Wind Power Generator Model Based on LS-SVM for Unbalanced Three-Phase Distribution System Power Flow Studies

Jarev Divinagracia, Russel John Gallano
Electrical and Electronics Engineering Institute
University of the Philippines Diliman
Quezon City, Philippines
jarevddivinagracia@gmail.com, russel.gallano@eee.upd.edu.ph

***Abstract*— Wind power generation is increasing around the world. However, increasing wind power penetration has a significant impact on the power system, in particular on distribution systems. Accurate wind generator models are needed for proper unbalanced load flow analysis. A non-linear equation-based model is the most accurate as it models, though requiring greater computational power and time. Machine learning-based models, such as Artificial Neural Networks, are accurate enough and are faster than non-linear models. This paper explores another machine learning algorithm–Least Squares Support Vector Machines (LS-SVM)–in wind generator modeling. It was shown that the LS-SVM model is more accurate than the artificial neural network (ANN) model and almost as accurate as the non-linear model. When the LS-SVM model was integrated to the load flow analysis of the IEEE 37-bus system, the solution had little to no deviation from the results using the non-linear model of wind generators.**

Keywords—wind power; generator model; support vector machines; unbalanced power flow;

I. Introduction

The rapid increase of wind power generation today has brought about new possibilities and problems on the electrical system. The growth of wind power comes in two fronts: large-scale wind farms connected to transmission systems, distributed wind generation connected to low voltage distribution systems. Increasing wind generation connected to distribution systems introduces many operational and control problems because of several characteristics of the system, like its radial structure and unbalanced loads and lines. Solving such problems would need the development of a robust and reliable computational wind generator model for use in power system studies.

Wind generator models used in power flows can be categorized into these three categories: constant power model, nonlinear model, and machine learning models. The paper by Feijoo et al presents the PQ model. The advantage of the PQ model is that the real power is calculated as a function of the wind speed for the first iteration of the power flow analysis, and from then on its value remains constant [1], [2]. These models are easy to integrate into power flow algorithms, making it the most popular wind modeling method out there today. However, for higher wind penetration levels this model may not give accurate power flow solutions due to over-simplification in assuming a constant output power [3].

Advanced nonlinear models, on the other hand, are built by cascading the constituent models of the turbine, generator and the converter. In these models, the power output of a wind generator depends on seven parameters: wind speed (U), three magnitudes of terminal Voltage (V_{ABC}) and three phase angles (δ_{ABC}). In solving these models we need to use an iterative process like Newton-Raphson. Having an iterative process to solve the models and another iterative process for the load flow analysis significantly slows the convergence time of the power flow solution [4].

The last category of wind generator model is designed using Artificial Neural Network (ANN) [5]. In a study by Opathella et al., ANN was used in order to construct a wind farm model that can predict the power output with wind speed and voltages as its inputs. The ANN model was shown to be faster than nonlinear three-phase models [6]. The disadvantage of ANN, however, is that different configurations need to be created manually and compared to find the one that yields the lowest error without sacrificing training time.

This study employs a more superior machine-learning algorithm, which is the Least Squares Support Vector Machine or LS-SVM. It is unlike neural networks, which increase in structure as the system increases. SVMs are also less prone to over-fitting, which is one of the drawbacks of using ANN. ANN models, in general, use empirical risk minimization, while SVM use structural risk minimization [7], [8]. For load flow studies and control in distribution systems, the algorithm to be used would need to have fast computation times and high accuracy. The reasons for these criteria are: (1) distribution systems (DS) tend to be large, greatly slowing down and impacting the number of load flow simulations and analysis that can be done; (2) DS have highly unbalanced loads and lines. Therefore, accuracy of the model is important to prevent unnecessary and erroneous power flows.

II. WIND GENERATOR MODELING USING LS-SVM

Due to its simple and robust nature, the LS-SVM will be used to model wind generation. The model is then integrated into the three-phase Backward-Forward (BF) sweep method to solve for the power flow parameters of an unbalanced three-phase distribution system having wind-power based distributed generators. The accuracy of the model is then compared to the results of a non-linear wind generator model through the root-mean-square error (RMSE) computation. The effects of varying the number of training data points on the variation of training times and RMSE were also explored to find a balance between optimal training time and accuracy.

A. LS-SVM Approach

Suykens proposed a modified version of the Support Vector Machine (SVM) called Least-Square SVM (LS-SVM). This modified version has faster learning speeds, higher accuracy and it is more suitable for nonlinear system identification [7].

A regression problem involving a nonlinear equation is made easier when mapped in a higher dimensional space. The kernel trick is used to construct the optimal hyper plane in a new space, called the feature space, without explicitly mapping the input points in the feature space itself [8]. In this study, the kernel function to be used shall be the Radial Basis Function (RBF) Gaussian Kernel, given in (1):

$$RBF: K(x, x_i) = exp\left(\frac{|x - x_i|^2}{2\sigma^2}\right) \quad (1)$$

where σ is the regularization parameter and is a positive real number [9].

The model is tuned using two methods: Coupled Simulated Annealing (CSA) and simplex method. CSA is characterized by its better optimization efficiency, since the algorithm has a reduced sensitivity to initialization parameters. It couples similar simulated annealing processes to steer the optimization process to the global optimum. Simplex method is a multidimensional unconstrained non-linear optimization. It uses pattern searches that compare the objective function values at several points in a small portion of the search space to find the minimum point or solution.

B. LS-SVM Parameters for the DFIG Model

In line with the main reference paper of Opathella et al., a data set is to be constructed. The data are represented as matrices where each row contains one data point [4].

For wind speeds, a series of 100 to 10000 random numbers (between 4m/s and 20m/s) are considered. For the voltage magnitudes, three series of 100 to 10000 random numbers between 0.94 and 1.06 were assumed. For the three voltage angles, three series of 100 to 10000 random numbers are generated between -5º and 5º. The output series is then computed given the required input series. The study would only cover the modeling of Type-3 DFIG WGs. Solving the model once for every data set will yield corresponding output data: PA, PB, PC. The input and output samples of data are normalized before use. The goal of normalization is the enhancement of the training speed.

For the tuning phase, CSA and simplex method is used. CSA determines suitable parameters given a certain criterion, while simplex performs a fine-tuning step on these parameters. After training the program with the sample input-output data set the next step is evaluating new data points for the model. While inserting new input data sets and acquiring the respective output data sets on the LS-SVM model, the same input data sets will also be fed to the nonlinear model made by A. Dadhania for validation and error analysis of the results generated by the LS-SVM model [10].

Root Means Square Error (RMSE) is be used to measure the difference between the results of the LS-SVM and nonlinear models. It is a widely used method in measuring the difference between values predicted by a model or an estimator and the values actually observed. It is an absolute measure of fit. Lower values of RMSE indicate better fit. The RMSE would be computed for a set of 100 test data points for wind speed, per phase voltage magnitude and angle. The formula is given in (2):

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(y_{SVM} - y_{NL})^2}{n}} \quad (2)$$

where y_{SVM} = SVM output, y_{NL} = non-linear model output, and n = total number of predictions.

III. ANALYSIS OF DFIG MODELING RESULTS

In the succeeding section, training data samples are the samples used in training the model and the test data samples are those used to test the accuracy of the model after having been trained.

Different models with different training data samples, from 100 to 10000, were created. In the validation stage, the results of the LS-SVM model will be compared to the results of the nonlinear model by simulating the nonlinear model and the LS-SVM model once, with the same input parameters, and by computing the RMSE between the two models.

A. One-Time Model Simulation

For the one-time simulation, the input data set inserted into the two models is presented in Table I below, with the results presented in Table II. Note that the LS-SVM model used contained 10000 training data samples. In Table II, the power outputs of the two models are identical. Showing that the LS-SVM has accurately modeled 3-phase wind generators with minimum error from the nonlinear models.

TABLE I. INPUT DATA SET FOR BOTH DFIG MODELS

Input Parameter	Value
Wind speed (m/s)	14
$V_A = V_B = V_C$	1.0 p.u.
$\angle V_A$	1.036º
$\angle V_B$	–118.964º
$\angle V_C$	121.036º

TABLE II. COMPARISON OF RESULTS FOR BOTH DFIG MODELS

Parameter	Nonlinear Model	LS-SVM Model
P_A	-445.72 kW	-445.72 kW
P_B	-445.96 kW	-445.96 kW
P_C	-445.92kW	-445.92kW

B. *Effect of Data Set on RMSE and Training Time*

RMSE is computed to determine the degree of error between the two models. Also, the training time for each data points will be compared. The graph presented below in Fig. 1 contains the said information and the breakeven point between the minimum training time and minimum RMSE is determined. There is a significant trade-off between the RMSE and the training time of the model. As the data points increase, the model becomes significantly longer to train but the degree of error decreases.

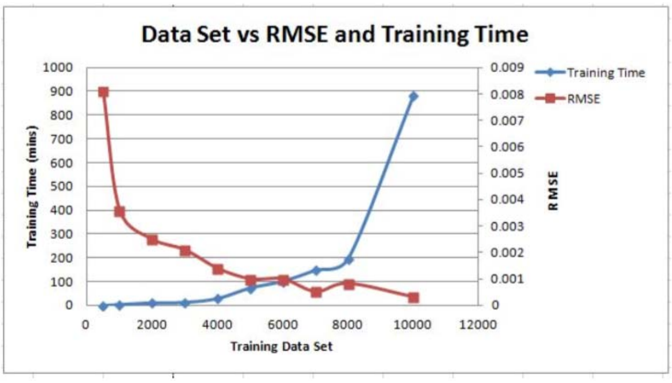

Fig. 1. Variation of training time and RMSE based on training data samples

A computation of RMSE with 10000 test data samples for the 10000 model yielded 4.5224e-4. Showing that the RMSE does not vary much with the number of inputs. The RMSE of the LS-SVM models were lower than that of the ANN model by Opathella et al. Their ANN model has an RMSE of 0.0138 on 10,000 training data points [3]. The LS-SVM model, on the other hand, already surpasses their RMSE on the 500 data point mark. This proves that LS-SVM is a much more superior machine modeling method than ANN.

Furthermore, the time it takes to execute the LS-SVM model once was compared with the results of the other models presented in the paper by Opathella et al. in [3], shown in Table III below.

TABLE III. COMPARISON OF RESULTS FOR BOTH DFIG MODELS

	Fixed PQ	Nonlinear	ANN	LS-SVM
Execution time (msec)	0.002	59.9713	4.9915	6.836

Opathella's paper did not indicate how the results for the other models were derived. Therefore, 10 trials were recorded then averaged for the execution time of the LS-SVM model. It shows that the ANN model is faster than the LS-SVM model by a few milliseconds. The researcher deemed it fit to use the 10,000 training data point model, because it already has a low, acceptable RMSE value. Furthermore, it makes model comparison easier, as the other paper also used 10,000 data points for their ANN model.

C. *Application of the Model to Load Flow Studies*

The LS-SVM wind generator model is then integrated to the IEEE 37-bus unbalanced three-phase distribution test system. The generator specifications are: the 1.5 MW rated doubly-fed induction generator (DFIG) wind generator will be connected to bus 775 of the system. The load flow method employed in this study is the three-phase backward-forward sweep power flow method.

There are two approaches in integrating the LS-SVM model into this power flow analysis. The first is the traditional approach, where the power flow is solved by making the DFIG model a fixed PQ load. The LS-SVM DFIG model is called at the start of the program and its power output will be fixed for the whole load flow process. This case is called LS-SVM LF1 in this study. In the second approach, the LS-SVM DFIG model is executed in each iteration of the power flow solution. The model updates the present value of currents, voltages, powers and losses on each phase. This case is pertained to as LS-SVM LF2 in this study.

The LS-SVM model is compared to the nonlinear model. The results shown in Fig. 2 to Fig. 4 contain the line-to-line voltages of each bus in per unit done with the LS-SVM DFIG model and the two approaches presented in Dadhania's paper. Table IV below shows the RMSE of the LS-SVM model compared to the nonlinear model.

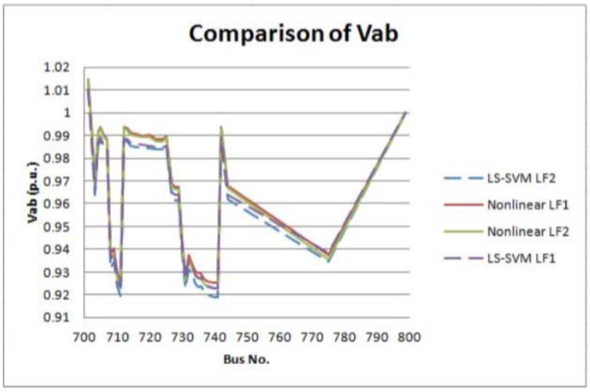

Fig. 2. Comparison of the profile for V_{AB}

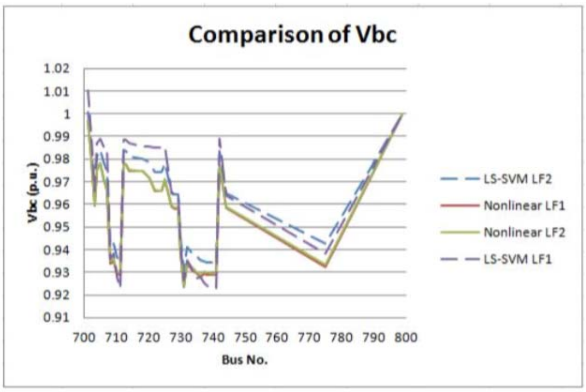

Fig. 3. Comparison of the profile for V_{BC}

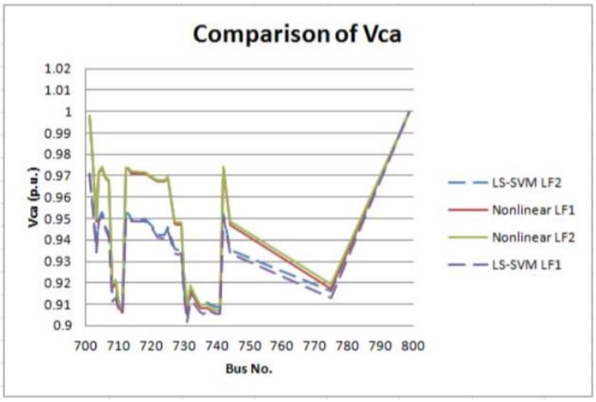

Fig. 4. Comparison of the profile for V_{CA}

TABLE IV. COMPARISON OF RESULTS FOR BOTH DFIG MODELS

LS-SVM Load Flow Approach	RMSE of LS-SVM model vs. Nonlinear model		
	V_{AB}	V_{BC}	V_{CA}
LS-SVM LF1	2.80e-4	2.12e-4	2.57e-4
LS-SVM LF2	6.97e-5	7.17e-5	1.58e-4

The load flow solutions show that there is little to no deviation between the nonlinear and LS-SVM models, provided that the same load flow method was done on both of them. The results prove that the computation time of the LS-SVM model is faster than when using the non-linear model. The first load flow integration approach of the LS-SVM was the fastest because power generated by the wind generator was only computed at the start and was fixed the whole load flow process. Then, the second load flow approach of the LS-SVM because the model computes the output power at a faster speed compared to the nonlinear model. These execution times are indicated in Table V.

TABLE V. EXECUTION TIMES

Model	Time (secs)
LS-SVM LF1	0.041790
LS-SVM LF2	0.355552
Nonlinear	0.813864

IV. CONCLUSIONS

This paper reports the development of LS-SVM DFIG models, their performance and integration to power systems. LS-SVM, a machine-learning algorithm, more widely used in classification problems was applied in wind generator modeling. These models were trained using accurate models of WGs and integrated in load flow studies in MATLAB. The LS-SVM DFIG models were designed to be flexible and generic. It can be trained to simulate any type or size of wind generators.

The proposed model has the same accuracy as nonlinear models, which is the most accurate model at present. It is also much more accurate than the ANN models because of its lower RMSE. In fact, the RMSE was already lower in models with lower training data samples. The RMSE for the 10000 training data sample model was so low that the power outputs for the nonlinear and LS-SVM model were almost perfectly identical. Although there is a significant trade-off between training time and RMSE, it can be decreased by using a much powerful computer. In terms of execution time, the ANN and LS-SVM differ only by a small amount. Thus, it can be seen that the LS-SVM model combines accuracy with fast computation times.

These models were subjected to load flow studies on a three-phase unbalanced system. Results acquired show that the LS-SVM model is favorable. The load flow solutions show that there is little to no deviation between the nonlinear and LS-SVM models, provided that the same load flow method was done on both of them. Furthermore, this model can be easily extended to wind farm modeling.

It is recommended to try other kernel functions which may yield much more favorable results. The modeling can also be extended to represent wind farms and distributed wind generators throughout the system. In reality, however, wind speed is not constant over the turbines of all WGs in a wind farm. The wind speed will change as the wind passes through the blades of a generator. A probabilistic approach can be used in order to take this into account.

REFERENCES

[1] A. Feijoo, "On PQ Models for Asynchronous Wind Turbines," in *Power Systems, IEEE Transactions on*, May 2013, pp. 24(4): 1890-1891.

[2] A. Feijoo and J. Cidras, "Modeling of wind farms in the load flow analysis," in *Power Systems, IEEE Transactions on*, Feb 2000, pp. 15(1): 110-115.

[3] C. Opathella, B.N. Singh, D. Cheng, and B. Venkatesh, "Intelligent Wind Generator Models for Power Flow Studies in PSSE and PSSSincal," in *Power Systems, IEEE Transactions on*, May 2013, pp. 28(2):1149-1159.

[4] C. Opathella, D. Cheng, and B. Venkatesh, "An Intelligent Wind Farm Model for Three-Phase Unbalanced Power Flow Studies," in *Engineering Technology and Technopreneurship (ICE2T), 2014 International Conference on*, Aug 2014, pp. 99-104.

[5] A. Ongsyiping. Simultaneous network recon guration and capacitor switching in distribution systems for loss reduction. In Undergraduate Project, pages 1 30, June 2015.

[6] D. Shirmohammadi, H.W. Hong, A. Semlyen, and G.X. Luo. A compensation-based power flow method for weakly meshed distribution and transmission networks. Power Systems, IEEE Transactions on, 3(2):753–762, May 1988.

[7] J. Suykens and et al, "Advances in Learning Theory: Methods, Models and Applications," in *NATO Advanced Study Institute on Learning Theory and Practice*, Belguim, Marck 2002, pp. 1-6.

[8] Mikko Parviainen. Radial basis function (rbf) and support vector machines (svm) networks. In Class lecture slides, pages 1–12, November 2012.

[9] J. Shawe-Taylor and N. Cristianini, Kernel Methods for Pattern Analysis, 2004.

[9] A. Dadhania, B. Venkatesh, Nasiff A. B., and V.K. Sood. Modeling of doubly fed induction generators for distribution system load flow analysis. In International Journal of Electrical Power and Energy Systems vol. 53, pages 576–583, Dec 2013.