Superconductivity 3 (2022) 100017

Contents lists available at ScienceDirect

Superconductivity

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

A statistical model for the design of rotary HTS flux pumps based on deep-learning neuron network

Zezhao Wen [a], Hongye Zhang [b,*], Mengyuan Tian [c], Francesco Grilli [d], Markus Mueller [a,*]

[a] School of Engineering, University of Edinburgh, Edinburgh EH9 3JL, United Kingdom
[b] School of Engineering, University of Manchester, Manchester M1 3BB, United Kingdom
[c] Engineering Department, University of Cambridge, Cambridge CB3 0FA, United Kingdom
[d] Institute for Technical Physics, Karlsruhe Institute of Technology, Karlsruhe 76131, Germany

ARTICLE INFO

Keywords:
HTS dynamo
Flux pump
Deep learning
H-A formulation
Output characteristics

ABSTRACT

Rotory high temperature superconducting (HTS) flux pumps can consistently generate a DC voltage by rotating magnets over superconducting tapes, and thus energize the circuit if a closed loop is formed. The voltage output is a crucial factor to reflect the performance of such an HTS flux pump, which is determined by a set of design specifications, and some of them have been investigated extensively in the current literature. However, no work has been done yet to study the HTS dynamo output voltage by efficiently integrating all the design parameters together. In this paper, a well-trained deep-learning neuron network (DNN) with back-propagation algorithms has been put forward and validated. The proposed DNN is capable of quantifying the output voltage of an HTS dynamo instantly with an overall accuracy of approximately 98% with respect to the simulated values with all design parameters explicitly specified. The model possesses a powerful ability to characterize the output behavior of HTS dynamos by considering multiple design parameters, e.g., airgap, superconductor tape width, operating frequency, remanent flux density, rotor radius, and permanent magnet width, which have covered all the typical design considerations. The output characteristics of an HTS dynamo against each of the design parameters have been successfully demonstrated using this model. Compared to conventional time-consuming finite element method (FEM) based numerical models, the proposed DNN model has the advantages of automatic learning, fast computation, as well as strong programmability. Therefore the DNN model can greatly facilitate the design and optimization process for HTS dynamos. An executable application has been developed accordingly based on the DNN model, which is believed to provide a useful tool for learners and designers of HTS dynamos.

1. Introduction

In order to develop large capacity and high-power density electric machines, strong magnetic fields are required from the rotor coils [1–4]. High temperature superconducting (HTS) machines have seen continuously increasing interest, given that HTS rotor coils are capable of generating ultra-high magnetic fields within limited space [5] and even demonstrating improved electrical stability compared to conventional machines [6]. However, due to the existence of flux creep [7] and substantial AC losses [8–10], it is challenging for HTS magnets to run in a self-maintained persistent current mode. The traditional approach to inject current into the rotor coils involves at least two current leads, which transport current from ambient temperature to the cryogenic environment. This imposes a considerable thermal load upon the cryogenic system and results in significant additional capital and operating cost [11].

Flux pumps have been regarded as a promising alternative solution for injecting sustainable current into a closed-circuit loop. Extensive research has been conducted worldwide in the past decades covering the theoretical analysis and experimental testing of different types of HTS flux pumps, including the rotary flux pump, linear flux pump, pulsed flux pump and transformer-rectifier flux pump. Details about each category can be found in a recent review paper [12]. Among those HTS flux pump variations, the HTS rotary flux pump, or so-called HTS dynamo, first proposed by Hoffman [13] has been a hot topic owing to its applicability to HTS machine design. The HTS dynamo employs permanent magnets (PMs) to provide travelling magnetic waveform to induce a time-averaged DC voltage without any

* Corresponding authors.
E-mail addresses: Hongye.Zhang@manchester.ac.uk (H. Zhang), Markus.Mueller@ed.ac.uk (M. Mueller).

https://doi.org/10.1016/j.supcon.2022.100017
Received 1 June 2022; Revised 11 July 2022; Accepted 11 July 2022
Available online 3 August 2022

physical contacts, which intuitively suits the standard structure of electric machines, e.g., a synchronous machine.

The open circuit voltage is one of the most important criteria to evaluate the performance of an HTS dynamo. Both qualitative and quantitative analysis aiming to investigate its output characteristics in different aspects have been done in existing literature [14–21], such as the output voltage dependence on airgap distance [22], the operating frequency [23], the width of the HTS tape [24] and the geometry of the magnet [25]. Nevertheless, most previous work focused on a specific parameter and its individual effect, but never achieved a best design for the HTS dynamo as it is challenging to combine all parameters together efficiently. The authors proved the feasibility of applying common machine-learning techniques to capture the output characteristics of an HTS dynamo in a fast and accurate manner in [26], which treats each of the design parameters as independent inputs and V_{oc} as output and link them together by a determined mathematical model. The model proposed in [26] is however limited by only considering the three previously considered design parameters, namely the air gap distance g, HTS tape width W_s and the remanent flux density B_r of the PM. In order to break through those limitations and expand our findings further, we developed an updated numerical model to generate data samples taking into account non-linear high frequency response f, rotor radius R_r and PM width W_m. With a full set of parameter inputs, all key design considerations of an HTS dynamo have been covered. Hence, the new proposed model is capable of efficiently describing the output behavior of HTS dynamo at full scale. Above all, it has also been demonstrated that by inversely applying the proposed deep learning based statistical model, the key parameters required to design an HTS dynamo with a specific output voltage can be conveniently quantified, i.e., a powerful design tool for rotary HTS flux pump has been provided in this paper.

2. Methodology

The work in this paper is divided into two parts, in which the first step is to construct a numerical model that can derive the open circuit voltage via finite element method (FEM) simulations. The derived values are then collected as data samples to feed into the second step, which is to apply an appropriate deep-learning framework to extract the underlying relationships among variables from the data sets. A complete DNN model is obtained, which is capable of predicting V_{oc} of an HTS dynamo by integrating all the six key parameters, without having to resort to complex numerical modelling.

2.1. Multi-layer model

Thanks to the efforts from researchers worldwide, a set of numerical models are available to implement different formulations for HTS modelling. Specifically in the field of HTS dynamo modelling [10,27], have summarized the most up-to-date models, which include coupled ***H-A*** formulation [28], ***H***-formulation with shell current [17–19], segregated ***H***-formulation [29], minimum electromagnetic entropy production (MEMEP) [30,31], coupled ***T-A*** formulation [32,33], integral equation [34], volume integral equation-based equivalent circuit [35]. Each of the listed formulations has been validated to be effective in modelling the electromagnetic characteristics for HTS dynamos.

In this paper, we adapted the benchmarked coupled ***H-A*** formulation as our base numerical model on the consideration of its ease of use, stable performance, and fast computational speed. The coupled ***H-A*** formulation proposed by Brambilla et al. [28] was initially introduced for modelling superconducting rotating machines, which is intuitively suitable for the HTS dynamo modelling in terms of handling the rotating magnets. In an ***H-A*** formulation-based model, the whole system is modelled, but separated into two domains with distinguished formulation: the ***H*** formulation is implemented in the region that contains the superconductors, while the ***A*** formulation is implemented in the region that does not contain superconductor (essentially the rest of the system). In addition, the model can be simplified by constraining the domain, where the magnetic vector potential ***A*** is directly solved, to a small region surrounding the superconductor, allowing the majority of the model to be solved with the magnetic scalar potential V_m, as illustrated in Fig. 1. By replacing the magnetic vector potential with scalar potential, substantial computation cost can be saved. Details about the basic implementation of the ***H-A*** formulation can be found in [27].

The non-linear electrical resistivity of superconductors is reflected by applying the 2D version of E-J power law in Cartesian coordinates (assuming the tape is infinitely long):

$$E_z = \rho J_z = E_c \left(\frac{J_z}{J_c(\boldsymbol{B})}\right)\left[\frac{|J_z|}{J_c(\boldsymbol{B})}\right]^{(n-1)} \tag{1}$$

where E_c and n are chosen as $10^{-4} V/m$ and 20 respectively as convention. $J_c(\boldsymbol{B})$ signifies the field dependent critical current density of the superconductors. In this work, we used the empirical function [36]:

$$J_c(\boldsymbol{B}) = \frac{J_{c0}}{\left(1+\frac{\sqrt{k^2 B_{para}{}^2 + B_{perp}{}^2}}{B_0}\right)^{\alpha}} \tag{2}$$

where B_{para} and B_{perp} represent the parallel and perpendicular components of the magnetic flux density with respect to the wide face of superconducting tape. J_{c0}, B_0, k and α are materials related constant coefficients, whose values are chosen as 23.583 GA/m^2, 169.4 mT, 0.1538 and 1.022 to match the experimental measurement in [17].

The time averaged open circuit voltage V_{oc} is derived by taking the time and surface integration of Eq. (1):

$$V_{oc} = -L\frac{1}{T}\int_T^{2T}\frac{1}{S}\iint E_z(x,y,t)\,ds\,dt \tag{3}$$

Please note that L is the effective depth of the magnet, and the time integration is performed in the second cycle to avoid any transient response that may occur in the first one.

Including the ***H-A*** formulation model, all the numerical models benchmarked in [27] followed a classical approximation that assumes the superconductor tape as a single layer structure. However, in [36] Zhang first put forward a multilayer numerical model for HTS coated conductors considering the impact of both the superconducting and non-superconducting components and modelled the electromagnetic loss in HTS coated conductors over a wide range of frequencies and found that, above a certain frequency, most magnetization losses occur in the copper stabilizers [37–42]. The multilayer model [43] has been adopted to study the frequency dependence of open circuit voltage of

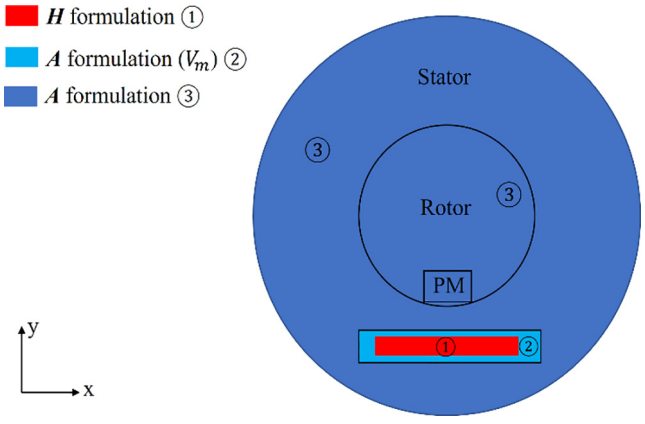

Fig. 1. Illustration of the domain division for the ***H-A*** formulation-based model with different state variables.

an HTS dynamo that has been recently investigated in [16], in which the non-linear frequency response has been attributed to the current interactions between different layers of the HTS tape. As a result, we adopted the multi-layer structure to model the HTS tape as shown in Fig. 2, enabling the model to include the effects of high frequency operation.

2.2. Deep-learning neuron network

Numerical models such as the ***H-A*** formulation described above are widely used to solve superconductivity problems, which can solve Maxwell equations combined with required physical laws and conditions through finite element method (FEM) simulations. It is well-recognized that FEM simulations can simulate the complete system operation while capturing details from the physical process involved, offering an efficient and low-cost alternative to experimental work. However, the utilization of FEM simulations is subjected to professional knowledge about the target problem, which requires abundant user interactions, and the efficiency deserves further improvement. Such an approach does not lend itself to the industrial design and optimization, where the focus is put on certain specifications that can reflect the performance of the device rather than the underlying physics.

Deep-learning is a neuron network based artificial intelligence method, which can represent unknown relations among a set of data samples in the form of a statistical model [44–46]. The multi-layer perceptron (MLP) is the most widely used neuron network, consisting of the input layer, output layer and hidden layer(s). Fig. 3 shows a full-connected neuron network, which has n hidden layers with m_n neurons in each of them. Neurons in each layer are connected through certain activation functions, so the given input and output can be linked together. For the problem concerned here, the number of neurons in the input and output layer are fixed to be six and one, respectively, in accordance with the six design parameters and one voltage output for an HTS dynamo. Given the relationships between the output and each of the inputs can be highly nonlinear, the rectified linear unit (ReLU) function, which possesses the superior ability to characterize the complex nonlinearity of the model and smoothen gradient propagation during the training, is employed as the activation function [47]:

$$ReLU(x) = \max(0, \boldsymbol{w}^T\boldsymbol{x} + b) \tag{4}$$

where $\boldsymbol{w}^T$ denotes the weight coefficients matrix and b signifies the bias coefficients.

The main purpose of the model is to predict the voltage, which can be considered as a regression task to predict a continuous variable, and hence the squared error (L2 norm) is adopted as the loss criterion to quantify the discrimination between every observation and prediction:

$$l(y, \hat{y}) = \{l_1, \cdots, l_N\}, l_N = (y_n - \widehat{y_n})^2 \tag{5}$$

Based on this loss function, the adaptive moment estimation (Adam) is applied as the optimisation solver to update weight coefficients matrix in Eq. (11) as follows [48]:

$$g_t = \nabla_\theta f_t(\theta_{t-1}) \tag{6}$$

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \tag{7}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \tag{8}$$

$$\widehat{m_t} = m_t / (1-\beta_1^t) \tag{9}$$

$$\widehat{v_t} = v_t / (1-\beta_2^t) \tag{10}$$

$$\theta_t = \theta_{t-1} - \alpha \widehat{m_t} / \left(\sqrt{\widehat{v_t}} + \varepsilon\right) \tag{11}$$

where f_t is the objective function and θ_t represents the parameters to be updated, while α, β and ε are set to be constant values as 0.001, 0.9 and 10^{-8}, respectively. The k-fold cross validation scheme is taken, which means the data set is divided into k mini-sets and the training process repeats k times by taking each of those mini-sets as the validation set. Thus, the adapted Nash-Sutcliffe model efficiency coefficient (NSE) is calculated to evaluate the performance of the trained model:

$$NSE = \frac{1}{k}\sum_{t=1}^{k}\left[1 - \frac{\sum_{i=1}^{p}\left(y_i - \widehat{y_i}\right)^2}{\sum_{i=1}^{p}\left(y_i - \bar{y}\right)^2}\right] \tag{12}$$

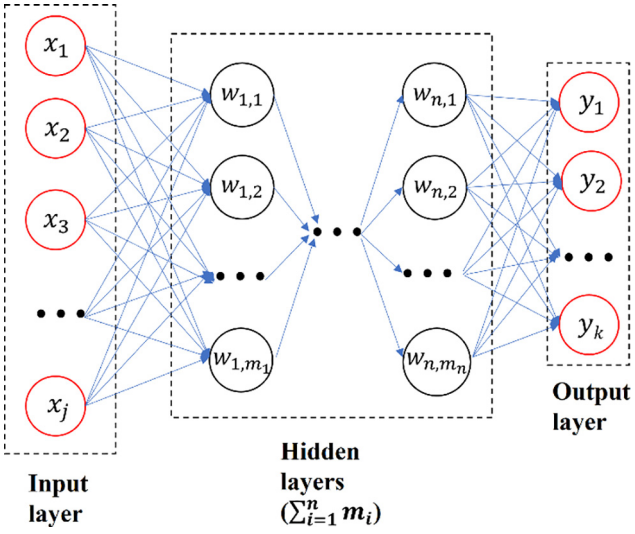

Fig. 3. Structure of a typical full-connected neuron network.

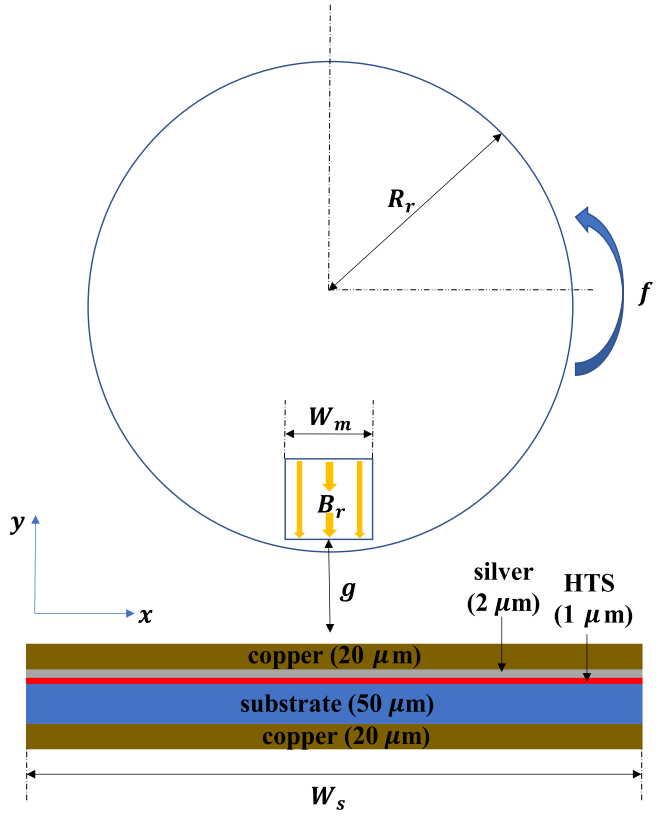

Fig. 2. Cross section schematic for the multilayer structure adopted in the ***H-A*** numerical model.

where k is the number of mini-sets (set to be 10), p is the number of data samples, $\bar{y}$ represents the average output, y_i and $\hat{y}_i$ indicate the observed and predicted output, respectively. The best NSE score that can be achieved is 1, which means that the model can predict all data samples perfectly with no errors. Otherwise, the closer the NSE score is to 1, the better accuracy the model will have. We have previously implemented the MLP with one hidden layer to predict the output characteristics for an HTS dynamo based on three input design parameters [26]. In this paper, we take advantage of the deep-learning approach enhanced by GPU acceleration to investigate more complicated neuron networks, then put forward an improved model, which can rapidly and accurately capture the output characteristics for HTS dynamo with all design parameters now included.

2.3. Data pre-processing

Each set of data required by the DNN training requires seven values, with six values for each of the input design parameters and one for the output open circuit voltage. In order to make the input data more effective and representative, two data schemes were chosen. The first one divides the value range for each parameter into 1000 small intervals, which essentially forms a dense data map that improves the resolution of the prediction model. Randomly combining the median of each interval k times, $1000 \times k$ sets of data samples can be obtained. The second one divides the value range for each parameter into 50 relatively large intervals, which allows more room for different parameter combinations under limited data samples, so that it is easier for the model to learn the general behavior from the data samples. By randomly combining the median of each interval k times, $50 \times k$ sets of data samples can be obtained. In this work, 4000 data samples have been generated following the first and second data chosen scheme, respectively, to form the data set required by the model training.

The input values propagate through the hidden layers simultaneously until they reach the output layer, so it is important that the values for each input neuron are closely comparable, otherwise the input neuron with a significantly higher value will dominate the training process and lead to poor performance for the model to predict the output when taking other input neurons into account. Therefore, the values for each input neuron were normalized into unit scale by the following transformation:

$$x_t = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{13}$$

Though it is commonly believed that the output values will not have impacts on the quality of the trained model, it should be noted that the output values in this study represent the open circuit voltage from an HTS dynamo, which can vary between zero to thousands of microvolts. In order to avoid any convergence difficulties due to the wide fluctuation of values, the logarithm for the output values is taken to constrain them in a narrow domain as follows:

$$y_t = \log_{10}(y + 1) \tag{14}$$

3. Results

3.1. Numerical model validation

In this work, an ***H-A*** formulation based numerical model is built to generate the data samples required for DNN model training. To verify this model, it was firstly utilized to calculate the time dependent equivalent voltage (surface integration of Eq. (1)) in comparison to the experimental measurement in [17], as plotted in Fig. 4. The waveforms show excellent agreement with each other, and their calculation results for Eq. (3) are 27.51 μV and 27.57 μV, confirming the effectiveness of this model. Note the results are obtained from the parameter setting, where g is 3.7 mm, W_s is 12 mm, f is 4.25 Hz, B_r is 1.25 T, R_r is 35 mm and W_m is 6 mm. Unless state explicitly, each parameter follows the same setting in results presented below.

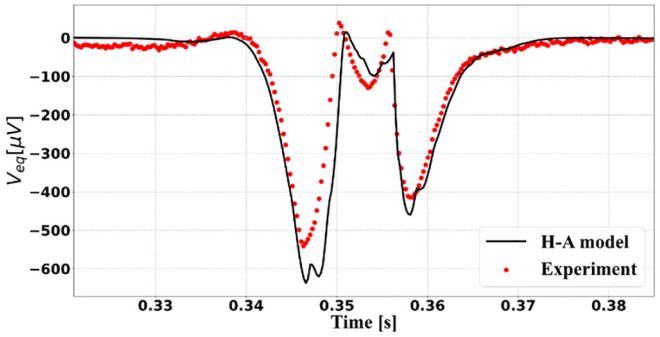

Fig. 4. Waveform profiles of the instantaneous voltage across the HTS tape for the permanent magnet past the HTS tape in the 2nd cycle, ignoring any initial transient effects.

3.2. DNN topology determination

In terms of constructing a MLP type neuron network, there are several influential factors, or so-called hyperparameters that have significant impact on the performance of the derived model. Some of these hyperparameters can be intuitively set according to the nature of the problem, such as the number of neurons in the input and output layer, and some of the others can follow the general rules of practice published in existing literature, such as the choice of proper activation function and loss function. However, exact guidelines about how to determine the structure of hidden layers are still lacking [49]. Researchers have put tremendous efforts in trying to provide a solution for this issue, including the Akaike's Information Criterion [50], Inverse test method [51] and some customized methods to find the optimal architecture for neuron network. Yet those methods are developed for specific problems, and it is hard to simply apply them to a new situation. In this paper, we determined the neuron network topology based on previous work and our understanding of this specific problem.

Theoretically, hidden layers greater than one can describe any arbitrary functions with arbitrary boundaries. Nevertheless, more hidden layers do not necessarily guarantee a better performance of the model. Besides, the increase of hidden layers will expand the model complexity and result in lower efficiency since more time is required to solve the model. Choldum et al., have stated in [52] that the number of hidden layers in the neuron network can be determined according to the main components suggested by the principal components analysis. In our case, the six input variables are pre-selected design parameters, which are independent to each other and no correlations exist among them. Hence, it is reasonable to limit the hidden layers for the neuron network below six. As for the number of neurons in the hidden layer, it also needs careful attention. On the one hand, insufficient neurons will make the model incapable of learning the underlying relations from the data sample, which is often referred to as 'under fitting'. On the other hand, excessive neurons will cause 'over fitting', where the model gets stuck in local optima and fails to learn the general behavior from the data samples. Considering the suggestions in [53–55] and the computing resources available, a randomized grid search was performed to test the number of neurons in each hidden layer ranges from 4 to 1024. A series of experiments with different DNN topology were then conducted, and the quality of those models were evaluated using their NSE scores. The best topology that achieved the highest NSE score for each number of hidden layers has been selected to be presented in Table 1.

Table 1
NSE scores for different DNN model topologies.

Number of neurons in each layer						NSE
1	2	3	4	5	6	
1020	NA	NA	NA	NA	NA	0.99601
808	292	NA	NA	NA	NA	0.99932
636	376	444	NA	NA	NA	0.99927
432	820	328	836	NA	NA	0.99908
400	816	856	348	400	NA	0.99910
968	328	160	696	128	92	0.99907

Yet a high NSE score does not necessarily mean that the model has good quality, because the data set prepared to train the model is limited. In order to ensure that the model has generalized prediction capability, the best trained DNN model was selected to predict a new set of data samples that are not covered in the training set. As shown in Fig. 5, all the predictions denoted by the red points locate in close proximity of the perfect match line, indicating very good performance of this model in the general case. The relative error in percentage for each prediction is plotted in Fig. 6, from where it can be seen that the maximum error for two predictions is about 12% out of all tested samples (total number of 183), while most of the predictions have an error rate less than 2% and the average error rate is maintained at 2.06%. Fig. 7 shows these prediction errors in more detail by dividing all the test samples into different intervals in terms of their proportions to the maximum value. It can be observed that high error rates are easier to occur for small predicted values, e.g. the 0 ~ 10% interval, which can be readily understood because small values are more vulnerable to the error measurement under relative criteria. Even including those outliers (identified by black cycles), the average error rate (marked by green triangles) for predictions in every interval are well below 3%. Besides, the demand for designing and optimizing HTS dynamo type of devices, in most cases, is to maximize the output voltage, for which we believe the applicability and effectiveness of this model will not be undermined.

3.3. Parameter grouping

As a demonstration of the proposed DNN model, we utilized the model to illustrate the output characteristics for each of the six parameters. Since it is not practical to visualize all the parameters in one single plot, the six parameters were manually classified into three groups. Firstly, both the frequency and the rotor radius determine how long

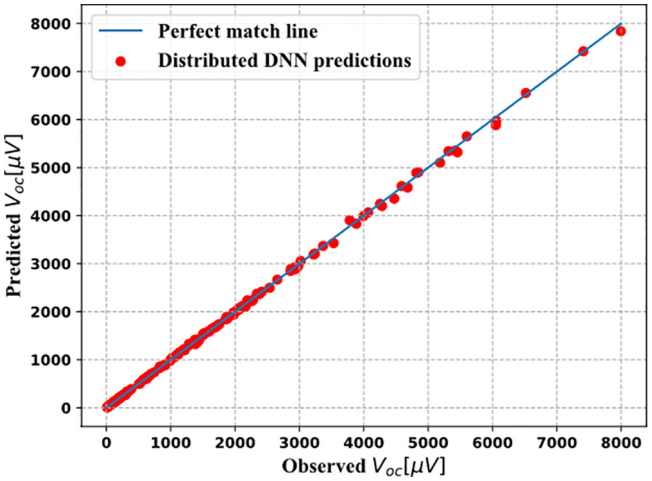

Fig. 5. Prediction results by the best trained DNN model.

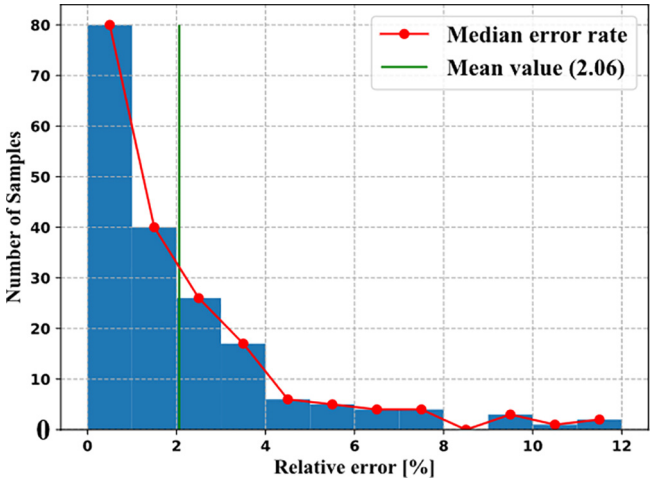

Fig. 6. Histogram plot of the error rates for predicted voltage.

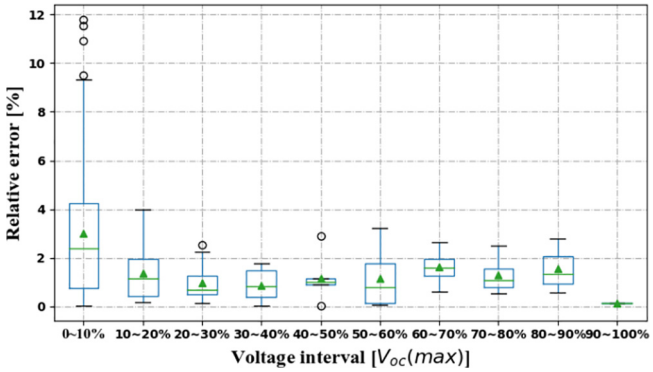

Fig. 7. Box Plot of the error rates for predicted voltage grouped by different intervals (proportional to the maximum prediction).

the HTS tape will experience the effective field provided by the rotating magnet in one complete cycle, so they are grouped together as the "duration group". Secondly, the magnetic field experienced by the HTS tape is directly controlled by the airgap distance and the PM remanent flux density, and so form the "field group". Lastly, the authors of [24] have pointed out that the HTS tape width relative to the PM width can make a difference on the output voltage by affecting some critical features, e.g. whether the applied field can be considered homogeneous: therefore we have grouped the tape and PM width together as the "width group". As it can be seen from Fig. 8(a), the smooth curved surface implies that the frequency response of the HTS dynamo tends to become non-linear when the frequency increases gradually (the turning point occurs at approximately 100 Hz), which is in accordance with the observation in [16]. In addition, the rotor radius response

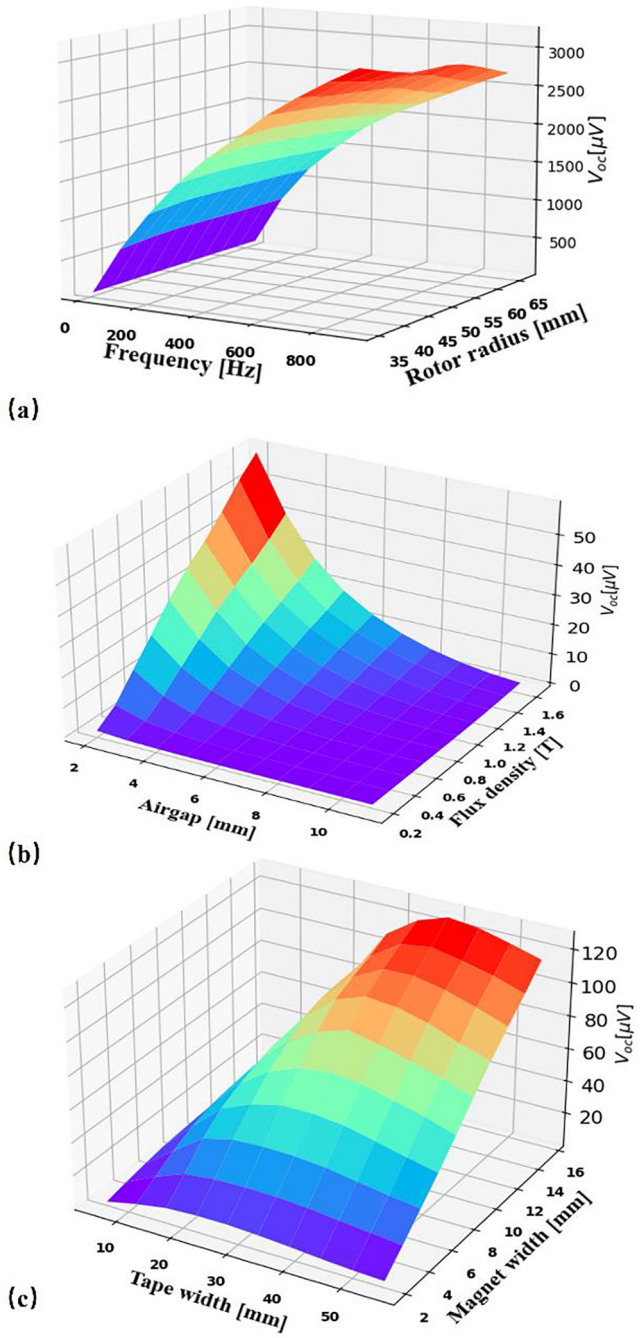

Fig. 8. Surface plots for different parameter groups. (a) "Duration group", (b) "field group", (c) "width group".

also shows what is expected: the output voltage decreases when the rotor gets larger. This is because a larger rotor will reduce the time interval during which the HTS tape is exposed to the effective PM magnetic field, which is essentially equivalent to a decrease in frequency. From Fig. 8(b) it can be observed that the peak output voltage can be achieved by having the smallest airgap and the largest remanent flux density at the same time. The saddle shape in Fig. 8(c) reflects the bilateral effects of the HTS tape width on the output voltage, namely the fact that the voltage increases with HTS tape width up to a certain point, after which it starts to decrease. Meanwhile, the PM width also plays a critical role on the output voltage, which increases monotonically with the PM width within the investigated range.

3.4. Optimal tape width

Combining our previous study in [26] and the work demonstrated in section 3.3, it is concluded that the individual impact of all design parameters on the output voltage can be described by a monotonic function, except for the HTS tape width. Since it has now been fully proved that there exists an optimal HTS tape width for the HTS dynamo to obtain maximum output voltage, it is worthwhile knowing whether the optimal width is influenced by other parameters. A series of parameter sweeps were conducted for each of the other five parameters by varying the HTS tape width, to identify the optimal HTS tape width for different configurations, with results shown in Fig. 9. According to the sweeping results, altering any of the parameters will change the optimal width correspondingly. An increase in PM remanent flux density, Fig. 9(b), and PM width, Fig. 9(e) both result in a considerably wider HTS tape required to obtain higher output voltage. The rotor radius has a relatively less influence on the optimal HTS tape, as shown in Fig. 9(d), which results in a difference less than 2 mm between the maximum optimal width of 32 mm and minimum optimal width of 30.5 mm. While its impact on the output voltage is also not significant as can be seen from Fig. 8(a), so that it is reasonable to ignore the rotor radius. The optimal HTS tape width tends to increase with air gap, as shown in Fig. 8(a). In particular, Fig. 9(c) indicates that the optimal tape width is inversely proportional to the frequency under low frequencies (roughly under 100 Hz), and then becomes positively proportional to the frequency at higher frequencies. Meanwhile, it has been revealed that either a decrease in the airgap or an increase in the operating frequency can help level up the output voltage. This observation can lead to a useful HTS dynamo design rule, namely that the device should be operated under small airgap distance and high frequency, because less HTS tape is required while obtaining higher output voltage.

As aforementioned, the width ratio of the PM and the HTS tape is a key factor in an HTS dynamo. In order to explore this factor in more detail, we defined the "width ratio" W_{rt} as:

$$W_{rt} = \frac{W_s}{W_m} \tag{15}$$

The output characteristics predicted by our DNN model for different W_{rt} comparison is presented in Fig. 10. Each of the solid lines with a distinct color represents the PM width response under a specific width ratio, and it is clear that increasing the PM width increases the output voltage if the width ratio is kept constant. However, for W_{rt} greater 6, the intersections imply that increasing the width ratio cannot guarantee that the voltage will also increase. This result proves the existence of the bilateral effect of the HTS tape width again and implies that the HTS tape width and the PM width should be considered separately while employing the width ratio between them is not sufficient to summarize their individual impacts in one variable. Moreover, Fig. 10 also provides a possible solution to avoid the limitation of the bilateral effect of the HTS tape width in HTS dynamo design, which is to increase the PM width further, so that the optimal HTS tape width can be effectively extended.

3.5. Case study

The DNN model proposed in this paper can bring direct benefits to those who want to obtain the output voltage for arbitrary configurations. When it comes to another typical design scenario, where the desired output voltage is specified but the design parameters are to be determined, this model also can be utilized as a powerful design tool to provide recommended parameters, according to the given specifications. As a demonstration of this useful function, several design conditions are assumed to perform a case study. In Case 1, there are no manual constraints on any of the design parameters, which means that the model is allowed to provide possible recommendations in the

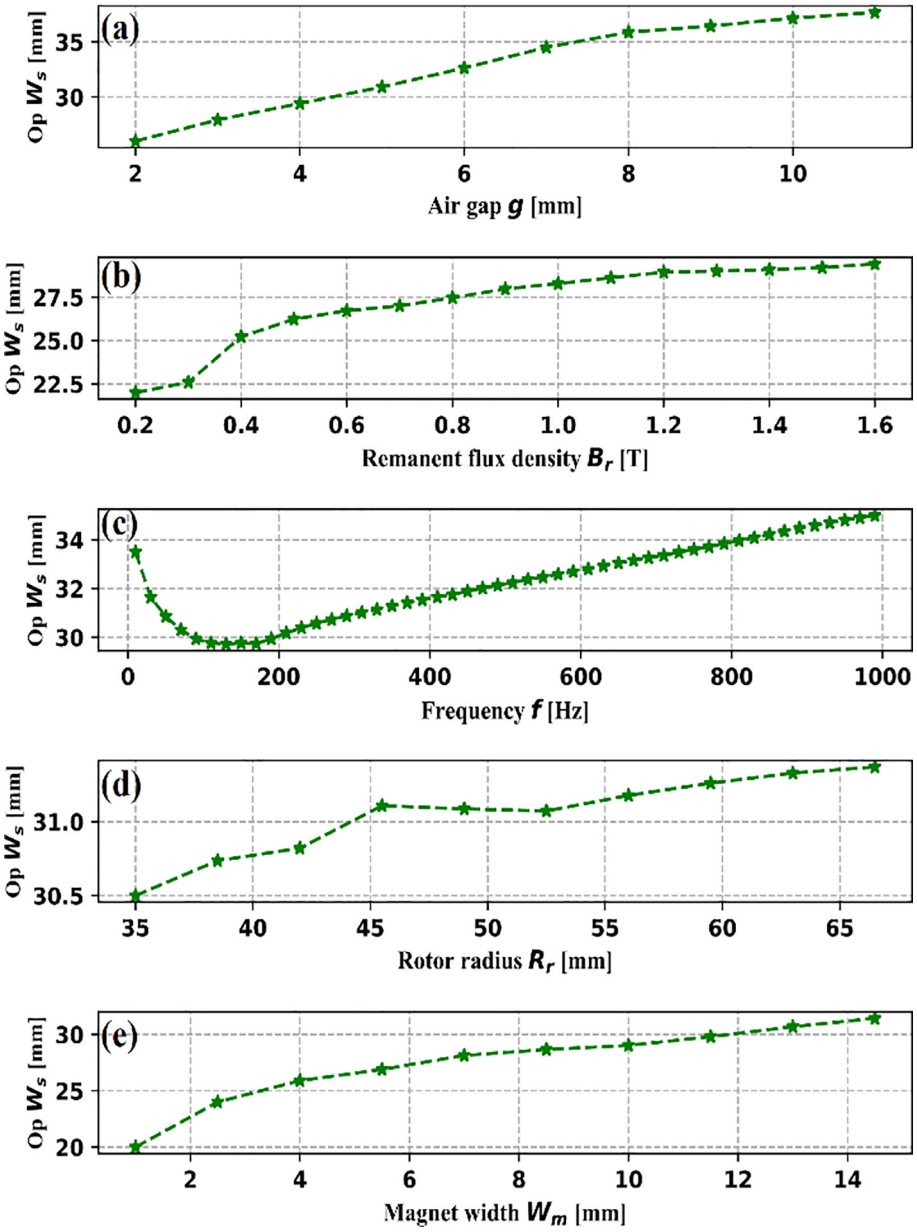

Fig. 9. Optimal HTS tape width for different (a) air gap, (b) remanent flux density, (c) frequency, (d) rotor radius, (e) PM width.

full default range for each of the parameters. In Case 2, the airgap is restricted between 3 and 5 mm. In Case 3, the frequency is restricted under 200 Hz while the permanent magnet width is restricted between 2 and 12 mm. For all the three cases described above, the desired output voltage is set to be 260 μV. It should be noted that there may exist various parameter combinations, which can result in the same output voltage. Hence, for each of the assumed design conditions, one set of the recommended design parameters is selected to run through the base numerical model to validate its correctness. The results are presented in Table 2, where the correctness is defined as:

$$correctness = 1 - \frac{|V_{simu} - V_{exp}|}{V_{exp}} \tag{16}$$

where V_{exp} identifies the desired voltage, V_{simu} identifies the output voltage calculated from the numerical simulations by setting the recommended parameters. It can be seen that an HTS dynamo with the recommended parameters for each case is capable of generating a output voltage that is close to the expected value, which demonstrates that the proposed model can provide fast guidance under specific design considerations by listing reliable recommended parameters.

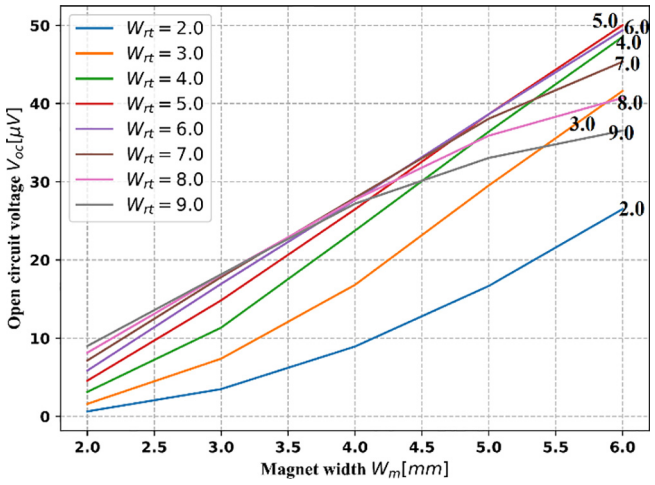

Fig. 10. Open circuit voltage characteristics against magnet width under various width ratio.

Table 2
Recommended parameters for each presumed condition.

	g/mm	W_s/mm	B_r/T	f/Hz	R_r/mm	W_m/mm	Correctness
Case 1	9.5	46	1.23	750.25	35	4.75	96.77%
Case 2	5	60	0.85	250.75	52.5	4.75	98.91%
Case 3	7	32	0.85	50.75	35	12	97.54%

4. Conclusions

In this paper, a novel statistical model has been proposed to efficiently predict the output voltage for a rotary HTS flux pump (HTS dynamo) with an overall accuracy of around 98% with respect to the FEM numerical models, by means of the deep learning neuron network. This model takes into account six design parameters, namely the airgap distance, HTS tape width, operating frequency, rotor radius, as well as remanent flux density and width of the PM, which cover all the typical design dimensions for HTS dynamos. The data set prepared to train this model was generated by a benchmarked ***H-A*** formulation based numerical model, by adding a multilayer structure and magnetic field sensitivity for the modelling of the HTS tape, to reflect the high frequency response and $J_c(\boldsymbol{B})$ dependence. Therefore, this model possesses a generalized ability to quantify the output voltage for an HTS dynamo according to its design parameters. It should be pointed out that despite the monotonic dependence of the output voltage on some parameters, e.g., the air gap and field density, the correlations between the output voltage and all these influential factors are predominantly non-linear, which is challenging to be quantitatively described by analytical equations. The proposed statistical model is capable of integrating all the studied parameters together to efficiently quantify the output voltage characteristics of a rotary HTS flux pump.

As a demonstration, the proposed model is utilized to investigate the output properties for each of the six parameters. It is observed that an increase in permanent magnet width can help to increase the effective width of the HTS tape, which can serve as a solution to avoid the bilateral effect observed in [26]. The results imply that properly increasing the operation frequency (without causing extra losses in the HTS and non-superconducting layers) can not only result in a higher output voltage, but also help decrease the width of HTS tape to achieve the maximum voltage. The case study shows that the model is also capable of generating the optimal design parameters under any specific design condition. In addition, though the model is developed specifically for the rotary HTS flux pumps described above, the methods and techniques demonstrated here can be flexibly extended to other more complicated scenarios. In a summary, a robust and reliable statistical model has been proposed in this paper, which opens the way to rapid and accurate design and optimization of HTS flux pumps.

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.

Acknowledgements

The authors would like to thank Dr. Chris Bumby and Dr. Ratu Mataira at the Victoria University of Wellington for providing experimental results to validate our numerical model. Hongye Zhang would like to give thanks to the support of 2021 IEEE Council on Superconductivity Graduate Study Fellowship in Applied Superconductivity. Mengyuan Tian would like to thank for the George and Lilian Schiff Foundation PhD scholarship provided by University of Cambridge to support her research.

Supplementary Materials

The DNN model proposed in this paper has been produced as an executable application to be freely downloaded here, which contains a detailed user guide and background information for the software development.

References

[1] Sung H-J, Kim G-H, Kim K, et al. Practical design of a 10 MW superconducting wind power generator considering weight issue. IEEE Trans Appl Supercond 2013;23(3):5201805.
[2] Sung H-J, Kim G-H, Kim K, et al. Design and comparative analysis of 10 MW class superconducting wind power generators according to different types of superconducting wires. Physica C 2013;494:255–61.
[3] Karmaker H, Ho M, Chen E, et al., Direct drive HTS wind generator design for commercial applications. pp. 491-495.
[4] Grilli F, Benkel T, Hänisch J, et al. Superconducting motors for aircraft propulsion: the Advanced Superconducting Motor Experimental Demonstrator project. J Phys: Conf Ser 2020;1590(1):012051.
[5] Weijers HW, Trociewitz UP, Markiewicz WD, et al. High field magnets with HTS conductors. IEEE Trans Appl Supercond 2010;20(3):576–82.
[6] Bumby CW, Badcock RA, Sung H-J, et al. Development of a brushless HTS exciter for a 10 kW HTS synchronous generator. Supercond Sci Technol 2016;29(2):024008.
[7] Feigel'man M, Geshkenbein V, Vinokur V. Flux creep and current relaxation in high-T c superconductors. Phys Rev B 1991;43(7):6263.
[8] Geng J, Zhang H, Li C, et al. Angular dependence of direct current decay in a closed YBCO double-pancake coil under external AC magnetic field and reduction by magnetic shielding. Supercond Sci Technol 2017;30(3):035022.
[9] Grilli F, Pardo E, Stenvall A, et al. Computation of losses in HTS under the action of varying magnetic fields and currents. IEEE Trans Appl Supercond 2014;24(1):78–110.
[10] Zhang H, Wen Z, Grilli F, et al. Alternating current loss of superconductors applied to superconducting electrical machines. Energies 2021;14(8):2234.
[11] Coombs T. Superconducting flux pumps. J Appl Phys 2019;125(23):230902.
[12] Wen Z, Zhang H, Mueller M. High temperature superconducting flux pumps for contactless energization. Crystals 2022;12(6):766.
[13] Hoffmann C, Pooke D, Caplin AD. Flux pump for HTS magnets. IEEE Trans Appl Supercond 2010;21(3):1628–31.
[14] Ghabeli A, Pardo E. Modeling of airgap influence on DC voltage generation in a dynamo-type flux pump. Supercond Sci Technol 2020;33(3):035008.
[15] Ghabeli A, Pardo E, Kapolka M. 3D modeling of a superconducting dynamo-type flux pump. Sci Rep 2021;11(1):1–12.
[16] Ainslie MD, Queval L, Mataira RC, et al. Modelling the frequency dependence of the open-circuit voltage of a high-T c superconducting dynamo. IEEE Trans Appl Supercond 2021;31(5):1–7.
[17] Mataira R, Ainslie M, Badcock R, et al. Origin of the DC output voltage from a high-Tc superconducting dynamo. Appl Phys Lett 2019;114(16):162601.
[18] Mataira R, Ainslie MD, Badcock R, et al. Modeling of stator versus magnet width effects in high-$ T_c $ superconducting dynamos. IEEE Trans Appl Supercond 2020;30(4):1–6.
[19] Mataira R, Ainslie M, Pantoja A, et al. Mechanism of the high-T c superconducting dynamo: models and experiment. Phys Rev Appl 2020;14(2):024012.
[20] Bumby C, Jiang Z, Storey J, et al. Anomalous open-circuit voltage from a high-T c superconducting dynamo. Appl Phys Lett 2016;108(12):122601.
[21] Ma J, Geng J, Gawith J, et al. Rotating permanent magnets based flux pump for HTS no-insulation coil. IEEE Trans Appl Supercond 2019;29(5):1–6.
[22] Jiang Z, Bumby CW, Badcock RA, et al. Impact of flux gap upon dynamic resistance of a rotating HTS flux pump. Supercond Sci Technol 2015;28(11):115008.
[23] Bumby CW, Phang S, Pantoja AE, et al. Frequency dependent behavior of a dynamo-type HTS flux pump. IEEE Trans Appl Supercond 2017;27(4):1–5.
[24] Pantoja AE, Jiang Z, Badcock RA, et al. Impact of stator wire width on output of a dynamo-type HTS flux pump. IEEE Trans Appl Supercond 2016;26(8):1–8.
[25] Badcock RA, Phang S, Pantoja AE, et al. Impact of magnet geometry on output of a dynamo-type HTS flux pump. IEEE Trans Appl Supercond 2017;27(4):1–5.
[26] Wen Z, Zhang H, Mueller M. Sensitivity analysis and machine learning modelling for the output characteristics of rotary HTS flux pumps. Supercond Sci Technol 2021;34(12):125019.
[27] Ainslie M, Grilli F, Quéval L, et al. A new benchmark problem for electromagnetic modelling of superconductors: the high-T c superconducting dynamo. Supercond Sci Technol 2020;33(10):105009.

[28] Brambilla R, Grilli F, Martini L, et al. A finite-element method framework for modeling rotating machines with superconducting windings. IEEE Trans Appl Supercond 2018;28(5):1–11.
[29] Quéval L, Liu K, Yang W, et al. Superconducting magnetic bearings simulation using an H-formulation finite element model. Supercond Sci Technol 2018;31(8):084001.
[30] Pardo E, Šouc J, Frolek L. Electromagnetic modelling of superconductors with a smooth current–voltage relation: variational principle and coils from a few turns to large magnets. Supercond Sci Technol 2015;28(4):044003.
[31] Pardo E, Kapolka M. 3D computation of non-linear eddy currents: Variational method and superconducting cubic bulk. J Comput Phys 2017;344:339–63.
[32] Zhang H, Zhang M, Yuan W. An efficient 3D finite element method model based on the T-A formulation for superconducting coated conductors. Supercond Sci Technol 2016;30(2):024005.
[33] Benkel T, Lao M, Liu Y, et al. T–A-formulation to model electrical machines with HTS coated conductor coils. IEEE Trans Appl Supercond 2020;30(6):1–7.
[34] Brambilla R, Grilli F, Nguyen DN, et al. AC losses in thin superconductors: the integral equation method applied to stacks and windings. Supercond Sci Technol 2009;22(7):075018.
[35] Morandi A, Fabbri M. A unified approach to the power law and the critical state modeling of superconductors in 2D. Supercond Sci Technol 2014;28(2):024004.
[36] Zhang H, Yao M, Kails K, et al. Modelling of electromagnetic loss in HTS coated conductors over a wide frequency band. Supercond Sci Technol 2020;33(2):025004.
[37] Zhang H, Mueller M. Electromagnetic properties of curved HTS trapped field stacks under high-frequency cross fields for high-speed rotating machines. Supercond Sci Technol 2021;34(4):045018.
[38] Zhang H, Machura P, Kails K, et al. Dynamic loss and magnetization loss of HTS coated conductors, stacks, and coils for high-speed synchronous machines. Supercond Sci Technol 2020;33(8):084008.
[39] Musso A, Breschi M, Ribani PL, et al. Analysis of AC loss contributions from different layers of HTS tapes using the A − V formulation model. IEEE Trans Appl Supercond 2021;31(2):1–11.
[40] Zhang H, Chen H, Jiang Z, et al. A full-range formulation for dynamic loss of high-temperature superconductor coated conductors. Supercond Sci Technol 2020;33(5):05LT01.
[41] Zhang H, Yao M, Jiang Z, et al. Dependence of dynamic loss on critical current and n-value of HTS coated conductors. IEEE Trans Appl Supercond 2019;29(8):1–7.
[42] Zhang H, Hao C, Xin Y, et al. Demarcation currents and corner field for dynamic resistance of HTS-coated conductors. IEEE Trans Appl Supercond 2020;30(8):1–5.
[43] Ma J, Geng J, Chan WK, et al. A temperature-dependent multilayer model for direct current carrying HTS coated-conductors under perpendicular AC magnetic fields. Supercond Sci Technol 2020;33(4):045007.
[44] Karlik B, Olgac AV. Performance analysis of various activation functions in generalized MLP architectures of neural networks. Int J Artif Intell Expert Syst 2011;1(4):111–22.
[45] Taud H, Mas J. Multilayer perceptron (MLP). In: Geomatic approaches for modeling land change scenarios. Springer; 2018. p. 451–5.
[46] Bisong E. The multilayer perceptron (MLP). In: Building machine learning and deep learning models on Google cloud platform. Springer; 2019. p. 401–5.
[47] Schmidt-Hieber J. Nonparametric regression using deep neural networks with ReLU activation function. Ann Stat 2020;48(4):1875–97.
[48] Kingma DP, Ba J, Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*, 2014.
[49] Cruz IA, Chuenchart W, Long F, et al. Application of machine learning in anaerobic digestion: Perspectives and challenges. Bioresour Technol 2022;345:126433.
[50] Akaike H. A new look at the statistical model identification. IEEE Trans Autom Control 1974;19(6):716–23.
[51] Panchal G, Ganatra A, Kosta Y, et al. Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. Int J Comput Theory Eng 2011;3(2):332–7.
[52] Choldun I, Santoso J, Surendro K, Determining the number of hidden layers in neural network by using principal component analysis. pp. 490-500.
[53] Karsoliya S. Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. Int J Eng Trends Technol 2012;3(6):714–7.
[54] Stathakis D. How many hidden layers and nodes? Int J Remote Sens 2009;30(8):2133–47.
[55] Vujicic T, Matijevic T, Ljucovic J, et al., Comparative analysis of methods for determining number of hidden neurons in artificial neural network. p. 219.