Modeling of quasistatic magnetic hysteresis with feed-forward neural networks

Dimitre Makaveev, a) Luc Dupré, Marc De Wulf, and Jan Melkebeek
Department of Electrical Power Engineering, Ghent University, Ghent B-9000, Belgium

A modeling technique for rate-independent (quasistatic) scalar magnetic hysteresis is presented, using neural networks. Based on the theory of dynamic systems and the wiping-out and congruency properties of the classical scalar Preisach hysteresis model, the choice of a feed-forward neural network model is motivated. The neural network input parameters at each time step are the corresponding magnetic field strength and memory state, thereby assuring accurate prediction of the change of magnetic induction. For rate-independent hysteresis, the current memory state can be determined by the last extreme magnetic field strength and induction values, kept in memory. The choice of a network training set is motivated and the performance of the network is illustrated for a test set not used during training. Very accurate prediction of both major and minor hysteresis loops is observed, proving that the neural network technique is suitable for hysteresis modeling. © 2001 American Institute of Physics. [DOI: 10.1063/1.1361268]

INTRODUCTION

A material hysteresis model that provides a description of the time-dependent relation between the magnetic field strength vector \( \mathbf{H}(t) \) and the magnetic induction vector \( \mathbf{B}(t) \) in ferromagnetic materials (e.g., SiFe) is required for an accurate finite element analysis of electromagnetic devices. In case of unidirectional magnetization [\( \mathbf{H}(t) \) and \( \mathbf{B}(t) \) are scalars, denoted \( H(t) \) and \( B(t) \)], one of the most widely used models for magnetic hysteresis is the Preisach model. 1 The classical Preisach model is very accurate in describing the quasistatic magnetic properties of electrical steel sheets under arbitrary unidirectional excitations. However, the extension of the model to two-dimensional (e.g., circular) magnetization, like that which occurs in lamination of rotating electrical machines, but also to a lesser extent in transformers, is not straightforward.1 “Black box” input–output mathematical models, based on artificial neural networks (ANNs), are emerging as a powerful modeling tool2 and could form an alternative to hysteresis modeling techniques like the Preisach model. In this article we present a neural network hysteresis model that offers the same accuracy as the classical scalar Preisach model. The neural network topology and input parameters are chosen based on the properties of the Preisach model and the theory of dynamic systems.

PROPERTIES OF THE CLASSICAL SCALAR PREISACH HYSTERESIS MODEL

Two important properties of the classical scalar Preisach model 1 are taken into account for the development of the neural network model. First, the Preisach model describes rate-independent (quasistatic) hysteresis. The hysteresis memory state at each time step \( k \) is determined only by the extreme values (maxima and minima) of the field strength history \( H_0, H_1, \ldots, H_{k-1}, H_k \), the input of the model (the index 0 denotes the initial demagnetized material state, with \( H_0 = 0 \), in discrete time notation). The change of the magnetic induction \( B_k \) (the model output) is determined from the current field value \( H_k \) and exactly one stored extreme field value \( H_{k}^{\text{extr}} \), at each time step \( k \) (Fig. 1). The relevant extreme value \( H_{k}^{\text{extr}} \) is the last one stored in memory. The closing of a minor loop deletes from memory the maximum and minimum \( H \) values associated with this minor loop (the wiping-out property 1). When no extreme values are present in memory, the virgin magnetization curve is followed. The relevant extreme value \( H_{k}^{\text{extr}} \) can be determined at each time step \( k \) by an algorithm that stores the encountered extreme values and implements the wiping-out property.

FIG. 1. Example of rate-independent hysteresis: (a) field strength history and corresponding extreme value; (b) \( BH \) loop.
The second important property of the classical Preisach model is the congruency property. It states that BH loops between two fixed extreme field values are independent of the induction level B. Combining the two properties, the Preisach model can be expressed mathematically as

\[ B_k = B_k^{\text{extr}} + f(H_k^{\text{extr}}, H_k, \text{FLAG}_k), \]  

(1)

where \( B_k^{\text{extr}} \) is the corresponding induction level at the relevant extreme field value \( H_k^{\text{extr}} \) and \( f \) is a nonlinear function of three variables. The variable \( \text{FLAG}_k \) can take only two different values, e.g., \(-1\) and \(1\), to distinguish between the virgin curve (no extreme values present in memory) and a hysteresis branch, respectively. Indeed, for the special case of \( H_k^{\text{extr}} = 0 \), there exist two distinct curves: \( B_k = B_k^{\text{extr}} + f_1(0, H_k) \) when following the virgin curve \( (B_k^{\text{extr}} = 0) \), and \( B_k = B_k^{\text{extr}} + f_2(0, H_k) \) when following a hysteresis branch (closing of a minor loop, \( B_k^{\text{extr}} \neq 0 \)). Note that Eq. (1) is similar to the description of the Preisach model with the Everett function \( Ev(H_1, H_2) \): \( 3 \)

\[ |\Delta B_k| = |B_k - B_k^{\text{extr}}| = Ev(H_k^{\text{extr}}, H_k) \]  

(2a)

or

\[ |B_k| = \frac{1}{2} Ev(-H_k, H_k), \]  

(2b)

depending whether a hysteresis branch or the virgin curve, respectively, is followed.

HYSTERESIS MODELING WITH FEED-FORWARD NEURAL NETWORKS

Artificial neural networks are black box mathematical tools that can be used for modeling nonlinear dynamic input–output relations. \( 2 \) In general, a magnetic hysteresis system can be represented as a nonlinear dynamic input-state-output relation (with memory), as follows:

\[ X_k = \Phi(X_{k-1}, H_{k-1}), \]  

(3a)

\[ B_k = \Psi(X_k, H_k), \]  

(3b)

with \( \Phi \) and \( \Psi \) nonlinear functions. Here \( X_k \) denotes the state of the system at the time \( k \), in the sense of the values of a set of system parameters that contains sufficient information to determine the output \( B_k \), when the input \( H_k \) is known. Note that even if the input-state equation [Eq. (3a)] is a dynamic relation, the output \( B_k \) can be derived from the state \( X_k \) and the input \( H_k \) using the static relation [Eq. (3b)].

By considering neural networks, a dynamic (recurrent) neural network could be used to model the entire input–output relation [Eq. (3)]. However, these types of networks are rather complicated and not sufficiently theoretically underlain yet. On the contrary, another type of neural network, the static feed-forward neural network (FFNN), (Fig. 2), is very well developed, with standard algorithms available. The outputs \( y_i \) of such a network are determined as weighed sums of its inputs \( u_j \), combined with a nonlinear sigmoidal activation function \( g \) (a linear activation \( g \) is used for the output layer):

\[ y_i = \sum_{j=0}^{M} w_{ij} g \left( \sum_{k=0}^{N} w_{kj} g \left( \sum_{l=0}^{n} w_{lj} u_l \right) \right) \]

\( l = 0, 1, \ldots, m \).

The weights \( w_{ij}^{(l)} \) are determined by network training. \( 2 \) A theorem states that any continuous nonlinear function \( f \) of an arbitrary number of variables can be approximated arbitrarily well over a compact interval by a multilayer FFNN consisting of one or more hidden layers, provided the number of hidden units is sufficiently large. \( 2 \) The FFNN can thus model static relations like the state-output relation [Eq. (3b)] with arbitrary accuracy. For modeling of dynamic systems, the system state should then be determined separately (if possible) and presented as an input to the FFNN. \( 4,5 \)

We investigate whether this approach can be used to model magnetic hysteresis. According to Eq. (1), the state of the system \( X_k \) in the sense of Eq. (3b) consists of the values of the three parameters: \( H_k^{\text{extr}}, B_k^{\text{extr}} \) and \( \text{FLAG}_k \). They can easily be determined at each time step \( k \) from the input and output history. FFNN with four inputs can therefore approximate the material behavior with arbitrary accuracy:

\[ B_k = \text{FFNN}(H_k^{\text{extr}}, B_k^{\text{extr}}, \text{FLAG}_k, H_k). \]  

(4)

FIG. 2. Feed-forward neural network with two hidden layers.

\[ y_i = \sum_{j=0}^{M} w_{ij} g \left( \sum_{k=0}^{N} w_{kj} g \left( \sum_{l=0}^{n} w_{lj} u_l \right) \right) \]

\( l = 0, 1, \ldots, m \).

The weights \( w_{ij}^{(l)} \) are determined by network training. \( 2 \) A theorem states that any continuous nonlinear function \( f \) of an arbitrary number of variables can be approximated arbitrarily well over a compact interval by a multilayer FFNN consisting of one or more hidden layers, provided the number of hidden units is sufficiently large. \( 2 \) The FFNN can thus model static relations like the state-output relation [Eq. (3b)] with arbitrary accuracy. For modeling of dynamic systems, the system state should then be determined separately (if possible) and presented as an input to the FFNN. \( 4,5 \)

We investigate whether this approach can be used to model magnetic hysteresis. According to Eq. (1), the state of the system \( X_k \) in the sense of Eq. (3b) consists of the values of the three parameters: \( H_k^{\text{extr}}, B_k^{\text{extr}} \) and \( \text{FLAG}_k \). They can easily be determined at each time step \( k \) from the input and output history. FFNN with four inputs can therefore approximate the material behavior with arbitrary accuracy:

\[ B_k = \text{FFNN}(H_k^{\text{extr}}, B_k^{\text{extr}}, \text{FLAG}_k, H_k). \]  

(4)

FIG. 3. Network training set: (a) network input; (b) target output.
However, with regard to computational efficiency and required training sets, it is more convenient to model only the nonlinear part of Eq. (1), and use

\[ B_k = B_k^{\text{extr}} + \text{FFNN}(H_k^{\text{extr}}, \text{FLAG}_k, H_k), \]  

or, equivalently

\[ \Delta B_k = B_k - B_k^{\text{extr}} = \text{FFNN}(H_k^{\text{extr}}, \text{FLAG}_k, H_k). \]

Note again the similarity with the Everett function description [Eq. (2)]. The static relation [Eq. (6)] can be approximated arbitrarily well by using a FFNN with three inputs derived at each time step \( k \) from the input field sequence. Equation (6) thus provides an arbitrarily accurate rate-independent hysteresis model, obeying the congruency property strictly.

RESULTS AND DISCUSSION

In order to use the neural network [Eq. (6)] for hysteresis loop prediction, it should be trained, i.e., the weights \( w_{ij}^{(k)} \) determined. A suitable training set of known input and output pairs is used for this task.\(^2\) The (measured) training set should span the whole range of possible network input values, because a neural network is capable of performing accurate nonlinear interpolations, but is not suitable for extrapolation. For a fixed maximum field strength \( H^{\text{max}} \), a possible training set providing all the necessary information is shown in Fig. 3. This set provides the value of the left-hand side expression in Eq. (6) for each possible physically meaningful combination of the network inputs \( H_k^{\text{extr}} \), \( \text{FLAG}_k \) and \( H_k \). Note that the virgin curve is included in the set, and \( \text{FLAG}_k \) could only equal \(-1\) if \( H_k^{\text{extr}} = 0 \). Instead of using a measured training set, for this article it was generated by the Preisach model, using a quasistatistically measured Everett function of a nonoriented 3 wt% SiFe material.\(^5\) A neural network with 2 hidden layers and 10 neurons in each hidden layer was trained for 1000 epochs (iterations) with the training set presented in Fig. 3 using the Levenberg–Marquardt training algorithm.\(^2\) The excellent correspondence between the neural network output and the Preisach model output for the training set from Fig. 3 is shown in Fig. 4. The neural network model was also tested with a test set shown in Fig. 5(a), which was not used during training. Figure 5(b) shows that both major and minor hysteresis loops can be predicted very accurately, thereby proving that the neural network technique is suitable for hysteresis modeling. The neural network performance as a function of the density of the training set and the number of hidden units, as well as its behavior for incomplete or noisy training sets, will be studied in the future. The extension of the technique to dynamic and vector hysteresis will be investigated as well.

ACKNOWLEDGMENTS

This work was supported financially by GOA Project No. 99-200/4, by research Project No. 39042099 of FWO-Vlaanderen, by STWW Project No. IWT 980357 and by IUAP Project No. 120C1597. One of the authors (L.D.) is a postdoctoral researcher at FWO-Vlaanderen.

---