Forecasting wavelet denoised global horizontal irradiance using attention-based long short term memory network: A case study of South Africa

---

By

Ndamulelo Innocent Nelwamondo
14011206

Supervisor(s):
Dr. Caston Sigauke
Dr. Alphonce Bere
Ms. Maduvhahafani Thanyani

submitted in partial fulfilment of the requirements for the degree of Master of Science Degree in e-Science

in the

Department of Mathematical and Computational Sciences
Faculty of Science, Engineering and Agriculture
University of Venda
Thohoyandou, Limpopo
South Africa

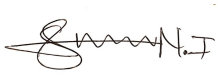
Abstract

Microgrids are becoming a crucial component of the electricity grid in dependability, economics, and environmental sustainability. Microgrids rely heavily on renewable energy sources. From an engineering standpoint, anticipating short-term solar generation is a critical challenge in microgrid planning and design. Anticipating solar power is heavily reliant on forecasting sun radiation. Short-term solar radiation forecasting may also be used to estimate the energy potentials of photovoltaic (PV) panels impacted by degradation rates. A comparison of multiple models, namely the Autoregressive Integrated Moving Average (ARIMA), Long Short Term Memory (LSTM), Attention-based LSTM and a hybrid Attention-based LSTM-ARIMA for forecasting 5-day ahead 1-minute solar radiation is performed in this study. The best model for forecasting Global Solar Radiation(GHI) from Richtersveld station is ARIMA with $MAE = 0.782$ and $RMSE = 1.271$, followed by hybrid model with $MAE = 4.120$ and $RMSE = 4.987$. For Stellenbosch University station, attention LSTM was the best with $MAE = 1.512$ and $RMSE = 1.640$, followed by hybrid with $MAE = 2.011$ and $RMSE = 2.511$. The hybrid attention-based LSTM-ARIMA model on the USAid Venda station was the best fitting model with $RMSE = 7.383$ and $MAE = 14.1293$, followed by LSTM with $MAE = 7.817$ and $RMSE = 8.444$. Comparing the results on non-wavelet denoised and wavelet denoised, models performed better on wavelet denoised data. ARIMA model was the best with $MAE = 0.194$ and $RMSE = 0.542$, followed by hybrid with $MAE = 2.176$ and $RMSE = 2.308$.

Keywords: Deep learning Algorithms; Global Horizontal Irradiance; Hybrid ARIMA-LSTM model; Solar Energy; Wavelet denoising.

Declaration

I, Ndamulelo Innocent Nelwamondo [student Number: 14011206], hereby declare that the mini-dissertation titled: "Forecasting wavelet denoised global horizontal irradiance using attention-based long short term memory network" for the Master of Science degree in e-Science at the University of Venda, hereby submitted by me, has not been submitted for any degree at this or any other university that it is my work in design and in execution, and that all reference material contained therein has been duly acknowledged.



Ndamulelo Innocent Nelwamondo
9 June 2022

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor Dr. Caston Sigauke for the dedication to my MSc in the field of the e-Science research project. With your guidance, I am sure I will go far with research. Dr. A. Bere and Ms. M. Tanyani, my co-supervisors, deserve special thanks for their patience and support.

My sincere thanks also go to my sponsor, The National e-Science Postgraduate Teaching and Training Platform(NEPTTP). They provided me an opportunity to join MSc in the e-Science programme and financial assistance. Without the precious support, I would not be conducting this project.

Contents

List of Figures

List of Tables

List of Abbreviations

AI	**A**rtificial **I**ntelligent
ANN	**A**rtificial **N**eural **N**etwork
ARIMA	**A**utoregressive **I**ntergrated **M**oving **A**verage
CNN	**C**onvolutional **N**eural **N**etwork
CWT	**C**ontinuous **W**avelet **T**ransform
DWT	**D**iscrete **W**avelet **T**ransform
EDA	**E**xplanatory **D**ata **A**nalysis
EMD	**E**mperical **M**ode **D**ecomposition
GHI	**G**lobal **H**orizontal **I**rradience
IMF	**I**ntermolecular **F**orce
kNN	**K**-nearest **N**eural **N**etwork
LSTM	**L**ong **S**hort **T**erm **M**emory
MAE	**M**ean **A**bsolute **E**rror
ML	**M**achine **L**earning
MLP	**M**ultilayer **P**ercetron **N**etwork
MODWT	**M**aximum **O**verlap **D**iscrete **W**avelet **T**ransform
NN	**N**eural **N**etwork
RF	**R**andom **F**orest
RMSE	**R**oot **M**ean **S**quare **E**rror
RNN	**R**ecurrent **N**eural **N**etwork
STFT	**S**hort **T**ime **F**ourier **T**ransform
SUN	**S**tellenbosch **U**niversity
SVM	**S**upport **V**ector **M**achine
RVD	**R**ichterveld
UNV	**U**Said **V**enda
WT	**W**avelet **T**ransform
PV	**P**hotovoltaic

Chapter 1

Introduction

1.1 Background

Because of increased understanding of environmental emissions and depleting fossil fuel reserves, renewable energy supplies have grown in value in the twenty-first century (Mutavhatsindi, Sigauke, and Mbuvha, 2020). Researchers are working hard to create a pollution-free world by proposing carbon-free solutions in various ways, including automobiles, clothing, home appliances, and other energy-consuming industries. Introducing the sun's radiation into the electrical system is crucial but complex. Forecasting errors will destabilise the supply-demand equilibrium and incur unnecessary costs. Consequently, reliably and efficiently predicting global horizontal irradiance (GHI) is a vital function of a photovoltaic installation.

1.2 Problem Statement

Under clear-sky conditions, solar radiation offers knowledge about the highest possible magnitude of the solar resource available at a place of interest. This global horizontal irradiance (GHI) time-series data can determine the limits of solar energy use in applications such as thermal and electrical energy generation. Thus, forecasting GHI can play a vital role in electrical power systems. If people use GHI estimation, which is wrong, there will be a strong impact on designing an electric power grid. Researchers and academicians in different fields may find this study essential for future research on GHI or solar energy.

1.3 Research Questions

The study seeks to answer two questions, which are:

(1) Between traditional LSTM, ARIMA, attention-based LSTM and hybrid attention-based LSTM-ARIMA models, which model is best suited for modelling and forecasting the GHI wavelet denoised time series data?

(2) Between the original GHI dataset and wavelet denoised GHI dataset, when do the models perform better?

1.4 Research Aim and Objectives

1.4.1 Research Aim

This research aims to develop a model for short-term forecasting of GHI.

To achieve this, several models (namely ARIMA, LSTM, Attention-based LSTM, and Attention-based LSTM-ARIMA) will be trained using the original GHI dataset and wavelet denoised GHI historical data to capture the relationships, trends and seasonality of GHI time series data.

1.4.2 Objectives

The objectives of the research are to:

- *remove* the noise from the GHI time series data using the wavelet transform,
- *predict* the GHI time sequence data using the LSTM,
- *predict* the GHI time sequence data using the ARIMA model,
- *predict* the GHI time sequence data using the attention-based LSTM-ARIMA model,
- *compare* the models.

1.5 Significance of the study

South Africa has seen 40% in load shedding and data shows it is still headed for a record year of power cuts. Thus, Solar forecasting allows grid operators to forecast and balance energy output and demand. Assuming the grid operator has a mix of producing assets at

their disposal, effective solar forecasting enables the operator to optimise how their controllable units are dispatched. Integration of solar energy in electricity generation will save South Africa from power cuts and emissions reduction since Eskom uses coal for electricity generation.

1.6 Definitions

1.6.1 Definitions

Terminology	**Definition**
Time series data	A collection of quantities in their order of occurrence in time.
GHI	The sum of the diffuse and direct solar radiation.
Wavelet	Small wave that goes up, and back down to zero.
Neural Network	Is a collection of algorithms that tries to detect underlying correlations in a batch of data using a technique that mimics how the human brain works.
Hybrid model	Combination of two or more models.
Attention-based model	Is a directed focus model which pays greater attention to certain factors when processing data.

1.7 Overview

The rest of the study is divided into four chapters. Chapter 2 discusses relevant topics in the field of study. Chapter 3 overviews how the data will be retrieved, cleaned, preprocessed and analysed. The chapter discusses the methods to be used in modelling data (Section 3.3) and measurement of their errors (Section 3.4.3). Chapter 4 gives the results of all the models and models comparison. Chapter 5 provides conclusions from the study and areas for future studies.

Chapter 2

Literature Review

2.1 Global Horizontal Irradiance

Renewable energy supplies are becoming increasingly valuable as environmental consciousness and energy demand rise (Renno, Petito, and Gatto, 2015), and will thus be significantly utilised in future power networks. Wind power plants and solar power plants, for example, must be interconnected and regulated by power system operators. Because variable renewable energy is non-dispatchable, it can influence system voltage and frequency, decrease load predictability, and cause reverse power flow (West et al., 2014). Solar irradiance forecasting is thus one of the important methods for integrating solar energy into electric power systems safely and reliably.

Solar irradiance displays time-dependent stochastic and fluctuant behaviour due to seasonal deviation and weather shifts. As solar power is introduced into an electric power grid, the erratic and irregular nature of solar irradiance triggers many serious issues, including voltage instability and low power quality (Alonso-Montesinos, Batlles, and Portillo, 2015). As a result, to ensure the future power grid's security and safety, it is important to keep track of solar irradiance (Urraca et al., 2016). However, several developing countries have inadequate or incomplete solar irradiance data due to the prohibitively high cost of visual equipment and maintenance (He and Yao, 2016).

2.2 Denoising methods

2.2.1 Introduction

In general, most time-series signals in actuality are noisy. Before doing a formal analysis, valid signals must be extracted. Linear filtering methods that have been used in the past will generate extra distortion since complex time series contain a high degree of nonlinearity.

Many novel denoising approaches have been presented in recent years to handle this type of noise removal problem (Wu, Jia, and Liu, 2021).

2.2.2 Wavelet denoising

The concept of wavelet threshold denoising is to use a wavelet transform to pass the signal. Furthermore, the noise's wavelet coefficient is less than the signals (Yang et al., 2021). Then, choose a suitable threshold and keep the wavelet coefficient bigger than it. Because the smaller one is called noise, it will be set to zero. Finally, use the inverse wavelet transform to produce denoised data.

2.2.3 Empirical Mode Decomposition (EMD)

EMD is an innovative method for dealing with non-stationary signals (Huang et al., 1998). Furthermore, linear and stationary signal analysis better captures the physical meaning of the signal than other time-frequency analysis approaches.

EMD aims to break down the signal based on time scale characteristics without establishing any fundamental functions. EMD does, however, have several flaws. According to the energy law, it immediately discards numerous small-order Intermolecular forces (IMFs). When filtering the noise, it will lose some essential signal information, especially if the signal comprises multiple sharp signals or unique points.

2.2.4 Singular Spectrum Analysis

The main principle of singular spectrum analysis is to perform singular value decomposition on the time series' trajectory matrix and partition the singular values into two big and small groups. The lower value group is thought to be in the noise section. The group with the higher value is called the normal part. After noise reduction, the normal group is rebuilt to create the time series. The major challenge in noise reduction based on single spectrum analysis is determining the border between noise components and usable components (Ospina, Newaz, and Faruque, 2019).

2.2.5 Conclusion

In summary, it is impossible to find time series data with no noise. For the prediction model to perform better or give accurate predictions, denoising is a crucial step in data preprocessing and must be done thoroughly. This study uses wavelet transform to denoise data as wavelet transform can work effectively on non-stationary time series. Wavelet analysis smoothes and filters the sun's speed, direction, and power while retaining the primary

trend of the time series. As a result, the wavelet is extremely beneficial for dealing with highly erratic GHI time series.

2.3 Deep Learning Algorithms

Deep learning methods are suitable for predictive modelling and capturing of long and short-term properties of non-linear time series and have been effectively utilised in the field of new energy (Wu, Jia, and Liu, 2021). For example, literature (Hochreiter and Schmidhuber, 1997) suggested LSTM, whose threshold structure overcomes the gradient problem of RNNs and has had an impact on time series prediction. For solar irradiance prediction, literature (Wang et al., 2018) presents a model based on wavelet transform, CNN, and LSTM.

Machine learning-based forecasting approaches have also been widely applied in various fields in recent years (Wang, Li, Liu, Mi, Shafie-Khah, and Catalão, 2018b). Non-linear regression models, such as the Artificial Neural Network (ANNs), the Support Vector Machine (SVM) (Jianwu and Wei, 2013), and the Markov chain, have been commonly applied in the field of solar forecasting. These non-linear regression models are often widely used in conjunction with classification models (Wang, Zhen, Wang, and Mi, 2018a).

Chih-Chiang (2017) used four forecasting models, namely, the multilayer perceptron (MLP), random forests (RF), k-nearest neighbours (kNN), and linear regression (LR), to model surface solar irradiance for Tainan City in Taiwan. MLP performed better than the others and LR was the worst-performing.

There are many divisions of the deep learning scheme, including LSTM, Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN), among others. Despite the excellent performance of deep learning algorithms, few studies have used deep learning approaches in day-ahead solar irradiance forecasting.

Mutavhatsindi, Sigauke, and Mbuvha (2020) in their paper applied long short-term memory (LSTM) networks, support vector regression (SVR) and feed-forward neural networks (FFNN) models on South African hourly solar irradiance data. Their findings show that the FFNN model was the best, with lower mean absolute error (MAE) and root mean square error (RMSE).

Chapter 3

Research Methodology

3.1 Introduction

The path researchers must take to perform their study is referred to as research methodology. It demonstrates how these researchers construct their problem and objectives and how they present their findings based on the data collected during the study period.

3.1.1 Wavelet Denoising

Wavelet is a type of localised function in the time and frequency domain that decomposes time series into extra elementary parts that hold time series-related information. For this purpose, wavelets transform will be used for noise removal in the GHI time series data. Data preprocessing will provide further details on the wavelet denoising process.

3.1.2 Deep Learning Algorithms

Artificial Intelligence (AI) for signal prediction entails utilising classic machine learning (ML) methods as well as deep learning models such as neural networks (Lu, Wang, and Xu, 2018). To forecast output values, ML algorithms receive and analyse input data. While being fed with more data, they enhance their performance; the higher the sample size, the better performance (Lu, Wang, and Xu, 2018). Neural networks (NN) are computing systems inspired by the human brain and are good at pattern recognition (Lu, Wang, and Xu, 2018).

3.1.3 ARIMA Model

To forecast the low volatility time series, an AutoRegressive Integrated Moving Average (ARIMA) model is utilised. The model is fit for ARIMA parameters p, d, and q. The differencing d is determined via Phillips Perron, Augmented Dickey-Fuller or Kwiatkowski Phillips Schmidt Shin.

3.1.4 Models Averaging

Artificial intelligence may be used in collaboration with statistical approaches and procedures. While statistics can cope with large volumes of data, AI is better at capturing relationships between data points. (Lu, Wang, and Xu, 2018). The predictions are averaged to generate the predicted 1-minute GHI time series data.

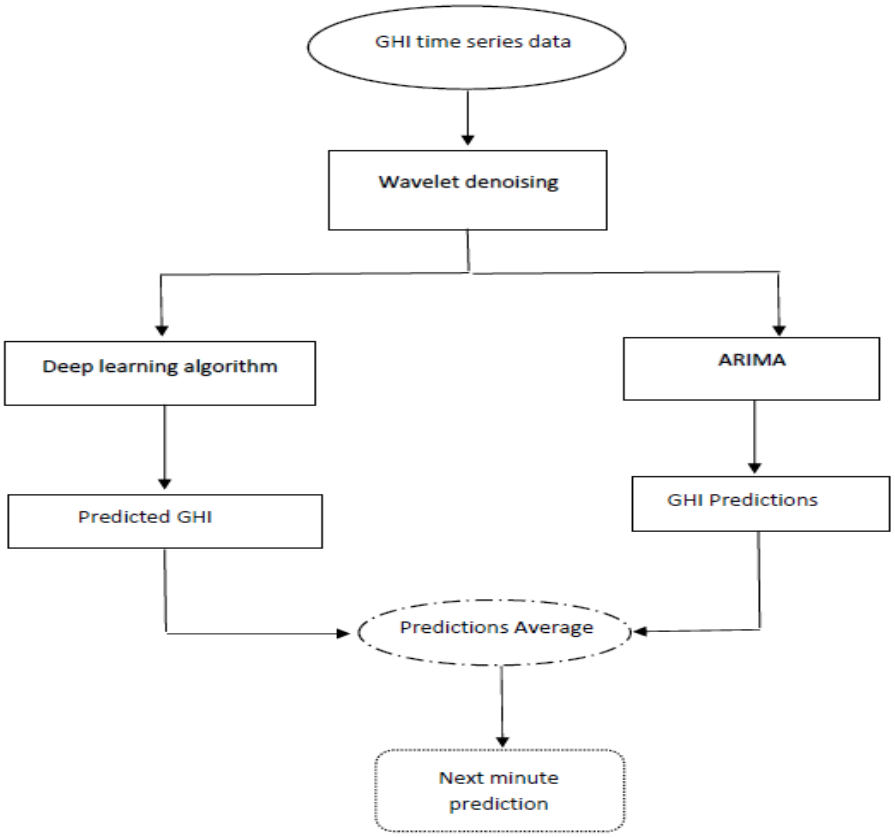


FIGURE 3.1: Research design (Attention-based LSTM-ARIMA hybrid model on wavelet denoised series).

Figure 3.1 shows the flow diagram of how the research will be conducted (i.e., the research design).

3.2 Data

The study uses secondary minute-averaged data from SAURAN website https://sauran.ac.za/ for three sites in South Africa, namely, UNV: USAid Venda, RVD: GIZ Richtersveld,

and SUN: Stellenbosch University. The paper will cover 1-minute time series data for five days from 23 March 2021 to 28 March 2021 for two stations (SUN and UNV), and 1-minute data for five days from 30 April 2020 to 4 May 2020 for the RVD station. The main reason for the selected periods was that they covered the most recent time and dates the stations last recorded the data. Table 3.1 gives the first two rows of the data for three sites' GHIs:

TABLE 3.1: Head of GHI historical data.

TmStamp	SUN	UNV	TmStamp	RVD
23/03/2021 00:00:00	0	0	30/04/2020 00:00:00	0
23/03/2021 00:01:00	0	0	30/04/2020 00:01:00	0

3.2.1 Data Preprocessing

There are many different methods one can use for time series data preprocessing. This research will focus on wavelet transformation due to their properties, which will be mentioned throughout this section.

Wavelet Transform

GHI data is monitored and evaluated for its behaviour using time signals. The first and most popular method for this is the Fourier transform, developed in 1807 by a French mathematician and physicist, Joseph Fourier (Fourier, 1807). The substitute method with appealing properties is the wavelet transform, which was first mentioned in a thesis by Alfred Haar in 1909 (Haar, 1909). A better knowledge of Fourier transform accompanies a good understanding of wavelet transform. Wavelet analysis may be done in numerous methods, including continuous wavelet transforms, discretised wavelet transforms, and genuine discrete wavelet transformations.

The Fourier transform only retrieves the global frequency content of a signal. Therefore, the Fourier transform is only useful for stationary and pseudo-stationary signals (Merry and Steinbuch, 2005). The Fourier transform does not give satisfactory results for signals that are highly non-stationary, noisy, and periodic (Sifuzzaman, Islam, and Ali, 2009).

The short-time Fourier transform (STFT) addresses the Fourier transform's restriction. A signal's frequency and temporal information can be extracted using the STFT. The STFT computes the Fourier transform of a windowed portion of the original signal that shifts along the time axis (Daubechies, 1990).

The selected window has a significant impact on the performance of the STFT analysis (Rioul and Vetterli, 1991). Although a small window provides adequate time resolution, distinct frequencies are not well identified. This resolution is not satisfactory (Sifuzzaman, Islam, and Ali, 2009). A high-frequency resolution is essential because low-frequency components frequently endure a long time. High-frequency components often arise in brief bursts, necessitating a higher temporal resolution (Merry and Steinbuch, 2005).

In comparison to the Fourier transform, the wavelet transform's analysing function may be selected with greater freedom, as it does not need the use of sine-forms. A wavelet function is a tiny wave that must be oscillatory in some form to differentiate across frequencies (Schneiders, van de Molengraft, and Steinbuch, 2001). The wavelet involves both the structure and the window study. Several types of wavelet functions are designed for the Continuous Wavelet Transform (CWT), each with unique features (Schneiders, van de Molengraft, and Steinbuch, 2001).

There are two types of wavelet transform, namely Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform(DWT). A brief description of the Discrete wavelet transform is given below, as it is the focus of the project:

Discrete Wavelet Transform

To perform wavelet analysis, the discrete wavelet transform (DWT) employs filter banks (Bullmore, 2002). The discrete wavelet transform decomposes the signal into wavelet coefficients, which may then be used to recreate the original signal (Merry and Steinbuch, 2005). The signal is represented by the wavelet coefficients in various frequency bands. The coefficients may be treated in a variety of ways, making the DWT more appealing than linear filtering (Percival and Walden, 2000).

To enhance the performance of the proposed models, version 4.0 of the Discrete wavelet transform (DWT), called the maximal overlap discrete wavelet transform (MODWT), is proposed in data preprocessing. The DWT and the MODWT draw on multi-resolution analysis to decompose a time series into lower and lower wavelet scales. In terms of multi-resolution research, the wavelet transform decomposes a time series into weighted moving average values ("smooths") and the information required to reconstruct the signal ("details") from the averages (Percival and Walden, 2000).

Wavelet Denoising Process

Wavelet denoising consists of three steps which are visible in Figure 3.2 and explained below:

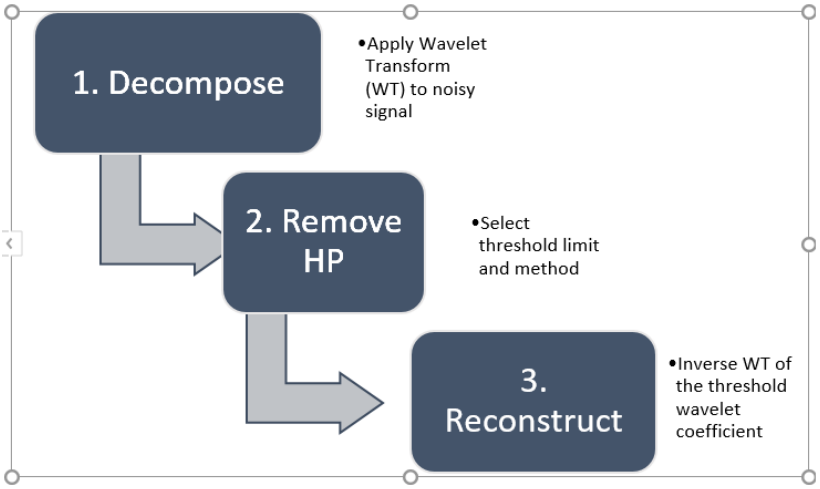


FIGURE 3.2: Wavelet denoising process diagram.

1. Apply wavelet transform to the noisy signal to produce the noisy wavelet coefficient to the level where GHI signal occurrence can be properly distinguished.
2. Select the appropriate threshold limit at each level and threshold method.

   Thresholding methods

   Soft thresholding - provides smoother results than hard thresholding.

   Hard thresholding - provides better peak preservation than the soft one.
3. Inverse wavelet transform of the thresholded wavelet coefficient to obtain a denoised signal

The denoised signal will then be partitioned to a training set of 75% and a test set of 25%. The description models for training and testing are given in Section 3.3.

3.3 Methods

The research methodology is employed to achieve the study's objectives. That will be focused on in this chapter. Furthermore, a detailed empirical framework of the proposed traditional and deep learning methods for the researcher to achieve the objectives will be

provided. The study suggests adopting and discussing ARIMA(p,d,q) and deep learning algorithmic frameworks to model and estimate GHI.

This is so due to the extensive use of ARIMA(p,d,q) models in related studies and that non-linear modelling techniques have gained much attention by academic researchers and electricity generation contributors like governments, corporations, enterprises, and dealers, lately, with neural networks assuming a prominent role. Neural Networks (NN) applications have shown an increase in power engineering for **live** systems and the results are promising. To support these claims, various performance figures are being quoted. Still, the absence of explicit models, due to the non-parametric nature of the approach, makes it difficult to assess the significance of the model estimated and the possibility that any short-term success is due to data mining (Achilleas and Apostolos-Paul, 1999).

3.3.1 Recurrent Neural Networks

RNNs are neural networks that are good at modelling sequence data. Recurrent neural networks are used in speech recognition, language translation, and stock predictions (*Illustrated Guide to Recurrent Neural Networks*). Recurrent neural networks allow information to persist as they are networks or loops. Figure 3.3 shows, the hidden network h_t, some input x_t and output y_t. A loop allows information to be passed from one step to the next.

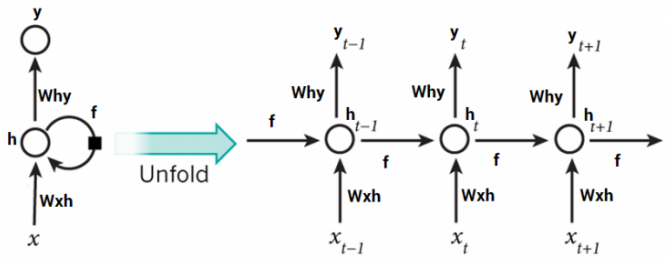


FIGURE 3.3: RNN model. Adapted from (*Illustrated Guide to Recurrent Neural Networks*)

$$h_t = f(W_{hh}h_{t-1} + W_{xh}x_t + b) \tag{3.1}$$

$$y_t = f(W_{hy}h_t + c), \tag{3.2}$$

where the equation of the hidden state h_t and output y_t is given by Equation 3.1 and 3.2 respectively. For which a non-linear function like $tanh$ or $ReLU$ are indicated by f, and $t = 1, ..., n$, where n is the sample size.

3.3.2 Long Short Term Memory

LSTM is a kind of RNN that was first introduced in 1997 (Hochreiter and Schmidhuber, 1997). In the case of LSTM architecture, LSTM cells replace the normally concealed layers. The cells are made up of several gates that may regulate input flow. Figure 3.4 shows the structure of an LSTM network, which can be formulated as follows:

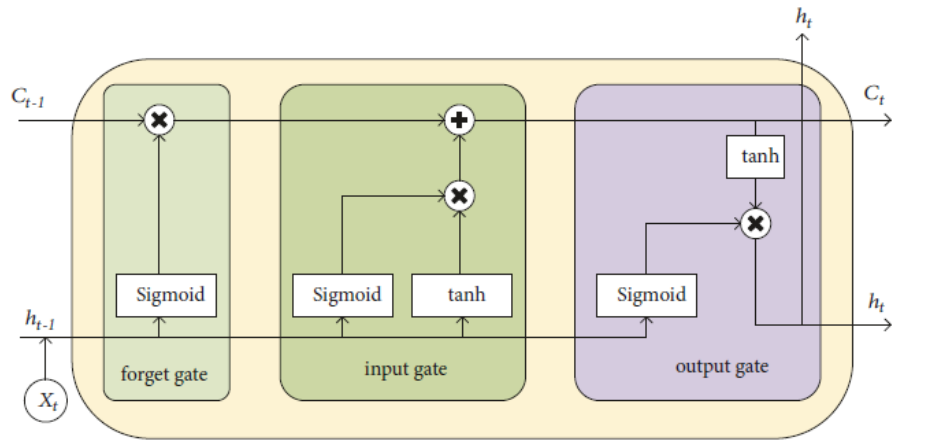


FIGURE 3.4: LSTM model. Adapted from (Yan, 2016).

$$f_t = \sigma(W_{t.}[x_t, h_{t-1} + b_f]) \tag{3.3}$$
$$i_t = \sigma(W_{i.}[x_t, h_{t-1} + b_i]) \tag{3.4}$$
$$\tilde{C}_t = tanh(W_{c.}[x_t, h_{t-1} + b_c]) \tag{3.5}$$
$$C_t = f_t * c_{t-1} + i_t * \tilde{C}_t \tag{3.6}$$
$$o_t = \sigma(W_{o.}[x_t, h_{t-1} + b_o]) \tag{3.7}$$
$$h_t = O_t * tanhC_t, \tag{3.8}$$

where x_t is the input value at each time t, h_t and h_{t-1} are hidden states of the LSTM, and C_t is the memory state. $sigmoid(\sigma)$ and $tanh$ are two types of the activation functions for three types of gating units: the input gate i_t, forget gate f_t, and output gate o_t (Selvin et al., 2017). W and b denote weight matrices and bias vectors, respectively.

3.3.3 Attention Based LSTM

By assigning enough attention to vital information, the attention mechanism highlights important local information (Qiu, Wang, and Zhou, 2020). Attention is a network design component that manages and quantifies interdependence: (1) between the input and output ports, and (2) within the input elements (*Attention mechanism*). The soft attention mechanism can be formulated as:

$$e_t = tanh(W_d[x_t, h_{t-1} + b_d]) \tag{3.9}$$

$$a_t = \frac{exp(e_t)}{\sum(exp(e_t))}, \tag{3.10}$$

where W_d is the weight matrix of the attention mechanism, indicating information that should be emphasised; e_t is the result of the first weighting calculation; d_d is the deviation of the attention mechanism; x_t is the input of the attention mechanism, which is now, the output of the LSTM hidden layer; and a_t is the final weight obtained by x_t.

3.3.4 Autoregressive Integrated Moving Average

The Autoregressive Integrated Moving Average (ARIMA) method depends on historical values in the series for forecasting. It was invented by Box and Jenkins (Box et al., 2015) and is a widely used forecasting model (Jang and Lee, 2018). The ARIMA model consists of the AutoRegressive terms (AR) and the Moving Average terms (MA).

Applying the lag operator denoted L, Autoregressive AR terms are lagged values of the dependent variable and refer to it as lag order p as the number of time lags. A non-seasonal AR(p) can be formulated as follow:

$$AR(p) : \phi(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p \tag{3.11}$$

Moving Average MA terms are lagged forecast errors in the predictions between actual past values and their predicted values and refer to it as the order of moving average q. A non-seasonal MA(q) can be formulated as follows:

$$MA(q) : \theta(L) = 1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q \tag{3.12}$$

The ARIMA model can be formulated as follows:

$$ARIMA(p, d, q) : \phi(L)(1 - L)^d r_t = \theta(L)\epsilon_t, \tag{3.13}$$

where L is the backward operator, r_t is the GHI time series for t = 1,..., n, ε_t is the white noise, and the d is the number of times that the observations are differenced.

3.3.5 Hybrid ARIMA-LSTM with Attention

Artificial intelligence and statistical methodologies and techniques can be integrated. While statistics make it possible to cope with large volumes of data, AI efficiently captures interconnections between data points. The hybrid model takes on the averages of all three models, i.e.:

$$Hybrid\_prediction = \frac{\sum_1^k (model_k\_prediction_i)}{number\_of\_models}, \tag{3.14}$$

where k is the number of models for the i's term.

3.4 Analysis

3.4.1 Wavelet Analysis

Wavelet decompositions will be performed using MODWT, wavelet filter of haar wavelet family, hard thresholding, and wavelet length of 2. Previously, both the discrete wavelet transform (DWT) and the maximum overlap discrete wavelet transform (MODWT) approaches were utilised to generate functional connectivity matrices in the literature (Vértes et al., 2012). Therefore, the L scale MODWT wavelet coefficient process is defined by Equation 3.15:

$$\overline{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l}, \quad t = 1, ..., n, \tag{3.15}$$

where the MODWT wavelet filter $\{\tilde{h}_{j,l}\}$ is based on a Haar wavelet filter of length L.

3.4.2 Model Specifications

ARIMA

The best fit for the ARIMA model was determined by the function autoarima() in python, which automatically finds the best ARIMA(p,d,q) parameters for the series, the Table 3.2 shows the ARIMA specification for the GHI's of all three sites.

TABLE 3.2: ARIMA Specifications

Site	**ARIMA(p,d,q)**
RVD	ARIMA(5,1,0)
SUN	ARIMA(5,1,0)
UNV	ARIMA(5,1,0)
UNV denoised	ARIMA(5,1,0)

LSTM and Attention Based LSTM

TABLE 3.3: LSTM specifications

Variables	**LSTM**	Attention-based LSTM
Units	100	25
Activation	—	relu
Dense	1	1
Dropout	0.15	—
Attention activation	—	sigmoid
Optimizer	adam	adam
Loss	mse	mse
Epoch	25	100
Batch size	70	70

Table 3.3 shows the model specifications for the LSTM model and attention-based LSTM. The same specifications were the same for all the four datasets of 3 different sites.

Hybrid Attention Based LSTM-ARIMA

The hybrid Attention-based LSTM-ARIMA comprises the aggregate of ARIMA, LSTM, and Attention-based LSTM.

3.4.3 Model Performance Metrics

The prediction results will be evaluated using the measures of errors, namely the root mean square error (RMSE) and mean absolute error (MAE), to find the best model, which will be used for predictions (Qiu, Wang, and Zhou, 2020). The smaller the RMSE and MAE, the closer the predicted return to the true return, and the better the model's fit. The formulation of the two performance metrics is as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}, \tag{3.16}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|(y_i - \hat{y}_i)\right|, \tag{3.17}$$

where N is the number of 1-minute GHI time series data, y_i is the actual GHI and $\hat{y}_i$ is the predicted GHI.

Chapter 4

Results and Discussions

4.1 Introduction

The chapter presents and analyses data obtained from https://sauran.ac.za/. The results on a 1-minute GHI signal will be provided. The analysis considers the objectives outlined in chapter 1 and the methodology discussed in Chapter 3. Matlab R2020a was used for wavelet denoising and Python 3.7 was used for analysing the non-wavelet and wavelet denoised GHI signal.

4.2 Data Source

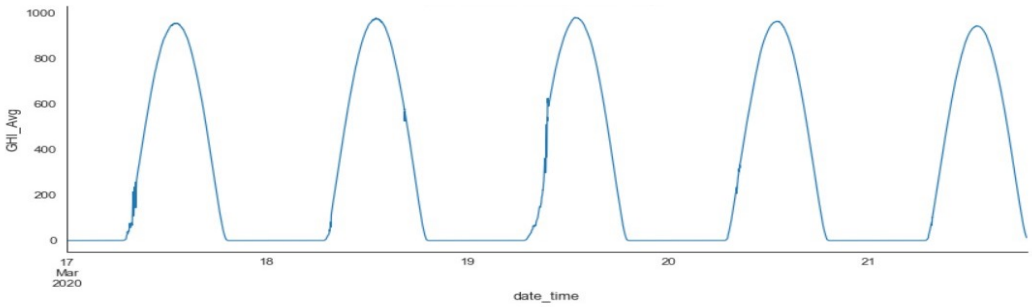
The study uses secondary 1-minute GHI signal data from https://sauran.ac.za/ over five days for three different sites.

4.3 Explanatory Data Analysis

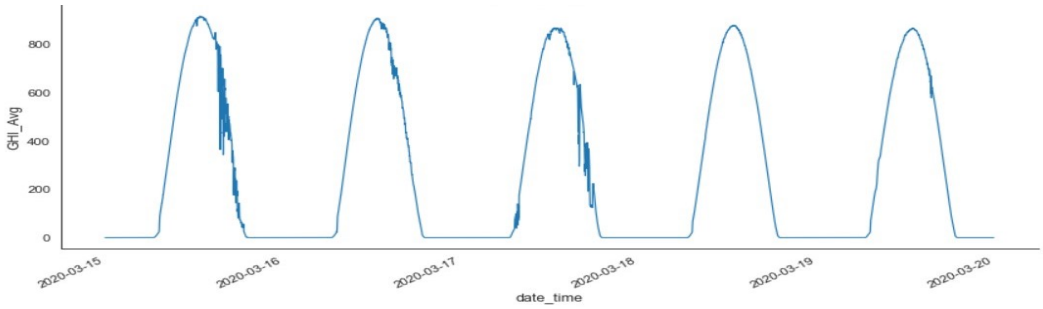
Explanatory data analysis (EDA) is an approach to analysing data sets and summarising their main characteristics or a critical step in analysing the data from an experiment (Tukey, 1977). Exploratory data analysis covers a preliminary selection of appropriate models, examinations of assumptions, and assisting statisticians in data exploration.

4.3.1 Time series plots

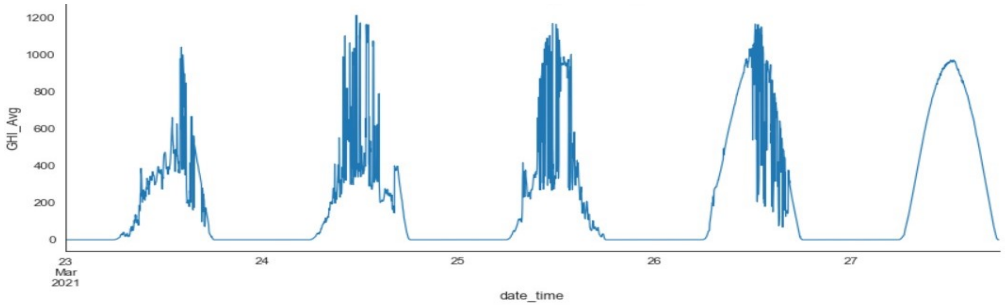
Time series charts study data patterns and behaviour across time and are frequently used to investigate a process modification's daily, weekly, and seasonal consequences. They provide a visual representation of the time series. Time-series graphs make it simple to examine data patterns.



(A) GHI for RVD station over time.



(B) GHI for SUN station over time.



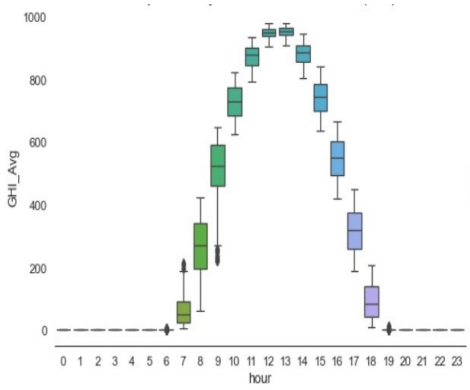
(C) GHI for UNV station over time.

FIGURE 4.1: Time series plots of GHI vs time for 3 sites.

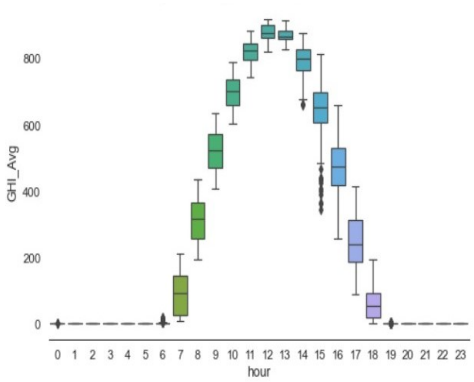
Figure 4.1 (A) gives the GHI time series plot for RVD: GIZ Richtersveld station, Figure 4.1 (B) gives the GHI time series plot for the SUN: Stellenbosch University station and Figure 4.1 (C) gives the GHI time series plot for the VEN: USAid Venda station, with a display of volatility clustering and shows that volatility occurs in bursts.

4.3.2 Box plots

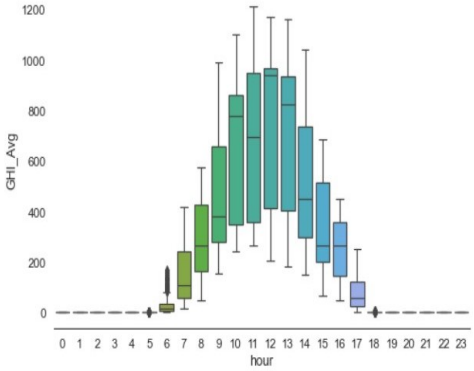
Box charts show data in batches. Conventionally, five values from a collection of data are used: the extremes(minimum and maximum), the upper(Q3) and lower(Q1) quartiles, and the median(Q2). These charts are becoming popular for exploratory data analysis and visual summaries for statisticians and non-statisticians.



(A) Hourly GHI for RVD.



(B) Hourly GHI for SUN.



(C) Hourly GHI for UNV.

FIGURE 4.2: Box plots of hourly GHIs for the 3 sites.

Figure 4.2 (A) gives the box plot of GHI for RVD: GIZ Richtersveld station, Figure 4.2 (B) gives the box plot of GHI for the SUN: Stellenbosch University station and Figure 4.2 (C) gives the box plot of GHI for the VEN: USAid Venda station. When comparing the boxplots for each hour and each site, it can be noticed that the median GHI is much higher at 12h00 than in other hours. This implies that 12h00 contains the peak GHI. From hour 00h00 to 05h00 am and hour 18h00 to 23h00 pm, in these hours, the stations recorded the lowest GHI of value 0. GHI increases from 06h00 am approaching 12h00 pm and decrease after

that. From Figures 4.1(A) and (B), it can be seen that the five-number summary(minimum, Q1, Q2, Q3, and maximum) are close to one another, whereas in Figure 4.2 (C) they are far apart This implies that the 1-minute time series GHI data from the UNV station is too much volatile, thus most models in volatile data tend to perform the worst.

Summary of descriptive statistics

Table 4.1 reports the key descriptive statistics of the GHI series. Descriptive statistics quantitatively describe or summarise features of the collected information. Descriptive statistics provides sample summaries of the sample. The sample summaries are the minimum, mean, maximum, standard deviations, skewness, kurtosis and sample size.

TABLE 4.1: 1 minute GHI data for 3 stations descriptive statistics.

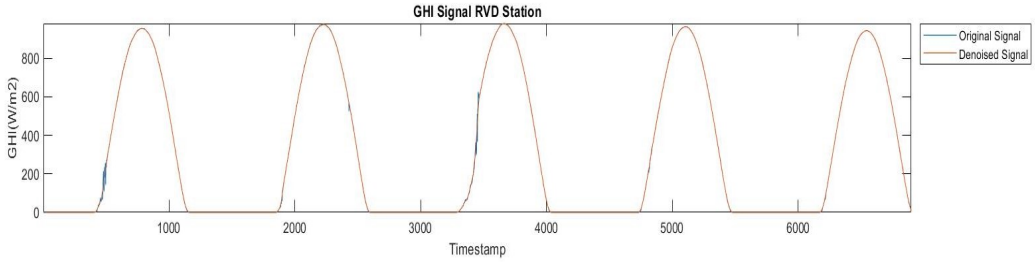
Variables	UNV	SUN	RVD
Min	0	0	0
Mean	213.4625	267.6134	301.3076
Std Dev	306.2724	337.1492	369.2255
Kurtosis	0.8023	-1.1099	-1.2020
Skewness	1.3976	0.7589	0.7025
Max	1211.8950	917.9128	980.0505
Range	1211.8950	917.9128	980.0505
n	6849	7189	6897

The kurtosis is a statistical measure used to describe the distribution and can also help explain the distribution's tails concerning the overall shape. Kurtosis has three categories which can be displayed by a data set, i.e. mesokurtic (kurt=3), leptokurtic (kurt>3), and platykurtic (kurt<3). The GHI signals have kurtosis values less than 3 for all three sites. Thus, they are platykurtic, implying that both the datasets have lighter tails than the normal distribution. All the datasets have positive skewness; RVD and SUN stations have moderate skewness, whereas UNV is highly skewed since it has skewness greater than 1. From the descriptive statistics, it can be seen that the UNV station dataset is noisy.

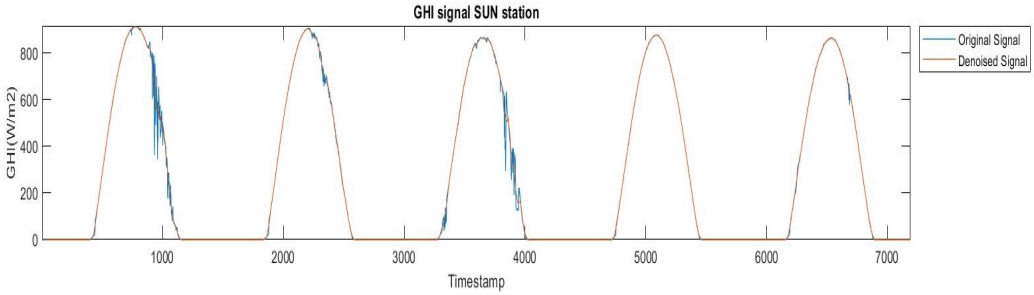
4.3.3 Data preprocessing

For solar energy generation planning, the model for GHI signal prediction can give valuable guidance to electricity suppliers, solar energy power suppliers and investors. However, due to excessive noise in GHI data, deep neural networks trained on the original data frequently fail to effectively forecast the GHI signal Liang et al., 2019. To address this issue,

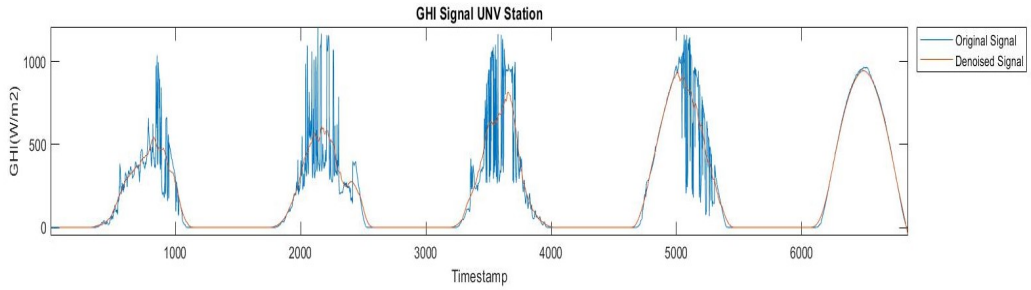
the training data was subjected to the wavelet threshold-denoising approach, which has been widely used in signal denoising Liang et al., 2019.



(A) RVD station Wavelet denoised vs original GHI over time.



(B) SUN station Wavelet denoised vs original GHI over time.



(C) UNV station Wavelet denoised vs original GHI over time.

FIGURE 4.3: Wavelet denoised vs original GHI over time for 3 sites.

Figure 4.3 shows the results from wavelet denoising, where a MODWT was applied with a soft thresholding method at the limit of level 2, and the haar family was used to denoise the data. The data still represents the original and once denoised, it is easy for the deep learning algorithm to capture the data's trends and relationships.

Descriptive statistics for denoised wavelet data(After preprocessing)

Table 4.2 reports the key descriptive statistics of the wavelet denoised GHI series. Since RVD and SUN GHI's show smooth curves, it is unnecessary to apply wavelets; thus, wavelets will be used to UNV GHI since the series has a lot of fluctuations.

TABLE 4.2: Descriptive statistics 1-minute wavelet denoised GHI data for UNV station.

Variables	Min	Mean	Std Dev	Kurtosis	Skewness	Max	n
Values	-49.8800	213.0866	282.5993	-0.0115	1.1277	946.2600	6849

The wavelet denoised GHI signal for the UNV site have a normal kurtosis value. Thus it is mesokurtic. Several models were considered to find the best model for the data to model the signal of GHI for the three sites.

4.4 Models comparison

4.4.1 Model results on original GHI signal(not denoised)

Results on UNV: USAid Venda station

TABLE 4.3: Original UNV GHI signal model results

Models	ARIMA	LSTM	ALSTM	Hydrid
MAE	8.855	7.818	9.317	7.383
RMSE	35.988	8.444	13.570	14.1293

The results in Table 4.3 show different models used to model and forecast GHI signal. The Hybrid Attention-based LSTM-ARIMA was the best model of all the models, followed by the LSTM model with little difference. This can be confirmed by the lower $MAE = 7.383$, and $RMSE = 14.1293$.

Figure 4.4 shows the respective GHI signal prediction results of the four prediction models on the GHI signal dataset. The black, red, green, and yellow lines represent the LSTM model, the ARIMA model, the attention LSTM model, and the proposed prediction model. The blue line indicates the actual GHI signal of the testing data time step. From figure 4.5, it can be seen that the prediction of the GHI signal of the hybrid model is the same as the original value in the trend, and the predicted value floats around the initial value.

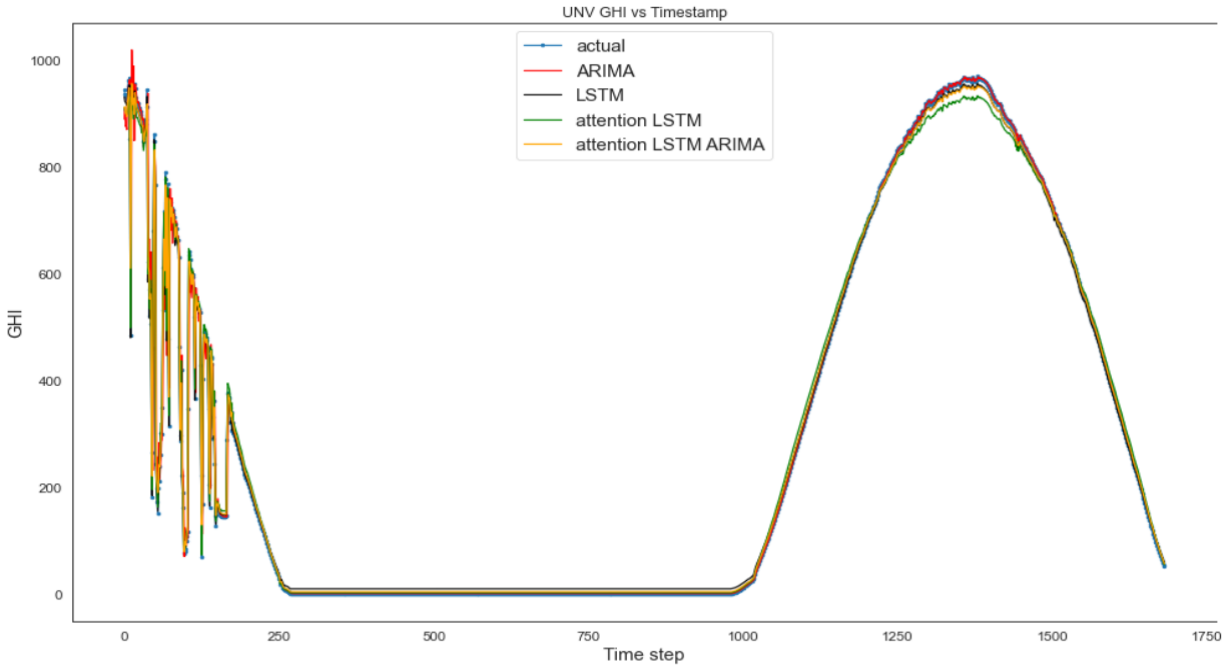


FIGURE 4.4: UNV GHI vs Timestamp of testing data predictions

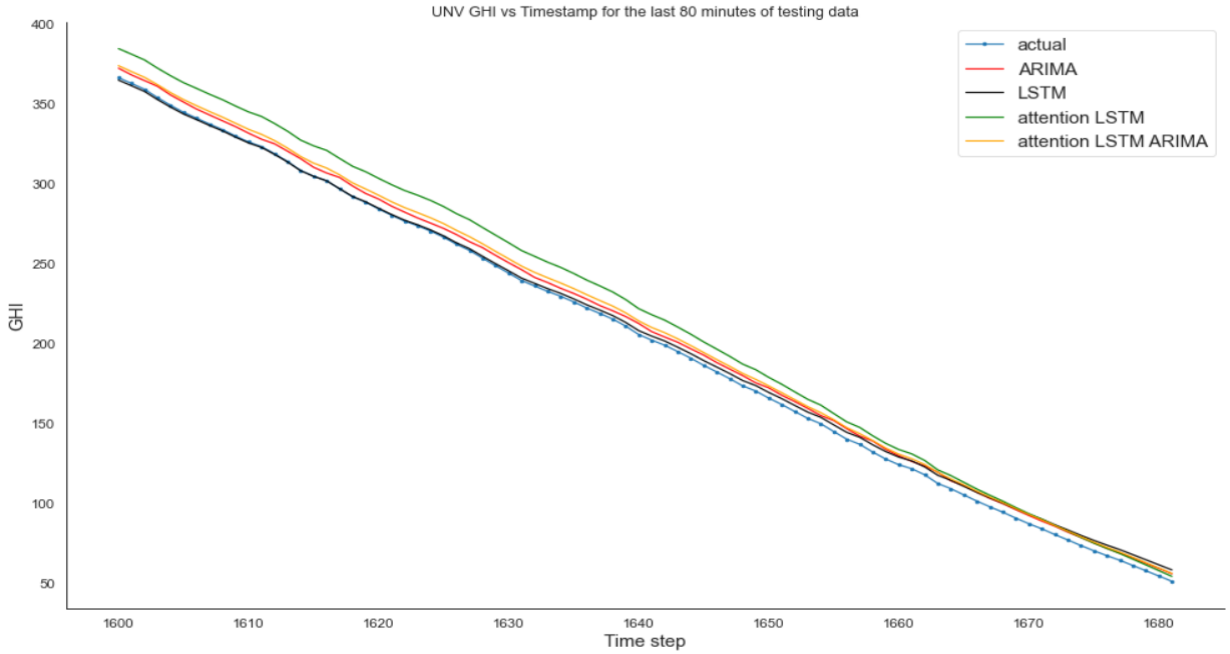


FIGURE 4.5: UNV GHI vs Timestamp for the subset of testing data

Results on RVD: GIZ Richtersveld

The results in table 4.4 show different types of models that were used to model and forecast GHI signal.

TABLE 4.4: Original RVD GHI signal model results

Models	ARIMA	LSTM	ALSTM	Hydrid
MAE	0.782	7.905	5.292	4.120
RMSE	1.271	9.969	7.314	4.987

In all the models, the ARIMA model was the best, followed by the hybrid Attention-based LSTM-ARIMA. This can be confirmed by the lower $MAE = 0.782$, and $RMSE = 1.271$.

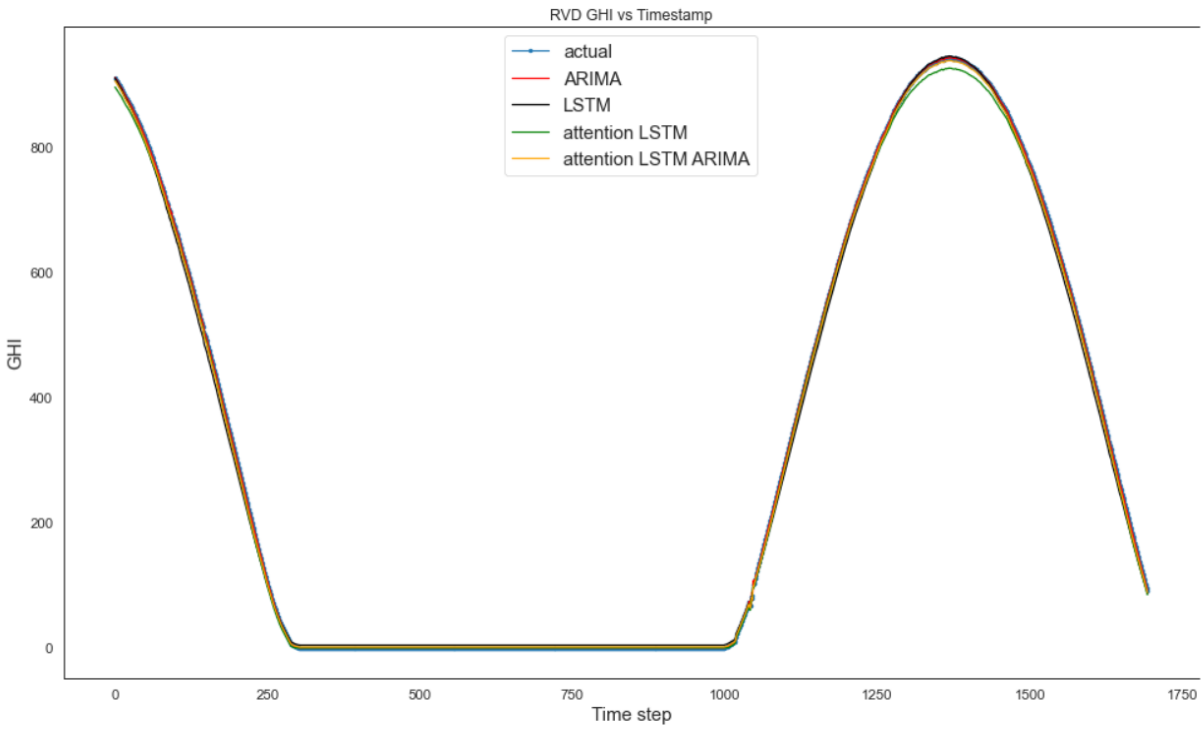


FIGURE 4.6: RVD GHI vs Timestamp for testing data and predictions

Figure 4.6 shows the respective GHI signal prediction results of the four prediction models on the GHI signal dataset. The black, red, green, and yellow lines represent the LSTM model, the ARIMA model, the attention LSTM model, and the hybrid model. The blue line indicates the actual GHI signal of the testing data.

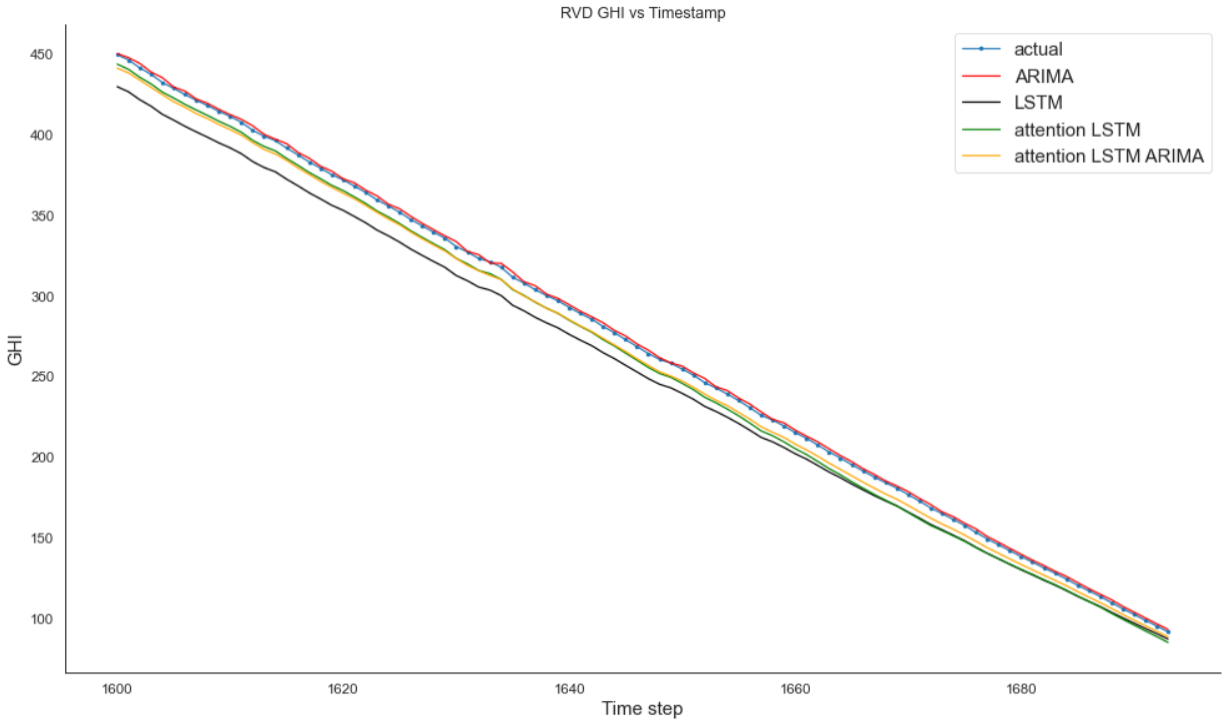


FIGURE 4.7: UNV GHI vs Timestamp for the subset of testing data

From Figure 4.7, it can be seen that the predictions of the GHI signal of the ARIMA model are the same as the original value in the trend, and the predicted value floats around the initial value.

Results on SUN: Stellenbosch University

The results in Table 4.5 show different types of models that were used to model and forecast GHI signal.

TABLE 4.5: Original SUN GHI signal model results

Models	ARIMA	LSTM	ALSTM	Hydrid
MAE	1.417	4.928	1.512	2.011
RMSE	3.726	6.235	1.640	2.511

In all the models, the Attention-based LSTM model was the best, followed by the hybrid Attention-based LSTM-ARIMA. This can be confirmed by the lower $MAE = 1.512$, and $RMSE = 1.640$.

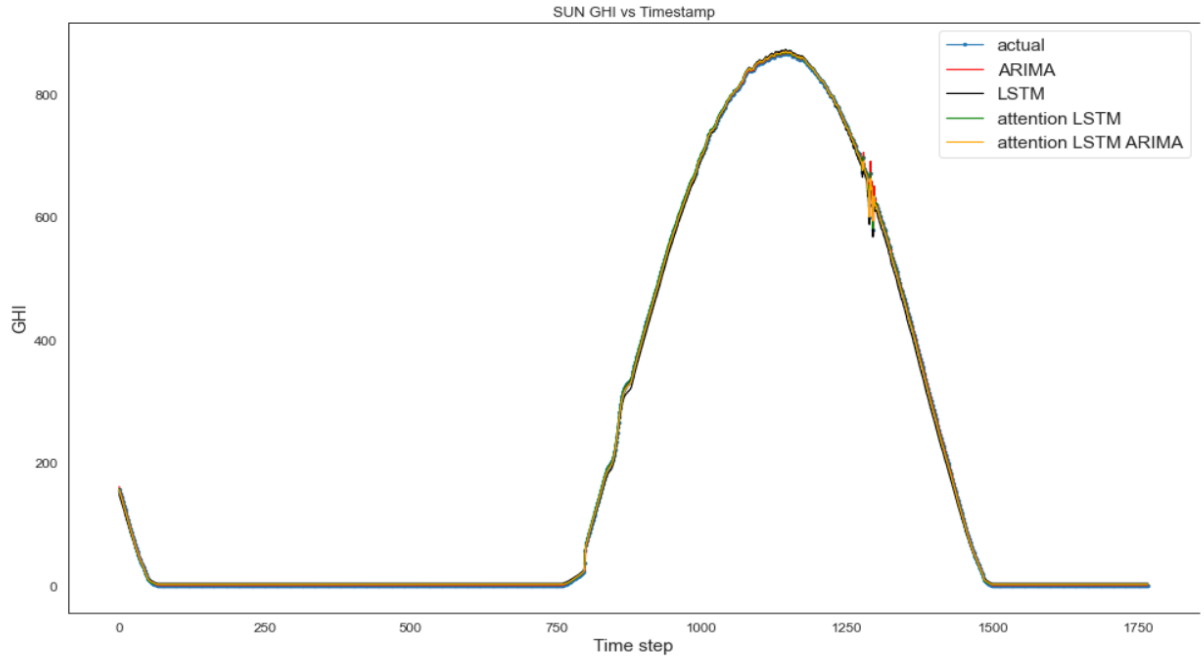


FIGURE 4.8: SUN GHI vs Timestamp for testing data and predictions

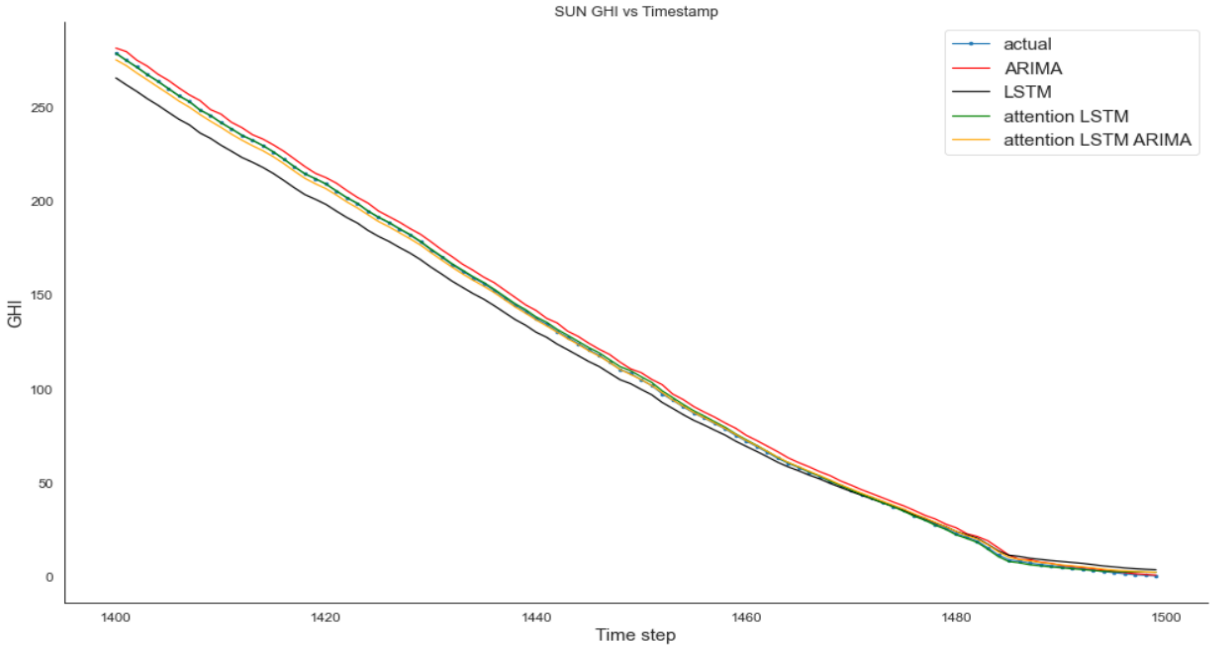


FIGURE 4.9: SUN GHI vs Timestamp for the subset of testing data

Figure 4.8 shows the respective GHI signal prediction results of the four prediction models on the GHI signal dataset. The black, red, green, and yellow lines represent the LSTM model, the ARIMA model, the attention LSTM model, and the hybrid model. The blue line indicates the actual GHI signal of the testing data.

From Figure 4.9, it can be seen that the prediction of the GHI signal of the ARIMA model is the same as the original value in the trend, and the predicted value floats around the initial value.

4.4.2 Model results on wavelet denoised GHI signal

The results in Table 4.6 show different types of models that were used to model and forecast GHI signal.

TABLE 4.6: Wavelet denoised UNV GHI signal model results

Models	ARIMA	LSTM	ALSTM	Hydrid
MAE	0.194	3.759	5.068	2.176
RMSE	0.542	4.950	5.778	2.308

In all the models, the ARIMA model was the best, followed by the hybrid Attention-based LSTM-ARIMA. This can be confirmed by the lower $MAE = 0.194$, and $RMSE = 0.542$.

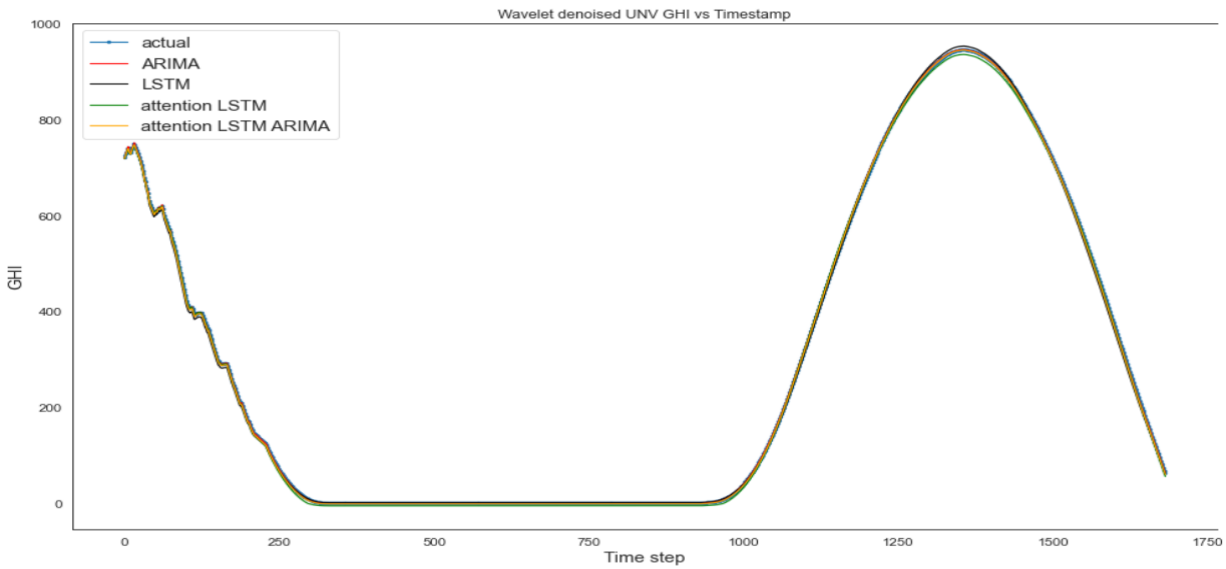


FIGURE 4.10: Wavelet denoised UNV GHI vs Timestamp for testing data and predictions

Figure 4.10 shows the respective GHI signal prediction results of the four prediction models on the GHI signal dataset. The black, red, green, and yellow lines represent the LSTM model, the ARIMA model, the attention LSTM model, and the hybrid model. The blue line indicates the actual GHI signal of the testing data.

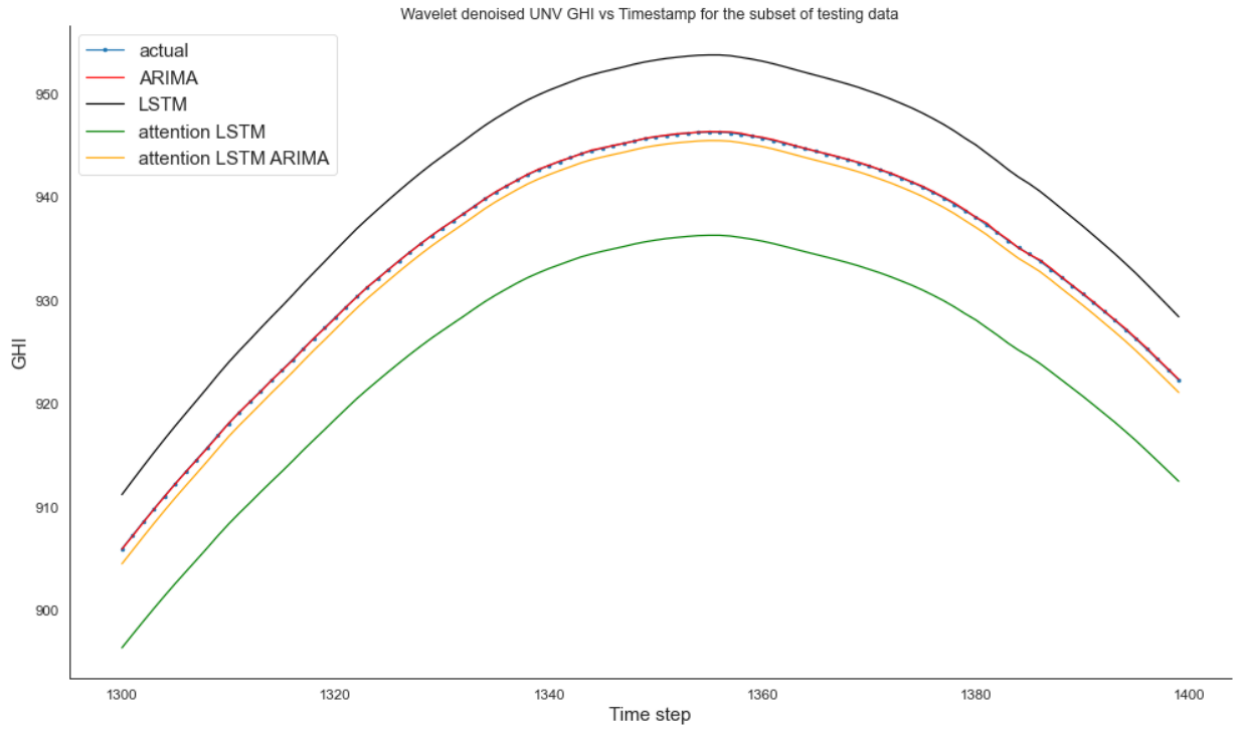


FIGURE 4.11: Wavelet denoised UNV GHI vs Timestamp for the subset of testing data

From Figure 4.11, it can be seen that the predictions of the GHI signal using the ARIMA model are the same as the original value in the trend, and the predicted value floats around the initial value.

4.4.3 Best Model Conclusions

UNV: USAid Venda station

The results from Table 4.3, supported by Figures 4.4 and 4.5 shows that the best model to forecast UNV GHI is the attention-based LSTM-ARIMA(hybrid). From the Figures 4.4 and 4.5, the line which is much closer to the original value is the one which represents the hybrid

model. This implies that the hybrid model best predicts and forecasts GHI from the UNV station.

RVD: GIZ Richtersveld station

The results from Table 4.4, supported by Figures 4.6 and 4.7 show that the best model to forecast RVD GHI is ARIMA. From the Figures 4.6 and 4.7, the line which is much closer to the original value is the one which represents the ARIMA model. This implies that the ARIMA model is the best model to predict and forecast GHI from the RVD station.

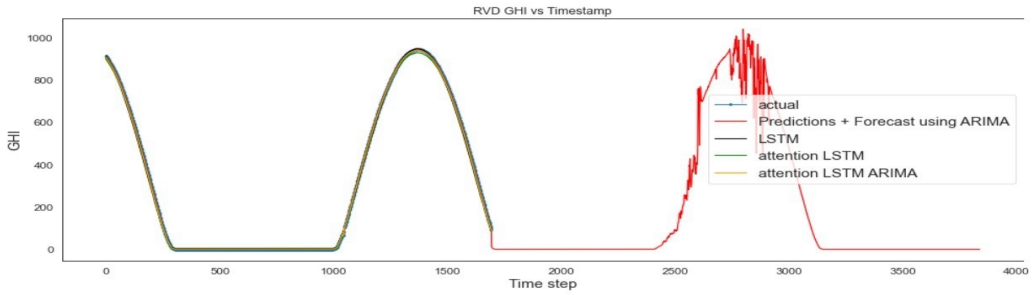
SUN: Stellenbosch University station

The results from Table 4.5, supported by Figures 4.8 and 4.9 shows that the best model to forecast RVD GHI is attention-based LSTM. From the Figures, 4.8 and 4.9, the line which is much closer to the original value is the one which represents the attention-based LSTM model. This implies that the best model to predict and forecast GHI from the SUN station is the attention-based LSTM model.

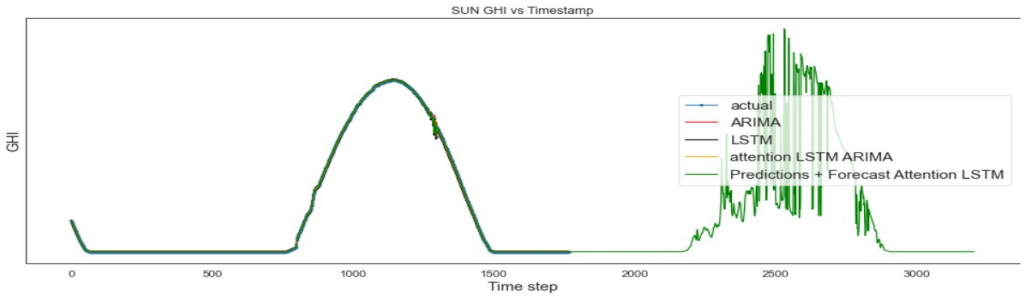
UNV: USAid Venda station(Wavelet denoised GHI)

The results from Table 4.6, supported by Figures 4.10 and 4.11 shows that the best model to forecast RVD GHI is ARIMA. From the Figures 4.10 and 4.11, the line which is much closer to the original value is the one which represents the ARIMA model. The best model to predict and forecast wavelet denoised GHI from UNV station is ARIMA.

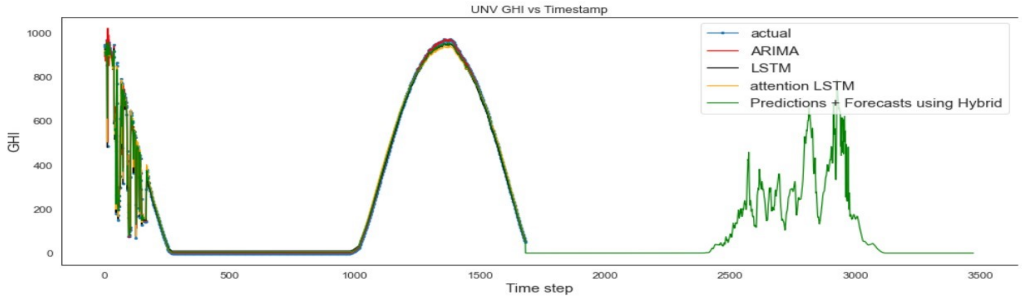
4.4.4 Best Model Future Forecasts



(A) 1 day forecast of GHI for RVD station over time.



(B) 1 day forecast of GHI for SUN station over time.



(C) 1 day forecast of GHI for UNV station over time.

FIGURE 4.12: Time series forecasts plots of GHI vs time for 3 sites.

Figure 4.12 shows the predicted values and future forecasts using the best model. From each of the Figures 4.12 (A)-(C), it can be seen that the best models were able to capture GHI trends and seasonality.

Chapter 5

Conclusions and future directions

5.1 Introduction

This chapter summarizes the research findings and proposes some recommendations. The study's limitations, including areas for future research, are also discussed in this chapter.

5.2 Study findings

This project explored the analysis of solar radiation. All the findings for this project are based on the GHI time-series data for three different sites in South Africa; namely, UNV (from 23 March 2021 to 28 March 2021), SUN (from 23 March 2021 to 28 March 2021) and RVD (from 30 April 2020 to 4 May 2020) stations retrieved from https://sauran.ac.za/. The main reason for the selected period was that it covers the recent timestamp with no null values.

Descriptive statistics indicate that the GHI signals for the three sites are not normally distributed. This study was carried out assuming univariate distributions and their skew variants.

An overall of 4 models, namely ARIMA, LSTM, attention-based LSTM, and hybrid attention-based LSTM-ARIMA, were used to forecast the signals of GHI in which the hybrid attention-based LSTM-ARIMA model on USAid Venda station was found to be the best fitting model with $RMSE = 7.383$ and $MAE = 14.1293$, followed by LSTM with $MAE = 7.817$ and $RMSE = 8.444$. Comparing the results on non-wavelet denoised and wavelet denoised, models performed better on wavelet denoised data. ARIMA model was the best with $MAE = 0.194$ and $RMSE = 0.542$.

The best model for forecasting GHI from GIZ Richtersveld station is ARIMA with $MAE = 0.782$ and $RMSE = 1.271$, followed by hybrid model with $MAE = 4.120$ and $RMSE =$

4.987. For Stellenbosch University station, attention LSTM was the best with $MAE = 1.512$ and $RMSE = 1.640$, followed by hybrid with $MAE = 2.011$ and $RMSE = 2.511$.

5.3 Limitations

The limitation of this study is that only four type models in a whole bunch of models were used and the comparison was made in four models.

5.4 Conclusion

5.4.1 UNV: USAid Venda station

The results from Table 4.3, supported by Figures 4.4 and 4.5 shows that the best model to forecast UNV GHI is the attention-based LSTM-ARIMA(hybrid). From the Figures 4.4 and 4.5, the line which is much closer to the original value is the one which represents the hybrid model. This implies that the hybrid model is the best to predict and forecast GHI from the UNV station.
On this station, an improved hybrid model (ARIMA-LSTM with attention) was used to have a better GHI prediction model. The performance was better than traditional methods. The experimental results illustrate that the data characteristics are of great significance to the performance of the whole model (Qiu, Wang, and Zhou, 2020).

5.4.2 RVD: GIZ Richtersveld station

The results from Table 4.4, supported by Figures 4.6 and 4.7 shows that the best model to forecast RVD GHI is ARIMA. From the Figures 4.6 and 4.7, the line which is much closer to the original value is the one which represents the ARIMA model. This implies that the ARIMA model is the best model to predict and forecast GHI from the RVD station.

5.4.3 SUN: Stellenbosch University station

The results from Table 4.5, supported by Figures 4.8 and 4.9 shows that the best model to forecast RVD GHI is attention-based LSTM. From the Figures 4.8 and 4.9, the line which is much closer to the original value is the one which represents the attention-based LSTM model. This implies that the best model to predict and forecast GHI from the SUN station is the attention-based LSTM model.

5.4.4 UNV: USAid Venda station(Wavelet denoised GHI)

The results from Table 4.6, supported by Figures 4.10 and 4.11 shows that the best model to forecast RVD GHI is ARIMA. From the Figures 4.10 and 4.11, the line which is much closer to the original value is the one which represents the ARIMA model. The best model to predict and forecast wavelet denoised GHI from UNV station is ARIMA. The prediction results for the original data technique without wavelet processing exhibit substantial fluctuations, and the fitting impact of the real data is weak (Qiu, Wang, and Zhou, 2020). The model's overall performance is impacted.

5.5 Areas of future study

Future research should look at forecasting GHI signals using a combination of several statistical, mathematical and Deep learning models (e.g., it can be a hybrid attention-based LSTM-GRU-ARIMA-GARCH on wavelet denoised GHI model). Also, future studies should look at modelling GHI for several sites to see which model is the best based on having the smallest RMSE, MSE, MAE and bigger R^2 on both or most stations. This will help check for the persistence of volatility in the renewable energy sector.

References

Achilleas, Z. and R. Apostolos-Paul (1999). "Principles of neural model identification, selection and adequacy: with applications to financial econometrics". In: Springer Science & Business Media.

Alonso-Montesinos, J., F.J. Batlles, and C. Portillo (2015). "Solar irradiance forecasting at one-minute intervals for different sky conditions using sky camera images". In: *Energy Conversion and Management* 105, pp. 1166–1177. ISSN: 0196-8904. DOI: https://doi.org/10.1016/j.enconman.2015.09.001. URL: https://www.sciencedirect.com/science/article/pii/S0196890415008365.

Attention mechanism. URL: https://blog.floydhub.com/attention-mechanism/.

Box, G.E.P, G.M Jenkins, G.C Reinsel, and G.M Ljung (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Bullmore, A. (2002). "The illustrated wavelet transform handbook, Institute of Physics Publishing". In: *Bristol, UK*.

Chih-Chiang, W. (2017). "Predictions of Surface Solar Radiation on Tilted Solar Panels using Machine Learning Models: A Case Study of Tainan City, Taiwan". In: *Energies* 10.10. ISSN: 1996-1073. DOI: 10.3390/en10101660. URL: https://www.mdpi.com/1996-1073/10/10/1660.

Daubechies, I. (1990). "The wavelet transform, time-frequency localization and signal analysis". In: *IEEE Transactions on Information Theory* 36.5, pp. 961–1005. URL: http://doi.org/0.1109/18.57199.

Fourier, J. (1807). "Sine and Cosine Series for an Arbitrary Function in Joseph Fourier 1768-1830 Ed. and annotated by I". In: *Grattan-Guinness The MIT Press, Cambridge, MA*.

Haar, A. (1909). "on the theory of orthogonal functional systems". In: Georg August University, Gottingen.

He, J. and D. Yao (2016). "A nonlinear support vector machine model with hard penalty function based on glowworm swarm optimization for forecasting daily global solar radiation". In: *Energy Conversion and Management* 126, pp. 991–1002. ISSN: 0196-8904. DOI: https://doi.org/10.1016/j.enconman.2016.08.069. URL: https://www.sciencedirect.com/science/article/pii/S0196890416307439.

Hochreiter, Sepp and Jürgen Schmidhuber (1997). "Long short-term memory". In: *Neural computation* 9.8, pp. 1735–1780.

Huang, Norden E, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H Liu (1998). "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis". In: *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences* 454.1971, pp. 903–995.

Illustrated Guide to Recurrent Neural Networks. URL: https://towardsdatascience.com/illustrated-guide-to-recurrent-neural-networks-79e5eb8049c9.

Jang, H. and J. Lee (2018). "An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information". In: *IEEE Access* 6, pp. 5427–5437. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2017.2779181.

Jianwu, Z. and Q. Wei (2013). "Short-term solar power prediction using a support vector machine". In: *Renewable Energy* 52, pp. 118–127. ISSN: 0960-1481. DOI: https://doi.org/10.1016/j.renene.2012.10.009. URL: https://www.sciencedirect.com/science/article/pii/S0960148112006465.

Liang, X., G. Zhaodi, S. Liling, H. Maowei, and C. Hanning (2019). "LSTM with wavelet transform based data preprocessing for stock price prediction". In: *Mathematical Problems in Engineering* 2019.

Lu, A., Z. Wang, and H. Xu (2018). "Stock prediction with deep learning framework". In: pp. 1–6. URL: http://cs230.stanford.edu/projects_fall_2018/reports/12450133.pdf.

Merry, R.J.E and M. Steinbuch (June 2005). "Wavelet theory and applications. literature study". In: vol. 41. 53, pp. 3–5. URL: http://www.mate.tue.nl/mate/pdfs/5500.pdf.

Mutavhatsindi, T., C. Sigauke, and R. Mbuvha (2020). "Forecasting Hourly Global Horizontal Solar Irradiance in South Africa Using Machine Learning Models". In: *IEEE Access* 8, pp. 198872–198885. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3034690.

Ospina, Juan, Alvi Newaz, and M Omar Faruque (2019). "Forecasting of PV plant output using hybrid wavelet-based LSTM-DNN structure model". In: *IET Renewable Power Generation* 13.7, pp. 1087–1095.

Percival, D.B and A.T Walden (2000). "The Maximal Overlap Discrete Wavelet Transform". In: *Wavelet Methods for Time Series Analysis*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, pp. 159–205. DOI: 10.1017/CBO9780511841040.006.

Qiu, J., B. Wang, and C. Zhou (2020). "Forecasting stock prices with long-short term memory neural network based on attention mechanism". In: *PloS one* 15.1. URL: https://doi.org/10.1371/journal.pone.0227222.

Qiu, J., B. Wang, and C. Zhou (Jan. 2020). "Forecasting stock prices with long-short term memory neural network based on attention mechanism". In: *PLOS ONE* 15.1, pp. 1–15.

DOI: 10.1371/journal.pone.0227222. URL: https://doi.org/10.1371/journal.pone.0227222.

Renno, C., F. Petito, and A. Gatto (2015). "Artificial neural network models for predicting the solar radiation as input of a concentrating photovoltaic system". In: *Energy Conversion and Management* 106, pp. 999–1012. ISSN: 0196-8904. DOI: https://doi.org/10.1016/j.enconman.2015.10.033. URL: https://www.sciencedirect.com/science/article/pii/S0196890415009553.

Rioul, O. and M. Vetterli (Oct. 1991). "Wavelets and signal processing". In: *IEEE Signal Processing Magazine* 8.4, pp. 14–38. URL: https://doi.org/10.1109/79.91217.

Schneiders, M.G.E, M.J.G van de Molengraft, and M. Steinbuch (2001). *Wavelets in control engineering*. Technische Universiteit Eindhoven.

Selvin, S., R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman (2017). "Stock price prediction using LSTM, RNN and CNN-sliding window model". In: *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1643–1647.

Sifuzzaman, M., M.R. Islam, and M.Z. Ali (Oct. 2009). "Application of Wavelet Transform and its Advantages Compared to Fourier Transform". In: *Journal of Physical Science* 13, pp. 121–134. URL: http://inet.vidyasagar.ac.in:8080/jspui/handle/123456789/779.

Sileyew, K.J. (Sept. 2019). "Research Design and Methodology[Online First]". In: *IntechOpen* 13, pp. 2–3. URL: https://www.intechopen.com/online-first/research-design-and-methodology.

Tukey, John W (1977). *Exploratory data analysis*. Vol. 2. Reading, MA.

Urraca, R., J. Antonanzas, M. Alia-Martinez, F.J. Martinez de Pison, and F. Antonanzas-Torres (2016). "Smart baseline models for solar irradiation forecasting". In: *Energy Conversion and Management* 108, pp. 539–548. ISSN: 0196-8904. DOI: https://doi.org/10.1016/j.enconman.2015.11.033. URL: https://www.sciencedirect.com/science/article/pii/S0196890415010535.

Vértes, E., A.F Alexander-Bloch, N. Gogtay, J.N Giedd, J.L Rapoport, and E.T Bullmore (2012). "Simple models of human brain functional networks". In: *Proceedings of the National Academy of Sciences* 109.15, pp. 5868–5873.

Wang, F., Z. Zhen, B. Wang, and Z. Mi (2018a). "Comparative Study on KNN and SVM Based Weather Classification Models for Day Ahead Short Term Solar PV Power Forecasting". In: *Applied Sciences* 8.1. ISSN: 2076-3417. DOI: 10.3390/app8010028. URL: https://www.mdpi.com/2076-3417/8/1/28.

Wang, F., K. Li, C. Liu, Z. Mi, M. Shafie-Khah, and J. P. S. Catalão (Nov. 2018b). "Synchronous Pattern Matching Principle-Based Residential Demand Response Baseline Estimation: Mechanism Analysis and Approach Description". In: *IEEE Transactions on Smart Grid* 9.6, pp. 6972–6985. ISSN: 1949-3061. DOI: 10.1109/TSG.2018.2824842.

Wang, Fei, Yili Yu, Zhanyao Zhang, Jie Li, Zhao Zhen, and Kangping Li (2018). "Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting". In: *applied sciences* 8.8, p. 1286.

West, Samuel R, Daniel Rowe, Saad Sayeef, and Adam Berry (2014). "Short-term irradiance forecasting using skycams: Motivation and development". In: *Solar Energy* 110, pp. 188–207.

Wu, Suwan, Li Jia, and Yining Liu (2021). "Ultra-short-term wind energy prediction based on wavelet denoising and multivariate LSTM". In: *2021 Power System and Green Energy Conference (PSGEC)*. IEEE, pp. 443–447.

Yan, S. (2016). *Understanding LSTM and its diagrams*. URL: https://medium.com/mlreview/understanding-lstm-and-its-diagrams-37e2f46f1714.

Yang, Fan, Yuancun Cui, Feng Wu, and Ridong Zhang (2021). "Fault Monitoring of Chemical Process Based on Sliding Window Wavelet DenoisingGLPP". In: *Processes* 9.1, p. 86.