Prediction of Turbulence Temporal Evolution in PANTA by Long-Short Term Memory Network

Masaomi AIZAWACARANZA, Makoto SASAKI, Hiroki MINAGAWA, Yuuki NAKAZAWA, Yoshitatsu LIU, Yuki JAJIMA, Yuichi KAWACHI[1,a)], Hiroyuki ARAKAWA[2)] and Kazuyuki HARA

College of Industrial Technology, Nihon University, Narashino 275-8575, Japan
[1)]*Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Kasuga 816-8580, Japan*
[2)]*Institute of Science and Engineering, Academic Assembly, Shimane University, Matsue 690-8504, Japan*

(Received 9 March 2022 / Accepted 23 March 2022)

Prediction of time evolution of multi-scale turbulence is performed by using Long-short term memory networks. The time series data is obtained by Langmuir probes in a linear magnetized plasma device, PANTA. The simultaneous prediction of high and low frequency components of turbulence is shown to be possible within several tens percent accuracy. The prediction accuracy depends on the initial network, which can be controlled by reducing the learning rate.

Keywords: deep learning, plasma, turbulence, data driven science, prediction, LSTM

DOI: 10.1585/pfr.17.1201048

Dynamical control of magnetized plasmas has been an active research topic in fusion research [1]. Prediction of abrupt plasma transport, such as edge localized modes (ELMs) [2] and avalanches [3], is one of critical issues. The abrupt phenomena are highly nonlinear and are often caused by micro-scale turbulence e.g. [4]. Therefore, the prediction of the multi-scale turbulence is crucial. On the other hands, in recent years, data-driven approaches, such as machine learning and advanced mode decomposition methods have shown remarkable development [1,5,6]. In this study, by using machine learning methods, temporal behaviors of turbulence are predicted, and its properties are reported.

The turbulence data analyzed in this study is obtained in a linear magnetized plasma device, PANTA [7,8]. Plasma discharges are operated by Argon gas with an RF helicon source (6 kW, 7 MHz) and with an external magnetic field of 0.13 T. Fluctuations of ion saturation current is measured by 64-channel Langmuir probes, which are aligned azimuthally at radial position of 4 cm from the center of the device. The following analysis is performed with a temporal evolution of ion saturation current obtained by a certain probe. In this discharge condition, several instabilities coexist, and their dynamical interaction determines the global system [9]: drift waves and axially symmetric mode, which could be a Kelvin-Helmholtz instability. The frequency of the drift wave is around 10 kHz, and that of the axially symmetric mode is 1 - 2 kHz, which shows the multi-scale turbulence. In this paper, as a first step toward predicting the intermittent plasma transport, both dynamics is predicted by a machine learning method.

The long-short term memory network (LSTM) is a kind of recurrent neural networks (RNNs), which have loop structures so as to store time series of data. The LSTM is an extension of the RNN to overcome an issue of parameter searching, called vanishing gradient problem [10]. Because of this, the LSTM can treat much longer time series of data compared to the conventional RNN. In this study, the LSTM is selected to predict the multi-scale turbulence, which contains long-period fluctuations as well as short-period ones. The LSTM Network consists of multiple units called LSTM Blocks in the hidden layer, which contain cells that store past information and three gate units (input gate units, forget gate units, and output gate units) that control the flow of information. The initial conditions of the network are set to 200 units and the learning rate of 0.005, which is a hyperparameter to tune the network [10]. The time series of 1600 points data, which corresponds to that for 16 ms, is used. The first 800 points data (8 ms) is used for the training and the network is established. Then, the next 800 points data (8 ms) is predicted by using the built LSTM blocks. By changing the initial network 50 times, and the prediction properties are statistically evaluated.

The predicted evolution by the LSTM is shown in Fig. 1. Top panel of Fig. 1 shows that the predicted data agrees well with the observed data for a few milliseconds in the beginning of the prediction. The dominant fluctuation is associated with the drift waves (high-frequency component), and the modulation of the fluctuation corresponds to the axially symmetric mode (low-frequency mode). The deviation of the prediction from the obser-

author's e-mail: sasaki.makoto@nihon-u.ac.jp
a) Current affiliation: National Institute for Fusion Science, National Institutes of Natural Sciences, Toki 509-5292, Japan

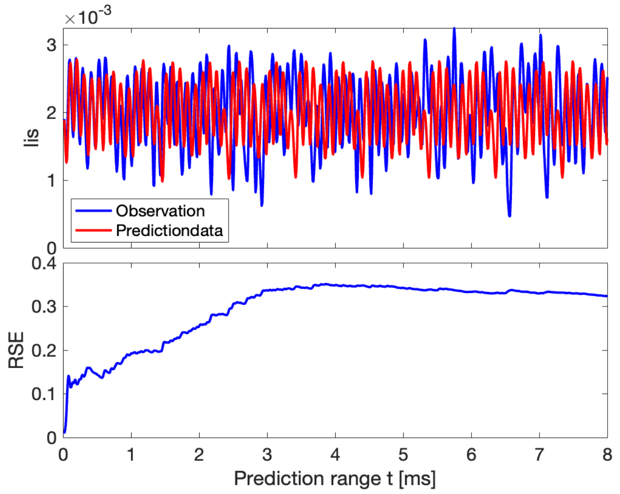

Fig. 1 Comparison between prediction and observation. Top panel shows the temporal evolutions of observation (blue) and prediction (red). Bottom panel illustrates the relative squared error defined by Eq. (1).

vation is evaluated by the relative squared error, defined as

$$\varepsilon(t) = \sqrt{\frac{\sum_{j=1}^{t} |y_{o,t} - y_{p,t}|^2}{\sum_{j=1}^{t} |y_{o,t}|^2}}, \quad (1)$$

where $y_{o,t}$ and $y_{p,t}$ denote the observation and prediction at time t, respectively. The error is shown in the bottom panel of Fig. 1. The prediction error increases in time with a typical time scale of millisecond, which is a typical period of the low frequency mode. The phase of the predicted temporal data for high-frequency component gradually deviates from the observation, which causes the increase of the prediction error.

To clarify the properties of the prediction by the LSTM, spectrum of each temporal data and the cross-spectrum between the time series of observed data and of the prediction is evaluated, which are shown in Fig. 2. In this analysis, the data at the beginning of the prediction, t < 2 ms, is used. Here, the light blue lines correspond to the predicted data with different initial networks, and the red solid line shows the statistical average of the predictions. As shown in the top panel of Fig. 2, the magnitude of the spectrum peaks with 10 kHz and 2 kHz including the side-bands of 8 kHz and 12 kHz of the observation and prediction agree well. The cross-spectrum, shown in the bottom panel of Fig. 2, indicates the high correlation at less than 2 kHz, and at around 10 kHz. Here, it is noted that the dependence of the prediction on the initial network condition is around few tens percent for the high-frequency component. It is found that both the low and high frequency components can be predicted simultaneously by the LSTM at least for a few milliseconds. It is noted that the prediction of the time evolution is possible within a similar accuracy for the different discharge data with similar turbulence, where the high frequency fluctuation is modulated by low frequency component.

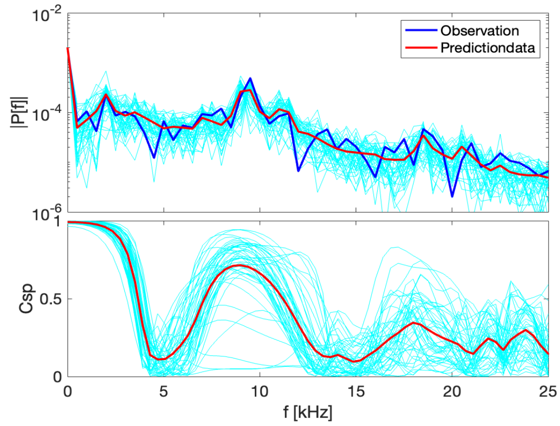

Fig. 2 Comparison between spectrum of prediction and observation. Top panel shows frequency spectrum of predictions average (red) and observation (blue). Bottom panel shows the cross-spectrum.

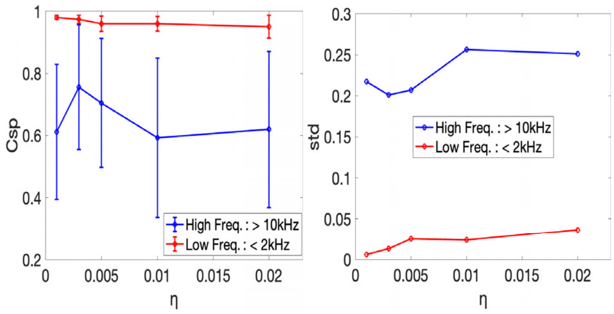

Fig. 3 Cross-Spectrum values and standard deviation for varying Learning rate. Left panel shows high frequency (blue) and low frequency (red) error bars of cross spectrum values. Right Panel illustrates standard deviation of left panel's error bars.

The variations of the prediction, dependent on the initial network, can be controlled by the learning rate η. By scanning the learning rate, the change of the values of the cross correlations at 2 kHz and 10 kHz are plotted in the left panel of Fig. 3, where the error bars are evaluated from the standard deviation of the results with 50 different initial conditions. The error bars are an index of the prediction variations. The prediction variation can be controlled by changing the learning rate, which is illustrated in the right panel of Fig. 3. Especially for the low-frequency component, the prediction variation becomes small with a smaller learning rate.

In conclusion, the prediction of the temporal evolution of the multi-scale turbulence, observed at a single spatial position, is performed by using the LSTM. The simultaneous prediction for the low and high frequency component is shown to be possible for several periods of the low frequency mode. In addition, the variations of the prediction, dependent on the initial network, can be suppressed with a small learning rate.

Acknowledgments

The authors acknowledge valuable discussions with Prof. S. Inagaki. This work was partly supported by JSPS KAKENHI Grant Numbers of JP21K03509, JP21K03513, the collaboration programs of NIFS (NIFS21KNST194, NIFS21KLPH049), the collaborative research program of RIAM Kyushu University (2021 S3-2) and Nihon University CIT fund for young Researchers.

[1] J.K. Harbeck *et al.*, Nature **568**, 526 (2019).
[2] A.W. Leonard *et al.*, Phys. Plasmas **21**, 090501 (2014).
[3] B.A. Carreras *et al.*, Phys. Rev. Lett. **80**, 4438 (1998).
[4] D.F. Kong *et al.*, Nucl. Fusion **58**, 034003 (2018).
[5] M. Sasaki *et al.*, Plasma Phys. Control. Fusion **61**, 112001 (2019).
[6] M. Sasaki *et al.*, Plasma Phys. Control. Fusion **63**, 025004 (2021).
[7] S. Inagaki *et al.*, Sci. Rep. **6**, 22189 (2016).
[8] F. Kin, Phys. Plasmas **26**, 042306 (2019).
[9] Y. Kawachi *et al.*, Phys. Plasmas **28**, 112302 (2021).
[10] S. Hochreiter and J. Schmidhuber, Neural Computation **9** (8), 1735 (1997).