \leq u \leq 42$ and $1 \leq v \leq 42$ represents the indices of two sub-series, and all of the 128 units in the multiple-FM layer are used for calculation. Not only the coupling effects

for different DoFs, but also those for high-frequency parts and low-frequency parts are modeled in this way.

IV. CONCLUSION

In this paper, we explored the possibilities of deep learning for ship motion prediction. We proposed to model the coupling effects through multiple FMs in a lightweight NN framework, mRNFM, and the prediction and online optimization can run in real time. Experiments and visualization verified the efficacy and the potential of our proposed method. High prediction accuracy is achieved in limited complexity, and the model can hopefully be deployed for lifelong learning on real ships.

Future works will focus on incorporating more sensor data, such as the wind speed and the propulsion power, which may improve the performance of mRNFM. More investigations will be done to fully exploit the potential of mRNFM.

Actually, the RNN can go much deeper [65], and the architecture can be much more complicated [66] [67]. There are also many other developed online learning techniques [68]. That said, these methods are not implemented in this paper because the great efforts for this can hardly add to the contributions of our paper. Instead, we provide the references for those who are interested in this, and encourage readers to explore more potential of the proposed mRNFM.

REFERENCES

- [1] H. Liang, H. Li, and D. Xu, "Nonlinear model predictive trajectory tracking control of underactuated marine vehicles: Theory and experiment," *IEEE Trans. Ind. Electron.*, vol. 68, no. 5, pp. 4238–4248, 2020.
- [2] A. Khan, C. Bil, and K. E. Marion, "Ship motion prediction for launch and recovery of air vehicles," in *Proc. OCEANS 2005 MTS/IEEE*, pp. 2795–2801. IEEE, 2005.
- [3] Y. Meng, W. Wang, H. Han, and M. Zhang, "A vision/radar/ins integrated guidance method for shipboard landing," *IEEE Trans. Ind. Electron.*, vol. 66, no. 11, pp. 8803–8810, 2019.
- [4] Y. Huang, M. Zhu, Z. Zheng, and K. H. Low, "Linear velocity-free visual servoing control for unmanned helicopter landing on a ship with visibility constraint," *IEEE Trans. Syst., Man, Cybern., Syst.*, DOI 10.1109/TSMC.2021.3062712, 2021.
- [5] P. J. From, J. T. Gravdahl, and P. Abbeel, "On the influence of ship motion prediction accuracy on motion planning and control of robotic manipulators on seaborne platforms," in *IEEE Int. Conf. Robot. Autom.*, pp. 5281–5288. IEEE, 2010.
- [6] S. Küchler, T. Mahl, J. Neupert, K. Schneider, and O. Sawodny, "Active control for an offshore crane using prediction of the vessel's motion," *IEEE/ASME Trans. Mechatronics*, vol. 16, no. 2, pp. 297–309, 2010.
- [7] O. Hızır, M. Kim, O. Turan, A. Day, A. Incecik, and Y. Lee, "Numerical studies on non-linearity of added resistance and ship motions of kvlcc2 in short and long waves," *Int. J. Nav. Archit. Ocean Eng.*, vol. 11, no. 1, pp. 143–153, 2019.
- [8] L. Zhou and F. Chen, "Chaotic motion of the parametrically excited roll motion for a class of ships in regular longitudinal waves," *Ocean Eng.*, vol. 195, p. 106729, 2020.
- [9] J. Du, C. Guo, S. Yu, and Y. Zhao, "Adaptive autopilot design of time-varying uncertain ships with completely unknown control coefficient," *IEEE J. Ocean. Eng.*, vol. 32, no. 2, pp. 346–352, 2007.
- [10] X. Cheng, G. Li, R. Skulstad, S. Chen, H. P. Hildre, and H. Zhang, "A neural-network-based sensitivity analysis approach for data-driven modeling of ship motion," *IEEE J. Ocean. Eng.*, vol. 45, no. 2, pp. 451–461, 2019.
- [11] P. Han, G. Li, X. Cheng, S. Skjong, and H. Zhang, "An uncertainty-aware hybrid approach for sea state estimation using ship motion responses," *IEEE Trans. Ind. Inform.*, 2021.
- [12] S. Chunyu, X. Zhang, and G. Zhang, "Nonlinear identification for 4 dof ship maneuvering modeling via full-scale trial data," *IEEE Trans. Ind. Electron.*, 2021.
- [13] C. Chen, K. Sasa, J. Prpić-Oršić, and T. Mizojiri, "Statistical analysis of waves' effects on ship navigation using high-resolution numerical wave simulation and shipboard measurements," *Ocean Eng.*, vol. 229, p. 108757, 2021.
- [14] M. A. Neves and C. A. Rodriguez, "On unstable ship motions resulting from strong non-linear coupling," *Ocean Eng.*, vol. 33, no. 14-15, pp. 1853–1883, 2006.
- [15] Y. Kim, B. Nam, D. Kim, and Y. Kim, "Study on coupling effects of ship motion and sloshing," *Ocean Eng.*, vol. 34, no. 16, pp. 2176–2187, 2007.
- [16] X. Zhao, R. Xu, and C. Kwan, "Ship-motion prediction: algorithms and simulation results," in *IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 5, pp. V–125. IEEE, 2004.
- [17] J. F. Fisac, A. K. Akametalu, M. N. Zeilinger, S. Kaynama, J. Gillula, and C. J. Tomlin, "A general safety framework for learning-based control in uncertain robotic systems," *IEEE Trans. Autom. Control*, vol. 64, no. 7, pp. 2737–2752, 2018.
- [18] M. Raghu, B. Poole, J. Kleinberg, S. Ganguli, and J. Sohl-Dickstein, "On the expressive power of deep neural networks," in *Int. Conf. Mach. Learn.*, pp. 2847–2854. PMLR, 2017.
- [19] X. Peng, B. Zhang, and L. Rong, "A robust unscented kalman filter and its application in estimating dynamic positioning ship motion states," *J. Mar. Sci. Technol.*, vol. 24, no. 4, pp. 1265–1279, 2019.
- [20] A. Somayajula and J. Falzarano, "Application of advanced system identification technique to extract roll damping from model tests in order to accurately predict roll motions," *Appl. Ocean Res.*, vol. 67, pp. 125–135, 2017.
- [21] H. Jiang, S. Duan, L. Huang, Y. Han, H. Yang, and Q. Ma, "Scale effects in ar model real-time ship motion prediction," *Ocean Eng.*, vol. 203, p. 107202, 2020.
- [22] T. Ozaki, "A bridge between nonlinear time series models and nonlinear stochastic dynamical systems: a local linearization approach," *Statistica Sinica*, pp. 113–135, 1992.
- [23] M. Zhu, A. Hahn, Y.-Q. Wen, and A. Bolles, "Identification-based simplified model of large container ships using support vector machines and artificial bee colony algorithm," *Appl. Ocean Res.*, vol. 68, pp. 249–261, 2017.
- [24] I. Senjanović, J. Parunov, and G. Ciprić, "Safety analysis of ship rolling in rough sea," *Chaos, Solitons & Fractals*, vol. 8, no. 4, pp. 659–680, 1997.
- [25] K. Sasa, K. Takeuchi, C. Chen, O. M. Faltinsen, J. Prpić-Oršić, M. Valčić, T. Mrakovčić, and N. Herai, "Evaluation of speed loss in bulk carriers with actual data from rough sea voyages," *Ocean Eng.*, vol. 187, p. 106162, 2019.
- [26] M.-W. Li, J. Geng, W.-C. Hong, and L.-D. Zhang, "Periodogram estimation based on lssvr-ccpsv compensation for forecasting ship motion," *Nonlinear Dyn.*, vol. 97, no. 4, pp. 2579–2594, 2019.
- [27] M.-W. Li, J. Geng, D.-F. Han, and T.-J. Zheng, "Ship motion prediction using dynamic seasonal rvsr with phase space reconstruction and the chaos adaptive efficient foa," *Neurocomputing*, vol. 174, pp. 661–680, 2016.
- [28] H. Fu, S. Liu, and F. Sun, "Ship motion prediction based on aga-lssvm," in *IEEE Int. Conf. Mechatron. Autom.*, pp. 202–206. IEEE, 2010.
- [29] T. Wang, G. Li, L. I. Hatledal, R. Skulstad, V. Aesoy, and H. Zhang, "Incorporating approximate dynamics into data-driven calibrator: A representative model for ship maneuvering prediction," *IEEE Trans. Ind. Inform.*, vol. 18, no. 3, pp. 1781–1789, 2021.
- [30] R. Skulstad, G. Li, T. I. Fossen, B. Vik, and H. Zhang, "A hybrid approach to motion prediction for ship docking—integration of a neural network model into the ship dynamic model," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2020.
- [31] J. Yin, N. Wang, and A. N. Perakis, "A real-time sequential ship roll prediction scheme based on adaptive sliding data window," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 12, pp. 2115–2125, 2017.
- [32] B.-G. Huang, Z.-J. Zou, and W.-W. Ding, "Online prediction of ship roll motion based on a coarse and fine tuning fixed grid wavelet network," *Ocean Eng.*, vol. 160, pp. 425–437, 2018.
- [33] H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent neural networks for time series forecasting: Current status and future directions," *Int. J. Forecast.*, vol. 37, no. 1, pp. 388–427, 2021.
- [34] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [35] K. Cho, B. van Merriënboer, C. Gulcehre, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," in *Conf. Empir. Methods Nat. Lang. Process.*, 2014.

- [36] T. Zhang, X.-Q. Zheng, and M.-X. Liu, "Multiscale attention-based lstm for ship motion prediction," *Ocean Eng.*, vol. 230, p. 109066, 2021.
- [37] Q. Sun, Z. Tang, J. Gao, and G. Zhang, "Short-term ship motion attitude prediction based on lstm and gpr," *Appl. Ocean Res.*, p. 102927, 2021.
- [38] M. H. Rashid, J. Zhang, and M. Zhao, "Real-time ship motion forecasting using deep learning," in *Int. Conf. Data Sci. Commun.*, pp. 1–5, 2021.
- [39] X. Peng, H. Dong, and B. Zhang, "Echo state network ship motion modeling prediction based on kalman filter," in *IEEE Int. Conf. Mechatron. Autom.*, pp. 95–100. IEEE, 2017.
- [40] I. Daubechies, *The wavelet transform, time-frequency localization and signal analysis*. Princeton University Press, 2009.
- [41] S. A. Billings and H.-L. Wei, "A new class of wavelet networks for nonlinear system identification," *IEEE Trans. Neural Netw.*, vol. 16, no. 4, pp. 862–874, 2005.
- [42] H. Zhao, S. Gao, Z. He, X. Zeng, W. Jin, and T. Li, "Identification of nonlinear dynamic system using a novel recurrent wavelet neural network based on the pipelined architecture," *IEEE Trans. Ind. Electron.*, vol. 61, no. 8, pp. 4171–4182, 2013.
- [43] N. Jin, Y. Zeng, K. Yan, and Z. Ji, "Multivariate air quality forecasting with nested lstm neural network," *IEEE Trans. Ind. Informat.*, 2021.
- [44] F. Duan, L. Dai, W. Chang, Z. Chen, C. Zhu, and W. Li, "semg-based identification of hand motion commands using wavelet neural network combined with discrete wavelet transform," *IEEE Trans. Ind. Electron.*, vol. 63, no. 3, pp. 1923–1934, 2015.
- [45] J. Wang, Z. Wang, J. Li, and J. Wu, "Multilevel wavelet decomposition network for interpretable time series analysis," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 2437–2446, 2018.
- [46] Z. Nie, F. Shen, D. Xu, and Q. Li, "An emd-svr model for short-term prediction of ship motion using mirror symmetry and svr algorithms to eliminate emd boundary effect," *Ocean Eng.*, vol. 217, p. 107927, 2020.
- [47] J. R. Saripilli and D. Sen, "Numerical studies on effects of slosh coupling on ship motions and derived slosh loads," *Appl. Ocean Res.*, vol. 76, pp. 71–87, 2018.
- [48] X.-s. Zhang, J.-h. Wang, and D.-c. Wan, "Numerical techniques for coupling hydrodynamic problems in ship and ocean engineering," *J. Hydrodyn.*, vol. 32, no. 2, pp. 212–233, 2020.
- [49] E. M. Lewandowski, *The dynamics of marine craft: maneuvering and seakeeping*, vol. 22. World scientific, 2004.
- [50] L. Chen, R. Cui, C. Yang, and W. Yan, "Adaptive neural network control of underactuated surface vessels with guaranteed transient performance: Theory and experimental results," *IEEE Trans. Ind. Electron.*, vol. 67, no. 5, pp. 4024–4035, 2019.
- [51] J. H. Friedman, "On bias, variance, 0/1-loss, and the curse-of-dimensionality," *Data Min. Knowl. Discov.*, vol. 1, no. 1, pp. 55–77, 1997.
- [52] M. J. Shensa *et al.*, "The discrete wavelet transform: wedding the a trous and mallat algorithms," *IEEE Trans. Signal Process.*, vol. 40, no. 10, pp. 2464–2482, 1992.
- [53] G. P. Nason and B. W. Silverman, "The stationary wavelet transform and some statistical applications," in *Wavelets and statistics*, pp. 281–299. Springer, 1995.
- [54] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," in *Int. Conf. Learn. Represent.*, 2016.
- [55] D. K. Ruch and P. J. Van Fleet, *Wavelet theory: an elementary approach with applications*. John Wiley & Sons, 2011.
- [56] A. v. d. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, "Wavenet: A generative model for raw audio," in *The 9th ISCA Speech Synthesis Workshop*, 2016.
- [57] S. Rendle, "Factorization machines," in *IEEE Int. Conf. Data Min.*, pp. 995–1000. IEEE, 2010.
- [58] X. He and T.-S. Chua, "Neural factorization machines for sparse predictive analytics," in *Proc. 40th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, pp. 355–364, 2017.
- [59] M. Blondel, A. Fujino, N. Ueda, and M. Ishihata, "Higher-order factorization machines," *Adv. Neural Inf. Process. Syst.*, vol. 29, pp. 3351–3359, 2016.
- [60] P. Hu, B. Chen, C. Zhang, and Q. Wu, "Correlation-averaging methods and kalman filter based parameter identification for a rotational inertial navigation system," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1321–1328, 2018.
- [62] J. Ding, V. Tarokh, and Y. Yang, "Model selection techniques: An overview," *IEEE Signal Process. Mag.*, vol. 35, no. 6, pp. 16–34, 2018.
- [61] P. Hu, P. Xu, B. Chen, and Q. Wu, "A self-calibration method for the installation errors of rotation axes based on the asynchronous rotation of rotational inertial navigation systems," *IEEE Trans. Ind. Electron.*, vol. 65, no. 4, pp. 3550–3558, 2017.
- [63] A. G. Baydin, R. Cornish, D. M. Rubio, M. Schmidt, and F. Wood, "Online learning rate adaptation with hypergradient descent," in *Int. Conf. Learn. Represent.*, 2018.
- [64] A. X. M. Chang and E. Culurciello, "Hardware accelerators for recurrent neural networks on fpga," in *IEEE Int. Symp. Circuits Syst.*, pp. 1–4. IEEE, 2017.
- [65] S. Li, W. Li, C. Cook, C. Zhu, and Y. Gao, "Independently recurrent neural network (indrn): Building a longer and deeper rnn," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 5457–5466, 2018.
- [66] P. Zhang, J. Xue, C. Lan, W. Zeng, Z. Gao, and N. Zheng, "Eleatt-rnn: Adding attentiveness to neurons in recurrent neural networks," *IEEE Trans. Image Process.*, vol. 29, pp. 1061–1073, 2019.
- [67] Y. Wang, Z. Gao, M. Long, J. Wang, and S. Y. Philip, "Predrn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning," in *Int. Conf. Mach. Learn.*, pp. 5123–5132. PMLR, 2018.
- [68] O. Marschall, K. Cho, and C. Savin, "A unified framework of online learning algorithms for training recurrent neural networks," *J. Mach. Learn. Res.*, 2020.

Chong Zhang is working toward the Undergraduate degree at Tsinghua University, Beijing, China, where he has been majoring in measurement, control technology and instruments.

His research interests include robot learning and control.

Peida Hu received the B.E. degree in mechanical design and automation from China University of Geosciences, Beijing, China, in 2003, the M.S. degree in mechanical design and automation from Beijing Institute of Technology, Beijing, China, in 2006, and the Ph.D. degree in instruments science and technology from Tsinghua University, Beijing, China, in 2011.

He is currently an Associate Professor with Tsinghua University. His research interests include attitude determination, attitude evaluation, and parameter identification for inertial navigation system and GNSS.

Rong Zhang received a B.E. in Mechanical Manufacturing and Automation in 1992, a MS in Precision Instrument and Mechanology in 1994, and a PhD in Precision Instrument and Mechanology from Tsinghua University in 2007.

He is currently a professor in Tsinghua University. His research interest includes MEMS inertial sensors and systems.

Qiuping Wu received the B.S., M.S., and Ph.D. degrees in precision instrumentation from Southeast University, Nanjing, China, in 1994, 1997, and 2000, respectively.

From 2000 to 2002, he was a Post-Doctoral Scholar with the Department of Precision Instrument, Tsinghua University, Beijing, China, where he is currently an Associate Professor. His research interests include inertial sensors, integrated navigation, and gravity gradiometers.

Anlan Yang received the B.E. degree in measurement, control technology and instruments from Tsinghua University, Beijing, China, in 2020. He is currently working towards the M.E. degree in Electronic and Information Engineering at Tsinghua University.

His research interest includes attitude determination and parameter identification for inertial navigation system.