Electric Power Systems Research 252 (2026) 112408

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Simulation of a closed-loop dc-dc converter using a physics-informed neural network-based model

Marc-Antoine Coulombe [a,*], Maxime Berger [a], Antoine Lesage-Landry [b]

[a] Department of Mathematics, Informatics, and Engineering, Université du Québec à Rimouski, Rimouski, Québec, Canada
[b] Department of Electrical Engineering, Polytechnique Montréal, GERAD & Mila, Montréal, Québec, Canada

ARTICLE INFO

Keywords:
Boost converter
Machine learning
Modelling, neural network
Power converters
Time-domain simulation

ABSTRACT

The growing reliance on power electronics introduces new challenges requiring detailed time-domain analyses with fast and accurate circuit simulation tools. Currently, commercial time-domain simulation software are mainly relying on physics-based methods to simulate power electronics. Recent work showed that data-driven and physics-informed learning methods can increase simulation speed with limited compromise on accuracy, but many challenges remain before deployment in commercial tools can be possible. In this paper, we propose a physics-informed bidirectional long-short term memory neural network (BiLSTM-PINN) model to simulate the time-domain response of a closed-loop dc-dc boost converter for various operating points, parameters, and perturbations. A physics-informed fully-connected neural network (FCNN) and a BiLSTM are also trained to establish a comparison. The three methods are then compared using step-response tests to assess their performance and limitations in terms of accuracy. The results show that the BiLSTM-PINN and BiLSTM models outperform the FCNN model by more than 9 and 4.5 times, respectively, in terms of median RMSE. Their standard deviation values are more than 2.6 and 1.7 smaller than the FCNN's, making them also more consistent. Those results illustrate that the proposed BiLSTM-PINN is a potential alternative to other physics-based or data-driven methods for power electronics simulations.

1. Introduction

The shift toward renewable sources of energy has accelerated, bringing a growing reliance on power electronics converters in power systems [1]. Their presence increases power system complexity and introduces new challenges related to stability, reliability, and safety. This requires the use of detailed time-domain analyses utilizing offline and real-time simulation tools to predict their impacts on power systems [2]. Fast and accurate simulation of power electronics remains challenging due to their time-varying structure, nonlinear response, and switching dynamics [3]. Moreover, recent applications require the use of real-time models and simulations, such as model predictive control and digital twins, further increasing the demand for faster circuit simulation [4–6]. Commercially available electromagnetic transient (EMT) simulation software predominantly relies on physics-based methods to translate power electronics converters into equations that can be solved with classical numerical techniques [7]. Computation time reduction is generally done by substituting a time-varying with an equivalent fixed topology using switching functions [8] or by using averaging techniques [9,10]. Recent works have shown that data-driven methods, physics-informed learning-based methods, and optimization techniques, can significantly increase the simulation speed while providing high-accuracy predictions, thus complementing traditional computational techniques [11,12]. Data-driven and physics-informed learning-based methods are also valued for their ability to model commercial power electronics converters in a blackbox manner, removing the need for prior knowledge of the system [13,14]. These emerging approaches are still in their early stage, such that their performance requires deeper evaluation to make them suitable for reliable integration into commercial simulation tools.

Among the different machine learning-based modelling approaches used to model power electronics converters, neural network dominates, with primarily two architectures: feedforward neural networks (FNNs) [15–17] and recurrent neural networks (RNNs) [6,11,13]. In [15], a model with multiple fully-connected neural networks (FCNNs) is developed to model an IGBT's switching transients on a field-programmable gate array (FPGA) for real-time simulation. The authors achieved real-time simulation with a timestep of 5 ns. However, the

* Corresponding author.
E-mail addresses: marc-antoine.coulombe@uqar.ca (M.-A. Coulombe), maxime_berger@uqar.ca (M. Berger), antoine.lesage-landry@polymtl.ca (A. Lesage-Landry).

https://doi.org/10.1016/j.epsr.2025.112408
Received 6 January 2025; Received in revised form 18 March 2025; Accepted 17 October 2025
Available online 5 November 2025
0378-7796/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

proposed model requires 150 and 500 neural networks for simulating the turn-on and turn-off transient waveforms, respectively. In [16], the authors combine an FCNN with Bayesian regularization backpropagation to model the transient response and random forests for the steady-state response. Using random forests alongside neural networks increases the model stability and reduces its variance, as the results obtained by decision trees are easier to interpret. An FCNN is used in [17] to predict intermediate latent states between two observable states. Then, a physical model is used to predict the observable states using predictions from the intermediate latent states enabling them to integrate physics into their model. This approach requires solving multiple equations for each prediction using an implicit Runge-Kutta method, which can increase the computation time for complex models. In [13], a long-short term memory (LSTM) neural network is used to model a dc-dc buck converter. Although the LSTM appears to perform better than other FNN-based architectures for modelling the converter response, limited quantitative assements are provided to compare performance. Which is required for reliable deployment in simulation tools. In [11], the output signal is decomposed into a transient and a periodic component modelled, respectively, by an FNN and an LSTM combined with a convolutional neural network. This approach reduces the complexity of the problem into two sub-problems. Reference [6] uses a nonlinear autoregressive exogenous network, which yields good prediction accuracy. However, this type of architecture is limited in its ability to capture long-term dependencies without using internal memories like LSTMs. To consider a longer sequence of values, a large input size is required, which reduces computational efficiency.

While not directly applied to model power electronics converters, other data-driven approaches are also considered in the power systems literature, e.g., physics-informed neural ordinary differential equations for modelling power transformer's dynamic thermal behaviour [18] or for predicting the nonlinear transient dynamics of a inverter under fault [19] and graph-based learning for high-frequency filter design or microgrids voltage prediction [20]. These approaches are promising alternatives for future studies on power electronics modelling.

Among the aforementioned articles on power electronics simulations, the following limitations have been observed:

- Some proposed solutions are difficult to generalize to multiple types of converters and different topologies;
- The assessment of data-driven models is limited, i.e., with little to no quantitative metrics, thus, preventing their reliable deployment in simulation tools for power system analysis;
- Some proposed solutions require significant computational time, making them difficult to apply in real-time simulations or without specialized hardware.

The main contributions of this paper are (i) designing a physics-informed bidirectional LSTM (BiLSTM-PINN) network architecture and to model the time-domain response of a closed-loop dc-dc boost converter and (ii) thoroughly benchmarking our model using step-response tests. The BiLSTM architecture stands out as one well studied and documented, and easy to implement. It can also run on graphics processing units (GPUs) or FPGAs for computation acceleration [21,22], making it readily accessible and practical for deployment in simulation software [23,24]. This architecture also takes into account both past and future values to make predictions, in contrast to RNNs or standard LSTMs, which only consider past values or FCNNs that are limited to present values. This property improves performance by utilizing the initial and final operating point values derived from the dc steady-state model to better reproduce the converter response. The inclusion of a physics-informed loss also helps mitigate prediction errors by penalizing results that do not adhere to physical principles during training. Computational time evaluation is left for future work as many sources of interference such as the programming language (compiled MATLAB® *vs.* interpreted Python), the solver(s) used for the physical model or the hardware used, can make the results unreliable. This further underscores the importance of first defining the model architecture and assessing the model's accuracy, as it is done in this work. However, we consider that function evaluation with neural networks should be faster than solving a complex system of time-varying equations as hinted in [11].

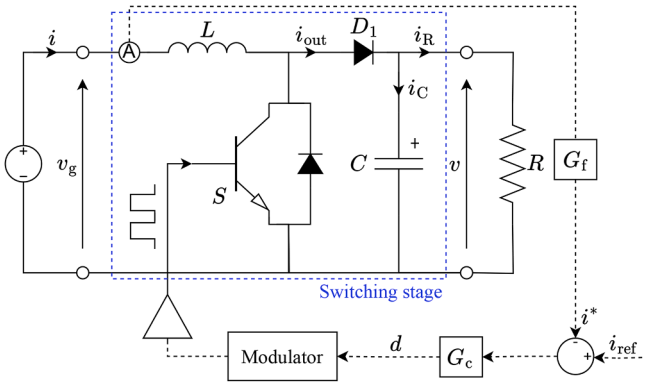


Fig. 1. Circuit topology of the current regulated boost converter.

In this paper, we train an FCNN, a BiLSTM, and a BiLSTM-PINN to model the input and output currents of a closed-loop dc-dc boost converter's large-signal averaged model. The validation of the FCNN and BiLSTM models with a detailed switching model is beyond the scope of this paper. By using an averaged model, we aim to reduce the complexity of the neural network, enabling a more generalizable methodology and reduced computation time. Additionally, this model helps us quantify the transient waveforms using common performance metrics. This model is utilized to generate a dataset using MATLAB® Simulink for various operating points, parameters, and perturbations. The dataset is then used to train an FCNN, a BiLSTM and, a BiLSTM-PINN, with hyperparameter tuning performed via `HyperOpt` [25]. Finally, the two architectures are compared using step-response tests to assess model performance.

The paper is organized as follows. Section II presents the dc-dc boost converter topology and the large-signal averaged model used to generate the training and testing datasets. Section 3 presents the physics-informed learning-based boost converter models. Section 4 discusses dataset generation and hyperparameter tuning, and validates the accuracy of our method via time-domain simulations in MATLAB® Simulink.

2. Boost converter model

The closed-loop boost converter topology used in this paper is shown in Fig. 1. The switch S is controlled by a pulse-width modulation (PWM) signal, introducing two distinct circuit states: switch S open and closed. The duty cycle $d(t)$ represents the proportion of time with the switch S closed during a complete period T_s, while the complementary duty cycle is given by $d'(t) = 1 - d(t)$. The input current i is the regulated variable. The transfer function G_f models the presence of a filter on the current measurement.

The circuit in Fig. 1 regulates the input current i to match the reference i_{ref} by calculating the error signal $\epsilon = i_{ref} - i^*$, where i^* is the filtered input current. The error signal is then fed into a PI controller, with a transfer function defined as:

$$G_c(s) = \frac{sK_p + K_i}{s},$$

where $K_p \in \mathbb{R}$ and $K_i \in \mathbb{R}$ are determined using the *algebra on the graph* technique [26] and further tuned to achieve the desired dynamic performance, with specified cross-over frequency $\omega_{\phi_m} = 2\pi f_{\phi_m}$ and phase-margin ϕ_m as explicited in Appendix. The filter and controller parameters for the converter model under validation are specified in Section 4.

In dc steady-state, the dc conversion ratio $M(D)$, the inductor's dc current I_L, the inductor's current ripple Δi_L and the capacitance's

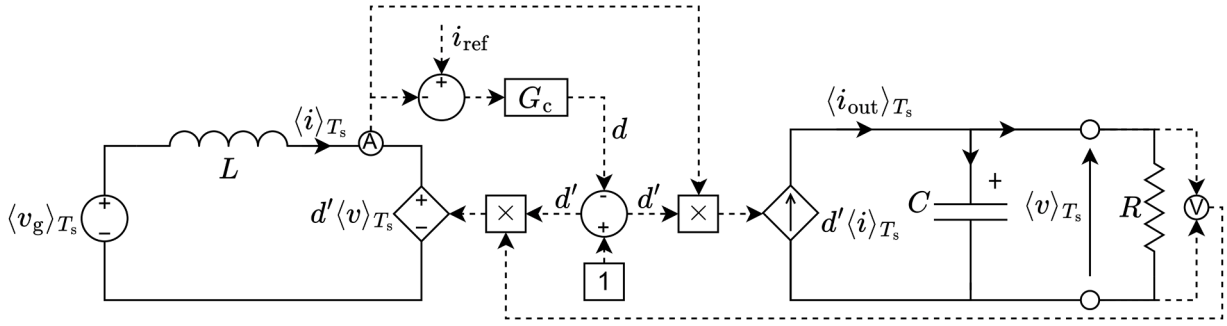


Fig. 2. Closed-loop large-signal averaged model for the boost converter.

voltage ripple Δv_C are derived as

$$M(D) = \frac{V}{V_g} = \frac{1}{1-D} \tag{1}$$

$$I_L = \frac{V}{(1-D)R} \tag{2}$$

$$\Delta i_L = \frac{DV_g}{2f_s L} \tag{3}$$

$$\Delta v_C = \frac{DV}{2f_s RC}, \tag{4}$$

where V_g and V are the dc input and output voltages, D is the duty cycle, and f_s is the switching frequency. We use (1)–(4) to describe the boost converter in initial and final steady-states which are used to initialize the BiLSTM models and the simulation parameters for the generation of the dataset in Sections 3 and 4, respectively.

We then use the classical dc-averaging technique for dc-dc converter [27] to simplify the switching stage (in blue colour in Fig. 1) leading to the following differential equations:

$$L \frac{\mathrm{d}\langle i \rangle_{T_s}}{\mathrm{d}t} = \langle v_g \rangle_{T_s} - d' \langle v \rangle_{T_s} \tag{5}$$

$$C \frac{\mathrm{d}\langle v \rangle_{T_s}}{\mathrm{d}t} = d' \langle i \rangle_{T_s} - \frac{\langle v \rangle_{T_s}}{R}, \tag{6}$$

where denotes $\langle \rangle_{T_s}$ the average value of a variable over T_s. The equivalent circuit derived from (5) and (6) is illustrated in Fig. 2. Our objective is to model the currents $\langle i \rangle_{T_s}$ and $\langle i_{\text{out}} \rangle_{T_s}$ in Fig. 2 with learning-based methods. The generation of the dataset to train learning-based models using this circuit is presented in Section 4.

3. Physics-informed learning-based boost converter models

We now develop our learning-based model of the boost converter. Let $\mathbf{x} \in \mathbb{R}^{18}$ and $\mathbf{y} \in \mathbb{R}^2$ be, respectively, the input (feature) and output (target) vector. We use the subscript $t \in \mathbb{N}$ to denote the time instance inputs/outputs are registered. We define the feature vector $\mathbf{x}_t$ as

$$\mathbf{x}_t = \text{vec} \begin{bmatrix} t & L & C \\ R & R_s & V_g \\ V(t=0) & \Delta v(t=0) & I_L(t=0) \\ \Delta i_L(t=0) & I_{\text{step}} & t_{\text{step}} \\ K_p & K_i & f_s \\ f_c & i_{\text{ref},t} & d_t \end{bmatrix}, \tag{7}$$

where vec is the vectorization operator. Various constants that characterize the converter are collected in (7), such as the inductance L, capacitance C, load resistance R, inductance resistance R_s, input voltage V_g, switching frequency f_s, and controller parameters K_i and K_p, as derived in Section 2. Next, additional features are used to define the initial state of the system, namely, the output voltage $V(t=0)$, the capacitance's voltage ripple $\Delta v(t=0)$, the inductor's current $I_L(t=0)$, and the inductor's current ripple $\Delta i_L(t=0)$. Then, time-varying features, viz., the desired current $i_{\text{ref},t}$ and the control signal d_t are included. Finally, I_{step} and t_{step} specify the amplitude and timing of the square disturbance on $i_{\text{ref},t}$.

In our setting, the entire system is assumed to be known, enabling the learning-based models to leverage comprehensive information about the circuit, its initial state, and its state at time t for accurate predictions. This assumption is compatible with settings encountered in power system simulation tools. Further explorations could be done to reduce the dimension of the feature vector $\mathbf{x}_t$ while minimizing the impact on accuracy. This is a topic for future work.

Next, we define the target vector at time t as

$$\mathbf{y}_t = \begin{bmatrix} \langle i \rangle_{T_s,t} & \langle i_{\text{out}} \rangle_{T_s,t} \end{bmatrix}.$$

The goal of the learning-based approach is to predict the input current $\langle i \rangle$ and the output current $\langle i_{\text{out}} \rangle$, derived using the previously defined large-signal averaged model. The decision to focus on current predictions is primarily to ensure signals operate on comparable time scales, hence simplifying the result analysis.

Let $\texttt{FCNN} : \mathbb{R}^{18} \to \mathbb{R}^2$ and $\texttt{BiLSTM} : \mathbb{R}^{18} \to \mathbb{R}^2$ be, respectively, the trained FCNN and BiLSTM with their corresponding collections of weights denoted by $\mathcal{W}_{\text{FCNN}}$ and $\mathcal{W}_{\text{BiLSTM}}$. We first model the converter with an FCNN, a common architecture in the machine learning literature. The input vector $\mathbf{x}_t$ is processed through the network, where a sequence of activation function and linear combination based weights $\mathbf{W} \in \mathcal{W}_{\text{FCNN}}$ compositions produces the prediction $\hat{\mathbf{y}}_{\text{FCNN},t}$ at the output layer. The FCNN is later used for benchmarking. We then utilize a BiLSTM to model the converter dynamics. To generate a prediction $\hat{\mathbf{y}}_{\text{BiLSTM},t}$, the BiLSTM network processes features from $\mathbf{x}_{t-k}$ to $\mathbf{x}_t$ and $\mathbf{x}_{t+k}$ to $\mathbf{x}_t$ through LSTM cells, as illustrated in Fig. 3. Here, $k \in \mathbb{N}$ denotes the sequence length, a hyperparameter that will be discussed in Section 4. The weights $\mathbf{W} \in \mathcal{W}_{\text{BiLSTM}}$ are the trainable parameters within each LSTM cell fully-connected layers. Because the LSTM cell follows a well-documented topology, we refer interested readers to, e.g., [13,28], for additional details. Finally, given observation $\mathbf{x}_t$, the resulting predictions $\hat{\mathbf{y}}_t$ are:

$$\hat{\mathbf{y}}_{\text{FCNN},t} = \texttt{FCNN}(\mathbf{x}_t; \mathcal{W}_{\text{FCNN}}) \tag{8}$$

$$\hat{\mathbf{y}}_{\text{BiLSTM},t} = \texttt{BiLSTM}(\mathbf{x}_t; \mathcal{W}_{\text{BiLSTM}}). \tag{9}$$

To achieve accurate predictions, the collections of weights $\mathcal{W}_{\text{FCNN}}$ and $\mathcal{W}_{\text{BiLSTM}}$ in (8) and (9) are computed to minimize the discrepancy between the predictions $\hat{\mathbf{y}}$ and the targets $\mathbf{y}$ through a loss function. Let the loss function be defined as

$$\mathcal{L} = \mathcal{L}_{\text{RMSE}} + \lambda \mathcal{L}_{\text{PBE}}, \tag{10}$$

where $\mathcal{L}$ combines the root-mean-squared error $\mathcal{L}_{\text{RMSE}}$ with the circuit power balance error $\mathcal{L}_{\text{PBE}}$ weighted by a constant $\lambda > 0$ which is to be tuned. The loss function is used to compute the gradient, which drives the backpropagation algorithm that trains both networks. By iteratively adjusting the collections of weights $\mathcal{W}_{\text{FCNN}}$ and $\mathcal{W}_{\text{BiLSTM}}$ using (10), the models progressively improve their accuracy. Specifically, the loss

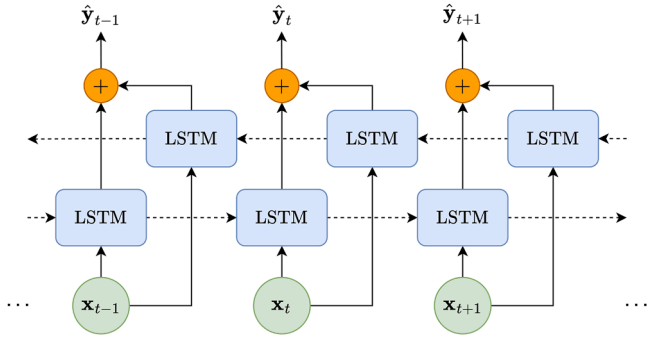


Fig. 3. Schematic of a bidirectional LSTM neural network.

function $\mathcal{L}_{\text{RMSE}}$ for a single curve is

$$\mathcal{L}_{\text{RMSE}} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\hat{\mathbf{y}}_t - \mathbf{y}_t\right)^2}, \tag{11}$$

where $n \in \mathbb{N}$ is the number of points on a curve related to the simulation duration and the step size. The power balance loss function $\mathcal{L}_{\text{PBE}}$ is then given by:

$$\mathcal{L}_{\text{PBE}} = \frac{1}{n}\sum_{t=1}^{n}\left|\langle P_{\text{in}}\rangle_{T_s,t} - \langle P_{\text{out}}\rangle_{T_s,t}\right|, \tag{12}$$

where the predicted currents $\hat{\mathbf{y}}_t$ are used to calculate, respectively, the predicted power entering and exiting the converter, defined as:

$$\langle P_{\text{in}}\rangle_{T_s,t} = d'_t\langle v\rangle_{T_s,t}\langle i\rangle_{T_s,t} \tag{13}$$

$$\langle P_{\text{out}}\rangle_{T_s,t} = \langle v\rangle_{T_s,t}\langle i_{\text{out}}\rangle_{T_s,t}\,. \tag{14}$$

In other words, we consider that the conservation of power applies while neglecting the converter inner losses. This assumption allows us to integrate physics into the BiLSTM loss function by penalizing predictions yielding a high-power imbalance between the converter's input and output, then leading to the BiLSTM-PINN (and FCNN) model.

4. Numerical results

We now present our numerical setting and then evaluate the performance of our learning-based models.

4.1. Numerical setting and training

To conduct the experiment, we first generate our dataset. This dataset is created by reproducing the lossless switching circuit in Fig. 1 and the large-signal averaged model shown in Fig. 2 in MATLAB® Simulink. The simulation parameters to create the dataset are provided in Table 1.

Key parameters of Fig. 2's circuit are defined, with their study range specified in a [lower bound, upper bound] format along with the step size for varying parameters such as C, L, R, V, and V_g. Other parameters, like f_s and R_s, remain constant for all simulations. During a simulation, a current step of amplitude I_{step} is applied to i_{ref} at time t_{step} based on a uniform distribution. The simulation itself is carried out from $t = 0$ to $t = 0.03$ with a fixed time step of 250×10^{-7}s.

In total, 565,950 operating points are simulated. Among these, cases that do not exhibit continuous conduction mode are excluded because they display response not accounted for by the large-signal averaged model. We also removed cases with an unfeasible PI controller (see Appendix). In addition, any cases where $|d| = 0.9$ are also rejected because the controller is programmed to saturate at this value, and saturation introduces non-linearities outside of the scope of this study. Approximately 10 % of the dataset is removed for these reasons. Additionally, the points corresponding to the first 0.01 second of each simulation are discarded to ensure the model reaches steady-state before introducing a perturbation. Finally, the dataset is split randomly using the holdout method with an 80%–20 % ratio.

Table 1
Simulations paramaters for the dataset generation.

Parameter	Interval	Step size
C (μF)	[1000, 7000]	1000
$d(t)$	± 0.9	–
f_c (kHz)	5	–
f_s (kHz)	70	–
f_{ϕ_m} (Hz)	255	–
I_{step} (A)	$\pm I_L(t=0)$	–
L (mH)	[1, 7]	1
R (Ω)	[1, 500]	10
R_s (Ω)	30×10^{-3}	–
V (V)	[200, 400]	10
V_g (V)	[100, 200]	10
t (s)	[0, 0.03]	250×10^{-7}
t_{step} (s)	[0.011, 0.021]	–
ϕ_m (°)	50	–

Table 2
Hyperparameters interval and tuned values.

Hyperparameter	Interval	FCNN	BiLSTM(-PINN)
Batch size	[128, 300]	294	207
Epochs	–	35	35
Learning rate	[1e^{-4}, 1e^{-2}]	6e^{-4}	1e^{-3}
Optimizer	–	ADAM	ADAM
Neurons/Cells (per hidden layer)	[100, 500]	304	165
Sequence length k	[10, 25]	–	25
Weights decay	[1e^{-5}, 1e^{-2}]	9e^{-3}	7e^{-4}
λ	[0, 1]	0.2	0 (0.2)

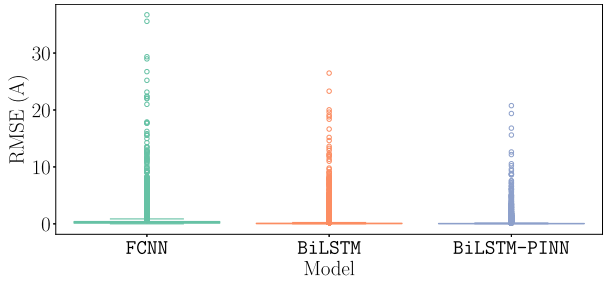
We identify suitable hyperparameters for our learning-based models using `HyperOpt` [25]. The resulting hyperparameters are presented in Table 2 together with the interval considered by `HyperOpt`. To define the intervals, preliminary trials were conducted to establish practical ranges, avoiding underfitting or overfitting. These issues were identified through an observable increase in the loss function $\mathcal{L}$. Subsequently, 30 optimization trials were performed using the tree-structured Parzen estimator (TPE) algorithm, with the objective of minimizing $\mathcal{L}$. The resulting values were rounded for simplicity. We chose the `ADAM` optimizer for all trials as it consistently provided better results compared to the alternatives [29]. From the hyperparameters described in Table 2, the FCNN, BiLSTM, and BiLSTM-PINN models are trained on 80 % of the dataset. Because our model is trained offline, we train multiple models simultaneously and select the one that delivers the best testing RMSE. Although this approach does not fully address sensitivity issues, it helps mitigate their impact by choosing the best-performing empirical model for deployment. Robustness of the model, the stability of outputs across different inputs, still remains to be established similarly to many of the neural network models. This is a topic for future work.

4.2. Test

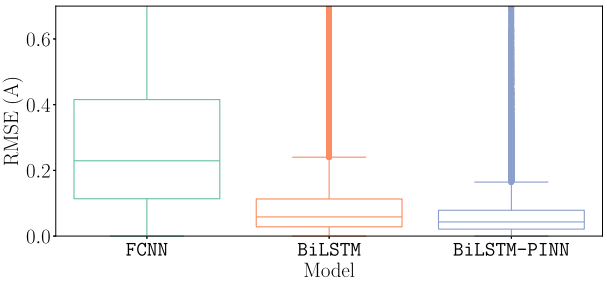
We then use the remaining 20 % of the dataset for out-of-sample testing purposes. For each predicted response curves, we calculate the RMSE using (11). The resulting box plot is illustrated at Fig. 4.

Fig. 4 presents the RMSE for the FCNN, BiLSTM, and BiLSTM-PINN models. From Fig. 4a, we can observe that some operating point's RMSE greatly diverges from the median value. This is especially the case for the FCNN. Fig. 4b further zooms on the box plot distribution. We remark that the BiLSTM-PINN outperforms the FCNN and the BiLSTM by achieving a lower median $\mathcal{L}_{\text{RMSE}}$ and distribution dispersion.

Next, we quantified the model prediction performance using the `stepinfo` function from MATLAB®. We refer readers to their



(a) RMSE distribution at the FCNN scale. Circles represent the outliers of each model on the test dataset.



(b) Quartiles comparison for both learning-based models representing quartiles and median values. The lowest and highest delimitations of the box are the 25^{th} and 75^{th} percentile, respectively. The vertical line inside the box represents the median and circles (appearing as a bold line) are the RMSE distribution outliers.

Fig. 4. RMSE box plot for the FCNN, BiLSTM, and BiLSTM-PINN models.

documentation for more details on the function and the evaluated metrics [30]. Table 3 compares all performance metrics for the FCNN and BiLSTM models.

Recall that $\mathbf{y}_t$ ($\hat{\mathbf{y}}_t$) consists of the input (predicted) current $\langle i \rangle_{T_s,t}$ ($\langle \hat{i} \rangle_{T_s,t}$) and the (predicted) output $\langle i_{\text{out}} \rangle_{T_s,t}$ ($\langle \hat{i}_{\text{out}} \rangle_{T_s,t}$). Let the total absolute prediction error be

$$e = e_{\text{input}} + e_{\text{output}},$$

where e_{input} and e_{output} are defined as the absolute input and output error, respectively, given by

$$e_{\text{input}} = \left| \langle \hat{i} \rangle_{T_s,t} - \langle i \rangle_{T_s,t} \right|$$

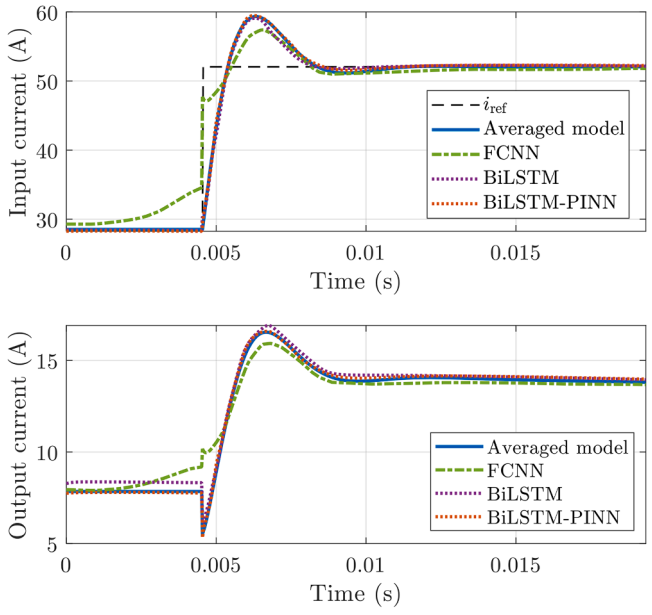
$$e_{\text{output}} = \left| \langle \hat{i}_{\text{out}} \rangle_{T_s,t} - \langle i_{\text{out}} \rangle_{T_s,t} \right|.$$

For each `stepinfo` characteristics, we report the mean, the standard deviation σ_{std}, and the median. Table 3 compares all performance metrics for the FCNN, BiLSTM, and BiLSTM-PINN models. Table 3 illustrates the superiority of BiLSTM-PINN over FCNN and BiLSTM on most metrics in terms of mean(e) and median(e). We remark that $\sigma_{\text{std}}(e)$ is often bigger than mean(e), demonstrating that outliers are, skewing the mean, similarly to the RMSE analysis in Fig. 4. In this case, the median might be more informative as an overall evaluation of the performance. Looking at the RMSE medians, we observe that the BiLSTM-PINN has a median error of 0.0256 A (≈ 9 times lower than the FCNN) in comparison to 0.0509 A (≈ 4.5 times lower than the FCNN), and 0.2311 A for the BiLSTM and the FCNN, respectively. We similarly notice that standard deviation values are 0.2782 A, 0.4236 A, and 0.7367 A for the BiLSTM-PINN, BiLSTM, and FCNN. The BiLSTM-PINN and BiLSTM standard deviation are approximately 2.6 and 1.7 times smaller than the FCNN, making the predictions not only generally more accurate, but also more consistent. Those results are consistent with [13], which showed that recurrent architecture tends to outperform non-recurrent ones. However, unlike [11], our physics-informed FCNN cannot reproduce as accurately the waveform when using $\mathcal{L}_{\text{PBE}}$ to integrate physics to the loss. Over all metrics, the BiLSTM-PINN main limitations are the overshoot and the undershoot. To better understand the problem, we extract four distinct cases from the test dataset, two with average performance and two with a high overshooting or undershooting error and plot their responses. Figs. 5 and 6, respectively, represent the two cases with average performance and with a high overshooting or undershooting error. The BiLSTM-PINN tends to follow accurately the averaged model in both cases while the FCNN exhibits some inaccuracies, e.g., Fig. 6. We notice that overshooting or undershooting error occurs when there is a small perturbations on i_{ref} and the final value is close to zero as observed in Fig. 6. These cases tend to produce outliers in terms of overshooting and undershooting errors given the division to a value close to zero, skewing the distribution and making the mean and standard deviation less representative. We omitted these values from Table 3 for this reason, keeping only the median values as more insightful in this case. This phenomenon is also amplified by the minimum and maximum values on $I_{\text{L}}(t = 0)$ in our dataset. For perturbations on i_{ref} close to $|I_{\text{L}}(t = 0)|$, a reduced number of operating points exist for training in our dataset due to the use of a uniform distribution, thus increasing the error of our model in these cases. In sum, the results illustrate that

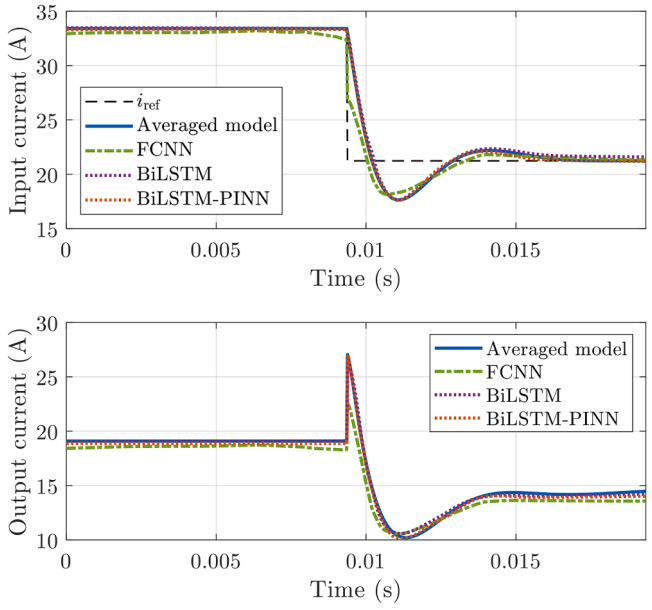
Table 3
Learning-based model performance metric comparison.

Metric	FCNN mean(e) ± $\sigma_{\text{std}}(e)$ median(e)	BiLSTM mean(e) ± $\sigma_{\text{std}}(e)$ median(e)	BiLSTM-PINN mean(e) ± $\sigma_{\text{std}}(e)$ median(e)
	0.3793 ± 0.7367	0.1161 ± 0.4236	0.0593 ± 0.2782
$\mathcal{L}_{\text{RMSE}}$ (A)	0.2311	0.0509	0.0256
	0.0004 ± 0.0009	7.7094e-5 ± 0.0004	3.5621e-5 ± 0.0003
Rise time (s)	0.0000e-5	0.0000e-5	0.0000e-5
	0.0036 ± 0.0037	0.0034 ± 0.0032	0.0022 ± 0.0021
Transient time (s)	0.0022	0.0024	0.0015
	0.0020 ± 0.0026	0.0016 ± 0.0023	0.0019 ± 0.0031
Settling time (s)	0.0010	0.0007	0.0003
	0.3131 ± 0.9630	0.1205 ± 1.5329	0.0685 ± 0.9676
Settling minimum (A)	0.1892	0.0438	0.0281
	0.5400 ± 1.6805	0.0747 ± 0.3949	0.1131 ± 0.8193
Settling maximum (A)	0.2034	0.0277	0.0299
	–	–	–
Overshoot (%)	7.4707	1.8219	1.7154
	–	–	–
Undershoot (%)	0.0000e-5	0.0000e-5	0.0000e-5
	0.5373 ± 1.6714	0.0741 ± 0.3933	0.1100 ± 0.7640
Peak (A)	0.2034	0.0277	0.0299
	0.0027 ± 0.0031	0.0017 ± 0.0029	0.0029 ± 0.0037
Peak time (s)	0.0008	0.0002	0.0005

our models may achieve good performance inside the parameter intervals defined in Table 2, but lose accuracy when operating closer to their minimum and maximum values due to having fewer available operating training points.

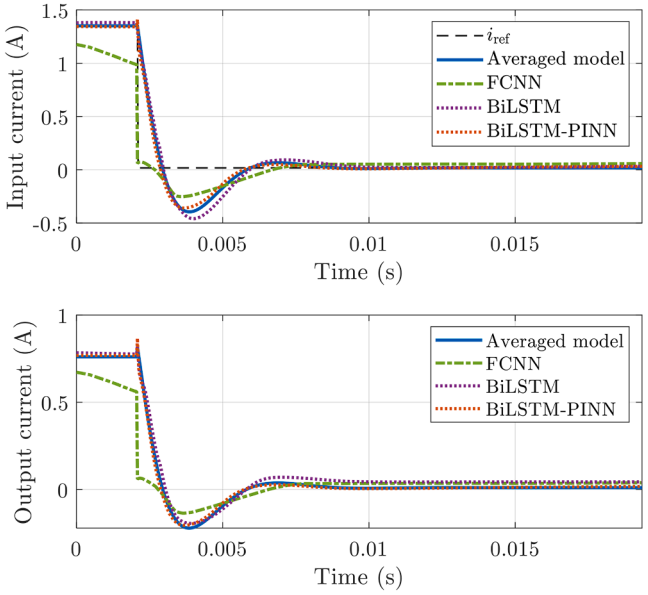


(a) Input and output current signals for ≈ 83% increase in i_{ref}.

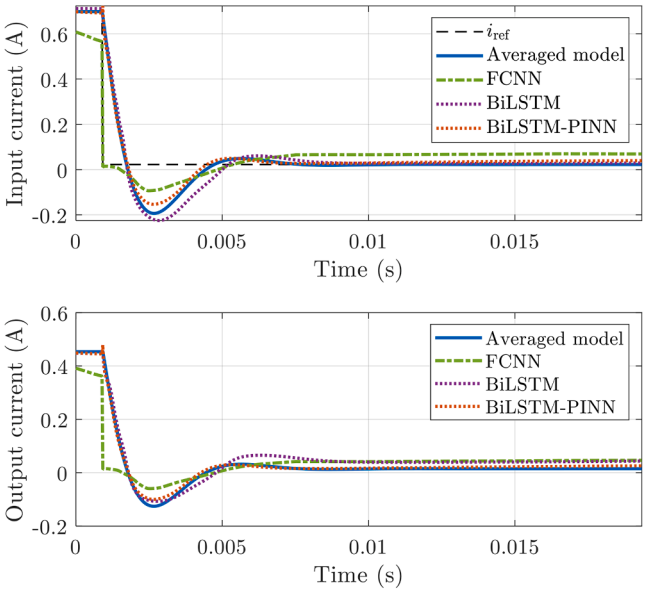


(b) Input and output current signals for ≈ 36% decrease in i_{ref}.

Fig. 5. Predictions for two operating points achieving close to or lower median performance in all metrics evaluated in Table 3.



(a) Input and output currents signals for ≈ 99% decrease on i_{ref}.



(b) Input and output currents signals for ≈ 98% decrease on i_{ref}.

Fig. 6. Predictions for two operating points resulting in some of the highest overshooting error (a) or undershooting error (b) among all operating points tested in Table 3 (outliers).

5. Conclusions

This work proposes a new physics-informed BiLSTM method to model power electronics converters. We evaluate the proposed BiLSTM-PINN by generating a dataset using the steady-state equations of a dc-dc boost converter along with a dc-averaged closed-loop model of the converter in MATLAB® Simulink. We also train a physics-informed FCNN and a data-driven BiLSTM to establish a performance comparison with a conventional non-recurrent neural network architecture. For both architectures, we tune the hyperparameters using `HyperOpt`. We then compare both methods by visualizing the RMSE distribution and quantifying their accuracy for various metrics using the `stepinfo` function from MATLAB®.

Our numerical study illustrates the ability of learning-based models to predict the response of power electronics converters. The BiLSTM-PINN model offers improved performance over the FCNN's and the data-driven BiLSTM and may be a good alternative to classical methods in some applications as a means of simplifying and accelerating real-time simulations. These results can improve the capabilities of commercial EMT simulation software by allowing the simulation of large-scale power systems through simplification of certain complex components, and for blackbox modelling of off-the-shelf converters.

Several limitations must be addressed before deployment in practical settings. Notably, the proposed model is based on an averaged large-signal representation, which overlooks high-frequency behaviour

in predicted signals and neglects non-idealities from diode D_1 and switch S. Moreover, the dataset relies on an unsaturated controller, limiting the validity of the model by excluding non-linearities in the control. However, our observations indicate that omitting K_p and K_i from the feature vector $\mathbf{x}_t$ did not significantly impact the performance metrics. This suggests that other controllers could be included into the dataset, as only the control signal remains. While we are confident that this methodology can be applied to other simple dc-dc topologies, such as the buck converter [11,13], extending it to more complex topologies is a topic of future work.

Future directions will focus on benchmarking the proposed method against physics-based approaches in terms of computational efficiency to quantify the potential gain in simulation speed. Expanding the framework to include other converter operating modes, such as protective states and fault-tolerant operation, is also a key direction to improve model generalization and extend its applications.

CRediT authorship contribution statement

Marc-Antoine Coulombe: Writing – original draft, Validation, Investigation, Conceptualization, Formal analysis, Visualization, Methodology; **Maxime Berger:** Supervision, Resources, Funding acquisition, Writing – review & editing, Software, Project administration, Conceptualization; **Antoine Lesage-Landry:** Project administration, Conceptualization, Writing – review & editing, Software, Funding acquisition, Supervision, Resources.

Data availability

The authors do not have permission to share data.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was funded by Mitacs, the Natural Sciences and Engineering Research Council of Canada (NSERC), and OPAL-RT Technologies under Grants ALLRP593314 & IT40898.

Appendix

We determine K_p and K_i using the *algebra on the graph* technique from [26]. To do so, we first set the cross-over frequency $\omega_{\phi_m} = 2\pi f_{\phi_m}$ where $f_{\phi_m} = 255$ Hz (more than 200 times smaller than the switching frequency f_s) and the phase-margin $\phi_m = 50°$ as a arbitrary compromise between stability margin, overshoot, and settling time. From those desired values, we first calculate the phase and magnitude of $G_c(s)$ given by $\angle G_c(\omega_{\phi_m}) = -(\angle G(\omega_{\phi_m}) + (180° - \phi_m))$ and $|G_c(\omega_{\phi_m})|_{dB} = -|G(\omega_{\phi_m})|_{dB}$. We then find the ratio α and compute K'_p of an intermediate controller with $K'_i = 1$ as

$$\alpha = \frac{K_i}{K_p} = \frac{\omega_{\phi_m}}{\tan\left(\angle G_c\left(\omega_{\phi_m}\right) + 90°\right)}$$

$$K'_p = \frac{1}{\alpha}.$$

We finally calculate the magnitude of $G'_c(s)$ to find K_p and K_i, and verify if a PI-controller with those parameters is feasible:

$$|G'_c\left(\omega_{\phi_m}\right)|_{dB} = 20\log\left[\frac{\sqrt{\left(\omega_{\phi_m} K'_p\right)^2 + \left(K'_i\right)^2}}{\omega_{\phi_m}}\right]$$

$$K_i = 10^{-\left|G\left(\omega_{\phi_m}\right)\right|_{dB} - \left|G'_c\left(\omega_{\phi_m}\right)\right|_{dB}/20}$$

$$K_p = \frac{K_i}{\alpha}.$$

References

[1] S. Peyghami, F. Blaabjerg, P. Palensky, Incorporating power electronic converters reliability into modern power system reliability analysis, IEEE J. Emerg. Sel. Top. Power Electron. 9 (2) (2021) 1668–1681. https://doi.org/10.1109/JESTPE.2020.2967216

[2] S. Peyghami, P. Palensky, F. Blaabjerg, An overview on the reliability of modern power electronic based power systems, IEEE Open J. Power Electron. 1 (2020) 34–50. https://doi.org/10.1109/OJPEL.2020.2973926

[3] J.A. Martinez-Velasco, Transient Analysis of Power Systems: Solution Techniques, Tools and Applications, John Wiley & Sons, 2014.

[4] S. Vazquez, J.I. Leon, L.G. Franquelo, J. Rodriguez, H.A. Young, A. Marquez, P. Zanchetta, Model predictive control: a review of its applications in power electronics, IEEE Ind. Electron. Mag. 8 (1) (2014) 16–31. https://doi.org/10.1109/MIE.2013.2290138

[5] F. Tao, H. Zhang, A. Liu, A.Y.C. Nee, Digital twin in industry: state-of-the-art, IEEE Trans. Ind. Inf. 15 (4) (2019) 2405–2415. https://doi.org/10.1109/TII.2018.2873186

[6] A. Wunderlich, E. Santi, Digital twin models of power electronic converters using dynamic neural networks, in: 2021 IEEE Applied Power Electronics Conference and Exposition (APEC), 2021, pp. 2369–2376. https://doi.org/10.1109/APEC42165.2021.9487201

[7] M. Berger, I. Kocar, H. Fortin-Blanchette, C. Lavertu, Large-signal modeling of three-phase dual active bridge converters for electromagnetic transient analysis in DC grids, J. Mod Power Syst. Clean Energy 7 (6) (2019) 1684–1696.

[8] L. Salazar, G. Joos, PSPICE Simulation of three-phase inverters by means of switching functions, IEEE Trans. Power Electron. 9 (1) (1994) 35–42.

[9] R.D. Middlebrook, S. Cuk, A general unified approach to modelling switching-converter power stages, in: 1976 IEEE Power Electronics Specialists Conference, 1976, pp. 18–34.

[10] S.R. Sanders, J.M. Noworolski, X.Z. Liu, G.C. Verghese, Generalized averaging method for power conversion circuits, IEEE Trans. Power Electron. 6 (2) (1991) 251–259.

[11] H. Ge, Z. Huang, Z. Huang, Modeling methodology based on fast and refined neural networks for non-isolated DC-DC converters with configurable parameter settings, IEEE J. Emerging Sel. Top. Circuits Syst. 13 (2) (2023) 617–628. https://doi.org/10.1109/JETCAS.2023.3251692

[12] G. Rojas-Dueñas, J.-R. Riba, M. Moreno-Eguilaz, Nonlinear least squares optimization for parametric identification of DC–DC converters, IEEE Trans. Power Electron. 36 (1) (2020) 654–661.

[13] P. Qashqai, K. Al-Haddad, R. Zgheib, Modeling power electronic converters using a method based on long-short term memory (LSTM) networks, in: IECON 2020 the 46Th Annual Conference of the IEEE Industrial Electronics Society, 2020, pp. 4697–4702. https://doi.org/10.1109/IECON43393.2020.9255375

[14] G. Rojas-Dueñas, J.-R. Riba, K. Kahalerras, M. Moreno-Eguilaz, A. Kadechkar, A. Gomez-Pau, Black-box modelling of a DC-DC buck converter based on a recurrent neural network, in: 2020 IEEE International Conference on Industrial Technology (ICIT), 2020, pp. 456–461. https://doi.org/10.1109/ICIT45562.2020.9067098

[15] Q. Li, H. Bai, E. Breaz, R. Roche, F. Gao, ANN-Aided Data-Driven IGBT switching transient modeling approach for FPGA-based real-time simulation of power converters, IEEE Trans. Transp. Electrif. 9 (1) (2023) 1166–1177. https://doi.org/10.1109/TTE.2022.3201656

[16] H.S. Krishnamoorthy, T. Narayanan Aayer, Machine learning based modeling of power electronic converters, in: 2019 IEEE Energy Conversion Congress and Exposition (ECCE), 2019, pp. 666–672. https://doi.org/10.1109/ECCE.2019.8912608

[17] S. Zhao, Y. Peng, Y. Zhang, H. Wang, Parameter estimation of power electronic converters with physics-informed machine learning, IEEE Trans. Power Electron. 37 (10) (2022) 11567–11578. https://doi.org/10.1109/TPEL.2022.3176468

[18] F. Bragone, K. Morozovska, P. Hilber, T. Laneryd, M. Luvisotto, Physics-informed neural networks for modelling power transformer's dynamic thermal behaviour, Electr. Power Syst. Res. 211 (2022) 108447. https://doi.org/{https://doi.org/10.1016/j.epsr.2022.108447}

[19] R. Nellikkath, I. Murzakhanov, S. Chatzivasileiadis, A. Venzke, M.K. Bakhshizadeh, Physics-informed neural networks for phase locked loop transient stability assessment, Electr. Power Syst. Res. 236 (2024) 110790. https://doi.org/https://doi.org/10.1016/j.epsr.2024.110790

[20] Y. Li, C. Xue, F. Zargari, Y.R. Li, From graph theory to graph neural networks (GNNs): the opportunities of GNNs in power electronics, IEEE Access 11 (2023) 145067–145084.

[21] D. Danopoulos, I. Stamoulias, G. Lentaris, D. Masouros, I. Kanaropoulos, A.K. Kakolyris, D. Soudris, LSTM Acceleration with FPGA and GPU devices for edge computing applications in B5G MEC, in: Embedded Computer Systems: Architectures, Modeling, and Simulation, Springer International Publishing, Cham, 2022, pp. 406–419.

[22] S.D. Nagarale, B.P. Patil, Accelerating AI-Based battery management system's SOC and SOH on FPGA, Appl. Comput. Intell. Soft Comput. 2023 (1) (2023) 2060808.

[23] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, Pytorch: an imperative style, high-performance deep learning library, Adv. Neural Inf. Process. Syst. 32 (2019).

[24] V. Rybalkin, A. Pappalardo, M.M. Ghaffar, G. Gambardella, N. Wehn, M. Blott, FINN-L: library extensions and design trade-off analysis for variable precision LSTM networks on FPGAs, in: 2018 28th International Conference on Field Programmable Logic and Applications (FPL), 2018, pp. 89–897. https://doi.org/10.1109/FPL.2018.00024

[25] J. Bergstra, D. Yamins, D. Cox, Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures, in: International Conference on Machine Learning, PMLR, 2013, pp. 115–123.

[26] D. Maksimovic, A.M. Stankovic, V.J. Thottuvelil, G.C. Verghese, Modeling and simulation of power electronic converters, Proc. IEEE 89 (6) (2001) 898–912. https://doi.org/10.1109/5.931486

[27] R.W. Erickson, D. Maksimovic, Fundamentals of Power Electronics, Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-43881-4_1

[28] G. Van Houdt, C. Mosquera, G. Nápoles, A review on the long short-term memory model, Artif. Intell. Rev. 53 (8) (2020) 5929–5955.

[29] D.P. Kingma, J.L. Ba, Adam: a method for stochastic gradient descent, in: ICLR: International Conference on Learning Representations, ICLR US., 2015, pp. 1–15.

[30] I. The MathWorks, MATLAB version: 24.1.0.2689473 (R2024a) Update 6, 2024, https://www.mathworks.com.