*35th EPS Conference on Plasma Phys. Hersonissos, 9 - 13 June 2008 ECA Vol.**32D**, P-1.065 (2008)*

Application of Neural Networks for Fast Tomographic Inversion on Wendelstein 7-X

H. Thomsen, P.J. Carvalho†, S. Gori, U. v. Toussaint, A. Weller, R. Coelho†, H. Fernandes†

Max-Planck-Institut für Plasmaphysik, EURATOM Association, Greifswald, Germany

†*Associação Euratom/IST Instituto de Plasmas e Fusão Nuclear, Instituto Superior Técnico 1049-001 Lisboa, Portugal*

Introduction

The Wendelstein 7-X (W7-X) device, presently under construction in Greifswald, will allow high power quasi-steady state discharges with durations of up to 30 min. Diagnostics should therefore be able to cope with large heat loads (requirement of active cooling) and drift compensation for a good signal measurement stability. Moreover, the recorded data might be used for controlling the discharge flow. For more advanced scenarios, a conditional branching into different discharge segments will be possible within the framework being developed by the W7-X software group [1]. For these scenarios online data analysis capabilities for the diagnostics are very favourable. On the other hand, the data rates of diagnostics with a high temporal and spatial resolution will make very fast data analysis methods necessary, either for identifying the interesting bits inside a huge amount of raw data, or – if fast data acquisition for the whole discharge is not feasible – online triggering capabilities for setting fast time windows for the data acquisition. One diagnostic with very high spatial and temporal resolution is the soft-Xray Multi Camera Tomography System (XMCTS) with 400 lines of sight (crossing through the plasma from 20 cameras located around the minor circumference) and 500 kHz bandwidth. The signals from the photo diode arrays will be digitized with 14 Bit at 1 MSample per second. The total data of this diagnostic alone are a huge 1400 GBytes per 30 min discharge. The system will allow a tomographic inversion of the radiation distribution in a poloidal plane of the plasma, giving information on plasma position and -shape as well as MHD-stability (considering pressure constancy on magnetic flux surfaces). Since the plasma shape has a specific form for different values of the plasma-beta, it is possible to derive this important quantity from the inversion. We note here that an obvious application of this technique for the case of tokamaks, namely the plasma position control, is not an issue for stellarators, due to their inherent stability of the magnetic configuration.

The method of choice for offline tomographic reconstruction will be based on regularized matrix inversion methods. Due to the iterative minimization involved, these methods will presumably be too slow (although there are interesting attempts in the community to speed up these methods [2]). In this paper we consider neural networks (NNs) for the fast reconstruction. For the training of NNs, a large amount of training data is necessary, comprising of the measurement data (input for the NNs) and the 2D radiation distribution of the plasma (corresponding output of the NNs).

Generation of training data for NNs

In order to achieve a robust training for the NNs, all possible (and probable) radiation distributions should be learned. We limited the training data set to one magnetic configuration: Since the confining field in a stellarator is applied by the external coils, it is possible to train different NNs for the various possible magnetic configurations and to select the appropriate one according to the coil currents. Several numerical software packages were involved for creating the radiation distributions (c.f. Fig. 1), including VMEC [3], a transport code [4] and the impurity radiation modeling code IONEQ [5]. The output of the latter code chain is a radiation profile.

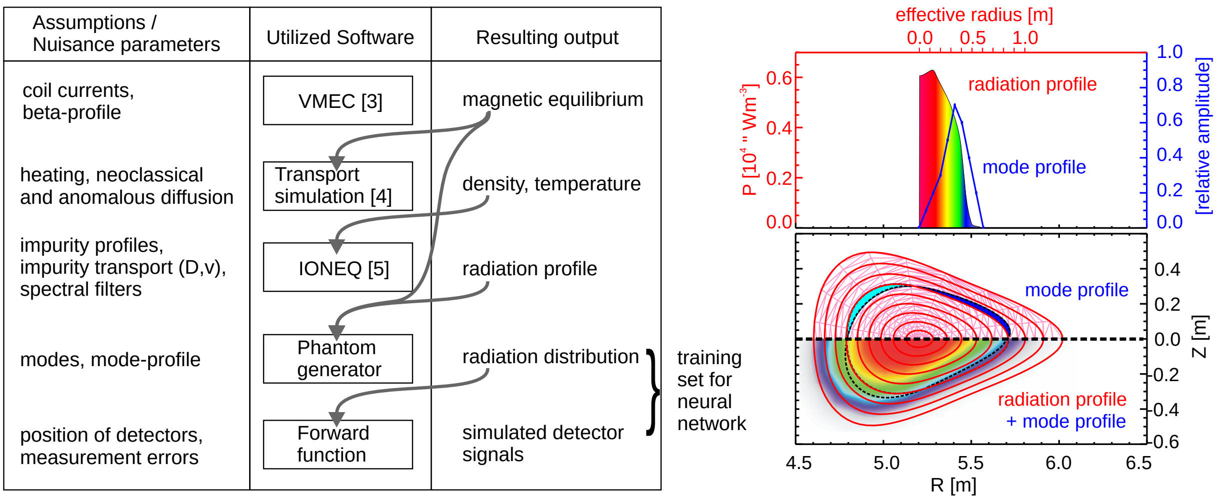

Figure 1: Left: Simulation software chain used to create the inputs and outputs for training the NNs. Dependencies and nuisance parameters are also listed. Right: m=2 mode perturbs the magnetic flux-surfaces (VMEC output). The radiation profile (from IONEQ) is mapped onto these perturbed flux-surfaces.

MHD modes are modeled as a perturbation in the magnetic flux surfaces, assuming mode number and mode location profiles as well as excursion amplitudes. The radiation profile is mapped on the thus perturbed magnetic flux surface grid (c.f. Fig. 1). The radiation distribution (output) is then interpolated on a 40x30 grid. A contribution matrix can be calculated via ray-tracing considering the etendues of the pinhole cameras for each diode. The inputs are then governed by means of a simple matrix-vector multiplication with the contribution matrix.

Neural networks

We utilize multi-layer-perceptron (MLP) networks with two hidden layers and feed-forward weights. The activation is via a tanh-function. The training and pruning is done unsupervised within an Bayesian statistics framework (utilizing the usual backpropagation algorithm for weight updating). The problem of finding the optimal number of hidden neurons is solved within this framework, using hyperplane priors [6]. However, the computational effort for training the NNs and finding the optimal one(s) is large: The optimization took several days on a 20 CPU Linux cluster for 140 training sets (times 4, noise added). The training sets comprise a beta-scan with a number of modes (m=0, m=2, m=3, m=4) for a single radiation and mode profile. In order to achieve a robustly trained NN (i.e., well-performing when confronted with unknown inputs, avoiding over-fitting), a Gaussian noise component with 5% relative amplitude was applied to the input data. For future improvements a more realistic noise modeling (correct noise statistics) should be considered.

Results

The trained neural network was validated with simulation data. In contrast to the training data, the mode profile location was varied. In order to test the generalization performance of the NN, two untrained magnetic configurations ($\beta = 2.7$ % and $\beta = 5.6$%) were used (cf. Fig. 2). The latter configuration can be seen as an extrapolation over the boundary of the trained

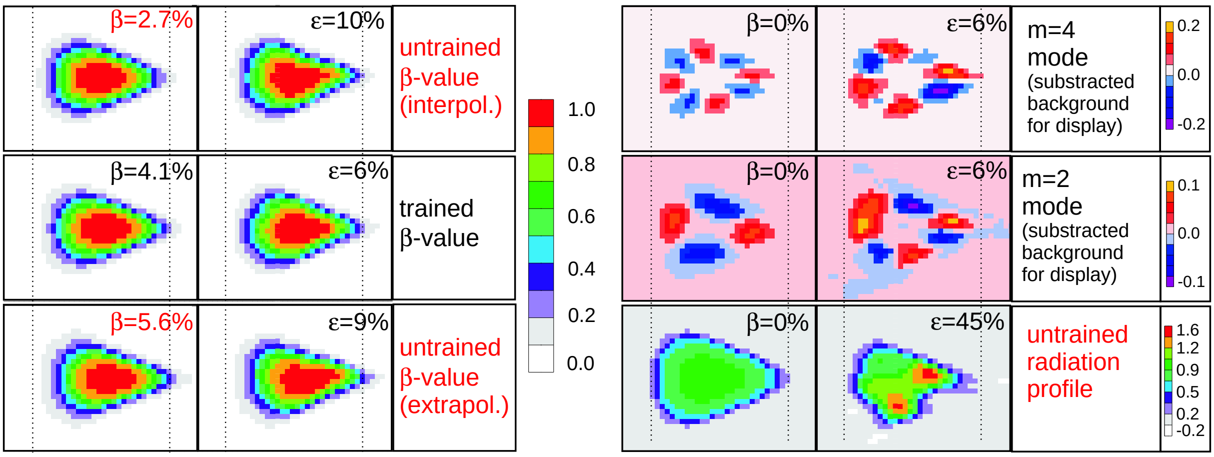

Figure 2: Left: comparison between phantom and reconstruction data for selected phantoms with different β-values. Right: comparison of mode reconstruction and a failing reconstruction for a phantom with a different (untrained) radiation profile.

parameter space, whereas the first configuration corresponds to an interpolation. As a measure for the deviation between phantom and reconstruction, the value ε is defined as

$$\varepsilon = \frac{\sum_{x,y}[g_{\text{phantom}}(x,y) - g_{\text{reconstruction}}(x,y)]^2}{\sum_{x,y}[g_{\text{phantom}}(x,y)]^2}, \qquad (1)$$

with $g(x,y)$ the radiation value at position (x,y) in the 2D grid. It should be noted that ε is only appropriate for cases where the initial phantoms are known. A very similar measure is possible for future real-time application, the reconstruction matrix needs to be back-calculated into the signal-space with the contribution matrix. The measured input signal from the diodes can be compared to the back-calculated signals and the reconstruction quality can thus be estimated. This method is only possible when reconstructing the radiation distribution, since the forward function is known and can be calculated with high speed. However, for NNs trained to output a single value (e.g., the β-value), this will not be possible. A second way to control the accuracy of the reconstruction is possible by comparing the output of two (or more) differently trained NNs. If measurements are carried out in an untrained parameter region of the NNs, the outputs will differ strongly.

Conclusions

Real time tomographic inversion requires very fast algorithms. Neural networks can be a solution for this challenge, since they provide a good generalization flexibility. A benefit is that not only the profile data can be reconstructed, but also relevant information can be learned directly, like the average plasma beta given a set of line-integrated measurements. However, the training and selection of a sufficiently general and robust network is a time consuming task. The NN optimization (including controlling the network complexity) can be automated within a Bayesian framework. The training data is generated using several simulation codes, including IONEQ for the radiation profile calculation and VMEC equilibrium for mapping the profiles into the real W7-X geometry. A possible extension of this attempt would be to integrate these models within a larger Bayesian framework in order to include profiles and further data from other diagnostics towards an integrated data analysis. Also, error bars for the reconstruction are feasible using errors on the inputs and calculating the error propagation through the various models.

Acknowledgement:

The first author (HT) would like to thank J. Geiger, Yu. Turkin and J. Svensson for fruitful discussions.

References

[1] J. Schacht et al. Fusion Engineering and Design **82**, 988 (2007).
[2] J. Mlynar, S. Coda, A. Degeling et al. Plasma Phys. Contr. Fusion **45**, 169 (2003).
[3] S. P Hirshman, W. I. van Rij, and P. Merkel, Comput. Phys. Commun. **43**, 143 (1986).
[4] Yu. Turkin, H. Maassberg, C. D. Beidler, et al. Fusion Sci. and Technol. **50**, 387 (2006).
[5] A. Weller, D. Pasini, A. W. Edwards, R. D. Gill, R. Granetz. JET Report JET-IR-(87)10 (1987).
[6] U. v. Toussaint, S. Gori and V. Dose. Neural Netw. **19**, 1550 (2006).