

CONFERENCE PRE-PRINT

**ESTIMATION OF PLASMA PARAMETERS BASED ON DISCHARGE SETTINGS
ON WEST**

Chenguang Wan
School of Physical and Mathematical Sciences, Nanyang Technological University
Singapore 637371, Singapore
Institute of Plasma Physics, Hefei Institutes of Physical Science, Chinese Academy of Sciences
Hefei, China
Email: chenguang.wan@ntu.edu.cn

Feda Almuhsen
CEA, IRFM
Saint Paul-lez-Durance, France

Philippe Moreau
CEA, IRFM
Saint Paul-lez-Durance, France

Rémy Nouailletas
CEA, IRFM
Saint Paul-lez-Durance, France

Zhisong Qu
School of Physical and Mathematical Sciences, Nanyang Technological University
Singapore, Singapore

Youngwoo Cho
School of Physical and Mathematical Sciences, Nanyang Technological University
Singapore, Singapore

Robin Varennes
School of Physical and Mathematical Sciences, Nanyang Technological University
Singapore, Singapore

Kyungtak Lim
School of Physical and Mathematical Sciences, Nanyang Technological University
Singapore 637371, Singapore

Kunpeng Li
School of Physical and Mathematical Sciences, Nanyang Technological University
Singapore, Singapore

Zhengping Luo
Institute of Plasma Physics, Hefei Institutes of Physical Science, Chinese Academy of Sciences
Hefei, China

Qiping Yuan
Institute of Plasma Physics, Hefei Institutes of Physical Science, Chinese Academy of Sciences
Hefei, China

Xavier Garbet
School of Physical and Mathematical Sciences, Nanyang Technological University
Singapore, Singapore
CEA, IRFM
F-13108 Saint Paul-lez-Durance, France

Abstract

Accurately predicting plasma behavior based on discharge configurations is essential for the safe and efficient operation of tokamak experiments. While physics-based integrated modeling codes provide valuable insights, their high computational cost limits their applicability for fast scenario design and control optimization. In this study, we propose a transformer-based machine learning model to predict key global plasma parameters on the WEST tokamak, including the normalized beta (β_n), toroidal beta (β_t), poloidal beta (β_p), plasma stored energy (W_{mhd}), safety factor at the magnetic axis (q_0), and safety factor at the 95% flux surface (q_{95}). The model uses only signals that can be specified or regulated before the discharge, such as magnetic coil currents, auxiliary heating power, plasma current reference, and line-averaged plasma density. Trained on 550 discharges from the WEST campaigns, the model demonstrates an average mean square error (MSE) loss of 0.026, an average coefficient of determination R^2 of 0.94, and achieves inference times on the order of 0.1 seconds. These results highlight the potential of data-driven surrogate models for assisting in discharge planning, scenario evaluation, and real-time control of tokamak plasmas.

1. INTRODUCTION

Any plasma discharge needs to be prepared in advance using a discharge schedule editor, in which all tokamak settings are defined. In order to optimize the tokamak operation, these settings must be assessed using a tokamak simulator in order to check the feasibility of the preset discharge scenario.

Predicting plasma behavior from the inputs taken from the discharge schedule editor in tokamaks has traditionally relied on first-principles physics codes, commonly referred to as "Integrated Modeling". These include ETS [1], PTRANSP [2], TSC [3], NICE [4], CRONOS [5], JINTRAC [6], METIS [7], ASTRA [8], RAPTOR [9], and TOPICS [10], etc. While these codes simulate key physical processes with high fidelity, their substantial computational cost makes them unsuitable for fast or real-time prediction tasks.

This is especially true when we need to quickly predict zero-dimensional parameters, such as plasma stored energy W_{mhd} , normalized beta β_n , safety factor q at magnetic axis q_0 , from a given set of coil current and the presetting parameters. The data-driven methods offer a way to bridge this gap by learning the input-output mapping from the experimental data and bypassing the need to solve first-principle equations every time. Because of this advantage, researchers in the fusion community have used data-driven methods to study several problems such as disruption prediction [11, 12], magnetic field control [13], surrogate model for Ion Temperature Gradient (ITG) mode [14], plasma magnetic measurement estimation [15], and so on.

Furthermore, some data-driven approaches have explored full discharge prediction [16, 17, 18]. These models do not rely solely on configurations but also incorporate signals such as gas puffing system data, which are hard to determine prior to a discharge and are influenced by the evolving plasma state during the discharge. This dependence constrains their generalizability and practical applicability.

In this study, we developed a transformer-based machine learning model that utilizes only signals that can be directly controlled or feedback-regulated. Compared to previous works [16, 17], our model does not rely on the state of any specific discharge process, thereby enhancing its practical applicability. Specifically, we successfully reproduced six key zero-dimensional global parameters on the Tungsten (W) Environment in Steady-State Tokamak (WEST) by inputting 19 signals, including poloidal field coil currents, auxiliary heating, plasma reference current, and plasma electron density measurements at the magnetic axis.

The remainder of this paper is organized as follows. Section 2 describes the dataset preparation and the analysis of the input and output signals. Section 3 outlines the machine learning methodology. Section 4 presents the model results along with the corresponding analysis. Finally, Section 5 presents a brief conclusion and the future direction.

2. DATASET

The WEST tokamak is equipped with a comprehensive suite of diagnostic systems, including magnetic diagnostics [19] and plasma state diagnostics [20, 21, 22], among others. In the present study, as shown in Table 1, the input signals include the power and corresponding phase of the Lower Hybrid Wave (LHW) Current Drive and Heating System, the power of the Ion Cyclotron Resonance Heating (ICRH) System, the reference plasma current, the currents in the Poloidal Field (PF) coils, and the real-time measured plasma line-average density. The output signals include normalized beta β_n , toroidal beta β_t , poloidal beta β_p , plasma stored energy W_{mhd} , safety factor q at magnetic axis q_0 and safety factor q at 95% flux surface. A total of 550 discharges, ranging from discharge #57381 to #60286, were initially selected from the WEST database. Discharges with durations less than 2 seconds

TABLE 1. THE LIST OF SIGNALS. A signal name ending with `_real` indicates the real diagnostic data during the tokamak discharge process. The signal name ending with `_ref` indicates that the data is reference data, i.e. nominal signals in the Discharge Control System (DCS) of WEST.

Signals	Physics meanings	# of channels	sampling rate
Output signals		6	
β_n	Normalized beta	1	1 - 12 Hz
β_t	Toroidal beta	1	1 - 12 Hz
β_p	Poloidal beta	1	1 - 12 Hz
W_{mhd}	Plasma stored energy	1	1 - 12 Hz
q_{95}	Safety factor q at 95% flux surface	1	1 - 12 Hz
q_0	Safety factor q at magnetic axis	1	1 - 12 Hz
Input signals		19	
LHW_real	Power of Lower Hybrid Wave (LHW) Current Drive and Heating System	4	1 kHz
ICRH_real	Power of Ion Cyclotron Resonance Heating (ICRH) System	3	1 kHz
I_p _ref	Reference of plasma current	1	1 kHz
n_e _real	Actual line-average electron density at magnetic axis	1	1 kHz
PF_real	Current of Poloidal Field (PF) coils	10	1 kHz

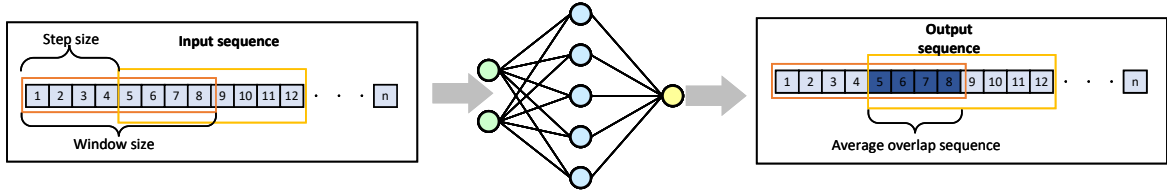


FIG. 1. The workflow of machine learning model in the present work

were excluded, as the ramp-up phase in the WEST tokamak typically takes about 2 seconds. Such short discharges are usually indicative of abnormal events and were not included in the analysis. All selected signals were uniformly resampled at 1 kHz from the start of each discharge and aligned to a common time base. The data were saved in HDF5 format on a discharge-by-discharge basis, resulting in approximately 15 gigabytes of original data.

3. METHODS

The machine learning model and its data processing workflow are shown in Fig. 1. Prior to model input, the raw data are smoothed using a simple moving average filter to suppress noise. To handle discharge sequences of different lengths, the data are segmented using a fixed-size sliding window of length 1024 and a step size of 512, resulting in a data augmentation ratio of 2. Because the output sequences overlap (as shown by the dark blue regions in Fig. 1), the overlapping segments are averaged to improve numerical stability and prediction accuracy. The 550 discharges were randomly divided into training (60%), validation (20%), and test (20%) sets.

Our machine learning model was developed using PyTorch on Red Hat Enterprise Linux 8, running on four A100 GPUs. During model training, we utilized the Bayesian algorithm [23] to perform the architectural hyperparameter search. Additionally, we experimented with various optimizers and regularization techniques, ultimately identifying the optimal set of hyperparameters, as shown in table 2. Inference time was further tested on both a consumer-grade RTX 3090 GPU and an A100 GPU, with both achieving similar inference times of ~ 0.1 seconds.

4. RESULTS

Fig. 2 shows the model prediction for three typical WEST discharges #57948, #58387, #59490. The duration of these discharges ranges from approximately 10 to 70 seconds, and they involve 2 types of auxiliary heating. Under these conditions, our model can successfully reproduce the entire discharge pulse from the ramp-up, through the

TABLE 2. OUR MODEL HYPERPARAMETERS

Hyperparameter	Explanation	Best value
η_t	Learning rate	1×10^{-3}
num_heads	Number of heads	8
num_layers	Number of layers	2
Optimizer	Optimizer type	Stochastic Gradient Descent (SGD)
$\eta_{dropout}$	Dropout rate	0.1
B	Batch size	16
W	Window size	1024
S	Step size	512
Filter	Filter type	Simple moving average
W_{filter}	Window size of the filter	11

flat-top, to the ramp-down phases. However, there is a slight discrepancy in the predictions of q_0 and q_{95} . We also observe the same phenomenon in statistical plots as shown in Fig. 3. These quantities exhibit lower accuracy because our input signals do not properly contain the pressure profiles and kinetic measurements. Our goal is to reproduce the tokamak response using signals that can be specified before discharge, these measurements should not be included in the input signals. If q_0 and q_{95} are not properly constrained, they may be unreliable due to large variances, as the model struggles to reconstruct these parameters.

As shown in Fig. 3, we quantitatively evaluate the model performance on the entire test set, which consists of 110 discharges. For each discharge, the target and predicted values are averaged over its duration. Overall, the model demonstrates strong performance across most output parameters, except for q_0 and q_{95} , where prediction accuracy is relatively lower. A small number of discharges show poor predictive performance. Detailed analysis reveals that these cases are mainly associated with fluctuations in magnetic field coil currents, events that are uncommon and not part of standard tokamak operational planning. Such cases deviate from our objective of reproducing plasma responses in conventional discharges and providing reliable references for tokamak scenario design, as these fluctuations are typically unintended and fall outside the scope of standard design procedures.

In addition, some discharges correspond to rare operational scenarios, such as discharge #60147, which involves the use of two valid auxiliary heating methods and they are very rare in our dataset. This discharge occurs only three times out of 550 discharges. All of these cases are in the test set. The model's generalization ability for such unusual and rare discharges remains limited. Handling rare cases is an inherent challenge for machine learning methods. To improve performance in future work, we plan to include more discharges from this scenario to enhance the model's exposure and learning. We will also explore some data augmentation for the rare and fluctuating discharges, or use some physics-informed method to enhance the ML model's generalization ability.

5. CONCLUSION AND DISCUSSION

In this study, we proposed a transformer-based machine learning model for forecasting the behavior of tokamak plasmas based on discharge configurations. The model takes as input the auxiliary heating power, magnetic field coil currents, plasma current, and line-averaged plasma density. These signals are typically designed or programmed prior to the actual discharge, enabling the model to predict the plasma's response to the intended operational settings.

The current model focuses on predicting six key global plasma parameters: normalized beta (β_n), toroidal beta (β_t), poloidal beta (β_p), plasma stored energy (W_{mhd}), safety factor at the magnetic axis (q_0), and safety factor at the 95% flux surface (q_{95}). These quantities serve as essential indicators of plasma performance, stability, and confinement quality. The model demonstrates promising predictive performance and offers potential applications in discharge planning and fast scenario evaluation for tokamak experiments.

The present work can be extended to predict additional plasma parameters, such as the loop voltage (V_{loop}) and internal inductance (l_i) or 1D profiles such as density and temperature, which are also important for characterizing plasma behavior and optimizing control strategies. However, a significant challenge lies in obtaining sufficient and diverse data to properly constrain the model and ensure its generalization to a broader range of operational scenarios.

Moreover, purely data-driven models exhibit strong sensitivity to the specific conditions of the device. Consequently, directly applying the model after major hardware modifications or transferring it to a different tokamak remains an unresolved challenge. To address these limitations, future research can explore hybrid approaches that integrate physical knowledge or constraints into the model architecture or training process. Incorporating physics-

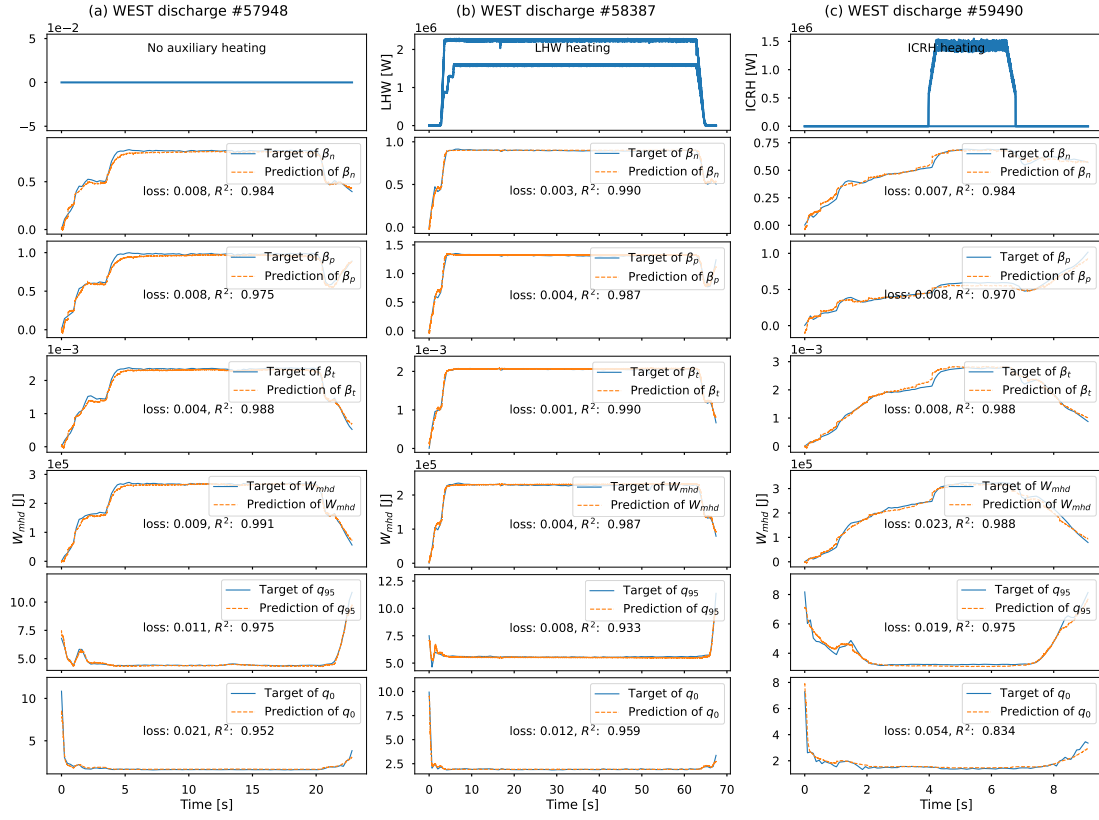


FIG. 2. Three types of typical discharge prediction in the test set. (a) The Ohmic heating-only discharge. (b) The discharge with LHW heating. (c) The discharge with ICRH heating. The targets represent experimental measurements, while the predictions correspond to the model estimations.

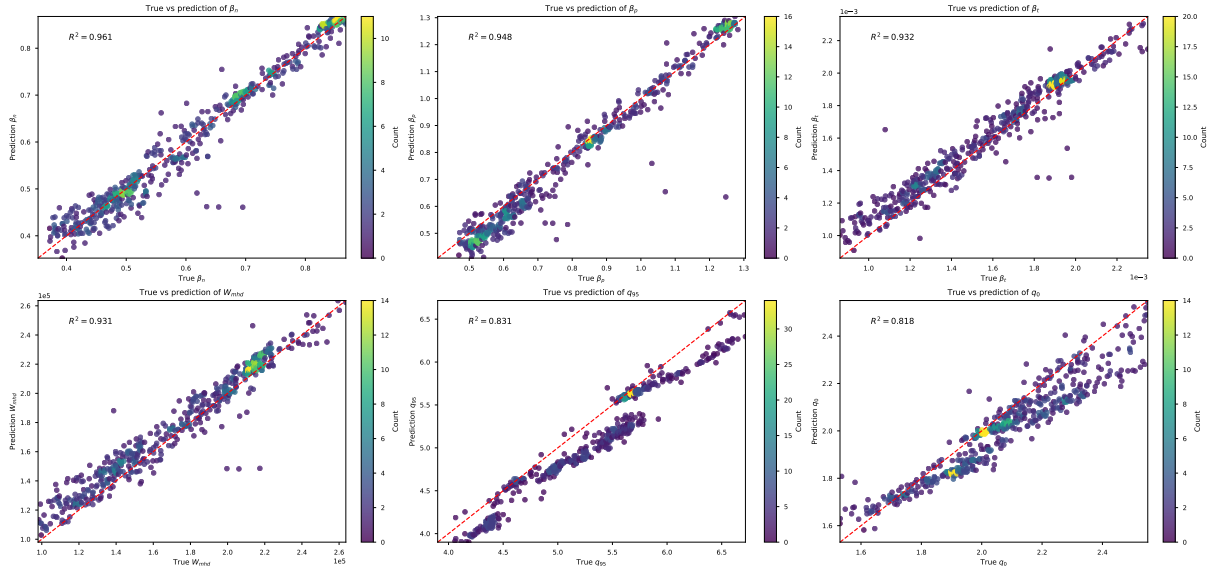


FIG. 3. Regression plots of the output signals. Both the target and predicted values are averaged over the duration of each discharge. Except for q_0 and q_{95} , the model achieves a coefficient of determination R^2 greater than 0.90. The performance for q_0 and q_{95} is slightly lower compared to the other signals.

informed methods has the potential to enhance the model's robustness and generalization capabilities, particularly for extrapolating to unfamiliar operational regimes.

Additionally, collecting datasets from multiple devices and adopting dimensionless representations for both input and output signals can further improve the cross-device applicability and generalization capability of the model. These approaches align with the broader goal of developing reliable, data-driven surrogate models that can support discharge design, scenario optimization, and real-time plasma control across present and future tokamak devices.

REFERENCES

- [1] G.L. Falchetto et al. "The European Integrated Tokamak Modelling (ITM) Effort: Achievements and First Physics Results". In: *Nuclear Fusion* 54.4 (Apr. 2014), p. 043018. ISSN: 0029-5515. DOI: 10.1088/0029-5515/54/4/043018.
- [2] R V Budny et al. "Predictions of {H}-Mode Performance in {ITER}". In: *Nuclear Fusion* 48.7 (2008), p. 75005. ISSN: 00295515. DOI: 10.1088/0029-5515/48/7/075005.
- [3] C E Kessel et al. "Long Pulse High Performance Plasma Scenario Development for the {National} {Spherical} {Torus} {Experiment}". In: *Physics of Plasmas* 13.5 (2006). ISSN: 1070664X. DOI: 10.1063/1.2177645.
- [4] Blaise Faugeras. "An Overview of the Numerical Methods for Tokamak Plasma Equilibrium Computation Implemented in the NICE Code". In: *Fusion Engineering and Design* 160 (Nov. 2020), p. 112020. ISSN: 09203796. DOI: 10.1016/j.fusengdes.2020.112020.
- [5] J.F. Artaud et al. "The CRONOS Suite of Codes for Integrated Tokamak Modelling". In: *Nuclear Fusion* 50.4 (Apr. 2010), p. 043001. ISSN: 0029-5515. DOI: 10.1088/0029-5515/50/4/043001.
- [6] Michele ROMANELLI et al. "JINTRAC: A System of Codes for Integrated Simulation of Tokamak Scenarios". In: *Plasma and Fusion Research* 9.SPECIALISSUE.2 (2014), pp. 3403023–3403023. ISSN: 1880-6821. DOI: 10.1585/pfr.9.3403023.
- [7] J F Artaud et al. "Metis: A Fast Integrated Tokamak Modelling Tool for Scenario Design". In: *Nuclear Fusion* 58.10 (Aug. 2018), p. 105001. ISSN: 17414326. DOI: 10.1088/1741-4326/aad5b1.
- [8] G V Pereverzew et al. "ASTRA. An Automatic System for Transport Analysis in a Tokamak." In: *IPP-Report IPP 5/98* (1991), p. 147.
- [9] F. Felici et al. "Real-Time Physics-Model-Based Simulation of the Current Density Profile in Tokamak Plasmas". In: *Nuclear Fusion* 51.8 (Aug. 2011), p. 083052. ISSN: 0029-5515, 1741-4326. DOI: 10.1088/0029-5515/51/8/083052.
- [10] N Hayashi and JET Team. "Advanced Tokamak Research with Integrated Modeling in {JT}-60 {Upgrade}". In: *Physics of Plasmas* 17.5 (2010). ISSN: 1070664X. DOI: 10.1063/1.3327917.
- [11] C. Rea et al. "Disruption Prediction Investigations Using Machine Learning Tools on DIII-D and Alcator C-Mod". In: *Plasma Physics and Controlled Fusion* 60.8 (2018). ISSN: 13616587. DOI: 10.1088/1361-6587/aac7fe.
- [12] Julian Kates-Harbeck, Alexey Svyatkovskiy, and William Tang. "Predicting Disruptive Instabilities in Controlled Fusion Plasmas through Deep Learning". In: *Nature* 568.7753 (Apr. 2019), pp. 526–531. ISSN: 0028-0836. DOI: 10.1038/s41586-019-1116-4.
- [13] Jonas Degraeve et al. "Magnetic Control of Tokamak Plasmas through Deep Reinforcement Learning". In: *Nature* 602.7897 (Feb. 2022), pp. 414–419. ISSN: 0028-0836. DOI: 10.1038/s41586-021-04301-9.
- [14] Chenguang Wan et al. "A High-Fidelity Surrogate Model for the Ion Temperature Gradient (ITG) Instability Using a Small Expensive Simulation Dataset". In: *Nuclear Fusion* 65.5 (Apr. 2025), p. 054001. ISSN: 0029-5515, 1741-4326. DOI: 10.1088/1741-4326/adc7c9.
- [15] Yunfei Ling et al. *PaMMA-Net: Plasmas Magnetic Measurement Evolution Based on Data-Driven Incremental Accumulative Prediction*. Jan. 2025. DOI: 10.48550/arXiv.2501.14003.
- [16] Chenguang Wan et al. "EAST Discharge Prediction without Integrating Simulation Results". In: *Nuclear Fusion* 62.12 (Dec. 2022), p. 126060. ISSN: 17414326. DOI: 10.1088/1741-4326/ac9c1a.
- [17] Chenguang Wan et al. "Experiment Data-Driven Modeling of Tokamak Discharge in EAST". In: *Nuclear Fusion* 61.6 (June 2021), p. 066015. ISSN: 0029-5515. DOI: 10.1088/1741-4326/abf419.

- [18] Ian Char et al. *Full Shot Predictions for the DIII-D Tokamak via Deep Recurrent Networks*. Apr. 2024. DOI: 10.48550/arXiv.2404.12416.
- [19] P. Moreau et al. “The New Magnetic Diagnostics in the WEST Tokamak”. In: *Review of Scientific Instruments* 89.10 (Oct. 2018), 10J109. ISSN: 0034-6748, 1089-7623. DOI: 10.1063/1.5036537.
- [20] Christophe Bouchand et al. “The Data Acquisition System of WEST’s New Thomson Scattering Diagnostics”. In: *IEEE Transactions on Nuclear Science* 72.3 (Mar. 2025), pp. 575–583. ISSN: 0018-9499, 1558-1578. DOI: 10.1109/TNS.2024.3470868.
- [21] O. Meyer et al. “Visible Spectroscopy Diagnostics for Tungsten Source Assessment in the WEST Tokamak: First Measurements”. In: *Review of Scientific Instruments* 89.10 (Oct. 2018), p. 10D105. ISSN: 0034-6748, 1089-7623. DOI: 10.1063/1.5035566.
- [22] C. Gil et al. “Renewal of the Interfero-Polarimeter Diagnostic for WEST”. In: *Fusion Engineering and Design* 140 (Mar. 2019), pp. 81–91. ISSN: 09203796. DOI: 10.1016/j.fusengdes.2019.02.003.
- [23] James Bergstra et al. “Algorithms for Hyper-Parameter Optimization”. In: *Advances in Neural Information Processing Systems*. Ed. by J Shawe-Taylor et al. Vol. 24. Granada, Spain: Curran Associates, Inc., 2011.