CHEMICAL ENGINEERING TRANSACTIONS Volume 25, 2011
Editors Jiří Jaromír Klemeš, Petar Sabev Varbanov, Hon Loong Lam
Copyright © 2011, AIDIC Servizi S.r.l., ISBN 978-88-95608-16-7 ISSN 1974-9791
DOI: 10.3303/CET1125127

Neural Networks and Support Vector Machine Models Applied to Energy Consumption Optimization in Semiautogeneous Grinding

Millaray Curilem[1], Gonzalo Acuña[2,*], Francisco Cubillos[3], Eduardo Vyhmeister[3]

[1]Departamento de Ingeniería Eléctrica, Universidad de la Frontera, [2]Departamento de Ingeniería Informática, Universidad de Santiago de Chile, [3]Departamento de Ingeniería Química, Universidad de Santiago de Chile
Ecuador 3659, Santiago 9170022, Chile, gonzalo.acuna@usach.cl

Semiautogenous (SAG) mills for ore grinding are large energy consumption equipments. The SAG energy consumption is strongly related to the fill level of the mill. Hence, on-line information of the mill fill level is a relevant state variable to monitor and drive in SAG operations. Unfortunately, due to the prevailing conditions in a SAG mill, it is difficult to measure and represent from first principle model the state of the mill fill level.
Alternative approaches to tackle this problem consist in designing appropriate data-driven models, such as Neural Networks (NN) and Support Vector Machine (SVM). In this paper, NN and a SVM (specifically a Least Square-SVM) are used as Nonlinear autoregressive with exogenous inputs (NARX) and Nonlinear autoregressive moving average with exogenous inputs (NARMAX) models for on-line estimation of the filling level of a SAG mill. Good performances of the developed models could allow implementation in SAG operation/control hence optimizing its energy consumption.

1. Introduction

Mineral processing and metal production is under increased pressure to reduce its energy consumption and greenhouse gases emition. Studies performed in different metal processes have shown that the crushing and grinding processes make the larger contribution to these negatives effects (Norgate and Haque, 2010; Wei and Craig, 2009). Semiautogeneous grinding (SAG) mill, due to different processing and designing variables, is one of the principal alternatives used for mineral size reduction (Salazar et al, 2009).
Given the previous statement, modeling and control of the grinding process is of considerable interest for the industry. Regretfully, the representation of interaction between grinding media and feedstock is complex, becoming very difficult to build adequate first principle models in order to perform forecasting, optimization or control.
To elaborate good dynamic models it is usually necessary a deep knowledge of the process. The representations obtained from first principle models are often too complex. Additionally, in the SAG mill process different variables, needed by the models, can be

Please cite this article as: Curilem M., Acuna G., Cubillos F. and Vyhmeister E., 2011, Neural networks and support vector machine models applied to energy consumption optimization in semiautogeneous grinding, Chemical Engineering Transactions, 25, 761-766
DOI: 10.3303/CET1125127

difficult to measure on-line therefore, assumptions or indirect variables are used to represent and control the grinding process, reducing model performance.

An alternative and fruitful approach to tackle these problems consists in designing appropriate data-driven models. In this sense, neural networks (NN) and support vector machines (SVM) have been proven in the last decade to be powerful tools for system modeling. Despite the successful results achieved with neural networks, there still remain unsolved a number of key issues such as: difficulty of choosing the number of hidden nodes, the overfitting problem, the existence of local minima solution, and poor generalization capabilities, among others. Alternatively, SVM has many advantages such as good generalization performance, fewer free parameters to be adjusted, and a convex optimization problem to be solved (Schölkopf, 2000).

SVM has been developed mainly for solving classification and static function approximation problems. Indeed, in the case of dynamic systems almost all the work that has been done concerning Support Vector Regression (SVR) is focused in series-parallel identification methods for NARX (Nonlinear autoregressive with exogenous inputs) modeling (Suykens, 2001). A more flexible tool for performing this support vector machine regression task is the least square support vector machine (LS-SVM) proposed by Suykens (2002). Indeed, neither SVR nor LS-SVM can easily be trained when coping with parallel identification methods for NARMAX type models. In this case NN still remain as the preferred tool.

In a recent work (Acuña and Curilem, 2009), a comparison between NARX SVR and NN models for the SAG process has been performed. In this paper, the performance of three dynamic models, built up with the use of a NARMAX-type NN, NARX-type NN and NARX-type LS-SVM are compared when acting as estimators of the SAG filling level, one of the most important state variables for SAG grinding operation.

2. Description of the Semiautogeneous Grinding Process

In SAG mills the grinding occurs by the falling action of the ore from a height close to the diameter of the mill. The SAG involves the addition of metallic grinding media to the mill, whose volumetric filling level varies from 4% to 15% of the mill's volume.

The aim during SAG operation is to work under conditions that imply the maximum installed power consumption; this means working under unstable operating conditions.

SAG operational control has shown to be complex. The experience gathered so far has allowed identifying the variables that affect, at a considerable level, the sag mill performance. Some of these variables are the primary crushers operational results (e.g. fragmentation degree, tonnage and particle size distribution), the feed water flow rate (which affects mill pulp viscosity, density, and transport), the design variables (e.g. lifter design, size), the grinding media, the processing speed, and total internal load. Among the most commonly controlled variables it is worth mentioning the product particle size, slurry level in the sump, discharge density, mill load, and mill power. The mill load and sump level are controlled to ensure stability whereas the particle size is controlled to maintain small variations in the set-point (Wei and Craig, 2009).

The mill load estimation is of considerable interest as controlled parameter since, in addition to the stability benefits obtained in its control, the mill load is also related to

mill energy consumption and energy efficiencies. A low filling level is related to unproductive grinding media interactions and harmful impacts in the mill shell. A high filling level beyond the point of maximum consumption leads to an overfilling condition and poor performance (Lameck et al., 2006).
Considering the previous information the SAG circuit operators must try to conjugate these factors in order to achieve first the stabilization of the operation, and then try to improve it (Magne et al., 1997). That is why it is important for them to have reliable and real-time information on the mill filling level.

3. Least Square Support Vector Machine

Support vector machines tackle classification and regression problems by nonlinearly mapping input data into high-dimensional feature spaces, wherein a linear decision surface is designed. SVR algorithms are based on the results of the statistical theory of learning given by Vapnik (1995), which introduces regression as the fitting of a tube of radius ν to the data. The decision boundary for determining the radius of the tube is given by a small subset of training examples called Support Vectors (SV). Vapnik's SVM Regression estimates the values of the support vector, w, to obtain the function of equation 1:

$$f(x)=(w\cdot x)+b \;\; with \; w,x\in R^N, b\in R \tag{1}$$

by introducing the so called ε-insensitive loss function shown in equation 2:

$$|y-f(x)|_\varepsilon = \max\{0,|y-f(x)-\varepsilon|\} \tag{2}$$

which does not penalize errors smaller than ε>0 (where ε corresponds to a value chosen a priori). The algorithm used for implementation and solution of the nonlinear regression problem in a space of higher dimensions are extensively covered somewhere else (Vapnik, 1995; Suykens 2001, 2002) and will not be covered in the present work. In the case of LS-SVM the problem is transformed in order to solve a linear Karush-Kuhn-Tucker system derived from the dual formulation of the non-linear regression problem (Suykens, 2002). The problem is simpler because there are no inequalities and a global and unique solution can be found when the matrix is full rank. The major drawback in this case is the lack of sparseness that results from the fact that all data points are taken as support vectors.

4. Identification of NARX LS-SVM and NN models

NARX model is the non-linear extension of the lineal ARX model given by equation 5:

$$y(k) = f(y(k-1),..,y(k-n),..u(k-1),...,u(k-m),..) + w(k) \tag{5}$$

with {w(k)} an independent, zero-mean, random variable sequence that represents the random error and uncertainty of the model. The associated predictor is given by equation 6:

$$\hat{y}(k) = \psi(y(k-1),...,y(k-n),.....,u(k-1),...,u(k-m)...) \quad (6)$$

To identify this predictor a series-parallel method for identifying the parameters of a LS-SVM or a NN can be used. To implement the NN-NARMAX models the ideas that are introduced in the work of Nerrand et al. (1993) were used (i.e. static neural network with external recurrences is equivalent to the neural networks with internal recurrences). Parallel identification methods using neural networks can be performed using a gradient descent algorithm. Better results can be obtained with a modified version using a 2nd order optimization method like Levenberg-Marquardt. The identification methods for neural networks correspond to the back-propagation-through-time algorithm proposed by Werbos (1990).

5. Results

5.1 Data selection, identification, and prediction

A set of 1000 examples was used for all the SVR and NN models training. Each example has the filling level and the bearing pressure at time t as inputs and the filling level at time t+1 as the output. Depending on the series-parallel or parallel identification structure used the identified model acts as an associated predictor of a NARX type model (case of LS-SVM) or as a NARMAX type model (case of NN). Once identified, each model (SVR and NN) was used for forecasting the SAG mill level (1000 new data-points) in a one-step-ahead (OSA prediction) and multiple-steps-ahead (MPO prediction) manner. The MPO predictions are much difficult than OSA predictions therefore the model that performed better as an MPO predictor was consider to be better (Leontaritis and Billings, 1985). The estimation error is quantified using the Matlab 7.1 MSE index.

5.2 LS-SVM training

The LS-SVM Matlab Toolbox developed by Suykens (2010) was used in this work. The LS-SVM model was implemented with a "RBF" kernel. Parameter C was increased by powers of two. The powers took values from -5 to 15. A 0.01 step was used. The best results were obtained for $C=2^{-1.92}$, what means that the solution requires a low complexity for the model to achieve a good generalization for MPO prediction. The parameter of the kernel function, called here sigma, also varied by powers of two, with the powers ranging from -5 to 10 with a 0.1 step. The best result was obtained for $sigma= 2^{-3.44}$.

5.3 NN training

The NN based System Identification Matlab Toolbox developed by M. Norgaard (2003) was used for training purposes with the Levenberg Marquardt optimization algorithm. One hidden layer architecture with two neurons in the input layer and one neuron in the output layer was chosen. Several output, input and error delays were tested. Finally, 1 output delay, 1 input delay, 2 error delays and 1 time delay were used. Three hidden neurons were used regarding a criterion of equivalent complexity with the 6 hidden neurons of the NN-NARX model (similar number of weights). NN-NARX

models were identified with a series-parallel method and used as MPO predictors just in the way as LS-SVM was used.

5.4 SVR and NN results

Figure 1 shows the results obtained when using LS-SVM and NN models as MPO NARX or NARMAX predictors. The mean square error (MSE) obtained for the MPO predictions were 3.5889, 1.0773, and 1.9201 for the NN-NARX, NN-NARMAX, and LS-SVM-NARX models, respectively. From these results it is seen that the NARMAX dynamic model performs better than both NARX models when acting as an MPO predictor, which is a rather difficult test to any dynamic model. It implies predicting the output variable (filling level) only from its initial value together with experimental input values of the other variable (bearing pressure).

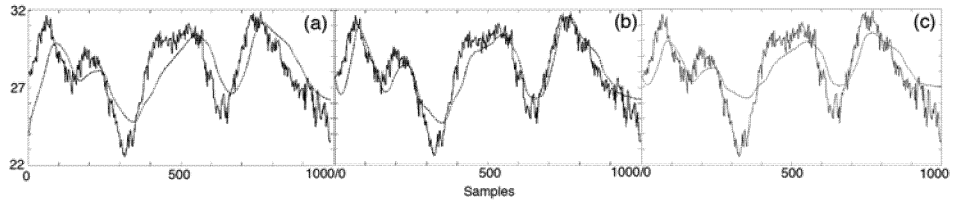

Figure 1. Estimations (dashed lines) and real data (continuous lines) for SAG filling level estimation (NN NARX (a); NN NARMAX (b); LS-SVM NARX (c)).

These results are coherent with the fact that NARMAX type models, even if their identification procedure is more complex, take into account more information about the process included in the delayed input errors. Concerning NARX models, LS-SVM performs better than NN thus confirming its good capacities including better generalization.

6. Conclusions

This paper presents LS-SVM and NN dynamic models acting as state estimators of the filling level variable of a semiautogenous grinding process.

For modeling the process LS-SVM NARX and NN acting as NARX or NARMAX associated predictors were used (i.e. identified with a series-parallel or parallel method). All models were used as MPO predictors. In this case the results show that the performance of the LS-SVM is better than the performance of the NN when acting as NARX models. However, the NN NARMAX dynamic model clearly outperforms both NARX models. This leads to the need to test the performance of SVR in a NARMAX configuration, which is the focus of the ongoing research.

Black box dynamic models for complex industrial processes like the SAG process shown in this work could be of great relevance to design good predictive control and estimate important process variables. Consequently, processes with better performance, lower undesirable operational conditions, and higher energy efficiencies can be obtained by using knowledge based estimation algorithms that have proven to be of great relevance for improving complex plant operation.

7. Acknowledgements

The authors would like to thank the financial aid from Fondecyt Project 1090316, the Research Direction of the La Frontera University and the Scientific and Technological Research Department of the Universidad de Santiago de Chile (DICYT).

References

Acuña, G. and Curilem, M., 2009, Comparison of neural networks and support vector machine dynamic models for state estimation in semiautogeneous mills, Lecture Notes in Computer Science, LNCS, (ISI-Proceeding). Lecture Notes in Computer Science, 5845, 478-487.

Lameck, N.S., Kiangi, K.K. and Moys, M.H., 2006, Effect of grinding media shapes on load behavior and mill power in a dry ball mill, Minerals Engineering, 19, 1357-1361.

Leontaritis, I.J. and Billings, S.A., 1985, Input-output parametric models for non-linear systems; Part 1: deterministic non-linear systems; Part 2: stochastic non-linear systems, International Journal of Control, 45, 1985, 303-344.

Magne, L., Valderrama, W. and Pontt, J., 1997, Conceptual vision and state of the semiautogenous mill technology, Revista Minerales, Instituto de Ingenieros de Minas de Chile. 52(218).

Nerrand, O., Roussel, P., Personnaz, L. and Dreyfus, G., 1993, Neural networks and non-linear adaptive filtering: unifying concepts and new algorithms, Neural Computation, 5, 165-199.

Norgaard, M., 2003, www.mathworks.com/matlabcentral/fileexchange/authors/1119, 02-06-2011.

Norgate, T. and Haque, N., 2010, Energy and greenhouse gas impact of mining and mineral processing operations, Journal of Cleaner Production, 18, 266-274.

Salazar, J.L., Magne, L., Acuña, G. and Cubillos, F., 2009, Dynamic modeling and simulation of semi-autogenous mills, Minerals Engineering, 22, 70-77.

Schölkopf, B., Smola, A., Williamson, R.C. and Bartlett P.L., 2000, New support vector algorithms, Neural Computation, 12, 1083-121.

Suykens, J., 2010, www.esat.kuleuven.be/sista/lssvmlab, 02-06-2011.

Suykens, J., 2001, Nonlinear modeling and support vector machines, Proceedings of the 18thIEEE. Instrum. and Measurement Technol. Conf., Budapest, Hungary, 287-294.

Suykens, J., Van Gestel, T., De Brabanter, J., De Moor, B. and Vandewalle, J., 2002, Least Squares Support Vector Machines, World Scientific, Singapore.

Vapnik, V., 1995, The nature of statistical learning theory, Springer, Verlag, New York.

Wei, D. and Craig, I.K., 2009, Grinding mill circuits – A survey of control and economic concerns, International Journal of Mineral Processing, 90, 56-66.

Werbos, P.J., 1990, Backpropagation through time: what it does and how to do it, Proceedings of the IEEE, Vol 78(10), 1150-1560.