Random Forest-Based Method for Micro Grid System in Energy Consumption Prediction

Murugananth Gopal Raj1*, C Pradip2, N Saju2, S V Tresa Sangeetha3
1Department of Electrical and Electronics Engineering, Ahalia School of Engineering and Technology, Kozhippara, Kerala, India
2Department of Electrical and Electronics Engineering, NSS College of Engineering, Palakkad, Kerala, India
3Department of Electrical and Electronics Engineering, University of Technology and Applied Science-Al Mudabah, Muladdah, Musanna, Sultanate of Oman
Email: *gmurugananth@gmail.com, #tresa@act.edu.om

Abstract. As a solution to mitigating rising energy needs, microgrids (MG) have arisen. But instead of microgrids are focused mainly on unconventional sources of energy. In their service, there is significant variability. Energy users will not know if their estimated load is long or short related to historical records. This paper aims to formulate a robust energy prediction of consumption in the microgrid system that uses random forest (RF) method theory as the mathematical framework. Effective MG energy forecast plays an essential role in power improvement MG efficacy. Comparing RF models with various parameter configurations and examining the parameters setting affects the model's estimation efficiency.

Keywords: MG, prediction of energy consumption, RF method.

1. Introduction
Electricity consumption prediction is now one of the new phenomena thanks to improvement in energy, power and control systems. For power controllers, the estimation of load demand is a daunting task. It is an old trend of research which preceded the growth of power plants and computer techniques [1-4]. At the beginning of the 21st century, clean natural resources and smart grids are now evolving. It is increasingly relevant.

Due to two major factors, load forecasts become tougher than ever. Firstly, in many countries, privatization and restructuring of power markets ensure that energy users can select their supplier from different operators. The second is that the strong penetration of unreliable energy enhances the great degree of instability due to their irregular actions into the grid, wind and solar. Market reform ensures that the power price varies, forcing consumers to buy when energy prices are low, thus requiring new prediction systems [5-6]. Short-term energy load projections use the most relevant historical evidence, from a few hours to a few weeks before the predicted day. Lately, research experiments used advance machine learning in short-term prediction models. The clean energy of the economy's manufacturing and residential areas is a significant economic growth guideline [7].

The traditional approach is to categories the predicted benefit into low, medium, and high stages, which customers can interpret. This research suggests an alternate approach used when predicted levels are appropriate for a predictive control framework rather than real values. In recent years power sector reform and renewable energy introduction have made the electricity load forecast more and more relevant. Advanced intelligent methods are built to address the coming difficulties and ensure
precise power predictions for various times [8-10]. The use of smart prediction algorithms is an effective method to deal with uncertainty and a core aspect of smart grids.

Load forecasts mean the projection of potential trends in a single apartment, local grid, an area, or even a whole country's electricity. This prediction is made over one or more time stages over duration called the forecast horizon. For load shipments, unit participation, and energy sharing decisions, a specific power forecast of one or multiple hours is of great significance [11-13]. For maintenance planning and energy efficiency strategies, forecasting loads for a prolonged period of time is beneficial [11-15].

This paper suggests a short-term forecast model based on the random forestry strategy to overcome the above challenges. This model is intended for a one-hour regular forecast in support of the Tunisian power system requirements such as a limited area, high humidity, no deregulation or irregular energy presence.

2. Methods

This section explains, discussed concepts and terminology which are used in every part of this paper. Throughout the area of computer engineering, energy use has been extensively researched for decades. Although energy is being used as a metric in machine education, most research still focuses on achieving high accuracy without any computational constraints. We present a review of the different approaches in general and machine learning applications to estimate energy consumption in this challenge. We also present state-of-the-art technical methods to approximate energy values along with two case studies that improve computer education energy usage.

The reliability improvement of the electricity is to be ensured from the consumer's viewpoint so that peak demand is compensated by the design of generation capacity. The reduction in the load profile of peak to average ratio is achieved through this design to reduce the more power plants requirement. This is obtained by two methods that smart pricing technique and optimal load control. Smart pricing technique is applied by utility side to manage the peak loads.

3. Random Forest

The Random Forest (RF) is a group method which combines several decision trees with prediction. The key theory is bagging; if a sample n is randomly picked from the training range, Sn is fitted to a backbone. This sample is known as bootstrap and is used to replace the sample.

A bootstrap analysis is collected by randomly choosing previous samples with Sn's substitution, and each measurement can pick 1=n. The independently dispersed random variable Hl represents this random collection. It selects multiple samples from the bagging process.

RF simple steps are approximately separated into four steps:
Step 1: Collecting n observations from a bootstrap sampling process;
Step 2: Choosing the c attributes of each entry feature by randomized selection and determining the optimal segmentation characteristics as a branch to create a regression tree;
Step 3: Repeated steps 1 and 2 for a maximum of m times to construct the m regress tree.
Step 4: Combining the outcomes of m trees for the final estimation function.

Random forest (RF) is a modified bagging that produces a large collection of independent trees and averages their results. RF construction depends on two hyperparameters: the number of forest trees and the number of characteristics to be considered when determining the right fraction. Figure 1 shows the tree structure in a random forest method.

A fresh bootstrapped DB dataset is created from training set D for each incremental tree. As the amount of tree in the forest grows, trees are more likely to collide. However, the benefit is that more ballots are cast in the training algorithm, minimizing the common error.

The number of trees has been shown to rise, and the theoretical threshold of the forest is reaching accuracy. The number of characteristics considered for each division influences the variety of trees. By taking into account all the characteristics at each break, each tree chooses its optimal global function, allowing similar trees. Reducing the number of features considered on each break raises the
probability that the best global function is omitted from the test subset of features. The goal is to build a combination of decision-making trees in different ways, contributing to various predictions. Figure 1 represents method of random forest.

Each of the trees produced by bagging is distributed equally, making it impossible to change other than to minimize variances. RF conducts a tree developmental process by randomly choosing an input variable to improve bagging without undue variation by reducing the association between the trees. The model is built according to an online learner using the RF characterized by resistance to parameter changes and intrinsic cross-validation.

4. Results
This removed anomalies from the data in the pre-processing layer. The knowledge is considered noise based on the intrinsic nature of the data collection, which influences interpretation by various external aspects. The results of the RF-based load scheduling and the energy consumption is predicted. The obtained results are verified that shows the distribution of energy consumption based on the hour, as illustrated in the figure. 2. The figure. 3 One-month energy consumption of Actual method versus RF. The Actual versus RF predicted results for one-year energy consumption is shown in the figure. 4.
5. Conclusion
A random forest classifier established was used to predict the level of energy. In the previous study, both approaches were based on a random forest optimization (RF), which underperformed other commonly used machine learning (ML) techniques. Similar preparation and assessment parameters were used in both systems. But as the number of their target levels increases, their estimates tend to decline, and the precision of the standard method is reached.

References
[1]. Ayón, X., Gruber, J. K., Hayes, B. P., Usaola, J and Prodanović, M., “An optimal day-ahead load scheduling approach based on the flexibility of aggregate demands”, Applied Energy, Vol. 198, Pp.1-11, 2017.
[2]. Liu, E., & Cheng, P. “Achieving privacy protection using distributed load scheduling: A randomised approach. IEEE Transactions on Smart Grid, Vol. 8, No.5, 2460-2473, 2017.
[3]. Rajkumar, S., Sundaram, C. S., Sethuraman, K and Muruganandhan, D., “Performance Analysis of Hub BLDC Motor Using Finite Element Analysis”, In 2019 ICSCAN, pp. 1-5, 2019
[4]. Anvari-Moghaddam, A., Guerrero, J. M., Vasquez, J. C., Monsef, H., & Rahimi-Kian, A., “Efficient energy management for a grid-tied residential microgrid”, *IET Generation, Transmission & Distribution*, Vol. 11, No.11, Pp. 2752-2761, 2017.

[5]. Wang, S., Janabi, A., & Wang, B., “Generalised Optimal SVPWM for the Switched-Capacitor Voltage Boost Converter”, ECCE, pp. 2708-2711, 2020.

[6]. Nosratabadi, S. M., Hooshmand, R. A and Gholipour, E., “A comprehensive review on microgrid and virtual power plant concepts employed for distributed energy resources scheduling in power systems”, *Renewable and Sustainable Energy Reviews*, Vol. 67, Pp. 341-363, 2017.

[7]. Ahmad, T., Chen, H., Huang, R., Yabin, G., Wang, J., Shair, J., and Kazim, M., “Supervised based machine learning models for short, medium and long-term energy prediction in distinct building environment”, *Energy*, Vol. 158, Pp.17-32, 2018.

[8]. Montavon, G., Hansen, K., Fazli, S., Rupp, M., Biegler, F., Ziehe, A and Müller, K. R., “Learning invariant representations of molecules for atomisation energy prediction”, *Advances in neural information processing systems*, Vol. 25, Pp. 440-448, 2012.

[9]. Sudharani, S., & Immanuel, D. G., “Mitigation of voltage sag/swell by dynamic voltage restorer using fuzzy-based particle swarm controller”, *Materials Today: Proceedings*, 2021.

[10]. Yadav, D., Parekh, U., Patel, K., Parmar, R and Prakash, M., “Application of Modified Three Phase Conduction Method to Minimise Torque Ripple in BLDC Motor”, *IJRAR*, vol. 6(1), 155-156, 2019.

[11]. Mahusudhana, G. R., Kumar, A and Ram, B. S., “Design of Neuro-fuzzy Inference distributed power flow controller for transient stability improvement. *Word Journal of Modeling and Simulation*, Vol. 10, No.4, Pp. 252-260, 2014.

[12]. Suresh, G., Kumar, N. C. S., Murugesan R and Mukunthan, P., “Multi-object and dynamic query-based CBIR system using DCT incorporated with HOG and HTF”, *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 9(2S3), 2019.

[13]. de Ves, E., Benavent, X., Coma, I and Ayala, G., “A novel dynamic multi-model relevance feedback procedure for content-based image retrieval”, *Neurocomputing*, Vol. 208, Pp.99-107, 2016.

[14]. Joseph, F and Subbiah, M., “Hybrid windowing adaptive FIR filter technique in underwater communication”, *International Journal of MC Square Scientific Research*, Vol. 10, No.2, Pp.17-21, 2018.

[15]. Memon, Li, J.P., Memon, I and Arain, Q.A., “GEO matching regions: multiple regions of interests using content-based image retrieval based on relative locations”, *Multimedia Tools and Applications*, Vol. 76, No.14, pp.15377-15411, 2017.