Data Mining For Multi-Criteria Energy Predictions

Kashif Gill and Dennis Moon

***Abstract*—We present a data mining technique for multi-criteria predictions of wind energy. A multi-criteria (MC) evolutionary computing method has been applied for the optimization of an Artificial Intelligence learning methodology "Support Vector Machines" (SVM). The multi-criteria SVM method is applied and tested on a dataset within North America, for predictions of wind energy using climate variables. The SVM training employs Swarm Intelligence method for multi-criteria optimization. The National Center for Environmental Prediction (NCEP)'s global reanalysis gridded dataset has been employed in this study. The gridded dataset for this particular application consists of 4- points each consisting of five variables. In order to study the impact of higher dimensions on the performance of SVM, Principal Component Analysis (PCA) is applied on the input data to reduce the dimensionality of the data. The results of multi-criteria SVM for the prediction of wind energy are reported with and without the pre-processing using PCA.**

***Index Terms*— Multi-Criteria optimization, Swarm Intelligence, Evolutionary computing, Principal Component Analysis, Support Vector Machines, Wind energy.**

I. INTRODUCTION

The current paper describes a method for training support vector machine (SVM) in applications for wind energy predictions at a wind-farm level. The proposed methodology employs a Multi-Objective Evolutionary Optimization approach for training the SVM. The SVM is a powerful learning algorithm developed by Vapnik primarily for classification problems and later was extended to deal with regression [4 and 5]. The method is well-suited for the operational predictions and forecasting of wind power, which is an important variable for power utility companies.

The Multi-Objective optimization approach differs from single objective in that the objective to be optimized is now a vector consisting of more than one objective. The current multi-objective methodology employs Swarm Intelligence based evolutionary computing multi-objective strategy called Multi-objective Particle Swarm Optimization (MOPSO) [6]. The PSO method has been developed for single objective optimization by R. C. Eberhart and J. Kennedy [1]. It has been later extended to solve multi-objective problems by various researchers including the method by [6].

The method originates from the swarm paradigm, called particle swarm, and is expected to provide the so-called global or near-global optimum. PSO is characterized by an adaptive algorithm based on a social-psychological metaphor [1] involving individuals who are interacting with one another in a social world. This sociocognitive view can be effectively applied to computationally intelligent systems [8]. The governing factor in PSO is that the individuals, or "particles," keep track of their best positions in the search space thus far obtained, and also the best positions obtained by their neighboring particles. The best position of an individual particle is called ''local best,'' and the best of the positions obtained by all the particles is called the "global best." Hence the global best is what all the particles tend to follow. The algorithmic details on PSO can be found in [1, 2, 3, and 6]. The approach in [6] presents a multiobjective framework for SVM optimization using MOPSO.

The multiobjective approach to the PSO algorithm is implemented by using the concept of Pareto ranks and defining the Pareto front on the objective function space. Mathematically, a Pareto optimal front is defined as follows: A decision vector $\vec{x}_1 \in S$ is called Pareto optimal if there does not exist another $\vec{x}_2 \in S$ that dominates it. Let $P \subseteq \Re^m$ be a set of vectors. The Pareto optimal front $P^* \subseteq P$ contains all vectors $\vec{x}_1 \in P$, which are not dominated by any vector $\vec{x}_2 \in P$:

$$P^* = \{\vec{x}_1 \in P \mid \nexists \vec{x}_2 \in P : \vec{x}_2 \prec \vec{x}_1\} \quad (1)$$

In the MOPSO algorithm, as devised here [6], the particles will follow the nearest neighboring member of the Pareto front based on the proximity in the objective function (solution) space. At the same time, the particles in the front will follow the best individual in the front, which is the median (middle particle) of the Pareto front. The

Manuscript received August 11, 2009. Kashif Gill and Dennis Moon are with WindLogics, Inc., St Paul, MN 55108 USA (corresponding author: 651-556-4289; fax: 651-556-4210; e-mail: kashif.gill@windlogics.com).

term 'follow' means assignments done for each particle in the population set to decide the direction and offset (velocity) in the subsequent iteration. These assignments are done based on the proximity in the objective function or solution space. The best individual is defined in a relative sense and may change from iteration to iteration depending upon the value of objective function.

The unique formulation of MOPSO helps it to avoid getting struck in local optima, when making a search in the multi-dimensional parameter domain. In the current research, the MOPSO is used to parameterize the three parameters of SVM namely; the trade-off or cost parameter 'C', the epsilon 'ε', and the kernel width 'γ'. The MOPSO method uses a population of parameter sets to compete against each other through a number of iterations in order to improve values of specified multi-objective criteria (objective functions) e.g., root mean square error (RMSE), bias, histogram error (BinRMSE), correlation, etc. The optimum parameter search is conducted in an intelligent manner by narrowing the desired regions of interest and avoids getting struck in local optima.

In our earlier efforts, a single objective optimization methodology has been employed for optimization of three SVM parameters. The approach was tested on a number of sites and results were encouraging. However, it has been noticed that using a single objective optimization method can result in sub-optimal predictions when looking at multiple objectives. The single objective formulation (using PSO) employed coefficient of determination (COD), as the only objective function, but it was noticed that the resulting distributions were highly distorted when compared to the observed distributions. The coefficient of determination (COD) is linearly related to RMSE and can range between –inf to 1; the value of 1 being a perfect fit.

$$COD = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}$$

where O stands for observed, P stands for predicted, $\bar{O}$ stands for mean of observed, and i is the index that goes from 1 to the length of time series 'n'.

It is shown in Figure 1 where a trade-off curve is presented between BinRMSE vs. COD. It can be noticed that COD value increases with the increase in BinRMSE value. The corresponding histograms are also shown in Figure 2 for each of the extreme ends (maximum COD and minimum BinRMSE) and the best compromise solution from the curve. The histograms shown in Figure 2 make it clear that the best COD (or RMSE) is the one with highest BinRMSE and indeed misses the extreme ends of the distribution. Thus no matter how tempting it is to achieve best COD value it does not cover the extreme ends of the distribution. On the other hand the best BinRMSE comes at the cost of lowest COD (or highest RMSE) and is not desired either. Thus it is required to have a multi-objective scheme that simultaneously minimize these objectives and provide a trade-off surface and therefore a compromise solution can be chosen between the two objectives. Figure 2 also shows the histogram for the best compromise solution which provides a decent distribution when compared to observed data.

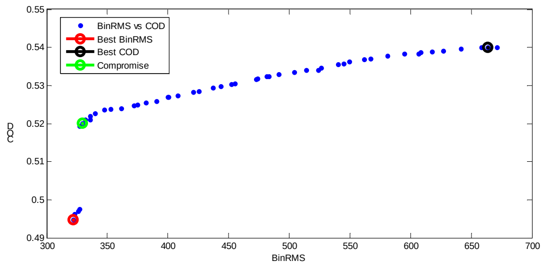

Figure 1: *Trade-off* curve between BinRMSE and COD

II. MATERIAL AND METHOD

The current procedures primarily employ SVM for building regression models for assessing and forecasting wind resources. The primary inputs to the SVM come from the National Center for Environmental Prediction (NCEP)'s reanalysis gridded data [7] centered on the wind farm location. The target is the measurements of wind farm aggregate power. The current training uses a k-fold cross validation scheme referred to as "Round-Robin strategy". The idea within the "Round-Robin" is to divide the available training data into two sets; use one for training and hold the other for testing the model. In this particular "Round-Robin strategy" data is divided into months. The training is done on all the months except one, and the testing is done on the hold-out month. The current operational methods employ manual calibration for the SVM parameters in assessment projects and a simple grid-based parameter search in forecasting applications. The goal in using MOPSO is to explore the regions of interest with respect to the specific multi-objective criteria in an efficient way. Another attractive feature of MOPSO is that it results in a so-called Pareto parameter space, which accounts for parameter uncertainty between the two objective functions. Thus the result is an ensemble of parameter sets cluttered around the so called global optimum w.r.t. the multi-objective space. This ensemble of parameter sets also gives tradeoffs on different objective criteria. The MOPSO-SVM method is tested on the data from an operational assessment site in North America. The results are compared with observed data using a number of evaluation criteria on the validation sets.

As stated above, the data from 4 NCEP's grid points each consisting of 5 variables (total 20 variables) is used.

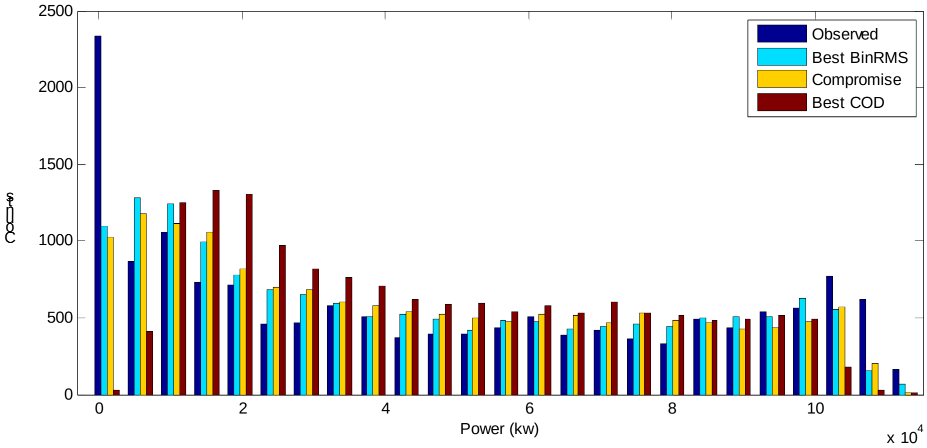

Figure 2: Histogram comparison between Observed, BinRMS best, COD best, and the best compromise solution

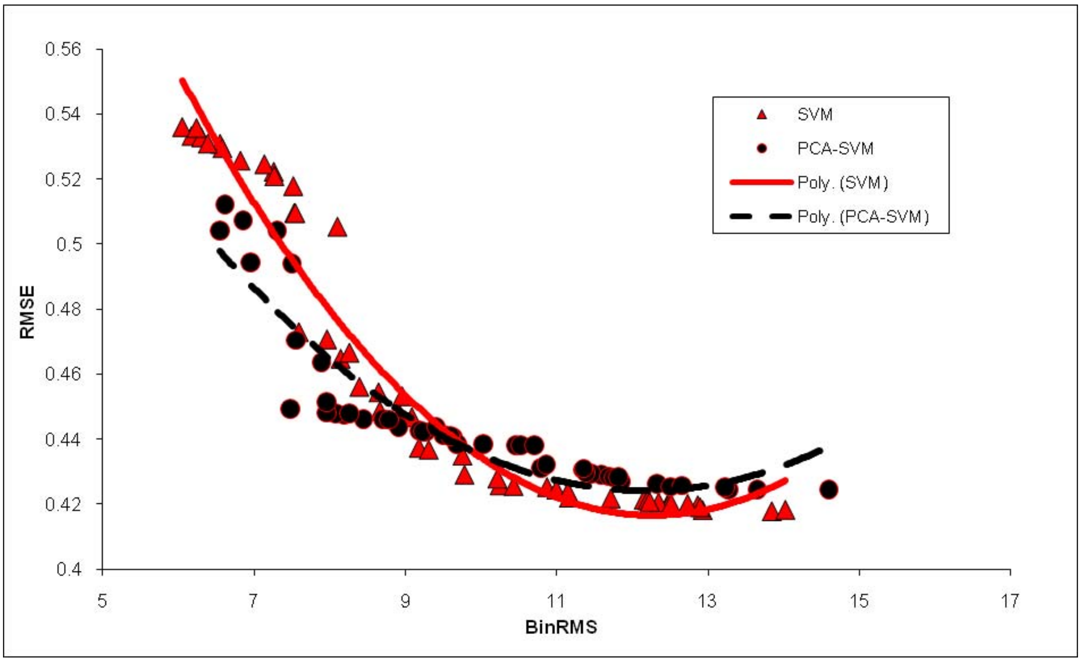

Figure 3: Trade-off curve between the two objectives BinRMS vs. RMSE for SVM and PCA-SVM

The problem therefore has a dimensionality of 20 and can pose a difficult task for SVM. In order to study this impact, Principal Component Analysis (PCA) is applied and the PCs explaining 95% of the variance are included as inputs to SVM. The comparison is made with SVM that does not use PCA as pre-processing step.

MOPSO require a population consisting of parameter sets to be evolved through a number of iterations competing against each other to obtain an optimum (minimum in this case) value for the BinRMSE and RMSE. In the current formulation, a 50 member population is evolved for 100 iterations within MOPSO for wind power predictions at the wind farm.

III. Results and Discussion

The MOPSO-SVM results are shown with and without PCA pre-processing. This wind farm site is located in Canada and has 29 months of energy data available. The MOPSO is used to train SVM over the available 17 months of training data (@ 6-hourly time resolution) using a “Round-Robin” cross-validation strategy. This gives an opportunity to train on 16 months and test the results on a ‘hold-out’ one month test set. By repeating the process for all the 17 months, gives a full 17 months of test data to compare against the observed. Since there is 29 months of data available for this site, a full one- year of data is used in validation (completely unseen data). The results that follow are the predictions on the validation set. The results are shown on the normalized (between -1 and 1) dataset.

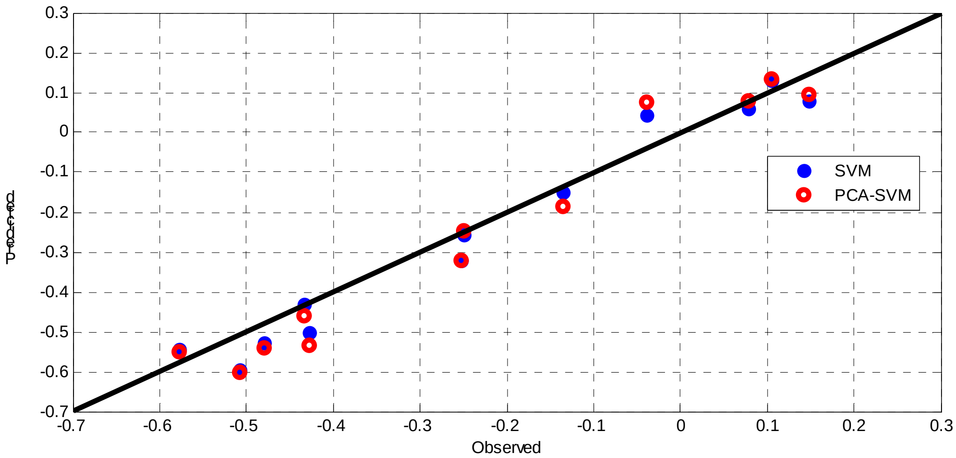

Figure 4: Scatter plot for the mean monthly wind energy data for SVM and PCA-SVM

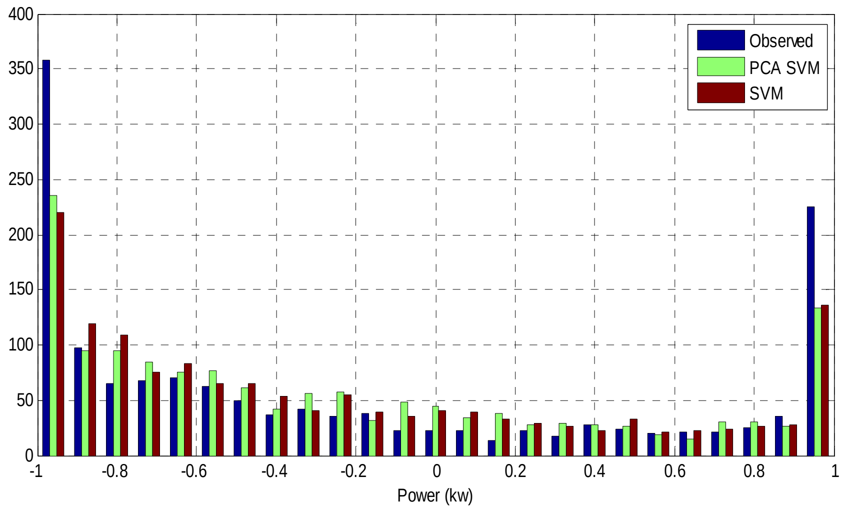

Figure 5: Histogram for SVM and PCA-SVM along with observed

The trade-off curve for MOPSO SVM optimization for the two objectives is shown in Figure 3. The trade-off between BinRMS vs. RMSE is shown for the SVM using original input compared against SVM using Principal Components (PCs) as inputs. A best fit line has also been shown for the two curves. It can be noticed that there is little difference between the two approaches. The PCA-SVM produced a better objective function result for BinRMS, where as simple SVM provided a better objective function result for RMSE.

Figure 4 shows the monthly mean wind power for the 12 months compared against the observed data. The results are shown for SVM prediction with and without the pre-processing using PCA. As stated above, there is a very little difference between the two approaches and a good fit has been found. It can be noticed that predictions are in reasonable agreement with the observed data. The results in Figure 5 show histogram of observed vs. the predicted wind power data at the 6-hourly time resolution (the prediction time step). The results are shown for SVM prediction with and without the pre-processing using PCA. It can be noticed that the distributions are well-maintained using MOPSO methodology and a reasonable agreement between observed and predicted power is evident from the figure.

A number of goodness-of-fit measures are evaluated in Table 1, which are monthly root mean square error (RMSE), monthly coefficient of determination (COD), instantaneous RMSE, instantaneous COD, and BinRMSE (histogram bin RMSE). The results in Table 1 are presented for SVM prediction with and without the pre-processing using PCA. Both monthly and instantaneous wind power are of significant interest, and thus are included in the current analysis. It can be noticed that not only monthly but also instantaneous power predictions are in close agreement with the observed.

Table 1: Wind energy goodness-of-fit

Goodness measure	MC-SVM	
	SVM	PCA-SVM
Monthly RMSE	0.053	0.064
Monthly COD	0.954	0.934
Instantaneous RMSE	0.380	0.382
Instantaneous COD	0.735	0.734
BinRMSE	37.758	36.491

IV. CONCLUSIONS

In the present paper, a multi-objective evolutionary computing method MOPSO is used to optimize the three parameters of SVM for wind energy predictions. The approach has been tested on data from a wind farm using NCEP's re-analysis grid data. The prediction strategy employs SVM with and without the pre-processing using PCA. A number of graphical and tabular results in the form of goodness-of-fit measures are presented for wind energy predictions. The SVM predictions at the wind farm level produced excellent agreement with the observed data for the validation set. The results for the two approaches are quite similar but SVM without any pre-processing using PCA produced slightly better results. Overall, the results have been encouraging and it is recommended to use MOPSO-SVM approach for other operational projects in the area of wind power predictions and forecasting. While further modifications and advancements are underway, the current procedure is sound enough to be applied in operational settings.

REFERENCES

[1] Eberhart, R. C., and J. Kennedy, A new optimizer using particle swarm theory, in Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 1995, MHS '95, 1995, pp. 39– 43, doi:10.1109/MHS.1995.494215, IEEE Press, Piscataway, N. J.
[2] Kennedy, J., and R. C. Eberhart, Particle swarm optimization, in Proceedings of IEEE International Conference on Neural Networks, IV, vol. 4, 1995, pp. 1942– 1948, doi:10.1109/ICNN.1995.488968, IEEE Press, Piscataway, N. J.
[3] Eberhart, R. C., R. W. Dobbins, and P. Simpson, Computational Intelligence PC Tools, Elsevier, New York, 1996.
[4] Vapnik, V., The Nature of Statistical Learning Theory, Springer, New York, 1995.
[5] Vapnik, V., Statistical Learning Theory, John Wiley, Hoboken, N. J., 1998.
[6] Gill, M. K., Y. H. Kaheil, A. Khalil, M. McKee, and L. Bastidas, Multiobjective particle swarm optimization for parameter estimation in hydrology, Water Resour. Res., 42, 2006, W07417, doi:10.1029/2005WR004528.
[7] Kalnay et al.,The NCEP/NCAR 40-year reanalysis project, Bull. Amer. Meteor. Soc., 77, 1996, 437-470.
[8] Kennedy, J., R. C. Eberhart, and Y. Shi, (2001). Swarm Intelligence.Morgan Kaufmann, San Francisco, CA.