Applied Energy 217 (2018) 422–439

Contents lists available at ScienceDirect

# Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

# Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting

Mohanad S. AL-Musaylh[a,b,*], Ravinesh C. Deo[a,*], Yan Li[a], Jan F. Adamowski[c]

[a] School of Agricultural, Computational and Environmental Sciences, Institute of Agriculture and Environment (IAg&E), University of Southern Queensland, QLD 4350, Australia
[b] Management Technical College, Southern Technical University, Basrah, Iraq
[c] Department of Bioresource Engineering, Faculty of Agricultural and Environmental Science, McGill University, Québec H9X 3V9, Canada

## HIGHLIGHTS

- Hybrid two-phase PSO-SVR is integrated with CEEMDAN multi-resolution tool for demand forecasting.
- ICEEMDAN-PSO-SVR is evaluated against single-phase hybrid and standalone models.
- ICEEMDAN-PSO-SVR outperforms several benchmark models at multiple-horizons.
- Two-phase hybrid model has potential applications in energy management systems.

## ARTICLE INFO

*Keywords:*
SVR
PSO
Improved CEEMDAN
Electricity demand
MARS
M5 model tree
Energy management system

## ABSTRACT

Real-time energy management systems that are designed to support consumer supply and demand spectrums of electrical energy continue to face challenges with respect to designing accurate and reliable real-time forecasts due to the stochasticity of model construction data and the model's inability to disseminate both the short- and the long-term electrical energy demand ($G$) predictions. Using real $G$ data from Queensland, Australia's second largest state, and employing the support vector regression (SVR) model integrated with an improved version of empirical mode decomposition with adaptive noise (ICEEMDAN) tool, this study aims to propose a novel hybrid model: ICEEMDAN-PSO-SVR. Optimization of the model's weights and biases was performed using the particle swarm optimization (PSO) algorithm. ICEEMDAN was applied to improve the hybrid model's forecasting accuracy, addressing non-linear and non-stationary issues in time series inputs by decomposing statistically significant historical $G$ data into intrinsic mode functions (IMF) and a residual component. The ICEEMDAN-PSO-SVR model was then individually constructed to forecast IMFs and the residual datasets and the final $G$ forecasts were obtained by aggregating the IMF and residual forecasted series. The performance of the ICEEMDAN-PSO-SVR technique was compared with alternative approaches: ICEEMDAN-multivariate adaptive regression spline (MARS) and ICEEMDAN-M5 model tree, as well as traditional modelling approaches: PSO-SVR, MARS and M5 model tree algorithms. To develop the models, data were partitioned into different subsets: training (70%), validation (15%), and testing (15%), and the tuned forecasting models with near global optimum solutions were applied and evaluated at multiple horizons: short-term (i.e., weekends, working days, whole weeks, and public holidays), and long-term (monthly). Statistical metrics including the root-mean square error (*RMSE*), mean absolute error (*MAE*) and their relative to observed means (*RRMSE* and *MAPE*), Willmott's Index (*WI*), the Legates and McCabe Index ($E_{LM}$) and Nash–Sutcliffe coefficients ($E_{NS}$), were used to assess model accuracy in the independent (testing) period. Empirical results showed that the ICEEMDAN-PSO-SVR model performed well for all forecasting horizons, outperforming the alternative comparison approaches: ICEEMDAN-MARS and ICEEMDAN-M5 model tree and the PSO-SVR, PSO-MARS and PSO-M5 model tree algorithm. Due to its high predictive utility, the two-phase ICEEMDAN-PSO-SVR hybrid model was particularly appropriate for whole week forecasts ($E_{NS}$ = 0.95, $MAPE$ = 0.89%, $RRMSE$ = 1.22%, and $E_{LM}$ = 0.79), and monthly forecasts ($E_{NS}$ = 0.70, $MAPE$ = 2.18%, $RRMSE$ = 3.18%, and $E_{LM}$ = 0.56). The excellent performance of the ICEEMDAN-PSO-SVR hybrid model indicates that the two-phase hybrid model should be explored for potential applications in real-time energy management systems.

* Corresponding authors at: School of Agricultural, Computational and Environmental Sciences, Institute of Agriculture and Environment (IAg&E), University of Southern Queensland, QLD 4350, Australia.
*E-mail addresses:* MohanadShakirKhalid.AL-Musaylh@usq.edu.au (M.S. AL-Musaylh), ravinesh.deo@usq.edu.au (R.C. Deo).

https://doi.org/10.1016/j.apenergy.2018.02.140
Received 13 November 2017; Received in revised form 13 February 2018; Accepted 21 February 2018

**Nomenclature**

| | |
|---|---|
| MW | megawatt |
| $G$ | electricity load (demand; mega watts) |
| MARS | multivariate adaptive regression splines |
| SVR | support vector regression |
| $RMSE$ | root-mean square error |
| $MAE$ | mean absolute error |
| $RRMSE$ | relative root-man square error, % |
| $MAPE$ | mean absolute percentage error, % |
| $WI$ | Willmott's index |
| $E_{NS}$ | Nash–Sutcliffe coefficient |
| $E_{LM}$ | Legates and McCabe Index |
| $\omega$ | weighting factor of PSO |
| $\omega_{min}$ and $\omega_{max}$ | the minimum and maximum of $\omega$ |
| RBF | radial basis function for SVR |
| ICEEMDAN | improved version of empirical mode decomposition with adaptive noise |
| $\varepsilon$ | loss function |
| $\sigma$ | kernel width for SVR model |
| $C$ | regulation for SVR model |
| PSO | particle swarm optimization |
| $GCV$ | generalized cross-validation |
| $T_{max}$ | maximum number of iterations in PSO |
| ECDF | empirical cumulative distribution function |
| IMF | intrinsic mode functions |
| ARIMA | autoregressive integrated moving average |
| ANN | artificial neural network |
| AEMO | australian energy market operator |
| DWT | discrete wavelet transform |
| $c_1$ and $c_2$ | PSO parameters |
| PACF | partial auto-correlation function |
| $MSE$ | mean square error |
| $R^2$ | coefficient of Determination |
| MODWT | maximum overlap discrete wavelet transform |
| EMD | empirical mode decomposition |
| WT | wavelet transforms |
| $N$ | the initial population of PSO |
| $G_i^{for}$ | $i$th forecasted value of $G$, MW) |
| $G_i^{obs}$ | $i$th observed value of $G$, MW) |
| $\overline{G^{for}}$ | the mean of forecasted values |
| $\overline{G^{obs}}$ | the mean of observed values |
| $\lvert FE \rvert$ | forecasted error statistics |
| EEMD | ensemble EMD |
| CEEMDAN | complete EEMD with adaptive noise |
| $r$ | correlation coefficient |
| VMD | variational mode decomposition |

## 1. Introduction

Electricity demand ($G$) forecasting can provide essential information that is likely to be utilized for energy transactions in competitive electricity markets [1–3]. Policies addressing energy distribution and pricing and providing energy security to a growing population requires accurate forecasting of $G$ data, especially for short-term periods (e.g., daily). Estimating $G$ is a very sensitive task as an error in under- or over-estimation of even just 1% can lead to millions of dollars in losses affecting the whole energy policy and management system [3–5]. As such, to estimate $G$, a very accurate near real-time (i.e., short-term), as well as a foresight (i.e., long-term), forecasting model is a useful tool.

In recent years, data-driven models, such as autoregressive integrated moving average (ARIMA) [6], artificial neural network (ANN) [7], support vector regression (SVR) [8], genetic algorithms, fuzzy logic, knowledge-based expert systems [9], M5 model tree [10,11] and multivariate adaptive regression splines (MARS) [12] have been widely adopted in energy demand forecasting studies. Based on structural risk minimization (SRM), the SVR model is able to reduce overfitting data through the minimization of expected error of a learning algorithm [13]. For example, the SVR model with a radial basis kernel function (RBF) has been used for $G$ forecasting [14]. The parameters of the SVR model can be selected by different optimization techniques, such as a grid search procedure [15], particle swarm optimization (PSO) [16] and a genetic algorithm [17]. The PSO algorithm can be considered an effective method to solve engineering challenges and can also be used to provide better performance when used to screen the near global optimum set of SVR parameters [18,19]. On the other hand, the MARS model is a fast and flexible statistical tool that can be developed to adopt a piecewise (linear or cubic) basis function [20,21]. In the literature, a significantly lower root-mean-square error (*RMSE)* was found for the MARS model when compared with the piecewise regression-based model used for $G$ forecasting [12]. A piecewise linear function in a M5 model tree [22] has also been used in different studies including wave [10] and solar energy studies [11].

However, these types of traditional machine learning methods often have challenges addressing non-stationary time series [22,23]. Non-stationary time series problems can be addressed by different model input data decomposition methods; for example, the discrete wavelet transform (DWT) [8], maximum overlap discrete wavelet transform (MODWT) [24], empirical mode decomposition (EMD) [25], ensemble EMD (EEMD) [26], complete EEMD with adaptive noise (CEEMDAN) [27] or improved CEEMDAN (termed hereafter as the 'ICEEMDAN' algorithm) [28]. These algorithms resolve the frequency components present in input series prior to using them in the model development process. These techniques are powerful tools as they can be used to decompose the original data into high and low frequency sub-series to address the issues of non-stationary, repeats/periods and jump-type perturbations before such data are utilized for prediction purposes.

In spite of the many applications of wavelet transforms (WT) (e.g., [29–34]), recent studies show major weaknesses in WT-based models, particularly in their forecasting ability, which is limited by the adoption of non-causal filters constructed with DWT algorithms. It should be noted that DWT can induce the decimation effect in model input sub-series coefficients, and therefore generate half the coefficients of the detailed signal at the current level, while the other half of the smooth version can be recursively processed by the high pass and low pass filters at a coarser temporal resolution [35]. Although the problem of the decimation effect in DWT can be solved by the more advanced MODWT algorithm, selection of the mother wavelet is a still major issue as there is no rule to select a near global optimum wavelet other than applying an iterative trial and error process [24]. However, there is an alternative decomposition tool available to address such issues: the self-adaptive EMD algorithm that splits data into several intrinsic mode functions (IMFs) and a residual data subset. While the frequent appearance of mode mixing in the EMD algorithm is problematic [26], it can be addressed by the EEMD-based model, which is able to obtain a true number of IMFs. However, when a signal is reconstructed using the EEMD process, different numbers of IMFs can be obtained which generates a new problem [27]. In order to resolve all of these issues, CEEMDAN was developed [27] to precisely reconstruct the original time series data and give a better spectral separation of the IMFs at a lower computational cost. Some of the residual noise within IMFs and the slower performance of the algorithm compared to the EEMD are two major issues associated with the CEEMDAN algorithm [28]. Hence, the ICEEMDAN algorithm was developed to address issues of model input decomposition [28].

Few studies (e.g. [16,36,37]) have applied the ICEEMDAN for