DOI: 10.48011/asba.v2i1.1570

Ensemble Grey and Black-box Nonlinear System Identification of a Positioning System

Walisson Chaves Ferreira Pinto *
Helon Vicente Hultmann Ayala *

* *Department of Mechanical Engineering, Pontifical Catholic University of Rio de Janeiro, RJ, Brazil (e-mails: walisson.chaves@gmail.com and helon@puc-rio.br).*

Abstract: In this work, grey and black-box approaches are used in order to model a electromechanical positioning system (EMPS). An ensemble model is then constructed by combining these two approaches, by using the predictions of both models in order to generate an improved estimated output. Four friction models, in their symmetric and asymmetric versions, namely (i) Coulomb model with finite slope at zero velocity and viscous friction, (ii) Coulomb model with viscous friction, (iii) Tustin friction model, (iv) Coulomb model with viscous friction and Stribeck effect were used to describe the dynamic behavior of the EMPS. The results have shown that the combination of grey and black-box models was able to perform better than the grey-box model and that the proposed friction models are also able to improve the relative error. This encourages further research on the application of the concept of ensemble model construction from machine learning to the nonlinear system identification context towards more accurate model construction.

Keywords: Mechanical and aerospace estimation; Nonlinear system identification; Grey box modelling.

1. INTRODUCTION

System identification involves a set of techniques responsible for building mathematical models of dynamic systems through observed/experimental data (Ljung, 2010). Factors such as the demand for models in process analysis and the practical limitations of first principles encourage the need for identification. The development of models bring benefits, since they can be applied for estimation/forecasting, control and monitoring. They may be used specially for simulations, which are cost and time-effective as well a safe alternative to experiments, they can be also used for testing new designs and strategies (Tangirala, 2015).

The main problem to be considered in system identification is to find a model structure that is adequate so that an suitable model can be built. It should be used prior knowledge and physical insight about the system for the task of selecting the model structure (Sjöberg et al., 1995). According to the amount of prior knowledge, models can be classified as white, grey and black-box. In the former, the model is perfectly known and its structure and parameters are entirely determined based on first principles. In the second, some physical insight is available, parameters must be determined from observed data though (Bohlin, 1994). Finally, in the black-box approach, a model is built based on observed data alone, it means that no physical insight is available or used. In this case the model describes the experimental data without any physical interpretation of its parameters (Sjöberg et al., 1995; Kerschen et al., 2006).

In the case of nonlinear black-box models, artificial neural networks (ANN) can be used for mapping nonlinearities, since they are universal approximators. We may cite a multi-layer perceptron ANN for example (Haykin, 2009).

There are many applications of a black-box approach. Irigoyen and Miñano (2013) used a nonlinear auto regressive with exogenous inputs (NARX) neural network in order to enhance cardiovascular rehabilitation therapies. The focus of the authors was the relationship between the required exercise (machine resistance) and the patient's heartbeat for an optimal training configuration. The model was efficient to reproduce the evolution of the heart rate in controlled cardiovascular aerobic training. Wunsch et al. (2018) applied a NARX ANN model for forecasting groundwater levels of three different sort of aquifers. Forecasts of lead times up to half a year were conducted. The results of their studies have shown that the NARX ANN models are well suited to perform groundwater predictions for uninfluenced observation wells in all aquifer types studied. Samara et al. (2013) introduced a time-dependent functional NARX methodology for the development of aircraft virtual sensors, the authors focused on the angle-of-attack for the main flight regimes. The main flight regimes (landing, take-off, clean flight) of a small commercial aircraft were considered for three different virtual sensor designs. Through a nonlinear 6 DOF simulation environment the performance of the developed virtual sensors was tested, they have shown to meet the design requirements, since they achieved simulations errors lower than the required. Tijani et al. (2014) proposed a

hybrid of conventional backpropagation training algorithm for the NARX network and multi-objective differential evolution algorithm for identification of the nonlinear dynamics of an unmanned small-scale helicopter from experimental flight data. This approach was able to yield a set of Pareto-solutions with optimal compromise between model accuracy and model complexity. In Worden et al. (2007) different physics-based and black-box approaches, such as Maxwell-Slip models and neural networks, were used in order to model the non-linear dependence of pre-sliding and sliding friction forces on displacement and velocity. The models have shown a suitable capability of friction prediction. An ensemble of the best models for prediction was also built, by this improvements in performance have been achieved.

An ensemble approach is composed of a set of combined models that act together to predict a response variable error (Ribeiro and dos Santos Coelho, 2019). Over the past two decades this approach has received increasing attention from the computational intelligence and machine learning community, since ensemble systems have proven to be very effective and versatile in many real-world problem domains and applications (Polikar, 2012). Some examples of application of the ensemble approach are load forecasting (Papadopoulos and Karakatsanis, 2015; Qiua et al., 2017; M.Saviozzi et al., 2017), output power forecast (Sperati et al., 2016; Raza et al., 2018) and fault detection (Nogoseke et al., 2017). Wong and Worden (2007) used a tribometer as an experimental rig for the modeling of dry friction using a NARX type shunting ANN model. They also constructed an ensemble model by combining the shunting neural network model with the physics based dynamic nonlinear regression Maxwell slip model, which achieved good accuracy in modelling the overall dynamical behavior. Zhang (2003) used three real data sets in order to explore a hybrid methodology that combines both ARIMA and ANN models, the aim was to take advantage of the unique strength of each model. The results indicate that the combined model can be an effective way to improve forecasting accuracy achieved by either of the models used separately.

Janot et al. (2017) and Brunot et al. (2018) used different approaches in order to estimate the dynamic parameters of a Electromechanical Positioning System (EMPS). In the former work the authors compared an inverse dynamic identification method with least squares and a state-dependent parameter method of nonlinear model estimation. In the latter, an automated instrumental variable method is used, the aim is the improvement of the usual IDIM-Instrumental Variable method with an identification of the remaining noise. Pinto and Ayala (2019) tested four symmetric and four asymmetric friction models to describe the nonlinear dynamics of the EMPS using a inverse dynamic model (IDM). System identification is also applied to the task of friction modeling, we may cite the work of Yoon and Trumper (2014) in which the Generalized Maxwell-Slip (GMS) friction model is adopted in order to study aspects of friction both in the sliding and presliding regimes. The authors developed a frequency-domain method to identify the model parameters based on the frictional resonances. Wang et al. (2004) presented an approach that uses a support vector machine regression for modeling friction for servo-motion systems as an alternative to static friction models. In the work of Huang and Chiu (2009) a LuGre friction model (de Wit et al., 1995) is used for the modeling of the friction dynamics of a piezoelectric actuating system. The authors employed a genetic algorithm approach in order to search for the optimal parameters of the LuGre friction model.

In this context, the present work applies grey and black-box identification as well an ensemble approach to the case study of a EMPS. In addition, four symmetric and four asymmetric friction models were compared. Since the aim of the work is to better simulate the system's response, the ensemble was built in order to achieve a more accurate response when compared to the pure grey-box approach. The results obtained indicate that: (i) the ensemble of both grey and black-box has the highest accuracy for prediction the system's output; (ii) models that smooth the discontinuity at zero velocity of Coulomb friction model have better performance. The contributions of the paper are thus: (i) the analysis of different friction models for grey-box modeling and (ii) the analysis of ensembles of hybrid grey and black-box models towards more precise simulation of a system involving friction, which are of great interest in positioning systems.

The rest of the paper is organized as follows. Section 2 brings a short description of the EMPS and the input and output data. In Section 3 a brief description of some friction models is provided, as well the methodology used for the formulation of the three approaches. The results are included in Section 4, while Section 5 gives conclusions.

2. CASE STUDY

The electromechanical positioning system (EMPS) is described in details by Janot et al. (2019) and all datasets used in this paper can be found on the Electro-Mechanical Positioning System's website as a kind contribution of the author. Basically, the EMPS is a standard configuration of a drive system for prismatic joint robots or machining tools. It consists of a DC motor equipped with an incremental encoder and low-friction high-precision positioning unit. Figure 1 illustrates the diagram of the EMPS, the proportional, derivative and drive gains are K_p, K_v and $g\tau$ respectively. The motor position and the reference of motor position are q_m and q_g respectively and w_q is the noise caused by the encoder. F_m is a perturbation encompassing the friction force and offset effects.

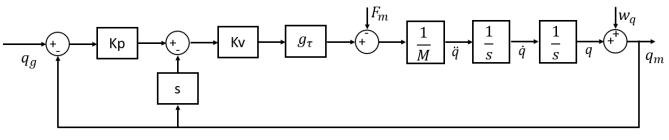

Figure 1. EMPS Diagram. Adapted from Janot et al. (2019).

Figure 2 shows the datasets corresponding to estimation and validation data of the EMPS. The force responsible for the movement of the load is computed by the product of the motor voltage (output of the controller) and the drive gain. The load's velocity and acceleration have to be calculated from the motor position (q_m) measured by the encoder. First, q_m is filtered so that the derivatives

can be calculated with finite differences. Measurements are performed during approximately 25 seconds using a sampling frequency of 1kHz.

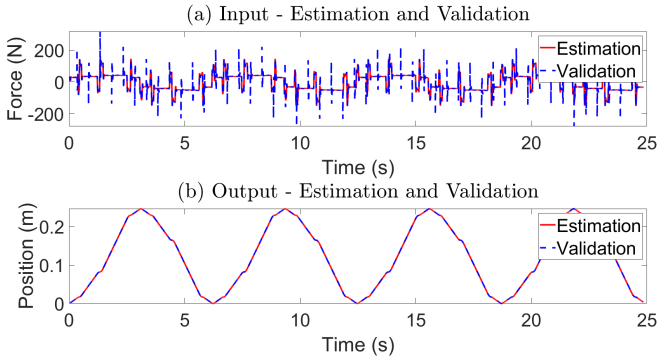

Figure 2. Input and output signals considered for modeling the EMPS.

3. METHODOLOGY

Figure 3 provides a general overview of the methodology employed for the formulation of the system's output. Firstly, a grey-box model (GBM) is used in order to perform the simulation of the system's response ($\widehat{y}_1$). Next, The error e_1 between the true system's output y and $\widehat{y}_1$ is modeled by a black-box approach using a NARX ANN model. By summing the modeled error ($\widehat{e}_1$) and the estimated output of the grey-box model the second estimated output ($\widehat{y}_2$) is generated. Finally, the ensemble (EGBB) is build by the combination of both grey and black-box models. It is given a weight to each single model so that a more accurate response can be estimated. Next we depict each of these three approaches separately.

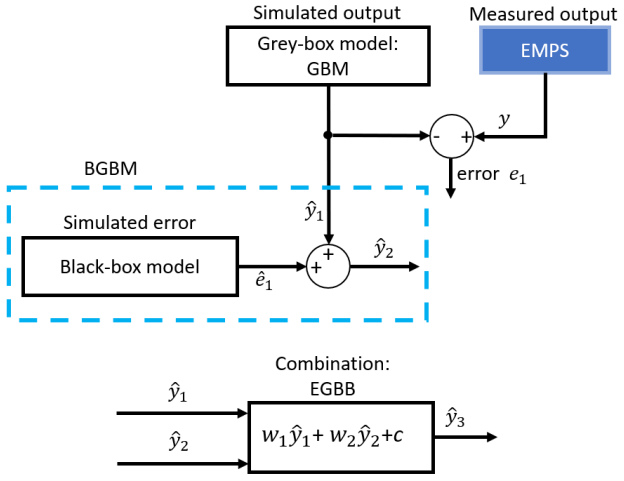

Figure 3. Methodological formulations for modeling the output of the EMPS. Approaches $\widehat{y}_1$, $\widehat{y}_2$ and $\widehat{y}_3$ are used in order to simulate the load's position of the EMPS.

3.1 Grey-box model (GBM)

Section 1 has already addressed some differences between white, black and grey-box approaches. In the latter, model sets have adjustable parameters with physical interpretation (Ljung, 1999), it gives more freedom to change the model allowing prior knowledge about the object of the modeling to be inserted (Bohlin, 2006).

An inverse dynamic model (IDM), which expresses the joint torque/force in terms of the joint position, velocity and acceleration (Khalil and Dombre, 2004) is used to describe the EMPS. The inverse dynamic model of the EMPS is given by (1):

$$\tau_{\mathbf{idm}} = M\ddot{\mathbf{q}} + F_f(\dot{\mathbf{q}}) + offset \tag{1}$$

where $\tau_{\mathbf{idm}}$ is the joint torque/force; M is the mass; F_f is the friction model; $\dot{\mathbf{q}}$ and $\ddot{\mathbf{q}}$ are the load velocity and acceleration respectively; and $offset$ is an offset of measurements.

We are interested in comparing different options for F_f. Some friction force models are presented and compared in the literature, we may cite the work of Marques et al. (2016). For the formulation of $F_f(\dot{\mathbf{q}})$ given in (1), some friction models were selected, namely: (i) Coulomb model with viscous friction, (ii) Coulomb model with finite slope at zero velocity and viscous friction, (iii) Tustin friction model, (iv) Coulomb model with viscous friction and Stribeck effect. Asymmetric and a symmetric version of each model were considered in the present work, they are respectively indicated as F_f^a and F_f^s. In the former different friction coefficients for positive and negative velocities are considered, but this consideration isn't applied to the latter.

The first friction model was presented by Coulomb (Coulomb, 1785), it states that the friction always opposes relative motion between contacting bodies and its magnitude is proportional to the normal contact force. Coulomb model can be described by (2) and Figure 4 represents the friction force as a function of the relative velocity.

$$\mathbf{F} = \begin{cases} F_C sign(\mathbf{v}) & \text{if } \|\mathbf{v}\| \neq 0 \\ min(\|\mathbf{F_e}\|, \mathbf{F_C}) sign(\mathbf{F_e}) & \text{if } \|\mathbf{v}\| = 0 \end{cases} \tag{2}$$

F_C denotes the magnitude of Coulomb friction, which is given by:

$$F_C = \mu \|\mathbf{F_N}\| \tag{3}$$

where $\mathbf{F_N}$ is the normal force of contact, μ is the coefficient of friction, $\mathbf{F_e}$ represents the resultant of the external forces acting on the reference body in the tangential direction of the contact, $\mathbf{v}$ is the relative tangential velocity of the body with respect to the other contacting surface.

Coulomb model with viscous friction Its symmetric and asymmetric versions are (4) and (5) respectively:

$$\mathbf{F_C^s} = F_V \dot{\mathbf{q}} + F_C sign(\dot{\mathbf{q}}) \tag{4}$$

$$\mathbf{F_C^a} = F_V^+ 0^+(\dot{\mathbf{q}}) + F_C^+ sign(0^+(\dot{\mathbf{q}})) + F_V^- 0^-(\dot{\mathbf{q}}) + F_C^- sign(0^-(\dot{\mathbf{q}})) \tag{5}$$

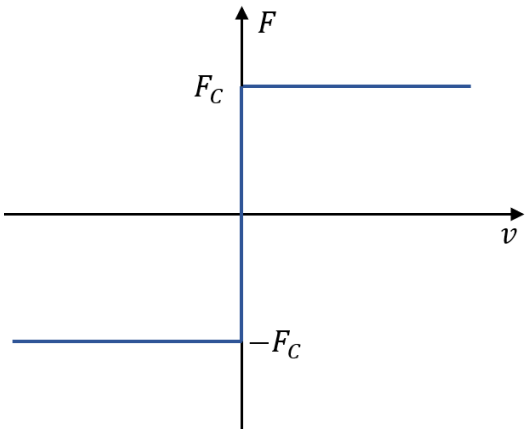

Figure 4. Representation of Coulomb friction for 1D case. Adapted from Marques et al. (2016)

where F_V represents the viscous friction coefficient; F_C^+, F_C^-, F_V^+, F_V^- represent, respectively, the magnitude of Coulomb friction and the viscous friction coefficient of the asymmetric model.

$0^+(\dot{\mathbf{q}})$ and $0^-(\dot{\mathbf{q}})$ denote mathematical operators, as given by (6) and (7). $0^+(\dot{\mathbf{q}})$ returns $\dot{\mathbf{q}}$ if $\dot{\mathbf{q}} > 0$ and 0 otherwise; $0^-(\dot{\mathbf{q}})$ returns -$\dot{\mathbf{q}}$ if $\dot{\mathbf{q}} < 0$ and 0 otherwise.

$$0^+(\dot{\mathbf{q}}) = \dot{\mathbf{q}}\,(1 + sign\,(\dot{\mathbf{q}}))\,/2 \tag{6}$$

$$0^-(\dot{\mathbf{q}}) = \dot{\mathbf{q}}\,(1 - sign\,(\dot{\mathbf{q}}))\,/2 \tag{7}$$

Coulomb model with finite slope at zero velocity and viscous friction The following model presented by Threlfall (1978) replaces the discontinuity at zero velocity of the Coulomb model by a finite slope model in order to smooth out its discontinuity.

$$\mathbf{F_F^s} = \begin{cases} F_V\dot{\mathbf{q}} + F_C\mathbf{K_Z}sign(\dot{\mathbf{q}}) & \text{if } \|\dot{\mathbf{q}}\| \leq v_0 \\ F_V\dot{\mathbf{q}} + F_C sign(\dot{\mathbf{q}}) & \text{if } \|\dot{\mathbf{q}}\| > v_0 \end{cases} \tag{8}$$

$\mathbf{K_Z}$ is defined by (9):

$$\mathbf{K_Z} = 1 - e^{-\frac{3\|\dot{\mathbf{q}}\|}{v_0}}/(1 - e^{-3}) \tag{9}$$

where v_0 is a tolerance velocity.

Tustin friction model According to Tao (2001) Tustin model is one of the best models describing friction force at a velocity close to zero. This model was proposed by Tustin (1947) and can be described by (10):

$$\mathbf{F_T^s} = F_V\dot{\mathbf{q}} + F_C sign(\dot{\mathbf{q}}) + (F_S - F_C)e^{-\frac{|\dot{\mathbf{q}}|}{v_s}} \tag{10}$$

F_S is the static friction coefficient and v_S is the Stribeck velocity.

Coulomb model with viscous friction and Stribeck effect This friction model was introduced by Bo and Pavelescu (1982) and takes into account the Coulomb, viscous, stiction and Stribeck friction effects. This model is given by (11).

$$\mathbf{F_V^s} = F_V\dot{\mathbf{q}} + \left(F_C + (F_S - F_C)e^{-\left(\frac{\|\dot{\mathbf{q}}\|}{v_s}\right)^{\delta_\sigma}}\right) sign(\dot{\mathbf{q}}) \tag{11}$$

where δ_σ is a factor that relies on the geometry of the contacting surfaces.

Similar to $\mathbf{F_C^s}$, models $\mathbf{F_F^s}$, $\mathbf{F_T^s}$ and $\mathbf{F_V^s}$ will also be evaluated in their asymmetric versions. They will be named as $\mathbf{F_F^a}$, $\mathbf{F_T^a}$ and $\mathbf{F_V^a}$.

The decision variables of the EMPS were estimated by a Least Squares (LS) algorithm, these parameters are given in Table 1. Since (4) and (5) are linear in relation to the parameters of the model, a QR factorization was used to solve the linear least squares problem, while the Levenberg-Marquardt Method was employed for the other models, which starts at a initial point and finds a minimum of the sum of squares (12) of the function passed in.

$$\min f(x) = \sum_{i=1}^{n}(y_i - f(x_i, \beta))^2 \tag{12}$$

where x_i is the independent variable and y_i is the values of measurements that expresses the joint torque/force. The optimization algorithm returns the vector of parameters (β) that best fits the model.

Table 1. Models and their parameters.

Model	Decision variables
F_F^s	M, F_V, F_C, v_0, *of*
F_F^a	M, F_V^+, F_C^+, F_V^-, F_C^-, v_0
F_C^s	M, F_V, F_C, *of*
F_C^a	M, F_V^+, F_C^+, F_V^-, F_C^-
F_T^s	M, F_C, F_S, F_V, v_S, *of*
F_T^a	M, F_C^+, F_S^+, F_V^+, F_C^-, F_S^-, F_V^-, v_S
F_V^s	M, F_V, F_C, F_S, v_S, δ_σ, *of*
F_V^a	M, F_V^+, F_C^+, F_S^+, F_V^-, F_C^-, F_S^-, v_S, δ_σ

3.2 Black and Grey-box model (BGBM)

Since the true output of the EMPS can be described by $\widehat{y}_1$ and e_1, as shown by figure 3, the aim here is to model e_1 using a NARX structure, given by (13), and sum the modeled error with $\widehat{y}_1$ in order to obtain a more accurate response than that found by the first approach. The input (u) and output (y) considered are respectively the EMPS force and the error e_1.

$$\begin{aligned} y(k) = F[&y(k-1), y(k-2), ..., y(k-n_y), \\ &u(k-d), u(k-d-1), ..., u(k-d-n_u)] \end{aligned} \tag{13}$$

where $y(k)$, $u(k)$ are the system output and input, respectively; n_y and n_u are respectively the maximum lags for the system output and input; F is some nonlinear function, and d is a time delay.

The model is essentially an expansion of past inputs and outputs and it is based on the linear ARX model (Billings, 2013), but the nonlinear ARX model uses nonlinear mapping F between the input and output data. In this paper, F is a wavelet network, which is similar to a neural network for the structure and the learning approach, but the former requires a smaller number of iterations for the

Table 2. Parameters of the symmetric and asymmetric friction models.

Model/ Parameter	F_C^a	F_C^s	F_F^a	F_F^s	F_T^a	F_T^s	F_V^a	F_V^s
M (kg)	95.1540	95.1089	95.154	95.109	95.547	95.681	95.153	95.105
$\mathrm{F}_V(N/ms^{-1})$	-	203.5034	-	203.5	-	203.36	-	200.85
$\mathrm{F}_C(N)$	-	20.393	-	20.393	-	17.023	-	20.639
$\mathrm{F}_S(N)$	-	-	-	-	-	17.023	-	18.968
$\mathrm{F}_K(N)$	-	-	-	-	-	-	-	-
Of (N)	-	-3.1648	-	-3.1648	-	-4.1234	-	-3.1458
$F_V^+(\mathrm{N/ms^{-1}})$	166.7061	-	166.71	-	178.48	-	166.1	-
F_V^- $(\mathrm{N/ms^{-1}})$	240.4236	-	240.42	-	218.67	-	240.19	-
F_C^+ (N)	20.1440	-	20.144	-	25.014	-	20.204	-
F_C^- (N)	20.6277	-	20.628	-	28.028	-	20.651	-
F_S^+ (N)	-	-	-	-	25.014	-	18.708	-
F_S^- (N)	-	-	-	-	21.578	-	20.058	-
δ_σ	-	-	-	-	-	-	1.8258	1.2375
v_0 (m/s)	-	-	0.0032239	0.0054125	-	-	-	-
v_S (m/s)	-	-	-	-	0.004156	0.006464	0.00063898	0.0054831

training phase (Postalcioglu and Becerikli, 2005). A family of wavelets can be generated by translations and dilations performed on a single fixed function called mother wavelet. A wavelet $\Phi_j(x)$ is derived from its $\phi(z_{jk})$ mother wavelet (Oussar et al., 1998), given by (14) :

$$\Phi_j(x) = \prod_{k=1}^{N_i} \phi(z_{jk}) \qquad \text{with} \qquad z_{jk} = \frac{x_k - m_{jk}}{d_{jk}} \tag{14}$$

where N_i is the number of inputs; m_j and d_j are the translation and dilation vectors, respectively.

$$y = \Psi(x) = \sum_{j=1}^{N_w} c_j \Phi_j + a_0 + \sum_{j=1}^{N_j} a_k x_k \tag{15}$$

The equation given by (15) represents a network with N_i inputs, a layer of N_w wavelets of dimension N_i, a bias term and a linear output neuron (y).

3.3 Building grey and black-box ensemble (EGBB)

The strategy adopted here is a linear regression to find the relationship between the true output of the EMPS and the two approaches $\widehat{y}_1$ and $\widehat{y}_2$ (predictors) given in subsections *3.1* and *3.2*. A weight is attributed to each different solution so that the approximation given by $\widehat{y}_3$ can be better than $\widehat{y}_1$ and $\widehat{y}_2$. The input arguments for the linear regression are the $\widehat{y}_1$ and $\widehat{y}_2$ as predictors and y as the response variable. A linear, first-order model applied to p explanatory variables $\mathbf{x}_j$ and a dependent variable $\mathbf{Y}$ that are observed on n individuals relies on the supposition that variables are related by the relation given by (16) (Draper and Smith, 1998).

$$\mathbf{Y} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1} \boldsymbol{X} + \boldsymbol{\epsilon} \tag{16}$$

where $\mathbf{Y}$ is the random centered vector; $\mathbf{X}$ is the fixed centered and standardized $(n \times p)$ matrix of predictor variables $\mathbf{X}_j$, $\boldsymbol{\beta_1}$ is the regression vector, $\boldsymbol{\beta_0}$ is a random vector consisting of identical elements and $\boldsymbol{\epsilon}$ is a random vector (error term). One purpose of regression is to give estimates b_j of the regression coefficients β_j.

3.4 Evaluation metric

The quality of the results can be assessed by computing the relative error ϵ:

$$\epsilon = 100 \frac{\|\mathbf{y} - \widehat{\mathbf{y}}\|}{\|\mathbf{y}\|} \tag{17}$$

where $\mathbf{y}$ is the $(N \times 1)$ vector of measurements and $\widehat{\mathbf{y}}$ is the $(N \times 1)$ vector of estimation.

4. RESULTS

In this section the results concerning the application of a grey and black-box modeling as well a ensemble approach for the estimation of the load's position of the EMPS are provided. The data provides estimation and validation data sets. The whole sytem was simulated in the software MATLAB® according to the diagram of Figure 1, *ode45* was used as the solver for the differential equation in order to simulate the load's position with the terms estimated by the grey-box approach. All necessary parameters for the simulation can be found in Table 2. The software MATLAB® was used to define the parameters of the black-box model, which are the orders *na*, *nb* and the number of neurons M of the wavelet network nonlinearity estimator that is used for the mapping of the models nonlinearity. The values of *na* and *nb* ranged from 2 until 15 and the number of neurons from 10 until 30.

Figure 5 shows the comparison between measured and the estimated position of the BGBM approach ($\widehat{\mathbf{y}}_\mathbf{2}$) of $\mathbf{F_V^s}$ and $\mathbf{F_T^s}$ models, which are considered as the best models, since they achieved higher relative error improvement. It can be seen that the estimated position is almost perfectly modeled by both models of the BGBM approach, showing their reasonable resemblance. The comparison was made taking in account the relative error of the friction model proposed in the benchmark (symmetric Coulomb model with viscous friction), and the models proposed here, see Table 6. In the benchmark, the relative errors are 0.013752% and 0.0080248% for the estimation and validation sets respectively.

Tables 3 - 5 show the parameters of all the models and their evaluation metrics. Considering the GBM approach

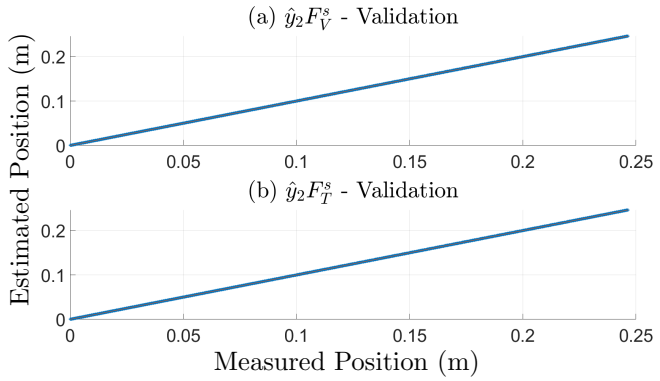

Figure 5. Comparison between the measured and estimated output given by models $\widehat{y}_2 F_V^s$ and $\widehat{y}_2 F_T^s$.

($\widehat{\mathbf{y}}_\mathbf{1}$), from the six proposed models, only $\widehat{y}_1 F_T^s$ is able to bring some error improvement considering estimation and validation sets. Table 4 shows clearly that the BGBM approach is the most suitable for better modeling the load's position of the EMPS. According to the metrics, $\widehat{y}_2 F_V^s$ and $\widehat{y}_2 F_C^s$ have, respectively, the lowest relative errors for estimation (ε_{Est}) with 0.0071464% and 0.0072538%. $\widehat{y}_2 F_T^s$ and $\widehat{y}_2 F_C^a$ with 0.0063456% and 0.0069405%, have respectively, the lowest relative errors for the validation set (ε_{Val}). For the estimation set, all models of $\widehat{\mathbf{y}}_\mathbf{2}$ perform better than the model proposed by the benchmark. We can see the same behavior at the validation set, except by $\widehat{y}_2 F_F^a$ and $\widehat{y}_2 F_F^s$ models. Table 5 shows the results of the EGBB approach ($\widehat{\mathbf{y}}_\mathbf{3}$). In this case, 50% of the models were able to perform better than the original friction model in the validation set. The best models for this approach are $\widehat{y}_3 F_T^s$ and $\widehat{y}_2 F_C^s$.

Table 6 shows the percentage of improvement achieved by each approach and for every model. All relative errors of every model are compared with the relative error achieved by $\widehat{y}_1 F_C^s$, which uses the friction model proposed in the benchmark. The model with the highest relative error improvement for the estimation set was $\widehat{y}_2 F_V^s$ (48.03%), followed by $\widehat{y}_2 F_C^s$ (47.25%) and $\widehat{y}_2 F_T^s$ (46.68%). For the validation set, $\widehat{y}_2 F_T^s$ had the highest error improvement percentage (20.92%), followed by $\widehat{y}_3 F_T^s$ and $\widehat{y}_2 F_C^a$ models with, 18.59% and 13.51% respectively.

The friction models proposed by the GBM approach has the lowest performance among all approaches. It is due to the number of parameters being estimated and also to the initial conditions, which have high influence on the performance of this approach. The difference between the relative errors of both asymmetric and symmetric models are not significant, but one should consider the fact that the asymmetric models have more parameters being estimated compared to the symmetric models. It leads us to do not consider one approach being better than the other one. BGBM approach has shown to be efficient on the error modeling, since only two of sixteen models were not able to perform better than the model proposed in the benchmark. Therefore, the EGBB approach has also proven to be effective, which achieved better performance on 50% of the models, among then $\widehat{y}_3 F_T^s$ achieved almost 20% of relative error improvement. Finally, the models

Table 3. Parameters and metrics of the GBM approach. $\widehat{y}_1 F_T^s$ and $\widehat{y}_1 F_V^s$ show relative error improvement for the estimation set compared to $\widehat{y}_1 F_C^s$, which is the friction model proposed by the benchmark. Except by $\widehat{y}_1 F_F^a$, $\widehat{y}_1 F_F^s$ and $\widehat{y}_1 F_V^s$ all other models show relative error improvement.

Model	Parameters	ε_{Est}(%)	ε_{Val}(%)
$\widehat{y}_1 F_C^a$	M=95.15540; F_V^+=166.7061; F_V^-=240.4236; F_C^+=20.1440; F_C^-=20.6277	0.0137580	0.0079672
$\widehat{y}_1 F_C^s$	M=95.1089; F_V=203.5034; F_C=20.3935; Of=-3.1648	0.0137522	0.0080248
$\widehat{y}_1 F_F^a$	M=95.154; F_V^+=166.71; F_V^-=240.42; F_C^+=20.144; F_C^-=20.628; v_o=0.0032239	0.0139793	0.0120720
$\widehat{y}_1 F_F^s$	M=95.109; F_V=203.5; F_C=20.393; Of=-3.1648 v_o=0.0054125	0.0139503	0.0121076
$\widehat{y}_1 F_T^a$	M=95.547; F_V^+=178.48; F_V^-=218.67; F_C^+=25.014; F_C^-=28.028; F_S^+=25.014; F_S^-=21.578; v_S=0.004156;	0.0149361	0.0079732
$\widehat{y}_1 F_T^s$	M=95.681; F_V=203.36; F_S=17.023; F_C=17.023; v_S=0.006464;Of=-4.1234	0.0125696	0.0079662
$\widehat{y}_1 F_V^a$	M=95.153; F_V^+=166.1; F_V^-=240.19; F_C^+=20.204; F_C^-=20.651; F_S^+=18.708; F_S^-=20.058; v_S=0.00063898 δ_σ=1.8258	0.0137678	0.0079584
$\widehat{y}_1 F_V^s$	M=95.105; F_V=200.85; F_C=20.639; F_S=18.968; Of=-3.1458; δ_σ=1.2375 v_S=0.0054831	0.0137384	0.0080252

Table 4. Parameters and metrics of the BGBM approach. It shows that $\widehat{y}_2 F_V^s$ and $\widehat{y}_2 F_T^s$ have the lowest relative errors for the estimation and validation sets respectively.

Model	na	nb	M	ε_{Est}(%)	ε_{Val}(%)
$\widehat{y}_2\ F_C^a$	5	13	10	0.0073630	0.0069405
$\widehat{y}_2\ F_C^s$	4	13	12	0.0072538	0.0070701
$\widehat{y}_2\ F_F^a$	3	13	14	0.0077180	0.0119043
$\widehat{y}_2\ F_F^s$	3	4	29	0.0081926	0.0108225
$\widehat{y}_2\ F_T^a$	6	3	10	0.0118427	0.0072983
$\widehat{y}_2\ F_T^s$	4	13	12	0.0073332	0.0063456
$\widehat{y}_2\ F_V^a$	4	6	15	0.0091737	0.0077414
$\widehat{y}_2\ F_V^s$	4	13	15	0.0071464	0.0073584

that smooth the discontinuity at zero velocity of Coulomb friction model have better performance.

5. CONCLUSION AND FUTURE WORKS

In this paper grey and black-box models and their ensembles have been employed in order to estimate the position of a load of an EMPS. According to the results, a significant relative error improvement can be achieved by modeling the estimation error using different friction models in a hybrid approach. An improvement higher than 20% was achieved combining the grey and black-box models. The GBM approach has the lowest performance among all approaches, it is due to the influence of the number of parameters being estimated and the initial conditions.

Table 5. Parameters and metrics of the third approach. Model $\widehat{y}_3 F_T^s$ and $\widehat{y}_2 F_C^s$ have lowest relative errors compared to the model proposed by the benchmark.

Model	c	$w1$	$w2$	$\varepsilon_{Val}(\%)$
$\widehat{y}_3\ F_C^a$	3.986E-06	0.0279	0.9721	0.0072700
$\widehat{y}_3\ F_C^s$	3.248E-07	0.0716	0.9284	0.0071669
$\widehat{y}_3\ F_F^a$	5.223E-06	0.1173	0.8826	0.0121304
$\widehat{y}_3\ F_F^s$	7.612E-06	-0.4233	1.4232	0.0118544
$\widehat{y}_3\ F_T^a$	1.866E-05	-0.0438	1.0438	0.0118841
$\widehat{y}_3\ F_T^s$	8.767E-08	0.1184	0.8816	0.0065332
$\widehat{y}_3\ F_V^a$	3.695E-06	-0.9112	1.9112	0.0085389
$\widehat{y}_3\ F_V^s$	8.294E-07	0.0683	0.9317	0.0073771

Table 6. General comparison between all approaches. It shows that $\widehat{y}_2 F_C^s$ and $\widehat{y}_2 F_T^s$ have the best performance for estimation and validation sets respectively.

Model	$\Delta_{Est}(\%)$	$\Delta_{Val}(\%)$	Model	$\Delta_{Est}(\%)$	$\Delta_{Val}(\%)$
$\widehat{y}_1\ F_C^a$	-0.04	0.72	$\widehat{y}_2 F_T^a$	13.89	9.05
$\widehat{y}_2\ F_C^a$	46.46	13.51	$\widehat{y}_3 F_T^a$	-	-22.76
$\widehat{y}_3\ F_C^a$	-	9.41	$\widehat{y}_1 F_T^s$	8.60	0.73
$\widehat{y}_2\ F_C^s$	47.25	11.90	$\widehat{y}_2 F_T^s$	46.68	20.92
$\widehat{y}_3\ F_C^s$	-	10.69	$\widehat{y}_3 F_T^s$	-	18.59
$\widehat{y}_1\ F_F^a$	-1.65	-50.43	$\widehat{y}_1 F_V^a$	-0.11	0.83
$\widehat{y}_2\ F_F^a$	43.88	-48.34	$\widehat{y}_2 F_V^a$	33.29	3.53
$\widehat{y}_3\ F_F^a$	-	-51.16	$\widehat{y}_3 F_V^a$	-	-6.4
$\widehat{y}_1\ F_F^s$	-1.44	-50.88	$\widehat{y}_1 F_V^s$	0.10	0.00
$\widehat{y}_2\ F_F^s$	40.43	-34.86	$\widehat{y}_2 F_V^s$	48.03	8.30
$\widehat{y}_3\ F_F^s$	-	-47.72	$\widehat{y}_3 F_V^s$	-	8.07
$\widehat{y}_1\ F_T^a$	-8.61	0.64			

Because of this and since there is no significant difference between the relative errors of the friction models of the GBM approach, it is not possible to conclude whether asymmetric or symmetric friction models are better to predict the position of the load.

The authors are also interested in using global search algorithms in order to estimated the dynamic parameters of the EMPS and simulate its velocity, acceleration and force, since this approach does not requires initial conditions to start the optimization process.

REFERENCES

(2001). *Adaptive Control of Nonsmooth Dynamic Systems*. Springer-Verlag, London, UK, 1 edition.

Billings, S.A. (2013). *Nonlinear system identification*. John Wiley Sons, Chichester, UK.

Bo, L. and Pavelescu, D. (1982). The friction-speed relation and its influence on the critical velocity of stick-slip motion. *Wear*, 82, 277–289.

Bohlin, T.P. (1994). A case study of grey box identification. *Automatica*, 30:2, 307–318.

Bohlin, T.P. (2006). *Practical Grey-box Process Identification - Theory and Applications*. Springer-Verlag, London, UK.

Brunot, M., Janot, A., and Carrillo, F. (2018). An automated instrumental variable method for rigid industrial robot identification. *IFAC-PapersOnLine*, 51:15, 431–436.

Coulomb, C.A. (1785). Théorie des machines simples, en ayant égard au frottement de leurs parties, et à la roideur des cordages. Paris, France.

de Wit, C.C., Olsson, H., Astrom, K., and Lischinsky, P. (1995). A new model for control of systems with friction. *IEEE Transactions on Automatic Control*, 40, 419–425.

Draper, N.R. and Smith, H. (1998). *Applied Regression Analysis*. John Wiley Sons, New York, UK, 3 edition.

Haykin, S. (2009). *Neural networks and learning machines*. Prentice Hall, Upper Saddle River, NJ,USA, 3 edition.

Huang, S.J. and Chiu, C.M. (2009). Optimal LuGre friction model identification based on genetic algorithm and sliding mode control of a piezoelectric-actuating table. *Transactions of the Institute of Measurement and Control*, 31, 181–203.

Irigoyen, E. and Miñano, G. (2013). A NARX neural network model for enhancing cardiovascular rehabilitation therapies. *Neurocomputing*, 109, 9–15.

Janot, A., Gautier, M., and Brunot, M. (2019). Data set and reference models of EMPS. *Workshop on Nonlinear System Identification Benchmarks,* Eindhoven, The Netherlands.

Janot, A., Young, P.C., and Gautier, M. (2017). Identification and control of electro-mechanical systems using state-dependent parameter estimation. *International Journal of Control*, 90, 643–660.

Kerschen, G., Worden, K., Vakakis, A.F., and Golinval, J.C. (2006). Past, present and future of nonlinear system identification in structural dynamics. *Mechanical Systems and Signal Processing*, 20, 505–592.

Khalil, W. and Dombre, E. (2004). *Modeling, Identification and Control of Robots*. Butterworth-Heinemann, London, UK, 1 edition.

Ljung, L. (1999). *System Identification - Theory for the user*. Prentice Hall, Upper Saddle River, NJ,USA, 2 edition.

Ljung, L. (2010). Perspectives on system identification. *Annual Reviews in Control*, 34, 1–12.

Marques, F., Flores, P., Claro, J.C.P., and Lankarani, H.M. (2016). A survey and comparison of several friction force models for dynamic analysis of multibody mechanical systems. *Nonlinear Dynamics*, 86:3, 1407–1443.

M.Saviozzi, S.Massucco, and F.Silvestro (2017). Implementation of advanced functionalities for distribution management systems: Load forecasting and modeling through artificial neural networks ensembles. *Journal of Electric Power Systems Research*, 167, 230–239.

Nogoseke, L.F., Andrade, G.H.B., and Boaretto, M.A.R. (2017). Investigating the use of extremely randomized trees, gradient boosting machine, k-nearest neighbors and their ensemble applied to fault detection. *Proceedings of the 24th ABCM International Congress of Mechanical Engineering,* Curitiba, Brazil.

Oussar, Y., I.Rivals, L.Personnaz, and G.Dreyfus (1998). Training wavelet networks for nonlinear dynamic input–output modeling. *Neurocomputing*, 20, 173–188.

Papadopoulos, S. and Karakatsanis, I. (2015). Short-term electricity load forecasting using time series and ensemble learning methods. *Proceedings of IEEE Power and Energy Conference,* Champaign, IL, USA, 1–6.

Pinto, W.C.F. and Ayala, H.V.H. (2019). Comparison of friction models for grey-box identification of an electromechanical positioning system. *Proceedings of the 24th Brazilian Symposium on Intelligent Automation - SBAI,* Ouro Preto, MG, Brazil, 2019.

Polikar, R. (2012). *Ensemble Machine Learning.* Springer, NY, USA.

Postalcioglu, S. and Becerikli, Y. (2005). Nonlinear system modeling using wavelet networks. *Proceedings of the International Symposium on Neural Networks,* Chongqing, China.

Qiua, X., Ren, Y., Suganthan, P.N., and Amaratunga, G.A.J. (2017). Empirical mode decomposition based ensemble deep learning for load demand time series forecasting. *Journal of Applied Soft Computing*, 54, 246–255.

Raza, M.Q., N.Mithulananthan, and Summerfield, A. (2018). Solar output power forecast using an ensemble framework with neural predictors and bayesian adaptive combination. *Journal of Solar Energy*, 166, 226–241.

Ribeiro, M.H.D.M. and dos Santos Coelho, L. (2019). Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. *Applied Soft Computing Journal-Article in Press.*

Samara, P.A., Sakellariou, J.S., Fouskitakis, G.N., Hios, J.D., and Fassois, S.D. (2013). Aircraft virtual sensor design via a time-dependent functional pooling NARX methodology. *Aerospace Science and Technology*, 29, 114–124.

Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P.Y., Hjalmarssont, H., and Juditsky, A. (1995). Nonlinear black-box modeling in system identification. *Automatica*, 31, 1691–1724.

Sperati, S., Alessandrini, S., and Monache, L.D. (2016). An application of the ECMWF ensemble prediction system for short-term solar power forecasting. *Journal of Solar Energy*, 133, 437–450.

Tangirala, A.K. (2015). *Pinciples of system identification.* CRC Press, Boca Raton, FL, USA.

Threlfall, D. (1978). The inclusion of coulomb friction in mechanisms programs with particular reference to dram au programme dram. *Mechanism and Machine Theory*, 13(4), 475 – 483.

Tijani, I.B., Akmeliawati, R., Legowo, A., and Budiyono, A. (2014). Nonlinear identification of a small scale unmanned helicopter using optimized NARX network with multiobjective differential evolution. *Engineering Applications of Artificial Intelligence*, 33, 99–115.

Tustin, A. (1947). The effects of backlash and of speed-dependent friction on the stability of closed-cycle control systems. *Journal of the Institution of Electrical Engineers - Part IIA: Automatic Regulators and Servo Mechanisms*, 94(1), 143–151.

Wang, G.L., Li, Y.F., and Bi, D.X. (2004). Support vector machine networks for friction modeling. *IEEE/ASME Transactions on Mechatronics*, 9(3), 601–606.

Wong, C. and Worden, K. (2007). Generalised NARX shunting neural network modelling of friction. *Mechanical Systems and Signal Processing*, 21, 553–572.

Worden, K., Wong, C.X., Parlitz, U., Hornstein, A., Engster, D., Tjahjowidodo, T., Al-Bender, F., D.Rizos, D., and Fassois, S.D. (2007). Identification of pre-sliding and sliding friction dynamics: Grey box and black-box models. *Mechanical Systems and Signal Processing*, 21:1, 514–534.

Wunsch, A., Liesch, T., and Broda, S. (2018). Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX). *Journal of Hydrology*, 567, 743–758.

Yoon, J.Y. and Trumper, D.L. (2014). Friction modeling, identification, and compensation based on friction hysteresis and dahl resonance. *Mechatronics*, 24, 734–741.

Zhang, G.P. (2003). Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50, 159–175.