A COMPARISON OF NEURAL NETWORK AND POLYNOMIAL MODELS FOR THE APPROXIMATION OF NON-LINEAR AND ANISOTROPIC FERROMAGNETIC MATERIALS

H Vande Sande, K Hameyer

Katholieke Universiteit Leuven, Belgium

INTRODUCTION

Polynomials fail to give suitable approximations for strongly non-linear functional mappings. In that case, neural networks can preferably be used. Over the past decade, their popularity steadily increased within various engineering disciplines. Here, it is shown that neural networks must not always be preferred over traditional polynomials. When modeling typical non-linear and anisotropic magnetic properties for e.g. finite element simulations, both approximations are fairly competitive.

MATERIAL DATA

The neural network model and the polynomial model are compared for two distinct material properties:

- The non-linear magnetization curve of a non grain-oriented steel. It is a smooth single input single output functional mapping $B = f(H)$, with B the magnetic induction [T] and H the magnetic field strength [A/m].
- The non-linear and anisotropic reluctivity curves of a grain-oriented steel (Fig. 1). It is a smooth single input double output functional mapping $\nu = f(B,\theta)$, with ν the reluctivity [Am/Vs] and θ the magnetization angle [°].

For accuracy reasons, the magnetic field strength and the reluctivity are transformed in a first step by a base ten logarithm. In a next step, all variables are normalized.

POLYNOMIAL MODELS

The non-linear magnetization curve is modeled by a polynomial of order N:

$$B = \sum_{n=0}^{N} a_n H^n \qquad (1)$$

By applying Horner's rule, it can be proven that this polynomial can be evaluated with $2N$ floating point operations (flops). The model has N+1 degrees of freedom (dofs). The number of flops per dof is denoted by the ratio R. For large N, this ratio approaches 2.

The non-linear and anisotropic magnetization curves are modeled by a two-dimensional polynomial of order N:

$$\nu = \sum_{n=0}^{N} \left(\sum_{m=0}^{N-n} a_{n,m} B^m \right) \theta^n \qquad (2)$$

This polynomial is evaluated similar to Horner's rule within $N^2 + 4N - 1$ flops. It contains $(N+1)(N+2)/2$ dofs. For large N, the ratio R approaches the value 2.

NEURAL NETWORK MODELS

Among the large amount of neural network types that can be used for regression purposes, the two-layer perceptron is selected for the comparison. According to Bishop (1), this perceptron is mathematically described as

$$B = \sum_{j=1}^{M} q_{1j} \phi\left(p_{j1} H + p_{j0}\right) + q_{10} \qquad (3)$$

and

$$\nu = \sum_{j=1}^{M} q_{1j} \phi\left(p_{j2} B + p_{j1} \theta + p_{j0}\right) + q_{10} \quad , \qquad (4)$$

for both problems respectively, with p and q the network weights. The function ϕ equals

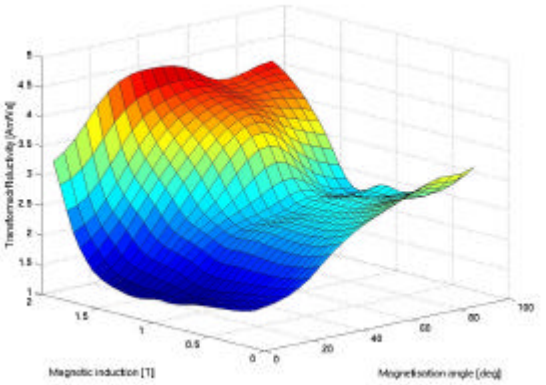

Figure 1: The non-linear and anisotropic reluctivity of a grain-oriented steel.

a. b.

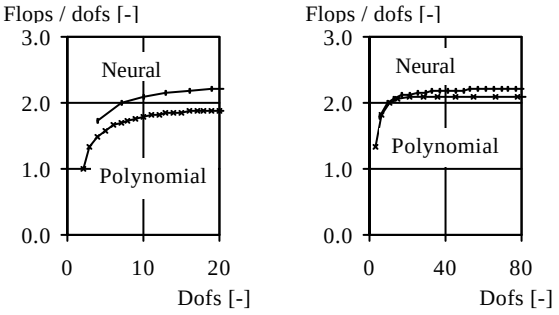

Figure 2: The number of flops per dof for
(a) the single input problem;
(b) the double input problem.

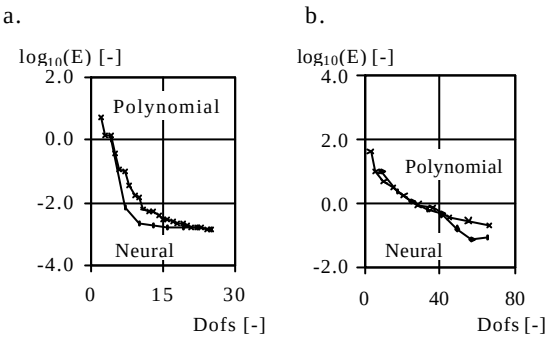

Figure 3: The sum-of-squares error for
(a) the single input problem;
(b) the double input problem.

$$\phi(a) = \frac{1}{1+e^{-a}} \tag{5}$$

It represents the non-linear transformation units or neurons in the network. Eqn. 3 has $3M+1$ dofs and is evaluated in $7M$ flops. The ratio R therefore approaches 2.33 for large values of M. Eqn. 4 is evaluated in $9M$ flops and contains $4M+1$ dofs. Here, the ratio R approaches 2.25 for large M.

DETERMINATION OF THE COEFFICIENTS

The coefficients of the polynomials are determined by solving an overdetermined system of equations using the least squares method, without constraints. The neural network is optimized by an unconstrained minimization of the sum-of-squares error

$$E(p,q) = \sum_{d=1}^{D} \left(y_d^{neur}(p,q) - y_d^{meas} \right)^2 \tag{6}$$

with respect to the network weights, as in Vande Sande et al (2). Here, D is the number of measurements, y_d^{neur} the network output for measurement d and y_d^{meas} the measured value. Several tests are performed and the best one is retained (1).

COMPARISON

For both models and both properties, figures 2a &b show the number of flops per dof, figures 3a & b the remaining sum-of-squares error and figures 4a & b a measure for the curvature C of the solution (1):

$$C = \sum_{d=1}^{D} \sum_{i=1}^{I} \left(\left. \frac{\partial^2 y}{\partial x_i^{\,2}} \right|_d \right)^2 , \tag{7}$$

with I the number of inputs and, y the model's output and x_i the i^{th} input.

Figure 2 reveals that neural networks require more flops for the same number of dofs, especially for the single input problem. Moreover, M of those flops are time consuming exponential function evaluations and the computer code for the evaluation of a perceptron is much longer than the code for a polynomial. On the other hand, the accuracy and the curvature of the neural

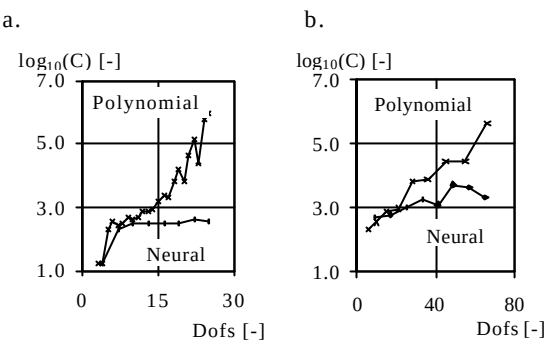

Figure 4: The curvature of the solution for
(a) the single input problem;
(b) the double input problem.

network models tend to be better than for polynomial models for the same number of dofs, as demonstrated by figures 3 & 4. The increase of the curvature towards a higher number of dofs is also more significant for polynomials.

CONCLUSIONS

A comparison between polynomial and neural network models for two important magnetic properties is presented. From a computational point of view, polynomials are easier to evaluate than neural networks are. Hence, for the same amount of computation time, polynomials can have slightly more dofs. This results in a lower sum-of-squares error for the polynomial, which may then be comparable with the neural network's error, depending on the specific case. If the curvature is not affected too much, the polynomial approximation should not be rejected a priori, particularly when the material model has to be evaluated frequently, e.g. in every element of a finite element mesh. Conversely, if the curvature of the solution should be kept low for numerical reasons, e.g. to ensure or to improve the convergence of an algorithm, as in (2) and Pahner et al. (3), the neural network model might be favorably used.

ACKNOWLEDGEMENTS

The authors are grateful to the Belgian "Fonds voor Wetenschappelijk Onderzoek Vlaanderen" for its financial support of this work and the Belgian Ministry of Scientific Research for granting the IUAP No. P4/20 on Coupled Problems in Electromagnetic Systems.

REFERENCES

1. Bishop,C.M.,1995,"Neural networks for pattern recognition",Oxford Univ.Press,New York,USA

2. Vande Sande,H.,Pahner,U.,De Gersem,H., Hameyer,K.,2001,CD Proc. of ISTET Linz,nr.117

3. Pahner,U,Belmans,R.,Hameyer,K.,1999,COMPEL, 18, 629-637.