Advanced Forecasting Algorithms for Renewable Power Systems

By

Meftah Elsaraiti

A Thesis Submitted to Saint Mary's University, Halifax, Nova Scotia
In Partial Fulfillment of the Requirements for the Degree of PhD in Applied Science

April 14, 2023, Halifax, Nova Scotia

Approved: Dr. Adel Merabet
Supervisor
Division of Engineering
Saint Mary's University, Halifax, Canada

Approved: Dr. Ehab Elsharkawi
Committee Member
Division of Engineering
Saint Mary's University, Halifax, Canada

Approved: Dr. Cristian Suteanu
Committee Member
Department of Environmental Science
Saint Mary's University, Halifax, Canada

Approved: Dr. Hamed Aly
Committee Member
Department of Electrical and Computer Engineering
Dalhousie University, Halifax, Canada

Approved: Dr. Bruce Smith
External Examiner
Department of Mathematics and Statistics
Dalhousie University, Halifax, Canada

Date: April 14, 2023

Abstract

Advanced Forecasting Algorithms for Renewable Power Systems

By Meftah Elsaraiti

Wind and solar power prediction is a challenging but important area of research. The thesis you described explores various statistical models and deep learning methods to improve the accuracy of wind speed and solar radiation predictions. The use of autoregressive integrated moving average (ARIMA) models, long short-term memory (LSTM) based recurrent neural network (RNN) models, and multilayer perceptron (MLP) neural networks were studied to predict future wind speed values and the performance of a photovoltaic (PV) system. The results showed that the proposed models can effectively improve the accuracy of wind speed and solar radiation prediction and that the LSTM network outperformed the MLP network in predicting solar radiation and energy for different time periods. It is important to note that the performance of the models may vary depending on the specific dataset used, the hyperparameters, and the model architecture. Therefore, it is essential to carefully tune these parameters to achieve the best possible performance. Accurately predicting the performance of a PV system at short time intervals is particularly important in the context of renewable energy sources, as it can help optimize the usage of these resources and improve overall efficiency. This research can contribute to the development of more accurate and reliable prediction models, which can lead to more efficient use of wind and solar power, reduce costs, and promote the adoption of renewable energy sources.

April 14, 2023

Acknowledgments

I would like to extend my sincere gratitude to my supervisor, Dr. Adel Merabet, for his essential guidance while I completed the fulfillment of my Ph.D. Your kindness, compassion, and belief in me encouraged me to continue on this journey. Your assistance and encouragement have always played a significant role in my progress and achievements. I am incredibly appreciative that I got to work with you. I would like to express my sincere gratitude to the honorable members of my committee, Dr. Ehab Elsharkawi, Dr. Cristian Suteanu, and Dr. Hamed Aly, for their advice and support throughout every phase of my research project, as well as to Dr. Bruce Smith, who served as my external examiner and provided illuminating feedback on my work. I also appreciate the help I got with my studies at Saint Mary's University (SMU) from the members of the Laboratory of Control Systems and Mechatronics and the Division of Engineering. I also want to thank my close friends for their spiritual guidance and support throughout my academic journey. They have been by my side even in finalizing this research project. I also greatly appreciate the support and encouragement I have received from all my loving brothers and sisters. I also incredibly appreciate the major support of my beloved wife and her unlimited support of me. She has always given me the will to succeed and the encouragement to go above and beyond whatever I believe I am capable of. I am also appreciative of all the family members who have supported me throughout the years and had a significant impact on my life. Finally, I extend my gratitude to the Libyan Ministry of Higher Education and to everyone who helped me along the way during my academic career.
Thank you very much, everyone!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABREVIATIONS

RES	Renewable energy sources
RNN	Recurrent neural network
LSTM	Long short-term memory
R^2	Coefficient of determination
MAE	Mean absolute error
MAPE	Mean absolute percentage error
RMSE	Root mean square error
ARIMA	Autoregressive Integrated Moving Average
ANN	Artificial neural networks
MLP	Multi-layer perceptron
MPPT	Maximum power point tracker
PV	Photovoltaic
PVPs	Photovoltaic plants
NREL	National renewable energy laboratory
SARIMA	Seasonal autoregressive integrated moving average
ARMA	Autoregressive moving average
AR	Autoregressive
MA	Moving average
SVM	Support-vector machines
NWP	Numerical weather prediction
MLR	Multiple linear regression
BRT	Bagged regression trees
WT-ARIMA	Wavelet Transform (WT) based ARIMA Autoregressive integrated moving average
NSCC	Nova Scotia Community College

LIST OF SYMBOLS

P	Mechanical power
A	Area of swept
v	Velocity of air
D	Dimeter of rotor
C_T	Thrust coefficient
C_p	Efficiency
β	Angle of inclination blades
ρ	Air density
λ	Speed of the wind wheel
oc	Open circuit
sc	Short circuit
FF	Fill Factor
η_c	Efficiency of a photovoltaic
γ	Solar irradiation coefficient
T_c	Operating temperature
ε	white noise

CHAPTER 1

Introduction

1.1 Renewable Energy and Current Statistics

Global primary energy consumption is continuously rising due to factors such as population growth, technological advancements, and increasing energy needs. Most power plants rely on non-renewable resources, contributing to the unsustainable use of these resources. As a result, the International Energy Agency projects that global energy consumption will reach 33.4 trillion kWh by 2030 and 41.3 trillion kWh by 2050. To meet this demand, the energy sector needs to undergo strategic changes, such as decentralizing generation, addressing supply and demand challenges, adopting smart grids, and addressing the issue of climate change. This has led to a worldwide effort to increase the use of solar and wind energy, which are among the most important types of alternative energy sources. Among the advantages that renewable energies have over conventional energies is the fact that they are inexhaustible and the operating costs, as well as the production of pollutants, particularly CO2 emissions, are very low. However, the main disadvantages of some renewable sources like wind and solar energy are their intermittency or volatility and the difficulty in predicting exactly how much energy will be produced. These two disadvantages are significant because, in order to meet the electrical network's operating codes, there must be a balance between the electricity consumed and the electricity generated. For example, if the electrical energy produced by wind energy presents variability in the power generated, there is non-dispatchable energy, and this can produce alterations in quality or supply. Unless some level of technological readiness and innovation receptivity in the economy is achieved, specific policy strategies will not be successful. Over the past decade, the depletion of fossil fuels and the growing demand for electricity have been high on the international agenda. Modern energy systems are undergoing a phase of major change as they move from a centralized nature that relies heavily on limited and polluting resources to clean and decentralized energy solutions in line with the 2016 Paris Climate Agreement to combat climate change and the depletion of natural resources, in addition to a cleaner, fairer, and safer energy future. As the transition to renewable energy becomes necessary as an alternative to traditional sources, it has become a priority in the electric power industry [1]. By adopting “green” power generation solutions, researchers continue to do their best to develop

more advanced and efficient technologies to make green power generation more cost-competitive with traditional energy sources, making it a more viable option for wider adoption [2]. The International Renewable Energy Agency (IRENA) reported an increase in the share of renewable energy in new generation capacity in its 2021 annual statistics. The world added 257 gigawatts of renewable energy, with solar and wind accounting for 91% of the total net additions. This indicates a growing trend towards renewable energy sources. The trend towards renewable energy sources promises to achieve carbon neutrality by 2050. The fastest growing renewable energy source is solar, which saw an increase of 133 GW (+19%), and wind power, which saw a growth of 93 GW (+13%). Hydroelectric capacity increased by 19 GW (+2%) and bioenergy by 10 GW (+8%), while geothermal energy increased by 1.6 GW. The dominance of wind and solar in renewable capacity expansion, along with the growth of geothermal, resulted in a significant annual increase in renewable energy generating capacity [3]. The renewable share of the annual growth in power capacity is shown in Figure 1.

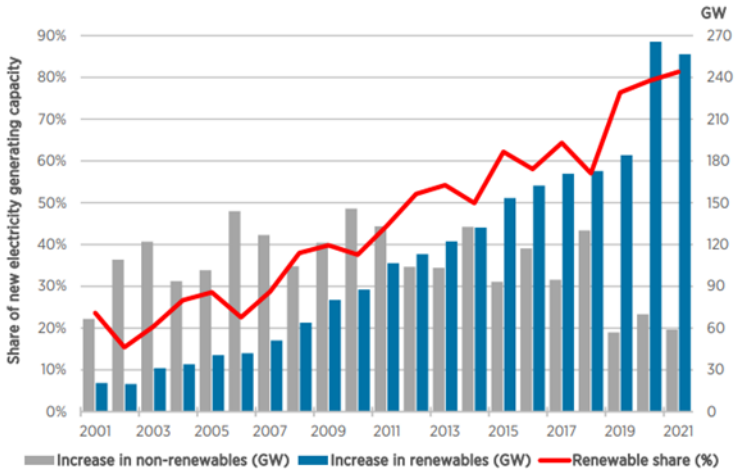

Figure 1. The renewable share of installed capacity, 2001-2021 [3].

Canada faces challenges in achieving carbon neutrality, but also has the potential to be a leader in the energy transition. The promotion of renewable energy sources and clean technology sectors is key to securing a stable energy supply while addressing climate change. Canada is just beginning to scratch the surface of its vast and untapped wind and solar resources. At the end of 2022, Canada had nearly 15 GW of installed wind capacity and more than 4 GW of primary solar capacity, for a total of more than 19 GW of installed renewable capacity across the country. Overall, the wind, solar, and energy storage sectors grew by 10.5% in 2022. Canada now has an installed capacity of

more than 19 GW of utility-scale wind and solar energy. Canada added more than 1.8 GW of new generation capacity in 2022, significantly more than last year's growth (1 GW in 2021) [4].

1.2 Research Motivations

The increased concern about preserving the environment and the need to seek new, less polluting energy sources have significantly boosted the use of energy sources considered clean, such as solar, wind, and tidal, among others [5]. Currently, the world is witnessing a transformation of the energy industry towards an increase in the share of renewable energy sources (RES), subject to the sustainable development of the economy, energy, and environment in the long term. Although researchers continue to improve the capacity of wind turbines [6] and the performance of solar panels [7]. Integration of a high percentage of renewable energy sources into electric power systems is difficult due to the stochastic and unpredictable nature of the generation of renewable energy sources, which increases the level of uncertainty in the processes of changing the parameters and conditions for the operation of energy systems, making it difficult to manage them [8, 9]. One of the biggest difficulties is its integration into the electrical system. Volatility and uncertainty are among the biggest challenges solar and wind power pose to the grid. Their large-scale integration into the electricity grid can have a negative impact on its stability [10]. Due to the importance that electric power generation from renewable energies provides to the electricity sector, the use of alternative energy sources has become a relevant topic all over the world. The dynamic changes in modern societies have made renewable energy sources a hot topic for an increasing number of researchers, driven by the need to reduce the negative impacts of climate change, air pollution, and dependence on limited natural resources. When these emerging problems are discussed, the proposed solution is to make efforts to shift away from methods of power generation that rely on fossil fuels, even if only slightly, and towards the use of environmentally friendly renewable energy sources. The most common ones are solar and wind energy as the energy source is free, and greenhouse gas emissions are zero [11]. The main issue with renewable energy is its inconsistency and difficulty in forecasting. In the case of solar energy, no two days will have the same energy production because it depends on irradiation, temperature, angle, and pollution. Although it will be possible to express the periodicity of the sun, in the case of wind energy, the wind will depend on the atmospheric pressure, the position of the storms, the area, and the height. Predicting the power output of renewable energy sources is one of the problems of the power

system. The predicted power of wind or solar radiation has forecast errors that depend on the time of year, climate, and geographic location [12]. Inaccurate forecasts of renewable energy variables are a significant source of error. Variable imprecision affects power system planning, such as generator scheduling. Therefore, these errors must be taken into account to improve the quality of the power system. Modeling errors in the forecast of renewable energy sources allows for errors to be taken into account in various algorithms. Renewable energies and their incorporation into the energy system made it possible to limit the exploitation of other, more limited, and polluting energies, such as those of fossil and nuclear origin. To facilitate this change, forecasting of renewable energy generation has been worked on, and in recent years, the advantages offered by artificial intelligence in terms of accuracy have been incorporated. Renewable energy generation forecasting is an ever-advancing and growing discipline. Although it has been carried out systematically for more than 20 years, greater precision is increasingly demanded, as are diverse time horizons (from the closest, in the next few minutes, to the farthest, in the next few months or even reaching years). An electrical network supplied mainly by intermittent RE sources requires the prediction of these sources in the management of the network in order to ensure its balance [13]. In an electrical network, it is essential to balance the supply and demand of electrical power [14]. The idea is to determine and implement a prediction method that best fits into the management of the microgrid by integrating renewable energy sources. In the literature, we find that there are several prediction methods for different prediction horizons, and these different methods have very diverse strengths and weaknesses and change according to the contexts or the environment of the prediction [15, 16]. This is why it is important to identify a prediction method that best fits the project. The prediction horizons are as follows: long-term, medium-term, short-term, and very short-term predictions. However, in this application, we want to predict renewable energy sources (wind and photovoltaic) in the context of decentralized network management. For this type of management, it is important to choose the prediction horizon that is essential for the management of the network. The time to acquire prediction information is critical for management to make appropriate decisions about alternative solutions. In addition, the data that will be predicted must be studied, and it is not always predictable, so it is important to determine the level of data that can be predicted. For this reason, new technologies and, specifically, artificial intelligence, have much to contribute to this field. Time series forecasting is an important tool for predicting future trends in renewable energy consumption and production. The correct selection

of a forecasting method is crucial to ensure accurate predictions and inform decision-making. As renewable energy becomes increasingly important, it is vital to have accurate forecasts to support the growth of the industry and help meet the demand for sustainable energy sources. Statistical methods are able to make predictions to a certain extent. In deep learning, a neural network is trained to extract features from data, such as patterns and relationships, and then use those features to make predictions. This allows deep learning to handle complex and non-linear relationships in data and make predictions with high accuracy. Moreover, deep learning models are highly flexible and can be adapted to different types of data, making them an attractive choice for time-series forecasting tasks. However, it is important to note that the success of deep learning models heavily depends on the quality and quantity of the data, as well as the proper hyperparameter tuning of the model. This has led to the development of deep learning algorithms, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks, which have shown remarkable results in the field of time-series forecasting, especially for areas such as energy consumption, stock prices, and weather forecasting. However, it is important to note that the use of deep learning in time-series forecasting is still in its early stages and further research and development are needed to fully leverage its potential [17]. However, an important feature of all the listed neural networks is their lack of memory about time dependencies in the data. They process each packet of input data independently, without saving the state between them, while human intelligence tends to perceive new information based on previous information, and constantly replenishes its experience as new information arrives. Artificial intelligence technology has been attracting attention in recent years. Deep learning technology is based on a neural network, but by using multiple layers of this neural network, it is possible to deepen the characteristics of the data and make it learn [18]. Recurrent neural networks (RNNs) have achieved more robust predictions for processing sequential data. However, this model suffers from long-term dependencies. Data cannot be properly processed because it is difficult to retain information from the past. The further apart the relevant information is from the context in which it is needed, the more difficult it is for the RNN to connect them. Recent research has focused heavily on learning algorithms for the deep architecture of the recurrent neural network (RNN) and its long short-term memory (LSTM), which have shown impressive results for modeling time series data in many fields compared to traditional statistical methods [19-21]. RNN has problems; that is, the more time series data it gets, the less effective it becomes. More specifically, deep learning

continues to learn with an algorithm called backpropagation, but the value of the RNN gets smaller and smaller and eventually vanishes. Long-term information is lost due to the vanishing gradient problem for simple RNNs [22]. RNN tends to face two issues, known as "explosive gradients" and "vanishing gradients." These issues are determined by the size of the gradient, which depends on the loss function along the error curve. Gradients are the values used to update the weight parameters of neural networks. The problem with a vanishing gradient is that the gradient shrinks as it propagates through time. When the gradient is too small, the algorithm no longer learns. Explosive gradients take place when the gradient is too large; this causes unstable model structuring [23]. This is solved with an LSTM-based RNN to control gradient growth in a type of RNN. The use of LSTM-RNN in predicting renewable energy is very little reported in the literature [24]. Compared with traditional approaches for analyzing and predicting time series, such as ARIMA and its variations, the LSTM approach proves to be much superior [25]. The LSTM has already proven to be well suited to perform the classification, processing, and prediction of time series with distinct time intervals. The LSTM model manages to obtain relevant characteristics from the input data and preserve this information for a period of time that can be long, and during the training process, a choice is made between excluding and preserving the data. The information is based on weight values. Therefore, the LSTM model has the ability to learn what it should preserve or remove from its memory [26, 27]. The impact of the different time intervals used to carry out the forecast has received very little study or analysis. In approaches using the LSTM model, the different configurations refer to the size of the time window and its time steps that are necessary for the model to be able to produce better results and their impact on the forecasting mechanism. Another point little mentioned and little explored in the literature is in relation to whether or not the use of meteorological data is advantageous as there are few studies that effectively compare the performance of the model used with the daring addition of the same [28]. The works found in the literature have very different approaches to solving the problem of forecasting the amount of energy generation from wind turbines and photovoltaic systems. It was possible to verify that in most of the works, there were higher difficulties and errors when dealing with days. For instance, the amount of energy produced by wind turbines is proportional to the variable speed of the wind. When it comes to solar energy, cloudy or rainy conditions are the same. On the other hand, although errors are still high in most methods, the use of LSTM-type recurrent neural networks is increasingly being highlighted due to the good results obtained. It was also

possible to see that some methods could be improved and had important loopholes worth investigating. Modification of the parameters of the LSTM network used and the variables used to obtain the prediction have been little explored. Despite the high variability of wind speeds and radiation on cloudy and rainy days, the results of studies using LSTM networks enhance the ability of this model to learn time-series behavior.

1.3 Objectives

The major goal of this work is to apply advanced renewable energy forecasting and scheduling technologies. This prior energy forecast is a key component of risk assessment, which includes confidence in estimated energy production, CO_2 emission countermeasure costs, lower equipment prices due to the expansion of solar and wind power, and schemes by energy providers for the contribution they make to generation systems. To achieve this, it is necessary to choose the appropriate forecasting model, ensure that the data collected is correct, accurate, reliable, consistent, and complete to be effective, and determine the forecast horizon to reach an accurate forecast with minimal errors. The specific objectives are:

1. To review and compare existing deep learning models for wind speed and solar energy prediction and their training strategies, focusing on the problem of vanishing gradients and explosions.
2. To investigate the causes of vanishing gradients and explosions in deep learning models for wind speed and solar energy prediction and their impact on prediction accuracy and training time.
3. To propose a deep learning model structure and training strategy that can address the problem of vanishing gradients and explosions and achieve high accuracy and fast training time for wind speed and solar energy prediction.
4. To evaluate the performance of the proposed model on different datasets and compare it with existing models.
5. To analyze the trade-off between the prediction accuracy and training time of the proposed model.

1.4 Methodology

The methodology used in this work consists of bibliographic research and experimental research. In the bibliographic research, a review of modern forecasting methods in time series related to wind and solar energy and a review of the literature were conducted in order to evaluate the findings of the developments of forecasting methods. The experimental research includes data acquisition and preparation, as well as the application of time series forecasting methods. The steps are detailed in next sub-sections.

1.4.1 Research Review

Scientific publications provide a platform for researchers to communicate the details of their research activities in the area of energy applications. These publications typically include the methods used, the results obtained, and the conclusions drawn from the research. Scientific publications are a means of sharing knowledge and contributing to the advancement of the field. They also allow for peer review, which is a critical part of the research process, ensuring the reliability and validity of the results obtained. The research review seeks to establish the theoretical work foundation. The sources researched are books, handouts, articles, printed or electronic magazines, websites, online courses, and others. The theoretical foundation is subdivided into two topics: 1) Wind and solar energy, and 2) a survey of contemporary time series forecasting techniques applied in the field of wind and solar energy and renewable energy systems, the following are addressed: Wind resources, mathematical equations, and physical concepts related to wind turbines; wind energy forecasting methods; the power curve for a modern wind turbine; solar radiation; the electrical efficiency of a photovoltaic system; estimation of the annual solar output of a photovoltaic system; units and tools for measuring wind speed and solar radiation (the latter is appropriate); define the parameters that will compose the data set used in training, validation, and testing of the prediction model. The subject of time series is already dealt with; Statistical Models: Concepts, Techniques, and Methods Most Used; Deep Learning: Artificial Neural Networks, Types of Neural Networks, especially LSTM-RNN. The literature review provides a useful guide for selecting appropriate technologies for research and improving the accuracy of wind speed and PV generation predictions. Wind speed prediction and photovoltaic (PV) generation are important areas of research in renewable energy. Time series and artificial intelligence (AI) techniques have been applied in these areas to improve forecast accuracy and

provide insights into renewable energy generation. Through this research, we seek to provide an overview of the latest research in these areas.

1.4.2 Experimental Research

Box-Jenkins methodology was applied, which is a traditional statistical modeling approach that involves three steps: identification, estimation, and diagnostic checking. The identification step involves selecting an appropriate model that can best explain the observed data. The estimation step involves estimating the parameters of the selected model using the available data. The diagnostic checking step involves evaluating how well the model fits and ensuring that the model assumptions are met. In addition to the Box-Jenkins methodology, the study also used auto-regression models that use only historical wind speed or solar data to obtain information for training the LSTM model. Auto-regressive models are statistical models that use past observations to predict future values of a variable. The study used several statistical performance measures, including the coefficient of determination (R2), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE), to evaluate and compare the performance of the models.

1.5 Main Contributions

The contributions made by this thesis to the fields of wind speed forecasting and solar energy forecasting can be summarized as follows:

1. Development of a deep learning model that can predict wind speed and solar energy with high accuracy and fast training time while addressing the problem of vanishing gradients and explosions.
2. Investigation of the causes of vanishing gradients and explosions in deep learning models for wind speed and solar energy prediction and their impact on prediction accuracy and training time.
3. Scaling the weights of the layers to have unit norm, which can help to prevent the gradients from becoming too small or too large during backpropagation.
4. Evaluation of the performance of the proposed model on different datasets and comparison with existing models.
5. Demonstration of the potential of the proposed model for real-time wind speed and solar energy prediction applications.

6. The focus of the PV power forecasting task for the days ahead was done at 30-min intervals using data previously collected over one year using LSTM networks, a special model for RNN with additional features for saving data sequences. This type of ANN connects past data with current data and compares them using a multi-layer neural network (MLP).

1.6 Scope

The projects make use of time series data on solar radiation and wind energy because these two renewable energy sources are the most widely used and the hottest topics in much of the current research. The Halifax, Canada, weather database was used as a source for a series of wind speed and solar energy to propose a high-accuracy and good forecasting stability wind and solar speed prediction system that can realise wind speed and solar energy data only as input data and output forecasted data of wind speed and solar energy generated in the short term from power plants in the future.

1.7 Organization

The rest of the thesis is structured as follows: Chapter 2 provides an overview of the basic concepts and characteristics of wind and solar energy; Chapter 3 explains the importance and methods of forecasting wind and solar energy and related works; Chapter 4 describes time series analysis methods and approaches; and Chapter 5 contains the results and discussions. Finally, Chapter 6 concludes the thesis and recommends some future research directions.

CHAPTER 2

BACKGROUND AND CHARACTERISTICS OF WIND AND SOLAR ENERGY

2.1 Introduction

The expectations indicate continuous growth in global energy consumption in the coming years, driven by both population and economic growth. On the other hand, fossil energy is limited, which is why it is currently dominant. The increase in its consumption also increases carbon dioxide emissions, and it is expected that the supply and demand balance of energy will become tighter on a global scale. As a result, according to the International Energy Agency (IEA) forecasts, global energy consumption for 2030 will be 33.4 trillion kWh and will increase by about 7.9 trillion kWh by 2050. In this sense, to ensure the increase of global demand, the energy sector requires major changes, namely a shift from centralized power grids to distributed power systems, network development using artificial intelligence, and significantly promoting the introduction of renewable energy (solar and wind energy). Only in this scenario will it be possible to radically reduce the cost of electricity. However, the main disadvantages of these sources are their intermittency or variability and the difficult predictability of the power available to be used. These two disadvantages are significant because in order to meet the electrical network's operating codes, there must be a balance between the electricity consumed and the electricity generated. In recent decades, there has been a great advance in forecasting systems, due to the enormous increase in computing power of current computers, which allow storing, analyzing, and relating large numbers of variables and their values in a very short time. That is why prediction algorithms are a highly developed and used tool today. The importance of predictions lies in the help they provide to plan and anticipate future values that will affect a system, help manage the acquisition of the necessary resources sufficiently in advance, or serve as a tool to maximize profitability by taking decisions that maximize the benefits of an activity. The first efforts to create prediction algorithms in the field of electrical systems were dedicated to the forecasting of electrical demand [29]. From the point of view of the electricity system operator, demand forecasts have helped to inform decisions such as the start-up of electricity production units sufficiently in advance or the maintenance schedule of elements of the electricity system. From the point of view of market agents, demand forecasting has been a tool used to optimize their energy offers to the electricity market with the aim of maximizing profitability. But the great boom that renewable energy sources have had in

recent years has forced a new forecasting problem to be considered: that of knowing sufficiently in advance the amount of energy that these plants are going to generate. Given that the greatest boom in the field of renewable energy has been in wind and solar power, greater efforts have been made to create reliable and efficient energy prediction tools that help integrate this form of energy into the network [30, 31].

2.2 Wind Energy

The installed capacity of renewable energy based on wind resources has been steadily growing in the electricity system in recent years. Wind power has the characteristics of strong randomness, obvious intermittent nature, large fluctuation range, irregular fluctuation frequency, and anti-peak regulation. These uncertain factors increase the potential risk of grid-stable operation and the difficulty of grid peak regulation [32]. Wind energy production is essential to ensuring the security of electricity supply in the generation mix and, together with photovoltaic energy, constitutes a necessary pillar to ensure the country's energy security. Wind energy is one of the renewable sources with the greatest potential in the world. In fact, wind power is the generation technology with the most installed power. However, its production is based on a source that is beyond human control, so its production shows great uncertainty. Wind energy is based on the use of the wind. The kinetic energy generated by the movement of the air mass is converted into mechanical energy, which turns the wind turbine blades, and finally into electrical energy [33].

2.2.1 Mathematical Equations and Physical Concepts Related to the Wind Turbine

The law of conservation of mass and momentum serves as the foundation for the basic theory guiding wind turbine operation. Albert Betz presented this theory in 1919, and it was published in his book “Wind Energy and its Extraction through Wind Mills” in 1926. Both wind turbines with horizontal and vertical axes can use the proposed theory [34].

Wind energy is the kinetic energy contained in the movement of an air mass. Power is equal to work divided by the change in time, given by the equation.

$$P = \frac{W}{\Delta t} \tag{2.1}$$

Kinetic energy is the work done by the air mass to be in motion, which is a form of energy that can be harnessed by wind turbines to generate electricity, as shown by Equation (2.2)

$$W = Ec = \frac{mv^2}{2} \tag{2.2}$$

The expression $\frac{1}{\Delta t}$ represents the time rate of change of the kinetic energy, or the power (P) generated by the air mass, as given by the formula:

$$P = \frac{Ec}{\Delta t} \tag{2.3}$$

Then

$$P = \frac{mv^2}{2\Delta t} \tag{2.4}$$

The power generated by a wind turbine is presented in equation (2.5):

$$P = \frac{mv^2}{2\Delta t} = \frac{m}{\Delta t}\frac{v^2}{2} = \frac{\dot{m}\quad v^2}{2} \tag{2.5}$$

Where, ṁ (kg) represents the mass flow of air that passes perpendicularly through section area S. This flow is described by:

$$\dot{m} = Q = \rho v A \tag{2.6}$$

Where, ρ is the density of air (kg/m^3), v is the speed of wind (m/s), and A is the area of the turbine rotor perpendicular to the wind direction (m^2).

Figure 2.1 illustrates the wind flow when passing through a turbine and helps us to understand how the wind behaves when passing through a rotor. The tube model presented was made by Betz.

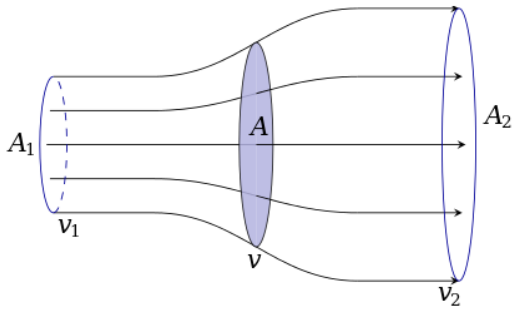

Figure 2.1. Diagram of Betz's tube showing air streamlines from the entrance to the exit (Reference: en.wikipedia.org)

The generation of electric energy in the generator is directly related to the transformation of the kinetic energy contained in the wind; therefore, the maximum power that can be collected from

the wind is that at which the rotor output speed is zero. However, if this occurs, there will be an accumulation of air at the rotor outlet, which will interrupt the airflow and, consequently, the generation of electrical energy. However, Betz physically demonstrated that in order to achieve the maximum theoretical power possible, the speed at the rotor's exit must be exactly equal to one-third the speed at the rotor's entrance, implying that two-thirds (2/3) of the kinetic energy contained in the mass of air passing through this rotor would be "captured" and converted into electrical energy [35].

The mechanical power available from the wind before it interacts with a wind turbine is given by the following formula in watts:

$$P = \frac{\rho v A v^2}{2} = \frac{\rho v^3 A}{2} \tag{2.7}$$

This power is proportional to the cube of the wind speed (v^3), the air density (ρ), and the area (A) swept by the wind turbine. The recoverable energy of the wind is associated with the kinetic energy as it passes through the turbine rotor. Equation (2.7) shows that the power contained in the wind varies with the cube of the wind speed, with the specific mass of the air, and with the area *A*, swept by the turbine blades used for the test, and cut by this mass flow of air (wind). In this way, the energy generated by the turbines is very sensitive to the wind speed that passes through the wind turbine blades (v^3). There is still a high sensitivity with respect to the rotor diameter, as shown in Equation (8).

$$A = \frac{\pi}{4} D^2 \tag{2.8}$$

Where, *A* is the cross-sectional area cut b y the airflow (m^2) proposed by Betz's theory; π is the constant; and *D* is the rotor diameter. The swept area of a Darrieus vertical-axis wind turbine can be difficult to calculate exactly due to the complex shape of the blades. However, one possible method for approximating the swept area of a Darrieus is to assume that each blade can be represented by a parabolic curve; the swept area can be stated as follows:

$$A = \frac{2}{3}\,(maximum\ rotor\ width\ at\ the\ center)(heigt\ of\ rotor) \tag{2.9}$$

Despite having only two or three thin blades with solidities ranging from 5 to 10%, the wind turbine effectively captures wind energy flowing across the entire swept area. The ratio of the solid area to the swept area of the blades is known as the solidity. The solidity ratio of the contemporary two-blade turbine is low. Because it uses less blade material to sweep wide areas, it is more cost-effective [36].

Equation (2.10) describes the upthrust force (*T*) that the wind applies to the blades in the direction of the wind:

$$T = \frac{1}{2}\rho A v^3 C_T \tag{2.10}$$

Where, C_T is the thrust coefficient.

Another very important parameter in wind turbine analysis is the power coefficient(C_p), called the efficiency measure for a wind turbine. According to physicist Albert Betz, who is responsible for investigating the operation and power of an ideal rotor, based on the amount of axial movement, there is a limit to the use of wind turbines. The best theoretical records of wind exploitation reach 59% (referring to the value of (C_p =0.59) [37]. This value suggests that an ideal rotor is made to work in such a way that the wind speed in the rotor is 2/3 of the wind speed, considered in its free path, where 8/9 of the kinetic energy contained in this airflow is lost. Among the effects that cause a decrease in the power coefficient, one can point out the rotation of the blades downstream of the rotor, the number of blades associated with the loss at the tip of the same, and aerodynamic resistance forces [38]. Power factor is a measure of the efficiency of a wind turbine in converting the kinetic energy of the wind into electrical power. It is defined as the ratio of the actual power output of the turbine to the power that would be generated if the turbine were able to capture all of the kinetic energy of the wind passing through its swept area. The power coefficient is denoted by the symbol C_P and is given by the equation:

The power coefficient measures the energy that can be produced in a turbine, disregarding any loss, in relation to the total energy contained in the wind that passes through it. The aerodynamic coefficient C_p for vertical axis wind turbines is related to the power extracted from the wind by the P_E turbine and the power contained in the wind P_V, and is given by the equation:

$$C_p = \frac{P_V}{P_E} \tag{2.11}$$

The kinetic energy contained in the wind cannot be fully captured by the turbine, since the air passing through the turbine rotor must be extravasated. Therefore, C_p is added to the power calculation such as [39]

$$P_E = \frac{1}{2} C_p \rho A v^3 \tag{2.12}$$

Where, p_e is the electrical power delivered by the wind turbine, ρ_{air} is the density of the air, S is the area it traverses, defined by the diameter of the rotor or the size of the blades, and v is the speed of the air current. C_p is the power coefficient, which indicates the degree of use of the energy contained in the air current, with a maximum theoretical value of 0.59 according to the Betz limit (that is, the conversion efficiency of wind turbines is no more than 59% of the energy provided by the wind) [34]. Using the formula for the area of a circle, the swept area of the turbine may be determined from the length of the blades such as

$$A = \pi r^2 \tag{2.13}$$

Where, r is the radius of circle represented by the blade length as shown in Figure 2.2.

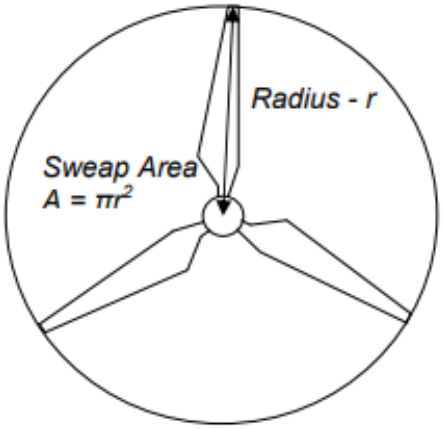

Figure 2.2. Wind energy swept area [40]

Wind power is proportional to wind area and wind speed cubed. When the wind speed fluctuates, the wind power fluctuates according to its cube and has a large fluctuation range. Natural winds do not always blow in a constant direction or speed, but they always fluctuate from time to time, and this causes them to be unstable. These fluctuations cause great difficulties for wind farm operators in scheduling the system and sending energy because of the lack of prior knowledge of

the schedule for its availability, and this makes for an accurate and reliable prediction of wind speed several hours in advance that has several benefits, including:

• The generation schedule can timely and effectively deal with wind generation.

• Mainly providing information to companies that install wind power generation facilities.

• To intensify competition in the field of renewable energies and provide greater progress in sustainable development, making it a wonderful supplier of clean and inexhaustible energy.

2.2.2 Wind Resource Assessment

Wind energy is a stochastic and spatiotemporal process that depends on several factors, including wind speed, direction, and frequency of occurrence. Wind speeds can vary greatly both spatially and temporally, depending on a variety of factors, such as local topography, land use, and weather patterns. The average wind speed is one of the most important factors that determine the wind energy potential at a site. It is typically measured at a height of 10 meters above ground level, although the hub height of wind turbines is typically much higher (on the order of 50 to 100 meters) in order to capture stronger and more consistent winds [41]. The process of converting the mechanical energy contained in the wind into electrical energy is subject to mechanical and electrical losses. These losses originate in mechanical components due to friction and in electrical components such as transformers, generators, cables, and electronic devices due to the Joule effect. Thus, the electric power effectively injected into the network is given by

$$P_E = \frac{1}{2} C_p(\lambda, \beta) \rho A v^3 \eta \tag{2.14}$$

Where, η is the yield of the conversion process, which includes both electrical and mechanical losses due to rotation, and $C_p(\lambda, \beta)$ is the performance coefficient, which depends on the angle of the inclination blades β, and the coefficient of the speed of the wind wheel λ.

2.2.3 Electric Power and Power Curves

Analytical methods can be used to determine the power curve of a wind turbine for various wind speeds, using the coefficient of performance $C_p(\lambda, \beta)$. Figure 2.3 shows the power curve for a contemporary wind turbine.

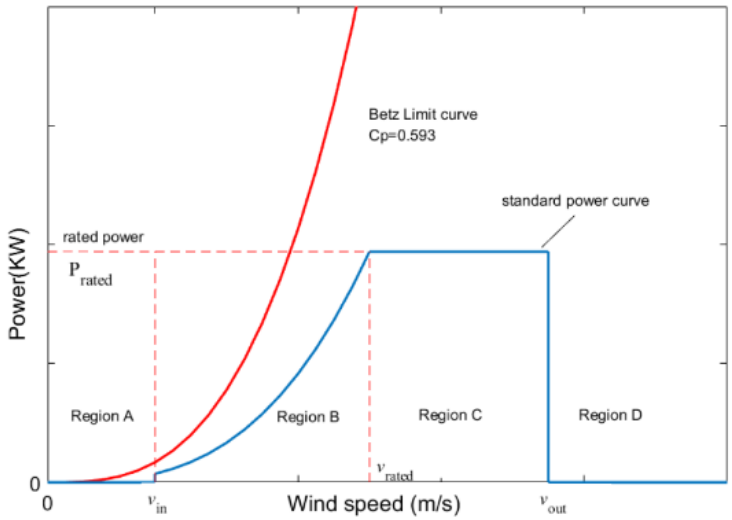

Figure 2.3. Power curve for a modern wind turbine [42].

When the wind speed in the first zone is less than a certain minimum known as the $v_{cut\ in}$, the power output is zero. For speeds below this value, the generator is not connected to the transmission shaft because the energy losses in the system for these wind speed values would be greater than the amount of energy generated. The second region, defined by $v_{cut\ in}$ and, v_{rated}, has a rapid increase in power production. The generator operating speed is controlled to maximize power generation C_p, for a given blade operating angle. The v_{rated} is the speed at which the generator operates at rated power. The algorithm responsible for this control performs tracking to adjust the Maximum Power Point Tracker (MPPT). Finally, the third turbine-operating region is the one between $v_{rated}$and, $v_{cut\ off}$where the generator operates at a constant speed with power limitation performed through step angle control. The wind speed that determines an emergency in wind turbine operation is the $v_{cut\ off}$, triggering the stop control. The purpose of this control is to prevent excessive mechanical loads from resulting in damage to the turbine structure. Some modern turbines reduce the power generated for wind speeds above $v_{cut\ off}$ so that the turbine can return to producing power as soon as the wind speed drops. For comparison, the curve P_{wind} in figure 2.3, the total power available in the wind.

2.3 Solar Energy

Today, solar energy is one of the most abundant and renewable sources of energy available. It can be easily obtained using photovoltaic panels, from small installations on rooftops to individual large photovoltaic plants (PVPs). Thanks to improvements in efficiency and affordability, recent

years have seen a rapid increase in the number of photovoltaic solar panels installed around the world. In addition, due to its environmental friendliness, the governments of many countries encourage the use of solar energy by creating the necessary incentives and support. For all these reasons, solar energy is expected to play a large role in the global energy supply in the near future. According to research data, by 2050, about 30% of the world's electricity will be supplied from photovoltaic systems [43]. Despite the fact that solar energy has several benefits over other traditional energy sources, the electricity produced by photovoltaic panels is highly variable because it depends not only on solar radiation and temperature but also on other meteorological parameters, such as wind speed, daylight hours, humidity, cloud cover, and precipitation. Solar energy is also an intermittent energy source, as it is only available during the daytime [44, 45]. Indeed, the large-scale integration of solar energy into the electricity grid is a challenging job due to its variability and discontinuity. Solar energy frequently experiences unforeseen shifts, which have a detrimental impact on grid balance and raise operational costs. For the grid operator and businesses supplying electricity from a photovoltaic system, a precise forecast of solar electricity generation is crucial due to its intermittent and uncontrollable nature. Creating an algorithm to ensure the stability of the power supply system is important to provide economic value from the solar system. The development of electricity generation forecasting techniques can lead to such stability and provide an approximation of future production. This allows utility companies to create a control mechanism to switch between the available energy sources present in a given combined station. It should be noted that the generation of solar energy has very different patterns. During sunny days, the photovoltaic system has a high power output. On cloudy and rainy days, the photovoltaic energy output is low and variable, with different generation curves for each of these conditions [46]. Figure 2.4 shows the half-hourly power profiles for three days.

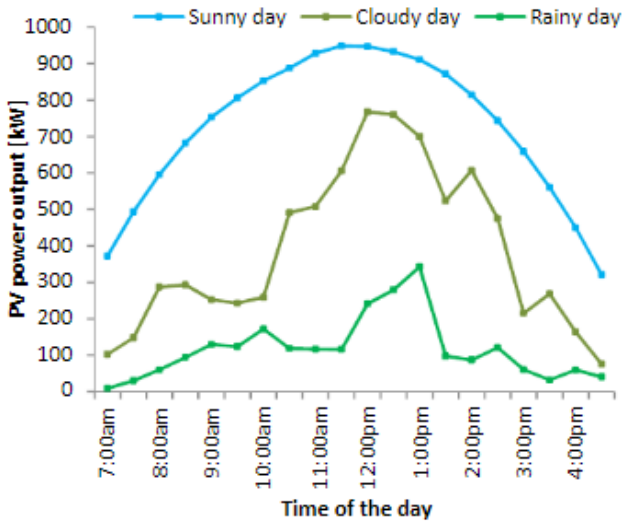

Figure. 2.4. Solar energy production for three separate days under different climatic conditions: a sunny day (20 April 2013), a cloudy day (15 April 2013), and a rainy day (13 April 2013) [47]. Copyright © 2016, IEEE

Solar irradiation has a high correlation with the generation of photovoltaic solar energy. Therefore, the historical series of photovoltaic production data is needed to have a reliable forecast of the energy to be generated and to monitor the performance of the photovoltaic panels. Detecting efficiency losses or adequately planning the distribution in smart energy networks can be done by comparing the generation estimate and the energy produced in a certain time interval [48].

2.3.1 Electric Efficiency of Photovoltaic System

Since the real electrical efficiency depends on the real operating conditions, it is necessary to mathematically evaluate the behavior and relationship of the efficiency with the factors that influence it. To carry out such an evaluation, a real electrical efficiency model that correlates the determinant variables to the real performance can be used. These models comprise the mathematical basis of most renewable energy simulation software. The effect of operating temperature on real electrical efficiency is one of the most representative loss factors and influences real-time efficiency during operation. Therefore, the operating temperature is one of the variables necessary for any electrical efficiency model. The definition of the most commonly used electrical efficiency model was established by [49, 50] and recently revisited by [51, 52], emphasizing the use of operating temperature models. The operating temperature of a photovoltaic cell can be related to the influence on electrical efficiency through current and voltage, which represents electrical power according to the equation [53]:

$$P_m = V_m i_m = V_{oc} i_{sc} FF \tag{2.15}$$

Where, P represents the power, i the electric current, and V the voltage in the cell. The symbol *m* indicates the highest value for the power produced in a module, while the values for *oc* refer to the open circuit, and *sc* refer to the short circuit, respectively.

The values of open circuit voltage and short circuit current are properties that characterize the cell and vary according to the temperature of the cells and the incident irradiation. Through Equation (2.15) we get the definition of the form factor (Fill Factor – FF), which represents a solar cell performance parameter, and is defined as the ratio between the maximum power point divided by the open circuit voltage, V_{oc}, and the short circuit current, i_{sc}. According to Equation (1.15), both V_{oc} and FF decrease with temperature due to the superposition of thermal excitation on the electrical properties of the semiconductor, while i_{sc} increases smoothly [54]. This behavior occurs in the same way on the efficiency of the cell or module, which is defined by Equation 2.16 .

$$\eta_c = \frac{P}{AG} \tag{2.16}$$

or in the format presented by [51].

$$\eta_c = \frac{V_{OC} i_{SC} FF}{G} \tag{2.17}$$

Where, A represents the cell area, G the irradiance and η_c The efficiency of a photovoltaic solar panel refers to how well it is able to convert sunlight into usable electrical energy.

The efficiency is typically expressed as a percentage and is determined by the ratio of the electrical output of the panel to the total energy that is absorbed by the panel from the sun when the solar cell is linked to the electrical circuit. Therefore, to produce maximum power, the cell must be operated under standard operating conditions (25 °C, spectral air mass, AM, of 1.5) [51].

Equations that correlate V_{OC} and i_{SC} with the cell operating temperature, T_c, are known based on several models proposed in the literature [55]. The influence of temperature on electrical efficiency is correlated in these models through the test parameters of each technology, including $I_{SC}, V_{OC}, I_m, V_m, FF, and\ \eta_c$. These parameters become, therefore, important for the correct dimensioning and operation of photovoltaic systems [54]. From these current and voltage models, the efficiency model proposed by [49] is defined, according to the equation:

$$\eta_c = \eta_{ref}[1 - \beta_{ref}(T_c - T_{ref}) + \gamma \log_{10} G] \tag{2.18}$$

Where, η_c is the actual electrical efficiency at operating temperature, T_c, and η_{ref} is the electrical efficiency at reference temperature, T_{ref} under the irradiance of 1000 *W/m²*. In this equation arise the temperature coefficient, β_{ref}, and the solar irradiation coefficient, γ, which are characteristic of the material properties and correlate the losses caused by the effect of temperature and incident irradiation, respectively. In equation (2.18), the last term can be set to zero, as demonstrated by [24], resulting in equation (2.19), which represents the traditional linear expression for the electrical efficiency of photovoltaic technology [53].

$$\eta_C = \eta_{ref}\left[1 - \beta_{ref}\left(T_C - T_{ref}\right)\right] \tag{2.19}$$

Once the electrical efficiency models are known, there is a need to obtain models for predicting the operating temperature of cells and modules.

2.3.2 Degradation of Photovoltaic System

The performance of modules can be gradually reduced over time due to several degradation factors, notably corrosion, debonding (loss of adhesion and contact), delamination, discoloration, fracture of cells or coating materials, as well as the degradation of components and semiconductors [56, 57]. Controlling and predicting the effects of module degradation is vital for maximizing the durability and performance of a photovoltaic system throughout its lifetime [58]. The causes of degradation are typically associated with adverse operating conditions that affect conversion efficiency during actual operation.

2.3.3 Estimation of the Annual Solar Energy Output of Photovoltaic Systems

Solar energy is without a doubt one of the most reliable and cleanest sources of renewable energy. Solar energy is the process of transforming solar radiation into electricity, either directly via the use of photovoltaic cells (PV), indirectly using concentrated solar power, or even both. In order to focus a vast region of sunlight into a narrow beam, concentrating solar power systems use solar tracking devices, lenses, or mirrors. Solar-powered photovoltaic panels work by harnessing photons from the sun to excite electrons in silicon cells, which in turn produce energy. Renewable energy sources could be supplied by this electricity [59]. PV batteries, or cells, are the smallest parts of systems. Cells produced using semiconductors such as silicon and amorphous silicon can be square, rectangular, or circular in shape. When the sun's rays hit the surface of the PV cell, energy is produced by the movement of electrons that break off. However, since the output voltages of the cells are low, modules are formed by connecting the cells in parallel or in series. Panels are created by combining these modules, and arrays are created by combining panels. Thus,

the output power is increased by the desired amount [60]. Figure 2.5 shows the design of a system for supplying usable solar energy via photovoltaic cells.

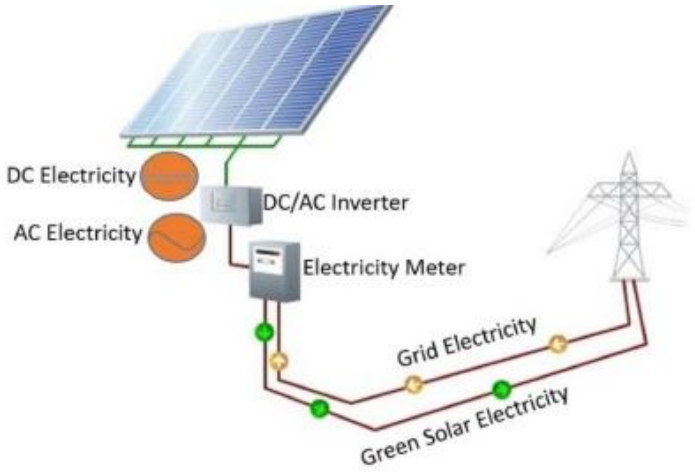

Figure 2.5. Photovoltaic system [61]

As it is an intermittent electrical energy source with many variables that influence energy generation, obtaining mathematical formulations that aim to estimate with good precision the energy from photovoltaic units becomes a great challenge for scholars. In the literature, methodologies can be found that associate the influence of solar irradiation and the nominal power of the photovoltaic units with factors such as the shadow caused by clouds [62, 63], the presence of dust in photovoltaic cells [64, 65], losses associated with the cabling of electrical equipment [66], the lifespan of solar panels [67], and cell temperature [68], among others. In the equation (2.20) that can be found in works done by [69], it can be seen that to determine the energy obtained by the solar plates E_s, in kilowatt-hours (kWh), the solar irradiation of the I_s environment, in kilowatt-hours per square meter (kWh/m^2), was used as variable, as was the installed capacity of the P_{FV} units, in kilowatt (kW), each year, the standard solar irradiance used in efficiency tests on I_P solar panels in the amount of $1kWh/m^2$, the efficiency of the solar panels η_{FV} (75%) , the temperature of the T_c cell, in degrees Celsius (^oC), and the temperature coefficient of the panels γ (considering the use of polycrystalline silicon cells, value of -0.38%/°C).

$$E_s = I_s * \frac{P_{FV}}{I_P} * \eta_{FV} * [1 + \gamma(T_c - 25)] \quad (2.20)$$

The National Renewable Energy Laboratory's (NREL) research has revealed that solar panels degrade at a median rate of roughly 0.5% per year, but that rate can be higher in hotter regions and rooftop systems. This degradation directly affects energy production over time [70].

The final equation used to estimate distributed photovoltaic solar generation is presented by

$$E_s = [P_{FV} \;\; P_{FV-1} \ldots \;\; P_{FV-n}] * \begin{bmatrix} (1-D)^0 \\ (1-D)^1 \\ \vdots \\ (1-D)^n \end{bmatrix} * \eta_{FV} * \frac{I_s}{I_P}[1 + \gamma(T_c - 25)] \qquad (2.20)$$

Where, *n* is the number of years/operation preceding the base year and *D* is the annual degradation factor of photovoltaic modules, which is 0.5% per year.

2.3.4 Influence of Radiation

Among the climatic factors that interfere with photovoltaic systems for the production of electricity, solar radiation is the main agent, having a direct influence on the voltage effect of photovoltaic cells [71]. Solar panels have an energy conversion efficiency that indicates how much electricity can be generated from solar energy. The amount of energy generated is steadily increasing due to the increase in solar radiation. The higher the level of radiation on the solar panel, the greater the open circuit voltage and short circuit current of the solar panel, and the greater the maximum output power [72, 73]. Figure 2.6 shows five simulated I-V (current versus voltage) and P-V (power versus voltage) curves for different radiation conditions incident on the panel. It is possible to observe in Figure 2 (a) that the increase in solar radiation considerably increases the short circuit current; however, it has less effect on the open circuit voltage. Figure 2 (b) illustrates how the maximum power increases as radiation levels rise.

Characteristic curves of the panels representing the variation of current by voltage (I x V) and power by voltage (P x V) allow evaluation of the behavior of the panels in order to obtain the maximum generative power state [75]. Recent research has concluded that the radiation variable is directly related to photovoltaic energy generation [76].

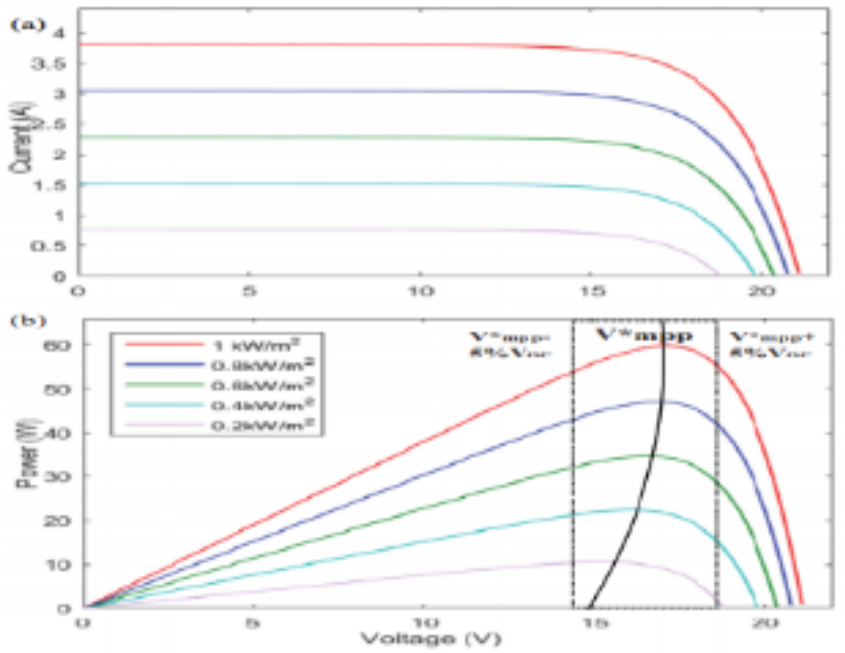

Figure 2.6: Characteristic curves at different radiance levels: (a) I-V curve (b) P-V curve [74]

CHAPTER 3

Forecasting Wind and Solar Energy and Related Works

3.1 The Future of Forecasting for Renewable Energy

Forecasting is one of the most difficult and necessary tasks of data analysis. The complexity of the forecasting process is associated with the need to analyze and evaluate large volumes of data, build and develop models that will correspond to the observed data, and further use these models to obtain future values. The peculiarity of forecasting is the acceptance of the fact that the data cannot be available in the future and that it is possible to evaluate them only thanks to the data that is already available, that is, the data collected from the events that have already taken place [77]. Electricity market forecasting is crucial. Accurate forecasts can help power companies and grid operators make informed decisions about power generation capacity and grid management. This can lead to a more efficient, cost-effective, and reliable electricity system. In energy policy, there are no simple solutions (no single energy source will solve all the problems) that are unlimited, cheap, and free of pollution. Due to the fact that most renewable energy sources, like solar and wind power, can be impacted by weather conditions and can be unpredictable, it is challenging to plan and forecast their generation. However, with advancements in technology and integration into the energy grid, renewable energy is becoming more competitive as an alternative to non-renewable resources and is seen as an important step towards a sustainable energy future [78]. Additionally, the energy generated from sources like wind and solar is highly indiscriminate. This position requires the application of advanced renewable energy scheduling and forecasting techniques [79]. AI and machine learning technologies have been increasingly utilized in weather forecasting to improve its accuracy. By working with meteorologists and incorporating large amounts of data, such as historical weather patterns and current conditions, machine learning algorithms can make more informed predictions about future weather. These technologies can also identify patterns and relationships in the data that may not be easily recognizable by humans, leading to more accurate forecasting. Machine learning is used for renewable energy forecasting by analyzing large amounts of data to detect patterns and make accurate predictions about future energy generation. This allows grid operators to better plan for the amount of renewable energy that will be needed at any given time. Different prediction models can be developed depending on the type of data being used and the time horizon of the forecast (i.e., short-term, medium-term, or

long-term). Short-term forecasts, which range from a few minutes to a few days, are important for making quick decisions on activating electric power generation units and can be used for daily operational purposes. Short-term forecast timeframes are such that changes in weather over a short period of time can be analyzed and used to predict what data will be used the next time. Renewable energy forecasting and scheduling require up-to-date and recent data as often as possible, and short-term forecasting achieves this. In this way, errors would be considerably minimized. Forecasting for renewable energy typically involves combining accurate weather predictions with the availability of plants and systems. Climate change and weather patterns can greatly impact the stability of renewable energy generation, as the amount of energy produced by sources like wind turbines and solar panels is directly related to weather factors such as wind speed and sunlight intensity. Therefore, accurate forecasting and scheduling of renewable energy is important to ensure efficient power generation and to help mitigate the impact of weather-related fluctuations [80]. Accurate forecasts can play a critical role in the development and financing of renewable energy projects. By providing more accurate estimates of the future performance of a project, accurate forecasts can help reduce the uncertainty and risk associated with renewable energy investments. The requirement for precise and reliable forecasts is more important than ever in the current environment of economic uncertainty and the ongoing energy transition. The success of these projects can have a substantial impact on the long-term viability of the energy sector.

3.2 Wind Energy Forecasting

Wind power is a non-programmable form of generation since energy is only produced when the wind blows, which can be highly variable even in the short term with the possibility of intermittence and large changes in short intervals of time. For this reason, it is difficult to know in advance and with sufficient precision the amount of wind energy that we can count on at any given time. This variability makes its operation especially complex, so its future production has to be estimated or forecasted, and this forecast of future power is inevitably affected by a prediction error or uncertainty. If the wind decreases, the power generated in the wind farms also decreases, and this lack of power must be replaced by other sources of generation with a sufficient reserve in magnitude and response speed so that the electricity demand is not affected. On other occasions, it may happen that not all the available wind energy production can be integrated into the system since wind energy is not generated according to consumer needs and it is necessary to reduce the supply of this energy source. For all this, the prediction of wind generation has become a key issue

to make the development and implementation of wind energy feasible and make possible its integration into the electrical system. From the point of view of wind generation or any other source of renewable energy, its forecast is useful both for the system operator and for market agents or park owners. Thus, the operator of the electricity system needs to know in advance the amount of wind energy that will be injected into the network to manage the power that the conventional plants must generate, with the aim of covering the total demand of the system. Meanwhile, market agents will be interested in knowing with the greatest possible certainty the power that their wind farms will generate in order to follow the most profitable strategies in the electricity market. In addition, the owners of wind farms will also be interested in knowing during what periods less power is expected to be generated at their facilities due to scheduled maintenance tasks. These predictions must have a prediction horizon that helps calculate the power production and improves the adjustment to the real demand curves. For the study of wind energy potential, the variables considered most important are wind speed and direction. Many experts mention that a measurement error of 1% can cause a deviation of around 2% in energy production [81]. Furthermore, no mathematical model, whether physical or numerical, provides a perfect and definitive solution [82]. Wind energy is a clean and renewable source of energy and has become an important alternative for generating electric power. Unlike non-renewable sources, wind energy is essentially limitless and produces no harmful emissions or waste [83]. With the increasing integration of wind energy into the energy mix, there is a growing need for research to improve and optimize its performance. One of the challenges of wind energy is that its output can fluctuate greatly due to changes in wind speed, which can be impacted by weather conditions. As a result, there is a need for research and development of new technologies and methods that can help better predict and manage fluctuations in wind energy production so that wind energy can be integrated and provide a reliable and sustainable source of power [84].

3.2.1 Wind Energy Forecasting Methods

The short-term prediction of wind power output from wind farms is an important and actively researched topic, with various prediction techniques available. The three main categories of prediction techniques are physical techniques, statistical techniques, and artificial intelligence techniques. Predicting wind power output is a complex task that requires the use of various techniques, each of which has its own strengths and limitations, and researchers continue to work to improve and refine these techniques to provide more accurate and reliable predictions of wind

power output. This involves developing new models, improving the quality and availability of data, and incorporating new insights and knowledge from the scientific community.

3.2.1.1 Physical Methods

Physical techniques use mathematical models to simulate the physical processes that affect wind power generation. Numerical weather prediction, as the core model of the physical method, has taken into account complex factors such as the topography of the selected area [85]. Some researchers proposed to introduce wind speed and direction as reference variables, combined with the wind power prediction model of the clustering method. In general, weather forecasts are based on observations and mathematical models that describe the dynamic and physical behavior of the atmosphere. The models consist of nonlinear partial differential equations that provide numerical solutions and are highly dependent on the initial conditions for which they were defined. Small errors in the initial data used in atmospheric studies can have a significant impact on the subsequent evolution of the atmosphere, and any slight disturbance in the starting conditions can quickly lead to large errors. This highlights the importance of having accurate and reliable data when studying the atmosphere and making predictions about its behavior to ensure that the results remain accurate and relevant over time [86]. Therefore, the two factors that influence the quality of a weather forecast are the uncertainty in the initial conditions of the atmosphere and the approximations used by the models to represent the physical processes that occur in it [87]. These two factors give rise to errors, which are amplified as the forecast time range advances and propagate, in turn, in the wind generation forecast models. However, due to complex computational factors and environmental constraints, the application of physical methods is greatly limited [88]. A case study based on weather data (atmospheric pressure, air temperature, relative humidity, and precipitation) to model monthly wind speed values is needed. The study must include analyzing the relationship between these variables and wind speed and estimating the relative importance of each variable [89].

3.2.1.2 Statistical Methods

Contrary to physical models, statistical models do not need system information but are based on statistical analysis of historical data, combined with linear or non-linear mathematical equations to make predictions. Autoregressive (AR) models, autoregressive moving average (ARMA) models, and autoregressive integrated moving average (ARIMA) models have all been used to learn statistical rules for wind power time series [90-92]. These types of models require little

computation time, are easy to formulate, and give good results for predictions at short-term horizons [93], which are useful for forecasting in certain industrial processes [94], and in the context of wind power forecasting, it provides reasonably good results for horizons up to 6 hours. At the same time, support vector machines (SVMs) and time series analysis methods are also applied to wind power forecasting [95-97]. Statistical methods make competitive predictions when the state is stable by making linearizing assumptions. However, the wind power series has random and intermittent characteristics, which makes the data very complex. These shallow models cannot extract the corresponding nonlinear features well [98], so there is still room for improvement in this type of method since it is difficult for traditional methods to extract deep features from wind power data [99].

3.2.1.3 Artificial Intelligence Technology

Deep learning methods have offered complex algorithms to analyse large amounts of data and make predictions based on patterns and relationships in the data, with the continuous development of artificial intelligence technology in recent years, in order to fully mine the historical data information of wind power [100]. More and more related artificial intelligence algorithms are being applied to the field of wind power prediction. These include prediction models combining stacked autoencoder (SAE) and backpropagation methods [101], and artificial neural networks (ANN) are also used to predict wind power time series [102]. Recurrent neural networks (RNNs), due to their recurrent design, are capable of learning highly nonlinear dynamic time information from sequences, have outstanding performance in the field of natural processing, and have been applied to many other time series tasks [103, 104]. To promote RNNs performance in various fields, a large number of studies have been proposed, including one on a very popular RNN variant, LSTM [105]. However, if LSTM is directly applied to the wind power prediction task, its ability to learn complex temporal and spatial patterns from the entire time series is limited. Along the time axis, model learning can easily lose long-term dependencies due to vanishing gradients [106]. A short-term wind power prediction model based on an integrated multi-scale LSTM is proposed. The prediction models are constructed through time series of different lengths, and the prediction results are fused to a certain extent. The problem of LSTM losing information dependence due to time series being too long is mitigated to some extent while retaining long-term rich time series information in the sequence data and time series information in the short time series that is less

disturbed by noise. The deep network constructed by the LSTM unit fully mines the implicit dependencies between the data in the time series and obtains a better short-term wind power prediction result.

3.3 Forecast of Photovoltaic Energy Generation

The increasing global energy crisis, along with concerns about climate change and global warming, have led to a growing emphasis on the development of renewable energy systems, including solar energy. The growth of the global installed capacity of solar energy has been steady in recent years, largely due to the many benefits it offers, such as its ability to be installed in a wide range of locations and its abundant and renewable source of energy, the sun. However, solar power generation has the characteristics of intermittency and volatility. When more and more solar power generation is integrated into the power grid, if the instability cannot be effectively predicted and controlled, the safe and stable operation of the power grid will face significant challenges. In this way, photovoltaic energy prediction is one of the alternatives to alleviate these adversities and allow optimized dispatch [107]. In order to meet the requirements of the automatic decision-making process, prediction models rely heavily on the spatial and temporal resolution of meteorological variability, the selection of input parameters, and training algorithms. Output power control is an important aspect of managing a renewable energy system, such as a solar photovoltaic system, as it helps balance the supply and demand of electricity. The output power of a photovoltaic system can vary greatly depending on various factors, including the amount of solar radiation received, temperature, insolation, and the angle of installation. These variations are due to the dynamic nature of atmospheric conditions such as weather and temperature, which can impact the performance of the photovoltaic system over time. By adjusting the output power, grid operators can help ensure that the supply of electricity matches the demand and maintain a stable electrical grid. [108]. The literature points out that accurate predictions of photovoltaic (PV) generation can bring many benefits to the power grid, including reduced variability, improved stability, increased penetration levels, lower maintenance costs, and better management of surplus or shortages of solar energy. These benefits help to make the integration of solar PV into the grid more efficient and effective, ultimately leading to a more sustainable energy system [109]. Forecasting the output power of a solar plant is a significant challenge for electric power departments. Improper predictions can result in difficulties with timely dispatch planning, decreased flexibility in managing the grid, and higher costs due to the need for increased reserve

capacity and the rotation of generation systems [110]. Accurate forecasting of solar energy helps system operators and plant management avoid fines that may result from discrepancies between projected energy and energy produced by optimizing resource use and making informed decisions [111]. Solar energy forecasting has evolved over time, and traditional time series methods such as Auto Regressive (AR), Moving Average (MA), and Auto Regressive Moving Average (ARMA) models have been commonly used to model linear dynamic structures in photovoltaic energy output prediction. However, nonlinear and non-stationary models, such as neural networks, support vector machines, and hybrid models, have been found to be more effective at capturing the complex patterns in the data, and they are increasingly being used for energy that requires more accurate production forecasting. These models can better non-linear relationships, handle non-stationary data, and provide improved accuracy in forecasting photovoltaic energy output [112, 113]. Predicting photovoltaic power generation is actually a prediction of non-stationary data flow, and its accuracy often depends on the extraction and processing methods of time series features. In recent years, scholars have divided the prediction methods of photovoltaic power generation into learning methods [114], physical model methods [115], and statistical methods [116, 117]. With the continuous development of machine learning and deep learning [118], the use of learning methods for photovoltaic power generation power prediction has gradually become popular [119-121], among which deep learning algorithms are the mainstream.

3.3.1 Uncertainty in Forecasting

The models used for short-term electricity production predictions (up to 48 or 72 hours of prediction horizon) can use, as inputs to them, the (meteorological) predictions generated by other models, the atmospheric models. These are, generally, properties of the national meteorological institutes, which provide predictions of the numerical values of the meteorological variables: radiation, temperature, atmospheric pressure, cloud cover, wind speed, and direction. From these numerical values, the result of the prediction of atmospheric models, and the prediction of the electrical power generated in a photovoltaic solar park, is calculated by statistical or physical approximations. But the prediction of electricity production carried out by these latest models has a major limitation: the prediction models of electricity production, in principle, cannot improve the (meteorological) predictions of atmospheric models. The sources of error in atmospheric models can be various: from poor modeling of the ground surface to incorrect initial conditions as a result of errors in the measurement process or interpretation of the meteorological variables that

define the starting state of the atmosphere. The first type of error, the one caused by poor modeling of the terrain surface, can be corrected statistically. The second type of error (incorrect initial conditions) cannot be corrected in advance and is responsible for important differences between the prediction values and the real ones, both in magnitude and in the instant of time in which they occur (the so-called "phase error" which refers to the expected variations or changes in the value of the meteorological variable that occur earlier or later than expected). In general, forecasting the electrical energy generated in the short term by a solar PV collector is a difficult task that is not exempt from important errors. But despite its difficulty, it turns out to be essential today, both from a technical and economic point of view.

3.3.2 Models Based on the Box-Jenkins Methodology or ARIMA Models

Auto-Regressive Integrated Moving Average (ARIMA) models are, without a doubt, the most widely used linear method to identify and build prediction models for univariate and stationary time series [122]. They have been used in numerous research areas; however, they have only recently been used to develop prediction models regarding photovoltaic solar production. Generally, the results obtained using these models are better the shorter the prediction horizon, although acceptable results are also obtained in the calculation of parameters that are less restrictive than the average hourly power produced, such as the average daily or monthly power. Several ARIMA models were compared with other more advanced techniques for different prediction horizons (5, 15, 30, and 60 min) using the global radiation at the moment before the prediction. The results reflect how it can be seen that the variability of meteorological conditions and cloud cover causes a non-seasonal behavior that makes its application difficult over longer horizons, but that in the very short term, the mastery of the behavior of the 24-hour cycles justifies the validity of auto-regressive models. Finally, in most of its tests, the smallest errors are made with the ARIMA models, reinforcing their validity for very short-term horizons [123]. A series of ARIMA models were developed to determine the average radiation (horizontal and inclined) with horizons of 10, 20, 30 minutes, and 1 hour for a given location in the United Kingdom. The results show that the efficiency of ARIMA decreases significantly with the lengthening of the prediction horizon [124].

3.3.3 Numerical Weather Prediction (NWP)

The advancements in computing power and technology have greatly influenced the relationship between meteorological conditions and power grid operation. The construction of large-scale high-

voltage and ultra-high-voltage power grids has also increased the need for accurate weather forecasting to prevent power outages and ensure stable grid operation. Numerical weather forecasting (NWP) has become increasingly important due to the close relationship between weather and the operation of power grids. NWP uses mathematical models and numerical algorithms to predict weather conditions by solving the equations that describe the evolution of various atmospheric variables such as radiation, temperature, wind speed and direction, humidity, and pressure [125]. The process starts with the analysis of the current state of the atmosphere using a very short-term forecast and adding the observations available to it in order to achieve a better description of the true current state of the atmosphere. From there, the computer model uses this information as the starting point to produce the forecast. The equations used by the computer model are based on the fundamental physical laws that govern the atmosphere and which are described in non-mathematical terms (advection, conservation of mass, continuity, hydrostatic, thermodynamic, state, and water vapor). The accuracy of numerical weather forecasting (NWP) depends on the assumptions and approximations used in the mathematical models and the numerical algorithms that are used to solve these equations. Different numerical models make different assumptions about the atmospheric processes and use different methods for solving the equations, leading to variations in the predicted weather conditions [126]. Not all atmospheric models work on the same space-time scale. Each atmospheric model tries to follow the evolution of the atmosphere at the scale that defines it. For example, for the prediction of radiation on the ground surface, the scale of the atmospheric model must be small: small spatially due to the influence that orography has on radiation, and small temporally due to the variability of the cloud cover (linked to the variability of wind). Numerical weather forecasting is still one of the most important projects in solar photovoltaic forecasting work. If the accuracy of numerical weather forecasts can be improved, the forecast error of solar photovoltaics will be effectively reduced. The variability of renewable energy sources, like solar and wind power, can make it challenging to integrate them into the grid in a reliable and consistent manner. This is because their output is dependent on weather conditions, which can be unpredictable and subject to sudden changes. To mitigate this issue, accurate forecasting methods are essential for effective grid integration [127]. There are several forecasting models for predicting future radiation, including those that include artificial intelligence technologies. One of the biggest problems that solar energy faces is climate variability, but artificial intelligence could obtain more accurate predictions and apply measures

so that photovoltaic installations perform to their greatest potential. In the electric power generation sector, machine learning techniques have gained great importance in the estimation of meteorological attributes from structured, semi-structured, and unstructured data sets. Although many of the stations are not taking solar brightness readings, there are some that do measure it, along with other climatic variables, in order to gain a better understanding of the environment. It was found that most modern researchers use deep learning to estimate solar radiation, as reported in preliminary work [128-130]. Depending on the type of input applied to the artificial intelligence model, forecasting techniques can be roughly classified as physical or statistical approaches. Physical methods model photovoltaics as a function of some independent variables, such as the characteristics of the photovoltaic cells, solar radiation, and the temperature of the cells. While statistical forecasting techniques base their predictions on historical data from the past, the artificial neural network algorithm is one of the most popular artificial intelligence algorithms for predicting solar production from meteorological factors since the artificial neural network method has an inherent ability to simulate non-linear, dynamic, and complex systems data [131]. A model based on networks called large short-term memory (LSTM) was presented for the prediction of solar irradiation for the next day using meteorological data. As inputs to the model, temperature, humidity, visibility, wind speed, weather type, and dew point are considered. The performance of this model was better than other models based on linear least square regression (LLSR) and feedforward neural networks (FFNN), with a root mean square error (RMSE) of less than 18.34% [132].

3.4 Related Works

According to one viewpoint, the development of green innovation necessitates good general framework conditions for innovation [133]. The analysis of time series data is fundamental for the control and management of electricity generation with renewable energy sources. Currently, renewable energy conversion systems are the main drivers for the development of time series analysis [134]. A time series is important to make predictions of future data and, with that, to carry out planning on how much energy it will be possible to dispatch in a given period and thus make more efficient management of the wind or solar power. Auto-Regressive Integrated Moving Average (ARIMA), seasonal ARIMA (SARIMA), and long short-term memory (LSTM) were

used to predict energy consumption in a daily database, and the results obtained show that the model LSTM shows better results [135].

Therefore, time series analysis of wind speed and solar irradiance data is extensively studied in order to maximize energy conversion. However, to predict future data, not only is the historical value of the objective variable sufficient to carry out a reliable analysis, but other factors with an implicit influence on the objective variable must also be considered [136]. In the analysis of time series problems, the LSTM method, which has been preferred by researchers in recent years, has provided significant improvements in the estimation consistency of complex and nonlinear data. The LSTM approach has been preferred in the literature due to the high accuracy of estimation in environmental time series problems, such as estimation of air pollution, water quality, and ozone gas density [137, 138]. Because wind speed is affected by temperature, solar radiation, humidity, and its direction, in addition to topographic conditions, when conducting analysis with many variables that are correlated with each other, the accuracy in forecasting objective variables improves due to the intrinsic correlation and historical characteristics between them [139]. The non-stationary and highly nonlinear characteristics of solar radiation time series make the artificial intelligence method better than the traditional statistical analysis method [140, 141]. Deep neural networks were used to estimate solar radiation values [142]. Artificial neural networks (ANNs) have been applied to obtain accurate short-term load predictions in photovoltaic (PV) systems and to investigate different levels of spatial aggregation. The results show that the energy flow underscores the benefits of time-series forecasting to support network operation [143]. The initial statistical analysis of wind speed appeared in the 1940s. In the last century, a research program was carried out in the United States, the purpose of which was to study potential sites for the construction of wind energy conversion systems. The first attempts to determine the mean hourly wind speed were based on Monte Carlo simulations, in which the expected wind speed distribution was given. However, because it did not account for the relationship between successive wind speed observations, the resulting prediction was inaccurate. Consequently, large periods of high and low wind speeds did not appear often enough. Further studies have attempted to use autocorrelation in the wind speed model [144, 145]. However, the approaches used in these studies were based on specific statistical assumptions inherent in wind speed data. Some researchers also neglected such important features as daily and seasonal fluctuations and the non-Gaussian shape of the wind speed distribution. A method has been proposed that included all the main characteristics of hourly

average wind speed, but it was applied to the study of only one case. To apply a true autoregressive model, it was necessary to use a power transformation. Thus, the non-Gaussian wind speed distribution was taken into account [146]. Other types of power transformations have been proposed, and data modeling has already been carried out using one of the autoregressive moving average models [147- 150]. It should be noted that these papers described short-term forecasts, while models with long-term memory were used to obtain reliable forecasts for the next 24-48 hours [151]. However, the analysis was limited to only four time series. LSTM networks have been used not only to make long-term predictions but also for abrupt changes (in less than 15 minutes) [152]. Modeling energy production is just as important as managing energy properly. As a result, deep learning models have been used to forecast power generation [153], as well as renewable energy resources such as wind [154], and solar [155, 156]. A statistical forecast analyzes relationships between variables based on historical information. One of the statistical methods that have been most successful in generating forecasts is that generated from artificial neural networks (ANN) [157]. Several ARIMA models are compared with other more advanced techniques to predict solar radiation with a short-term prediction horizon [158]. The performance of artificial neural networks was compared with ARIMA, and the results concluded that the ARIMA model is better for short-term periods (10 minutes, 1 hour, 2 hours, and 4 hours) [159]. Artificial neural networks have been compared to the seasonal ARIMA (SARIMA). The result was that the SARIMA model is better at predicting wind speeds and has a better response to structural changes in the wind regime [160]. An analysis of the energy consumption dataset was developed to generate a prediction model using ARIMA [161]. The ARIMA algorithm was used to predict radiation in six cities. The results show that there are advantages in the performance of ARIMA over ANN in the case of data time-steps greater than 5 minutes; otherwise, in high data resolutions (between 1 minute and 4 minutes), ANN presents better behavior. This leads to the conclusion that the choice of model depends directly on the input dataset. ARIMA better models the daily radiation cycle in long-term datasets. This means that a training stage is not necessary due to the linear behavior of the data in the database. The computational cost is also evaluated in these works. It is shown that ARIMA has a lower computational cost, in part due to the linear approximations with which it develops the estimate [162, 163]. Regarding short-term wind energy forecasts using statistical methods, the ARMA model was applied to wind speed, and the results showed that the time-series data and the model were successful in the 1-hour predictions and a set of wind farm data [164]. A

method based on the ARIMA model has been proposed, and the results were compared with persistence methods [165]. The ARMA model has been used for short-term wind energy forecasting, and successful results have been obtained [166]. In order to forecast wind speed, a SARIMA model was developed, and the model was evaluated using three example datasets. The results showed that the prediction error one hour before the prediction was less than 16%, which is an acceptable result considering the short data collection time [167]. Hourly wind speeds were predicted using both the ARMA and ANN models. The result of the comparison between them showed that their performance is very close, and both models give results with prediction accuracy up to a few hours ahead [168]. A LSTM model has been proposed to predict PV power. When the proposed model is compared to two other models, it is found to have the best performance, with an RMSE value of less than 21% [169]. The LSTM model combined with a deep neural network was used to predict the load and the PV energy generated in a smart network. The performance of the model is compared to that of other models, and a satisfactory result is obtained with a MAPE value of 7.43% [170]. An ultra-short-term wind power prediction model based on a long short-term memory (LSTM) network was proposed. Because the LSTM network has short-term memory ability, it can reflect the impact of the output value of the past moment on the value of the current moment. Compared with ARMA, ANN, SVM, and other models, its prediction accuracy is significantly improved [171]. The LSTM method was used to estimate the hourly solar radiation values, and the weather forecast values for the same time period were used as input data [172]. Wind power forecasting remains challenging due to the variability and unpredictability of weather conditions, making it necessary to constantly improve and refine forecasting methods [173]. ANN and ARIMA models are commonly used for short-term wind speed forecasting. Both methods have shown good results in short-term wind speed forecasting, but their accuracy may decrease over longer forecasting horizons [174]. ARIMA is a statistical method used to model and forecast time series data; it has proven to be an effective method for short-term wind speed prediction and is widely used [175]. In a study on the iterative wavelet transform based ARIMA model for very short-term wind speed prediction, the lack of ARIMA and WT-ARIMA models, which are recently popular techniques for short-term and very short-term estimation of wind speed [176], was investigated. Long- short term memory-based deep learning architectures were applied for wind energy prediction [177,178]. Studies have indeed shown that LSTMs often outperform traditional deep neural networks in terms of forecasting accuracy for wind energy. However, as with any

forecasting method, the results may vary depending on the specific data set and problem at hand, and it may be necessary to experiment with different methods to find the best fit [179]. For wind energy time series forecasting, LSTM, one of the most popular models among RNNs, has been proposed and compared to the ARIMA model [180]. According to the comparisons between the performances of ARIMA, ARIMAX, and simple LSTM models for wind energy forecasting, the ARIMAX model can perform similarly or even better than simple LSTM models, which are in the lead, followed by the ARIMAX model and then the ARIMA model, respectively. Model selection is influenced by a number of factors, such as computing complexity, interpretability, and the ability to handle non-static and non-linear data [181]. There are several statistical measures used to determine the accuracy of the ARIMA model for forecasting wind speed and how well it is able to predict future values based on past data [182]. The LSTM-RNN model was applied to forecast wind power from 1 to 24 hours ahead, using past wind power measurements as input and forecasting future wind power output based on this historical data [183]. Recurrent neural network models have been evaluated for the accuracy of wind speed prediction, and compared with ARIMA models at the univariate and multivariate levels, RNN achieved better results [184]. Using historical wind speed data and auto-regression for a two-year period, the Long-Term Memory (LSTM) model can be used to forecast short-term spatiotemporal wind speed for multiple locations. The use of a numerical weather prediction (NWP) model to update the forecast every six hours can further improve the accuracy of the predictions, especially for horizons as long as fifteen days using LSTM for two or three hours [185]. Mathematical modeling has been widely used in renewable energy systems, specifically for solar forecasting, and has shown improved results compared to traditional statistical methods [186]. In general, LSTMs are powerful models for sequential data and can handle noise well [187]. Deep learning algorithms have shown great results for classification and regression issues. This is due to supervised learning, in which the model is trained on labeled data and the internal parameters are adjusted to minimize the prediction error through backpropagation and gradient descent [188]. The application of artificial intelligence to various aspects of solar energy systems remains an active area of research and development and has recently been used in modeling, simulation, and control processes to improve the efficiency of solar panel systems [189]. The LSTM model showed better performance compared to the ARIMA model in predicting the use of solar energy. The data included various factors such as energy structure, energy intensity, economic activity, and population, which have an impact on changes

in solar energy consumption [190]. LSTM is a type of artificial intelligence that has gained popularity in recent years as part of the third AI boom. It is commonly used in short-term predictions of wind speed due to its ability to store and use previous information to make predictions [191, 188]. LSTM networks have been shown to be effective in predicting wind speeds in 24-hour wind farms, according to some studies. LSTM results showed more accurate performance compared to traditional methods such as MLP (multilayer perceptrons) and other neural networks in many sequence prediction tasks [192]. Deep learning based on neural networks is being actively introduced for solar radiation applications, and some studies have shown that the combination of multilayer perceptron (MLP) and linear regression can result in better performance in forecasting solar energy compared to some other methods. According to these studies, the MLP model achieved a high coefficient of determination (R^2) of 97.7% and a low root mean square error (RMSE) of 0.033, which indicates its effectiveness in this application [193]. The LSTM-RNN has been proven to be a more effective method for predicting the output power of solar PV systems using hourly data sets over the course of a year. Compared to multiple linear regression (MLR), bagged regression trees (BRT), and neural networks (NNs) methods, the LSTM networks provide lower predictive error [194]. LSTM neural networks are commonly used in the medium to forecast time series data, such as solar and wind energy. It recorded lower errors compared to SVM and persistence models [195]. When compared to traditional methods, the use of artificial neural networks, specifically the multilayer perceptron (MLP) structure, has proven to be an effective tool in predicting solar radiation. MLP is a feedforward neural network that is widely used in various applications, including regression and classification tasks [196].

CHAPTER 4

Time Series Analysis Techniques and Methodologies

4.1 Introduction

The analysis of time series data is fundamental for the control and management of electricity generation with renewable energy sources. Currently, renewable energy conversion systems are the main drivers for the development of time series analysis [197]. A time series is important to make predictions of future data and, with that, to carry out planning of how much energy it will be possible to dispatch in a given period and, thus, make more efficient management of the wind or photovoltaic park in question [198]. Therefore, time series analysis of wind speed and solar irradiance data is extensively studied in order to maximize energy conversion. However, to predict future data, not only is the historical value of the objective variable sufficient to carry out a reliable analysis, but other factors with an implicit influence on the objective variable must also be considered [199]. Wind speed can be influenced by temperature, solar radiation, humidity, and direction, in addition to topographic conditions [199]. When analyzes are performed with several variables that are correlated with each other, the accuracy in forecasting the objective variables is improved due to the intrinsic correlation and the historical characteristics between them [200]. Several works that do time series analyses to forecast future data have been published in recent years. Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), and Long Short-Term Memory (LSTM) were used to predict energy consumption in a daily database, and the results obtained show that the model LSTM shows better results [198]. An analysis of an energy consumption dataset was developed to create a forecasting model using ARIMA [201]. Artificial Neural Networks (ANNs) have been proposed for accurate short-term load predictions in photovoltaic (PV) systems and to investigate different levels of spatial aggregation. The results show that power flow confirms the benefits of time-series prediction to support network operation. For intelligent control and management of wind turbines, it is important to obtain wind speed information in advance to allow stable operation of the power system [202].

4.2 Statistical Analysis of Time Series and Their Components

A time series is a sequence of observed values of some variable produced at regular intervals of time. If we take the length of such an interval as a unit of time (year, quarter, month, day), then it can be assumed that consecutive observations $X_1, X_2, \dots, X_n$ are made at moments $t = 1,2, \dots, n$ [202]. The introduction of a time scale into a time series significantly distinguishes it from a simple (random) sample of statistical data. The key feature of the time series is the binding of values (measurements) to the corresponding moments of time. Time series analysis pursues two main goals: determining its structure (nature) and forecasting future values of the series based on current and past measurements [204]. Both are closely related. Solving the first task is necessary for building a mathematical model of the time series, its correct identification, and its formalization. The mathematical model will become, in a way, a laboratory for the study of time series and a foundation for relatively accurate (with an acceptable margin of error) predictions for the series. The time series has four main components: trend, seasonality, cyclicality, and irregular components. These characteristics distinguish a time series from ordinary observed data [205]. Linear trend estimation is a statistical method that is used to understand the structure of data by fitting a straight line to the data points (Figure 4.1. An example of a growing trend). The goal of linear trend estimation is to identify any underlying patterns or trends in the data, such as increasing or decreasing trends over time, which can then be used to make predictions about future data points. This method is widely used in various fields to forecast future trends and make informed decisions based on the data. The simplicity and interpretability of linear trend estimation make it a popular choice for analyzing and predicting time series data [204].

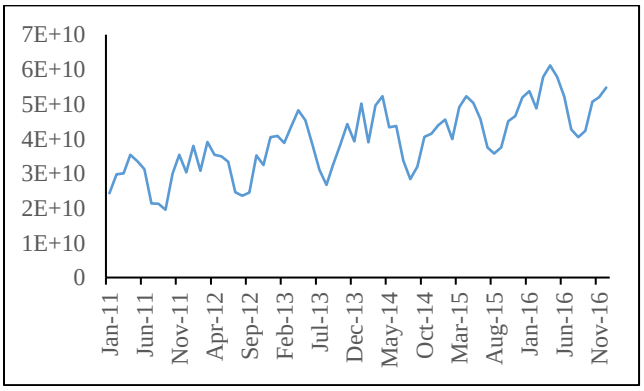

Figure 4.1. Actual U.S. Wind Power Monthly Consumption from January 2011 to December 2016 (An example of a growing trend).

A good understanding of the data prior to model training is essential for building an appropriate model. The amount of labeled data used for training also has a significant impact on determining the quality of the model. In general, more labeled data leads to a better model. A time series, represented as a simple proportional function, can be plotted on a graph with time on the x-axis and the value of the series on the y-axis. This is the easiest way to start the analysis of a time series and allows for the discovery of the most important characteristics and components of a series. Given a set of time instants *t* and data values observed for those time instants, the values of *a* and *b* are chosen such that:

$$\sum_t [y_t - (\hat{a}t + \hat{b})]^2 \tag{4.1}$$

In this case, the trend line is located at + b, thus the objective is to minimize the sum of squared deviations between the observed data points and the trend line represented by the model, so this optimization problem can always be solved in closed form [10]. Time series analysis aims to study patterns and trends in data collected over time with the goal of understanding and predicting future behavior. Time series analysis considers various factors such as seasonality, trend, and fluctuations to identify the underlying structure of the data and make predictions. The technique is widely used in several fields. The seasonal component of the additive time series model is calculated as follows:

$$S = Y - (T + C + I) \tag{4.2}$$

Where, *S* is the seasonal values, *Y* is the real values of time series data, *T* is the trend values, *C* is the cyclical values, and *I* is the irregular values. Figure 4.2 shows the decomposition of additive time series.

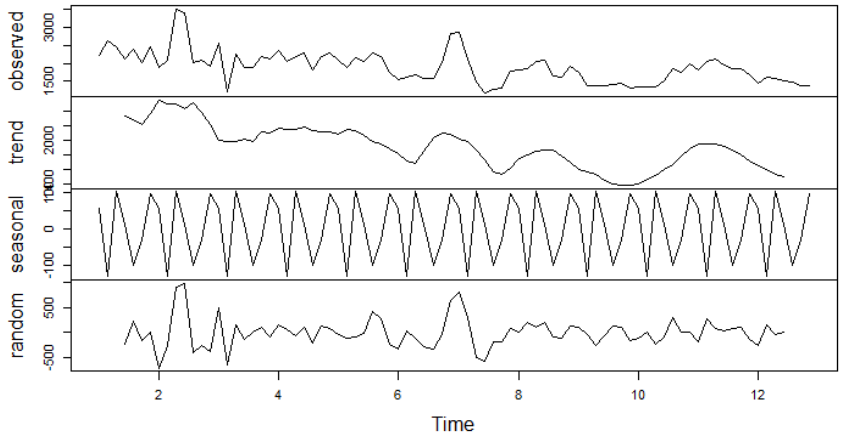

Figure 4.2. Decomposition of additive time series

The cyclical component is reflected in non-seasonal fluctuations in the data. The duration of these fluctuations is usually at least two years. The cyclical behavior differs from the seasonal behavior in that if the oscillations do not have a definite period, then they are cyclical; if the oscillations do not have a fixed period, then they are cyclical; if the period is fixed and related to some aspect of the calendar, then the model is seasonal. Irregular time series store timestamps for each element rather than storing offsets because the interval between each element can be of different lengths. Invalid elements are stored as the next element by default and cannot be null. Modern time series analysis has evolved beyond traditional linear methods and now incorporates both statistical and dynamic techniques to analyze the complex and nonlinear behavior of systems that were previously not possible. The basis of stochastic data processing methods is the series y_n and the noise is a sequence of uncorrelated, equally distributed random variables ξ_i with zero mean. After that can write:

$$y_n = F(y_{n-1}, \dots, y_{n-m}, \xi_n, \dots, \xi_{n-k}) \tag{4.3}$$

Where, *k* and *m* are some finite numbers.

4.3 Mathematical Models of Time Series

4.3.1 Autoregressive Forecasting Models

Autoregressive is a time series model in which the current value linearly depends on the previous values of the same series, X_t, can be explained in terms of *p* past values $X_{t-1}, X_{t-2}, \dots, X_{t-p}$, where *p* determines the number of lags needed to predict a current value. The moving average (MA) and autoregressive (AR) models are widely used for time series analysis and prediction. The MA model uses the past values of the time series to model the error term and make predictions about the future values. Further, the AR model uses the past values of the time series itself as explanatory variables to make predictions about the future values. Both models can be used to identify various features of the time series, such as trends, seasonality, and irregular behavior. The characteristics of the time series and the goals of the analysis determine whether to use the MA or AR models [206]. The autoregressive model of order (*p*) is given by:

$$X_t = c + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \varepsilon_t \tag{4.4}$$

Expressed in terms of the lag operator

$$\left(1 - \emptyset_1 L - \emptyset_2 L^2 - \cdots - \emptyset_p L^p\right) X_t = \varepsilon_t \tag{4.5}$$

$$\emptyset_p(L)X_t = \varepsilon_t \quad (4.6)$$

Where, ε_t is the white noise process and $\emptyset_1, \emptyset_2, \dots, \emptyset_p$ are the model parameters [207].

4.3.1.1 Autoregressive Process of Order 1: AR(1)

In processes, AR(1) the variable X_t is determined only by the passed value, that is X_{t-1}.

$$X_t = \emptyset X_{t-1} + \varepsilon_t \quad (4.7)$$

Where, ε_t is a white noise process with mean zero, and constant variance σ^2 and $\emptyset$ is the parameter. Assuming also that the process is unpredictable, that is, the future does not include the past. To verify that the model AR(1) is stationary for any value of the parameter, it is necessary to test the following conditions.

a) Stationary in mean

$$E(X_t) = E(X_t = \emptyset X_{t-1} + \varepsilon_t) = \emptyset E(X_{t-1}) \quad (4.8)$$

For the process to be stationary the mean must be constant and finite in time, this implies that:

$$E(X_t) = \emptyset E(X_t) \quad (4.9)$$

$$(1 - \emptyset)E(X_t) = 0 \quad (4.10)$$

$$E(X_t) = \frac{0}{1 - \emptyset} = 0 \quad (4.11)$$

Therefore, for the process to be stationary, the parameter $\emptyset \neq 1$.

b) Stationary in covariance:

For a process AR(1) to be stationary, the variance has to be constant and finite in time:

$$\gamma_0 = E(X_t - E(X_t))^2 = E(\emptyset X_{t-1} + \varepsilon_t - 0)^2 = \emptyset^2 V(X_{t-1}) + \sigma^2 \quad (4.12)$$

Given the autocorrelation of the process

$$E(X_{t-1}\varepsilon_t) = E[(X_{t-1} - 0)(\varepsilon_t - 0)] = cov(X_{t-1}\varepsilon_t) = 0 \quad (4.13)$$

Under the assumption that the process is stationary,

$$E(X_{t-1})^2 = V(X_{t-1}) = V(X_t) = \gamma_0 \quad (4.14)$$

Therefore $\gamma_0 = \emptyset\gamma_0 + \sigma^2$, then $\gamma_0 = \frac{\sigma^2}{1-\emptyset^2}$

For a process to be stationary, constant variance and finite, it is necessary that $|\emptyset| < 1$

The order autocovariance function *k* is:

$$\gamma_k = E\big(X_t - E(X_t)\big)\big(X_{t-k} - E(X_{t-k})\big) = E[(\emptyset X_{t-1} + \varepsilon_t)X_{t-k}] \quad (4.15)$$

$$\gamma_k = \emptyset E(X_t X_{t-k}) + E(\varepsilon_t X_{t-k}) = \emptyset\gamma_{k-1} \quad (4.16)$$

So that:

$$\gamma_1 = \emptyset\gamma_0$$

$$\gamma_2 = \emptyset\gamma_1$$

$$\gamma_3 = \emptyset\gamma_2$$

$$\vdots$$

It can be concluded, therefore, that the process AR(1) is stationary if and only if $|\emptyset| < 1$. The autocovariance function of a stationary process AR(1) is:

$$\gamma_k = \begin{cases} \frac{\sigma^2}{1-\emptyset^2} & k = 0 \\ \emptyset\gamma_{k-1} & k > 1 \end{cases} \quad (4.17)$$

The autocorrelation coefficients of a stationary process $AR(1)$ are:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\emptyset\gamma_{k-1}}{\gamma_0} = \emptyset\rho_{k-1} \quad (4.18)$$

The autocorrelation function of a stationary process AR(1) is:

$$\rho_k = \begin{cases} 1 & k = 0 \\ \emptyset\rho_{k-1} & k > 1 \end{cases} \quad (4.19)$$

It can be easily shown that the Autocorrelation Function of a model AR(1) is an exponential function.

$$\rho_1 = \emptyset\rho_0 = \emptyset$$

$$\rho_2 = \emptyset\rho_1 = \emptyset^2$$

$$\rho_3 = \emptyset\rho_2 = \emptyset^3$$

$$\vdots$$

So that:

$$\rho_k = \emptyset^k, where\ k = 1,2,3, ... \tag{4.19}$$

An alternate way of writing the model AR(1) is as follows:

$$X_t = \emptyset X_{t-1} + \varepsilon_t \tag{4.20}$$

$$(1 - \emptyset L)X_t = \varepsilon_t \tag{4.21}$$

The quotient $\frac{1}{1-\emptyset L}$ can be expressed as an infinite polynomial, that is:

$$\frac{1}{1-\emptyset L} = 1 + \emptyset L + \emptyset^2 L^2 + \cdots \tag{4.22}$$

Substituting we have

$$X_t = \frac{1}{1-\emptyset L}\varepsilon_t = (1 + \emptyset L + \emptyset^2 L^2 + \cdots)\varepsilon_t \tag{4.23}$$

$$X_t = \varepsilon_t + \emptyset\varepsilon_{t-1} + \emptyset^2\varepsilon_{t-2} + \emptyset^2\varepsilon_{t-2} + \emptyset^3\varepsilon_{t-3} + \cdots \tag{4.24}$$

Therefore, model AR(1) is a constrained version of a general moving average model. From the results obtained, we can point out the most relevant characteristics of the model AR(1): it is stationary only if the autoregressive coefficient $\emptyset$ is less than one in absolute value ($|\emptyset| < 1$). If this condition is not met, the model is non-stationary, meaning that its statistical properties change over time. Another important feature of the AR(1)model is that the correlogram, which is a plot of the autocorrelation function (ACF) against the lag, will decay exponentially towards zero with all positive values if $\emptyset > 0$ or it will alternate signs, starting with a negative value, if $\emptyset < 0$. In addition, the model is invertible if its ACF has no spikes outside the confidence bounds, which means that the model can be expressed as a finite weighted sum of past and current error terms. The partial autocorrelation function (PACF) of the AR(1) model is zero for all lags greater than one, which coincides with the autoregressive coefficient of the model. This means that the current

value of the time series is only related to its immediate past value and not to any other previous value.

4.3.1.2 Autoregressive Process of Order 2: AR(2)

In processes, $AR(2)$ the variable X_t is determined only by the passed value, that is.

$$X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \varepsilon_t \tag{4.25}$$

Where, ε_t is white noise.

Assuming stationarity, the characteristics of the process are:

a) Mean

$$E[(1 - \emptyset_1 - \emptyset_2 L^2)X_t] = E[\varepsilon_t] \tag{4.26}$$

$$(1 - \emptyset_1 L - \emptyset_2 L^2)E[X_t] = 0 \tag{4.27}$$

$$E[X_t] = \frac{0}{1 - \emptyset_1 L - \emptyset_2 L^2} = 0 \tag{4.28}$$

b) Autocovariance function:

$$\gamma_0 = E(X_t - E(X_t))^2 = E(X_t)^2 = E(\emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \varepsilon_t)^2 \tag{4.29}$$

$$\gamma_0 = \emptyset_1^2 \gamma_0 + \emptyset_2^2 \gamma_0 + \sigma^2 + 2\emptyset_1 \emptyset_2 \gamma_1 \tag{4.30}$$

$$(1 - \emptyset_1^2 - \emptyset_2^2)\gamma_0 = \sigma^2 + 2\emptyset_1 \emptyset_2 \gamma_1 \tag{4.31}$$

$$(1 - \emptyset_1^2 - \emptyset_2^2)\gamma_0 = \sigma^2 + 2\emptyset_1 \emptyset_2 \gamma_1 \tag{4.31}$$

$$\gamma_0 = \frac{\sigma^2 + 2\emptyset_1 \emptyset_2 \gamma_1}{1 - \emptyset_1^2 - \emptyset_2^2} \tag{4.32}$$

$$\gamma_1 = E\big(X_t - E(X_t)\big)\big(X_{t-1} - E(X_{t-1})\big) = E(X_t X_{t-1}) \tag{4.33}$$

$$\gamma_1 = E[(\emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \varepsilon_t)X_{t-1}] \tag{4.34}$$

$$\gamma_1 = \emptyset_1 E(X_{t-1})^2 + \emptyset_2 E(X_{t-2} X_{t-1}) + E(\varepsilon_t X_{t-1}) \tag{4.35}$$

$$\gamma_1 = \emptyset_1 \gamma_0 + \emptyset_2 \gamma_1 \tag{4.36}$$

$$\gamma_1 = \frac{\emptyset_1 \gamma_0}{1 - \emptyset_2} \tag{4.37}$$

γ_0 and γ_1 give the first two autocovariances as a function of the parameters $\emptyset_1$ and $\emptyset_2$, and σ^2 the variance of the white noise.

The autocovariances of order k, for everything $k > 1$ are:

$$\gamma_k = E\big(X_t - E(X_t)\big)\big(X_{t-k} - E(X_{t-k})\big) = E(X_t X_{t-k}) \tag{4.38}$$

$$\gamma_k = E[(\emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \varepsilon_t) X_{t-k}] \tag{4.39}$$

$$\gamma_k = \emptyset_1 \gamma_{k-1} + \emptyset_2 \gamma_{k-2} \tag{4.40}$$

The autocovariance function of a model $AR(2)$ is:

$$\rho_k = \begin{cases} \gamma_0, & k = 0 \\ \gamma_1, & k = 1 \\ \emptyset_1 \gamma_{k-1} + \emptyset_2 \gamma_{k-2}, & k > 1 \end{cases} \tag{4.41}$$

c) Autocorrelation coefficient

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\emptyset_1 \gamma_{k-1} + \emptyset_2 \gamma_{k-2}}{\gamma_0} = \emptyset_1 \rho_{k-1} + \emptyset_2 \rho_{k-2} \; for \; k = 1,2,3, \ldots \tag{4.42}$$

A general way of writing the autocorrelation coefficients is:

$$\rho_k = \begin{cases} 1 & k = 0 \\ \rho_k = \emptyset_1 \rho_{k-1} + \emptyset_2 \rho_{k-2}, & k > 0 \end{cases} \tag{4.43}$$

Stationarity conditions

Model AR(1): $X_t = \emptyset X_{t-1} + \varepsilon_t, then \; (1 - \emptyset L) X_t = \varepsilon_t$

Autoregressive polynomial: $\emptyset_1(L) = 1 - \emptyset L, the \; roots \; of \; 1 - \emptyset L = 0$ are:

$$L = \frac{1}{\emptyset} \tag{4.44}$$

The stationarity condition of the model AR(1) is:

$$|L| = \left|\frac{1}{\emptyset}\right| > 0, \text{ then } |\emptyset| < 1 \tag{4.45}$$

Model $AR(2)$: $X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \varepsilon_t, then \; (1 - \emptyset_1 L - \emptyset_2 L^2) X_t = \varepsilon_t$

Autoregressive polynomial: $\emptyset_2(L) = 1 - \emptyset_1 L - \emptyset_2 L^2, the \; roots \; of \; 1 - \emptyset_1 L - \emptyset_2 L^2 = 0$ are:

$$L_1, L_2 = \frac{\emptyset_1 \mp \sqrt{\emptyset_1^2 + 4\emptyset_2}}{-2\emptyset_2} \tag{4.46}$$

The stationarity condition of the model AR(2) is:

$$|L_1| = \left|\frac{\emptyset_1+\sqrt{\emptyset_1^2+4\emptyset_2}}{-2\emptyset_2}\right| > 1, \text{ and } |L_2| = \left|\frac{\emptyset_1-\sqrt{\emptyset_1^2+4\emptyset_2}}{-2\emptyset_2}\right| > 1 \quad (4.47)$$

If $\sqrt{\emptyset_1^2 + 4\emptyset_2} > 0$, the roots are real, and are complex when < 0.

4.3.2 Integrated Process *I*(*d*)

In time series analysis, it is often observed that many real-world time series exhibit non-stationarity, meaning that their statistical properties change over time. One way to deal with non-stationarity is to difference the time series, which involves computing the differences between consecutive observations. This transformation can stabilize the mean of the series and make it stationary. A time series that has been differentiated d times to achieve stationarity is called an integrated process of order d, denoted by *I(d).* The ARMA (*p, q*) model can be applied to this differentiated series to give rise to the ARIMA (*p, d, q*) model.

4.3.3 Moving Average Process $\boldsymbol{MA}(\boldsymbol{q})$

The moving average (MA) model models the link between an observation and the residual errors from a moving average model applied to lagged observations. The order, *q*, stands for the number of lagged errors used in the model and is usually estimated using the time series ACF (autocorrelation function) plot. The MA(q) method can be expressed in the following way:

$$X_t = \theta_0 - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q} - \varepsilon_t \quad (4.48)$$

Expressed in terms of the delay factor polynomial, which is:

$$X_t = \left(1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q\right)\varepsilon_t \quad (4.49)$$

$$X_t = \theta_q(L)\varepsilon_t \quad (4.50)$$

Where, ε_t is a white noise error term, and $\theta_1, \theta_2, \ldots, \theta_q$ are the model parameters.

Moving Average Process of order 1: MA(1)

The moving average models determine the value of X_t depending on the current innovation and its first lag, that is:

$$X_t = \varepsilon_t - \theta_1\varepsilon_{t-1} \quad (4.51)$$

It is expressed as a polynomial function of the delay factor, which is:

$$X_t = (1 - \theta)\varepsilon_t \quad (4.52)$$

$$X_t = \theta_1(L)\varepsilon_t \tag{4.53}$$

Where, ε_t is a white noise process and θ is the parameter.

a) Stationary in mean

$$E(X_t) = E(\varepsilon_t - \theta\varepsilon_{t-1}) \tag{4.54}$$
$$E(X_t) = E(\varepsilon_t) - \theta E(\varepsilon_{t-1}) \tag{4.55}$$
$$E(X_t) = 0 \tag{4.56}$$

It is stationary on the mean for all values of the parameter.

b) Stationary in covariance

$$\gamma_0 = E(X_t - E(X_t))^2 = E(X_t)^2 = E(\varepsilon_t - \theta\varepsilon_{t-1})^2 \tag{4.57}$$
$$\gamma_0 = E(\varepsilon_t)^2 + \theta^2(\varepsilon_{t-1})^2 - 2\theta E(\varepsilon_t\varepsilon_{t-1}) = \sigma^2 + \theta^2\sigma^2 - 0 \tag{4.58}$$
$$\gamma_0 = (1+\theta^2)\sigma^2 < \infty \tag{4.59}$$

The autocovariance for γ_1 and γ_2 is:

$$\gamma_1 = E\big(X_t - E(X_t)\big)\big(X_{t-1} - E(X_{t-1})\big) = E(\varepsilon_t - \theta\varepsilon_{t-1})\,(\varepsilon_{t-1} - \theta\varepsilon_{t-2}) \tag{4.60}$$
$$\gamma_1 = E(\varepsilon_t\varepsilon_{t-1}) - \theta E(\varepsilon_{t-1})^2 - \theta E(\varepsilon_t\varepsilon_{t-2}) + \theta^2 E\,(\varepsilon_{t-1}\varepsilon_{t-2}) = -\theta\sigma^2 < \infty \tag{4.61}$$
$$\gamma_2 = E\big(X_t - E(X_t)\big)\big(X_{t-2} - E(X_{t-2})\big) = E(\varepsilon_t - \theta\varepsilon_{t-1})\,(\varepsilon_{t-2} - \theta\varepsilon_{t-3}) \tag{4.62}$$
$$\gamma_2 = E(\varepsilon_t\varepsilon_{t-2}) - \theta E(\varepsilon_{t-1}\varepsilon_{t-2}) - \theta E(\varepsilon_t\varepsilon_{t-3}) + \theta^2 E\,(\varepsilon_{t-1}\varepsilon_{t-3}) = 0 \tag{4.63}$$

A general form of the autocovariance function is:

$$\rho_k = \begin{cases} \gamma_0 = (1-\theta^2)\sigma^2 & k = 0 \\ \gamma_1 = -\theta\,\sigma^2 & k = 1 \\ \gamma_2 = 0 & k > 1 \end{cases} \tag{4.64}$$

The autocovariance function is finite and depends only on k but not on time, this for any value of the parameter θ. This implies that it is not necessary to put constraints on the parameter θ so that the $\mathrm{MA}(1)$ be stationary.

The autocorrelation function of a process $\mathrm{MA}(1)$ is:

$$\rho_k = \begin{cases} 1 & k = 0 \\ \dfrac{-\theta}{1+\theta^2} & k = 1 \\ 0 & k > 1 \end{cases} \tag{4.65}$$

Moving Average Process of order 2: MA(2)

Consider the order 2 moving average model:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \tag{4.66}$$

Where, the parameters are θ_1 and θ_2, also ε_t is a white noise process. This process is stationary for any value of θ_1 and θ_2.

The most important characteristics are:

a) Stationary in mean

$$E(X_t) = E(\varepsilon_t - \theta_1\, \varepsilon_{t-1} - \theta_2\, \varepsilon_{t-2}) = 0 \tag{4.67}$$

b) Autocovariance function $\gamma_{k,}\ k = 0,1,2,3, \ldots$

$$\gamma_0 = E(X_t - E(X_t))^2 = E(X_t)^2 \tag{4.68}$$
$$\gamma_0 = E(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2})^2 \tag{4.69}$$
$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma^2 \tag{4.70}$$
$$\gamma_1 = E\big[\big(X_t - E(X_t)\big)\big(X_{t-1} - E(X_t)\big)\big] = E(X_t X_{t-1}) \tag{4.71}$$
$$\gamma_1 = E[(\varepsilon_t - \theta_1\, \varepsilon_{t-1} - \theta_2\, \varepsilon_{t-2})(\varepsilon_{t-1} - \theta_1 \varepsilon_{t-2} - \theta_2 \varepsilon_{t-3})] \tag{4.72}$$
$$\gamma_1 = (-\theta_1 + \theta_1\theta_2)\sigma^2 \tag{4.73}$$
$$\gamma_2 = E\big[\big(X_t - E(X_t)\big)\big(X_{t-2} - E(X_t)\big)\big] = E(X_t X_{t-2}) \tag{4.74}$$
$$\gamma_2 = E[(\varepsilon_t - \theta_1\, \varepsilon_{t-1} - \theta_2\, \varepsilon_{t-2})(\varepsilon_{t-2} - \theta_1 \varepsilon_{t-3} - \theta_2 \varepsilon_{t-4})] \tag{4.75}$$
$$\gamma_2 = -\theta_2\sigma^2 \tag{4.76}$$
$$\gamma_3 = E\big[\big(X_t - E(X_t)\big)\big(X_{t-3} - E(X_t)\big)\big] = E(X_t X_{t-3}) \tag{4.77}$$
$$\gamma_2 = E[(\varepsilon_t - \theta_1\, \varepsilon_{t-1} - \theta_2\, \varepsilon_{t-2})(\varepsilon_{t-3} - \theta_1 \varepsilon_{t-4} - \theta_2 \varepsilon_{t-5})] \tag{4.78}$$
$$\gamma_3 = 0 \tag{4.79}$$

In summary, the autocovariances of MA(2) are:

$$\rho_k = \begin{cases} \gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma^2 & k = 0 \\ \gamma_1 = (-\theta_1 + \theta_1\theta_2)\sigma^2 & k = 1 \\ \gamma_2 = -\theta_2\sigma^2 & k = 2 \\ \gamma_3 = 0 & k > 2 \end{cases} \tag{4.80}$$

The autocorrelation functions are given by:

$$\rho_k = \begin{cases} \rho_1 = \dfrac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2} & k = 1 \\ \rho_2 = \dfrac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} & k = 2 \\ \rho_3 = 0 & k > 2 \end{cases} \tag{4.81}$$

4.3.4 Autoregressive Moving Averages Process ARMA($\boldsymbol{p}, \boldsymbol{q}$)

The Auto-Regressive Moving Average (ARMA) model combines the autoregressive (AR) and moving average (MA) models to provide a comprehensive representation of a stationary random process. The AR part of the model represents the autoregressive relationship between the current and past values of the time series, while the MA part represents the relationship between the current value and the past residual errors. The general ARMA model was first proposed by Whittle in his 1951 dissertation, and it has since become an important tool in time series analysis [208]. The ARMA model is used to describe and forecast future values of a time series Y_t. The AR part represents the relationship between Y_t and its past values up to a maximum delay of p, while the MA part represents the error term as a linear combination of error terms occurring simultaneously and at different periods in the past [206]. The order of the ARMA model is denoted by the ARMA (p, q) model, where p denotes the number of moving average and autoregressive components, respectively. The AR (p) model and the MA (q) model are two models included in the ARMA (p, q) model that work together to produce a thorough description of the time series. The ARMA model can be expressed as follows:

$$X_t = c + \emptyset_1 X_{t-1} + \cdots + \emptyset_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{4.82}$$

Where, ε_t is the error term at time t, and c, $\emptyset_1, \ldots, \emptyset_p, \theta_1, \ldots, \theta_q$ are the model parameters. The stationarity of this process depends entirely on the autoregressive part, and its reversibility depends on the moving average part. In this case, the stationarity is the same as in the AR(p) model, and its reversibility is the same as in the MA(q) model (AR is always reversible, MA is always a stationary process). To obtain an ARMA model, it is necessary to combine the MA process with the linear differential equation. If the homogeneous part of the difference equation includes p delays and the *MA* part includes q delays, the *ARMA* model is expressed. Here, if q = 0, it is the AR(p) process,

and if p = 0, it is the MA(q) process. In terms of the lag operator, the ARMA model can be expressed as follows:

$$\left(1 - \emptyset_1 L - \emptyset_2 L^2 - \cdots - \emptyset_p L^p\right) X_t = \left(1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q\right) \varepsilon_t \quad (4.83)$$

$$\emptyset_p(L) X_t = \theta_q(L) \varepsilon_t \quad (4.84)$$

Where, $\emptyset_p(L)$ is the autoregressive polynomial, $\theta_q(L)$ is the moving average polynomial. If the process is stationary, its representation is MA(∞).

$X_t = \frac{\theta_q(L)}{\emptyset_p(L)} \varepsilon_t$, then $X_t = \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \varphi_3 \varepsilon_{t-3}$ ...

If the process is invertible, a representation is AR(∞)

$\frac{\emptyset_p(L)}{\theta_q(L)} X_t = \varepsilon_t$ then $X_t = \varepsilon_t + \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \pi_3 Y_{t-3}$ ...

The weights of representation MA(∞), as in the form AR(∞), are constrained to depend on the finite vector of model parameters ARMA(p, q): $\emptyset_1, \ldots, \emptyset_p, \theta_1, \ldots, \theta_q$.

One way to determine if an ARMA process is stationary is by examining the roots of its autoregressive polynomial, $\emptyset_p(L)$ where L is the lag operator. If the modulus of all the roots of the autoregressive polynomial is less than 1, then the ARMA process is stationary. In other words, if all the roots of the autoregressive polynomial are outside the unit circle, then the ARMA process is stationary and can be used to model a time series that exhibits a stable mean and variance over time. The stationarity conditions, of the ARMA(p, q) model, are imposed by the autoregressive part, since the finite moving average part is always stationary.

- An autoregressive moving average process ARMA(p, q) is invertible if and only if the modulus of the roots of the moving average polynomial $\theta_q(L)$ are outside the unary circle.

The invertibility conditions, of the ARMA(p, q) model, are imposed by the moving average part, since the autoregressive part is always invertible, because it is always directly written in autoregressive form.

The ARMA(p, q) models will always share the characteristics of the model AR(p) and MA(q),

This is because it contains both structures at the same time. The model ARMA(p, q) has zero mean, variance constant and finite and an infinite autocorrelation function. The autocorrelation function is infinite, rapidly decreasing toward zero.

Autoregressive moving average process of order (1,1): ARMA(1,1)

Consider the ARMA(1,1) model, where X_t is determined as a function of its past to first lag, contemporary innovation, and past innovation to lag 1.

$$X_t = \emptyset_1 X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1} \tag{4.85}$$

Where, ε_t is the error term at time *t* and ∅, θ are the parameters of the model.

To check the stationarity of the model, the roots of the autoregressive polynomial are calculated:

$1 - \emptyset L = 0$, where $|L| = \left|\frac{1}{\emptyset}\right|$ that is $|\emptyset| < 1$

To check the invertibility condition of the model, the roots of the moving average polynomial are calculated:

$1 - \theta L = 0$, where $|L| = \left|\frac{1}{\theta}\right|$ that is $|\theta| < 1$

Characteristics of a stationary $ARMA(1,1)$ process

a) Mean

$$E(X_t) = E(\emptyset X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}) = \emptyset E(X_{t-1}) \tag{4.86}$$

$$E(X_t) = 0 \tag{4.87}$$

b) Autocovariance function

$$\gamma_0 = E(X_t - E(X_t))^2 = E(X_t)^2 \tag{4.88}$$

$$\gamma_0 = E(\emptyset X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1})^2 \tag{4.89}$$

$$\gamma_0 = \frac{(1+\theta-2\theta\emptyset)\sigma^2}{1-\emptyset^2} \tag{4.90}$$

$$\gamma_1 = E\left[\left(X_t - E(X_t)\right)\left(X_{t-1} - E(X_t)\right)\right] = E(X_t X_{t-1}) \tag{4.91}$$

$$\gamma_1 = E[(\emptyset_1 X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1})X_{t-1}] \tag{4.92}$$

$$\gamma_1 = \emptyset_1\gamma_0 - \theta\sigma^2 \tag{4.93}$$

$$\gamma_2 = E\left[\left(X_t - E(X_t)\right)\left(X_{t-2} - E(X_{t-2})\right)\right] = E(X_t X_{t-2}) \tag{4.94}$$

$$\gamma_2 = \emptyset\gamma_1 \tag{4.95}$$

Summarizing the autocovariances of ARMA(1,1)

$$\gamma_k = \begin{cases} \gamma_0 = \frac{(1+\theta-2\theta\emptyset)\sigma^2}{1-\emptyset^2} & k = 0 \\ \gamma_1 = \emptyset_1\gamma_0 + \theta\sigma^2 & k = 1 \\ \gamma_k = \emptyset\gamma_{k-1} & k > 1 \end{cases} \tag{4.96}$$

The autocorrelation function of ARMA(1,1) is

$$\rho_k = \begin{cases} \rho_1 = \emptyset - \frac{\theta\sigma^2}{\gamma_0} & k = 0 \\ \rho_k = \emptyset\rho_{k-1} & k > 1 \end{cases} \quad (4.97)$$

4.3.5 Autoregressive Integrated Moving Average Process ARIMA$(\boldsymbol{p}, \boldsymbol{d}, \boldsymbol{q})$

Not all time series are stationary; some show level changes over time or the variance is not constant. ARIMA is a statistical model used for analyzing and forecasting time series data that is non-stationary. The I component (symbolized by d) is added to make the series stationary by differencing. The AR component, often known as p, is created by correlating the most recent values of a data series with its prior values. The MA component, q, is created by correlating the most recent values of a random error term with its previous values. By combining these components, ARIMA can capture both the trend and the seasonality in the data, making it a powerful tool for time series analysis and forecasting. ARIMA stands for "AutoRegressive Integrated Moving Average," a time series forecasting model that considers past values, differences between values, and residual errors. The ARIMA (*p*, *d*, *q*) model is set by three parameters: *p* represents the order of the autoregressive component, *d* is the degree of differencing needed to render the time series stationary, and *q* represents the order of the moving average component, where the parameters *p*, *d*, and ($q \geq 0$). The model uses past values, differences, and residual errors to make predictions about future values in the time series [122]. The form is stated algebraically as follows:

$$X_t^d = c + \emptyset_1 X_{t-1}^d + \cdots + \emptyset_p X_{t-p}^d + \theta_1 \varepsilon_{t-1}^d + \cdots + \theta_q \varepsilon_{t-q}^d + \varepsilon_t^d \quad (4.98)$$

Expressed in the form of the lag operator polynomial, the ARIMA(p, d, q) model is

$$\Phi(L)(1-L)^d X_t = c + \Theta(L)\varepsilon_t \quad (4.99)$$

Where, X_t^d is the series of order differences d, ε_t^d is a white noise process, and $c, \emptyset_1, \dots, \emptyset_p, \theta_1, \dots, \theta_q$ are the model parameters.

In order to model such processes, first of all, stability must be ensured. The Box-Jenkins methodology for the construction of ARIMA(p, d, q) models is carried out iteratively through a process in which four stages can be distinguished. Figure 4.1 illustrates the Box-Jenkins for ARIMA technique schematically.

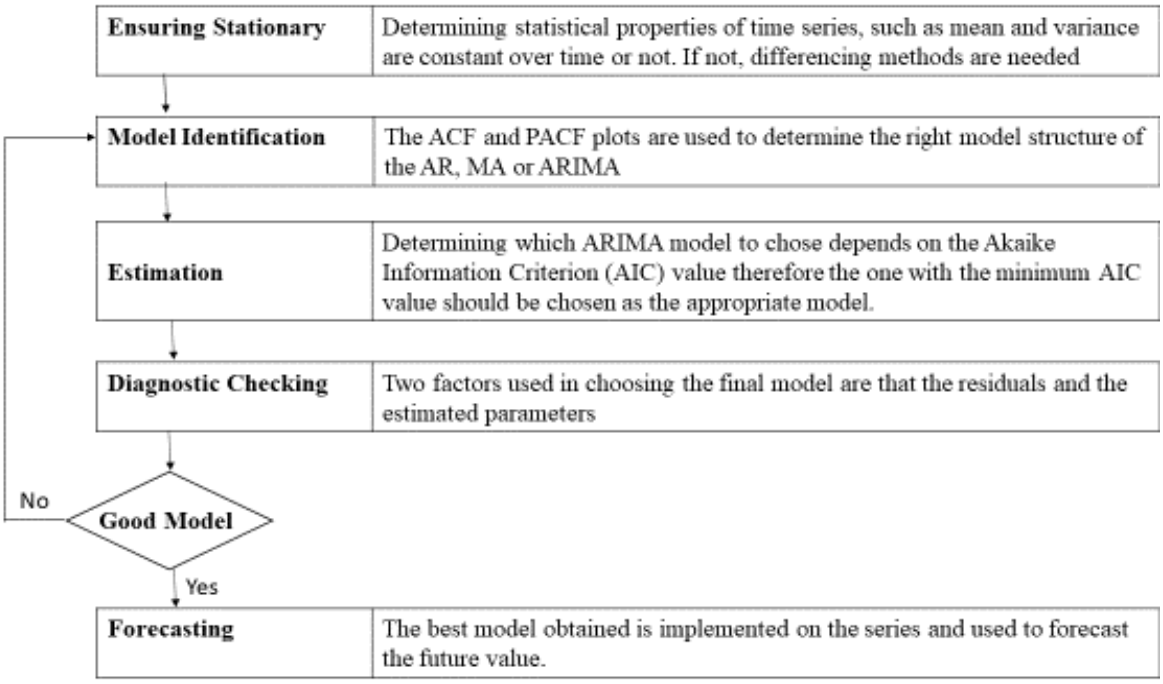

Figure 4.3 Schematic representation of the Box-Jenkins for ARIMA approach.

4.3.6. Seasonal Auto-Regressive Integrated Moving Average ARIMA$(p, d, q)(P, D, Q)_S$

SARIMA is a type of statistical model that analyzes the seasonal patterns that may be present in the data. It combines the ARIMA model with a seasonal component, allowing for the modeling of seasonal patterns in the data. The model is specified using six parameters: *p*, *d*, and *q* for the ARIMA component, and *P*, *D*, and *Q* for the seasonal component where *p* and *P* are the orders of the AR component, *d* and *D* are the orders of differencing, and *q* and *Q* are the orders of the MA component. The S in SARIMA represents the number of seasons in each cycle. The general structure of the model is expressed as follows:

$$X_t = c + \emptyset_1 X_{t-1} + \cdots + \emptyset_p X_{t-p} + \theta_1 X_{t-s} + \cdots + \theta_p X_{t-ps} + \varepsilon_t + \varphi_1 \varepsilon_{t-1} - \cdots - \varphi_q \varepsilon_{t-q} - \vartheta_1 \varepsilon_{t-s} \ldots - \ \vartheta_Q \varepsilon_{t-Qs} \quad (4.100)$$

The parameters $\emptyset_1, \ldots, \emptyset_p, \theta_1, \ldots, \theta_p, \varphi_1, \ldots, \varphi_q, \vartheta_1, \ldots, \vartheta_Q \ and \ \varepsilon_t \sim N(0, \sigma^2)$, where σ^2 is the variance.

4.4 Prediction of Time Series with Neural Networks

The Box-Jenkins method is a powerful tool for analyzing time series data and can be applied to a wide range of situations. However, it does require a strong understanding of statistical methods and may not be appropriate for all types of time series data. However, many time series contain

nonlinear relationships as well as linear relationships. Different methods are needed to model this nonlinear relationship. Artificial neural networks (ANNs), which can model both linear and nonlinear relationships depending on the feature of the activation function in their structure, have become more popular for solving modeling problems. Neural networks are a type of machine learning algorithm that is inspired by the structure and function of the human brain. Artificial neural networks are information processing models, inspired by the functioning of the brain. They have the ability to learn from experience. Neural networks have shown efficiency in the study of different problems for which traditional methods have not produced good results [208]. Artificial neural networks are indeed based on simple mathematical models that are inspired by the structure and function of the human brain. A neural network can be viewed as a network of neurons organized in layers, in which the predictors (or inputs) form the lowest layer and the predictions (outputs) form the highest layer. Between both layers, there may be intermediate layers with hidden neurons. This hidden intermediate layer is the one that allows a non-linear relationship between the inputs and the outputs, allowing the model more degrees of freedom (since non-linear activation functions are used) (if only the input and output layers were present, it would be a regression linear (simplest model)). An artificial neuron consists of a set of inputs (x), a set of synaptic weights, corresponding to each neuron input (w), an aggregation function (Σ), an activation function (f), and a set of outputs.

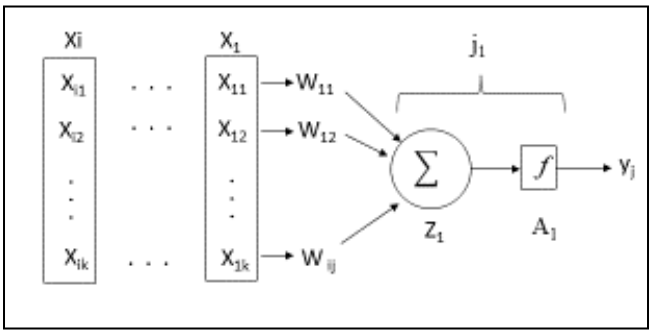

Figure 4.4 Schematic representation of neurons [209].

In a neural network, each neuron is going to receive a set of inputs weighted by a specific value, to which an activation function is applied once the sum of their values has been made in order to obtain an output. The inputs correspond to the input data if it is in the first layer, or to the outputs of the other neurons if they are in the hidden or final layers. In addition, there is an additional input called a threshold, or bias that provides an additional degree of freedom to the model. The weights are real values that multiply the inputs. Initially, the weights are initialized randomly, and they are

updated by means of training algorithms on the observed data that minimize a cost function, so they will represent the knowledge of the network. Networks are series of structures where each layer is a function of the layer that precedes it. They can be defined such that:

$$a^k = W^k h^{k-1} + b^k, k > 0, \quad (4.101)$$

$$h^k = g^k(a^k), k > 0, \quad (4.102)$$

Where, W^k is the vector of weights for layer k with dimension $(n_k \times n_{k-1})$ with n being the number of neurons in layer k. $b^k(n_k \times 1)$ is the vector of biases for layer k, $a^k(n_k \times 1)$ is the vector resulting from the operation on layer k, and $h^k(n_k \times 1)$ the vector of activations resulting from said layer after applying an activation function on it. Activation functions are nonlinear functions that enable the network to learn complex patterns in data. The goal of activation functions is to convert the output of a linear transformation, as shown in equation (4.102), into a nonlinear transformation that allows complex patterns in the input data to be captured. There are a number of characteristics that activation functions must have. The necessary ones are those of continuity throughout the defined domain and of being a non-decreasing monotonic class function. There are some other desirable characteristics for the formulation of a good activation function: the characteristic of being continuously differentiable, as well as the fact that the derivative is monotonic in order not to have problems with optimization in the form of a gradient since functions where the derivative is nonexistent can cause errors in training progress. Another desirable property is the approximation of the identity near the origin, that is, functions that, at the origin, have any constant value but a value of the derivative that is different from zero. Taking the theory to a more pragmatic level, functions that do not have a high computational cost are needed, since neural networks are inherently expensive in terms of training time. This can, for example, be optimized by trying to find an expression for the derivative that can be expressed in terms of the value of the function without a derivative. Another function is to model the output to be predicted in the form of a probability distribution; that is why we usually use a series of different activations for the hidden layers and others for the output layer, making these last activations dependent on the problem that is being solved. Some of the best-known activation functions are:

4.4.1 Linear Activation Function

The input has a directly proportional relationship with the linear activation function, which means it can increase or decrease its output. This allows for the linear activation function to solve the issue of zero regression in the binary step function. Additionally, the linear activation function can

output any value between negative infinity and positive infinity, making it more flexible than the binary step function, which is limited to only two output values (0 or 1). The identifier may take the form of:

$$F(z) = az \tag{4.103}$$

where the user may choose any constant value for variable *a*.

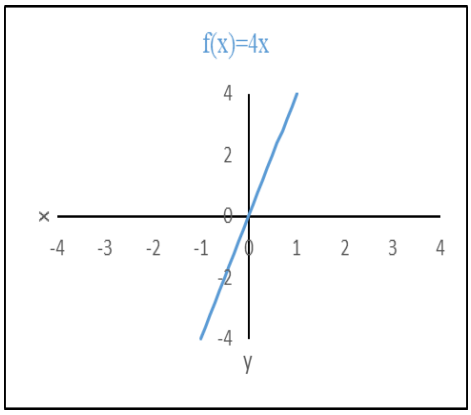

Figure 4.5 Linear activation function.

The derivative of the function *f(x)* in this case is not zero but rather equal to the value of the chosen constant. The gradient is not zero but rather a constant value irrespective of the input value *x*, indicating that the weights and biases will be changed throughout the backpropagation phase even if the updating factor remains the same. Since the neural network would not reduce the error because the gradient would be the same for each iteration, adopting a linear function is not particularly advantageous. In addition, the network will not be able to extract complex patterns from the data. Therefore, linear functions are suitable for straightforward tasks and situations where interpretation is necessary [211].

4.4.2 Sigmoid Activation Function

The sigmoid function is a non-linear function that is continuously differentiable, monotonic, and has a fixed range between 0 and 1. However, as you mentioned, the sigmoid function can suffer from the vanishing gradients problem, which can make it difficult for deep neural networks to learn. The sigmoid function is defined as:

$$f\ (x) = \frac{1}{1+e^{-x}} \tag{4.104}$$

Where, x is the input to the function and $f(x)$ is the output of the function. The sigmoid function is commonly used as an activation function in the neurons of artificial neural networks, especially in the context of binary classification problems. The derivative of the sigmoid function can be computed as:

$$f'(x) = f(x) * (1 - F(x)) \tag{4.105}$$

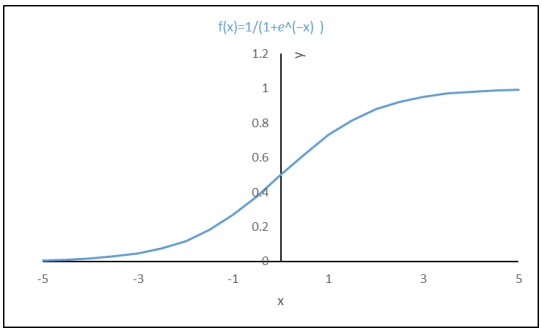

Figure 4.6 Sigmoid Function.

4.4.3 Tanh Function

Like the sigmoid function, the hyperbolic tangent function is continuous and differentiable, and its outputs are bounded between -1 and 1. The hyperbolic tangent function also has an advantage over the sigmoid function in that its derivative is steeper around the origin, which means that the gradient does not saturate as quickly as the sigmoid function for inputs with large absolute values. However, like the sigmoid function, the hyperbolic tangent function can still saturate at very high or low values, which can lead to vanishing gradients in deep architectures. It can be defined as:

$$f(x) = 2sigmoid(2x) - 1 \tag{4.106}$$

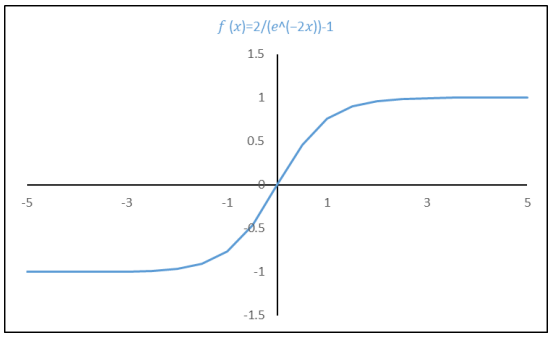

Figure 4.7 Tanh Function.

4.4.4 Rectified Linear Activation ReLU Function

ReLU function is widely used in practice and can be defined by

$$f(x) = \max(0, x) \tag{4.107}$$

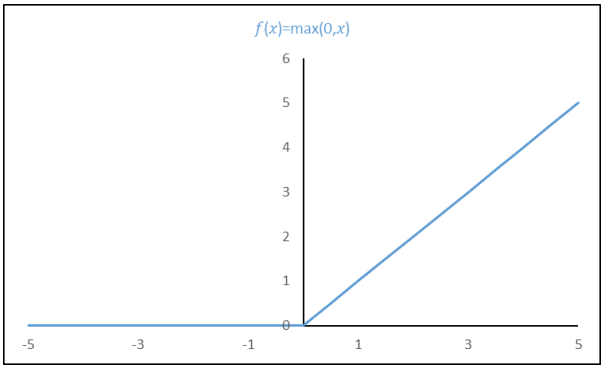

Figure 4.8 ReLU Activation Function.

The activation function in neural networks is computationally efficient and can help address the problem of vanishing gradients that can occur with other activation functions. Even so, it has problems with activations in regions where the input is negative ($x < 0$) or where the ReLU gradient is 0, because the weights are not updated during the process, which implies that those neurons go to a non-activation state. During the training, In addition to the fact that the activation can acquire very high values for certain inputs, a normalization prior to the use of architectures based on ReLU activations is necessary, as is an initialization of the weights in accordance with their characteristics.

4.4.5 Leaky ReLU Function

It is an improvement to ReLU function. It can be expressed mathematically as:

$$f(x) = \begin{Bmatrix} x & x>0 \\ \alpha x & x\leq 0 \end{Bmatrix} \tag{4.108}$$

It is a variant where it is intended to solve the problem of "dead" neurons in ReLU due to zero gradients with a zone for negative inputs with a decreasing gradient controlled by a constant of magnitude α.

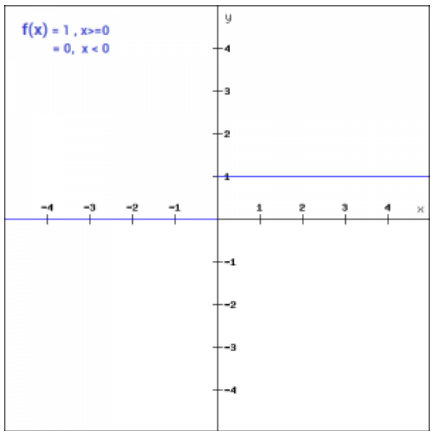

Figure 4.9 Leaky ReLU Function.

4.4.6 Softmax Activation Function

Softmax is a generalization of the sigmoid function and is used when the categories to be modeled are mutually exclusive, that is, each instance belongs to only one class. When modeling non-exclusive classes, the sigmoid activation can be used, which produces a probability for each class that does not have to add up to unity, which is useful in multilabel problems. Also, it can use the previous activations to model outputs in certain ranges. This flexibility provided by neural networks is one of the reasons why they are widely used in practise and in the field of pattern recognition. Softmax has a function that represents a probability distribution:

$$f(x) = \frac{e^{x_j}}{\sum_j e^{x_j}} \tag{4.109}$$

interpreting the result as probabilities of the type $y_j(j) = p(class = j|x)$.

In time series prediction problems, the goal is typically to predict a continuous value at some point in the future given a sequence of historical data. In this context, a common approach is to use a neural network with a linear output layer, which is appropriate for regression problems where the goal is to predict a continuous value and can be viewed as a conditional Gaussian distribution with mean and variance given by the output of the network like:

$$p(t|x) = N(t: Wh + b) \tag{4.110}$$

Where, *t* is the target variable at a future time point, *x* is the historical data, *W* is the weight matrix, *h* is the output of the hidden layer of the neural network, *b* is the bias vector, and *N*( ) is the Gaussian distribution. The mean of the Gaussian distribution is given by the linear combination $Wh + b$, which is the output of the neural network, and the variance is not explicitly specified in

this expression. In practice, the variance is often assumed to be constant or modeled as a function of the input *x* or some other parameters.

4.4.7 Choosing the Right Activation Function

Numerous factors need to be taken into account for better performance and fewer inaccurate outcomes, including the number of hidden layers in a network, training techniques, hyperparameter tweaking, etc. One of the most important considerations is the activation function. Selecting the appropriate activation function for any given activity may be a laborious process that calls for extensive study and investigation. The decision of which activation function to use is context-dependent, or it depends on the task that needs to be completed, rather than being governed by any general guidelines. Depending on the sort of system that is being created, several activation functions each offer advantages and disadvantages of their own.

4.5 Recurrent Neural Networks (RNN)

Artificial intelligence has been widely used for predictions of temporal sequence data. RNN was used and presented satisfactory results [212]. The idea behind the architecture of recurrent neural networks is that each time step shares the same progressive (i.e., current) internal state, and the neural network has information not only about the current time step but also about the previous one. Such an implementation is achieved, for example, by establishing a joint connection between hidden layers and receiving a signal from the previous time step to the hidden layers of the current time step, as shown in Figure 4.10.

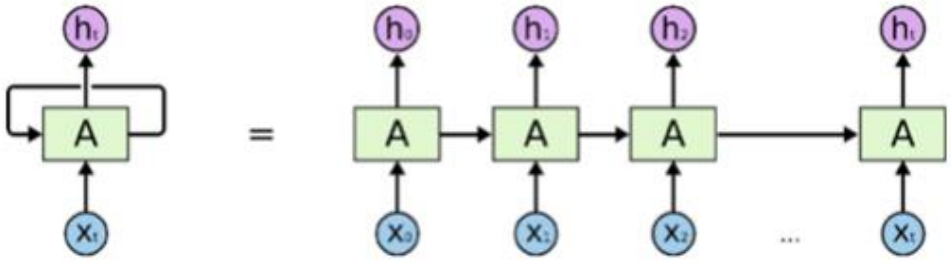

Figure 4.10 RNN Traditional network. [213]

The connection between the listed elements can be formalized as follows:

$$h(t) = f(h(t-1), X(t), \theta) \tag{4.111}$$

Where, f is the direct propagation function, $h(t)$ is the state at time t, $X(t)$ is a vector of input signals, and θ is a vector of parameters.

Like most neural networks, recurrent neural networks were described quite a long time ago. The vanishing gradient problem refers to the situation in deep neural networks where the gradients of the parameters, which are used to update the weights during training, become very small, causing the optimization (minimizing the error) process to stall. This can occur because the gradients are multiplied many times during backpropagation, leading to an exponential decay of their magnitude. As a result, the weights are not updated effectively, leading to slow or ineffective learning. The problem of the vanishing gradient arises because, by passing through the layers of the neural network and propagating in the network, vectors are subject to multiplication operations that significantly increase or decrease their value, especially in recurrent neural networks. Due to the fact that the composition of connections between the layers of neural networks is carried out using a multiplication operation, the derivatives in these cases are vulnerable to exaggeration of the value (the so-called "explosion") or, vice versa, to an excessive decrease in the value and disappearance. If it gets into a situation with an excessive increase in the gradient, then such a gradient is called "explosive," and in this case, each weight will be too large and the training of the neural network will be impossible. The gradients of these weights saturate at the highest end, meaning they are considered too powerful. However, "explosive" gradients can be solved relatively easily because they can be truncated or dissected. Also, the described neural network has a significant drawback: it does not have the ability to get an estimate for events that happened a long time ago; that is, speaking in terms of hidden layers, the influence of the hidden layers of the recurrent neural network decreases over time, which concludes that this type of neural network can be used for forecasting and analysis of time series that do not contain a large amount of data. This is because the gradient of the loss function decays exponentially with time [214]. Deepened multi-layer neural networks are now called "deep learning." However, the vanishing gradient problem remains a challenge for complex models. To use a recurrent network to solve the given task of forecasting and time series analysis, it must be modified to be able to receive data from hidden layers that have an impact on the current internal state, record and store this data in some way, and then operate on or draw conclusions about these data and use the received information to build a forecast. Recurrent neural networks that have this modification are called "neural networks with long short-term memory" and are denoted as LSTM.

4.6 Neural Networks with Long Short-Term Memory

Long short-term memory networks (LSTM) are a special type of recurrent neural networks (RNN) to overcome the vanishing gradient for nonsequential data [215, 216]. The most obvious difference between LSTM and RNN is that LSTM addresses the vanishing gradient problem by introducing memory cells that allow information to be stored and processed over a long period of time. Memory blocks are more complex and intelligent than traditional neurons. The vanishing and exploding gradient problems are common in traditional RNNs and can be addressed by using the LSTM architecture, which uses memory cells to maintain information over a longer period of time and avoid the gradients from either vanishing or exploding during backpropagation through time. The vanishing and exploding gradient problems can occur in RNNs when matrix weights are either too small or too large, respectively. When the matrix weights are small (less than 1.0), learning becomes very slow or stops working because the incident gradient signal is very small, which tends to disappear, with the result that it is difficult to learn long-term dependencies in the data. If the matrix weights are large (greater than 1.0), the learning will never converge because the gradient signal is too large, and this leads to the so-called gradient explosion. To address these problems, the LSTM architecture was developed, which uses memory cells to maintain information over a longer period of time and avoid the vanishing or exploding gradient problems during backpropagation through time. LSTM was introduced in 1995 as an extension of RNN and has since become a popular solution for processing sequential data in many applications [217]. LSTM has more power to learn from experiences in time series analysis than the recurrent neural network [218]. LSTM is realized in Figure 4.11 by replacing the middle layer unit of an RNN with a block with memory and three gates [219].

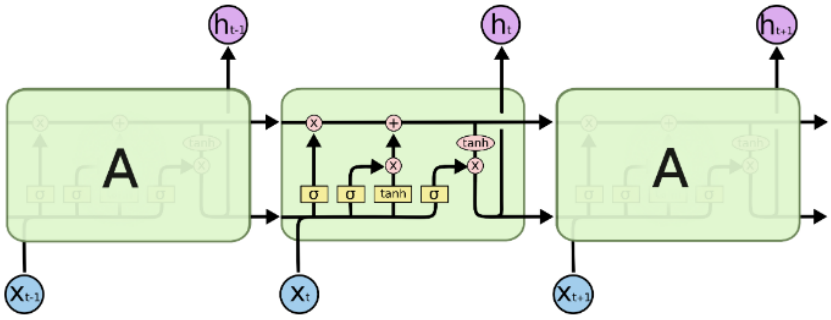

Figure 4.11 Four interconnected layers form the repeating module in an LSTM.

The memory cell of the LSTM combines a sigmoid layer and gates with the ability to add and erase information from the cell state, consisting of a multiplier (see Figure 4.12).

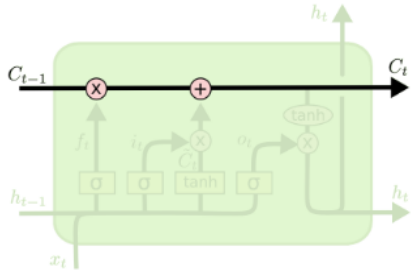

Figure 4.12 Memory cell

In the context of Figures, 4.11/4.12 the memory cells are indeed a type of hidden module. In an LSTM network, the memory cells serve as an internal memory that allows the network to selectively remember or forget information from previous inputs. The number of training sequences, or the size of the training dataset, can also influence the optimal number of hidden modules in a neural network. A larger training dataset may allow for more complex models with a larger number of hidden modules, while a smaller dataset may require a simpler model with fewer hidden modules to avoid overfitting. However, the relationship between the number of training sequences and the number of hidden modules is not straightforward and may depend on other factors such as the complexity of the problem and the specific architecture of the network.

The gates in an LSTM network play a crucial role in allowing the network to selectively add or delete information from its memory. These operations are carefully regulated by internal structures called gates. These gates allow information to be added or removed from memory at a given time. They are made up of a one-layer neural network (normally with a sigmoid as the activation function) together with an arithmetic operation (multiplication or addition), as in Figure 4.13.

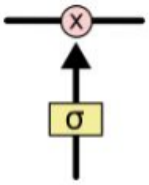

Figure 4.13 Gates

The sigmoid layer describes what information to give to the components by converting the information into a number between 0 and 1. If the output is 1, all information is passed; if it is 0, no information is allowed to pass. The memory cell in LSTM consists of three different gates. These are the forget gate that decides whether to keep the information, the input gate that activates the updated information from the last cell, and the output gate that chooses what information to

output to the next cell. In LSTM, information is allowed to flow along the series while the information remains unchanged to maintain integrity for a long time.

Next, the internal operation and the flow of information in an LSTM will be explained gradually:

- The LSTM decides what data to remove from memory as its initial step. The door called "Forget gate" makes this selection. As can be seen in Figure 4.14, h_{t-1} and x_t are concatenated, and the result will be the input to the small network that makes up the forget gate. The result of the network is used to decide if the state of the memory will be left as it is or if it will be altered by deleting some element.

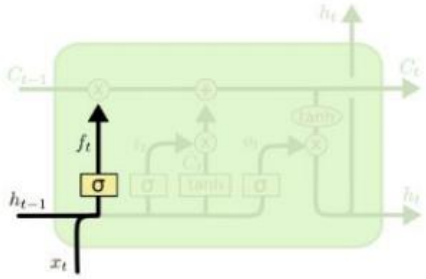

Figure 4.14 Forget gate

$$f_t = \sigma\left(W_f . [h_{t-1}, x_t] + b_f\right) \tag{4.112}$$

- The next stage is choosing which new data should be kept in the memory state. This is divided into two parts. First, the "input gate layer" of the sigmoid layer provides the values that need to be updated. The new candidate values are then created by the *tanh* layer into a vector $\hat{C}_t$ of the new candidate values, which can be added to the state. The next step is to combine the two to update the state.

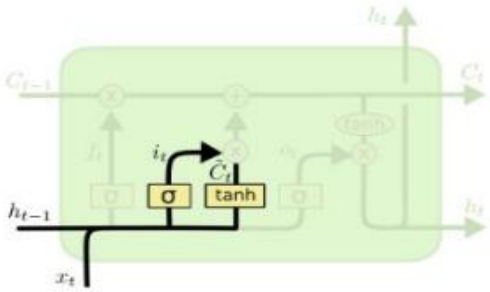

Figure 4.15 Input gate

$$i_t = \sigma(W_i . [h_{t-1}, x_t] + b_i) \tag{4.113}$$

$$\widehat{C}_t = \tanh(W_c . [h_{t-1}, x_t] + b_c) \tag{4.114}$$

At this point, the memory state can be updated based on the previous memory state C_{t-1} and the new input for that time step C_t. The forget gate f_t controls the extent to which previous information is forgotten or erased from memory, and the input gate i_t controls the extent to which new information is added to memory. These gates are used to compute the new memory state. By multiplying the previous state by the f_t, and forgetting a certain amount of information as a function of the f_t. Then i is added to memory, and it is updated with new $i_t * \hat{C}_i$ values that may be useful in the future.

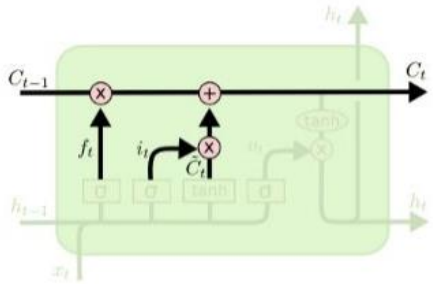

Figure 4.16 Memory update

$$C_t = f_t * C_{t-1} + i_t * \hat{C}_t \tag{4.115}$$

Finally, need to decide what the LSTM cell output will be. This will be obtained as the product of two elements. The first of them will be the output of the sigmoid network, which will be used to decide which elements of the memory will be combined. The second element will be the filtering of data from memory by a tanh (to push the values to be between -1 and 1). These two elements will multiply, resulting in the new output of the cell.

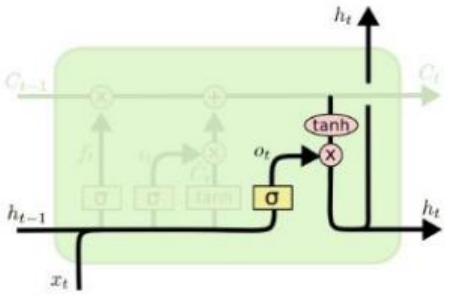

Figure 4.17 Output

$$O_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \tag{4.116}$$

$$h_t = o_t * \tanh(C_t) \tag{4.117}$$

It makes sense that the learning phase of LSTMs takes longer than that of a traditional neural network or RNN due to its complex structure, particularly the existence of three weight matrices (forget gate, input gate, and memory calculation weight). However, an LSTM performs far better. In reality, due to their ability to manage long-term dependency, LSTMs are now one of the most popular recurrent network types, with amazing results.

4.7 Forecast Evaluation Measures

To evaluate the accuracy level and the performance of the model and network, several popular statistical metrics are available.

4.7.1 Mean Absolute Error (MAE)

In time series analysis, the mean absolute error (MAE) is a widely used indicator of forecast error. It is determined by averaging the total set of data absolute discrepancies between the original and anticipated values. A lower MAE value denotes a better forecast because the MAE is a linear measure of error and is always positive. The following formula can be used to calculate the MAE:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|o_i - p_i| \tag{4.118}$$

Where, N is the number of data points in the data set and Σ represents the sum of the absolute differences between the actual and predicted values.

4.7.2 Root Mean Square Error (RMSE)

The square root of the mean square of all errors refers to the root mean square error (RMSE). In both statistics and machine learning, the use of the RMSE in regression is highly popular, and it is regarded as a superior all-purpose error metric for numerical predictions. The higher the RMSE, the greater the error, while the lower the RMSE, the better the model is at making accurate predictions. A characteristic of the RMSE is that the errors are squared before being averaged, which results in a much bigger weight being ascribed to larger errors. Therefore, an error of 10 is 100 times worse than the error of 1. It is determined by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(o_i - p_i)^2}{N}} \tag{4.119}$$

RMSE is based on the sample size (N) and the observation scale and displays the magnitude of the estimation error [220].

4.7.3 Mean Absolute Percentage Error (MAPE)

In time series forecasting, MAPE is frequently employed. It is represented as a percentage and derived by:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{o_i - p_i}{o_i}\right| \quad (4.120)$$

MAPE has been used as a mean of comparison between several algorithms for predictions. It is considered highly accurate when the value is less than 10% and a reasonable forecast when the value is 11-20% [221].

4.7.4 Coefficient of Determination (R^2)

The coefficient of determination is used to calculate the strength of the linear relationship between the actual and predicted values of the models. In terms of the equation, it looks like this:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(o_i - p_i)^2}{\sum_{i=1}^{N}(o_i - \hat{o}_i)^2} \quad (4.121)$$

The closer the value is to one, the better the model performs. When the coefficient is equal to 0.80, the regression line should contain 80% of the data points. A larger value indicates greater observations' quality of fit [221].

For all measures, N is the total sample count used to determine the statistical assessment criteria, o_i observation value, p_i is the forecasted and $\hat{o}_i$ is the average of the actual observation values.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Introduction

This chapter presents the results obtained after applying the models used in this study to the collected data. It is divided into four sections relating to different experimental conditions. The most objective possible conclusions are drawn in order to highlight the strengths and weaknesses of the applied models depending on the situations in which they were used. The first section of this chapter covers topics such as time series analysis and forecasting of wind speed data. The second section discusses the results of wind speed prediction using long-term memory recurrent neural networks. A comparison of the efficiency of ARIMA and LSTM models for wind speed prediction and validation is presented in the third section. The final section relates to the use of deep learning techniques for solar energy forecasting. The results obtained in each section were analyzed and compared in order to identify the most effective models for prediction in different conditions. Results were evaluated based on several performance measures. These metrics allow a quantitative assessment of the accuracy of models and help determine which models provide the most accurate predictions. The results of this study provide valuable insights into the use of different models to predict wind speed and solar energy and can help develop more accurate and effective models for this application in the future.

5.2 Analysis of Time Series and Forecasting of Wind Speed Data

The time series under study is composed of 60 data samples and represents the average monthly wind speed from January 2015 to December 2019 in Chester, a Canadian village located in a part of the Nova Scotia municipality region, as a source of information to allow testing of the model used in this analysis. The data was tested in the RStudio statistical data processing program, and the program code was written in the R programming language, a free and open source software development environment widely used as statistical data analysis software that has become a de facto standard for statistical programs. Using R for time series analysis can accomplish two goals: the first is to determine the nature of the series, which includes selecting the main components of the time series and studying them in depth, and the second is to predict future values of the time

series using past values [222]. Figure 5.1 shows the average monthly wind speed behavior used in this analysis.

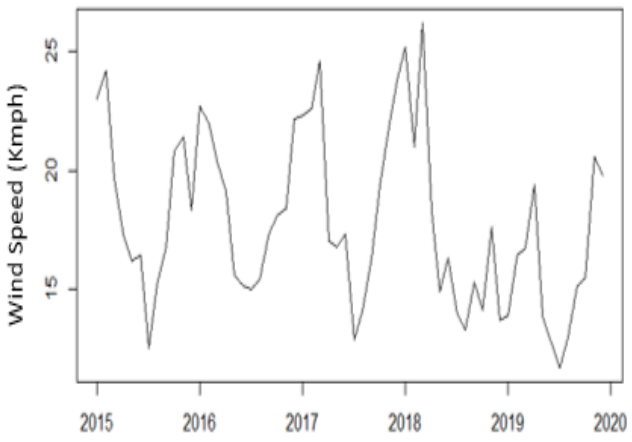

Figure. 5.1. Time series of the wind speed in Chester

The series of data were modeled through the application of descriptive statistics concepts to analyze how they behave; among the applied concepts are the arithmetic average, used to calculate the general average of the wind speed in the studied period, and the standard deviation, used to compare differences between sets of values. The minimum and maximum values of the monthly average wind speed are also recorded, as shown in table 5.1

Table 5.1: Statistical parameters of the time series.

Sample Size (N)	Mean	St. Dev.	Min	Max
60	17.754	3.601	11.7	26.3

In Figure 5.1, seasonal behavior is observed in the time series, but there are no trends. There is a regular succession of "peaks and troughs" that usually occur due to fluctuations associated with the seasons, which are repeated every year (with greater or lesser intensity). Time series that have a trend and/or seasonality are not stationary; hence, it is important to utilize the right methods in these cases. It can be checked by decomposing the time series. In the additive model, a function built into R that can be used to decompose a time series into its directional, seasonal, and random components, seasonal variations are represented by seasonal indices, or seasonal factors, one for each period in which the year is divided (if the series is registered monthly, there are 12 indices). The seasonal indices modify the trend when they are added, and if all indices are close to or exactly equal to zero, then the seasonal components do not seem to have a great effect on the series; if the

indices are substantially different from zero, both positive and negative, the trend value will be modified by them, indicating the influence of seasonal components in the series.

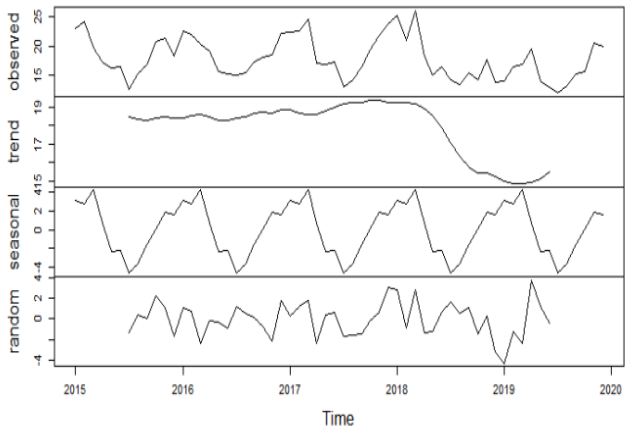

Figure 5.2. Decomposition of additive time series

If the graphs in Figure 5.2 are carefully examined, it can be determined whether the data fit all of the ARIMA modeling requirements, particularly invariance and seasonality. This leads to choosing the Dickey-Fuller test. The test results show that the data is non-stationary since the P value is 0.057, which is higher than 0.05. Therefore, that leads to transforming the series in Figure 5.1. When using the first differences, it turns out that the series still displays the trend, and this leads to the use of the second differences to remove the trend component and makes the series stationary in mean and variance, as shown in Figure 5.3

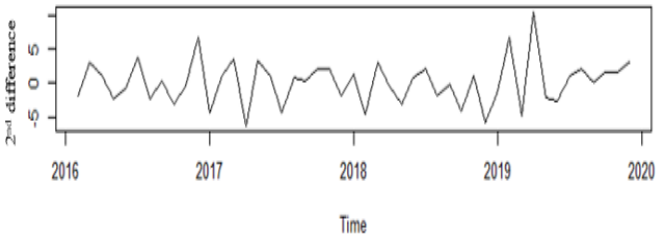

Figure 5.3. 2nd difference of time series

Note the scale of the graph. The values oscillate around zero: if greater than zero, they indicate components that increase the trend; if smaller, they decrease it.

The process of defining the structure of an ARIMA model involves making the time series stationary and then analyzing the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the difference series to determine the number of autoregressive (AR) and moving average (MA) terms needed. The highest correlation lag in the ACF plot is used to

determine the number of AR terms (p), while the PACF plot is used to determine the number of MA terms (q). The three variables, p, d (number of differences), and q, are positive integers that indicate the arrangement of the autoregressive, integrated, and moving average terms in the ARIMA model. The analysis of the ACF and PACF plots plays a crucial role in choosing the appropriate ARIMA model for the data. From the plot of (ACF) and (PACF) of the difference series, as shown in Figures 5.4 and 5.5, it is possible to determine the required AR and MA terms selected to define the model.

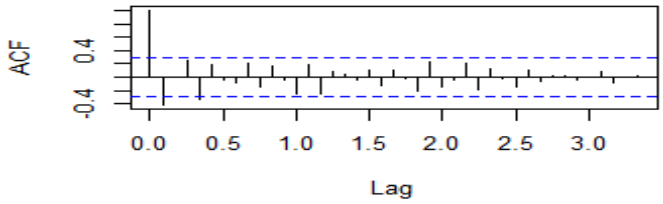

Figure 5.4. Autocorrelation function.

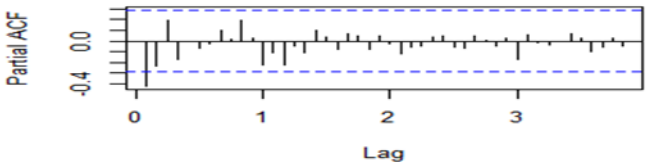

Figure 5.5. Partial autocorrelation function.

The amount of AR and MA terms required to define the model can be inferred from the plots of the difference series ACF and PACF, as shown in Figures 5.4 and 5.5From the information, we conclude that the initial time series was first converted into a second difference series, and then the ACF and PACF plots of the second difference series were used to determine the possible AR and MA terms of the SARIMA model. The specific SARIMA model is $(1,2,1)(2,2,1)_{12}$, which indicates that the model has an AR (1) term, a seasonal AR (2) term with a period of 12 (i.e. monthly seasonal), and MA (1), and seasonal MA (1) with a duration of 12. Table 2 presents the candidate models and the corresponding statistical criteria for each model to determine the best model for predicting future wind speed. Based on these criteria, SARIMA $(1,1,2)(2,1,1)_{12}$ was found to be the best predictive structure for Chester wind speed.

Table 5.2: RMSE, MAE and MAPE for 5 SARIMA model chosen

Model	RMSE	MAE	MAPE
$(2,2,1)(2,1,1)_{12}$	2.156	1.41	8.085
$(1,1,2)(2,1,1)_{12}$	1.968	1.295	7.396
$(1,2,1)(2,1,1)_{12}$	2.248	1.553	8.99
$(1,1,1)(2,1,1)_{12}$	2.175	1.438	8.24
$(1,2,1)(2,2,1)_{12}$	2.66	1.546	9.473

Note that the MAPE value is less than 10%, and in this case, it is considered highly accurate. If the proposed model is adequate, the residuals must be "close" to the errors and, therefore, must have a zero mean, constant variance, and be approximately uncorrelated (white noise), It is a random signal (a stochastic process) that is characterized in that its signal values at two different instants of time are not statistically correlated. Consequently, its power Spectrum density is constant. This means that the signal contains all the frequencies, and they all have the same power. The auto-correlation function (FAC) of the residuals is a very helpful tool for confirming the presumption that the errors are white noise. Figure 5.6 shows the random distribution of standardized residuals and does not follow a particular pattern.

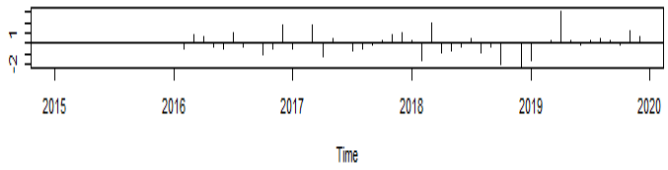

Figure 5.6. Standardized Residuals

To rule out structural breaks. The auto correlation function (ACF) for the model showed no signs of correlation since the mean of the residuals is close to zero and all of the spikes were within the standard error bars as shown in Figure 5.7.

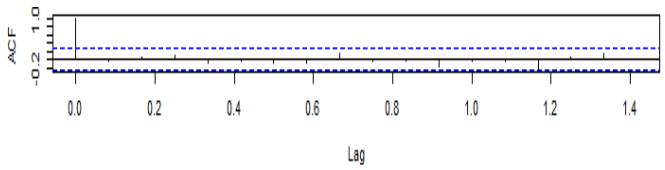

Figure 5.7. ACF of Residuals

Figure 5.8 shows the results of the Ljung-Box test, which tests the hypothesis of whether the residuals of a time series model are independently distributed or not. The p-values in the figure indicate the likelihood of observing the test statistic under the assumption of independence. If the p-value is greater than 0.05, the residuals are most likely independent, and the null hypothesis is not rejected. It is clear from the figure that the p-value of the test is above the blue dashed line at

0.05. This indicates that the residuals of the time series model are independently distributed, which is a common assumption made when structuring the model.

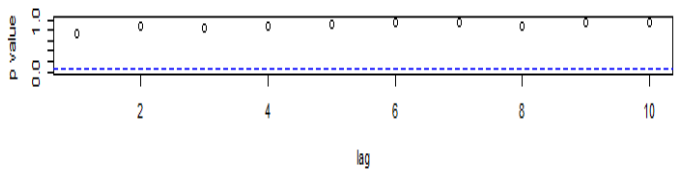

Figure 5.8. P values Ljung-Box statistic

Figure 5.9 shows the histogram of the residuals; it is noted that it is symmetrical and has an almost normal distribution form.

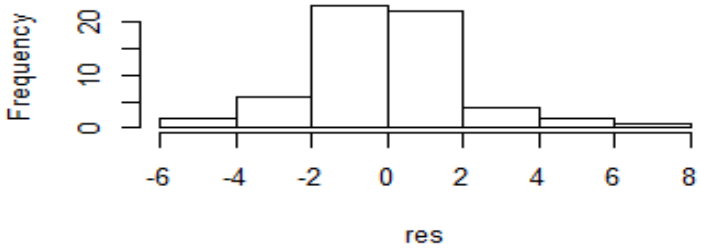

Figure 5.9. Histogram of residuals

The wind speed forecast using the ARIMA method is presented in Figure 5.10 for a time series that corresponds to the wind speed observations over a period of six years. Five years were used in the modeling, and the year ahead was used to confirm the accuracy of the prediction made using the method. Shaded areas in the figure represent 95% confidence intervals in the prediction. Seasonal ARIMA models proved highly effective in producing modified time series on the observed data for the study areas, as they could be managed, for example, to represent the realistic characteristics of the local seasonality and their high accuracy, which provided accuracy in predictions.

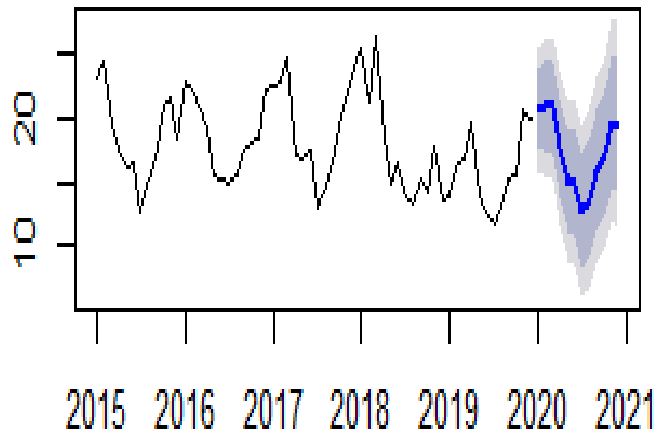

Figure 5.10. Forecasts from ARIMA (1, 1, 2) $(2, 1, 1)_{12}$

A summary of the results of this section is due to the great importance of wind speed when it comes to optimizing wind utilization. SARIMA technology was used to calculate wind speed forecasts for the Chester region. The results showed that the SARIMA(1,1,2)(2,1,1) structure had the best prediction accuracy of the models considered based on time series data from 2015 to 2019. To select the best time series model, RMSE, MAE, and MAPE were taken into account. The MAPE model score is 7.396%; this is an indication that the model is classified as highly reliable, and decision makers can benefit from it in developing plans. In general, this model is adaptable, can deal with seasonality, and only needs the latest model from a time series. However, it has disadvantages in that it is unlikely to perform well on long-term predictions, requires a lot of computational power, and can become subjective, requiring a high level of analysis of basic statistics.

5.3 Application of LSTM-RNN to Forecast Wind Speed

Although wind power has many advantages, the biggest difficulty is that wind speed is not constant over time. For this reason, it is very important to forecast the wind speed and, consequently, to know the amount of clean energy that can be produced. However, the task of forecasting wind speed is not an easy one, especially when you want to forecast a long horizon. Forecast models that use artificial intelligence have become popular in the last few decades as serious contenders for classical statistical forecasting models. In this section, the focus is on the usage of the LSTM algorithm for forecasting wind speed, along with a comparison of the method's prediction efficiency and accuracy when applied to various wind speed time series used for training and testing. The proposed model was applied to measure wind speed data from a station at Halifax

Dockyard in Nova Scotia. Figure 11 displays wind speed data measured at an altitude of 3.80 m as the source for two separate seasons, spring (March 2015) and summer (July 2015). For both seasons, data were collected at hourly intervals, using 576 read observations at 24 days and 168 reads at 7 days for the training and test sets, respectively.

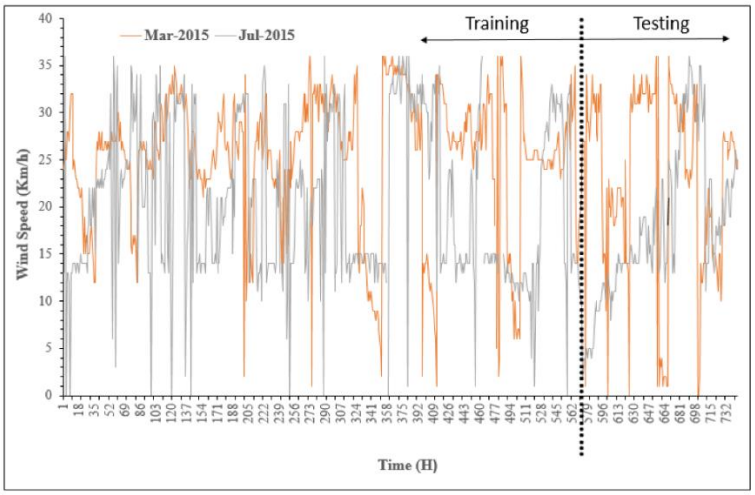

Figure 5.11. Time series of the wind speed in Halifax, Canada.

MATLAB software was used for the training process of the LSTM, which is a powerful type of recurrent neural network used in machine learning and designed to handle sequence dependencies. The LSTM was designed to train the sequence regression network, whose responses are training sequences with changing values in a one-time step. That is, using the most recent time steps, the LSTM network learns to predict the next time step. This is referred to as the window, and its size is an adjustable parameter for each issue. The choice of window size depends on the specific characteristics of the data and the modeling objectives and may require some trial and error experimentation to find the optimal value. For example, the current time step (t) and two previous time steps (t-1 and t-2) are used as input features to predict the value for the next time step (t+1). To implement the model, different window sizes ranging from 24 to 168 time steps (corresponding to 1 hour to 1 week of data) were chosen. The purpose of this study is to thoroughly and methodically evaluate the effectiveness and application of the suggested model, which is intended to forecast two series of wind speed data for two distinct seasons with two different meteorological characteristics, namely spring (March 2015) and summer (July 2015), respectively. Each data series is divided into 1–576 (24 days) observations and 577–744 (7 days), respectively, as the

training and test sets. Training data were standardized to have a zero mean and unit variance at prediction time to prevent training divergence. The best training parameter to obtain the lowest RMSE is found using an initial learning rate of 0.005. Figures 5.12 and 5.13 show the comparison of observed values with predicted values of hourly wind speed series collected in the spring (1–31 March 2015) and summer (1–31 July 2015), respectively, for evaluation of the LSTM, which was trained once (the value of the previous prediction) and reused to make a prediction for each time step between predictions. This means that no updates are made after the model has successfully fitted the training data. For this reason, the model in this instance is referred to as the "fixed model." Options for LSTM network training were chosen for 200 hidden modules. The maximum number of iterations is limited to 250, and the initial learning rate is set at 0.005. To stop the gradients from exploding, the gradient threshold is set to 1. After 125 epochs, the learning rate is reduced by a factor of 0.2 (full passes through the training data).

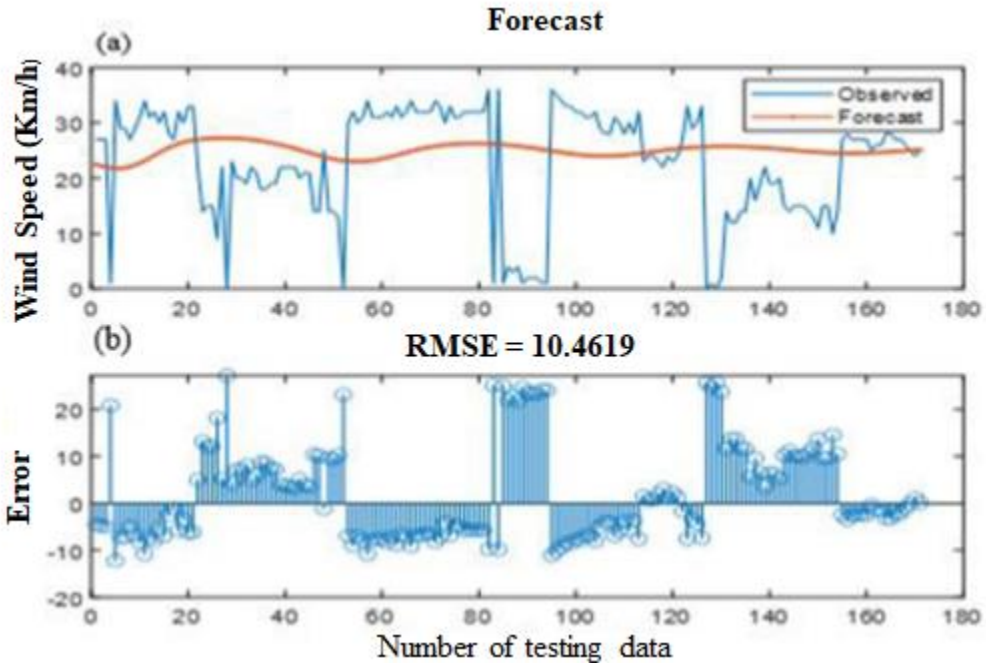

Figure 5.12. Wind speed prediction using LSTM for Spring (March 2015) data: (a) Comparison of real values and forecast values; (b) Error measurements on the RMSE scale.

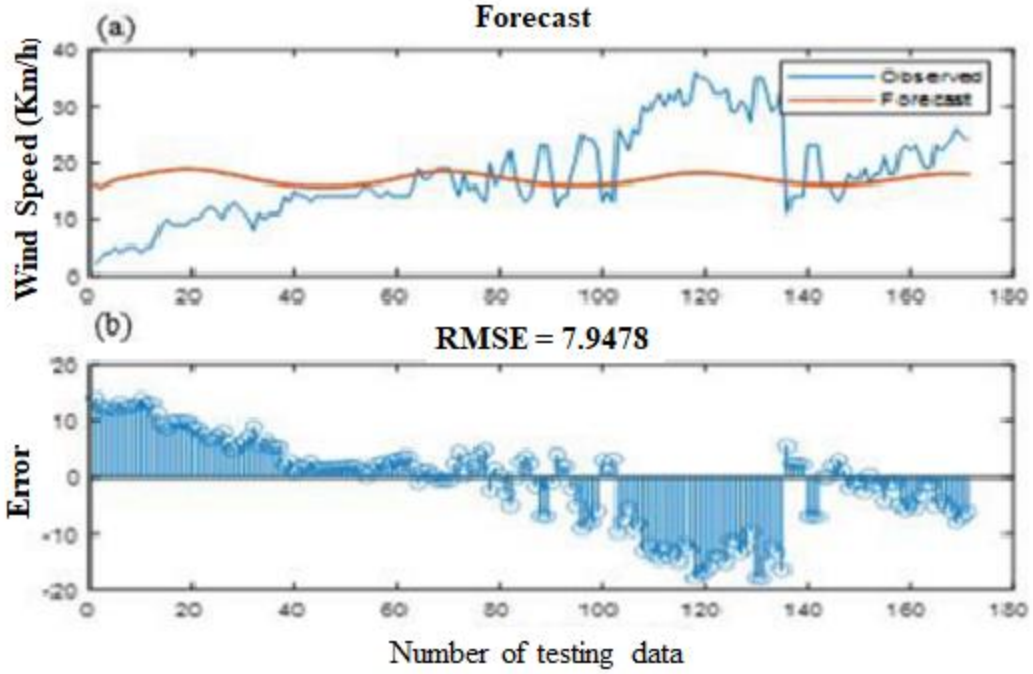

Figure 5.13. Wind speed prediction of existing LSTM for Summer (July 2015) data: (a) Comparison of observed values and forecast values; (b) Error measurements on the RMSE scale.

In both Figures 5.14 and 5.15, the LSTM state is updated with the most recent information for the time series prediction by updating the state values with the forecasted values from the test set and making them available to the model, so it can forecast on the following time step and update the state of the network after each prediction. At each time step, the memory state of the LSTM network is updated based on the previous memory state and the new input for that time step. The forget gate f_t controls the extent to which previous information is forgotten or erased from memory, and the input gate i_t controls the extent to which new information is added to memory. These gates are used to compute the new memory state. By incorporating the previous cell state, C_{t-1}, into the input, forget, and output gates, the LSTM can better capture the dependencies and patterns in the time series data. This makes it more effective at handling complex sequences and longer-term dependencies. By comparing the predicted values with the actual values in the test data set, the error score provides a summary of how well the model is able to capture the underlying trends in the data. The root mean square error (RMSE) is used because it penalizes large errors and results in a score in the same units as the predicted data, which is wind speed per hour. By calculating the error score of the modified LSTM, the RMSE value declined in the spring (March 2015) series by 4.5845 and in the summer (July 2015) series by 4.9392, due to their different characteristics. Based

on the results of this study and the results of many previous studies of different prediction models, the research is still ongoing to develop a model that gives results with the same accuracy for predicting data with different characteristics.

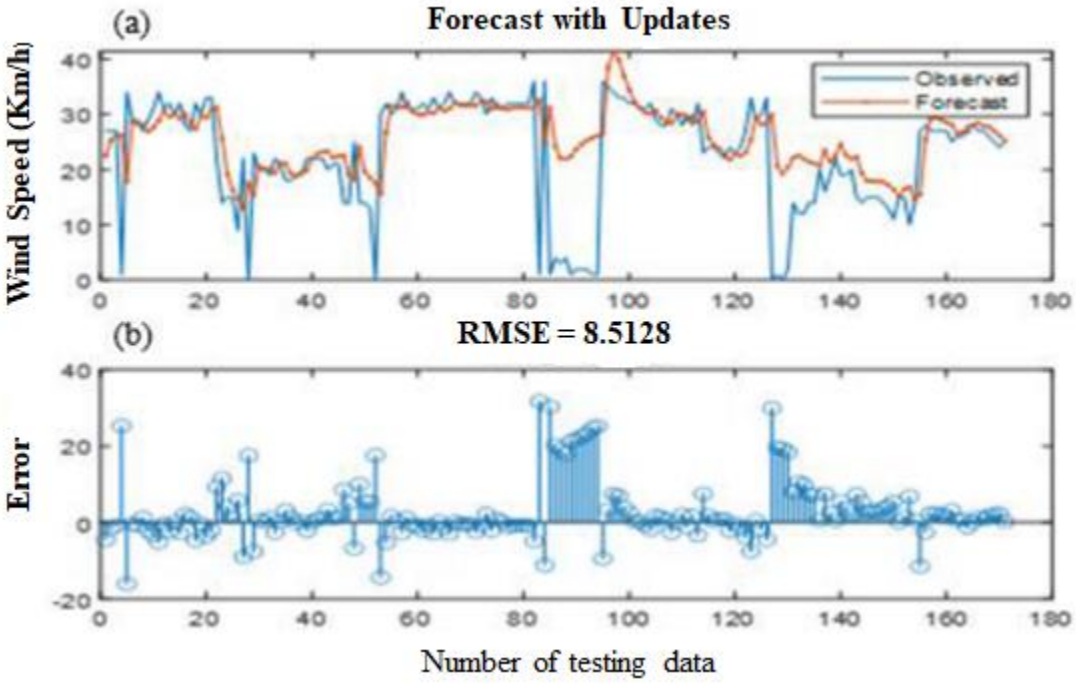

Figure 5.14. Wind speed prediction applied LSTM model for spring (March 2015) data: (a) Comparison of actual values and forecast values; (b) Error measurements on the RMSE scale.

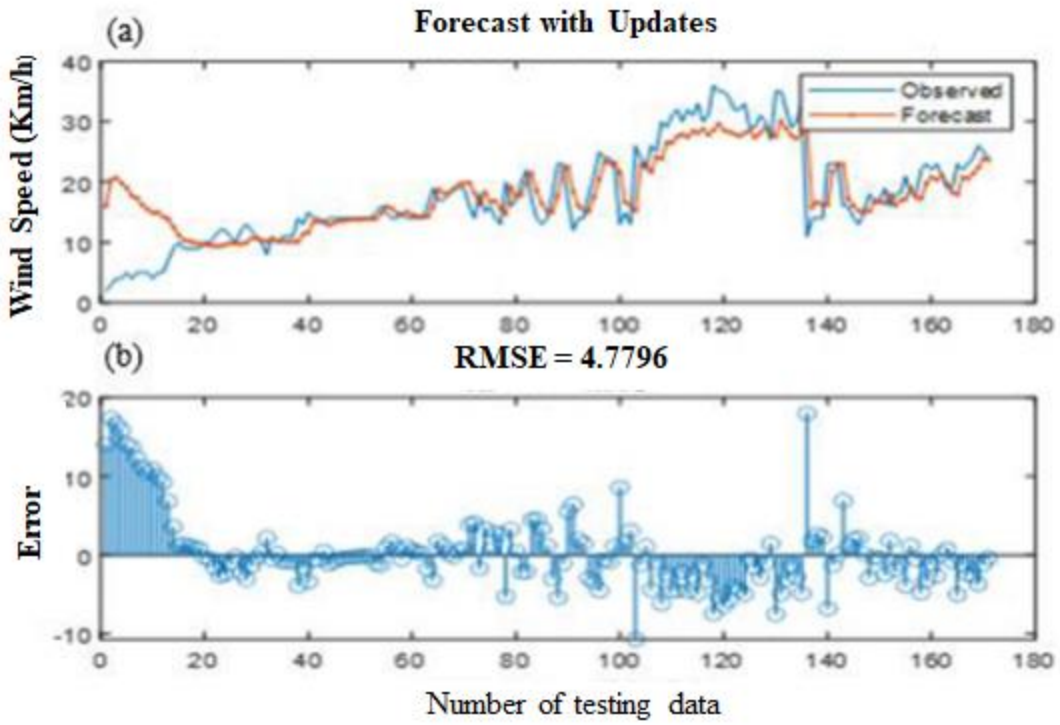

Figure 5.15. Wind speed prediction of proposed LSTM for summer (July 2015) data: (a) Comparison of acual values and forecast values; (b) Error measurements on the RMSE scale.

Table 5.3 shows the prediction error measurements using the RMSE metric for the two (July 2015) and (March 2015) data series when applying the modified LSTM model proposed in this work.

Table 5.3. Error measurements of LSTM model on the RMSE scale.

Time Series		
RMSE	Spring (1–31 March 2015)	Summer (1–31 July 2015).
	8.5128	4.7796

This work focuses on updated long term memory (LSTM), which has been suggested for predicting wind speed. In the case of forecasting wind speed, the observed value updates the network state rather than gates, allowing the real value of the time steps between forecasts to be determined. This is due to the fact that each time the LSTM runs, the state of the cell changes input, forget, and output. A statistical measure known as the root mean square error (RMSE) was used to assess the effectiveness of the model. The results of the model showed an increased ability to predict wind speed.

5.4 Comparison of the ARIMA and LSTM Predictive Models and the Accuracy of Wind Speed Prediction

In this section, wind speed (km/h) data for the Halifax region has been prepared for the period May 1, 2021, to June 20, 2021 (https://climate.weather.gc.ca/historical_data/search_historic_data_e), in order to capture wind speed series characteristics at hourly time scales. Wind speed time series, such as those shown in Figure 5.16, are characterized by basic knowledge of statistics such as time-varying mean, variance, and normal distribution, which are common properties of non-stationary time series. The signal cannot easily yield typical patterns, so this form of data requires extra attention and understanding.

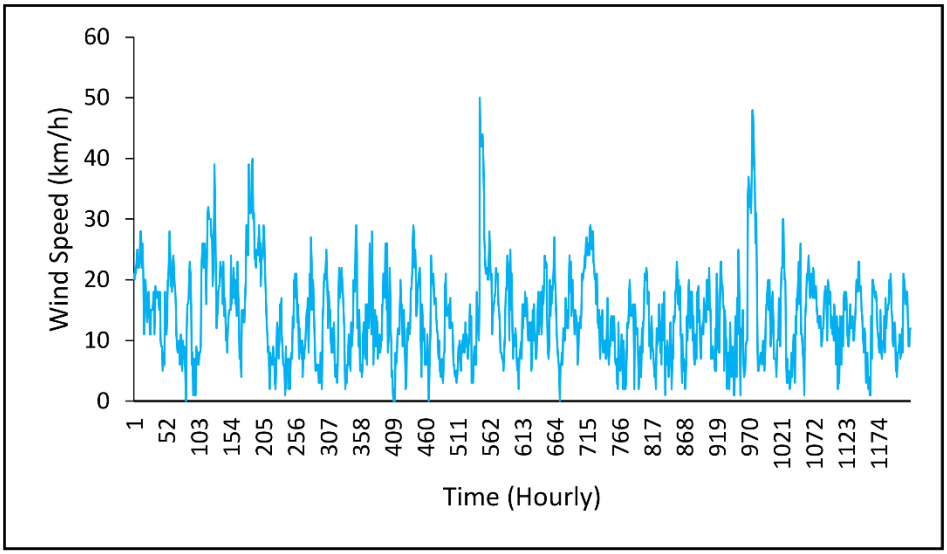

Figure 5.16. The length of the time history of the time series.

ARIMA and LSTM models are compared for forecasting hourly wind speed data in Halifax, Canada.

5.4.1 ARIMA Model

The investigation period to determine the structure of the ARIMA model extends from May 1 to June 20, 2021. The total hourly wind speed time series data was 1224 points, divided into the first 1200 points for building the model and the remaining 24 points for forecasting and evaluating the performance of the model. The model parameters p of AR were determined as the autoregressive part order, and the q-order of MA represented the moving average part order in Figures 5.17 and 5.18, respectively, which represented the ACF and PACF graphs of the wind speed data. From

Figure 5.18, which represents the PACF plot, we notice that the cut-off is at lag 2, which is the order of AR and is acceptable for the actual data.

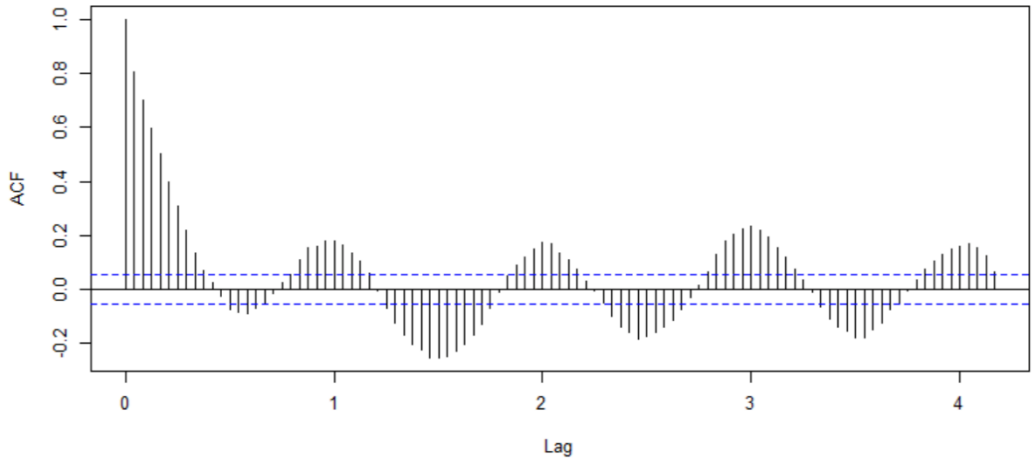

Figure 5.17. Autocorrelation functions for real wind speed data.

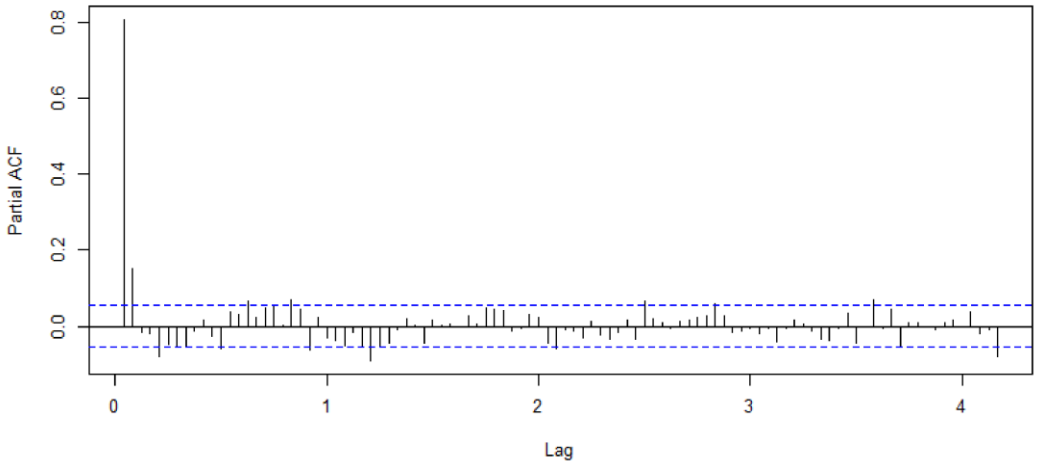

Figure 5.18. Partial autocorrelation functions for real wind speed data.

The series in Figure 5.16 displays recurrent patterns with easily discernible cycles. The underlying processes of interest may exhibit periodic activity, and their identification will depend on the speed or frequency of oscillations that define the behaviour of the main series. The series shows two main types of oscillations: cyclical oscillations, which are ups and downs in the form of sinusoidal waves, and a slower frequency that appears to repeat itself frequently. In general, non-stationary data is difficult to model or predict. Results using non-stationary time series are misleading because they can indicate a relationship between two variables when neither is present. The characteristics of the data are changed from variable variance and a mean that does not remain

closed or revert to a long-run average over time to a fixed long-run average with a constant variance that is independent of time by this transformation. From the autocorrelation function (ACF), it can be concluded that the data series has a nonstationary nature, and the reason for this is the tendency of its values to deteriorate slowly, making it necessary to convert it to a stationary one. This is done by taking the first difference of the original series data, as shown in Figure 5.19.

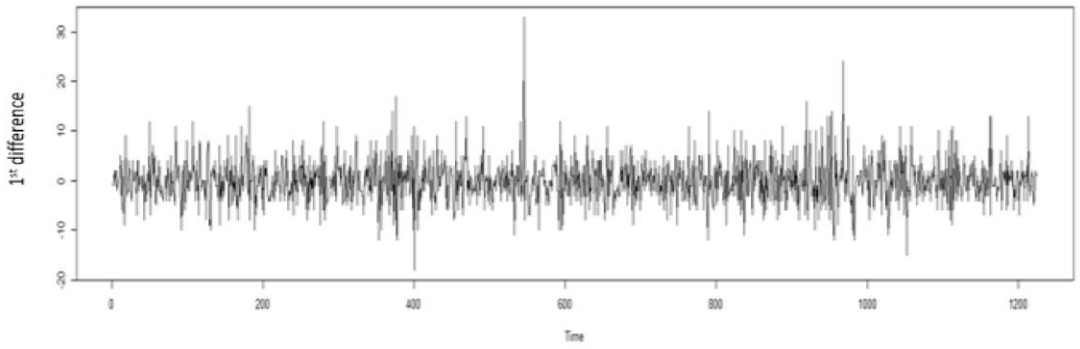

Figure 5.19. First difference of wind speed time series.

To build an ARIMA model, the data must become stationary, as was done by taking the first difference and determining the extent to which the model needs AR or MA terms to correct any autocorrelation that remains in a different series. The values for the parameters p and q obtained from partial autocorrelation and autocorrelation analysis are not accurate. However, the value of the d parameter was considered to be 1, based on what the studies showed because the values of the time series under study must be stationary. The integrated wind speed data are shown in Figures 5.20 and 5.21 ACF and PACF, respectively. It is possible that some of the ARMA (*p*, *q*) models with p and q taking values 1 or 2 will generate large PACF values at a 6 hour lag, but it is not guaranteed.

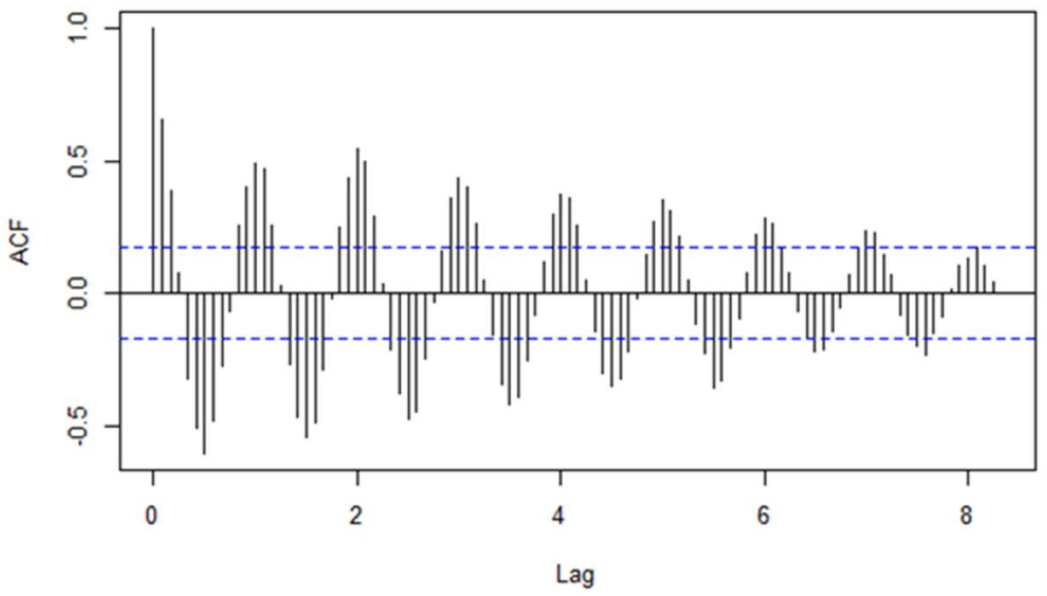

Figure 5.20. Autocorrelation functions for integrated wind speed data

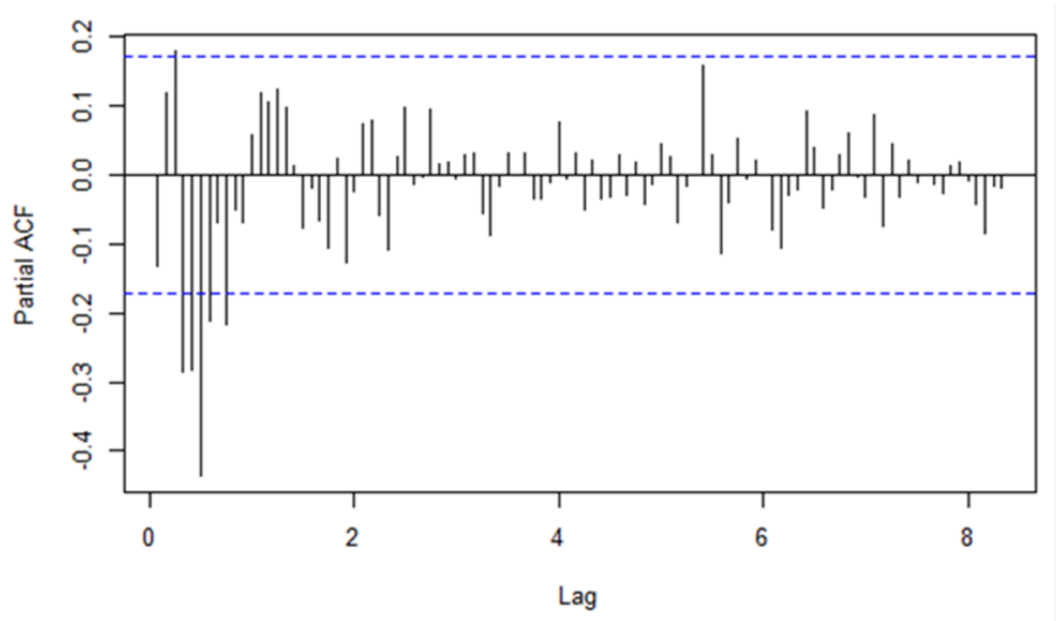

Figure 5.21. Partial autocorrelation functions for integrated wind speed data.

The values of the p and q parameters were chosen to obtain the optimal model. For each set of parameters, a new model was created. To compare the models to one another, RMSE was used. The results are shown in Table 5.4.

Table 5.4. RMSE values for various combinations of orders AR (p) and MA (q).

AR(p)					
MA (q)		0	1	2	3
	0	6.6490	4.4601	4.4275	4.4101
	1	4.4149	4.4101	4.2975	4.2912
	2	4.4149	4.4037	4.2904	4.2996

It is clear from table 5.4 that the lowest value was recorded by the RMSE with the model structure of the ARIMA (2,1,2), and this indicates that it is the optimal model for prediction. Wind speed predictions for the 24 hours ahead were made using the ARIMA (2,1,2) model. Actual wind speed data and forecasted values using the ARIMA (2,1,2) model are shown in Figure 5.22.

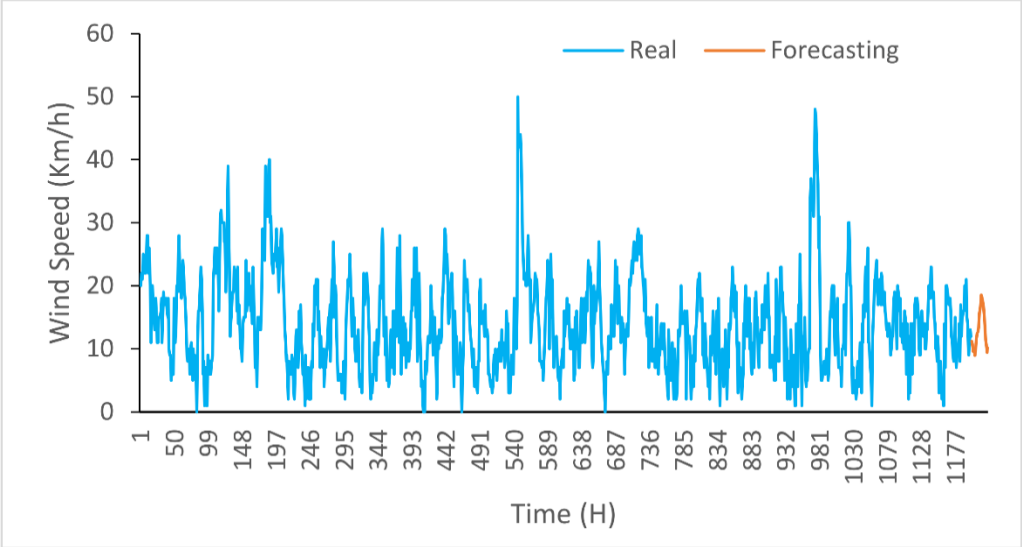

Figure 5.22. Forecasts from ARIMA (2,1,2).

5.4.2 LSTM Model

In this section, the focus was on developing an LSTM regression network to predict wind speed using the MATLAB 2019b environment. The training parameters were optimized to achieve the best results, including a maximum number of 1000 epochs and an initial learning rate of 0.01, which was reduced by 0.2 after 125 epochs. To avoid the problem of gradient explosion, the gradient threshold was set to 1. The resulting model was called a static model because no updates

were made once it fit the training data set accurately. The optimal learning rate was found by testing different initial learning rates on a logarithmic scale, with the best training parameters resulting in the lowest RMSE and loss with a learning rate of 0.01. Figure 5.23 shows that the initial learning rate affects the training process of the network. If the learning rate is small, the training time will increase and may not reach the optimal point, while if it is high, the training results may become suboptimal or divergent. The large degree of weight change can cause the loss function to worsen. In summary, the results of network training improved, and the accuracy of the model increased when using a new initial learning rate of 0.01 in a 24 time-step test. This demonstrates the exceptional performance of LSTM in sequencing long time series data, resulting in a lower RMSE value.

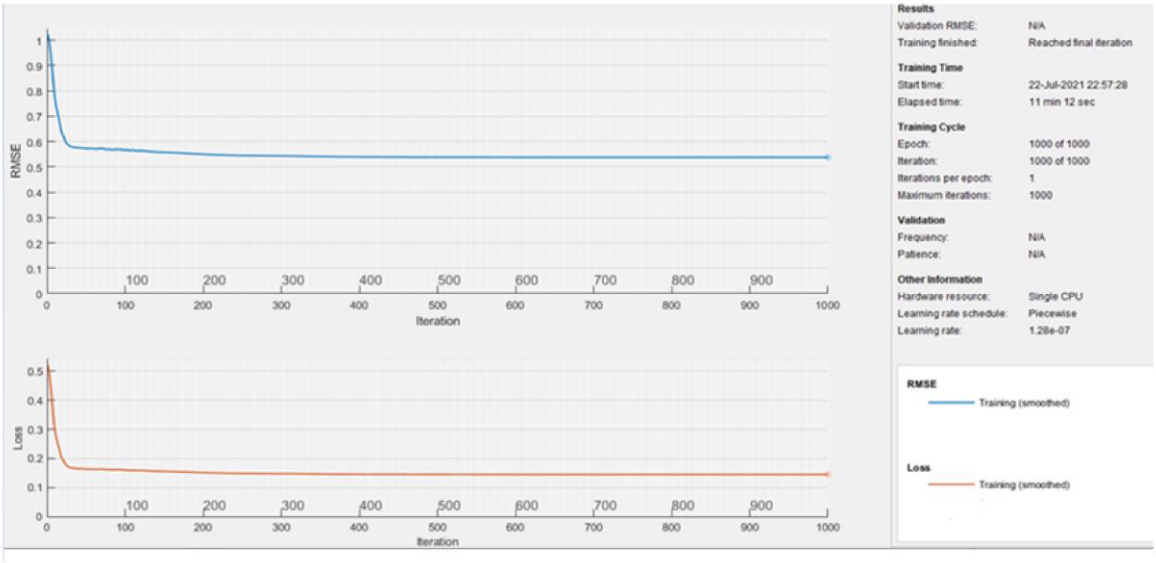

Figure 5.23. Training process with learning rate 0.01 and time step test as 24.

LSTM differs from RNN in the hidden layer mechanism. During LSTM-RNN training, the update gates determine how much information is retained from the last state and how much information is taken from previous layers. The structure of the model is built by all the training data, and after validating each prediction, the model is updated. Thus, the model is ready to add two training epochs before proceeding with the next predictions. The RMSE is recalculated after canceling the prediction using the previously calculated mean and standard deviation. Figure 5.24 presents the results of a 24-step forecast of wind speed. It is noted that the forecasted data of the model in all training epochs is very similar to the real data, with no fluctuations that were too sharp, referring to the RMSE as the best overall test. This indicates that gradient explosion or a vanishing gradient

rarely occurs with the LSTM method. In addition, results were forecasted to fall within a reasonable range with no observable rise or fall.

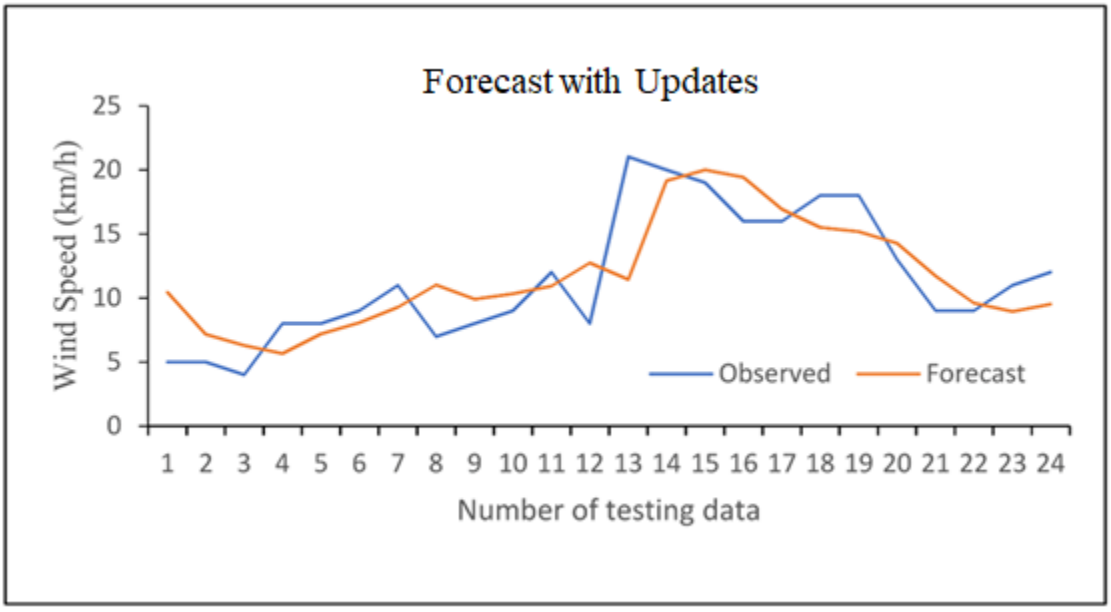

Figure 5.24. Prediction result with 24 time steps.

Figure 5.25 shows the RMSE of all forecasting points of 24-step wind speed.

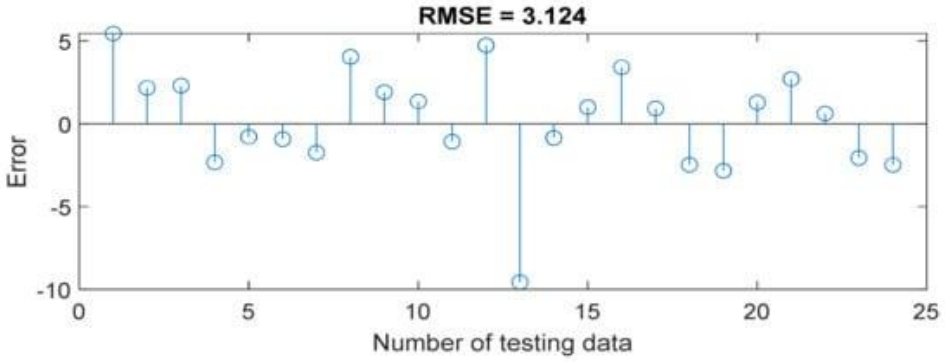

Figure 5.25. RMSE result with 24 time steps.

From the results presented in Table 5.5, it is evident that the LSTM model has the best prediction accuracy in terms of RMSE and MAE compared to the ARIMA model.

Table 5.5. Summary of test statistical errors.

Models	RMSE	MAE
ARIMA	3.423	2.772
LSTM	3.124	2.457

In Figure 5.26, the LSTM and ARIMA models are compared to forecast hourly wind speed data for Halifax, and it is clear that the LSTM model is more accurate than the ARIMA model in tracking the real data.

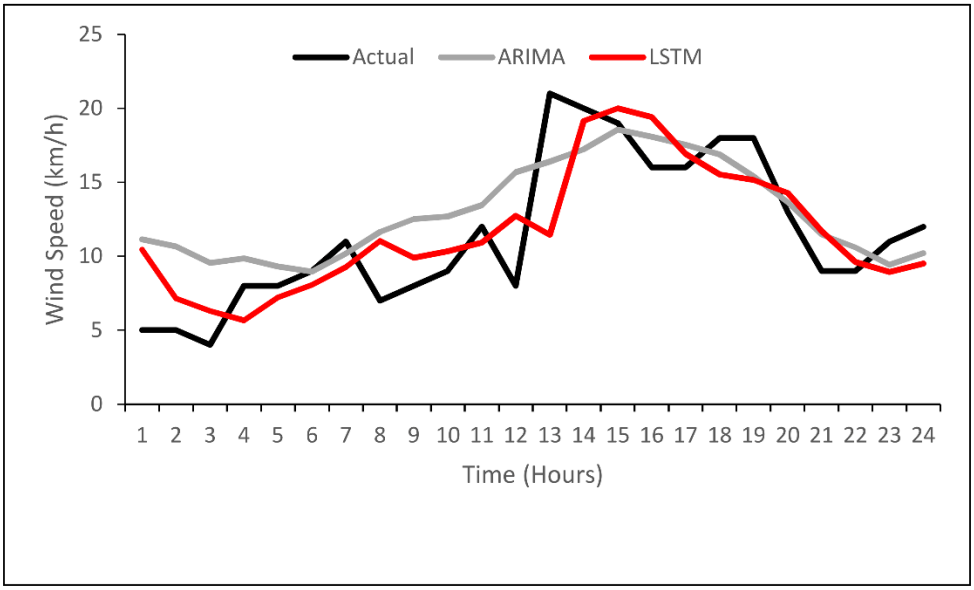

Figure 5.26. Twenty-four hours ahead measured value and forecasted value of wind speed by ARIMA and LSTM models.

After careful study of the results, it can be revealed that the algorithm based on the proposed LSTM model has an advantage for obtaining a better quality of time series prediction compared to the ARIMA model with the same data set, according to different evaluation scales. In comparison to other models, this has become a frequent focus in forecasting operations by researchers. As we know, normal RNNs can handle short term dependencies, but it is not possible to learn long term sequences. LSTM can be implemented with deep learning to obtain results suitable for such sequences. A study of the literature revealed that in previous academic investigations, the ARIMA model achieved superior results with a smaller amount of data. However, the abundance of data in the models proposed for this study shows that algorithms based on deep learning, such as LSTM, have more successful results than traditional methods, such as the ARIMA model.

5.5 Solar Power Forecasting Using Deep Learning Techniques

In this work, LSTM-RNN is proposed for its ability to handle features of time series data. It has a memory unit in the hidden layer that can remember the unit shape, allowing for long-term storage of data and making it suitable for dealing with long-term Photovoltaic (PV) data. The model has

been trained on 43,800 measured daylight hours with a 5-minute solar interval for the Halifax region, covering the days from January 1, 2017 to December 31, 2017, with training sequences at least 5 minutes long. The data is available from the Nova Scotia Community College (NSCC) for the Photovoltaic (PV) forecasting task for the days ahead in short time intervals (30 minutes), based on data previously recorded over one year. Figure 5.27 displays the PV data, measured from 8 a.m. to 5 p.m. with an interval of 5 minutes.

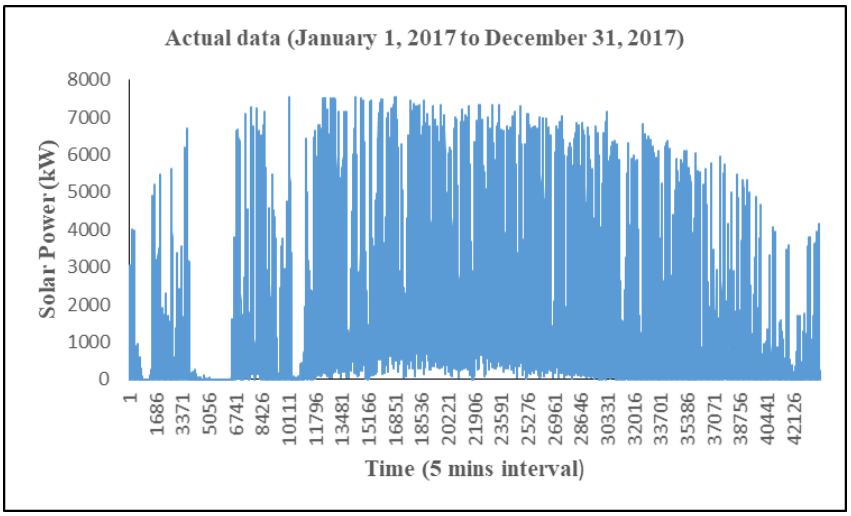

Figure 5.27. Shows the original photoelectric data

This study used a long short-term memory (LSTM) algorithm to make predictions of photovoltaic data for two days (December 31 and June 30) every 30 minutes for a total of 7,300 values. The training data consisted of 7,280 data points, and the last 20 data points were used as a test set to evaluate the effectiveness of the proposed method. The LSTM was trained using MATLAB R2019b with an initial learning rate of 0.01 and a maximum epoch of 1000. The method to determine the optimal number of epochs was based on monitoring the performance of the model on the validation set during training and stopping when the validation error starts to increase (indicating overfitting). The data used in the case study was collected for one year during daylight hours (8 a.m. - 5 p.m.) at 5-minute intervals. This work summarizes the photovoltaic data at 30-minute intervals to make half-hour predictions for the day ahead. The actual daily power generation is compared with the forecast in Figures 5.28 and 5.29, where Figure 5.28 shows the results for a day in December and Figure 5.29 shows the results for a day in June. The results indicate that the forecast is accurate and closely matches the actual data, particularly during the

winter period when the system generated relatively little power. In the summer, the LSTM model performed better than the MLP model.

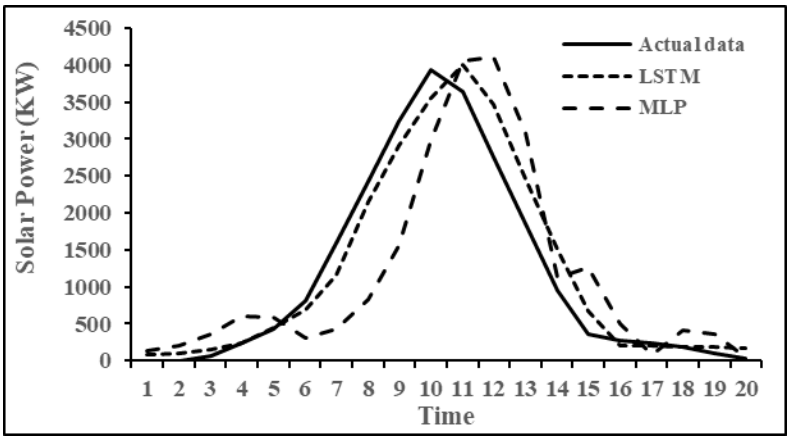

Figure 5.28. Shows one day 30 minutes ahead (December 31).

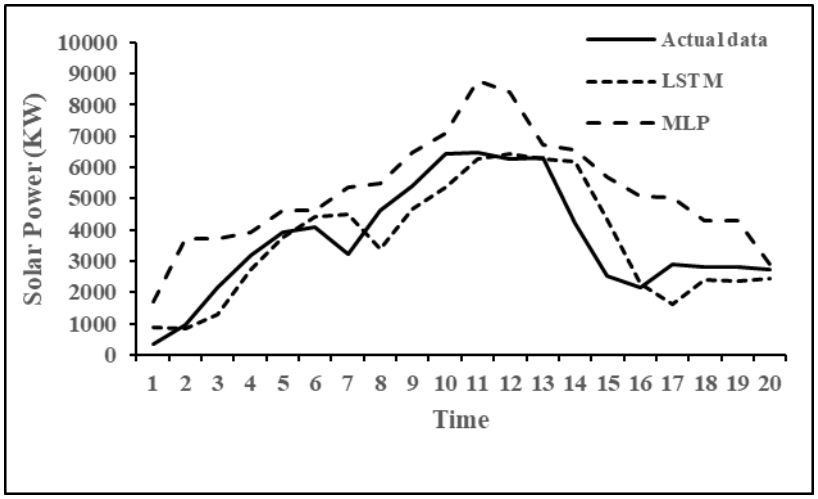

Figure 5.29. Shows one day 30 minutes ahead (June 30).

Based on the information provided in tables 5.6 and 5.7, the LSTM model has outperformed the MLP model in all performance criteria on both winter and summer days.

Table 5.6. Performance assessment for the winter day

	MAE	MAPE	RMSE	R^2
LSTM	236.35	2.17	317.4	0.999
MLP	600.17	4.158	799.9	0.9381

Table 5.7. Performance assessment for the summer day

	MAE	MAPE	RMSE	R^2
LSTM	676.34	0.275	883.5	0.791
MLP	1545.5	0.755	1780.1	0.745

The graph in Figure 5.30 compares the projected values from the LSTM model with the actual values of the 30-minute solar forecast for June. In conclusion, the LSTM model shows good consistency with actual values and has the ability to capture the trend in solar forecasts with minimal fluctuations. According to the results, the model can generalise effectively and avoid overfitting, gradient exploding, or a vanishing gradient problem.

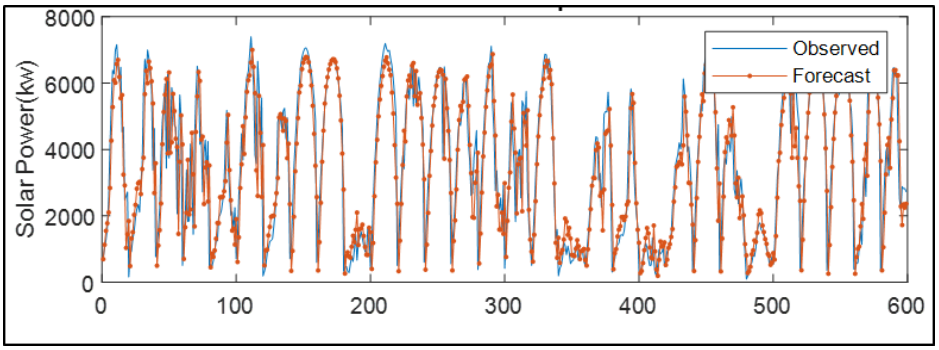

Figure 5.30. Prediction result of June (30 minutes interval)

Figures 5.31(a) and (b) show the results based on the June 2017 solar forecast using the deep learning network LSTM. The first shows the ratio of mean daily solar energy during daylight hours from 8 a.m. to 5 p.m. in June (actual and projected) to time, and the second shows the change in prediction error (RMSE) over time, respectively. The figure demonstrates how well the forecast of the model results matches reality. However, it varies because of a number of atmospheric variables, including water vapour, clouds, and pollutants.

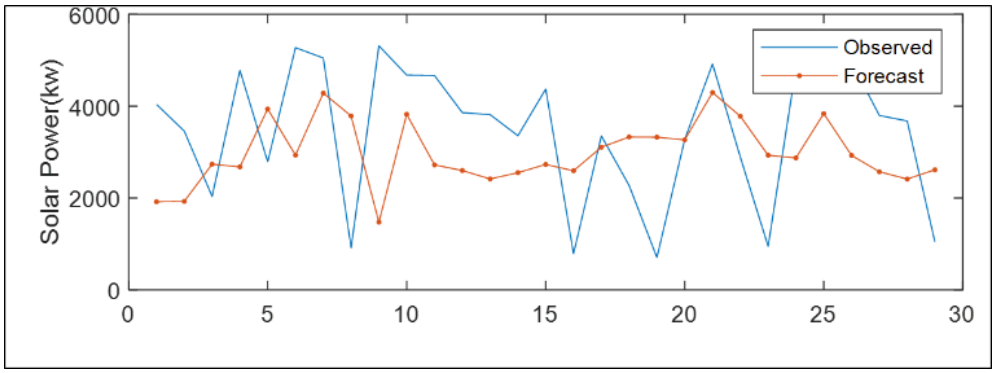

Figure 5.31 (a). Forecast results and actual values in June.

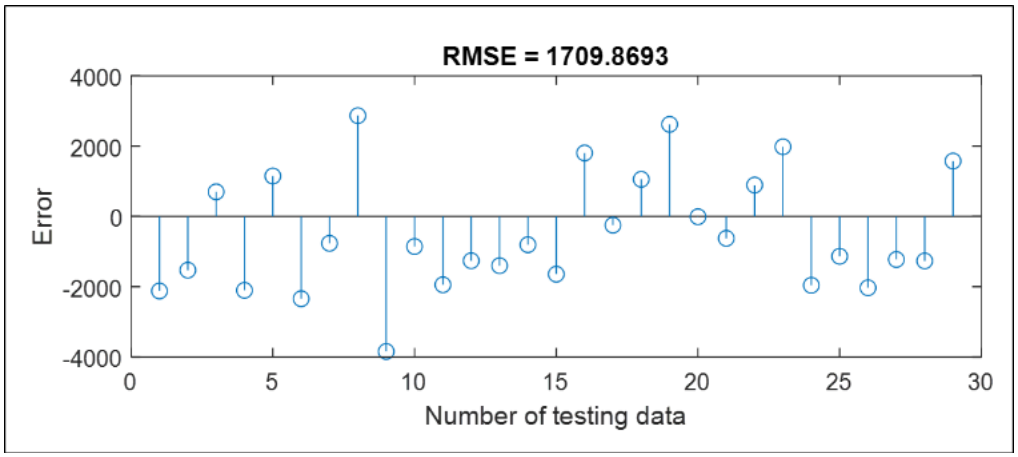

Figure 5.31(b). RMSE result in June.

The results highlight how crucial the time series were in training the models. The quality of the data and its presentation in an appropriate format to train the model is one of the factors that has an effective role in the efficiency obtained by the prediction model. However, the contribution of the different time series to the resulting prediction error has not been analyzed yet. While a particular performance of the model may be acceptable with certain time series data, it may be poor with other time series whose data has different characteristics. As a result, it is practical to take into account a variety of models while searching for reliable predictions. However, a variety of time series forecasting methods can be applied based on the methodology outlined in this paper. In summary of this section, the task of forecasting PV power for the days ahead at 30-min intervals was highlighted. In order to assess and examine the accuracy and performance of the suggested LSTM model, they were compared with those of the MLP model, which is often employed in literature. According to the comparison results, the proposed LSTM model provided values with sufficient accuracy for each category of days for all MAE, MAPE, RMSE, and R2 performance measures. Future photovoltaic power plants will be able to operate more effectively thanks to the proposed reliability of the model data. It seems that combining the concepts of artificial intelligence and energy efficiency promises a bright future, particularly in terms of enhancing energy sustainability, significantly reducing carbon dioxide, and digitizing the electricity sector.

CHAPTER 6

Conclusion and Future Work

6.1 Conclusion

Renewable energy sources are becoming increasingly popular, including solar, wind, hydropower, and geothermal energy, due to their environmental friendliness and sustainability. Unlike fossil fuels, which emit greenhouse gases and contribute to global warming, with the rapid depletion of non-renewable energy sources, it is becoming increasingly important to explore new ways to harness renewable energy and make use of it in a more efficient and sustainable manner. The development of renewable energy forms is crucial for reducing the dependence on finite fossil fuels, mitigating the impacts of climate change, and promoting sustainable economic growth. As environmentally friendly new energy sources, wind and solar energy have become the clean and renewable energy sources with the most potential, the fastest development, and the most mature technology among all new energy sources due to their advantages of easy access and low cost. However, because wind and solar are extremely unstable energy sources and one of the most difficult meteorological elements to forecast, it has become critical to conduct ongoing research to improve the accuracy of wind speed and solar energy forecasts. The main aim of this thesis was to address the challenges posed by the fluctuating nature of renewable energy sources and improve the accuracy of forecasting in these areas. By focusing on wind speed, solar irradiance, and wind power ramp forecasting, the thesis aimed to contribute to the development of more effective and efficient methods for harnessing renewable energy and integrating it into the power grid. The results of the research conducted for this thesis will have important implications for the effective and sustainable development of renewable energy sources, and it will help address some of the challenges associated with their integration into the power grid. Time series modeling and prediction can be broadly categorized into two approaches: traditional methods and deep learning-based methods. Traditional methods, such as the ARIMA family of models, have been developed and widely used for many years and are relatively simple to implement. Deep learning-based methods, such as LSTMs, have been shown to outperform traditional methods in terms of time series prediction accuracy. Despite this, deep learning models are generally more complex and require more computing resources. They also require more specialized knowledge for design and

implementation. On the other hand, traditional models are easier to construct and require fewer computational resources. The choice between traditional and deep learning methods will depend on the specific requirements of the problem at hand and the available resources. The study topic was introduced in Chapter 1, which served as the introduction, and previous studies of the issue highlighted in this dissertation were examined. In addition, the need for further research is substantiated. In particular, section 1.1 describes the growing need for renewable energy sources and presents statistics on the use of wind and solar energy worldwide and Canada's installed wind and solar energy capacity. Section 1.2 is devoted to the main problem that arises when wind and solar energy are integrated into the electrical system, where volatility and uncertainty are among the biggest challenges they pose to the grid and reveal the need for reliable prediction models. Section 1.3 determines the main objective of this work, which is a method for forecasting renewable energy sources. Section 1.4 outlines the methodology adopted in this work which reviews bibliographic research and empirical research. In the bibliographic search, a review of the latest time-series forecasting methods in the context of wind and solar energy and a review of the literature are undertaken in order to assess the state of the art. Experimental research explains the proposed methods, data sources, and processing. Section 1.5 focuses on the contributions of this thesis to the latest developments in wind speed and solar forecasting. The second chapter provides the background and characteristics of wind and solar energy. In particular, Section 2.2 describes the conversion and future of wind energy and its impact on reducing CO_2 levels. Section 2.2.1 discusses mathematical equations and physical concepts related to wind turbines. Sections 2.2.2 and 2.2.3 explain the electric power and power curves for a modern wind turbine. Section 2.3 is a brief explanation of solar energy and the important role it plays in sustainable development on a global scale. Section 2.3.1 shows the mathematical average of the effect of the operating temperature of a photovoltaic cell on electrical efficiency through current and voltage, which represent electrical power. Section 2.3.2 gives a brief definition of the degradation of a photovoltaic system. Section 2.3.3, describes mathematical formulas for estimating the annual solar energy output from the photovoltaic system. Section 2.3.4 shows a simplified explanation of the climatic factors that interfere with the photovoltaic systems to produce electricity and an illustration of the characteristic curves of the panels that represent the change of current with voltage and power with voltage. The third chapter discusses wind and solar energy forecasting as well as related works. Section 3.1 deals with a simplified explanation of the future of renewable

energy forecasting. Sections 3.2 and 3.3 are explanations of wind energy forecasting and solar energy forecasting respectively, and the methods used for those. Section 3.4 highlights the previous work related to the study and the developments reached in the prediction models and their shortcomings to provide a basis on which this work is based. Chapter 4 was dedicated to analyzing time series analysis methods and approaches. Section 4.1 contains a brief introduction. Section 4.2 describes the statistical analysis of time series and their components. Sections 4.4, 4.5, and 4.6 are focused on using neural networks for time series prediction. Section 4.4 covers the use of neural networks for this purpose. Section 4.5 focuses on the use of recurrent neural networks, and section 4.6 focuses on the use of long short-term memory (LSTM) networks specifically. Section 4.7 reviews the most important metrics used to assess forecast accuracy. Chapter 5 presents and discusses the results of the research in graphs and tables. In conclusion, according to the results obtained, it can be concluded that a large amount of data in the models generated in this study shows that algorithms based on deep learning, such as LSTM, perform better than the ARIMA model for predicting wind speed. The results of the proposed LSTM model for solar forecasting show improved performance in all performance parameters, including MAE, MAPE, RMSE, and R2, for each category of days. This suggests that the proposed model is a reliable and effective tool for predicting solar energy production. The integration of AI and energy efficiency concepts has the potential to play a significant role in promoting sustainability, reducing carbon emissions, and digitizing the electricity sector. The results of the simulation tests indicate that the proposed LSTM model has the potential to make a positive impact and help PV plants operate more efficiently in the future.

6.2 Future Work

The need for research in the field of renewable energy is expected to grow in the future as the demand for sustainable energy sources increases. With the increasing size and complexity of renewable energy data, the importance of efficient, effective feature extraction and selection methods will become even more critical. As the renewable energy sector continues to expand, it will be necessary to develop new and innovative methods for handling the large amounts of data generated by renewable energy sources in order to effectively harness and integrate them into the power grid. The ongoing development and improvement of feature extraction and selection methods will be critical to the continued growth and success of the renewable energy sector. Time series exist in all aspects of the objective world, and many data points can be regarded as time

series data. With the rapid development of computer technology and statistical methods, the methods of analyzing time series are changing with each passing day. The choice of model and the quality of the forecasting effect depend on the characteristics of the data itself. Time series forecasting research has been given more and more attention by more and more researchers. Forecasting is the basis of decision-making, so time series forecasting has always been a research field with strong practicability and an important reference value for decision-makers. Decision-making is the continuation of prediction, and the purpose of prediction, to provide a basis for decision-making. The more scientific and reasonable the prediction is the more correct and reliable the decision will be. Although many scholars have done a lot of research on the forecasting of time series, it is still necessary to continuously improve the model to improve the forecasting performance and accuracy, especially with the rapid development of various machine learning algorithms. In view of the fact that the prediction effect of time series sometimes cannot meet the accuracy requirements of practical applications, this paper proposes a new model combining the traditional analysis model and a machine learning model to improve the accuracy of sequence prediction. In order to further improve the prediction effect of the time series model, we highly recommend doing this study again with a new model that combines a traditional analysis model and a machine learning model with more real data to improve sequence prediction accuracy.

LIST OF REFERENCES

1. Mohseni S, Brent AC, Burmester D. Off-Grid Multi-Carrier Microgrid Design Optimisation: The Case of Rakiura–Stewart Island, Aotearoa–New Zealand. Energies. 2021 Oct 11;14(20):6522.
2. Tailor R, Čonka Z, Kolcun M, Beňa Ľ. Electrical Energy Flow Algorithm for Household, Street and Battery Charging in Smart Street Development. Energies. 2021 Jun 23;14(13):3771.
3. Renewable capacity highlights (IRENA) 11 April 2022. Available online https://www.irena.org/publications/2022/Apr/Renewable-Capacity-Statistics-2022
4. Canadian Renewable Energy Association. CanREA's 2050 vision: powering Canada's journey to net-zero. Canadian Renewable Energy Association, Ottawa, Ontario, Canada. 2022.
5. Cui S, Wang YW, Lin X, Xiao JW. Residential virtual power plant with photovoltaic output forecasting and demand response. Asian Journal of Control. 2019 Jul;21(4):1906-17.
6. Council GW. GWEC Global Wind Report 2019. Global Wind Energy Council: Bonn, Germany. 2017.
7. Bullis K. A Cheap Material Boosts Solar Cells by 50 Percent. Materials News, 30 Jan. 2015.
8. Bremnes JB. A comparison of a few statistical models for making quantile wind power forecasts. Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology. 2006 Jan;9(1-2):3-11.
9. Mellit A, Kalogirou SA. Artificial intelligence techniques for photovoltaic applications: A review. Progress in energy and combustion science. 2008 Oct 1;34(5):574-632.
10. Patel MR, Beik O. Wind and solar power systems: design, analysis, and operation. CRC press; 2021 Mar 23.
11. Kumar M. Social, economic, and environmental impacts of renewable energy resources. Wind solar hybrid renewable energy system. 2020 Jan 21;1.
12. Zhang ZS, Sun YZ, Lin J, Cheng L, Li GJ. Versatile distribution of wind power output for a given forecast value. In2012 IEEE Power and Energy Society General Meeting 2012 Jul 22 (pp. 1-7). IEEE.
13. Goutard E. Renewable energy resources in energy management systems. In2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe) 2010 Oct 11 (pp. 1-6). IEEE.
14. De Vos K, Driesen J. Balancing management mechanisms for intermittent power sources—A case study for wind power in Belgium. In2009 6th International Conference on the European Energy Market 2009 May 27 (pp. 1-6). IEEE.
15. Soman SS, Zareipour H, Malik O, Mandal P. A review of wind power and wind speed forecasting methods with different time horizons. InNorth American power symposium 2010 2010 Sep 26 (pp. 1-8). IEEE.
16. Wu YK, Hong JS. A literature review of wind forecasting technology in the world. 2007 IEEE Lausanne Power Tech. 2007 Jul 1:504-9.

17. Lee D, Baldick R. Short-term wind power ensemble prediction based on Gaussian processes and neural networks. IEEE Transactions on Smart Grid. 2013 Sep 23;5(1):501-10.
18. Yadav AP, Kumar A, Behera L. RNN based solar radiation forecasting using adaptive learning rate. In International Conference on Swarm, Evolutionary, and Memetic Computing 2013 Dec 19 (pp. 442-452). Springer, Cham.
19. Mora E, Cifuentes J, Marulanda G. Short-term forecasting of wind energy: A comparison of deep learning frameworks. Energies. 2021 Nov 26;14(23):7943.
20. Lee D, Baldick R. Short-term wind power ensemble prediction based on Gaussian processes and neural networks. IEEE Transactions on Smart Grid. 2013 Sep 23;5(1):501-10.
21. Haykin SS, Haykin SS, editors. Kalman filtering and neural networks. New York: Wiley; 2001 Sep.
22. Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J. LSTM: A search space odyssey. IEEE transactions on neural networks and learning systems. 2016 Jul 8;28(10):2222-32.
23. Phi M. Illustrated guide to lstm's and gru's: A step by step explanation, 2018. URL https://towardsdatascience. com/illustrated-guide-to-lstms-and-gru-sa-step-by-step-explanation-44e9eb85bf21.
24. Feng C, Cui M, Hodge BM, Zhang J. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. Applied Energy. 2017 Mar 15;190:1245-57.
25. Siami-Namini S, Tavakoli N, Namin AS. A comparison of ARIMA and LSTM in forecasting time series. In2018 17th IEEE international conference on machine learning and applications (ICMLA) 2018 Dec 17 (pp. 1394-1401). IEEE.
26. Sherstinsky A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena. 2020 Mar 1;404:132306.
27. Siami-Namini S, Tavakoli N, Namin AS. The performance of LSTM and BiLSTM in forecasting time series. In2019 IEEE International Conference on Big Data (Big Data) 2019 Dec 9 (pp. 3285-3292). IEEE.
28. Alkhayat G, Mehmood R. A review and taxonomy of wind and solar energy forecasting methods based on deep learning. Energy and AI. 2021 Jun 1;4:100060.
29. Bunn, D. and Farmer, E.D., 1985. Comparative models for electrical load forecasting.
30. Thor SE, Weis-Taylor P. Long-term research and development needs for wind energy for the time frame 2000–2020. Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology. 2002 Jan;5(1):73-5.
31. Giebel G, Landberg L, Kariniotakis G, Brownsword R. State-of-the-art Methods and software tools for short-term prediction of wind energy production. InEWEC 2003 (European Wind Energy Conference and exhibition) 2003 Jun 16.
32. Jun LI, Xue DU. Application of sparse Gaussian process in short-term wind power probability prediction. Electric Machines and Control. 2019;23(8):67-77.
33. Anaya-Lara O, Jenkins N, Ekanayake JB, Cartwright P, Hughes M. Wind energy generation: modelling and control. John Wiley & Sons; 2011 Aug 24.
34. Ragheb M, Ragheb AM. Wind turbines theory-the betz equation and optimal rotor tip speed ratio. Fundamental and advanced topics in wind power. 2011 Jun 20;1(1):19-38.

35. Chaudhuri A, Datta R, Kumar MP, Davim JP, Pramanik S. Energy conversion strategies for wind energy system: Electrical, mechanical and material aspects. Materials. 2022 Feb 7;15(3):1232.
36. Patel MR, Beik O. Wind and solar power systems: design, analysis, and operation. CRC press; 2021 Mar 23.
37. Tala H, Patel H, Sapra RR, Gharte JR. Simulations of small scale straight blade Darrieus wind turbine using latest CAE techniques to get optimum power output. International Journal of Advance Foundation and Research in Science & Engineering (IJAFRSE). 2014 Oct;1(5):46-58.
38. Lim SC. Feasibility study of wind energy harvesting at TELCO tower in Malaysia. 670216917. 2019.
39. Kusiak A, Zheng H, Song Z. On-line monitoring of power curves. Renewable Energy. 2009 Jun 1;34(6):1487-93.
40. Ghenai C, Sargsyan N. Wind Energy. InPaths to Sustainable Energy 2010 Dec 30. IntechOpen.
41. Breeze P. Wind power generation. Academic Press; 2016 Jan 21.
42. Adedipe, T.A., 2018. Impact of forest types on wind power (Master's Thesis) Lappeenranta University of Technology, Finland.
43. Choi Y, Rayl J, Tammineedi C, Brownson JR. PV Analyst: Coupling ArcGIS with TRNSYS to assess distributed photovoltaic potential in urban areas. Solar Energy. 2011 Nov 1;85(11):2924 39.
44. Şen Z. Solar energy in progress and future research trends. Progress in energy and combustion science. 2004 Jan 1;30(4):367-416.
45. Choi Y, Suh J, Kim SM. GIS-based solar radiation mapping, site evaluation, and potential assessment: A review. Applied Sciences. 2019 May 13;9(9):1960.
46. Panagea IS, Tsanis IK, Koutroulis AG, Grillakis MG. Climate change impact on photovoltaic energy output: the case of Greece. Advances in Meteorology. 2014 Jan 1;2014.
47. Rana M, Koprinska I, Agelidis VG. Solar power forecasting using weather type clustering and ensembles of neural networks. In2016 International Joint Conference on Neural Networks (IJCNN) 2016 Jul 24 (pp. 4962-4969). IEEE.
48. Cui S, Wang YW, Lin X, Xiao JW. Residential virtual power plant with photovoltaic output forecasting and demand response. Asian Journal of Control. 2019 Jul;21(4):1906-17.
49. Evans DL, Florschuetz LW. Cost studies on terrestrial photovoltaic power systems with sunlight concentration. Solar Energy. 1977 Jan 1;19(3):255-62.
50. Evans DL. Simplified method for predicting photovoltaic array output. Solar energy. 1981 Jan 1;27(6):555-60.
51. Skoplaki E, Palyvos JA. On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations. Solar energy. 2009 May 1;83(5):614-24.
52. Singh P, Ravindra NM. Temperature dependence of solar cell performance—an analysis. Solar energy materials and solar cells. 2012 Jun 1;101:36-45.
53. Skoplaki EP, Palyvos JA. Operating temperature of photovoltaic modules: A survey of pertinent correlations. Renewable energy. 2009 Jan 1;34(1):23-9.

54. Skoplaki E, Boudouvis AG, Palyvos JA. A simple correlation for the operating temperature of photovoltaic modules of arbitrary mounting. Solar energy materials and solar cells. 2008 Nov 1;92(11):1393-402.
55. De Soto W, Klein SA, Beckman WA. Improvement and validation of a model for photovoltaic array performance. Solar energy. 2006 Jan 1;80(1):78-88.
56. Sharma V, Chandel SS. Performance and degradation analysis for long term reliability of solar photovoltaic systems: A review. Renewable and sustainable energy reviews. 2013 Nov 1;27:753-67.
57. Ndiaye A, Charki A, Kobi A, Kébé CM, Ndiaye PA, Sambou V. Degradations of silicon photovoltaic modules: A literature review. Solar Energy. 2013 Oct 1;96:140-51.
58. Jordan DC, Kurtz SR. Photovoltaic degradation rates—an analytical review. Progress in photovoltaics: Research and Applications. 2013 Jan;21(1):12-29.
59. Rajgure, M.K. and Shahakar, D.A. Design and Development of Solar Panel Cleaning Machine. International Research Journal of Engineering and Technology (IRJET),2021, Volume: 08 Issue: 06:2621-2625.
60. Green MA. Developments in crystalline silicon solar cells. Sol. Cell Mater. Wiley, West Sussex, United Kingdom. 2014 Jan 16:65-84.
61. Al-Najideen MI, Alrwashdeh SS. Design of a solar photovoltaic system to cover the electricity demand for the faculty of Engineering-Mu'tah University in Jordan. Resource-Efficient Technologies. 2017 Dec 1;3(4):440-5.
62. Grady WM, Libby L. A cloud shadow model and tracker suitable for studying the impact of high-penetration PV on power systems. In2012 IEEE Energytech 2012 May 29 (pp. 1-6). IEEE.
63. Ebad M, Grady WM. A cloud shadow model for analysis of solar photovoltaic power variability in high-penetration PV distribution networks. In2016 IEEE Power and Energy Society General Meeting (PESGM) 2016 Jul 17 (pp. 1-5). IEEE.
64. Liu F, Zhang Z, Zhao Y, Zhu Z, Pan W, Wang L, Bin X. A method of calculating the daily output power reduction of PV modules due to dust deposition on its surface. IEEE Journal of Photovoltaics. 2019 Mar 25;9(3):881-7.
65. Liangyu M, Jintuo L, Rui M, Jing L, Weiliang L. Output power attenuation rate prediction for photovoltaic panels considering dust deposition in hazy weather. In2017 36th Chinese Control Conference (CCC) 2017 Jul 26 (pp. 4099-4102). IEEE.
66. Desmet J, Debruyne C, Vanalme J, Vandevelde L. Power injection by distributed generation and the influence of harmonic load conditions. InIEEE PES General Meeting 2010 Jul 25 (pp. 1-6). IEEE.
67. Jordan DC, Kurtz SR. Photovoltaic degradation rates—an analytical review. Progress in photovoltaics: Research and Applications. 2013 Jan;21(1):12-29.
68. Riffonneau Y, Bacha S, Barruel F, Ploix S. Optimal power flow management for grid connected PV systems with batteries. IEEE Transactions on sustainable energy. 2011 Feb 17;2(3):309-20.
69. Hasan H, Munawar MR, Siregar RH. Neural network-based solar irradiance forecast for peak load management of grid-connected microgrid with photovoltaic distributed generation. In2017 International Conference on Electrical Engineering and Informatics (ICELTICs) 2017 Oct 18 (pp. 87-90). IEEE.
70. Mow B. STAT FAQs Part 2: Lifetime of PV Panels. NREL Blog, April. 2018 Apr;23:2018.

71. Zhu H, Li X, Sun Q, Nie L, Yao J, Zhao G. A power prediction method for photovoltaic power plant based on wavelet decomposition and artificial neural networks. Energies. 2015 Dec 24;9(1):11.
72. Ye JY, Reindl T, Aberle AG, Walsh TM. Effect of solar spectrum on the performance of various thin-film PV module technologies in tropical Singapore. IEEE Journal of Photovoltaics. 2014 Jul 11;4(5):1268-74.
73. Koehl M, Heck M, Wiesmeier S, Wirth J. Modeling of the nominal operating cell temperature based on outdoor weathering. Solar Energy Materials and Solar Cells. 2011 Jul 1;95(7):1638-46.
74. Ali HG, Vilanova R, Pelez-Restrepo J. Perturb & Observe based Adaptive Sliding Mode MPPT Control of Solar Photovoltaic System. In2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe) 2020 Jun 9 (pp. 1-6). IEEE.
75. Islam MA, Kassim NM, Alkahtani AA, Amin N. Assessing the impact of spectral irradiance on the performance of different photovoltaic technologies.
76. Yin J, Molini A, Porporato A. Impacts of solar intermittency on future photovoltaic reliability. Nature communications. 2020 Sep 22;11(1):1-9.
77. Brownlee J. Time series prediction with lstm recurrent neural networks in python with keras. Machine Learning Mastery. 2016 Jul 21;18.
78. Cleary K, Palmer K. Renewables 101: Integrating Renewable Energy. Resources into the Grid. 2020. Resources for the Future
79. Chernyakhovskiy I. Forecasting Wind and Solar Generation: Improving System Operations, Greening the Grid. National Renewable Energy Lab.(NREL), Golden, CO (United States); 2016 Jan 1.
80. Fischer A, Montuelle L, Mougeot M, Picard D. Statistical learning for wind power: A modeling and stability study towards forecasting. Wind Energy. 2017 Dec;20(12):2037-47.
81. Chen C, Pang Y. Exploring Machine Learning Techniques for Smart Drainage System. In2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService) 2019 Apr 4 (pp. 63-70). IEEE.
82. Optis M, Perr-Sauer J. The importance of atmospheric turbulence and stability in machine-learning models of wind farm power production. Renewable and Sustainable Energy Reviews. 2019 Sep 1;112:27-41.
83. Feng C, Cui M, Hodge BM, Zhang J. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. Applied Energy. 2017 Mar 15;190:1245-57.
84. Giebel G, Landberg L, Kariniotakis G, Brownsword R. State-of-the-art Methods and software tools for short-term prediction of wind energy production. InEWEC 2003 (European Wind Energy Conference and exhibition) 2003 Jun 16.
85. Khalid M, Savkin AV. A method for short-term wind power prediction with multiple observation points. IEEE Transactions on Power Systems. 2012 Jan 31;27(2):579-86.
86. Serendipity Initial value vs. boundary value problems. 2010. Available at: https://www.easterbrook.ca/steve/2010/01/initial-value-vs-boundary-value-problems/
87. Shouman, Enas Raafat Maamoun. "Wind Power Forecasting Models." (2022).
88. Yan J, Liu YQ, Zhang H. Dynamic wind power probabilistic forecasting based on wind scenario recognition. Modern Electr. Power. 2016 Apr;33(2):51-8.

89. Ghorbani MA, Khatibi R, Hosseini B, Bilgili M. Relative importance of parameters affecting wind speed prediction using artificial neural networks. Theoretical and Applied Climatology. 2013 Oct;114(1):107-14.
90. Kavasseri RG, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renewable Energy. 2009 May 1;34(5):1388-93.
91. Chen P, Pedersen T, Bak-Jensen B, Chen Z. ARIMA-based time series model of stochastic wind power generation. IEEE transactions on power systems. 2009 Nov 24;25(2):667-76.
92. Erdem E, Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction. Applied Energy. 2011 Apr 1;88(4):1405-14.
93. Jung S, Kwon SD. Weighted error functions in artificial neural networks for improved wind energy potential estimation. Applied energy. 2013 Nov 1;111:778-90.
94. Makridakis S, Wheelwright SC, Hyndman RJ. Forecasting methods and applications. John wiley & sons; 2008.
95. Ding Z, Yang P, Yang X, Zhang Z. Wind power prediction method based on sequential time clustering support vector machine. Dianli Xitong Zidonghua(Automation of Electric Power Systems). 2012;36(14) 131-
96. Yang J, Zhang L, Wang MQ, Han X. Wind-power fore-casting error simulation considering output level and self-correlation. Electric Power Automation Equipment. 2017;37(09):96-102.
97. Shi KP, Qiao Y, Zhao W, Huang SL, Liu ZJ, Guo L. Short-term wind power prediction based on entropy association information mining of historical data. Autom. Electr. Power Syst. 2017 Feb 10;41(3):13-8.
98. Hu Q H, Zhang R J, Zhou Y C. Transfer learning for short-term wind speed prediction with deep neural networks[J]. Renewable Energy, 2016, 85: 83-95. DOI:10.1016/j.renene.2015.06.034
99. Lv Y, Duan Y, Kang W, Li Z, Wang FY. Traffic flow prediction with big data: a deep learning approach. IEEE Transactions on Intelligent Transportation Systems. 2014 Sep 9;16(2):865-73.
100. Ju Y, Sun G, Chen Q, Zhang M, Zhu H, Rehman MU. A model combining convolutional neural network and LightGBM algorithm for ultra-short-term wind power forecasting. Ieee Access. 2019 Feb 27;7:28309-18.
101. Jiao R, Huang X, Ma X, Han L, Tian W. A model combining stacked auto encoder and back propagation algorithm for short-term wind power forecasting. Ieee Access. 2018 Mar 22;6:17851-8.
102. Cui M, Ke D, Sun Y, Gan D, Zhang J, Hodge BM. Wind power ramp event forecasting using a stochastic scenario generation method. IEEE Transactions on sustainable energy. 2015 Jan 23;6(2):422-33.
103. Dasgupta S, Osogami T. Nonlinear dynamic Boltzmann machines for time-series prediction//The 31 st AAAI Conference on Artificial Intelligence. San Francisco, USA: AAAI, 2017, 31(1): 1833-1839.
104. Jain A, Kumar A M. Hybrid neural network models for hydrologic time series forecasting. Applied Soft Computing, 2007, 7(2): 585-592.
105. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation, 1997, 9(8): 1735-1780.

106. Lai G, Chang WC, Yang Y, Liu H. Modeling long-and short-term temporal patterns with deep neural networks. InThe 41st international ACM SIGIR conference on research & development in information retrieval 2018 Jun 27 (pp. 95-104).
107. Wan C, Zhao J, Song Y, Xu Z, Lin J, Hu Z. Photovoltaic and solar power forecasting for smart grid energy management. CSEE Journal of Power and Energy Systems. 2015 Dec;1(4):38-46.
108. Wang G, Su Y, Shu L. One-day-ahead daily power forecasting of photovoltaic systems based on partial functional linear regression models. Renewable Energy. 2016 Oct 1;96:469-78.
109. De Giorgi MG, Congedo PM, Malvoni M. Photovoltaic power forecasting using statistical methods: impact of weather data. IET Science, Measurement & Technology. 2014 May;8(3):90-7.
110. Zeng J, Qiao W. Short-term solar power prediction using a support vector machine. Renewable energy. 2013 Apr 1;52:118-27.
111. Antonanzas J, Osorio N, Escobar R, Urraca R, Martinez-de-Pison FJ, Antonanzas-Torres F. Review of photovoltaic power forecasting. Solar energy. 2016 Oct 15;136:78-111.
112. De Leone R, Pietrini M, Giovannelli A. Photovoltaic energy production forecast using support vector regression. Neural Computing and Applications. 2015 Nov;26(8):1955-62.
113. Cococcioni M, D'Andrea E, Lazzerini B. One day-ahead forecasting of energy production in solar photovoltaic installations: An empirical study. Intelligent Decision Technologies. 2012 Jan 1;6(3):197-210.
114. Yadav HK, Pal Y, Tripathi MM. Photovoltaic power forecasting methods in smart power grid. In2015 annual IEEE India conference (INDICON) 2015 Dec 17 (pp. 1-6). IEEE.
115. Cheng WY, Liu Y, Bourgeois AJ, Wu Y, Haupt SE. Short-term wind forecast of a data assimilation/weather forecasting system with wind turbine anemometer measurement assimilation. Renewable Energy. 2017 Jul 1;107:340-51.
116. Guan L, Zhao Q, Zhou B, Lyu Y, Zhao W, Yao W. Multi-scale clustering analysis based modeling of photovoltaic power characteristics and its application in prediction. Automation of Electric Power Systems. 2018;42(15):24-30.
117. Masseran N, Razali AM, Ibrahim K, Latif MT. Fitting a mixture of von Mises distributions in order to model data on wind direction in Peninsular Malaysia. Energy Conversion and Management. 2013 Aug 1;72:94-102.
118. Goodfellow I, Bengio Y, Courville A. Deep learning. MIT press; 2016 Nov 10.
119. Cui M, Ke D, Sun Y, Gan D, Zhang J, Hodge BM. Wind power ramp event forecasting using a stochastic scenario generation method. IEEE Transactions on sustainable energy. 2015 Jan 23;6(2):422-33.
120. Lai G, Chang WC, Yang Y, Liu H. Modeling long-and short-term temporal patterns with deep neural networks. In the 41st international ACM SIGIR conference on research & development in information retrieval 2018 Jun 27 (pp. 95-104).
121. Guo H M, Wu B J, Jiang X R, et al. Research on detection method of modal power using photonic lantern[J]. Acta Optica Sinica, 2022, 42(1): 0106003
122. Box GE, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. John Wiley & Sons; 2015 May 29.
123. Kärner O. ARIMA representation for daily solar irradiance and surface air temperature time series. Journal of Atmospheric and Solar-Terrestrial Physics. 2009 Jun 1;71(8-9):841-7.

124. Craggs C, Conway EM, Pearsall NM. Statistical investigation of the optimal averaging time for solar irradiance on horizontal and vertical surfaces in the UK. Solar Energy. 2000 Feb 1;68(2):179-87.
125. Kalnay E. Atmospheric modeling, data assimilation and predictability. Cambridge university press; 2003.
126. Pu Z, Kalnay E. Numerical weather prediction basics: Models, numerical methods, and data assimilation. InHandbook of hydrometeorological ensemble forecasting 2019 (pp. 67-97). Springer Berlin Heidelberg.
127. Wolff B, Kühnert J, Lorenz E, Kramer O, Heinemann D. Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data. Solar Energy. 2016 Oct 1;135:197-208.
128. Rahimikhoob A. Estimating global solar radiation using artificial neural network and air temperature data in a semi-arid environment. Renewable energy. 2010 Sep 1;35(9):2131-5.
129. Premalatha N, Valan Arasu A. Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. Journal of applied research and technology. 2016 Jun;14(3):206-14.
130. Qazi A, Fayaz H, Wadi A, Raj RG, Rahim NA, Khan WA. The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review. Journal of cleaner production. 2015 Oct 1;104:1-2.
131. Reikard G, Hansen C. Forecasting solar irradiance at short horizons: Frequency and time domain models. Renewable energy. 2019 May 1;135:1270-90.
132. Qing X, Niu Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy. 2018 Apr 1;148:461-8.
133. Guinot J, Barghouti Z, Chiva R. Understanding Green Innovation: A Conceptual Framework. Sustainability. 2022 Jan;14(10):5787.
134. Jin X, Yu X, Wang X, Bai Y, Su T, Kong J. Prediction for Time Series with CNN and LSTM. InProceedings of the 11th International Conference on Modelling, Identification and Control (ICMIC2019) 2020 (pp. 631-641). Springer, Singapore.
135. Dubey AK, Kumar A, García-Díaz V, Sharma AK, Kanhaiya K. Study and analysis of SARIMA and LSTM in forecasting time series data. Sustainable Energy Technologies and Assessments. 2021 Oct 1;47:101474.
136. Liang S, Nguyen L, Jin F. A multi-variable stacked long-short term memory network for wind speed forecasting. In2018 IEEE international conference on big data (Big Data) 2018 Dec 10 (pp. 4561-4564). IEEE.
137. Wen C, Liu S, Yao X, Peng L, Li X, Hu Y, Chi T. A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. Science of the total environment. 2019 Mar 1;654:1091-9.
138. Kim K, Kim DK, Noh J, Kim M. Stable forecasting of environmental time series via long short term memory recurrent neural network. IEEE Access. 2018 Dec 4;6:75216-28.
139. Du S, Li T, Yang Y, Horng SJ. Deep air quality forecasting using hybrid deep learning framework. IEEE Transactions on Knowledge and Data Engineering. 2019 Nov 20;33(6):2412-24.

140. Bae KY, Jang HS, Sung DK. Hourly solar irradiance prediction based on support vector machine and its error analysis. IEEE Transactions on Power Systems. 2016 May 20;32(2):935-45.
141. Huang Y, Cao C, Gu H. Short-term photovoltaic power generation forecasting scheme based on IKFCM and multi-mode social spider optimization SVR. Power System Protection and Control. 2018;46(24):96-103.
142. Alzahrani A, Shamsi P, Dagli C, Ferdowsi M. Solar irradiance forecasting using deep neural networks. Procedia Computer Science. 2017 Jan 1;114:304-13.
143. Zufferey T, Ulbig A, Koch S, Hug G. Forecasting of smart meter time series based on neural networks. InInternational workshop on data analytics for renewable energy integration 2016 Sep 23 (pp. 10-21). Springer, Cham.
144. Chou K.C., Corotis R.B., Simulation of hourly wind speed and array wind power, Solar Energy 26 (1981) 199-212.
145. Goh T.N., Nathan G.K., A statistical methodology for study of wind characteristics from a close array of stations, Wind Engineering 3 (1979) 197-206.
146. Brown B.G., Katz R.W., Murphy A.H., Time Series Models to Simulate and Forecast Wind Speed and Wind Power, Journal of Climate and Applied Meteorology 23 (1984) 1184-1195.
147. Daniel A.R., Chen A.A., Stochastic simulation and forecasting of hourly average wind speed sequences in Jamaica, Solar Energy 46 (1991) 1-11.
148. Poggi P., Muselli M., Notton G., Cristofari C., Louche A., Forecasting and simulating wind speed in Corsica by using an autoregressive model, Energy Conversion and Management 44 (2003) 3177-3196.
149. Nfaoui H., Buret J., Saygh A.A.M., Stochastic simulation of hourly average wind speed sequences in Tangiers, Solar Energy 56 (1996) 301-314.
150. Torres J.L., García A., De Blas M., De Francisco A., Forecast of hourly average wind speed with ARMA models in Navarre (Spain), Solar Energy 79 (2005) 65-77.
151. Kavasseri RG, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renewable Energy. 2009 May 1;34(5):1388-93.
152. Li H, Liu H, Ji H, Zhang S, Li P. Ultra-short-term load demand forecast model framework based on deep learning. Energies. 2020 Jan;13(18):4900.
153. Qu Z, Xu J, Wang Z, Chi R, Liu H. Prediction of electricity generation from a combined cycle power plant based on a stacking ensemble and its hyperparameter optimization with a grid-search method. Energy. 2021 Jul 15;227:120309.
154. Mishra S, Bordin C, Taharaguchi K, Palu I. Comparison of deep learning models for multivariate prediction of time series wind power generation and temperature. Energy Reports. 2020 Feb 1;6:273-86.
155. Xu C, Wang J, Zhang J, Li X. Anomaly detection of power consumption in yarn spinning using transfer learning. Computers & Industrial Engineering. 2021 Feb 1;152:107015.
156. Liu H, Gao Q, Ma P. Photovoltaic generation power prediction research based on high quality context ontology and gated recurrent neural network. Sustainable Energy Technologies and Assessments. 2021 Jun 1;45:101191
157. Saberian A, Hizam H, Radzi MA, Ab Kadir MZ, Mirzaei M. Modelling and prediction of photovoltaic power output using artificial neural networks. International journal of Photoenergy. 2014 Apr 6;2014.

158. Kärner O. ARIMA representation for daily solar irradiance and surface air temperature time series. Journal of Atmospheric and Solar-Terrestrial Physics. 2009 Jun 1;71(8-9):841-7.
159. Palomares-Salas JC, De La Rosa JJ, Ramiro JG, Melgar J, Aguera A, Moreno A. ARIMA vs. Neural networks for wind speed forecasting. In2009 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications 2009 May 11 (pp. 129-133). IEEE.
160. Cadenas E, Rivera W. Wind speed forecasting in the south coast of Oaxaca, Mexico. Renewable energy. 2007 Oct 1;32(12):2116-28.
161. Rashid MH. AMI smart meter big data analytics for time series of electricity consumption. In2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE) 2018 Aug 1 (pp. 1771-1776). IEEE.
162. Reikard G. Predicting solar radiation at high resolutions: A comparison of time series forecasts. Solar energy. 2009 Mar 1;83(3):342-9.
163. Yang D, Jirutitijaroen P, Walsh WM. Hourly solar irradiance time series forecasting using cloud cover index. Solar Energy. 2012 Dec 1;86(12):3531-43.
164. Rajagopalan S, Santoso S. Wind power forecasting and error analysis using the autoregressive moving average modeling. In2009 IEEE Power & Energy Society General Meeting 2009 Jul 26 (pp. 1-6). IEEE.
165. Abdelaziz AY, Rahman MA, El-Khayat M, Hakim M. Short term wind power forecasting using autoregressive integrated moving average modeling. InProceedings of the 15th International Middle East Power Systems Conference, Alexandria, Egypt 2012 Dec 23 (pp. 23-25).
166. Wang J, Zhou Q, Zhang X. Wind power forecasting based on time series ARMA model. InIOP Conference Series: Earth and Environmental Science 2018 Dec 1 (Vol. 199, No. 2, p. 022015). IOP Publishing.
167. Tyass I, Bellat A, Raihani A, Mansouri K, Khalili T. Wind speed prediction based on seasonal ARIMA model. InE3S Web of Conferences 2022 (Vol. 336, p. 00034). EDP Sciences.
168. Chellali F. Short-Term Wind Forecasting in Adrar, Algeria, Using a Combined System. Engineering Proceedings. 2023 Jan 13;29(1):11.
169. Zhang J, Verschae R, Nobuhara S, Lalonde JF. Deep photovoltaic nowcasting. Solar Energy. 2018 Dec 1;176:267-76.
170. Wen L, Zhou K, Yang S, Lu X. Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting. Energy. 2019 Mar 15;171:1053-65.
171. Zhu Q, Li H, Wang Z, Chen JF, Wang BJ. Short-term wind power forecasting based on LSTM. Power System Technology. 2017 Dec;41(12):3797-802.
172. Qing X, Niu Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy. 2018 Apr 1;148:461-8.
173. Liu, H.; Mi, X.; Li, Y. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short-term memory neural network and Elman neural network. Energy Convers. Manag. 2018, 156, 498–514.
174. Madhiarasan, M. Accurate prediction of different forecast horizons wind speed using a recursive radial basis function neural network. Prot. Control. Mod. Power Syst. 2020, 5, 22.

175. Tian, Z.; Wang, G.; Ren, Y. Short-term wind speed forecasting based on autoregressive moving average with echo state network compensation. Wind. Eng. 2020, 44, 152–167.
176. Yu R, Gao J, Yu M, Lu W, Xu T, Zhao M, Zhang J, Zhang R, Zhang Z. LSTM-EFG for wind power forecasting based on sequential correlation features. Future Generation Computer Systems. 2019 Apr 1;93:33-42.
177. Qin Y, Li K, Liang Z, Lee B, Zhang F, Gu Y, Zhang L, Wu F, Rodriguez D. Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal. Applied energy. 2019 Feb 15;236:262-72
178. Singh SN, Mohapatra A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. Renewable energy. 2019 Jun 1;136:758-68.
179. Wu, W.; Chen, K.; Qiao, Y.; Lu, Z. Probabilistic short-term wind power forecasting based on deep neu-ral networks. In Proceedings of the 2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Beijing, China, 16 October 2016; pp. 1–8.
180. Sandhu, K.S.; Nair, A.R. A comparative study of ARIMA and RNN for short term wind speed forecasting. In Proceedings of the 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 6 July 2019; pp. 1–7.
181. Werngren, S. Comparison of Different Machine Learning Models for Wind Turbine Power Predictions; Uppsala University: Uppsala, Sweden, 2018.
182. Elsaraiti, M.; Merabet, A.; Al-Durra, A. Time Series Analysis and Forecasting of Wind Speed Data. In Proceedings of the IEEE Industry Applications Society Annual Meeting, Baltimore, MD, USA, 29 September–3 October 2019; pp. 1–5.
183. Kumar, D.; Mathur, H.D.; Bhanot, S.; Bansal, R.C. Forecasting of solar and wind power using LSTM RNN for load frequency control in isolated microgrid. Int. J. Model. Simul. 2020.
184. Cao, Q.; Ewing, B.T.; Thompson, M.A. Forecasting wind speed with recurrent neural networks. Eur. J. Oper. Res. 2012, 221, 148–154.
185. Özen, C.; Kaplan, O.; Özcan, C.; Dinç, U. Short Term Wind Speed Forecast by Using Long Short Term Memory. In Proceedings of the 9th International Symposium on Atmospheric Sciences ATMOS 2019, Istanbul, Turkey, 23–26 October 2019.
186. Kalogirou SA. Artificial neural networks in renewable energy systems applications: a review. Renewable and sustainable energy reviews. 2001 Dec 1;5(4):373-401.
187. Elsaraiti M, Merabet A. A comparative analysis of the arima and lstm predictive models and their effectiveness for predicting wind speed. Energies. 2021 Oct 18;14(20):6782.
188. Young SR, Rose DC, Karnowski TP, Lim SH, Patton RM. Optimizing deep learning hyper-parameters through an evolutionary algorithm. InProceedings of the workshop on machine learning in high-performance computing environments 2015 Nov 15 (pp. 1-5).
189. Kalogirou SA. Solar thermal power systems. Solar Energy Engineering. 2009:521-52.
190. Chen J, Yu J, Song M, Valdmanis V. Factor decomposition and prediction of solar energy consumption in the United States. Journal of Cleaner Production. 2019 Oct 10;234:1210-20.
191. Elsaraiti M, Merabet A. Application of long-short-term-memory recurrent neural networks to forecast wind speed. Applied Sciences. 2021 Mar 8;11(5):2387.

192. Bottieau J, Vallée F, De Grève Z, Toubeau JF. Leveraging provision of frequency regulation services from wind generation by improving day-ahead predictions using LSTM neural networks. In2018 IEEE International Energy Conference (ENERGYCON) 2018 Jun 3 (pp. 1-6). IEEE.
193. Rabehi A, Guermoui M, Lalmi D. Hybrid models for global solar radiation prediction: a case study. International Journal of Ambient Energy. 2020 Jan 2;41(1):31-40.
194. Abdel-Nasser M, Mahmoud K. Accurate photovoltaic power forecasting models using deep LSTM-RNN. Neural Computing and Applications. 2019 Jul;31(7):2727-40.
195. Han S, Qiao YH, Yan J, Liu YQ, Li L, Wang Z. Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network. Applied energy. 2019 Apr 1;239:181-91.
196. Diagne M, David M, Lauret P, Boland J, Schmutz N. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renewable and Sustainable Energy Reviews. 2013 Nov 1;27:65-76.
197. Jin X, Yu X, Wang X, Bai Y, Su T, Kong J. Prediction for Time Series with CNN and LSTM. InProceedings of the 11th International Conference on Modelling, Identification and Control (ICMIC2019) 2020 (pp. 631-641). Springer, Singapore.
198. Dubey AK, Kumar A, García-Díaz V, Sharma AK, Kanhaiya K. Study and analysis of SARIMA and LSTM in forecasting time series data. Sustainable Energy Technologies and Assessments. 2021 Oct 1;47:101474.
199. Liang S, Nguyen L, Jin F. A multi-variable stacked long-short term memory network for wind speed forecasting. In2018 IEEE international conference on big data (Big Data) 2018 Dec 10 (pp. 4561-4564). IEEE.
200. Du SS, Zhai X, Poczos B, Singh A. Gradient descent provably optimizes over-parameterized neural networks. arXiv preprint arXiv:1810.02054. 2018 Oct 4.
201. Rashid MH. AMI smart meter big data analytics for time series of electricity consumption. In2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE) 2018 Aug 1 (pp. 1771-1776). IEEE.
202. Zufferey T, Ulbig A, Koch S, Hug G. Forecasting of smart meter time series based on neural networks. InInternational workshop on data analytics for renewable energy integration 2016 Sep 23 (pp. 10-21). Springer, Cham.
203. Pole A, West M, Harrison J. Applied Bayesian forecasting and time series analysis. Chapman and Hall/CRC; 2018 Oct 8.
204. Kalimoldayev M, Drozdenko A, Koplyk I, Marinich T, Abdildayeva A, Zhukabayeva T. Analysis of modern approaches for the prediction of electric energy consumption. Open Engineering. 2020 Jan 1;10(1):350-61.
205. Gourieroux C, Monfort A, Christian G. Time series and dynamic models. Cambridge University Press; 1997.
206. Cools M, Moons E, Wets G. Investigating the variability in daily traffic counts through use of ARIMAX and SARIMAX models: assessing the effect of holidays on two site locations. Transportation research record. 2009;2136(1):57-66.
207. Singh AP, Gaur MK, Agrawal S, Kasdekar DK. Time Series Model Forecasting of Boot Using Holt, Winter and Decomposition Method. Journal of Industrial Safety Engineering. 2015;2(2):23-31.

208. Han P, Wang PX, Zhang SY. Drought forecasting based on the remote sensing data using ARIMA models. Mathematical and computer modelling. 2010 Jun 1;51(11-12):1398-403.
209. Zhang G, Patuwo BE, Hu MY. Forecasting with artificial neural networks:: The state of the art. International journal of forecasting. 1998 Mar 1;14(1):35-62.
210. Elsaraiti M, Merabet A. Solar power forecasting using deep learning techniques. IEEE Access. 2022 Mar 17;10:31692-8.
211. Sharma S, Sharma S, Athaiya A. Activation functions in neural networks. Towards Data Sci. 2017 Sep;6(12):310-6.
212. Muralikrishna A, Vieira LE, dos Santos RD, Almeida AP. Total Solar Irradiance Forecasting with Keras Recurrent Neural Networks. InInternational Conference on Computational Science and Its Applications 2020 Jul 1 (pp. 255-269). Springer, Cham.
213. Olah, C. 2015. http://colah.github.io/posts/2015-08-Understanding-LSTMs/, Erişim Tarihi: 16.02.2021.
214. Sako K, Mpinda BN, Rodrigues PC. Neural Networks for Financial Time Series Forecasting. Entropy. 2022 May 7;24(5):657.
215. Erichson NB, Azencot O, Queiruga A, Hodgkinson L, Mahoney MW. Lipschitz recurrent neural networks. arXiv preprint arXiv:2006.12070. 2020 Jun 22.
216. Malinović NS, Predić BB, Roganović M. Multilayer long short-term memory (LSTM) neural networks in time series analysis. In2020 55th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST) 2020 Sep 10 (pp. 11-14). IEEE.
217. Hochreiter, S., Schmidhuber, J. 1997. Long short-term memory. Neural computation, 9(8): 1735-1780.
218. Yu, R., Gao, J., Yu, M., Lu, W., Xu, T., Zhao, M., Zhang, Z. 2019. LSTM-EFG for wind power forecasting based on sequential correlation features. Future Generation Computer Systems, 93: 33-42.
219. Gers FA, Schraudolph NN, Schmidhuber J. Learning precise timing with LSTM recurrent networks. Journal of machine learning research. 2002;3(Aug):115-43.
220. Vakili, M., Sabbagh-Yazdi, S.R., Kalhor, K, et al. Using artificial neural networks for prediction of global solar radiation in Tehran considering particulate matter air pollution. Energy Procedia 2015, 74: 1205-1212.
221. Fan, C., Xiao, F, Zhao, Y. A short-term building cooling load prediction method using deep learning algorithms. Applied energy 2017, 195: 222-233.
222. Mochalov VP, Bratchenko NY, Yakovlev SV. Analytical model of object request broker based on Corba standard. InJournal of Physics: Conference Series 2018 May 1 (Vol. 1015, No. 2, p. 022012). IOP Publishing.