Discourse on Data Mining Applications to Design Renewable Energy Systems

Atika Qazi, H. Fayaz, and Ram Gopal Raj

***Abstract*—**Sun is the most powerful source of energy. Sunlight, or solar energy, can be used for heating, lighting and cooling buildings, generating electricity, water heating, and a variety of industrial processes. It is one of the most important clean, renewable energy resources, which comes directly from the sun in the form of radiation. For the development of successful solar systems there is immense need of solar radiation prediction. In recent years data mining and related applications are excessively used in predicting solar radiation and solar energy system design. The aim of this work is to give an overview of such predictive data mining techniques. This paper also highlights the importance of solar energy systems in terms of clean environment.

***Keywords*—**Artificial neural networks (ANN), Fossil fuels, Renewable energy, Solar systems.

I. Introduction

RENEWABLE energy resources, such as wind, solar and hydropower, offer clean alternatives to fossil fuels. Mostly renewable energy resources come either directly or indirectly from the sun. For example, wind energy is produced from wind blow that produced from sub heat, participates to the growth of plants that are used for biomass energy, and plays a key role in evaporation cycle and precipitation. Sun light is freely available useful resource of renewable energy that reduces operating costs, greenhouse gas (GhG) emissions, and other pollutants as well.

The marginal economic and environmental benefits associated with additional solar thus depend on the operating characteristics and emission intensities of the units displaced on either the operating or build margin [1]. Solar electricity generation is non-dispatch able as it cannot be turned on and off when needed but works precisely in the presence of sun light, consequently it demands of solar radiation prediction.

Due to the strong increase of solar power generation, the predictions of incoming solar energy are acquiring more importance. It is necessary to predict the amount of energy which will be produced, up to 72 h before, and deviations of energy production are strongly penalized [2]. Recent studies of solar radiation have shown that instantaneous solar radiation exhibits a distinct bimodal character associated with clear and cloudy states. This suggests that many solar systems may simply be modeled to operate in an on/off fashion corresponding to clear/cloudy time intervals. It is shown that the average-daily solar system performance may be calculated from the product of clear-sky solar performance and the average time fraction of clear sky [3]. Therefore it is indeed important to predict the solar radiations in different time intervals by applying such techniques that gives maximum prediction accuracy.

Atika Qazi is with Faculty of Computer Science and Information Technology, University of Malaya, Lembah Pantai, 50603 Kuala Lumpur, Malaysia. (Corresponding author e-mail: atika@siswa.um.edu.my).

Fayaz Hussain is with UM Power Energy Dedicated Advanced Centre (UMPEDAC), Level 4, Wisma R&D, University of Malaya, Jalan Pantai Baharu, 59990 Kuala Lumpur.

Ram Gopal Raj is with the Faculty of Computer Science and Information Technology, University of Malaya, Lembah Pantai, 50603 Kuala Lumpur, Malaysia.

II. Back Ground

Influential facilities of solar systems reduced the environmental impacts of combustion used in fossil fuel power generation, such as impacts from greenhouse gases and other air pollution emissions. Unlike fossil fuel power generating facilities, solar facilities have very low air emissions of air pollutants such as sulfur dioxide, nitrogen oxides, carbon monoxide, volatile organic compounds, and the greenhouse gas carbon dioxide during operations. To get the maximum benefits of solar systems, accurate prediction of solar irradiance is required. For this many applications are introduced for predication of solar radiation; however data mining applications are considered widely.

Data mining is the process of discovering the hidden patterns and analyzing the data from diverse aspects and summarizing it into practical knowledge. It can be used to enhance revenues, decrease cost or sometimes both. The use of data mining in prediction and manufacturing began in the 1990's [4] . An ID3 algorithm was generalized by Irani et.al [5] , where under diverse and overall conditions outcome of the future experiments is predicted. It is a process in the database that is used to find out and reveal the previously unknown, concealed, significant ad useful patterns [6], [7].

Prediction is the eventual aim of predictive data mining. The predictive data mining is the usually used in business applications to predict some response of interest through a statistical or artificial neural network (ANNs) model or set of models for forecasts. These predictive techniques include: Bagging (Voting, Averaging), Boosting, Stacking (Stacked Generalizations), and Meta-Learning. Artificial neural networks (ANNs) have arisen as advanced data mining tools in cases where other techniques may not produce acceptable predictive results[8]. As the term implies, neural networks

have a biologically inspired modeling capability, but are essentially statistical modeling tool [9]. Data Mining is frequently regarded as combination of statistics, AI (artificial intelligence), and data base research.

Artificial neural network were originated from studying and understanding the behavior of complicated neural network of living creatures of human brain. Using ANN techniques has become visible in for modeling and estimation global solar radiation data. Many researcher like [10],[11], [12],[13] used this technique with different parameters for input like: longitude ,latitude, number of day, temperature, Altitude, humidity, mean diffuse radiation, mean beam radiation, month, humidity, relative humidity ,sun shine hours and wind speed, long wave length, rainfall. However, neural networks are good at fitting functions. In fact, there is proof that a fairly simple neural network can fit any practical function.

III. Data Mining For Solar Energy Systmes

There are mainly two types of solar energy systems used such as (1) thermal and (2) electrical. Solar thermal systems are often used for space heating, space cooling, process heat generation and water heating. Solar power is converted into electricity by two ways (1) directly through photovoltaic (PV) and (2) indirectly by concentrated solar power (CSP). Because of the inconsistent incident solar energy the performance of the solar systems is not constant. Therefore the prediction of solar systems becomes necessary for electrical and thermal loads. There are many methods to predict such systems but recently artificial neural networks (ANNs) have achieved extensive attention. The ANN techniques are mostly embedded in many solar systems such as, hot water generation, fault diagnostic system, solar-assisted air-conditioning system and refrigeration system for ice production etc. The proposed ANN modeling is used for generating hot water from solar energy.

The system has proved 40% efficiency by using solar radiation according to experiment. Additionally,18kW maximum power was supplied to the system at noon and 6kW minimum in the afternoon [14].Soteris Kalogirou et al. [15] design a fault diagnostic system (FDS) using ANN for solar water heater (SWH) that consist of prediction module, a residual calculator and the diagnosis module. The system can predict collector faults and faults in insulation of the pipes. The various input values for faults of the system are used to validate the system. S. Rosieka, and F.J. Batlles [16] have proposed Artificial Neural Networks (ANN) to model a solar-assisted air-conditioning system fitted based on absorption chiller and provided only with solar collectors. By using ANN model the results are accurate at satisfactory level. M. Laidi and S. Hanini [17] developed a solar intermittent refrigeration system for ice production working with Activated carbon (AC)/methanol pair .The prediction using the ANN achieved a very small error and the proposed interface is easy to be used.

IV. Growth Rate Of Ann

To see the growth of using ANN over the time; we prepared an informal survey through"www.sciencedirct.com" which is the famous research engine for scientific paper. Two keywords have been used, that are"ANN" and "solar radiation". However, we limited the search for just engineering journal paper and for the period 2001 to 2013.It has been shown, that there are 633 papers were written using these concepts. Figure 1 shows the ranking of publications through the selected period which shows increase in publication over time.

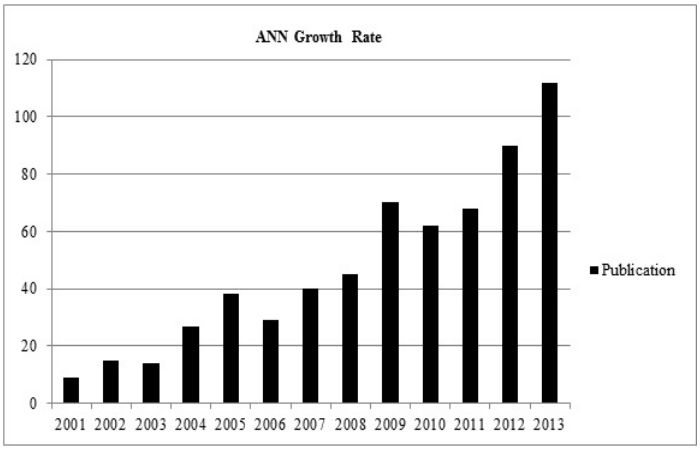

Fig. 1 Artificial Neural Network usage for solar radiation prediction

The renewable energy resources prediction used prediction error. The prediction error is measured by different metrics presented in table below.

TABLE I
Nomenclature

Abbreviation	Description
BP	Back propagation
MAPE	Mean absolute percentage error
MBE	Mean bias error
R	Correlation coefficient
RMSE	Root mean square error
NRMSE	Normalized root mean square error
Lat	Latitude
Lon	Longitude
Alt	Altitude
Mth	Month
Avg	Averag
WS	Wind speed
LW	Long wave length
RH	Relative humidity
AR	Air temperature
TEMP	Temperature
MOY	Month of year/time
ATM	Atmospheric pressure
DOM	Day of month
SCG	Scale conjugate gradient
RP	Resilient propagation

V. Fossil Fuels And Environmental Pollution

Fossil fuels are formed by natural processes such as anaerobic decomposition of buried dead organisms. The age of the organisms and their resulting fossil fuels is typically millions of years, and sometimes exceeds 650 million years [18]. Fossil fuels such as coal, crude oil and natural gas are commonly used all over the world to produce

energy. The use of fossil fuels promotes serious environmental concerns. The burning of fossil fuels produces around 21.3 billion tones (21.3gigatonnes) of carbon dioxide (CO2) per year [19]. Carbon dioxide is one of the greenhouse gases that enhances radiation forcing and contributes to global warming, causing the average surface temperature of the Earth to rise in response, which cause major adverse effects. To clean the environment from pollution a global awareness towards generation of renewable energy is therefore required.

To replace the fossil fuels with clean energy renewable resources such as, wind, sun, water and geothermal heat are considered to be sustainable to produced energy without harming environment. These renewable energy resources warrant being the future of power by reducing global warming.

VI. Conclusion

Globally, many events have highlighted the need of environmentally sound energy resources. The solar energy is one these resources that is also a barrier against fossil fuels. Prediction of solar radiation intensity is required for cost-effective PV sizing and intelligent energy systems. It is concluded that ANN-based prediction offers greater accuracy, therefore ANN is much more dependable and demanding in the domain of renewable energy resource predication. The study demonstrates that ANN models predict solar radiation more accurately than statistical, conventional, linear, non-linear and fuzzy logic models. In future work, we aim to carry out in depth research on acceptance of solar systems in Malaysia.

References

[1] E. Baker, M. Fowlie, D. Lemoine, and S. S. Reynolds, "The economics of solar electricity," resource, vol. 5, 2013.

[2] L. Martín, L. F. Zarzalejo, J. Polo, A. Navarro, R. Marchante, and M. Cony, "Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning," Solar Energy, vol. 84, pp. 1772-1781, 10// 2010.

[3] H. Suehrcke and P. G. McCormick, "A performance prediction method for solar energy systems," Solar Energy, vol. 48, pp. 169-175, // 1992.

[4] M. Lee, "The knowledge-based factory," Artificial intelligence in Engineering, vol. 8, pp. 109-125, 1993. http://dx.doi.org/10.1016/0954-1810(93)90021-7

[5] K. B. Irani, J. Cheng, U. M. Fayyad, and Z. Qian, "Applying machine learning to semiconductor manufacturing," iEEE Expert, vol. 8, pp. 41-47, 1993. http://dx.doi.org/10.1109/64.193054

[6] S. Appavu and R. Rajaram, "Knowledge-based system for text classification using ID6NB algorithm," Knowledge-based systems, vol. 22, pp. 1-7, 2009. http://dx.doi.org/10.1016/j.knosys.2008.04.006

[7] S. Appavu, R. Rajaram, M. Muthupandian, G. Athiappan, and K. Kashmeera, "Data mining based intelligent analysis of threatening e-mail," Knowledge-Based Systems, vol. 22, pp. 392-393, 2009. http://dx.doi.org/10.1016/j.knosys.2009.02.002

[8] J. P. Bigus, Data mining with neural networks: solving business problems from application development to decision support: McGraw-Hill, Inc., 1996.

[9] J. Han, M. Kamber, and J. Pei, Data mining: concepts and techniques: Morgan kaufmann, 2006.

[10] A. Sözen and E. Arcakliog˘lu, "Effect of relative humidity on solar potential," Applied energy, vol. 82, pp. 345-367, 2005. http://dx.doi.org/10.1016/j.apenergy.2004.12.001

[11] J. Bosch, G. Lopez, and F. Batlles, "Daily solar irradiation estimation over a mountainous area using artificial neural networks," Renewable Energy, vol. 33, pp. 1622-1628, 2008. http://dx.doi.org/10.1016/j.renene.2007.09.012

[12] S. Alam, S. Kaushik, and S. Garg, "Assessment of diffuse solar energy under general sky condition using artificial neural network," Applied Energy, vol. 86, pp. 554-564, 2009. http://dx.doi.org/10.1016/j.apenergy.2008.09.004

[13] M. Journée and C. Bertrand, "Improving the spatio-temporal distribution of surface solar radiation data by merging ground and satellite measurements," Remote Sensing of Environment, vol. 114, pp. 2692-2704, 2010. http://dx.doi.org/10.1016/j.rse.2010.06.010

[14] C. Cetiner, F. Halici, H. Cacur, and I. Taymaz, "Generating hot water by solar energy and application of neural network," Applied Thermal Engineering, vol. 25, pp. 1337-1348, 6// 2005.

[15] S. Kalogirou, S. Lalot, G. Florides, and B. Desmet, "Development of a neural network-based fault diagnostic system for solar thermal applications," Solar Energy, vol. 82, pp. 164-172, 2// 2008.

[16] S. Rosiek and F. J. Batlles, "Modelling a solar-assisted air-conditioning system installed in CIESOL building using an artificial neural network," Renewable Energy, vol. 35, pp. 2894-2901, 12// 2010.

[17] M. Laidi and S. Hanini, "Optimal solar COP prediction of a solar-assisted adsorption refrigeration system working with activated carbon/methanol as working pairs using direct and inverse artificial neural network," International Journal of Refrigeration, vol. 36, pp. 247-257, 1// 2013.

[18] P. Mann, L. Gahagan, and M. B. Gordon, "Tectonic setting of the world's giant oil and gas fields," 2003.

[19] R. Watson, H. Rodhe, H. Oeschger, and U. Siegenthaler, "Greenhouse gases and aerosols," Climate change: the IPCC scientific assessment, vol. 1, p. 17, 1990