OPEN ACCESS
IOP Publishing | International Atomic Energy Agency
Nuclear Fusion
Nucl. Fusion **63** (2023) 086020 (9pp)
https://doi.org/10.1088/1741-4326/acdf00

Reconstruction of tokamak plasma safety factor profile using deep learning

Xishuo Wei[1,5], **Shuying Sun**[2,5], **William Tang**[3], **Zhihong Lin**[1], **Hongfei Du**[4] and **Ge Dong**[3,4,5,*]

[1] University of California, Irvine, CA, United States of America
[2] Fusion Simulation Center, Peking University, Beijing 100871, China
[3] Princeton Plasma Physics Laboratory, Princeton, NJ, United States of America
[4] Energy Singularity, Shanghai, China

E-mail: dongge@energysingularity.cn

Received 3 February 2023, revised 24 May 2023
Accepted for publication 16 June 2023
Published 28 June 2023

Abstract

The motional Stark effect (MSE) diagnostic has been a standard measurement for the magnetic field line pitch angle in tokamaks that are equipped with neutral beams. However, the MSE data are not always available due to experimental constraints, especially in future devices without neutral beams. Here we develop a deep-learning based model (SGTC-QR) that can reconstruct the safety factor profile without the MSE diagnostic to mimic the traditional equilibrium reconstruction with the MSE constraint. The model demonstrates promising performance, and the sub-millisecond inference time is compatible with the real-time plasma control system.

Keywords: deep learning, tokamak equilibrium reconstruction, safety profile, artificial intelligence

(Some figures may appear in colour only in the online journal)

1. Introduction

In tokamak operations, accurate equilibrium reconstruction is essential for reliable real-time control and realistic post-shot instability analysis. The safety factor (q) profile, which defines the magnetic field line pitch angle, is an important quantity to be calculated in equilibrium reconstruction. The popular reconstruction method, for example, that is used in the equilibrium fitting (EFIT) code [1], uses the iterative algorithm to minimize the total error between computed and measured values of some key parameters such as the magnetic field line pitch angle [2]. The final reconstruction and the subsequent physical analysis can depend sensitively on the capability of measurements. EFIT traditionally used data from external magnetic measurements. For increased accuracy, internal measurements such as the magnetic pitch angle from the motional Stark effect (MSE) [3] can be used to dramatically improve the accuracy of the equilibrium reconstructions. In this paper, we have followed the naming convention at DIII-D National Fusion Facility for convenience, where the magnetics-only EFIT reconstructions is called EFIT01 and the magnetics + MSE EFIT reconstructions is called EFIT02. The equilibria reconstructed by the EFIT01 and EFIT02 can lead to different predictions for magnetohydrodynamics (MHD) instability for the same discharge. For example, figure 1 shows a comparison of the gyrokinetic toroidal code (GTC) [4, 5] simulation results of internal kink mode using the equilibria reconstructed by the EFIT01 and EFIT02, as compared in figures 1(*a*) and (*b*), respectively, for the same shot (#168973) at an identical time (4676 ms). Using the same set of simulation control parameters (such as the time step size and grid number), only equilibrium from the EFIT02 reconstruction with the MSE constraint leads to unstable current driven instability, which is more consistent with the low frequency long-lived modes observed from the Mirnov coil

[5] These authors contributed equally to this work.
[*] Author to whom any correspondence should be addressed.

measurements, as shown in figure 1(*g*). The reason for this distinct difference in the current driven kink instability is that the q profile from the EFIT02 is less than 1 near the axis, while the minimum q value from the EFIT01 is larger than 1 as shown in figure 1(*a*). This example demonstrates that the reconstruction from the EFIT02 (or other EFIT versions using more constraints) is preferable to the EFIT01 for an accurate physics analysis. A limitation of the EFIT02 is that the reconstruction relies heavily on the measurement of the magnetic pitch angle and radial electric field using the MSE diagnostic, which relies on specific reference neutral beam settings [6, 7]. There is a significant portion of DIII-D shots where the MSE diagnostic is unavailable, or only available for limited periods of time during operation for the real-time EFIT ('rtEFIT') [8] equilibrium reconstruction. Future fusion pilot plants might need neutral beams solely for diagnostic purposes, bringing additional unnecessary costs [9]. To address this long-standing problem, it is desirable to build a model which uses available diagnostic signals to reconstruct the q profile as accurately as that in EFIT02.

The desired model should have three features. Firstly, the model prediction should be accurate and performs better than EFIT01. Secondly, it should be capable of using as many real-time diagnostic signals as possible. We notice that many signals may correlate to the q-profile, but there is no easy way to express explicitly any single signal (or a simple combination of a group of signals) as a constraint of the Grad–Shafranov solver. For example, the plasma rotation profile combined with distinguishing features of current-driven instabilities and Alfven Avalanches from the spectrogram of magnetic signals can be correlated with constraints in the q profile, such as the minimum q and the radial location of the $q = 1$ surface. Finally, the model prediction should be fast. There have been many studies aimed at fast equilibrium reconstruction, but it is generally hard to achieve computational speed and the integration of physics-based information for improving accuracy at the same time. Multiple acceleration methods have been implemented in the EFIT framework, i.e. rtEFIT [8], P-EFIT [10], and the machine learning model EFIT-AI [11]. The time scale of the q-profile reconstruction should also be on the order of or shorter than 1 ms to be compatible with the plasma control system (PCS). Traditional methods to improve the accuracy of q-profile reconstruction like the multi-step EFIT reconstruction based on measured plasma instabilities such as that described in [12] require time-costly manual investigations for post-shot analysis. In this work, we have developed a deep-learning based model, which serves as the q-profile reconstruction module of the surrogate GTC framework (SGTC-QR), to recover the q-profile reconstructed with the MSE constraint (EFIT02). We demonstrate that the SGTC-QR result is better than the traditional method when MSE is not considered as a constraint (EFIT01). We should also point out that the 'kinetic-EFIT' that uses kinetic profiles as constraints provides more accurate equilibrium including q-profile [2, 13, 14]. In this paper, we demonstrate the q-profile reconstruction workflow using the EFIT01 and EFIT02 data, but it is also interesting to study the machine learning algorithm to reproduce the kinetic-EFIT q-profile in the future.

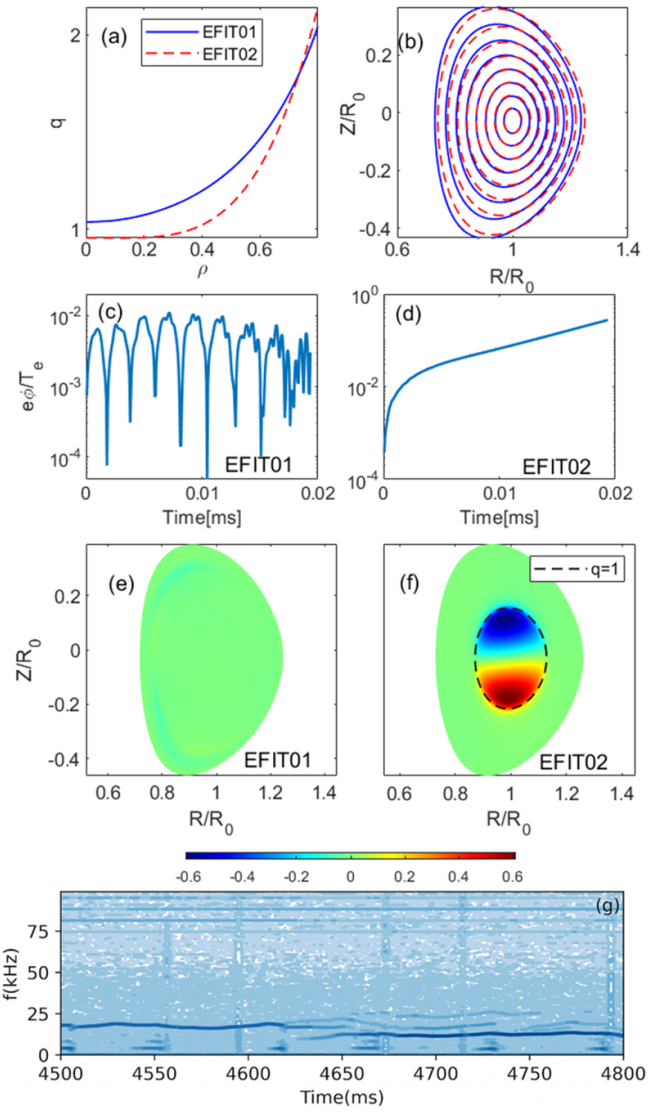

Figure 1. Comparison of GTC simulation results using equilibria reconstructed by the EFIT01 and EFIT02 for DIII-D shot #168973 at time 4676 ms. The q profile (panel (*a*)) and shape of flux surfaces (panel (*b*)) are compared. Panels (*c*) and (*d*) are the time history of $m = 1, n = 1$ mode, and exponential linear growth is observed only in the simulation result with EFIT02 equilibrium shown in panel (*d*). Panels (*e*) and (*f*) are poloidal mode structures of the perturbed electrostatic potential at the end of the simulation. Panel (*e*) only exhibits noise level fluctuations while panel (*f*) shows the unstable kink mode structure. Panel (*g*) shows the Mirnov coil spectrogram from 4500 ms to 4800 ms in shot #168973, where low-frequency long-lived perturbations can be observed.

In this work, we have built and tested a statistical model based on a deep neural network, which takes as inputs the measured plasma perturbation signals, combined with the real-time EFIT results without the MSE constraints, to predict the q profile. The model is trained on about 12 000 DIII-D EFIT02 reconstruction data. The output q profile is close to the EFIT02 results, which confirms that the MSE constraints are implicitly included in the trained model. The sensitivity test on input parameters has been carried out, showing the relative importance of each parameter to the accurate q reconstruction. This

model has been implemented as the SGTC q-Reconstruction module (SGTC-QR). The outputs of the SGTC-QR provide more accurate plasma equilibrium information for more accurate instability analysis and predictive models in the PCS, such as the SGTC MHD instability simulator [15] and fusion recurrent neural network (FRNN) disruption predictors [16]. The inference time of the pre-trained surrogate models from the SGTC-QR will be compatible with the PCS.

Machine learning and deep learning methods have previously been introduced for equilibrium reconstruction by solving the Grad–Shafranov equation [17, 18]. More recently, the EFIT-AI model [11] and the deep-reinforcement learning control [19] have demonstrated the capability for computing-intensive, hard-to-model problems. We should note the difference between the current work and the previous deep learning-based Grad–Shafranov solver [11, 17, 18]. The previous equilibrium reconstruction focused on finding the surrogate model of a known physical equation, i.e. the Grad–Shafranov equation, with given constraints. In this work, we focus on the q-profile and investigate whether it is possible to obtain a higher-quality q-profile reconstruction without the important constraint of the MSE signal. The methodology of this paper is similar to the recent works that use machine learning methods to predict the equilibrium related quantities and the last closed flux surface (LCFS) [20, 21]. In other words, the machine learning algorithm tries to recover the information contained in the MSE measurements from other measurements in the absence of a physical model, looking to the potential of reducing the expensive physical diagnostics in the planning and design of fusion pilot plants. In the future, the recovery of other measurements can also be studied. A complete equilibrium reconstruction, including the flux surface shape, will also be investigated.

In section 2, we describe the building and training details of the SGTC-QR. The testing results of SGTC-QR and an example of physical analysis based on the predicted q profile are presented in section 3. The conclusions and future works are presented in section 4.

2. Workflow and data for q profile reconstruction

The workflow of SGTC-QR is shown in figure 2. The real-time measurements provide some global equilibrium parameters such as the total current, the stored energy, MSE (if available), and the perturbation signal such as the electron cyclotron emission (ECE) and magnetic perturbations. Through the offline or real-time EFIT, the 2D equilibrium is reconstructed, including the flux surface shape, q-profile, current profile, pressure profile, etc. When the MSE measurement is not available, we can incorporate all available equilibrium and perturbation signals in the SGTC-QR module to predict the best q profile. Then these equilibrium profiles from SGTC-QR and the 2D equilibrium quantities from EFIT are taken as input by the stability models or predictors like SGTC or FRNN. The stability estimations act as feedback to the control systems. And with the knowledge of real-time instability information, the control system can actively modulate or mitigate the instabilities.

The architecture of the SGTC-QR is shown in the lower panel of figure 2. We select 16 diagnostic signals (listed in table 1) as input for the SGTC-QR. The 0D and 1D signals we have selected are widely accepted as the most important parameters to determine the tokamak operation, the equilibrium quantities, and the pedestal shape. For each time step, the electron density profile n_e and electron temperature profile T_e are 1D functions on the N_r grid points of the normalized square root of toroidal flux ρ. The magnetic perturbations are a function of frequency ω and the other signals are scalar variables. The n_e and T_e are combined into a $N_r \times N_r$ 2D feature and fed into a set of N_{c1} convolutional layers. The δB acts as a 1D feature of size N_ω, and feeds into another set of N_{c2} convolutional layers. The scalar variables are fed in a set of N_{fl} fully connected layers. The output of the convolutional layers and the fully connected layers are stacked together to go through N_{f2} fully connected layers. The final output layer has the size of N_r, which corresponds to the q profile on grid points of ρ. We use the convolutional neural networks for this task mainly for two reasons, (1) it is proved that the neural networks can approximate any functions given sufficient neural units [22]. And (2) the convolutional layers consist of small filters that are sensitive to various local patterns. A filter is activated when it sees a certain local pattern. Multiple convolutional layers will capture the global and local patterns from spatial or temporal signals or transform the features to signals. In this paper, the convolutional layers are used to capture the relevant spatial information from the 1D radial profiles and transform the latent representation to the radial q-profile.

The data for training and testing SGTC-QR are randomly selected from the DIII-D shots #125000–180844. In total, 13 491 shots are picked, and 4521 of them are disruptive. A total of 6780 shots are chosen for the training data set, with 2267 of them being disruptive; 3339 for the validation data set with 1122 of them being disruptive and another 3372 for the testing data set with 1132 of them being disruptive. The equilibrium slices are evenly spaced and selected every 25 ms from these shots, as a result, each of them typically contains 110 time-sliced equilibriums on average. The models are trained and tested in the PyTorch framework [23]. The ensemble method is used to reduce stochastic errors. We build 40 different networks, with different learning rate schedulers, and different hidden layer depths and widths. The mean-square-error between the output q profile and the EFIT02 q profile is taken as the loss function. Note that the training does not involve any output from the EFIT01. The training process uses the stochastic gradient descent optimizer, with a minibatch size of 4. The 40 networks are trained parallelly on the same training data set. The whole training takes less than 1 h for one epoch when training using one node with 128 CPUs and 4 GPUs on NERSC's Perlmutter computer. The training process would be terminated when the average validation loss of the 40 networks no longer drops in 1000 epochs to avoid overfitting.

3. Results

After the training is complete, five networks with the best validation accuracy are chosen, and the average value of

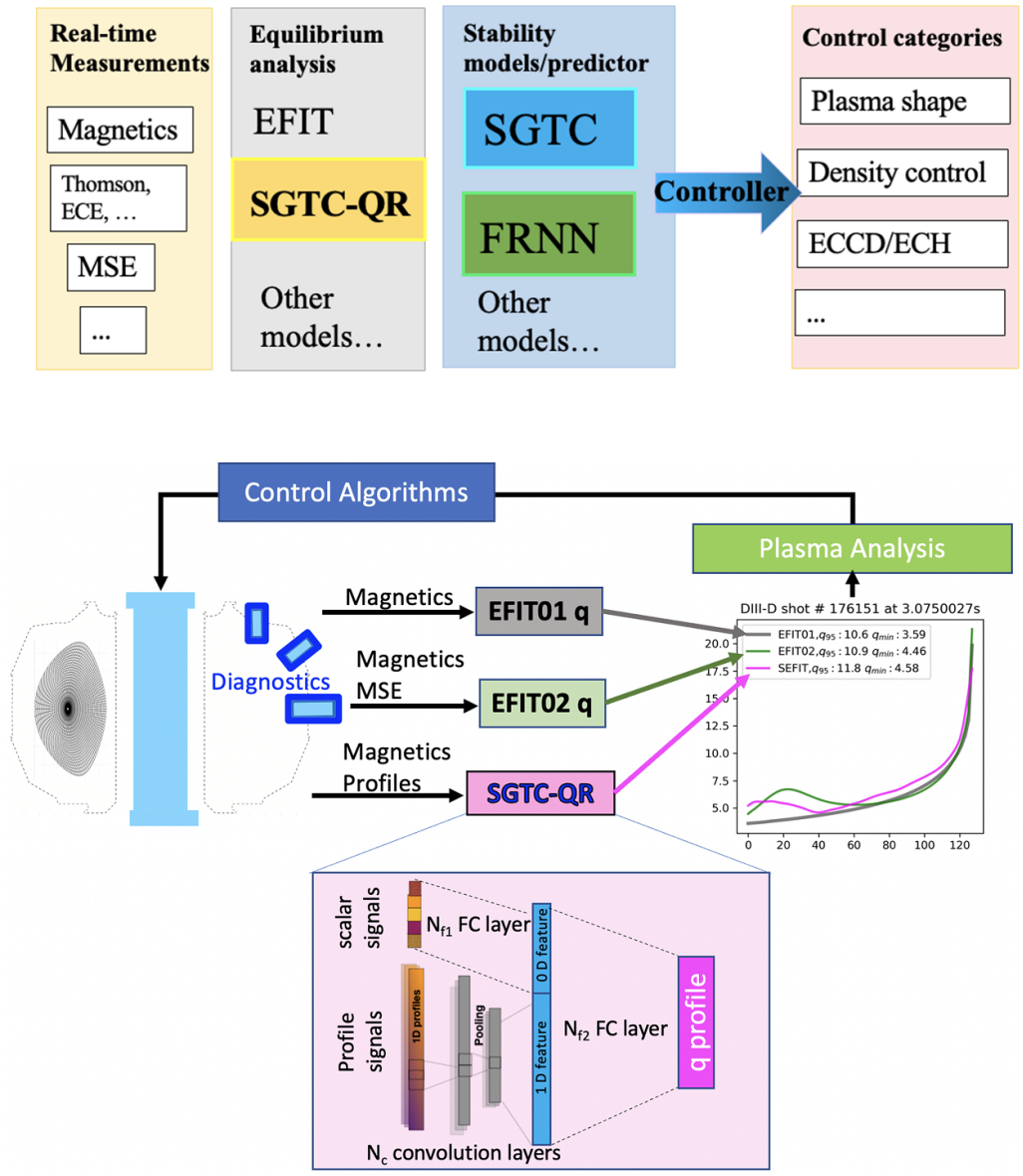

Figure 2. Upper panel shows SGTC workflow. Lower panel shows SGTC-QR workflow.

their outputs is taken as the prediction for the q profile. The average prediction time is less than 1 ms, meaning this model is compatible with the real-time PCSs. Figure 3 shows the differences between the reconstructed q profile from the SGTC-QR model and EFIT02, which is taken as the reference. The average mean-square-error of SGTC-QR is 0.165, lower than the average mean-square-error of the EFIT01 of 0.232 also using EFIT02 as the reference. The median mean-square-error of SGTC-QR (0.064) is also lower than that of EFIT01 (0.080). The distribution of the mean-square-error of the q profile in figure 3 shows that the SGTC-QR tends to have smaller errors. The comparison shows that in general the SGTC-QR reconstructed q profile is more accurate than that by the EFIT01. q_{95}, and $q_{\min}$ are two important features in the q profile that can affect the physics near the edge and core in tokamaks. In figure 4 we compare these two parameters between EFIT01 and SGTC-QR reconstructions, using the EFIT02 results as reference true value. The deviation of the points from the $y = x$ line shows the difference of reconstructed q_{95} or $q_{\min}$ from the EFIT02 results. From the two upper panels of figure 4, we see that the q_{95} difference of SGTC-QR with the EFIT02 is comparable to that of the EFIT01. In the lower two panels of figure 4, the SGTC-QR $q_{\min}$ values are closer to the EFIT02 than that of the EFIT01, which indicates that the reconstruction in the core region is more accurate. This is also demonstrated by the radial error distribution in figure 5. The errors in each of the 128 radial points are calculated for all shots in the test dataset and plotted in figure 5. It clearly shows that both the average and median of the errors from the SGTC-QR reconstruction are lower than the EFIT01 when $\rho < 0.8$. The reason for the large error near the edge is as follows. The q_{95} is taken as a constraint in the EFIT01 and EFIT02 reconstructions, which enforces the q_{95} value to be close to the measured one during the iteration of the EFIT calculation.

Table 1. Data used in training, validation, and testing of the SGTC-QR. ρ is normalized toroidal flux and ω is time frequency. The pedestal height is set to 0 for L-mode shots.

Signal description	Symbol	Usage in SGTC-QR	Shot number
Safety factor profile from EFIT02	$q(\rho)$	Output	*12 000 equilibriums from shot # 125000–180844*
Safety factor profile from EFIT01		Not used	
Electron temperature profile	$T_e(\rho)$	Input	
Electron density profile	$n_e(\rho)$		
Mirnov spectrogram	$\delta B(\omega)$		
Internal inductance	L_i		
Plasma density	n_{e0}		
Safety factor at 95% radial domain from EFIT01	q_{95}		
Plasma current	I_p		
Ratio of thermal to magnetic pressure	β		
Plasma stored energy	W_{MHD}		
Input beam power	P_{in}		
Input beam torque	T_{in}		
Pressure profile pedestal height	P_{ped}		
Temperature profile pedestal height	T_{ped}		
Density profile pedestal height	n_{ped}		
Temperature profile pedestal width	Δ_T		
Density profile pedestal width	Δ_n		

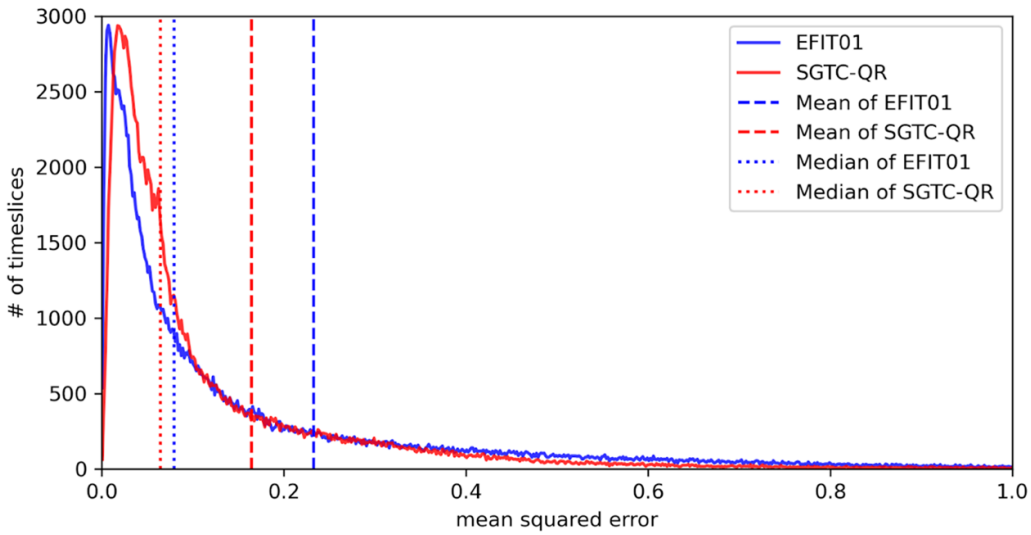

Figure 3. Comparison of the mean squared error of EFIT01 q profile and SGTC-QR q profile, using EFIT02 q profile as true value. Data are shown for the test dataset with 164 782 time slices from 3372 shots.

On the other hand, SGTC-QR only takes the q_{95} as an input parameter, without using it as a constraint. Furthermore, the q value at the separatrix is singular (infinity), and the reconstruction in the EFIT02 (and of course also in the EFIT01) is inaccurate near the separatrix. Therefore, it is very difficult for the deep-learning method to find the pattern of the q profile from the EFIT02 near the separatrix. In fact, the errors rapidly increase when ρ approaches 1, especially at the last three radial points, which are omitted in figure 5 (and included in the calculation of figure 3). If we ignore the region for the last three points, the error of the q profile from SGTC-QR would be even smaller than that shown in figure 3.

From the above results, we conclude that the reconstructed q profile outperforms the EFIT01 and successfully incorporates the MSE constraint implicitly. The underlying reason for the successful reconstruction is revealed through a sensitivity study. In this study, we contaminate one of the inputs by directly setting it to 0 each time, and then measure the prediction accuracy on the testing data set. The increase of the median and average mean-square-error value from the original one reflects the relative importance of each input parameter. The result is shown in figure 6, where the title of each bar labels the contaminated input. The lines are arranged in the order of decreasing average mean-square-error, and all the errors are normalized so that the median and average mean-square-error of the original SGTC-QR prediction is 1. We also show the error from the EFIT01 q profile reconstruction in the figure. Figure 6 shows the most important input for successful q reconstruction is the plasma current, the internal

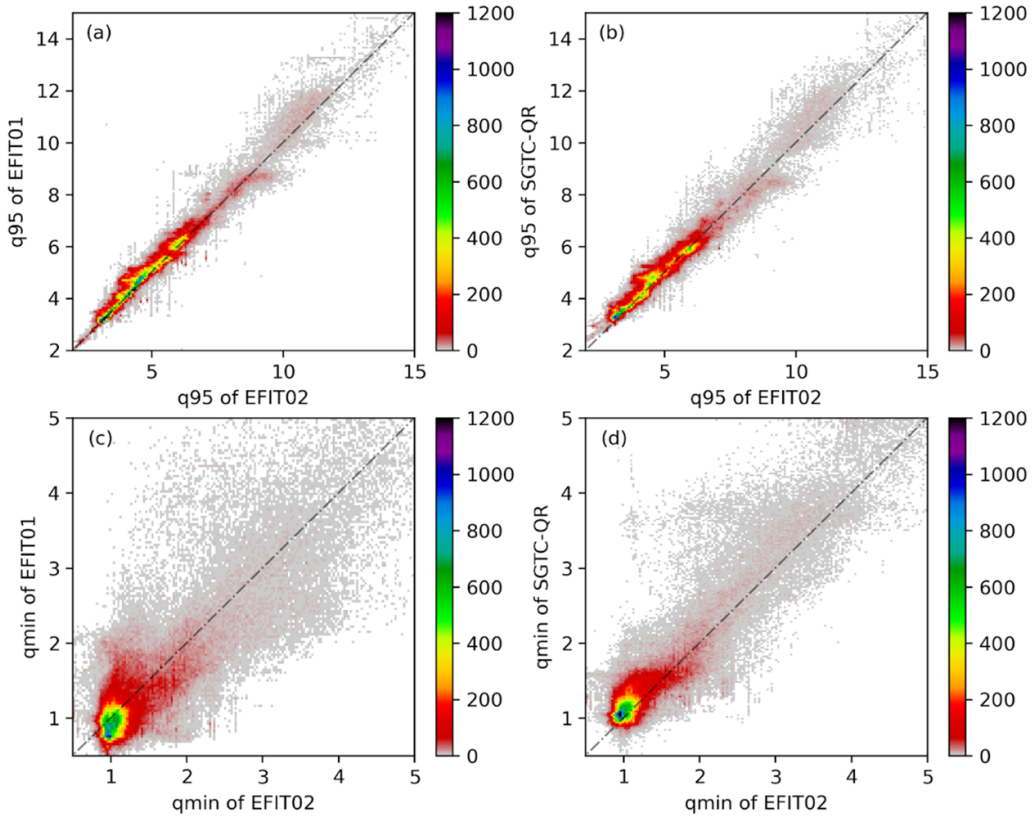

Figure 4. Comparison of q_{95} and $q_{\min}$ given by the EFIT01 and SGTC-QR, using q_{95} and $q_{\min}$ from the EFIT02 as the reference. Data are shown for the testing dataset with 164 782 time slices from 3372 shots. Colors indicate the number of points.

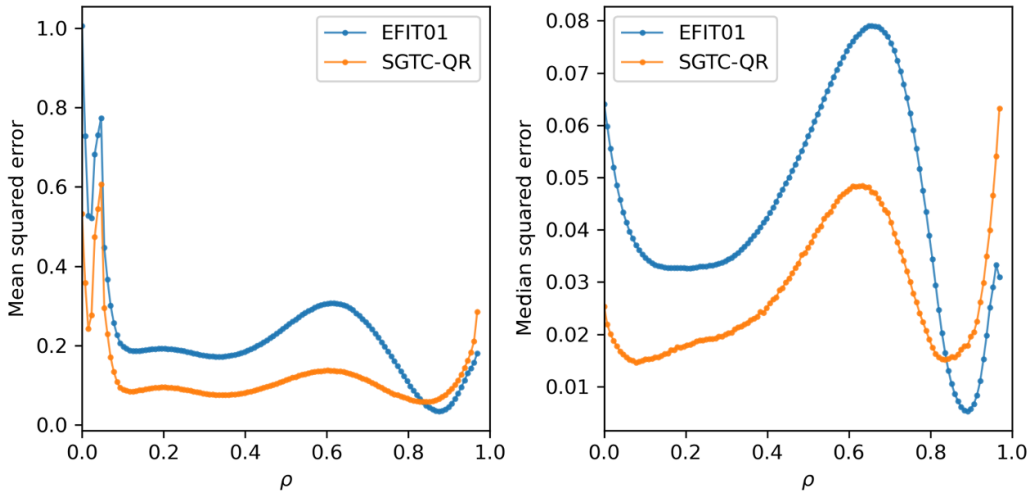

Figure 5. Comparison of mean-square-error of the EFIT01 and SGTC-QR at each radial point. The q-profile from EFIT02 is used as the reference true value (the last three radial grid points are not shown).

inductance, and the normalized beta. The importance of these factors to the q profile makes physical sense because they directly affect the current and pressure profile, and directly determine the solution of the Grad–Shafranov equation that determines the q profile. On the other hand, we should point out that the 'less importance' does not necessarily imply less physical relation between the signals and the q-profile. There could be redundancy among the input signals and the inclusion of one relevant signal may not increase the accuracy of QR significantly with the presence of all other signals. It will be an interesting future study to find an algorithm to reveal the correlation of different signals and the rigorous physical dependence of q-profile on each of the signals. It is remarkable that no additional input signal significantly deteriorates the performance of SGTC-QR, demonstrating the advantage of deep-learning based methods that have much stronger capability of drawing useful information from large size of input data. It is also worth noting that the Mirnov coil signal is not

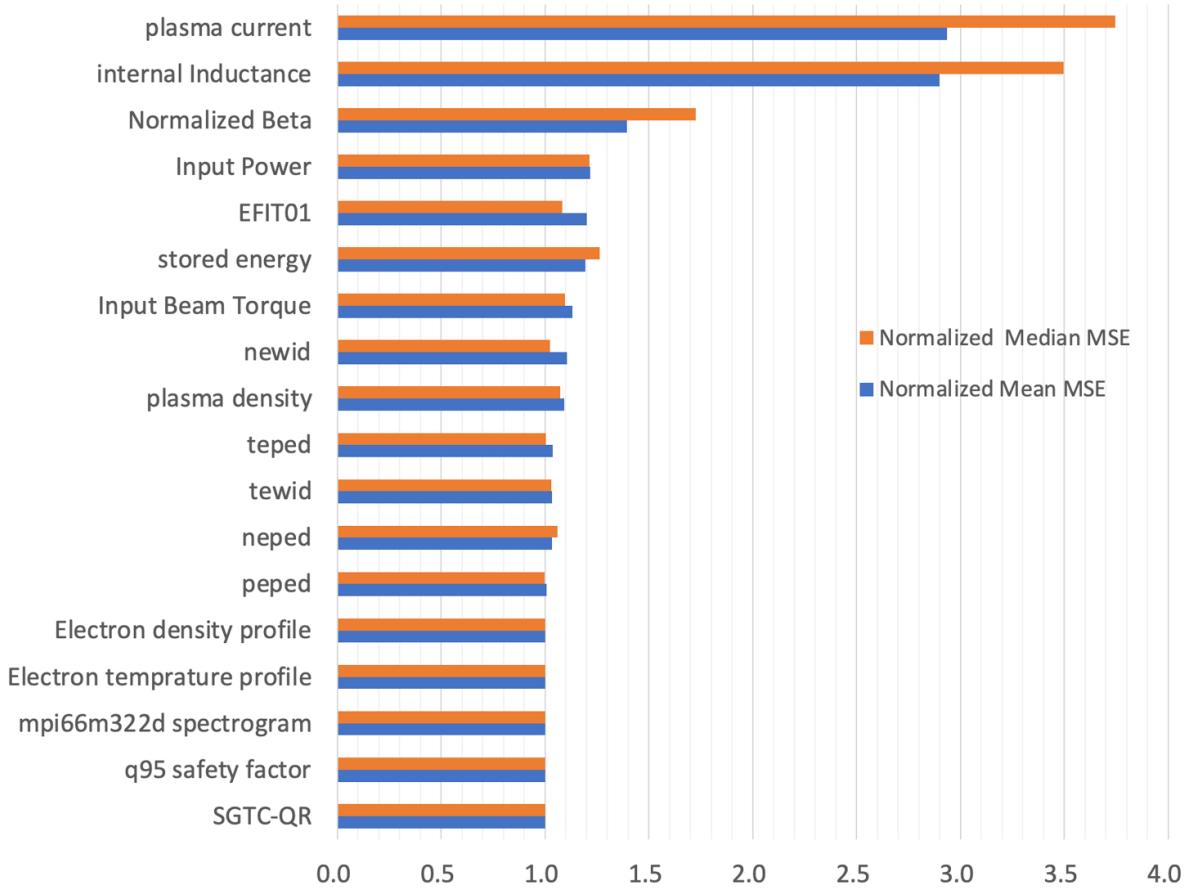

Figure 6. Sensitivity study on testing data set. All mean-square-error data are normalized to the results of the SGTC-QR model. Orange/blue bars show the normalized median/mean mean-squared-error when the labeled signal is contaminated. Longer bar indicates higher sensitivity in the marked input signal.

very important to the q-profile reconstruction. One possibility is that the dominant fluctuations shown from the Mirnov coils may come from the instability that does not strongly depend on the q-profile and the 0D magnetic fluctuations are just not enough to correlate with the q-profile data. Another possibility is that the Mirnov signal is redundant when other signals are taken as input of the model. This result implies that we should add more input signals in the future like spatial distribution of the fluctuations and study the redundancy in the input features that determine the q-profile. It is also notable that the q_{95} value and the magnetic fluctuation signal do not make much difference in the improvement of the reconstruction. This indicates that the SGTC-QR does not take the information from the q_{95} factor efficiently, and it potentially causes an increase in the error in the edge region. In principle, we can add q_{95} as a term in the loss function or increase the weight of the edge region in the loss function to increase the accuracy in the edge region. But for the neural network architectures we have used in this study, we have found a tradeoff between the accuracy in the core region and the edge region. Because the SGTC modules developed previously and in the near future will mainly focus on instabilities in the core including Alfven eigenmodes, kink mode, and neo-classical tearing modes (NTM), here we prefer to keep the core accuracy. A possible way to improve the precision of the construction is to use two different neural networks to predict the q-profile in the core region and edge region, respectively. We can still use the current neural network structure and training method in the core region. And the model for the edge reason needs re-design and increase the penalty of q_{95} error during training. Finally, a method is needed to connect the two regions smoothly to form the complete q-profile. We will leave this as an open issue, and develop a separate model to target the q profile reconstruction near the edge region in future work. In the future, we will also perform the uncertainty quantification to show the quality of the reconstruction due to the defect in the training data at different radial locations.

In figure 7, we show the q-profile reconstruction for two typical cases that appeared in previous publications. The line for SGTC-QR is from the model with the best validation score among all 40 models. The left panel is the equilibrium in the benchmark case for the internal kink mode [5], in which the EFIT01 reconstruction is invalid. Both EFIT02 and SGTC-QR show normal shear, and the q profile near the axis is very similar, which is important for calculating the correct kink drive. The right panel is the equilibrium in the benchmark case for the reversed-shear Alfven eigenmode [24]. The EFIT02 q profile shows normal shear, but the SGTC-QR shows reversed shear,

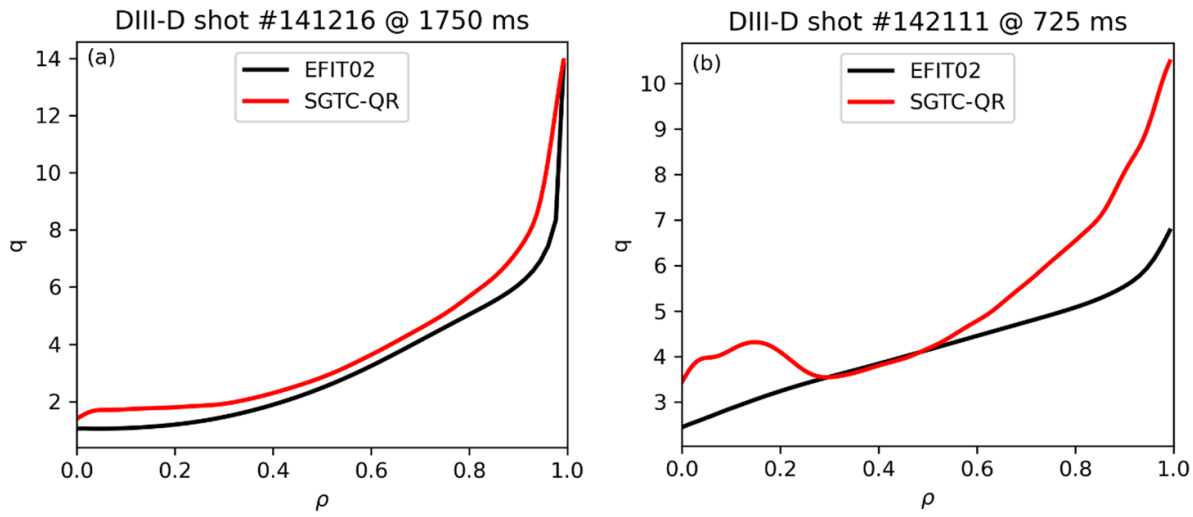

Figure 7. Comparison of q profile reconstructions for shot #141216 [5] and # 142111 [24].

with $q_{\min} \approx 3.1$. The inconsistency between the two q profiles does not mean the failure of the machine learning model. More accurate analysis with human calibration indicates that the q-profile should have a reversed shear with $q_{\min} = 3.18$ [24]. In this case, the EFIT02 q profile deviates from the q distribution that is considered more physical by human experts. By learning on the whole training dataset, the robust machine learning model successfully predicts the correct q profile according to the learned data distribution. In this sense, the SGTC-QR has the potential to warn and correct the q profile if the q profile reconstruction in the test dataset is inaccurate and deviates from the typical data distribution. On the other hand, the SGTC-QR q-profile in figure 7(*b*) shows the irregular shape near the axis. It is unclear at this point whether this shape comes from the pattern in the EFIT02 data, or it is from the prediction with large uncertainty of the machine learning model. The uncertainty quantification in the future upgraded version of SGTC-QR model will reveal the reason of similar irregular q-profile shape and the confidence we should have on the reconstruction at certain radial locations.

4. Conclusions and future study

In this work, we have designed and built a deep-learning based model for accurately reconstructing the safety factor profile, and tested the model for DIII-D experimental data. A total of 26 400 equilibria from DIII-D shots were selected for the training and testing of the model. Using the 0D and 1D signals, the model is trained to recover the q profile from the EFIT02 reconstruction. The model has demonstrated promising capability in terms of speed and accuracy. The typical prediction time is less than 1 ms, and the predicted q profile is close to that from the EFIT02. The correlation between the input signals and the MSE constraint has been found and included during the training process. This tool can be used to generate the q profile for the shots where MSE is unavailable. The SGTC-QR module has been implemented in the SGTC framework and the whole framework can provide abundant real-time equilibrium and perturbation information for the PCS in fusion devices. Several improvements can be made in future work. We will include more input features and explore other carefully designed loss function to improve the reconstruction accuracy to meet the engineering criteria, especially in the edge region. We will use the more accurate kinetic-EFIT constructed q-profile as reference and explore the relations between this q-profile and the input features. We can extend the current model to include more constraints and predict the self-consistent 2D equilibrium, pressure profile, and the current profile without solving the Grad–Shafranov equation. An end-to-end 2D or 3D equilibrium reconstruction model will be built in the same manner as in this work. A deep learning model to build Boozer coordinate system based on the constructed equilibrium could also be interesting. We will explore the efficient way to transfer of the learned QR model to other tokamaks with acceptable errors.

Acknowledgments

This work is supported by the U.S. Department of Energy (DOE) SciDAC Project ISEP and used resources of the Oak Ridge Leadership Computing Facility at Oak Ridge National Laboratory (DOE Contract No. DE-AC05-00OR22725) and the National Energy Research Scientific Computing Center (DOE Contract No. DE-AC02-05CH11231). This work is partially based upon work using the DIII-D National Fusion Facility, a DOE Office of Science user facility, under Awards DE-FC02-04ER54698 and DE-FG02-95ER54309. This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness,

ORCID iDs

Xishuo Wei https://orcid.org/0000-0001-7486-0407
Shuying Sun https://orcid.org/0000-0002-3991-872X

References

[1] Lao L.L., St. John H., Stambaugh R.D., Kellman A.G. and Pfeiffer W. 1985 *Nucl. Fusion* **25** 1611
[2] Lao L.L., Ferron J.R., Groebner R.J., Howl W., St. John H., Strait E.J. and Taylor T.S. 1990 *Nucl. Fusion* **30** 1035
[3] Wróblewski D. and Lao L.L. 1992 *Rev. Sci. Instrum.* **63** 5140
[4] Lin Z., Hahm T.S., Lee W.W., Tang W.M. and White R.B. 1998 *Science* **281** 1835
[5] Brochard G. *et al* 2022 *Nucl. Fusion* **62** 036021
[6] Holcomb C.T., Makowski M.A., Jayakumar R.J., Allen S.A., Ellis R.M., Geer R., Behne D., Morris K.L., Seppala L.G. and Moller J.M. 2006 *Rev. Sci. Instrum.* **77** 10E506
[7] Holcomb C.T., Makowski M.A., Allen S.L., Meyer W.H. and Van Zeeland M.A. 2008 *Rev. Sci. Instrum.* **79** 10F518
[8] Ferron J.R., Walker M.L., Lao L.L., John H.E.S., Humphreys D.A. and Leuer J.A. 1998 *Nucl. Fusion* **38** 1055
[9] Creely A.J. *et al* 2020 *J. Plasma Phys.* **86** 865860502
[10] Huang Y. *et al* 2020 *Nucl. Fusion* **60** 076023
[11] Lao L.L. *et al* 2022 *Plasma Phys. Control. Fusion* **64** 074001
[12] Spong D. *et al* 2012 *Phys. Plasmas* **19** 082511
[13] Turnbull A.D. *et al* 2002 *Nucl. Fusion* **42** 917
[14] Xing Z.A. *et al* 2021 *Fusion Eng. Des.* **163** 112163
[15] Dong G., Wei X., Bao J., Brochard G., Lin Z. and Tang W. 2021 *Nucl. Fusion* **61** 126061
[16] Kates-Harbeck J., Svyatkovskiy A. and Tang W. 2019 *Nature* **568** 526–31
[17] van Milligen B.P., Tribaldos V. and Jiménez J.A. 1995 *Phys. Rev. Lett.* **75** 3594
[18] Joung S. *et al* 2019 *Nucl. Fusion* **60** 016034
[19] Degrave J. *et al* 2022 *Nature* **602** 414–9
[20] Wan C., Yu Z., Pau A., Liu X. and Li J. 2022 *Nucl. Fusion* **62** 126060
[21] Wan C., Yu Z., Pau A., Sauter O., Liu X., Yuan Q. and Li J. 2023 *Nucl. Fusion* **63** 056019
[22] Hornik K., Stinchcombe M. and White H. 1989 *Neural Netw.* **2** 359–66
[23] Paszke A. *et al* 2019 PyTorch: an imperative style, high-performance deep learning library *Advances in Neural Information Processing Systems* vol 32 (Red Hook, NY: Curran Associates, Inc.) pp 8024–35 (available at: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf)
[24] Deng W., Lin Z., Holod I., Wang Z., Xiao Y. and Zhang H. 2012 *Nucl. Fusion* **52** 043006