

MULTIVARIATE TIME SERIES ANALYSIS OF SOLAR IRRADIANCE FOR
PHOTOVOLTAIC SYSTEMS: THE HYBRIDIZATION OF NARX AND LSTM
MODELS

OSAMA GAMAL MAHMOUD IBRAHIM MOTIR

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Master of Data Science

Faculty of computing
Universiti Teknologi Malaysia

NOVEMBER 2024

TABLE OF CONTENTS

	TITLE	PAGE
	TABLE OF CONTENTS	ii
	LIST OF FIGURES	iii
CHAPTER 1	INTRODUCTION	1
1.1	Research Background	1
1.2	Problem Statement	4
1.3	Research Goal	5
1.4	Research Objectives	5
1.5	Scope of The Project	5
REFERENCES		7

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
Figure 1.1	Architecture of an MG [1].	2

CHAPTER 1

INTRODUCTION

1.1 Research Background

Degradation of the environment is the most difficult challenge in nations that have been detribalized, and the argument that it is not a concern in countries that have not been industrialized is no longer relevant [2]. The high emissions of greenhouse gases (GHGs) are having an effect on both industrialized and non-industrialized nations all over the world. The rise in energy consumption that is a direct result of the expansion of the human population is the root cause of environmental deterioration. In 2023, the CO₂ emissions record had shown that the reading of CO₂ had reached 37.4 Gt [3] where is considered the highest record. Which led many countries to focusing on replacing the traditional power grid like Coal-fired power station where they use the concept of coal-combustion to generate electricity, to the renewable energy sources which is more sustain and safe on the environment.

For this reason, the development of clean and renewable energy sources has emerged as a crucial approach for the economic and social sustainability of development for every nation on the planet [4]. The idea of a microgrid was proposed by the researchers in order to resolve the conflict that exists between grid generation and distributed generation, as well as to make the most of the benefits that distributed generation offers in terms of the economics, energy, and the environment.

Microgrids, localized power systems that can operate independently or in conjunction with the traditional power grid structure as shown in Figure 1.1, represent a compact yet fully operational grid system, functioning within a confined geographical region [1]. Typically, a Microgrid (MG) is characterized as a combined energy provision system composed of various Distributed Generations (DGs), energy storage components, loads, and devices for monitoring and protection. It functions as an

independent, self-sustaining system that interfaces with the broader grid as a unified, regulated entity [5]. This system is endowed with the ability for self-regulation, protective measures, and overall management, all while satisfying the customer's expectations for reliable power supply and high-quality electricity.

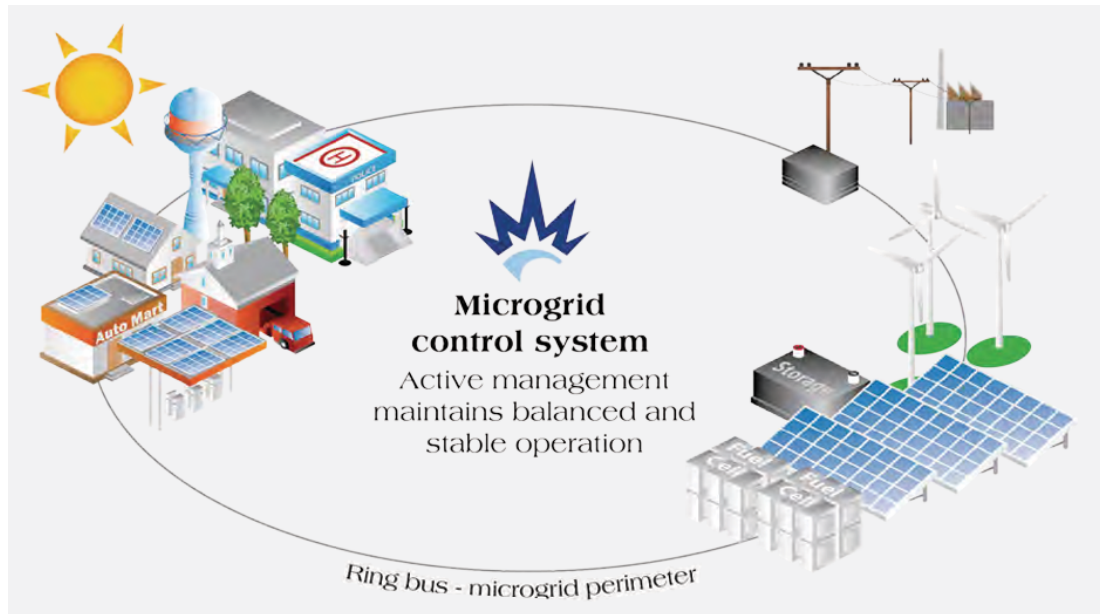


Figure 1.1 Architecture of an MG [1].

The shift to renewable energy is a worldwide movement in the twenty-first century. Solar energy is driving this change because of its huge resource potential. Photovoltaic (PV) energy offers significant availability and extended longevity to stakeholders. As well as there are many technologies that had been developed in field of solar energy in terms of improving the power efficiency (PV cell) also the development reach to the storage system where it can help in improving the isolated system and the hybrid system [6]. However, photovoltaic energy exhibits several constraints about inadequate power stability and subpar power quality. Photovoltaic plants are consistently affected by meteorological factors like cloud cover, wind velocity, and temperature fluctuations. Moreover, the primary impetus for energy production, namely irradiation, is only accessible during daylight hours [7].

Consequently, forecasting models have been extensively used for photovoltaic systems to predict produced photovoltaic power on one hand and load demand on the other, so facilitating intelligent demand response and efficient energy management.

For the purpose of forecasting the solar irradiance, researchers are employing a wide range of methodologies. These methods may be generally classified into four categories: data-driven approaches, image-based approaches, numerical weather prediction (NWP) models, and hybrid approaches [8]. For the purpose of forecasting the solar time series, data-driven or statistical and machine learning methodologies make use of the historical data that has been observed or felt in the past. Images obtained from sky/shadow cameras or satellites are utilised by image-based techniques, which then integrate this information with past data in order to eventually anticipate the irradiance. For the purpose of predicting the irradiance, NWP models make use of computer-based programs. There are a variety of hybrid or combination models, which are diverse combinations of two or more of the types that were originally listed. However many methods facing difficulty to accurately forecast the solar irradiance due to many factors.

Machine learning (ML) approaches have been suggested up until this point in order to comprehend the behavior of selected feature patterns that are constantly changing over time in order to reconstruct a clear vision about future values [9]. In layman's terms, these methods examine the past (the data that was input) in order to make predictions about the future (the behavior of the system). Direct and indirect PVPF are the components that make up Photovoltaic power forecasting (PVPF). On the one hand, direct PVPF evaluates the meteorological data from the past in order to forecast the amount of electricity that may be generated by PV. The indirect PVPF, on the other hand, employs an approach that consists of two stages. When it comes to the initial stage, the solar irradiance is anticipated to be the PV power aspect that is most dependent on it. Within the second step, the determination of the power generated by photovoltaic cells is based on mathematical relationships [10]. According to the findings presented in [11], it has been demonstrated that the second technique is superior in terms of accuracy, versatility, costability, and the amount of computing effort that it requires. According to the authors of [12], hybrid models can obtain competitive outcomes in comparison to the approaches that are considered to be state-

of-the-art by merging two or more single machine learning models from different sources.

According to the approaches of forecasting that are now in use, there are a few handles that continue to require additional mastery. To begin, the range of prediction for weather forecast data is restricted, and the historical time series of solar power is non-stationary, dynamic, and nonperiodic. These characteristics make it impossible for typical artificial intelligence systems to comprehend the data. In the second place, the input-output prediction patterns of the current modes are investigated from the point of view of statistical analysis. This approach either disregards the effect of other connected elements or necessitates the collection of high-quality data from many associated factors, which restricts the practical applicability. In conclusion, the complex non-linearity of the photovoltaic time series over different forecasting horizons exists between univariate time steps and among the variables that are significant.

To address the current challenges and achieve precise photovoltaic power forecasting, the hypothesis of the proposed method in this paper is primarily based on the following considerations where the current research indicates that NARX effectively manages nonlinear relationships, while LSTM excels in extracting temporal features, and the attention mechanism mitigates the issues of distraction. hence, the hybridization of NARX and LSTM is proposed in this study.

1.2 Problem Statement

The total capacity of renewable energy sources around the globe reached 3,381 gigawatts in the year 2023. With a capacity of 1053 gigawatts (GW), solar energy has emerged as the second most widely adopted of these renewable sources for energy production. Malaysia is positioned as top four nations in Southeast Asia in terms of its significant solar energy capacity. In the present moment, renewable energy sources provide 13.3% of Malaysia's overall energy capacity, it is anticipated that Malaysia would attain 18.2% of its capacity to generate electricity from renewable sources by 2025, and 70% by 2050. Consequently, the amount of integration of the

renewable energy with the main grid will increase exponentially which will increase the complexity in terms of grid management, and stability within the grid due to the difficulty in renewable energy forecasting. Therefore, to achieve the integration with RE, an accurate and strong forecasting model is required for predicting the solar energy.

1.3 Research Goal

The aim of the project is to present a hybrid Nonlinear Autoregressive Network with Exogenous Inputs Long Short-Term Memory (NARXLSTM) model to meet the growing need for precise daily solar irradiance predictions, especially under volatile weather conditions in Malaysia.

1.4 Research Objectives

- (a) To develop a hybrid machine learning model combining LSTM and NARX neural networks, focusing on improving the precision of daily solar irradiance predictions.
- (b) To analyze the influence of various weather parameters, determining their individual and collective impact on the forecasting model's performance and.
- (c) To develop a dashboard to display forecasts and analyze trends.

1.5 Scope of The Project

- Python will be used as the main tool to develop and analyze the model evaluation.
- The model performance will be evaluated by using historical weather data that had collected from Johor, Malaysia environments.

- The project will focus on constructing a hybrid model that combines the strengths of Nonlinear Autoregressive with Exogenous Inputs (NARX) and Long Short-Term Memory (LSTM) networks to forecast solar irradiance.
- The model will integrate various weather parameters which are Air Temperature, Cloud Attenuation, Precipitation Rate, Dewpoint Temperature, Surface Pressure, Precipitable Water, Relative Humidity, Wind Speed, Wind Direction

REFERENCES

1. Bank, A. D. *Handbook on Microgrids for Power Quality and Connectivity*. Asian Development Bank Institute. 2020.
2. Yoldaş, Y., Önen, A., Muyeen, S., Vasilakos, A. V. and Alan, I. Enhancing smart grid with microgrids: Challenges and opportunities. *Renewable and Sustainable Energy Reviews*, 2017. 72: 205–214.
3. Raimi, D., Zhu, Y., Newell, R. G. and Prest, B. C. Global Energy Outlook 2024: Peaks or Plateaus, 2024.
4. Suman, G. K., Guerrero, J. M. and Roy, O. P. Optimisation of solar/wind/bio-generator/diesel/battery based microgrids for rural areas: A PSO-GWO approach. *Sustainable cities and society*, 2021. 67: 102723.
5. Sepasi, S., Talichet, C. and Pramanik, A. S. Power quality in microgrids: A critical review of fundamentals, standards, and case studies. *IEEE Access*, 2023. 11: 108493–108531.
6. Dutta, S., Li, Y., Venkataraman, A., Costa, L. M., Jiang, T., Plana, R., Tordjman, P., Choo, F. H., Foo, C. F. and Puttgen, H. B. Load and renewable energy forecasting for a microgrid using persistence technique. *Energy Procedia*, 2017. 143: 617–622.
7. Shakya, A., Michael, S., Saunders, C., Armstrong, D., Pandey, P., Chalise, S. and Tonkoski, R. Solar irradiance forecasting in remote microgrids using markov switching model. *IEEE Transactions on sustainable Energy*, 2016. 8(3): 895–905.
8. Guermoui, M., Melgani, F., Gairaa, K. and Mekhalfi, M. L. A comprehensive review of hybrid models for solar radiation forecasting. *Journal of Cleaner Production*, 2020. 258: 120357.
9. Massaoudi, M., Refaat, S. S., Chihi, I., Trabelsi, M., Abu-Rub, H. and Oueslati, F. S. Short-term electric load forecasting based on data-driven deep learning techniques. *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*. IEEE. 2020. 2565–2570.

10. Massaoudi, M., Chihi, I., Sidhom, L., Trabelsi, M. and Oueslati, F. S. Medium and long-term parametric temperature forecasting using real meteorological data. *IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society*. IEEE. 2019, vol. 1. 2402–2407.
11. Lei, Z. and Yang, Y.-b. Research on data mining algorithm for regional photovoltaic generation. *Advanced Hybrid Information Processing: Third EAI International Conference, ADHIP 2019, Nanjing, China, September 21–22, 2019, Proceedings, Part I*. Springer. 2019. 429–438.
12. Akhter, M. N., Mekhilef, S., Mokhlis, H. and Mohamed Shah, N. Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. *IET Renewable Power Generation*, 2019. 13(7): 1009–1023.