ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOST044

FORTUNE TELLING OR PHYSICS PREDICTION? DEEP LEARNING FOR ON-LINE KICKER TEMPERATURE FORECASTING

F. M. Velotti, * M.J. Barnes, B. Goddard, I. Revuelta
CERN, Geneva, Switzerland

Abstract

The injection kicker system MKP of the Super Proton Synchrotron SPS at CERN is composed of 4 kicker tanks. The MKP-L tank provides additional kick needed to inject 26 GeV Large Hadron Collider LHC 25 ns type beams. This device has been a limiting factor for operation with high intensity, due to the magnet's broadband beam coupling impedance and consequent beam induced heating. To optimise the usage of the SPS and avoid idle (kicker cooling) time, studies were conducted to develop a recurrent deep learning model that could predict the measured temperature evolution of the MKP-L, using the beam conditions and temperature history as input. In a second stage, the ferrite temperature is also estimated putting together the external temperature predictions from accurate thermo-mechanical simulations of the kicker magnet. In this paper, the methodology is described and details of the neural network architecture used, together with the implementation of an ad-hoc loss function, are given. The results applied to the SPS 2021 operational data are presented.

INTRODUCTION

The restart in 2021 after the Long Shutdown (LS) 2 of the CERN accelerator complex was the first year where all the upgrades of the LHC Injector Upgrade (LIU) project were deployed with the goal of achieving the High Luminosity (HL) LHC brightness requirements [1, 2]. The brightness increase is achieved by doubling the intensity per bunch of the beams and reducing the transverse emittances [3].

The SPS injection system comprises a septum system, MSI, and a kicker system, MKP. The kicker system is composed of four tanks, The first three are used to inject 14 GeV beams, and the last one, the so-called MKP-L, aids to inject the 26 GeV beam for LHC physics.

The MKP-L is one of the main limiting systems to the maximum storable beam intensity in the SPS. Due to beam induced heating via broadband coupling impedance, the MKP-L temperature rises at a rate much higher than all other MKP kickers, risking the ferrite to reach the Curie temperature and inducing significant out-gassing. This translates in reduced availability of the system, as machine operation has to stop to restore safe conditions. Such stops are in the order of many hours due to the large thermal inertia of the MKP-L kicker modules.

In order to optimise the machine time and to avoid idle time, we propose a data-driven model to estimate the MKP-L temperature readings starting from beam parameters and expected operational scenarios. The limitation of such an approach are the temperature probe locations, as they are not on the ferrite directly. In this paper we summarise the purely data-driven model and its application. We present the thermo-mechanical simulation studies that aim to model the full heat transfer from the beam induced power deposition in the ferrite, and finally we give an overview on how to combine the two approaches using PINN.

* francesco.maria.velotti@cern.ch

Brief Physical Model Description

The MKP-L showed a factor of three to four larger temperature increase with high intensity beam operation than the other MKP kickers [4]. This is due to the large real longitudinal beam impedance of the MKP-L. The beam induced power loss can be written as:

$$\Delta W = (f_0 e I_b N_b)^2 \sum_{k=-\infty}^{\infty} \left(|\Lambda(k\omega_0)|^2 \Re\left[Z_{\parallel}(k\omega_0)\right]\right) \quad (1)$$

where f_0 is the revolution frequency, $I_b N_b$ is total number of particles in the accelerator, $\Lambda(k\omega_0)$ is the normalised beam spectrum and $\Re\left[Z_{\parallel}(k\omega_0)\right]$ is the real part of the longitudinal impedance of the kicker. From simple considerations, the temperature variation due to ΔW power loss is [5]:

$$\frac{d}{dt}T = \frac{\Delta W}{F_{cool} C_{th}} \quad (2)$$

where C_{th} is the thermal capacitance and F_{cool} is the cooling factor. In order to extend the calculation to a different location in a kicker module, one considers a simplified model of the heat propagation:

$$\frac{\partial T}{\partial t} = \frac{k}{\rho C_p}\frac{\partial^2 T}{\partial x^2} + \frac{\Delta W(x,t)}{\rho C_p} \quad (3)$$

where k is the thermal conductivity, ρ is the density of the material and C_p is the specific heat capacity. Solving Eq. (1) and Eq.(3), one can compare data from the temperature probe, as measured in the SPS, with expected temperature of the same location in the MKP-L module.

LONG SHORT TERM MEMORY MODEL FOR TEMPERATURE PREDICTIONS

Deep neural networks are not designed to deal with time-series, where causality imposes well determined constraints. To solve this problem, the proposed architecture was the so-called recurrent neural network (RNN [6]) which exploits recursion to deal with sequences with time dependence. Due to practical issues [7], pure RNN are not very common anymore and they have been replaced with Long Short Term Memory (LSTM) networks [8].

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOST044

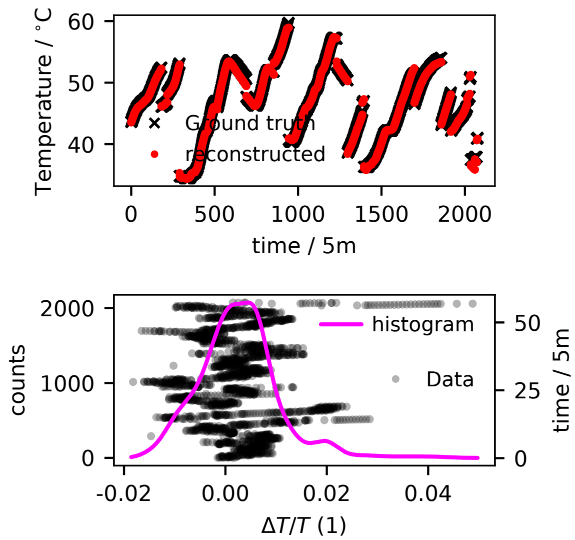

Figure 1: (Top) Temperature measured and reconstructed at the PT100 sensor using the described LSTM-based NN. (Bottom) Relative error of the reconstruction when predicting for the whole period length.

Classic time-series forecasting use multi-feature inputs from the past to forecast the future evolution of those quantities. In our case, and in most of particle accelerators problems, we are interested in feeding the model with exogenous inputs of the future time intervals to predict. This can be seen as the forcing term of our Initial Value Problem (IVP). Also, we are interested in predicting very far ahead in time and hence exposing the LSTM layers to their own noise is fundamental.

Data Structure and Model Architecture

In order to build a complete dataset the following beam characteristics were selected from the accelerator logging system: integrated intensity for 25 ns bunch spacing beams, mean and minimum bunch length and peak bunch intensity, all averaged along 5 min of machine operation. The quantities, together with the measured temperature, are arranged to form input sequences of 40 time steps long. The sequence to predict is finally 30 time steps long: $\hat{Y} = NN(X); X \in t(-40, 0]; \hat{Y} \in t[1, 30]$. The framework used to design and train the NN is *PyTorch* [9]. The network architecture that better suited the problem and the amount of data available for training is composed by two LSTM layers with 170 units, every layer is followed by a dropout layer with 50 % probability and finally a linear layer for the output prediction. The loss function used for training the NN weights is calculated comparing the whole output sequence and not just the single time step prediction.

PREDICTION OF MKP-L MEASURED TEMPERATURE

The data used for training and validation spans from 2015 to 2018 high intensity runs. All data were aggregated in periods, which were chosen in case of breaks of operation longer than 6 h. These periods are then used to test the NN predictions for sequences which are much longer than those used at training time.

The full dataset adds up to about 2500 data points which were split in 90% training and 10% validation. Such a large imbalance was chosen because the final testing is done on prediction of sequences longer than NN was trained on, as just discussed. There is a large difference in amount of data between heating and cooling due to the way data were collected - this is one of the main concerns for the result of the trained model.

The result of the forward pass of the NN on each of the periods available in the full dataset is shown in Fig. 1. Even for long periods, which extend more than five times the sequence length used in training, the NN successfully reconstructs the measured temperature. The error is less than 4 °C also for completely unseen data during training.

TOWARDS FERRITE TEMPERATURE ESTIMATION

The MKP-L injection kickers are of transmission line type, consisting of 22 cells per module [4]. A cell is constructed from a C-core ferrite sandwiched between HV plates. Ground plates interleave with the HV plates to form parallel plate capacitors. The thermal contact conductance (TCC) between ferrite and HV plates is an important parameter for the extraction of heat energy from the ferrite. As mentioned in the introduction, there is strong interest to predict the temperature of the ferrite so that SPS operation can be fine-tuned to avoid the ferrite reaching its Curie temperature, as there is a relatively long thermal time-constant for cool-down. Several MKP modules have PT100 temperature sensors mounted on a side-plate, which is at ground potential. The approach was, for a given beam and thus predicted power deposition in the ferrite yoke, model the non-linear beam induced power deposition [10] using ANSYS and compare the measured and predicted temperatures in the location of the PT100. A balance between precision and computational resources must be reached, as a large number of simulations are carried out to tune the parameters of the model. Thus, the model was kept as simple as feasible, neglecting all non-critical parts for thermal analysis. The non-linear beam induced power deposition, in the ferrite yokes, results in the first yokes to see beam being heated the most [10]: hence, the first three cells are modelled in ANSYS.

Building the Model: Tuning the Parameters and Difficulties

The mesh, especially at thermal contacts between surfaces of material (e.g. ferrite and HV plate), must be carefully defined, as the programme does not always define it correctly by default. Measurements to determine the emissivity of various surfaces and the TCC were carried out. For the TCC, the value was measured in air. However, in reality the TCC in vacuum is likely lower – but the measurement serves

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOST044

as reference for an initial value in the ANSYS simulations, to be subsequently tuned by iterating simulations. Before iterating the model, the influence of the TCC upon the rate of cooling of the ferrite, in a simplified model, was verified analytically.

During the iteration process, the relative influence of parameters was determined. In terms of heat radiation, small changes to the emissivity of the plates (HV and GND) resulted in the predicted temperature, at the position of the PT100, changing significantly. Deriving an accurate value of TCC between the HV plates and the ferrite yokes is a challenge: the temperature of the PT100 location is not sensitive to this TCC, whereas the temperature of the ferrite is sensitive (Fig. 2).

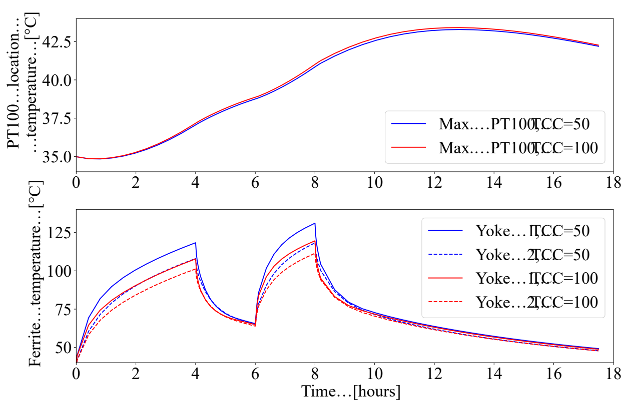

Figure 2: Influence of TCC between HV plate and ferrite. Top: temperature simulated by ANSYS at PT100 location. Bottom: temperature predicted for first two ferrite yokes.

Capacitive pickups (CPUs) are installed at the input and output of the kicker magnet: a CPU faces the HV input and output busbars, respectively. The electrical delay time of the module is determined from the measured CPU signals: the delay of each cell is proportional to the square root of the cell's inductance. If the ferrite is at the Curie temperature its relative permeability reduces to one [11]. Thus, the delay from the CPUs can be used as a diagnostic [12, 13] and provides input to the ANSYS simulations as it helps to set a minimum value for TCC between a ferrite and HV plates. Figure 3 shows a plot of the delay of a central cell (i.e. neglecting end effects) versus the real relative permeability of the ferrite of the cell. The delay depends upon this permeability and thus how close the ferrite temperature is to its Curie point: the imaginary relative permeability is also temperature dependent [11] and, thus, the beam induced losses will also decrease as the Curie temperature is approached. Hence, during SPS operation, the ferrite permeability may not reduce to one, but instead reach a state where the beam induced losses balance the radiated and conducted thermal energy.

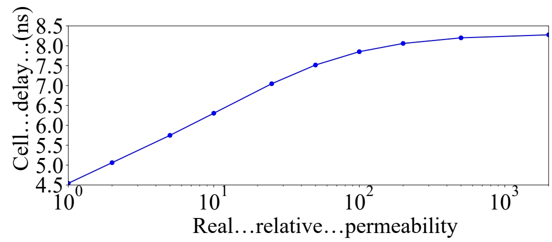

Figure 3: Cell delay versus relative permeability of ferrite.

PHYSICS INFORMED NEURAL NETWORK EXTENSION

The time-consuming ANSYS simulations could be replaced with a Physics Informed Neural Network (PINN), following the very promising results shown for both the NN and the pure ANSYS simulations. The idea is to use a very similar NN architecture to model the temperature prediction and then regularise the training using a physics-informed loss function. It will be composed by different terms: reconstruction loss of temperature, initial and boundary conditions (as simulated in a specific case with ANSYS) and the PDE loss. The specific heat capacity, the TCC and the other system-dependent terms will be either input from measurements or represented with dedicated NN. In this way, a complete surrogate model will be available to assess the ferrite and the PT100 temperature for well defined beam conditions. Also, once the model will be fully verified, an interlock directly on the predicted ferrite temperature could be envisaged.

CONCLUSION AND OUTLOOK

The MKP-L is one of the main limiting elements to store high intensity per bunch in the SPS due to its beam coupling impedance. To predict the temperature evolution of this system, an LSTM-based NN was trained using available data of high intensity SPS operation. Such a model was then used in operation to calculate the time needed for cool-down.

To estimate the actual ferrite temperature from the available measurements, a full ANSYS model of the MKP-L was developed and initial results were presented. The model is capable to precisely reproduce the measured temperature, but the uncertainty on the TCC translates in too many solutions available for the ferrite temperature evolution. To bridge between ANSYS simulations and the empirical NN, PINN is proposed. Work is still ongoing to finalise such a model. Dedicated measurements to help the estimation of TCC and other system parameters are foreseen, which will also be used to constraint and validate the PINN predictions.

The approach presented is very general and hence applications to other systems, such as LHC injection kicker MKI and SPS extraction septum ZS are being evaluated.

REFERENCES

[1] O. Aberle *et al.*, "High-Luminosity Large Hadron Collider (HL-LHC): Technical design report", CERN Yellow Reports:

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-TUPOST044

Monographs, Geneva: CERN, 2020.

[2] H. Damerau, A. Funken, R. Garoby, S. Gilardoni, B. Goddard, K. Hanke, A. Lombardi, D. Manglunki, M. Meddahi, B. Mikulec, G. Rumolo, E. Shaposhnikova, M. Vretenar, and J. Coupard, "LHC Injectors Upgrade, Technical Design Report", Dec. 2014.

[3] V. Kain *et al.*, "Achievements and Performance Prospects of the Upgraded LHC Injectors", presented at the IPAC'22, Bangkok, Thailand, Jun. 2022, paper WEIYGD1, this conference.

[4] M. J. Barnes *et al.*, "Mitigation of High Voltage Breakdown of the Beam Screen of a CERN SPS Injection Kicker Magnet", presented at the IPAC'22, Bangkok, Thailand, Jun. 2022, paper THPOTK043, this conference.

[5] C. Zannini, "Electromagnetic Simulation of CERN accelerator Components and Experimental Applications," Mar. 2013, Presented 15 Apr. 2013.

[6] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation," Tech. rep. ICS 8504, California Univ. San Diego La Jolla Inst. for Cognitive Science, San Diego, California, USA, 1985.

[7] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 6, pp. 107–116, 1998. doi:10.1142/S0218488598000094

[8] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997. doi:10.1162/neco.1997.9.8.1735

[9] A. Paszke, S. Gross *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, pp. 8024–8035, Curran Associates, Inc., 2019.

[10] M. J. Barnes, O. Bjorkqvist, and K. Kodama, "Beam Induced Power Deposition in CERN SPS Injection Kickers", in *Proc. IPAC'21*, Campinas, Brazil, May 2021, pp. 3490–3493. doi:10.18429/JACoW-IPAC2021-WEPAB342

[11] A. Chmielinska, B. Popovic, M. J. Barnes, F. Caspers, and C. Vollinger, "Measurements of electromagnetic properties of ferrites as a function of frequency and temperature," *Journal of Physics: Conference Series*, vol. 1067, p. 082018, 2018. doi:10.1088/1742-6596/1067/8/082018

[12] M. J. Barnes, L. Ducimetière, N. Garrel, B. Goddard, V. Mertens, and W. J. M. Weterings, "Analysis of Ferrite Heating of the LHC Injection Kickers and Proposals for Future Reduction of Temperature", in *Proc. IPAC'12*, New Orleans, LA, USA, May 2012, paper TUPPR090, pp. 2038–2040.

[13] A. Chmielinska, "Optimization of the Beam Screen for the FCC Injection Kicker Magnets," EPFL, Lausanne, France, 2019. doi:10.5075/epfl-thesis-8079