ISSN 1674-1056
http://cpb.iphy.ac.cn
http://iopscience.iop.org/cpb
中国物理 B
Chinese Physics
中国物理 Chinese Physics B

Estimation of plasma equilibrium parameters via a neural network approach

Zi-Jian Zhu(朱子健)[1,2], Yong Guo(郭勇)[2], Fei Yang(杨飞)[3], Bing-Jia Xiao(肖炳甲)[1,2], Jian-Gang Li(李建刚)[1,2]
Citation: Chin. Phys. B . 2019, 28(12): 125204 . **doi:** 10.1088/1674-1056/ab55d1
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Influence analysis of symmetry on capsule in six-cylinder-port hohlraum

You Zou(邹游), Wudi Zheng(郑无敌), Xin Li(李欣)
Chin. Phys. B . 2019, 28(3): 035203 . **doi:** 10.1088/1674-1056/28/3/035203

Rayleigh-Taylor instability at spherical interfaces of incompressible fluids

Hong-Yu Guo(郭宏宇), Li-Feng Wang(王立锋), Wen-Hua Ye(叶文华), Jun-Feng Wu(吴俊峰), Ying-Jun Li(李英骏), Wei-Yan Zhang(张维岩)
Chin. Phys. B . 2018, 27(2): 025206 . **doi:** 10.1088/1674-1056/27/2/025206

Rayleigh-Taylor instability of multi-fluid layers in cylindrical geometry

Hong-Yu Guo(郭宏宇), Li-Feng Wang(王立锋), Wen-Hua Ye(叶文华), Jun-Feng Wu(吴俊峰), Wei-Yan Zhang(张维岩)
Chin. Phys. B . 2017, 26(12): 125202 . **doi:** 10.1088/1674-1056/26/12/125202

Magneto-Rayleigh–Taylor instability in compressible Z-pinch liner plasmas

Xue Yang(杨学), De-Long Xiao(肖德龙), Ning Ding(丁宁), Jie Liu(刘杰)
Chin. Phys. B . 2017, 26(7): 075202 . **doi:** 10.1088/1674-1056/26/7/075202

Surface diffuse discharge mechanism of well-aligned atmospheric pressure microplasma arrays

Ren-Wu Zhou(周仁武), Ru-Sen Zhou(周儒森), Jin-Xing Zhuang(庄金星), Jiang-Wei Li(李江炜), Mao-Dong Chen(陈茂冬), Xian-Hui Zhang(张先徽), Dong-Ping Liu(刘东平), Kostya (Ken) Ostrikov, Si-Ze Yang(杨思泽)
Chin. Phys. B . 2016, 25(4): 045202 . **doi:** 10.1088/1674-1056/25/4/045202

Estimation of plasma equilibrium parameters via a neural network approach*

Zi-Jian Zhu(朱子健)[1,2], Yong Guo(郭勇)[2], Fei Yang(杨飞)[3,†], Bing-Jia Xiao(肖炳甲)[1,2], and Jian-Gang Li(李建刚)[1,2]

[1] *Department of Engineering and Applied Physics, University of Science and Technology of China, Hefei* 230026, *China*
[2] *Institute of Plasma Physics, Chinese Academy of Sciences, Hefei* 230031, *China*
[3] *Department of Medical Information Engineering, Anhui Medical University, Hefei* 230026, *China*

(Received 25 August 2019; revised manuscript received 22 October 2019; published online 29 November 2019)

Plasma equilibrium parameters such as position, X-point, internal inductance, and poloidal beta are essential information for efficient and safe operation of tokamak. In this work, the artificial neural network is used to establish a non-linear relationship between the measured diagnostic signals and selected equilibrium parameters. The estimation process is split into a preliminary classification of the kind of equilibrium (limiter or divertor) and subsequent inference of the equilibrium parameters. The training and testing datasets are generated by the tokamak simulation code (TSC), which has been benchmarked with the EAST experimental data. The noise immunity of the inference model is tested. Adding noise to model inputs during training process is proved to have a certain ability for maintaining performance.

Keywords: neural network, fusion, plasma equilibrium, noise immunity

PACS: 52.55.Fa, 07.05.Mh, 52.70.–m, 52.25.Xz **DOI:** 10.1088/1674-1056/ab55d1

1. Introduction

The tokamak device, which uses powerful magnetic fields to confine the high-temperature plasma, is one of several types of magnetic confinement facilities being developed to produce controlled thermonuclear fusion power. Estimation of fusion plasma equilibrium parameters from many diagnostic signals is essential for efficient and safe operation of tokamak. To extract the plasma equilibrium parameters such as position, triangularity, internal inductance, and poloidal beta, several schemes have been proposed. The originally used schemes are the first-principles approaches based on a multifilament model or a control surface model for plasma.[1,2] The currents flowing in the filaments or the current density distribution on the control surface are chosen in such a way as to best fit the measured fields and fluxes. In the development afterward, the plasma is also schematized by means of a set of finite element shape functions.[3] It has an advantage over the multifilament model since the field singularities are automatically eliminated. However, the calculation amount is also increased. Via using the graphic processing unit (GPU), the fast equilibrium reconstruction has been implemented in Experimental and Advanced Superconducting Tokamak (EAST).[4]

Recently, a new statistical approach is proposed to make use of a database that contains a large number of equilibria of the plasma and the corresponding measurements. Via applying this dataset statistical method, it is possible to find a function relating the plasma parameters to the external measurements.[5,6] As one of the big-data-driven statistical methods, the artificial neural network (ANN) has shown its effectiveness in many applications in fusion research. The network estimation approach has no iteration process during prediction once the model has been trained well. We do not have to concern about the computational burden for real-time processing. And the need for the physical interpretation of the measurements will probably be much less important in future reactors.[7] Under this background, a typical neural network approach has been introduced to establish the required nonlinear mapping between the measured diagnostic signals and a finite set of parameters describing the plasma equilibrium. The technique has been successfully applied to single-null diverted discharge equilibria of DIII-D tokamak after having checked its practicability in the simple test case of a circular plasma.[8] The aim of the present paper is to take a step forward in the application of the neural network approach for the identification of limiter and divertor plasma equilibria. In the remainder of this paper, we investigate the applicability of the new technique. In Section 2, the method of the artificial neural network is introduced, and the database generation procedure is briefly described. In Section 3, the details of the NN-based plasma equilibrium parameter estimation are depicted. And the results analysis and discussion are given in Section 4. In the end, summary and conclusions are drawn in Section 5.

*Project supported by the National Magnetic Confinement Fusion Energy R&D Program of China (Grant No. 2018YFE0302100), the National Key Research and Development Program of China (Grant Nos. 2017YFE0300500 and 2017YFE0300501), the National Natural Science Foundation of China (Grant Nos. 11575245, 11805236, and 11905256), and Young and Middle-aged Academic Back-bone Finance Fund from Anhui Medical University.

†Corresponding author. E-mail: yangfei@ahmu.edu.cn

© 2019 Chinese Physical Society and IOP Publishing Ltd

2. Method and database

2.1. Neural network method for classification and regression

The original artificial neural network comes from the perceptron model.[9] An artificial neural network is composed of activated functions of neurons, the network topology, connection weights, and the threshold of neurons. Generally speaking, when the network topology is fixed, the output is affected by changes of the connection weights. The learning procedure is to minimize the cost function by changing the weights and bias. With the help of the error back-propagation (EBP) algorithm,[10] the artificial neural network can be more and more widely used to solve classification and regression problems. Most of characteristics of classification and regression neural networks are similar except for the output layer. The activation function for the output layer would be chosen according to the type of task. In this work, we divide the equilibria into two classes (i.e., limiter and divertor). For this binary classification problem, the logistic-sigmoid function can be used with binomial cross-entropy as the cost function. And then, for the regression problem (i.e., real-value inference), the linear function is used in the output layer and the regularly used cost function is the mean squared error (MSE). One of the regression neural networks used in our study is presented in Fig. 1, with 88 inputs, 20 neurons in the hidden layer, and 10 neurons in the output layer. Another regression neural network is designed with 88 inputs, 20 neurons in the hidden layer, and 7 neurons in the output layer. The classification neural network has the same inputs and hidden layer while it has one neuron in the output layer.

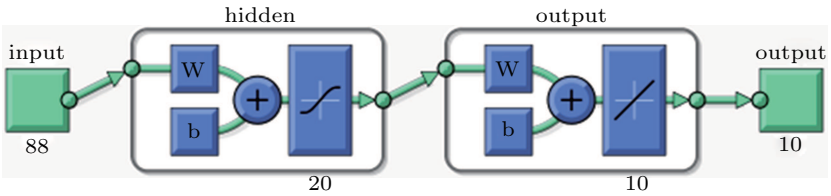


Fig. 1. Topology of a neural network for regression task.

2.2. Generation of training and testing database

EAST is one of the key fusion research projects in China. The datasets used during the neural network training and testing phase come from EAST discharge simulations using the tokamak simulation code (TSC). The TSC studies the evolution of magnetic field in a rectangular computational domain using the Maxwell MHD equations for the plasma, coupled with the boundary conditions to the circuit equations for the poloidal field (PF) coils.[11,12] It could generate evolution data during the whole discharge process at a time interval of 1 millisecond. The high time density and stability of the data make it suitable for incipient neural network model training. The geometry of EAST is shown in Fig. 2. The diagnostic signals including magnetic probes, flux loops, poloidal field coil (PF-coil) currents, and plasma current are used as possible input parameters in the neural network model. This kind of input set has been proved sufficient for plasma equilibrium reconstruction when using the EFIT method, thus we choose similar signals as the neural network inputs. As present in Table 1, ten plasma equilibrium parameters are selected as the possible outputs. The scope of the plasma shape and position has been chosen so as to cover the equilibrium parameter range experienced during most normal discharges. In our case, the total number of computed equilibria is 16466. Ninety percent of the whole dataset is used during the classification network training process. We use 4980 limiter equilibria and 6384 lower X-point divertor equilibria for parameter inference model training respectively. One-tenth of equilibria is reserved for testing. It should be noted that the neural network model has its own suitable field. For the new equilibrium which is out of the range of training datasets, we need to train a new neural network model or adjust the normalization parameters. For the key inference model training and testing, the composition and subdivision of the dataset of the simulated equilibria are denoted in Table 2.

Table 1. Output parameter ranges.

Dimension outputs	Range/m
Z of magnetic axis (Z_m)	−0.077 to 0.096
R of magnetic axis (R_m)	1.694 to 1.948
Major radius (R_0)	1.685 to 1.930
Minor radius (a)	0.321 to 0.507
R of X-point (R_X)	1.604 to 1.668
Z of X-point (Z_X)	−0.899 to −0.745
Dimensionless outputs	**Range**
Safety factor (q_0)	0.91 to 4.95
Poloidal beta (β_p)	0.01 to 0.73
Internal inductance (l_i)	0.30 to 0.83
Elongation (κ)	0.88 to 1.86

Table 2. The composition and subdivision of the dataset for regression task.

	Total	Training	Testing
Limiter equilibria	4980	4482	498
Divertor equilibria	6384	5746	638

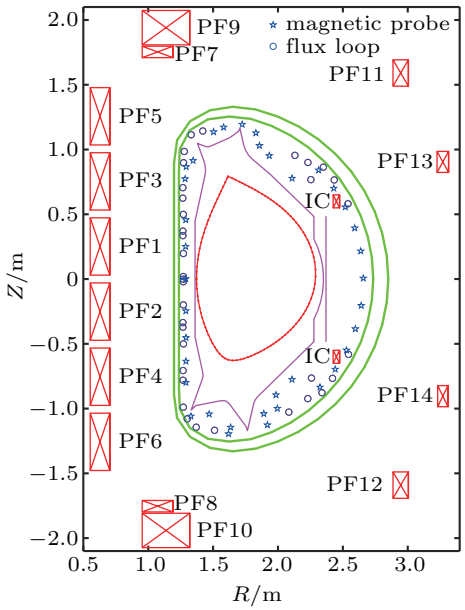


Fig. 2. EAST geometry and distribution of electromagnetic diagnosis.

3. Plasma equilibrium parameters estimation

As already pointed out, the estimation process is split into a preliminary classification of the kind of equilibrium (limiter or divertor) and subsequent inference of the equilibrium parameters. Both the classification network and the regression network have the same input parameters, i.e., the above mentioned 38 magnetic probes, 35 flux loops, 14 PF-coil currents, and plasma current. A single hidden layer is proved sufficient for these two tasks. We designed 20 neurons for the hidden layer, and the output layer depends on different tasks.

In the first step, the classification network shows perfect ability to distinguish between the limiter and the divertor configurations on the basis of electromagnetic measurements. The final train performance valued by the binomial cross-entropy function is up to 1.09×10^{-6}. The classification network should give an output equal either to zero or one for limiter or divertor configurations, respectively. In the test dataset, we round the elements of outputs to the nearest integer. All the 498 limiter configurations and 1148 divertor configurations are classified correctly. The results achieved are extremely satisfactory. It has been discussed in Ref. [7] that the ability of ANN to form correct decision regions between the two classes is probably mainly related to the flux loops located near the limiter. Someone can try to use a much smaller number of measurements if there is a demand.

After that, for the subsequent step of inferring the plasma equilibrium parameters, we trained the regression neural network respectively for the limiter and divertor configurations. The R and Z coordinates of the magnetic axis, internal inductance, poloidal beta, major and minor radii, and elongation are chosen as outputs of the limiter plasma parameters inference, and for the divertor parameters inference, we added the safety factor at the magnetic axis, R and Z coordinates of the X point into the model outputs. Table 3 summarizes the global performance of the inference model in terms of MSE, and we list the inference accuracies achieved for the individual plasma parameters in Tables 4 and 5 for the limiter and the divertor cases, respectively. The mean absolute errors between the prediction values and target values are less than 1 mm or 1% for the dimensional and dimensionless parameters, respectively. The max absolute error and relative error also show good inference performance of the regression neural network model. The inference performance of the limiter case is a little bit worse than that of the divertor case, which is probably caused by the stronger non-linear relationship in the plasma ramp-up phase. In these models, all the 88 available electromagnetic measurements are used as the inputs. The prediction parameters are compared with the target values using test datasets.

Table 3. Summary of the global performance of the inference model in terms of MSE.

Total mean square error	Training	Testing
Regression NN (divertor)	2.30×10^{-6}	2.20×10^{-6}
Regression NN (limiter)	1.49×10^{-5}	1.65×10^{-5}

Table 4. The inference accuracies achieved for the individual plasma parameters for the divertor case.

Equilibrium parameters (regression NN outputs)	Mean absolute error/%	Max absolute error/%	Relative error/%
Z of magnetic axis (Z_m)/m	0.04	0.54	0.21
R of magnetic axis (R_m)/m	0.03	0.63	0.12
Safety factor (q_0)	0.15	2.30	0.04
Poloidal beta (β_p)	0.10	1.41	0.14
Internal inductance (l_i)	0.08	0.68	0.14
Major radius (R_0)/m	0.03	0.51	0.13
Minor radius (a)/m	0.05	0.32	0.27
Elongation (κ)	0.19	2.31	0.19
R of X-point (R_X)/m	0.05	0.46	0.70
Z of X-point (Z_X)/m	0.04	0.41	0.26

Table 5. The inference accuracies achieved for the individual plasma parameters for the limiter case.

Equilibrium parameters (regression NN outputs)	Mean absolute error/%	Max absolute error/%	Relative error/%
Z of magnetic axis (Z_m)/m	0.13	0.82	0.73
R of magnetic axis (R_m)/m	0.10	0.66	0.39
Poloidal beta (β_p)	0.53	4.37	0.74
Internal inductance (l_i)	0.37	3.27	0.69
Major radius (R_0)/m	0.09	1.07	0.37
Minor radius (a)/m	0.1	1.91	0.53
Elongation (κ)	0.13	0.96	0.13

Figure 3 shows the comparison of divertor plasma equilibrium parameters and their deviations. The limiter inference result is drawn in Fig. 4. For all the inference parameters, the target and prediction values coincide well and the errors are pretty small. Basically, for both the limiter and divertor plasmas, a generally good estimation for the plasma equilibrium parameters can be obtained via the neural network model. Reference [7] investigated the inference model with reduced inputs. The inference performance was found a little decreased but still acceptable when using only half sensors as the inputs. We can also choose fewer measurements for reduced model training according to actual accuracy requirements and available diagnostic signals. Here, we just show the best performance with sufficient inputs. The robustness to noise and possible faults is analyzed and discussed in the next section using the divertor equilibrium parameter inference model with sufficient inputs.

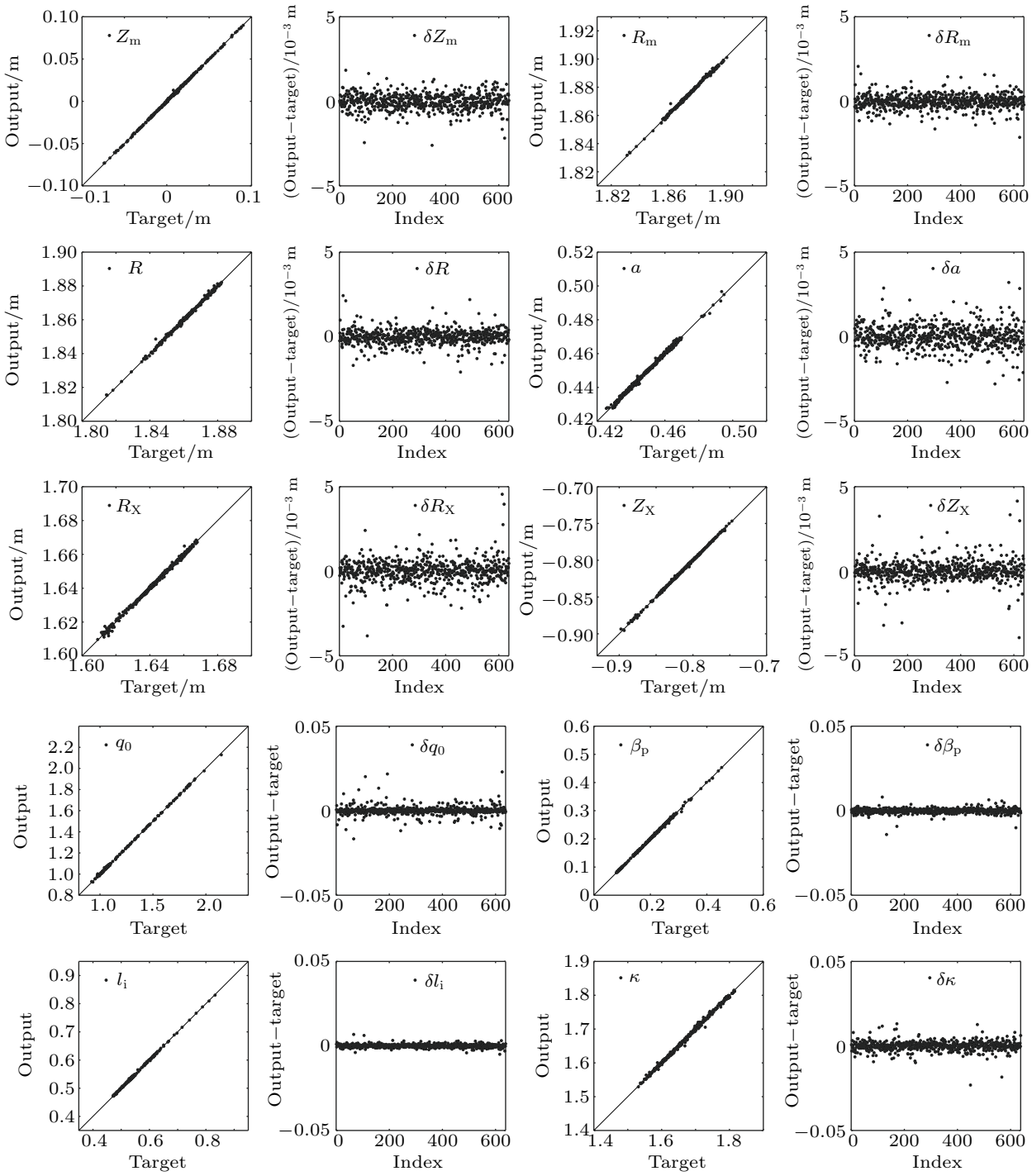


Fig. 3. Divertor prediction values and target values, and their deviations.

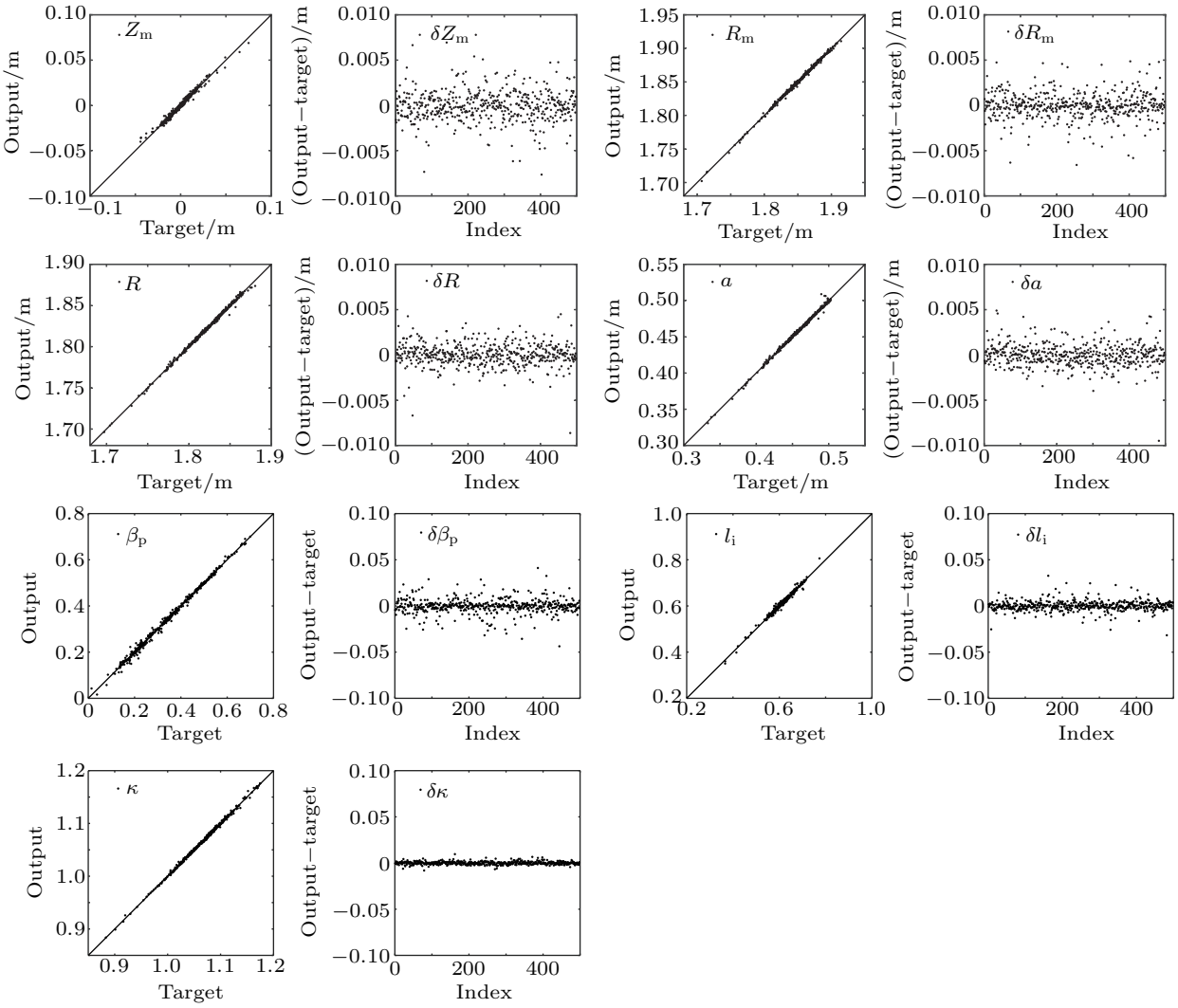


Fig. 4. Limiter prediction values and target values, and their deviations.

4. Results analysis and discussion

In real experiments, the electromagnetic measurements, which are used as the model inputs, may contain noise or faults. In order to assess the robustness of the estimation, we have also attempted to analyze the behavior of the neural network model when the input values are affected by possible faults and measurement noise. The assumed experimental noise level for the EAST electromagnetic measurements is less than 3%. We have trained and tested the regression neural network model with different noise amplitudes from 0 to 3%. As shown in Fig. 5(a), the color bar presents the logarithm value of test MSE performance of a trained model. The top line shows the performance degradation of the noiseless-trained model. We can discover that the model trained with no input noise cannot maintain its performance when there is noise in the test input. In the red frame, when the models are trained with input noise from 0.1% to 0.7%, the noise immunity of these models is improved and keeps the logarithm value of total test MSE performance still smaller than −5. It is the best performance region of the regression neural network model. Adding noise to model training inputs also has a negative influence on the model performance. We can see in Fig. 5(b), the blue line shows the noiseless-test performance degradation when the model is trained with noise from 0 to 3%. The logarithm value of models' noise-test performance is decreased as the noise in the training inputs is increased. Meanwhile, the yellow line shows the mean values of the test results with 3% noise and the red line denotes the test results with noise from 0.1% to 3%, which indicates that the noise immunity of the model is improved. We need to make a trade between the model total performance and its noise immunity according to the actual application. In Fig. 6, the inference values of R and Z coordinates of the magnetic axis (we only choose a part of divertor equilibrium parameters due to limited space) are compared with the target values. The models, which are trained with noiseless and 0.5% noised inputs, are both tested with 0.5% noised inputs. The red symbols present the good prediction of the noise-trained model while the bad prediction of the noiseless-trained model is denoted with black symbols.

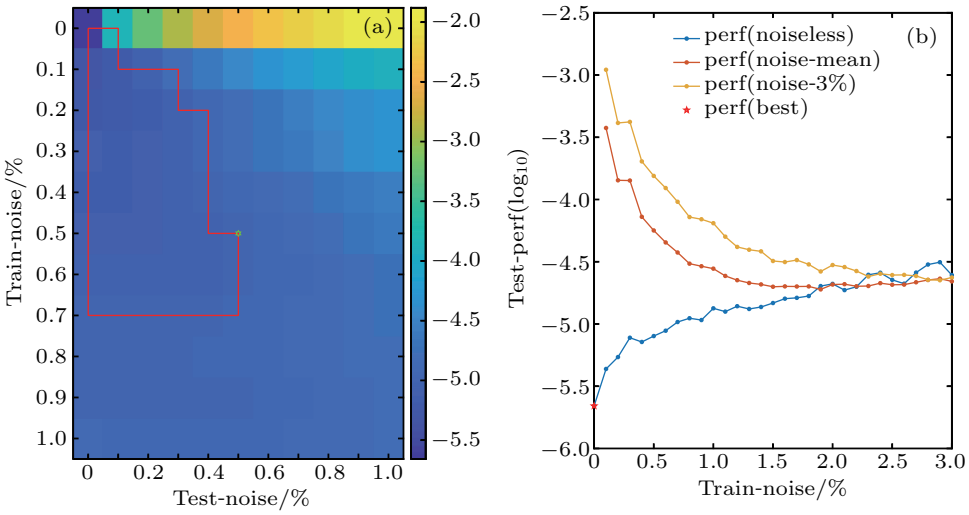


Fig. 5. Test performance comparison for the model trained with different noised inputs.

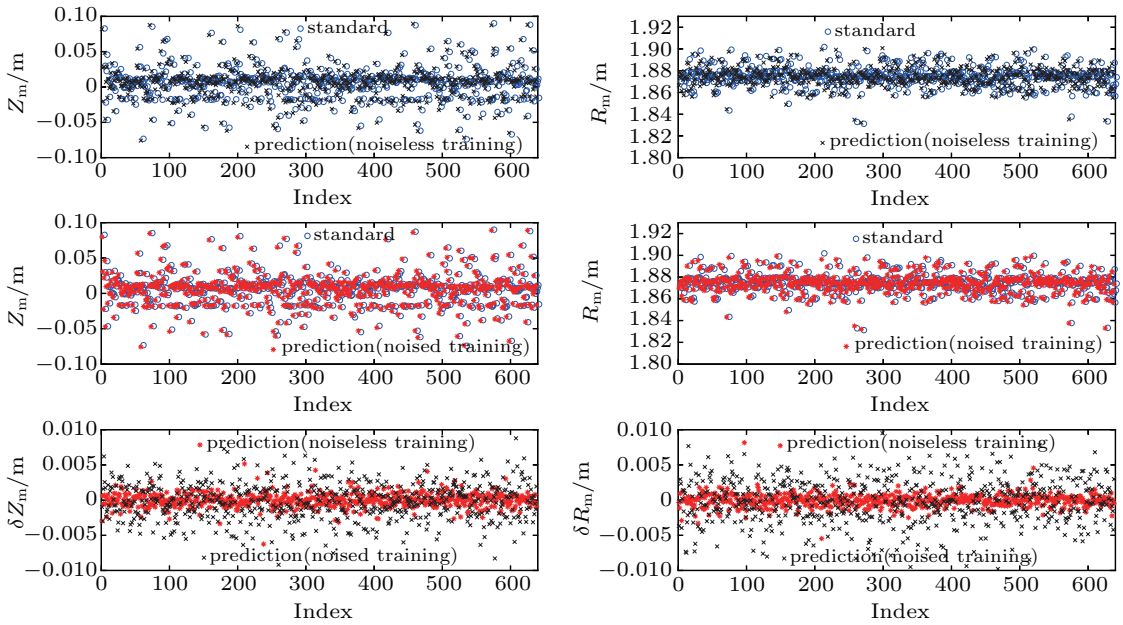


Fig. 6. Divertor parameter inference; prediction values and target values of R and Z coordinates of the magnetic axis; noiseless/0.5% noised training, test with noised inputs.

Table 6. Random faults and noised training, final model performance.

Total mean square error, regression NN (divertor, 3 random faults, 0.5% input noise)		Training 1.04×10^{-5}	Testing 1.16×10^{-5}
Equilibrium parameters	Mean absolute error/%	Max absolute error/%	Relative error/%
Z of magnetic axis (Z_m)/m	0.07	0.46	0.39
R of magnetic axis (R_m)/m	0.06	0.54	0.23
Safety factor (q_0)	0.30	3.74	0.07
Poloidal beta (β_p)	0.25	1.54	0.35
Internal inductance (l_i)	0.16	0.80	0.31
Major radius (R_0)/m	0.06	0.50	0.24
Minor radius (a)/m	0.07	0.48	0.39
Elongation (κ)	0.61	3.35	0.62
R of X-point (R_X)/m	0.06	0.37	0.99
Z of X-point (Z_X)/m	0.06	1.05	0.39

We have also trained a model with three measurements absent while 0.5% percent noise in training inputs and testing inputs is still assumed (green star in Fig. 5(a)). The three absent measurements are randomly chosen in the magnetic probes, flux loops, and PF-coil currents, respectively. It is a primary test for possible faults in measurements. We retrained

the model, and the final model performance is tested with input data which has noise and faults. Its global performance and the inference accuracies achieved for the individual plasma parameters are shown in Table 6. The mean absolute error of the dimension parameter is less than 1 mm while the mean absolute error of dimensionless parameter is less than 1%. The max absolute error and relative error also are satisfactory small.

5. Summary and conclusions

In the end, the speed of the inference model based on neural network is discussed as it aims to control application. After the model training has been finished, the topology and parameters of the neural network are settled. For plasma equilibrium parameters inference with presupposed inputs, there are several algebraic calculation processes such as normalization, matrix multiplication, hyperbolic tangent function, and anti-normalization. The total inference time of the neural network model is up to an order of 1 millisecond when running it on an ordinary laptop. Neural network estimation without iteration process is faster than EFIT and some other iterative solvers when they are implemented by traditional CPU computer. With the improvement of computer performance, the elapsed time can be reduced. For the EAST plasma control system, the experimental requirement for shape control is about 1 ms and the neural network model can be promisingly used for it. In case much shorter execution time is required, someone can use a parallel architecture or direct hardware for the implementation of the neural network model.

In this paper, the classification neural network and the regression neural network are both used for estimation of plasma equilibrium parameters. We split the estimation process into a preliminary equilibrium classification and a subsequent inference of the equilibrium parameters. The classification network correctly classifies all the divertor and limiter equilibria in the test dataset. For the parameter inference task, the regression network also shows good performance even with measurement noise and possible faults. The model trained with 0.5% noise is chosen after we trade the global inference performance of the model and its noise immunity. The possible faults in measurements are also tested in the end. The neural network model is proved to be robust with a small number of measurements absent. In conclusion, the method looks very promising especially for the real-time estimation of the plasma equilibrium parameters.

Acknowledgments

The authors are grateful to Dr. Y Huang and Dr. Z Luo for useful suggestions, patient guidance, and technical assistance. There is also much gratitude to the EAST team for having made available detailed technical information about the machine geometry and on relevant studies on the related dataset of computed equilibria.

References

[1] Feneberg W, Lackner K and Martin P 1984 *Comput. Phys. Commun.* **31** 143
[2] Swain D and Neilson G 1982 *Nucl. Fusion* **22** 1015
[3] Hofmann F and Tonetti G 1988 *Nucl. Fusion* **28** 519
[4] Huang Y, Xiao B J and Luo Z P 2017 *Chin. Phys. B* **26** 085204
[5] Calcagno S, Greco A, Morabito F C and Versaci M 2006 *The 2006 IEEE International Joint Conference on Neural Network Proceedings* Vancouver, BC, pp. 835–842
[6] Matsukawa M, Hosogane N and Ninomiya H 1992 *Plasma. Phys. Contr. F* **34** 907
[7] Coccorese E, Morabito C and Martone R 1994 *Nucl. Fusion* **34** 1349
[8] Lister J B and Schnurrenberger H 1991 *Nucl. Fusion* **31** 1291
[9] Rosenblatt F 1958 *Psychol. Rev.* **65** 386
[10] Rumelhart D E, Hinton G E and Williams R J 1986 *Nature* **323** 533
[11] Guo Y, Xiao B, Wu B and Liu C 2012 *Plasma. Phys. Contr. F* **54** 085022
[12] Guo Y, Xiao B J, Liu L, Yang F, Wang Y H and Qiu Q L 2016 *Chin. Phys. B* **25** 115201