Gray-Box Modeling of Mechanical Loads for Electric Drive Systems using Neural Networks

Miguel Vélez-Reyes and Roberto Rivera-Sampayo
Center for Power Electronics Systems
University of Puerto Rico-Mayagüez
P.O. Box 9048 Mayagüez, PR 00681-9048
PUERTO RICO, USA

Abstract— Modeling of Electric Motors coupled to complex nonlinear mechanical system is a challenging task. However, high performance control of electric drives systems requires accurate models of mechanical loads. Gray-Box models using neural networks are used here to identify the electrical machine-mechanical load system of an electric drive. The use of this approach for electric drive identification is illustrated with a DC motor drive system driving an unknown static load. Simulation results are presented to demonstrate the capability of the proposed approach.

Key-words: - Neural Networks, gray-box modeling, electric drives, system identification

1. Introduction

Modern control strategies for electric drives require accurate knowledge of the machine parameters and the dynamic characteristics of the actuated load. In general, we have a good understanding of the dynamics of the electrical drive while good models of the mechanical load are not available because of poor knowledge or load complexity.

A methodology to overcome these problems based on gray-box modeling is presented in this work. In gray-box modeling, a system is partitioned into two components. One component is well understood from physics principles. This is the case of the electrical model of an electric drive. The second component of the model is unknown or partially unknown. To model this part of the system, a black-box model is used. In this work an integrated gray-box model is proposed developed and studied in the modeling of a DC motor drive system. Neural networks are used to construct the black-box model for the mechanical load. The reason to use neural networks is their adaptive and nonlinear nature.

The work is presented here as follows. First, the concept of gray-box models is introduced. Then the case study is presented and simulation studies are used to illustrate the underlying concepts. The problems found and the tools used are discussed. Finally, the concluding points and future work are presented.

2. Gray-Box Modeling

In modeling complex dynamic systems, scientists and engineers often face the fact that physics-based models of all system components are not available due to poor understanding of the component dynamics or their complexity. The gray-box modeling methodology helps to address this problem by combining prior physical knowledge with the simplicity of black box model structure [1] to take as much advantage as possible of the available physical information. We will describe the gray-box model structure used in this research next.

Let us assume that the system of interest can be modeled by a state space representation

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \tag{1}$$

Let us further assume that the system can be split in two additive components,

$$\dot{\mathbf{x}}(t) = \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t)) + \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t)) \tag{2}$$

where $g(\cdot)$ is the known part of the system and $h(\cdot)$ is the unknown part. For identification purposes,

the unknown component is substituted by a black box model leading to the gray-box model structure

$$\dot{\mathbf{x}}(t) = \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t)) + BB(\mathbf{x}(t), \mathbf{u}(t)) \qquad (3)$$

where **BB**(x(t), u(t)) is the black box component of the model. Other architectures are possible. Parameter estimation is applied to (3) using any standard parameter estimation method. In this paper, we use neural networks for the black box component of the gray-box model.

3. Case Study

3.1 DC Motor Drive System

Our main interest is to study the capabilities of the gray-box modeling approach for electric drive modeling. For this purpose, the identification of a permanent magnet DC motor driving a nonlinear static load is studied. A schematic of the system is presented in Figure 1.

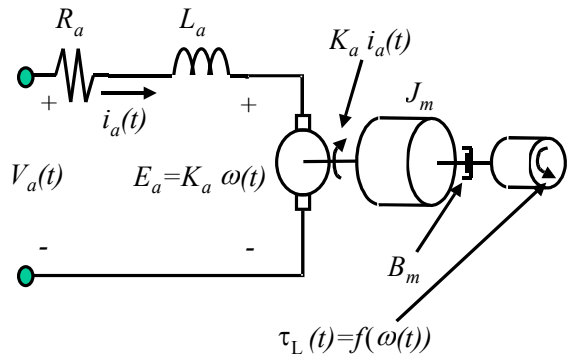

Fig. 1. Schematic of the DC drive system

Table 1: Expressions for case study loads.

Case	f(ω)
Fan Load	$\tau_L(\omega) = \mu\,\mathrm{sgn}(\omega)\omega^2$
Nonlinear Friction	$\tau_L(\omega) = \mathrm{sgn}(\omega)\left(\mu \tan^{-1}\left(\frac{\alpha}{\lvert\omega\rvert}\right) + \beta\right)$

here R_a, L_a, are the armature resistance and inductance, and K_a is the back EMF/torque constant; $i_a(t)$, and $v_a(t)$ are the armature current, and voltage; J_m, and B_m represent the inertia, and viscous friction coefficients; $\tau_L(\omega)$ is the speed dependent load torque; and $\omega(t)$ is the rotor speed.

The state space model for the DC motor drive system is:

$$\begin{aligned} \frac{di_a(t)}{dt} &= -\frac{R_a}{L_a} i_a(t) - \frac{K_a}{L_a}\omega(t) + \frac{1}{L_a} v_a(t) \\ \frac{d\omega(t)}{dt} &= \frac{K_a}{J_m} i_a(t) - \frac{B_m}{J_m}\omega(t) - f(\omega(t)) \end{aligned} \qquad (4)$$

In our study, we will consider two different loads in which $\tau_L(\omega)$ will be function of the speed but this dependency will be assumed to be unknown. The expressions for the three loads under consideration are presented in Table 1. These loads are found in typical applications of electric drives. The load parameters in all the cases are such that the load torque equals rated torque at rated speed. The load parameters are given in Table 2.

Table 2: Load Parameters.

Case	Parameter Values
Nonlinear Friction	μ=1 N-m, α=1 rad/sec, β=3 N-m
Fan Load	μ=0.039 N-m s^2/rad^2

3.2 The Gray-Box Model

In our modeling work, the static load will be assumed to be an unknown static function of the speed. Based on (3), a gray-box model for the drive is given by.

$$\begin{aligned} \frac{di_a(t)}{dt} &= -\frac{R_a}{L_a} i_a(t) - \frac{K_a}{L_a}\omega(t) + \frac{1}{L_a} v_a(t) \\ \frac{d\omega(t)}{dt} &= \frac{K_a}{J_m} i_a(t) - \frac{B_m}{J_m}\omega(t) - \frac{1}{J_m} NN(\omega(t)) \end{aligned} \qquad (5)$$

where NN(ω) is the black box model using the neural network.. The selection of the neural network model structure for this case is intuitive. The speed dependent loads under consideration are odd functions that suggest the use of neural networks with hyperbolic tangent basis functions.

3.3 Nonlinear Least Squares Parameter Estimation

Physical parameter estimates and network weights are computed by minimizing a quadratic cost function. The nonlinear least-squares parameter estimation problem can be formulated as the optimization problem:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\theta} S(\boldsymbol{\theta}) \tag{6}$$

where $\hat{\boldsymbol{\theta}}$ is the estimate of $\boldsymbol{\theta}$ and

$$S(\theta) = \sum_{i=1}^{n} r_i(\boldsymbol{\theta})^2 = \sum_{i=1}^{n} [y_i - f_i(\boldsymbol{\theta})]^2 = \|\mathbf{y} - \mathbf{f}(\boldsymbol{\theta})\|^2 \tag{7}$$

is the quadratic cost function that measures the two norm of the model prediction error. The standard method to compute the parameter estimate is given by the Gauss-Newton method.

The Gauss Newton method is an iterative method of the form

$$\hat{\boldsymbol{\theta}}^{(a+1)} = \hat{\boldsymbol{\theta}}^{(a)} + \gamma^{(a)}\mathbf{p}^{(a)} \tag{8}$$

where $\hat{\theta}^{(a)}$ is the estimate at the a-th iteration, $\gamma^{(a)}$ is the step size, and $p^{(a)}$ is the Gauss-Newton search direction computed by solving the linear least squares problem

$$\mathbf{p}^{(a)} = \arg\min_{\mathbf{p}} \left\| \mathbf{r}^{(a)} - \mathbf{J}^{(a)}\mathbf{p} \right\| \tag{9}$$

where $\mathbf{r}^{(a)}$ is the residual vector and $\mathbf{J}^{(a)}$ is the Jacobian at the a-th iteration given by

$$\mathbf{r}^{(a)} = \mathbf{y} - \mathbf{f}\left(\hat{\boldsymbol{\theta}}^{(a)}\right), \ \mathbf{J}^{(a)} = J(\hat{\boldsymbol{\theta}}^{(a)}) = \left.\frac{\partial \mathbf{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\theta=\hat{\theta}^{(a)}} \tag{10}$$

A problem during the parameter estimation problem is that of ill conditioning due to the overparameterization (large number of weights) of the neural network that results in a nearly singular Jacobian. To deal with this difficulty, we used the truncated singular value decomposition of the Jacobian where small singular values are set equal to zero when solving (9). This results in improved convergence and more stable results. For derivations of the Jacobian and details on ill-conditioning problems see [7,8] for more details.

4. Simulation Results

For the case of study, the parameters of the DC drive systems are given it Table 3 [5]. The motor is a 220 V, 550 rpm and 1hp permanent magnet DC motor. The experiment is conducted in open loop with no current loops or speed feedback. The armature excitation voltage is shown in Figure 2. This provides an operating speed range from ±57 rad/sec.

For experimental purposes the DC drive is simulated under the rated operation conditions using the Matlab™ software package. The simulated data is collected and used in the estimation of the model parameters.

We will examine 4 candidate models based on (5) where the mechanical equation is modified as follows

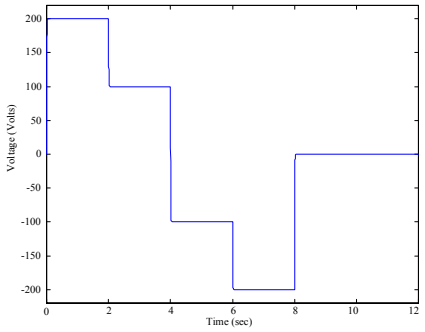

Fig. 2: Excitation Voltage for Open Loop Experiments.

Explicit viscous friction term:

$$\frac{d\omega(t)}{dt} = \frac{K_a}{J_m} i_a(t) - \frac{B_m}{J_m}\omega(t) - \frac{1}{J_m} NN(\omega(t)) \tag{11}$$

All load torque embedded in the neural network:

$$\frac{d\omega(t)}{dt} = \frac{K_a}{J_m} i_a(t) - \frac{1}{J_m} NN(\omega(t)) \tag{12}$$

We also considered 2 and 3 layer feedforward neural networks given by

$$NN(\omega) = \mathbf{W}_2 \tanh(\mathbf{W}_1\omega + \mathbf{b}_1) + \mathbf{b}_2 \tag{13}$$

$$NN(\omega) = \mathbf{W}_3 \tanh(\mathbf{W}_2 \tanh(\mathbf{W}_1\omega + \mathbf{b}_1) + \mathbf{b}_2) + \mathbf{b}_3 \tag{14}$$

where $\mathbf{W}_i$ and $\mathbf{b}_i$ represent the ith layer weight matrix and bias vector.

The 49 elements of the gray model parameter vector is given by

$$\boldsymbol{\theta} = \left[\hat{R}_a \ \ \hat{L}_a \ \ \hat{K}_a \ \ \hat{B}_m \ \ \hat{J}_m \ \ rsW_n \ \ \ldots \ \ rsW_1 \ \ rsd_n \ \ \ldots \ \ rsd_1\right]^T$$

where the bias and weight matrices of the network are arranged as a row string vectors $rs_{(\cdot)}$.

Combination of (11) and (13) is model A, (11) and (13) is model B, (12) and (13) is model C, and (12) and (14) is model D.

The parameters were computed using the Gauss Newton method modified with the Truncated Singular Value Decomposition described previously. A maximum of 11 iterations was set as a stopping criterion. Our results were compared in terms of the error in the estimate of the physical parameters and in the approximation of the neural network to the nonlinear load curve.

Table 3: PARAMETERS OF THE MOTOR AND INITIALIZATION VALUES.

	Actual Value	Initial Estimate
Armature Resistance R_a	7.56 Ω	6.4260 Ω
Armature Inductance L_a	0.055 H	0.0468 H
Torque Coefficient K_a	3.475 N-m/A	3.0233 N-m/A
Viscous Friction Coeff. B_m	0.03475 N-m-s	0.0313 N-m-s
Inertia J_m	0.06 Kg-m^2	0.0612 Kg-m^2

4.1 Initialization

An important part of the identification is initialization of the network to reduce the initial approximation error. This was done by pre-training the network to data taken during steady-state operation and with prior estimates of K_a and B_m.

In all cases, the physical parameters were initialized to the values shown in Table 3.

4.2 Fan Load

The estimated values for this case are shown in Table 4. We only show the estimates for the physical parameters, which are compared to actual values as away to evaluate the goodness of the identified model. We can see that the only bad estimate was that of B_m. It is our belief that this is because the network tries to fit the viscous torque as part of the load torque causing errors in that parameter estimate.

Figures 3 and 4 compare the load torque characteristic with the identified characteristic using the neural network plus viscous damping combination as in models A and B and the network by itself as in models C and D. In both cases the identified load shows good agreement within the range of speeds where the system was excited ±57 rad/sec. Figures 5 and 6 show validation results where the performance of the identified model and actual system are compared in a transient different from that used for identification. Models B and D have the worst performance in terms of prediction error. Models B and D are the ones where a three layer network is used.

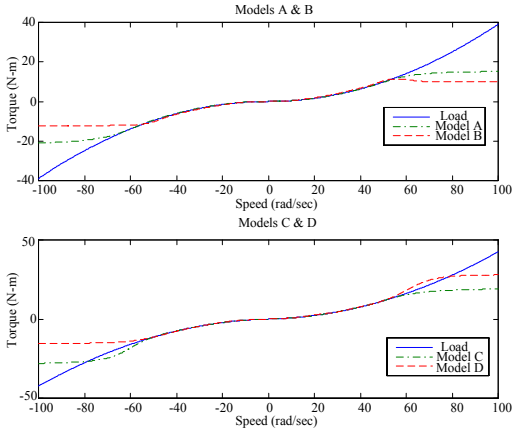

Fig. 3: Estimated and actual fan characteristics.

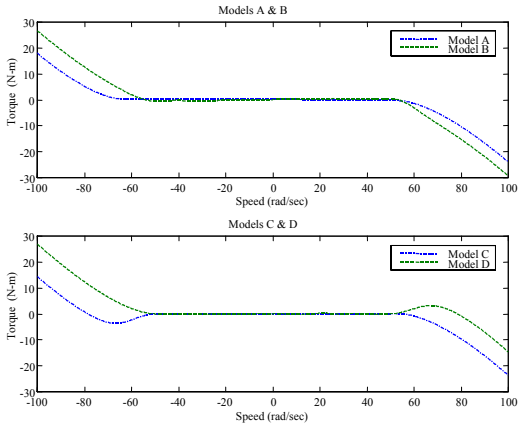

Fig. 4: Estimation error for fan characteristic.

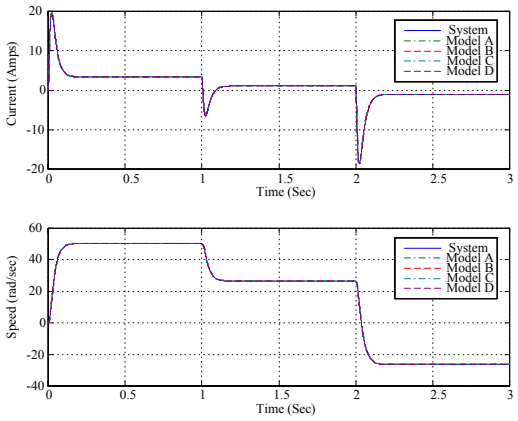

Fig. 5: Validation results for identified drive with fan load.

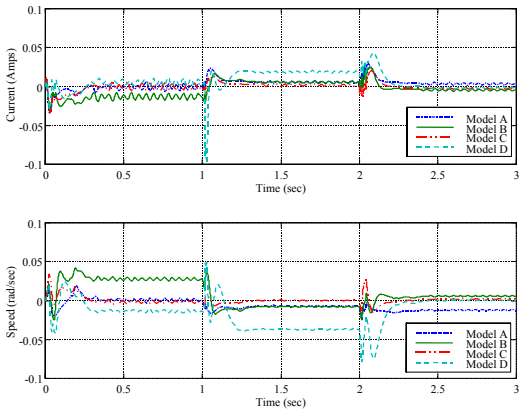

Fig. 6: Model prediction error with validation data.

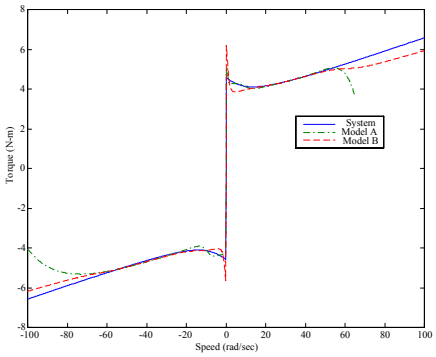

Fig. 7: Estimated versus actual friction load characteristic.

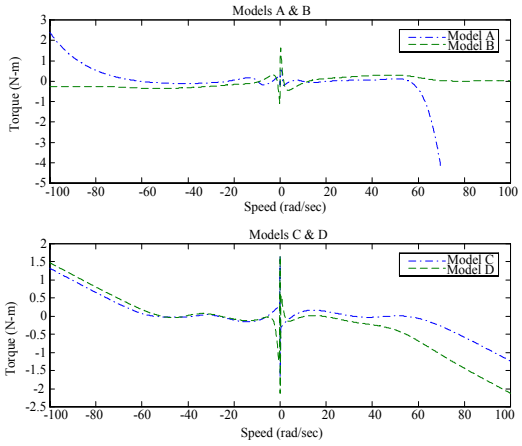

Fig. 8: Error in friction load characteristic estimation.

4.3 Nonlinear Friction Load

The physical parameter estimates for this case are shown in Table 5. Results are similar to those for the fan load. We observe low errors in all estimates except B_m. The estimated load characteristic and corresponding estimation error are shown in Fig. 7 and 8. This is a harder model to estimate because of the discontinuity at the origin. There is where the largest errors occur in all cases. The best-fit overall is achieved by model B, which is the model with the largest number of parameters. Model B includes viscous friction and a three-layer network. It is reasonable to expect that the larger number of degrees of freedom can handle the modeling of friction better. Validation runs shown in Figs. 9 and 10 show in general relatively good agreement for all models with best performance for model B and worst performance for model D.

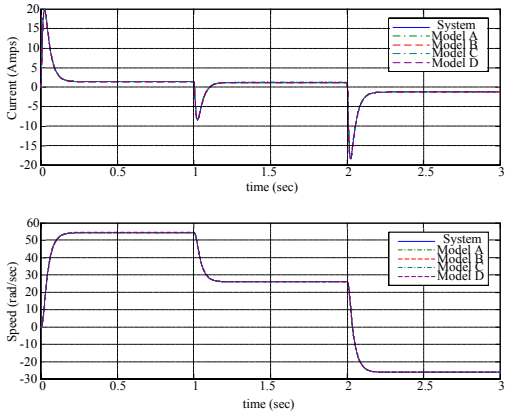

Fig. 9: Validation results for friction load case.

4.4 Comparison between model structures

All models in general did a good job in fitting the data. Two aspects are important to evaluate the closeness to actual system structure achieved by including the viscous friction term and the higher number of degrees of freedom achieved by 2 or three layers. When comparing structure including the viscous friction term always resulted in a model with better characteristics A performs better than C and B performs better than D regardless of the load, which points out to the fact that the closer the model structure to the actual physical structure, the better should be the model performance.

In terms of the number of degrees of freedom, we get mixed results depending on the characteristics of the load. In the case of the continuous load, the extra number of degrees of freedom seems to worsen the identification: A does better than B and C does better than D. In the case of the discontinuous load, we get mixed results where B does better than A but C does better than D. It will be beneficial to investigate how pruning methods can help in this kind of situation.

Table 4: Physical Parameter Estimates for the Fan Load Case.

Model	$\hat{R}_a$	%err	$\hat{L}_a$	%err	$\hat{K}_a$	%err	$\hat{B}_m$	%err	$\hat{J}_m$	%err
A	7.586	0.347	0.05484	0.291	3.4733	0.0484	0.03476	7.815	0.0677	0.3812
B	7.583	0.311	0.05482	0.320	3.4735	0.0426	0.0232	33.12	0.0677	0.3812
C	7.581	0.289	0.05480	0.359	3.4736	0.0398	-	-	0.0678	0.2925
D	7.581	0.289	0.05486	0.237	3.4736	0.0398	-	-	0.0678	0.2925

Table 5: Physical Parameter Estimates for the Friction Load Case.

Model	$\hat{R}_a$	%err	$\hat{L}_a$	%err	$\hat{K}_a$	%err	$\hat{B}_m$	%err	$\hat{J}_m$	%err
A	7.580	0.277	0.05494	0.117	3.4744	0.0177	0.03344	3.753	0.0678	0.330
B	7.581	0.282	0.0550	0.010	3.4744	0.0162	0.2795	19.55	0.0677	0.4319
C	7.5807	0.273	0.0548	0.273	3.4743	0.019	-	-	0.0679	0.2251
D	7.5848	0.328	0.0549	0.216	3.4742	0.021	-	-	0.0678	0.3451

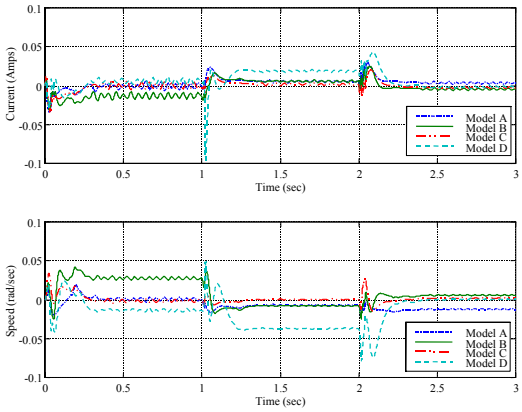

Fig. 10: Error in model prediction for validation case.

5. Conclusions

The modeling approach presented illustrates the potential of gray-box modeling for identification of electric drives with unknown mechanical loads. The convenience of this model is that the identified physical parameters give information about the physical system while the neural network based black-box model part would allow to model different types of loads independently of their actual form and therefore can be used for self-tuning or self-calibration of electric drives. A self-calibration procedure would be able to tune drive performance for multiple mechanical loads with no need for detailed load modeling.

6. Acknowledgements

This work was supported primarily by the ERC Program of the National Science Foundation under Award Number EEC-9731677.

References

[1] R. Rico-Martinez, J.S. Anderson and I. G. Kebrekidis, Continuous-time nonlinear signal processing: a neural network based approach for gray box identification, *Proc. IEEE Workshop on Neural Networks for Signal Processing,* pp. 596 – 605, Oct. 1994.

[2] F. Girosi, and T. Poggio, Neural Networks and best approximation property, *Biological Cybernetics*, 63, pp.169-176.

[3] H. Khalil, *Nonlinear Systems*, 3rd ed., Prentice Hall, 2002.

[4] J.W. Brewer, Kronecker products and matrix calculus in system theory, *IEEE Trans. Circuit Syst.*, 25, pp. 772-781, Sept. 1978.

[5] S. Weerasooriya and M.A. El-Sarkawi, Identification and control of a dc motor using back-propagation neural networks, *IEEE Trans. on Energy Conversion*. 6, pp. 663-669, Dec, 1991.

[6] J. Eriksson, *Optimization and Regularization of Nonlinear Least Squares Problems*, Ph.D. Thesis, Department of Computing Science, Umea University, Sweden, 1996.

[7] R. Rivera and M. Vélez-Reyes, Gray-Box Modeling of Electric Drive Systems using Neural Networks, *Proceedings of the IEEE Conference on Control Applications*, September 5-7, 2001, Mexico.

[8] R. Rivera, *Gray Box Modeling of Electric Drives using Neural Networks*. Master of Science Thesis, University of Puerto Rico Mayagüez, December 2001.