[Applied Computing Journal](https://doi.org/10.52098/acj.202140) https://doi.org/10.52098/acj.202140
ISSN: 2788-9688, Volume 1, Issue 3, pp 133-140, July 2021
Received: 5/8/2021 Revised: 8/8/2021 Accepted: 8/8/2021

Implementation of Big Data Analytics for Simulating, Predicting and Optimizing the Solar Energy Production

Abstract

The notable developments in renewable energy facilities and resources help reduce the cost of production and increase production capacity. Therefore, developers in renewable energy evaluate the overall performance of the various equipment, methods, and structure and then determine the optimal variables for the design of energy production systems. Variables include equipment characteristics and quality, geographical location, and climatic variables such as solar irradiance, temperature, humidity, dust, etc. This paper investigated and reviewed the current big data methods and tools in solar energy production. It discusses the comprehensive two-stage design and evaluation for examining the optimal structure for renewable energy systems. In the design stage, technical and economic aspects are discussed based on a robust analysis of all input/output variables for determining the highest performance. Next, assess and evaluate the effectiveness of each method under different circumstances conditions. Then convert each qualitative indicator into a quantitative measure using extensive data analysis methods to determine the overall performance of the various qualitative variables. The paper also provides an in-depth analysis of the mathematical techniques used in measuring the efficiency of the renewable energy production system and discussing future axes of work in the field of specific energy.

***Keywords**:* Big Data; machine learning; solar energy; ANN; model optimization;

1. Introduction

Student Big Data has become a significant interest research area for both academic and business organizations. It describes the use of large, hard-to-manage volumes of data. It comprises the analysis of the three V's (Volume, Velocity, and Variety) for insights that lead to better decisions and strategic business moves (Yousif & Saini, 2020). On the other hand, Solar energy is an inherent energy source that uses light and heat to apply photovoltaic and thermal energy systems. Big data techniques help promote solar/ thermal energy generation for efficient and reliable usage of renewable sources. Energy plays a vital role in the economy and environment (Singh et al., 2021). Therefore, the evaluation of energy production, capacity factor, and cost of power is a potential research focus. Figure 1 shows the growth in renewable electricity generation globally.

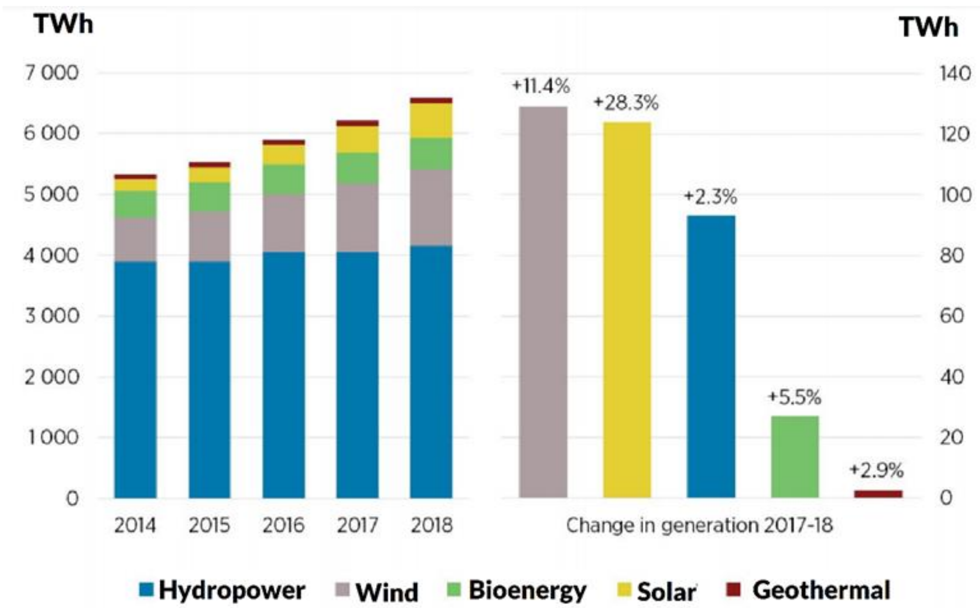

Figure 1. growth in renewable electricity generation globally (IRENA, 2020)

In It shows that the renewable electricity production in 2018 reaches 376 TWh, higher than in 2017, increasing 6.1%. The generation growth of Solar and wind grow strongly by 28% and 11% to become the third-largest source of renewable electricity generation in 2018 (IRENA, 2020).

The transformation of renewable energy technology into the global economy will boost technological innovations to enhance power production and reduce prices. This requires the application of new algorithms and methods to increase the performance of solar energy production (khan I, 2021). Also, to find more efficient systems to store and transfer energy generated from the grid-connected system to sources of consumption quickly and cheaply. Applying big data analysis based on machine learning and artificial intelligence will efficiently help analyze problems and find practical solutions (He et al., 20202). This work aims to present modern technologies and explore practical solutions that help increase solar and thermal energy production performance to guide new researchers.

2. Challenges in Big Data analysis

Specialists working in Big Data analysis specify many challenges in exploring large data sets when extracting information to build relationships between different data types (Ren et al., 2021). Difficulties in Big Data analysis include data capturing, warehouse, exploration, sharing, accounting, management, and visualization. Big Data management aims to ensure the reliability of the data and provide simple ways to access and control the stored data securely. One of the most crucial challenges in dealing with Big Data is collecting data from different sources and integrating them into a standard format (Devaraj et l., 2021). Also, finding the efficient model for storing the generated power to reduce storage size. Another significant challenge is the efficient management of Big Data to accelerate the extraction of information and facilitate the transfer of information at a low cost. The demand for renewable energy is increasing; therefore, it is requisite to recognize the factors that prevent its scalability. Today renewable energy achieves several benefits, but industries will fail without efficient forecasting and intelligent scheduling of the resources (Fan Z, 2021). Hence, examining appropriate data analytics and machine learning techniques can help simulate and predict weather conditions efficiently and improve productivity. Incorporating these technologies will eventually lead to the development of the renewable energy sector, which is an important factor in the economy of every country.

3. Big Data Techniques for Renewable Energy

This work aims to investigate and review the current big data methods and tools in solar energy production. It should focus on the comprehensive two-stage design and evaluation for examining the optimal variables and structure for renewable energy systems (Akbari-Dibavar et al., 2021). In the design stage, technical and economic aspects are discussed based on a robust analysis of all input/output variables for determining the highest performance. Next, assess and evaluate the effectiveness of each method under different circumstances conditions for optimizing the renewable energy systems.

The photovoltaic/thermal (PVT) is a hybrid system that combines solar thermal and electrical power simultaneously (Gül & Akyüz, 2021). The performance of PV systems is influenced significantly by various weather conditions, such as (Solar Irradiance (SR), Temperature (TM), Relative Humidity (RH), Wind Speed (WS), Wind Direction (WD), Sunshine Hours (SH), Evaporation (EVA), Pressure (PRE), Cloud (CL), Cell Temperature (TC), Fluid Temperature (FT), etc.

Several fields of research could be investigated that includes:

- Power Data production and forecasting.
- Resource management.
- Storage management.
- Predicting and optimization of future figures.

Big Data are mostly based on machine learning techniques can include Artificial Neural Network (ANN), Support Vector Machine (SVM), Linear/ Logistic regression (LR) (Yousif & Fekihal, 2012).

These approaches help for examining and enhancing renewable energy production could include the following techniques:

- **Supervised Learning**, for example (Multilayer perceptron (MLP), Generalized Regression Neural Network (GRNN), Learning Vector Quantization (LVQ).
- **Unsupervised Learning**, for example (Recurrent Neural Network (RNN), Self-Organizing Feature Map (SOFM), Principal Component Analysis (PCA).
- **Semi-supervised Learning** (semi-supervised Support Vector Machines (S3VMs)).
- Reinforcement Learning is based on Markov decision process (MDP)
- **Deep Learning** such as convolutional neural networks (CNN), Long Short-Term Memory Networks (LSTM), and Generative Adversarial Networks (GANs).

Table 1 presents and explores several studies that deploys Artificial Neural Network (ANN) for estimating, simulating, predicting, and optimizing the solar energy production.

Table 1. Review of studies that deploys Artificial Neural Network (ANN) for solar energy production.

Reference	Location	ANN Method	Estimation/ Simulation / Prediction / Optimization
(Jabar & Hussein, 2021)	Oman	MLP, SOFM, SVM	Prediction & Optimization
(Zayed et al., 2021)	USA, Arizona	RVFL, PSO, SSO, WOA.	Prediction & Optimization
(Al-Waeli, et al., 2020)	Malaysia	MLP, SOFM	Estimation & Prediction
(Li, Y et al., 2020)	Hong Kong	FFMLP- LM	Simulation & Prediction
(Al-Waeli, et al., 2019a)	Malaysia	MLP, SOFM	Simulation & Prediction
(Kazem et al., 2019a)	Oman	TLRN, FRNN	Simulation & Prediction
(Alnaqi et al., 2019)	Iran/ Kermanshah	FFMLP, PSO	Prediction & Optimization
(Kazem et al., 2019)	Oman	TLRN, FRNN	Simulation & Prediction
(Al-Waeli, et al., 2019b)	Malaysia	MLP	Prediction & Optimization
(Yousif et al., 2019a)	Oman	PMM	Prediction
(Ghimire et al., 2019)	Australia /Queensland	MLP, SVR, GPML, GP	Prediction
(Yousif et al., 2019b)	Oman	Comparative study	Simulation
(Behera et al., 2018)	Latitude: 20º25, Longitude:85º80	SLFN, PSO	Prediction & Optimization
(Ahmad et al., 2018)	France / Chambéry	RF, ET, DT, SVR	Prediction & Optimization
(Al-Waeli, et al., 2018)	Malaysia	MLP, SOFM, SVM	Estimation & Prediction
(Kalani et al., 2017)	Iran	MLP, PBF-NN, ANFIS	Simulation & Prediction
(Yousif et al., 2017)	Oman	MLP, SOFM, SVM, PMM	Simulation & Prediction
(Bassam et al., 2017)	Mexico	ANNFIS	Prediction & Optimization
(Kazem & Yousif, 2017)	Oman	MLP, GFF-NN, SOFM, SVM	Prediction
(Panthee & Jha, 2016)	Nepal/ Kathmandu	FF-NN	Prediction
(Kazem & Yousif, 2016)	Oman	MLP, SVM	Prediction
(Kazem et al., 2016)	Oman	SVM	Estimation & Prediction
(Gunasekar et al., 2015)	India/ Coimbatore	FF-NN	Prediction
(Mohanraj et al., 2015)	India	FF-NN, RBF-NN, GRNN	Prediction
(Ceylan et al., 2014)	Turkey Kutahya, Usak, Afyon	MPL	Prediction
(Al-Shamisi et al., 2014)	UAE Al-Ain	MLP / RBF-NN	Estimation
(Alzahrani et al., 2014)	USA Maries County	NARX	Prediction
(Assi et al., 2013)	UAE Al-Ain	MLP / RBF-NN	Prediction
(Ahmed et al., 2013)	Eygpt, Qena	FF-NN	Estimation
(Hasni et al., 2012)	Algeria/ Bechar	MLP -LM	Estimation
(Premalatha & Arasu, 2012)	India / Tamilnadu	FFMLP	Estimation
(Asl et al., 2011)	Iran Dezful	MLP	Prediction
(Qin et al., 2011)	China/ Tibetan Plateau	MLP-LM	Estimation

Random Vector Functional Link (RVFL); Particle Swarm Optimization (PSO); Spherical Search Optimization (SSO), Whale Optimization Algorithm (WOA); Levenberg-Marquardt algorithm (LM); Polynomial Mathematical Model (PMM); Gaussian Process Machine Learning (GPML); Genetic Programing (GP); Single layer feed-forward network (SLFN); Feedforward Multilayer Perceptron (FFMLP); Recurrent Neural Network (RNN); Full Recurrent Neural Network (FRNN); Nonlinear Autoregressive Network (NARX); Support Vector Machine (SVM); Multilayer Perceptron (MLP); Self-Organizing Feature Map (SOFM); Time Lagged Recurrent Network (TLRN); Random Forest (RF); Extra Trees (ET); Decision Trees (DT); Polynomial basis function (PBF-NN); Adaptive Neuro Fuzzy Inference System (ANFIS); Generalized Feed-Forward Neural Networks (GFF-NN); Generalized Regression Neural Network (GRNN); Radial Basis Function Neural Network (RBF-NN); Feed-Forward Neural Networks (FF-NN);

4. Conclusion

Several essential technologies are helping to create an efficient, productive, and sustainable future globally. Solar energy and big data are two primary methods of these technologies. They will change how energy is purchased, marketed, distributed, and used responsibly. Advanced technologies such as big data and analytics have an enormous effect on every aspect of energy production and today's modern world. The aggregate of big data and solar energy has also supported the creation of other brand-new energy distribution models. Solar panels are affordable enough for some companies and customers to buy outright. In other cases, having access to the correct data can help identify issues where additional agreements are more appropriate for the parties involved. In these ways and more, big data and solar power make an ideal match that will continue paying significant interest. Solar panels let us tap into an almost inexhaustible source of energy. Big data analysis helps us manage and enhance power production and its cost-effectively.

The following are two recommendations:

a) One of the most significant problems of managing electric power generation is the combination of different types of renewable energy (wind, solar cells, thermal units, etc.) and knowing the exact amount of production during different year periods. The increase and decrease in production will generate a storage defect that may exceed the current storage capacity.

- The solution is to find efficient mathematical models that can predict wind and solar energy production, which helps reduce uncertainty about the changing renewable energy. Predictive models will help network operators commit extreme events more efficiently, reducing system balancing costs.

b) Most of the energy production network systems differ depending on the type of tools used, location, and various weather conditions such as (Solar Irradiance (SR), Temperature (TM), Relative Humidity (RH), Wind Speed (WS), Wind Direction (WD), Sunshine Hours (SH), Evaporation (EVA), Pressure (PRE), Cloud (CL), Cell Temperature (TC), Fluid Temperature (FT), etc.

- So, there is a significant need to find optimal solutions to address the integration between these different variables, which is a major branch of research and development.

Acknowledgment

The research leading to these results has received no Research Project Grant Funding.

References

[1]. Ahmad, M. W., Reynolds, J., & Rezgui, Y. (2018). Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. Journal of cleaner production, 203, 810-821.

[2]. Ahmed, E. A., & Adam, M. E. N. (2013). Estimate of global solar radiation by using artificial neural network in Qena, Upper Egypt. Journal of Clean energy technologies, 1(2), 148-150.

[3]. Alnaqi, A. A., Moayedi, H., Shahsavar, A., & Nguyen, T. K. (2019). Prediction of energetic performance of a building integrated photovoltaic/thermal system thorough artificial neural network and hybrid particle swarm optimization models. Energy conversion and management, 183, 137-148.

[4]. Al-Shamisi, M., Assi, A., & Hejase, H. (2014). Estimation of global solar radiation using artificial neural networks in Abu Dhabi City, United Arab Emirates. Journal of solar energy engineering, 136(2).

[5]. Al-Waeli, A. H., Kazem, H. A., Yousif, J. H., Chaichan, M. T., & Sopian, K. (2019). Mathematical and neural network models for predicting the electrical performance of a PV/T system. International Journal of Energy Research, 43(14), 8100-8117.

[6]. Al-Waeli, A. H., Kazem, H. A., Yousif, J. H., Chaichan, M. T., & Sopian, K. (2020). Mathematical and neural network modeling for predicting and analyzing of nanofluid-nano PCM photovoltaic thermal systems performance. Renewable Energy, 145, 963-980.

[7]. Al-Waeli, A. H., Sopian, K., Kazem, H. A., Yousif, J. H., Chaichan, M. T., Ibrahim, A., ... & Ruslan, M. H. (2018). Comparison of prediction methods of PV/T nanofluid and nano-PCM system using a measured dataset and artificial neural network. Solar Energy, 162, 378-396.

[8]. Al-Waeli, A. H., Sopian, K., Yousif, J. H., Kazem, H. A., Boland, J., & Chaichan, M. T. (2019). Artificial neural network modeling and analysis of photovoltaic/thermal system based on the experimental study. Energy Conversion and Management, 186, 368-379.

[9]. Alzahrani, A., Kimball, J. W., & Dagli, C. (2014). Predicting solar irradiance using time series neural networks. Procedia Computer Science, 36, 623-628.

[10]. Akbari-Dibavar, A., Tabar, V. S., Zadeh, S. G., & Nourollahi, R. (2021). Two-stage robust energy management of a hybrid charging station integrated with the photovoltaic system. International Journal of Hydrogen Energy, 46(24), 12701-12714.

[11]. Asl, S. F. Z., Karami, A., Ashari, G., Behrang, A., Assareh, A., & Hedayat, N. (2011). Daily global solar radiation modeling using multi-layer perceptron (MLP) neural networks. World Academy of Science, Engineering and Technology, 79, 740-742.

[12]. Assi, A. H., Al-Shamisi, M. H., Hejase, H. A., & Haddad, A. (2013, October). Prediction of global solar radiation in UAE using artificial neural networks. In 2013 International Conference on Renewable Energy Research and Applications (ICRERA) (pp. 196-200). IEEE.

[13]. Bassam, A., May Tzuc, O., Escalante Soberanis, M., Ricalde, L. J., & Cruz, B. (2017). Temperature estimation for photovoltaic array using an adaptive neuro fuzzy inference system. Sustainability, 9(8), 1399.

[14]. Behera, M. K., Majumder, I., & Nayak, N. (2018). Solar photovoltaic power forecasting using optimized modified extreme learning machine technique. Engineering Science and Technology, an International Journal, 21(3), 428-438.

[15]. Ceylan, İ., Erkaymaz, O., Gedik, E., & Gürel, A. E. (2014). The prediction of photovoltaic module temperature with artificial neural networks. Case Studies in Thermal Engineering, 3, 11-20.

[16]. Devaraj, J., Madurai Elavarasan, R., Shafiullah, G. M., Jamal, T., & Khan, I. (2021). A holistic review on energy forecasting using big data and deep learning models. International Journal of Energy Research.

[17]. Fan, Z. (2021). Fault detection of energy-aware grid systems in big data environment. International Journal of Communication Systems, e4755.

[18]. Ghimire, S., Deo, R. C., Downs, N. J., & Raj, N. (2019). Global solar radiation prediction by ANN integrated with European Centre for medium range weather forecast fields in solar rich cities of Queensland Australia. Journal of cleaner production, 216, 288-310.

[19]. Gül, M., & Akyüz, E. (2021). Energy and exergy analysis of solar photovoltaic thermal system: experimental and numerical verification. International Journal of Exergy, 34(1), 16-28.

[20]. Gunasekar, N., Mohanraj, M., & Velmurugan, V. (2015). Artificial neural network modeling of a photovoltaic-thermal evaporator of solar assisted heat pumps. Energy, 93, 908-922.

[21]. Hasni, A., Sehli, A., Draoui, B., Bassou, A., & Amieur, B. (2012). Estimating global solar radiation using artificial neural network and climate data in the south-western region of Algeria. Energy Procedia, 18, 531-537.

[22]. He, Y., Henze, J., & Sick, B. (2020). Continuous Learning of Deep Neural Networks to Improve Forecasts for Regional Energy Markets. IFAC-PapersOnLine, 53(2), 12175-12182.

[23]. IRENA, 2020. [Accessed 1/8/2021] https://www.irena.org/Statistics

[24]. Kalani, H., Sardarabadi, M., & Passandideh-Fard, M. (2017). Using artificial neural network models and particle swarm optimization for manner prediction of a photovoltaic thermal nanofluid based collector. Applied Thermal Engineering, 113, 1170-1177.

[25]. Kazem, H. A., & Yousif, J. H. (2017). Comparison of prediction methods of photovoltaic power system production using a measured dataset. Energy Conversion and Management, 148, 1070-1081.

[26]. Kazem, H. A., Chaichan, M. T., & Yousif, J. H. (2019). Evaluation of oscillatory flow Photovoltaic/Thermal system in Oman. International Journal of Computation and Applied Sciences IJOCAAS, VI (1), 429-436.

[27]. Kazem, H. A., Yousif, J. H., & Chaichan, M. T. (2016). Modeling of daily solar energy system prediction using support vector machine for Oman. International Journal of Applied Engineering Research, 11(20), 10166-10172.
[28]. Kazem, H. A., Yousif, J., Chaichan, M. T., & Al-Waeli, A. H. (2019). Experimental and deep learning artificial neural network approach for evaluating grid-connected photovoltaic systems. International Journal of Energy Research, 43(14), 8572-8591.
[29]. Khan, I. (2020). Impacts of energy decentralization viewed through the lens of the energy cultures framework: Solar home systems in the developing economies. Renewable and Sustainable Energy Reviews, 119, 109576.
[30]. Li, Y., Nord, N., Zhang, N., & Zhou, C. (2020). An ANN-based optimization approach of building energy systems: Case study of swimming pool. Journal of Cleaner Production, 277, 124029.
[31]. Mohanraj, M., Jayaraj, S., & Muraleedharan, C. (2015). Applications of artificial neural networks for thermal analysis of heat exchangers–a review. International Journal of Thermal Sciences, 90, 150-172.
[32]. Panthee, K., & Jha, A. K. (2016). Estimation of global solar radiation using Artificial Neural Network in Kathmandu, Nepal. International Journal of Engineering Research and Science, 2(1), 62-68.
[33]. Premalatha, N., & Arasu, A. V. (2012). Estimation of global solar radiation in India using artificial neural network. Int. J. Eng. Sci. Adv. Technol, 2(6), 1715-1721.
[34]. Qin, J., Chen, Z., Yang, K., Liang, S., & Tang, W. (2011). Estimation of monthly-mean daily global solar radiation based on MODIS and TRMM products. Applied energy, 88(7), 2480-2489.
[35]. Ren, D., Guo, X., & Li, C. (2021). Research on big data analysis model of multi energy power generation considering pollutant emission—Empirical analysis from Shanxi Province. Journal of Cleaner Production, 316, 128154.
[36]. Singh, A. K., Kumar, P., Srivastava, B. K., Yadav, R. K., Singh, M., Kumar, A., ... & Prasad, R. (2021). Comparative observation of energy matrix and environ-economy for solar still systems. Materials Today: Proceedings.
[37]. Yousif, J. H., Al-Balushi, H. A., Kazem, H. A., & Chaichan, M. T. (2019). Analysis and forecasting of weather conditions in Oman for renewable energy applications. Case Studies in Thermal Engineering, 13, 100355.
[38]. Yousif, J. H., & Fekihal, M. A. (2012). Neural approach for determining mental health problems. Journal of Computing, 4(1), 6-11.
[39]. Yousif, J. H., & Kazem, H. A. (2016). Modeling of daily solar energy system prediction using soft computing methods for Oman. Research Journal of Applied Sciences, Engineering and Technology, 13(3), 237-244.
[40]. Yousif, J. H., & Kazem, H. A. (2021). Prediction and evaluation of photovoltaic-thermal energy systems production using artificial neural network and experimental dataset. Case Studies in Thermal Engineering, 101297.
[41]. Yousif, J. H., Kazem, H. A., & Boland, J. (2017). Predictive models for photovoltaic electricity production in hot weather conditions. Energies, 10(7), 971.
[42]. Yousif, J. H., Kazem, H. A., Alattar, N. N., & Elhassan, I. I. (2019). A comparison study based on artificial neural network for assessing PV/T solar energy production. Case Studies in Thermal Engineering, 13, 100407.
[43]. Yousif, J. H., & Saini, D. K. (2020). Big Data Analysis on Smart Tools and Techniques. In Cyber Defense Mechanisms (pp. 111-130). CRC Press.
[44]. Zayed, M. E., Zhao, J., Li, W., Elsheikh, A. H., Abd Elaziz, M., Yousri, D., ... & Mingxi, Z. (2021). Predicting the performance of solar dish Stirling power plant using a hybrid random vector functional link/chimp optimization model. Solar Energy, 222, 1-17.