ISSN: 0974-5823 Vol. 7 (Special Issue 5, April-May 2022)

International Journal of Mechanical Engineering

Power Consumption prediction using Random Forest model

Waleed Abdu Zogaan

Department of Computer Science, Faculty of Computer Science and Information Technology, Jazan University, Jazan, 45142, Saudi Arabia

ABSTRACT

In the era of the energy crisis, Power Consumption (PC) plays a critical role in a global economy due to the imbalance between energy production and demand. Machine learning (ML) models have been widely recognized as a precise and computationally effective solution in prediction, which can help energy managers to control power systems better and improve energy usage. In this paper, three different models comprising Support Vector Machine (SVM), Artificial Neural Network (ANN), and Random Forest (RF) are employed to predict PC. An open-source dataset is used to validate the efficiency of the models, and different performance measures are employed to evaluate the effectiveness of the models. Experimental results show that RF model gets the least errors and the highest correlation between measured and predicted PC during training and test phases. In addition, the required training and testing time by the RF is also the smallest compared to the other two models. RF model can provide a practical and accurate solution for PC prediction.

Keywords: Energy, power consumption, machine learning, prediction, Artificial intelligence.

INTRODUCTION

Industrial modernization has constantly increased the energy demand of our economy over decades, and the energy supply is limited. PC presents a significant slice of energy use in the world, which plays a crucial role economy, and therefore, it is vitally important to manage this problem. PCBased on historical data, PC prediction in the early stages would be an essential tool for energy suppliers, managers, and society. ML can help design a robust and accurate prediction model for PC based on a given historical data, which is important for energy management. ML models have become popular for predicting PC due to their capabilities to learn from a vast amount of historical energy data and provide energy managers with more accurate results.

In recent years, PC prediction has gained special attention in academia [1–5]. Several ML models have been reported to predict PC using ANN [6, 7, 8]. In [9], a novel Neural Network (NN) approach is presented to predict PC. they usedNNbased Genetic Algorithm denoted by NNGA and the Neural Network based Particle Swarm Optimization (NNPSO) for optimizing the weight of the NN. The proposed approach achieved better outcomes than the CNN model for PC usinga real-time dataset obtained from pecan street. In another work [10], five ML models, including Multiple Regression (MR), SVM, ANN, Deep Neural Network (DNN), and Genetic Programming (GP), are compared to predict PC in a building. The findings showed that the ANN model performs better than the other models. In [11], feedforward back-propagation ANN and RF are investigated to predict the PC of a hotel in Madrid, Spain. The results indicated that both models are equally applicable for PC prediction.

In [12], SVM with Radial Basis Function (RBF) kernel function is employed to predict PC in hotels. The results showed that the SVM model effectively predicts actual energy usage in the hotel for improving the hotel's operations by reducing PC. In [13], ANN and SVM are compared for PC prediction and the results indicated that both models have similar prediction power. In [14], PCk-fold cross-validation and grid search methods are used to enhance the efficiency of the SVM in predicting PC of a building in China. The authors conclused that the optimized SVM model can effectively predicts building PC with a good prediction accuracy. In [15], a Genetic Algorithm (GA) is used to fine-tune SVM's parameters to predict electricity PC. The proposed combined model showed promising improvements in the predictive accuracy compared to the other models used in their study. In [16], SVM is used to predict PC using data collected from SCADA Office Intelligent Context Awareness Management (SOICAM) system. The obtained results confirmed that SVM has an excellent ability to predict PC than ANN.

PCIn [17], different ML models comprising ANN, RF, Decision Tree (DT), and SVM for regression (SVR) with RBF kernel function are assessed for PC prediction. The results showed that RF performed the best. A hybrid PC prediction model in [18], first classified data by pattern similarity using the DT method. It uses two models, RF and Multilayer Perceptron, to select the best one with a higher prediction capability using three different power distribution networks of Tetouan city located in Morocco. The results of comparative analysis using three university building clusters showed that PCthe RF model is more capable of predicting PC than other models. In [19], various regression models are used to predict urban area electrical energy demand in an urban area located in Sydney. The results showed that the RF model attained the best predeiciton results for the demand short-term electrical energy.

This work aims to investigate the best and most accurate model for predicting PC using commonly available measurements. Three different models comprising SVM, ANN, and RF predict PC using an open-source dataset. These models are evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), coefficient of determination (R2), Pearson's correlation coefficient, and computation time for training and testing phases. The rest of this paper is structured as follows: Section 2 describes the dataset used to validate the models and PCprovides an overview of the models used for PC prediction. Section 3 presents the statistical metrics used to evaluate the models. A comparative analysis of all the models is presented in section 4 and finally, section 5 provides the conclusion and future work of the paper.

MATERIAL AND METHOD

Study case

A dataset collected from the Supervisory Control and Data Acquisition System (SCADA) of Amendis, a public service operator in charge of distributing drinking water and electricity since 2002 is used. It is freely available on UCI Machine Learning Repository [20], and it was collected every 10 minutes for the period of: January–December 2017. The dataset does not have missing values and comprises 52417 instances and five different attributes including the date, time, temperature, humidity, wind speed, general diffuse flows, diffuse flows, and PC of three distribution zones. Variation PCs is shown in Figure 1.

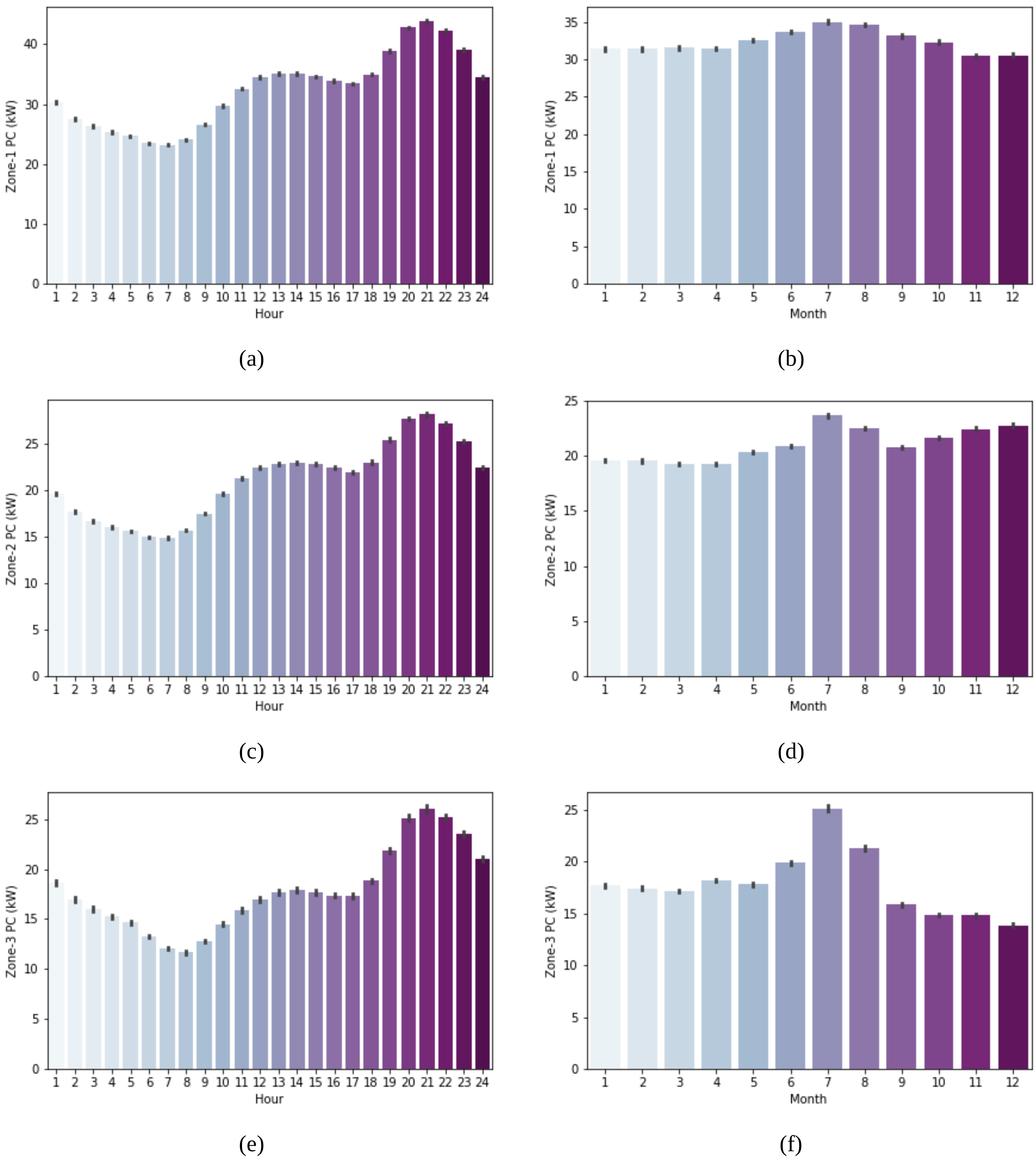


Figure. 1: Variation of PC for all three zones w.r.t. hour of the day (a, c, e) and month of the year (b, d, f).

Figure 2 shows the correlation between different attributes pairs. It can be seen from the figure that only temperature is moderately correlated with PC of different zones while other attributes show very small correlation with PC.

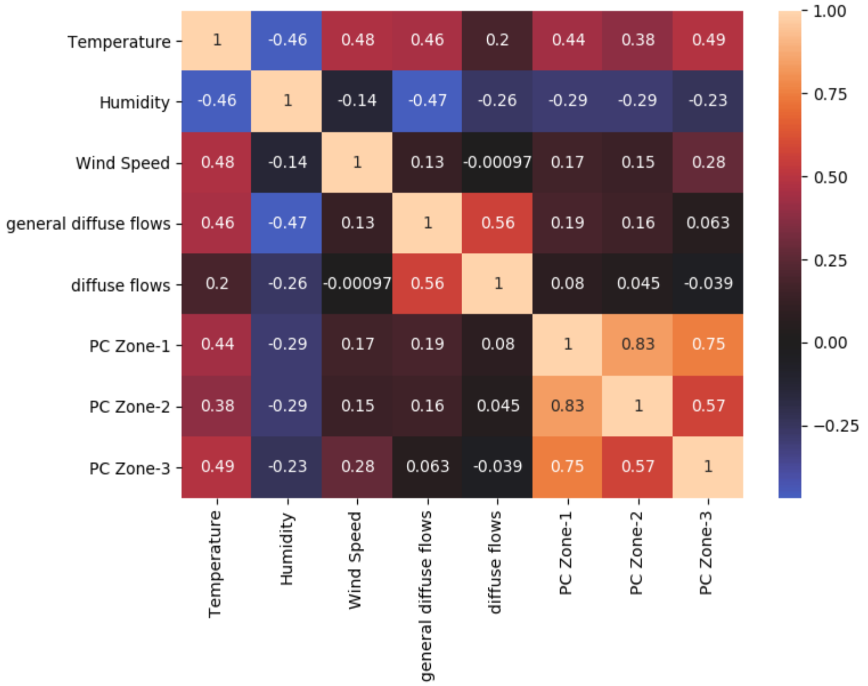


Figure. 2: Correlation matrix for attribute pairs in the PC dataset.

Furthermore, the dataset is preprocessed to generate five new attributes including hour, minute, quarter of the year, day of month, day of year, day of week. A revised correlation matrix with additional features is shown in Figure 3. It can be seen that newly added feature 'hour' is well correlated with PC of all three zones while other are moderately correlated with PC of at least one of the zones.

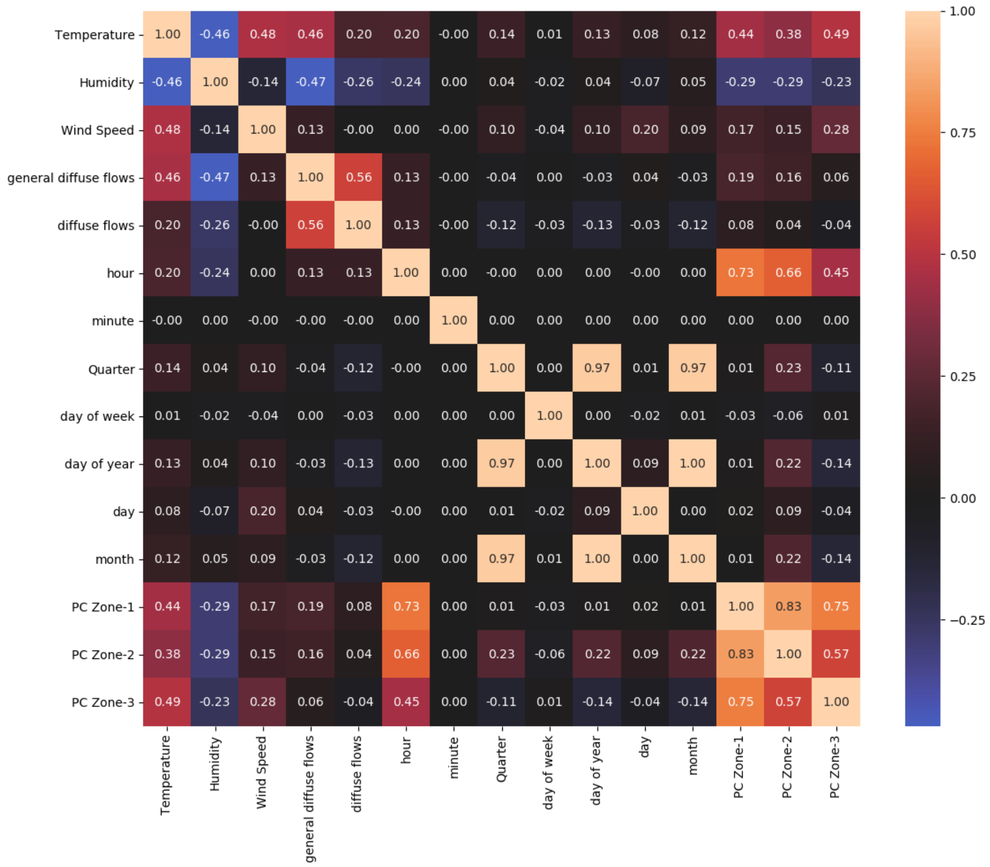


Figure. 3: Revised correlation matrix for attribute pairs in the PC dataset with additional features.

ML MODELS

Support Vector Machine (SVM)

SVM algorithm is commonly employed to solve various classification and regression problems [21]. It constructs an N-dimensional hyperplane to classify the data into two groups. Its ability to handle complex parameters and the flexible use of kernel function that enhances its ability to convert data into high dimensional space. In this work, RBF as a kernel function is used for the SVM model.

Suppose the training sample set $G = \{x_I, y_i\}_i^N$ is given, where x_I, is ith input vector of the data sample, y_I is the ith target value, and N is the total number of data samples. SVM approximates the function as [21]:

$$f(x) = w\varphi(x) + b, \quad (1)$$

where $\varphi(x)$ is used to map x_I into a higher dimensional feature-space, w is a normal vector, and b is a scalar

$$\begin{cases} \min \frac{1}{2} [\![w]\!]^2 + C\sum_{i=1}^{N} \zeta_i, \\ s.t., \ y_i(w^T x_i + b) \geq 1 - \zeta_i, \\ \zeta_i \geq 0, \end{cases} \quad (2)$$

where C is the cost of penalty and ζ_i is a positive slack variable.

The hyperplane function can be written as:

$$f(x_i) = \sum_{n-1}^{N} \alpha_n y_n K(x_n, x_i) + b \quad (3)$$

where. A_n, y_n, and$x_n$are Lagrange multiplier, membership class label, and the support vector for nth class and $K(x_n, x_i)$ is the kernel function.

Artificial Neural Network (ANN)

ANN, inspired by the human brain [22], it is excellent to learn the nonlinear relationship between input and output features. Multilayer Perceptron (MLP) architecture of ANN comprises an input layer, an output layer, and one or more hidden layers. Each hidden layer consists of multiple neurons activated by a nonlinear function that helps to transform a weighted linear combination of outputs from the previous layer. The dataset specifies the number of neurons in input and output layers. In this work, the MLP-ANN model uses ten input nodes corresponding to dataset attributes and one output to predict PC. All input attributes and output PC values are normalized individually to have zero mean and unit variance. These statistics are recorded to apply an inverse transformation to predict PC values. ANN with one hidden layers (5 neurons) and rectified linear unit activation function is used in this work. The model is trained using a back-propagation algorithm with an Adam optimizer to minimize the prediction error.

$$y = \text{sigmoid}(Whd * \text{relu}(Win * X + Bin) + Bhd) \quad (4)$$

where, W_{hd} and W_{in} are weight matrices for hidden and inout layers, B_{hd} and B_{in} are bias vectors for hidden and inout layers, X is an inout vector and y is output class prediction.

Random Forest (RF)

RF, initially introduced by [23], is an ensemble classifier. It gained popularity in scientific and engineering applications due to its training speed, suitability for regression and prediction problems, and efficiency in handling complex datasets [24]. RF generates many non-pruned DT and then it aggregates their results using a majority voting and each. tree is constructed from bootstrap data drawn from the training data to increase the diversity of the trees. On the other hand, the samples not involved in the construction phase refer to “Out-Of-Bag” (OOB)data. The algorithm internally uses this OOB data as validation data during the training phase. An RF model with ten ensemble trees with a minimum of ten samples per node is used to optimize PC prediction. The RF prediction model can be presented as [23] :

$$f_{rf}^{N}(x) = \frac{1}{N}\sum_{n=1}^{N} T_{ree}(x), \qquad x = x_1, x_2, \cdots x_p \quad (5)$$

where N represents the average number of regression trees built by RF, x is a p-dimensional vector of inputs and T_{ree} refers to DT.

PERFORMANCE METRICS

Four different statistical measures are used to assess the performance of ML models. These metrics are selected due to their suitability for PC prediction in earlier works [14, 16, 25]. The metrics are defined as follows:

$$\boldsymbol{MAE} = \frac{1}{N}\sum_{i=1}^{N}|m_i - p_i| \quad 6$$

$$\boldsymbol{MSE} = \frac{1}{N}\sum_{i=1}^{N}(m_i - p_i)^2 \quad 7$$

$$\boldsymbol{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(m_i - p_i)^2} \quad 8$$

$$\boldsymbol{R^2} = 1 - \frac{\sum_{\forall i}(m_i - p_i)^2}{\sum_{\forall i}(m_i - \mu_m)^2} \quad 9$$

where N is the number of observations, m_i is the ith measured PC value, p_i is ith PC value predicted by the model and μ_m is the mean of measured PC values.

EXPERIMENTAL RESULTS AND DISCUSSION

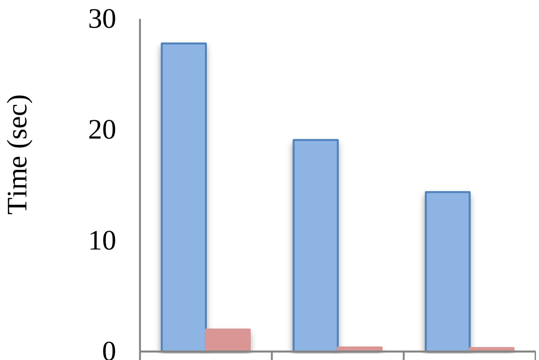
The capability of SVM, ANN, and RF models for PC prediction are assessed using statistical metrics and visualization. Table 1 provides the model's performance metrics for both the training and testing datasets. The methods are implemented using a Python-based ML environment on Windows 7 with a 2 Duo CPU running on a 3.13 GHz PC with 44.25 GB RAM.

Table 1. The statistical performance of PC prediction using 5-fold cross validation

Zone	Model	Dataset	MAE	MSE	RMSE	R^2
1	SVM	Mean	0.7073	1.1952	1.0925	0.9765
		Std.	0.0166	0.0879	0.0403	0.0016
	ANN	Mean	2.8441	13.5639	3.6584	0.7335
		Std.	0.3485	3.1387	0.4244	0.0616
	RF	Mean	0.2166	0.0987	0.3140	0.9013
		Std.	0.0038	0.0046	0.0073	0.0050
2	SVM	Mean	0.5132	0.6189	0.7866	0.9772
		Std	0.0091	0.2378	0.1500	0.0011
	ANN	Mean	2.3411	9.1341	3.0192	0.6633
		Std	0.1165	0.7990	0.1358	0.0306
	RF	Mean	0.2265	0.0994	0.3152	0.9005
		Std	0.0037	0.0034	0.0053	0.0042
3	SVM	Mean	0.4367	0.5144	0.7167	0.9883
		Std	0.0055	0.0038	0.0075	0.0041
	ANN	Mean	2.8906	14.3779	3.7858	0.6722
		Std	0.1979	1.5973	0.2136	0.0361
	RF	Mean	0.2161	0.0930	0.3049	0.9069
		Std	0.0041	0.0046	0.0027	0.0010

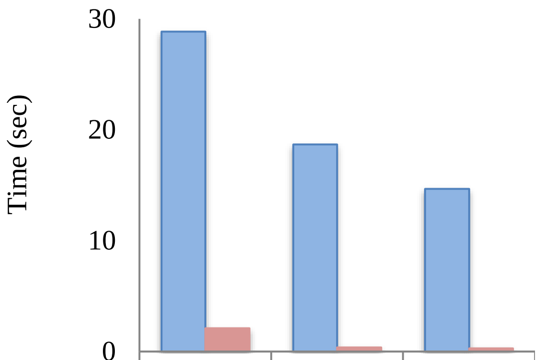
Std refers to standard deviation

Based on the statistical results provided in Table 1, RF performed better than the other models in terms of mean and Std. The RF model obtained lowest errors mean and Std of MAE, MSE, RMSE on zone 1, 2 and 3. The smallest MAE of 0.2161, MSE of 0.0930, RMSE of 0.3049, and and highest R2 of 0.9069 are obtained for zone 3. These results confirm the capability of the RF model in predicting PC compared to the other models. RF needs the least time during the training phase than SVM and ANN models, as shown in Figure 4. It can be seen from figure 4 that the RF model required less time in the training and testing phases compared to the other mdoels for the three zones is slightly higher than GB, the performance is significantly improved. These results prove the effectiveness of the RF model over the other comparable models for PC prediction.



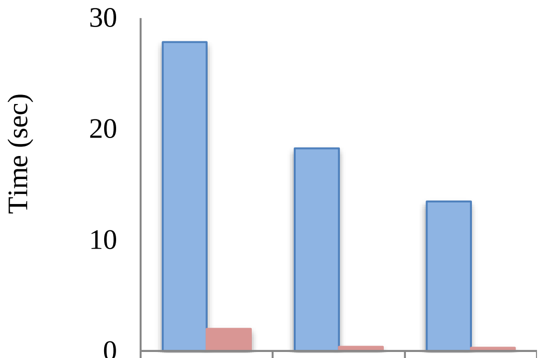
	SVM	ANN	RF
Training	27.7903	19.0743	14.3767
Testing	2.0033	0.3685	0.3108

(a)



	SVM	ANN	RF
Training	28.8584	18.6761	14.6763
Testing	2.1056	0.3737	0.3067

(b)

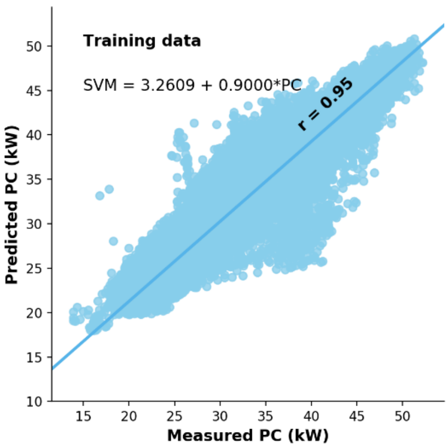


	SVM	ANN	RF
Training	27.8222	18.2671	13.4751
Testing	2.0015	0.3928	0.3032

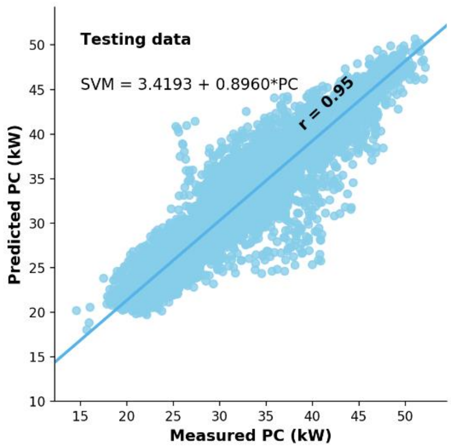
(c)

Figure 4. Computational time (in seconds) for different zones': (a) zone 1, (b) zone 2, and (c) zoen 3 using all the models

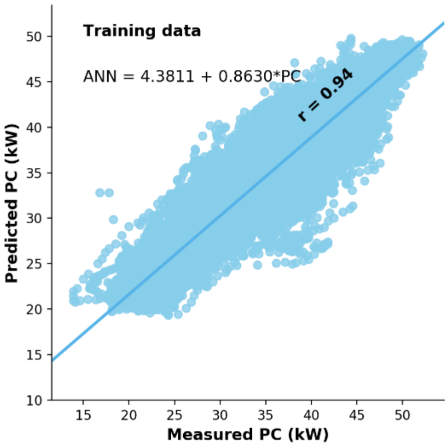
The relationship between the predicted PC values using the models and the measured values are illustrated in scatter plots for training and testing datasets. The linear regression of predicted and original PC is shown with a line, and Pearson's correlation coefficient (r) are shown in figure 5,6 and 7 for the zone 1, 2 and 3 respectively . From figures 5, 6 and 7, the RF model has the highest correlation of 1 and 0.99 for the training and testing on the zone 1 and 2 respectively and it attaeined the highest correlation on zone 3. The PC prediction by the RF is much closer to measured data during tainting and testing for all the three zone . Both SVM and ANN showe almost equal correlation for both the training and testing phases.



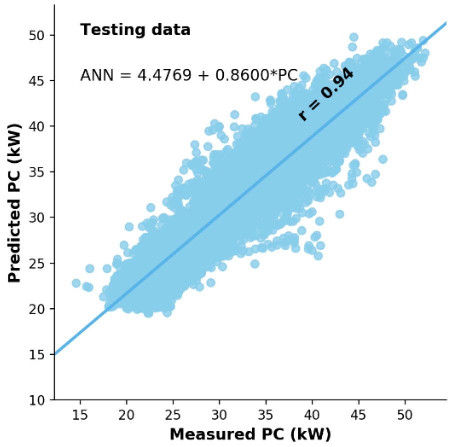
(a)



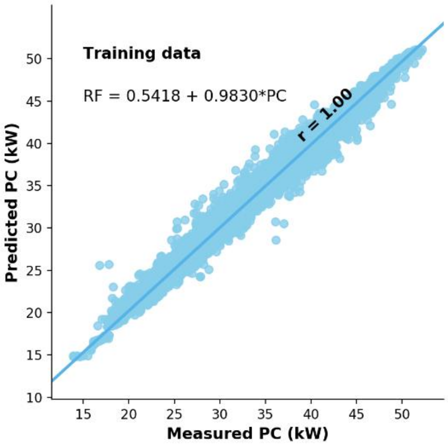
(b)



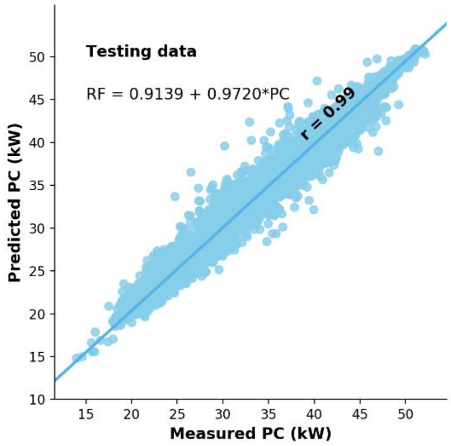
(c)



(d)

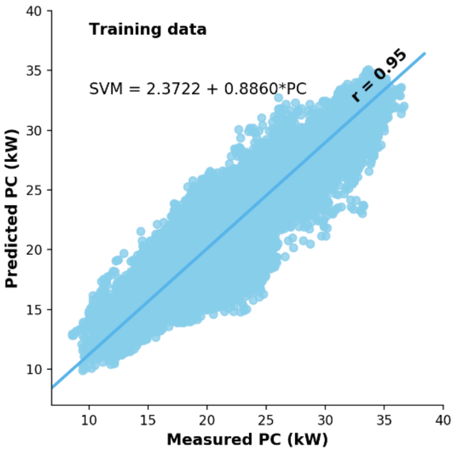


(e)

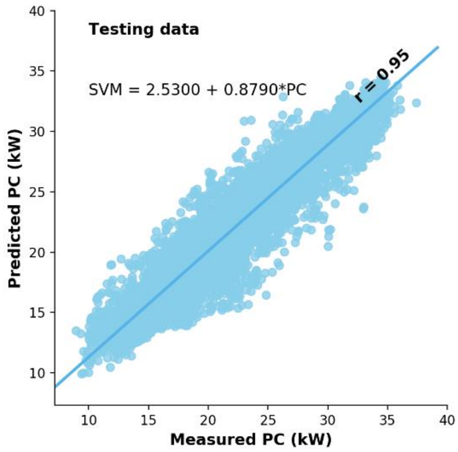


(f)

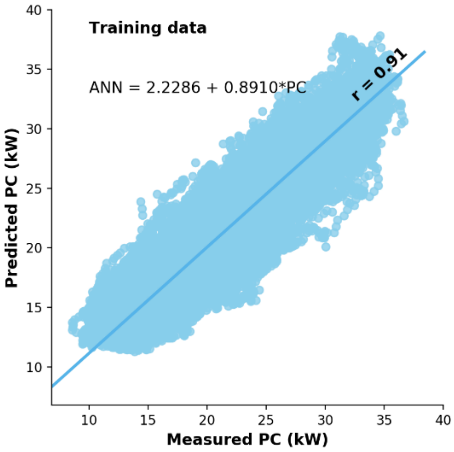
Figure 5. Scatter-plots performance of the Zone-1 PC prediction by the SVM((a)-(b)), ANN((c)-(d)), and RF ((e)-(f)) for training (left column) and testing (right column)



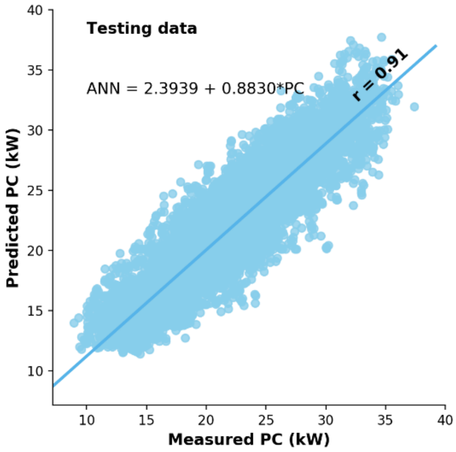
(a)



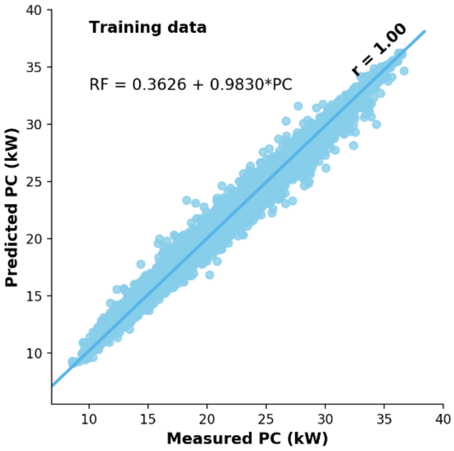
(b)



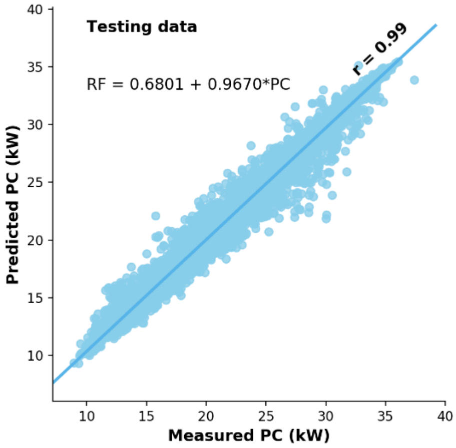
(c)



(d)

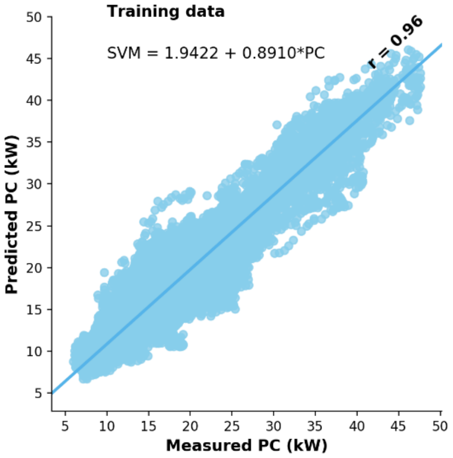


(e)

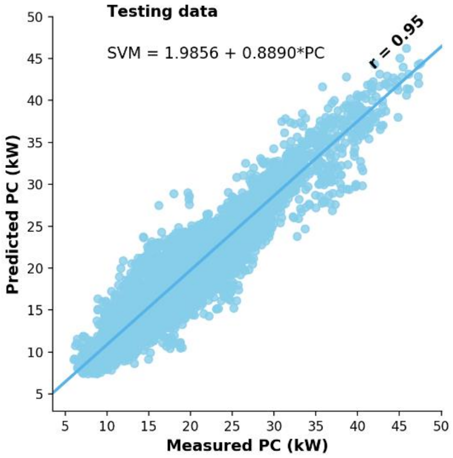


(f)

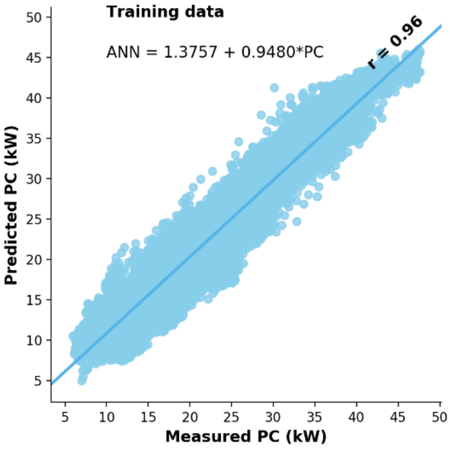
Figure 6. Scatter-plots performance of the Zone-2 PC prediction by the SVM((a)-(b)), ANN((c)-(d)), and RF ((e)-(f)) for training (left column) and testing (right column)



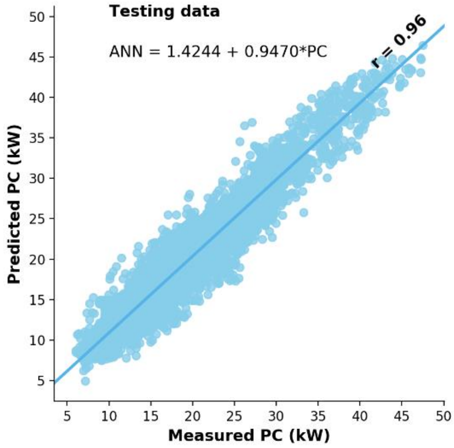
(a)



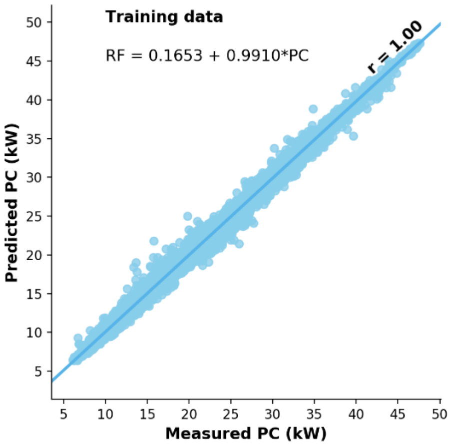
(b)



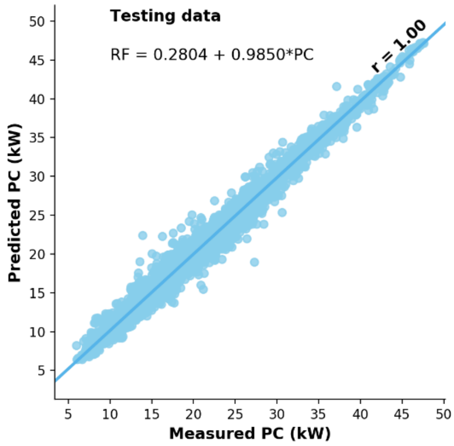
(c)



(d)



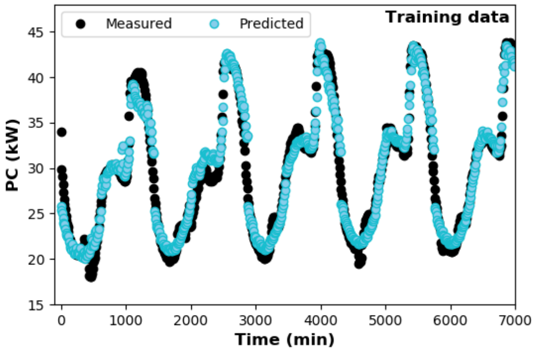
(e)



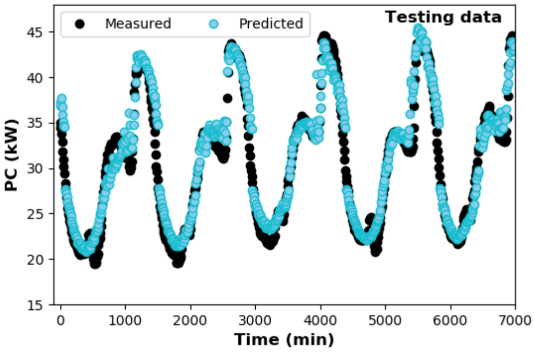
(f)

Figure 7. Scatter-plots performance of the Zone-3 PC prediction by the SVM((a)-(b)), ANN((c)-(d)), and RF ((e)-(f)) for training (left column) and testing (right column)

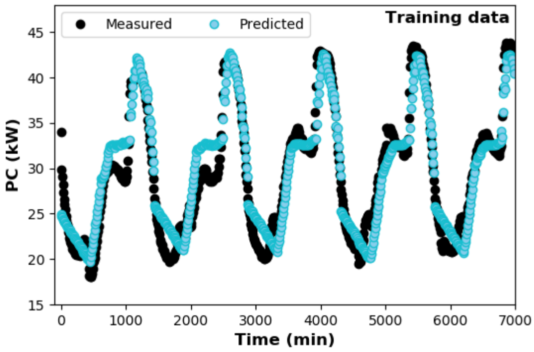
Figures 8, 9 and 10 illustrate the measured and predicted PC by the models over the time for the zones 1, 2 and 3 during both the training and testing phases. Results show an excellent PC output is achieved by the RF model compared to the other models.

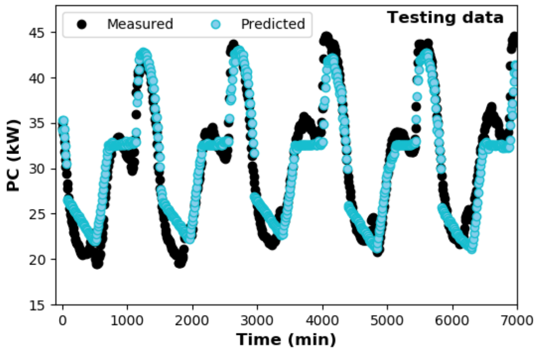


(a)

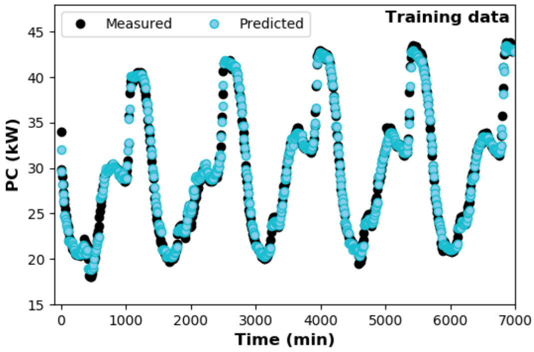


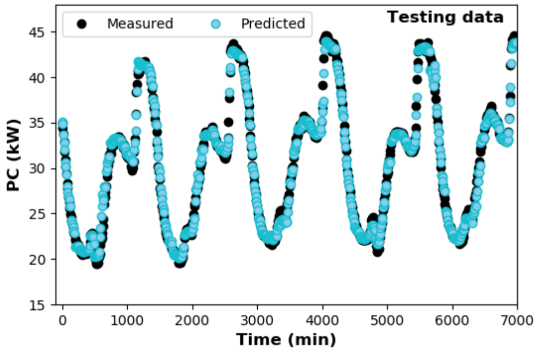
(b)





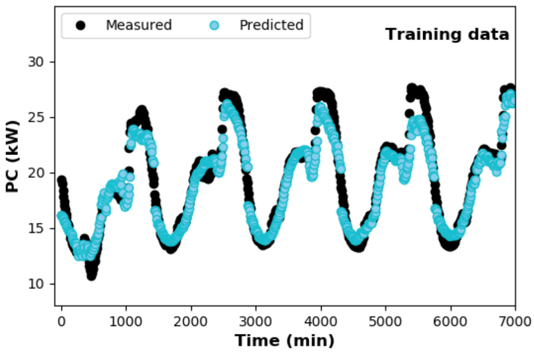
(c) (d)

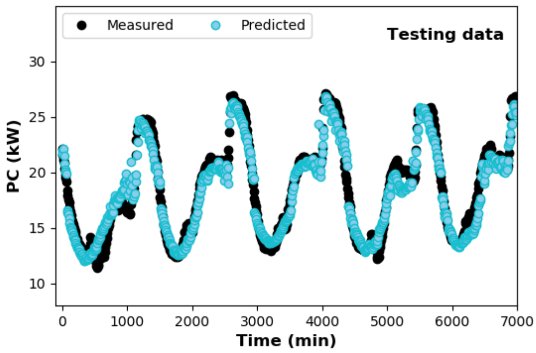




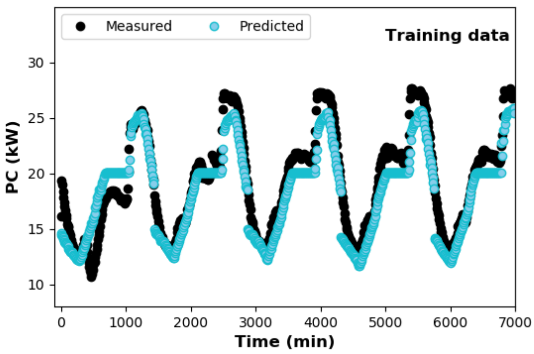
(e) (f)

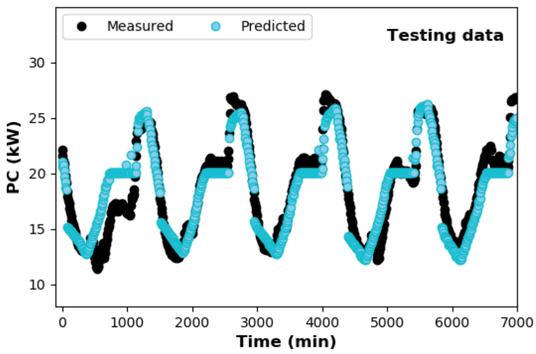
Figure 8. Zone-1 PC prediction over time by the SVM((a)-(b)), ANN((c)-(d)), and RF ((e)-(f)) for training (left column) and testing (right column)



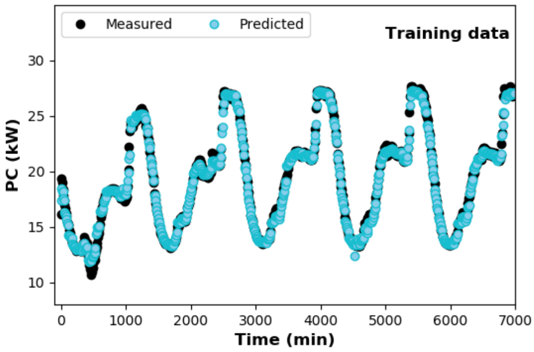


(a) (b)

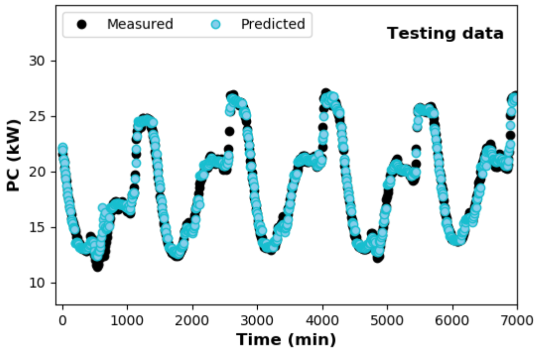




(c) (d)

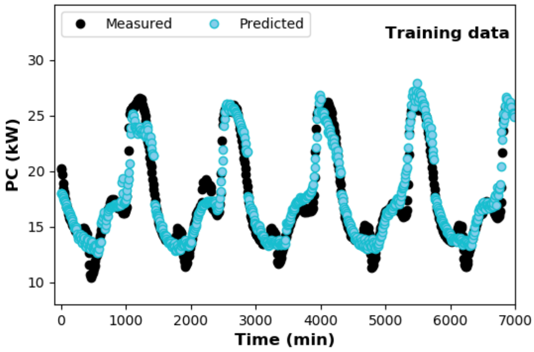


(e)

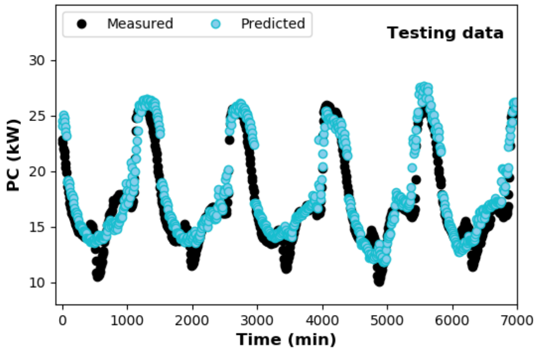


(f)

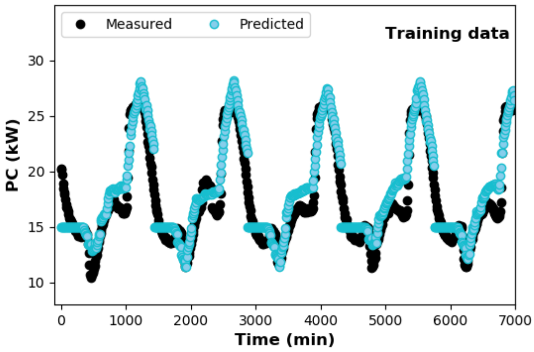
Figure 9. Zone-2 PC prediction over time by the SVM((a)-(b)), ANN((c)-(d)), and RF ((e)-(f)) for training (left column) and testing (right column)



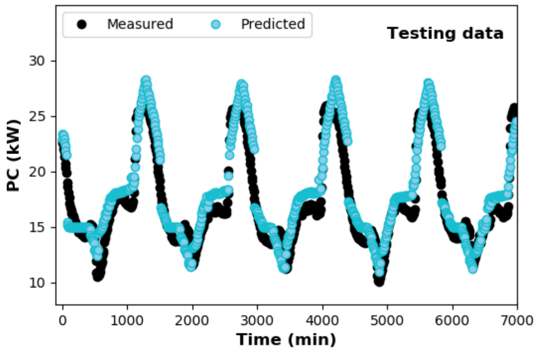
(a)



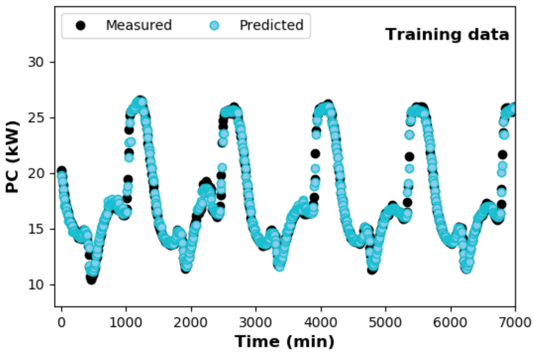
(b)



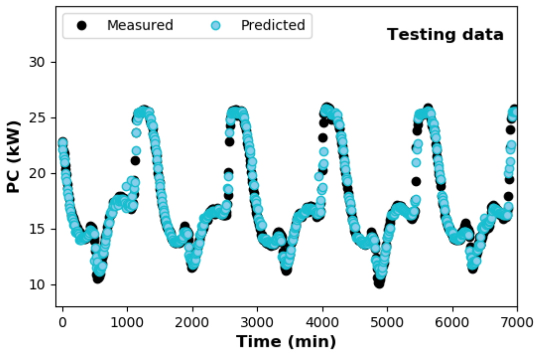
(c)



(d)



(e)



(f)

Figure 10. Zone-3 PC prediction over time by the SVM((a)-(b)), ANN((c)-(d)), and RF ((e)-(f)) for training (left column) and testing (right column)

The evidence provided using performance metrics and prediction plots confirm the RF model's potential to provide robust PC predictions compared to the other used models. The low computational time in the training and testing also confirms the model's suitability for real-time PC prediction.

CONCLUSION AND FUTURE WORKS

This work aimed to find the suitable and powerful model for prediction PC using ML techniques. Three ML models are investigated: SVM, ANN, and RF for PC prediction. Five statistical metrics, MAE, MSE, RMSE, R2, and Pearson's correlation coefficient, are employed due to their suability to evaluate prediction capability of the SVM, ANN and RF models. The results showed that the RF model provided the most reliable and accurate PC prediction followed by SVM and ANN models. The results also showed that the RF required the least computational time than other models during the training and testing phases. From the results, it can be concluded that RF is a robust model, and it has great potential to predict PC. In future, further invesitigation is required to improve RF by applying metaheuristic algorithms to select the most effective inputs for the RF model

References

1. Zhong, H., Wang, J., Jia, H., Mu, Y., & Lv, S. (2019). Vector field-based support vector regression for building energy consumption prediction. Applied Energy, 242, 403-414..
2. Lei, R., & Yin, J. (2022). Prediction method of energy consumption for high building based on LMBP neural network. Energy Reports, 8, 1236-1248..
3. Moon, J., Park, S., Rho, S., & Hwang, E. (2022). Robust building energy consumption forecasting using an online learning approach with R ranger. Journal of Building Engineering, 47, 103851.
4. Paudel, S., Elmitri, M., Couturier, S., Nguyen, P. H., Kamphuis, R., Lacarrière, B., & Le Corre, O. (2017). A relevant data selection method for energy consumption prediction of low energy building based on support vector machine. Energy and Buildings, 138, 240-256..
5. Liu, Y., Chen, H., Zhang, L., Wu, X., & Wang, X. J. (2020). Energy consumption prediction and diagnosis of public buildings based on support vector machine learning: A case study in China. Journal of Cleaner Production, 272, 122542.
6. Wong, S. L., Wan, K. K., & Lam, T. N. (2010). Artificial neural networks for energy analysis of office buildings with daylighting. Applied Energy, 87(2), 551-557.
7. Aydinalp-Koksal, M., & Ugursal, V. I. (2008). Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector. Applied energy, 85(4), 271-296.
8. Azadeh, A. G. H. A. D. E. R. I., Ghaderi, S. F., & Sohrabkhani, S. (2008). Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. Energy Conversion and management, 49(8), 2272-2278.
9. Muralitharan, K., Sakthivel, R., & Vishnuvarthan, R. (2018). Neural network based optimization approach for energy demand prediction in smart grid. Neurocomputing, 273, 199-208..
10. Amber, K. P., Ahmad, R., Aslam, M. W., Kousar, A., Usman, M., & Khan, M. S. (2018). Intelligent techniques for forecasting electricity consumption of buildings. Energy, 157, 886-893.
11. Ahmad, M. W., Mourshed, M., & Rezgui, Y. (2017). Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. Energy and buildings, 147, 77-89.
12. Shao, M., Wang, X., Bu, Z., Chen, X., & Wang, Y. (2020). Prediction of energy consumption in hotel buildings via support vector machines. Sustainable Cities and Society, 57, 102128.
13. Liu, Z., Wu, D., Liu, Y., Han, Z., Lun, L., Gao, J., ... & Cao, G. (2019). Accuracy analyses and model comparison of machine learning adopted in building energy consumption prediction. Energy Exploration & Exploitation, 37(4), 1426-1451.
14. Ma, Z., Ye, C., Li, H., & Ma, W. (2018). Applying support vector machines to predict building energy consumption in China. Energy Procedia, 152, 780-786.
15. Xu, L., Hou, L., Zhu, Z., Li, Y., Liu, J., Lei, T., & Wu, X. (2021). Mid-term prediction of electrical energy consumption for crude oil pipelines using a hybrid algorithm of support vector machine and genetic algorithm. Energy, 222, 119955.
16. Vinagre, E., Pinto, T., Ramos, S., Vale, Z., & Corchado, J. M. (2016, September). Electrical energy consumption forecast using support vector machines. In 2016 27th International Workshop on Database and Expert Systems Applications (DEXA) (pp. 171-175). IEEE.
17. Salam, A., & El Hibaoui, A. (2018, December). Comparison of Machine Learning Algorithms for the Power Consumption Prediction:-Case Study of Tetouan city–. In 2018 6th International Renewable and Sustainable Energy Conference (IRSEC) (pp. 1-5). IEEE.

18. Moon, J., Kim, Y., Son, M., & Hwang, E. (2018). Hybrid short-term load forecasting scheme using random forest and multilayer perceptron. Energies, 11(12), 3283.
19. Johannesen, N. J., Kolhe, M., & Goodwin, M. (2019). Relative evaluation of regression tools for urban area electrical energy demand forecasting. Journal of cleaner production, 218, 555-564.
20. UCI Machine Learning Repositor, obtained from: UCI Machine Learning Repository: Power consumption of Tetouan city Data Set . Accessed on 10th April 2022.
21. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.
22. Hinton, G. E. (1992). How neural networks learn from experience. Scientific American, 267(3), 144-151.
23. Jiang, R., Tang, W., Wu, X., & Fu, W. (2009). A random forest approach to the detection of epistatic interactions in case-control studies. BMC bioinformatics, 10(1), 1-12.
24. Genuer, R., Poggi, J. M., & Tuleau-Malot, C. (2010). Variable selection using random forests. Pattern recognition letters, 31(14), 2225-2236.
25. Ngo, N. T., Truong, T. T. H., Truong, N. S., Pham, A. D., Huynh, N. T., Pham, T. M., & Pham, V. H. S. (2022). Proposing a hybrid metaheuristic optimization algorithm and machine learning model for energy use forecast in non-residential buildings. Scientific Reports, 12(1), 1-18.