Thunderstorm Severity Prediction for South Africa between 2015 to 2016 using LSTM Model Variants and Remote Sensing Instruments

---

Yaseen Essa
Student Number: 0101198N

Supervisors:
Dr Ritesh Ajoodha
Dr Hugh G. P. Hunt

A dissertation in partial fulfilment for the degree of Master of Science in the field of e-Science

at

the University of the Witwatersrand, Johannesburg

1 June 2022

Declaration

I, Yaseen Essa, declare that this research report is my own, unaided work. It is being submitted for the degree of Master of Science in the field of e-Science at the the University of the Witwatersrand, Johannesburg. It has not been submitted for any degree or examination at any other university.

Yaseen Essa
Student Number: 0101198N
1 June 2022

Abstract

Lightning is one of the leading causes of electrical outages in South Africa, and the most severe weather-related killer in the country. Unfortunately, quantitative lightning prediction remains challenging. In this research, we evaluate the accuracy of LSTM neural network model variants on lightning frequency prediction using remote sensing instruments.

These LSTM model variants are the LSTM-Fully Connected (LSTM-FC), the Convolutional-Neural-Network-LSTM (CNN-LSTM) and the Convolutional-LSTM (ConvLSTM) models. Both the CNN-LSTM model and the ConvLSTM model recognize spatiotemporal features. The data consists of lightning detection data from the South African Lightning Detection Network (SALDN) and weather-feature information from the South African Weather Service. We forecast lighting frequency, every hour, between December-2013 and March-2016 within the predicted area. The models used in this research were trained on data between July-2008 to November-2013. All models were trained to minimize Mean Squared Error but were evaluated on Mean Absolute Error (MAE flashes.hr^{-1}). We also varied models based on input datasets: SALDN-only, SAWS-only and SALDN+SAWS datasets.

We found that the CNN-LSTM model (MAE=51) performed best amongst LSTM model variants (LSTM-FC MAE=67; ConvLSTM MAE=86). When models were evaluated between input datasets, we found that SALDN only (MAE= 59) outperformed SAWS only and SALDN+SAWS (SAWS MAE=74; SAWS+SALDN MAE=70). We conclude that CNN-LSTM models outperform prediction accuracy for thunderstorm severity compared with ConvLSTM and FC-LSTM models but consideration on input data is required.

Acknowledgements

We acknowledge the support in part by the National Research Foundation of South Africa (Grant numbers: 121835 and 98244) and the DSI-NICIS National e-Science Postgraduate Teaching and Training Platform (NEPTTP). Note that the conclusions and opinions are those of the author only and not endorsed by the NEPTTP.

We also thank the South African Weather Service, specifically to Morné Gijben and Rydall Jardine, for the use of the SAWS and SALDN datasets.

Contents

List of Figures

List of Tables

List of Abbreviations

CNN	**C**onvolutional **N**eural **N**etwork
ConvLSTM	**Conv**olutional **LSTM**
LSTM	**L**ong **S**hort **T**erm **M**emory (Neural Network).
RNN	**R**ecurrent **N**eural **N**etwork
SALDN	**S**outh **A**frican **L**ightning **D**etection **N**etwork
SAWS	**S**outh **A**frican **W**eather **S**ervice.

Definitions

Lightning flash (lightning discharge, lightning). These terms are used interchangeably to describe from the point before a lightning leader emerges until all subsequent discharges and/or strokes.

Lightning stroke. Refers to each time the attachment process occurs. A lightning flash often have numerous lightning strokes.

Cloud-to-ground lightning. This refers to lightning between cloud and the ground. It may occur either in the direction of cloud-to-ground or ground-to-cloud.

LSTM neural network. A Recurrent Neural Network variant that has the functionality to filter inputs from recurrent connections.

Thunderstorm severity. The quantity of lightning flashes per unit time per area.

Chapter 1

Introduction

A lightning strike fascinates and frightens in a split of a second. Lightning is more than just a geophysical phenomena, lightning has also been associated with deity. These include the Greek God Zeus, Norse God Thor, and Hindu God Indra. As human understanding of lightning increased, we have rationalised lightning and turned our attention on how to manage the effects of lightning.

In general, most lightning flashes occur between clouds which is termed intra-cloud (IC) lightning. However, Cloud-to-Ground (CG) lightning is of more interest to us due to CG lightnings' greater influence on human activity. The ratio of IC/CG varies based on the lightning characteristics, location and also weather conditions. Additionally, the IC/CG flash ratio may vary during a thunderstorm [24]. For South Africa, Maseko (2021) estimated an IC/CG ratio of 3 - the most commonly used ratio for the mid-latitude region [23].

1.1 Problem Statement

The destructive effects of lightning has anguished humans for millennia. In modern times, lightnings has deleterious effects on wind turbines [27], electrical infrastructure and electrical equipment [19]. Most of this damage is caused by cloud-t0-ground lightning flashes

Lightning has an increased impact in South Africa as the country has a relatively high lightning density, especially in the North-Eastern area of the country (Figure 1.1). Here, lightning causes about twenty-percent of all electrical distribution outages [3]. Further it is estimated that, on average, 264-people die from lightning strikes in South Africa yearly [18].

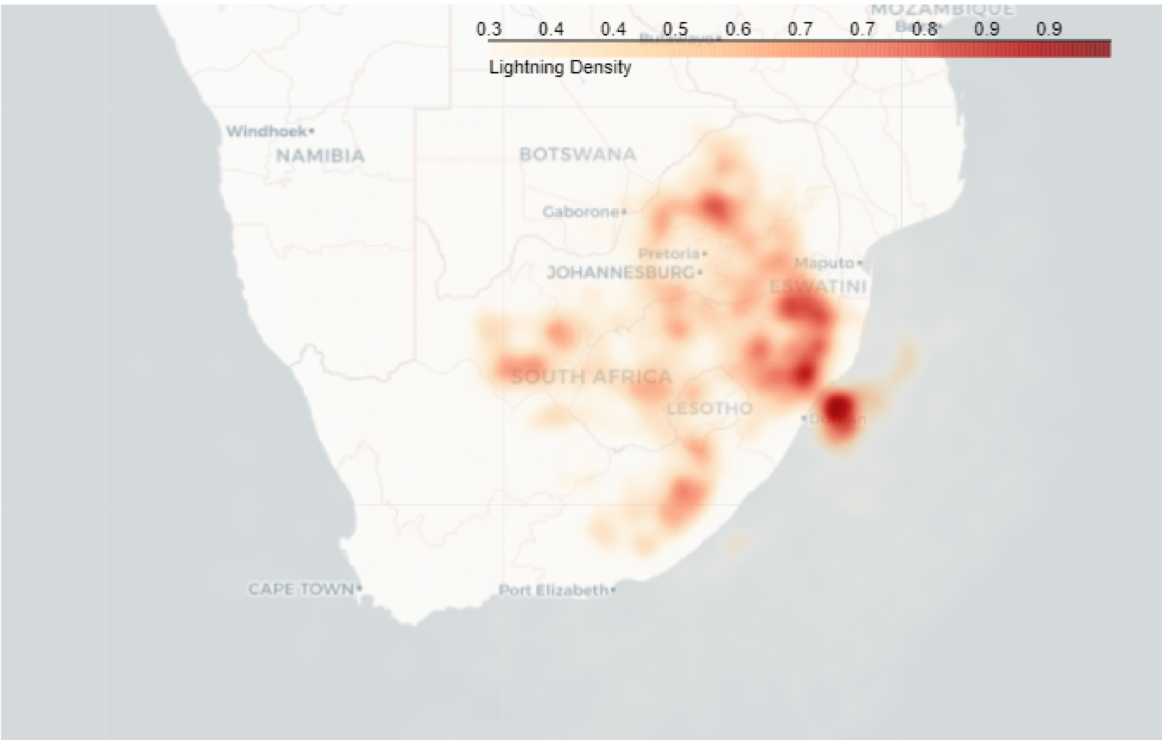

FIGURE 1.1: Heatmap of lightning density in South Africa for year 2015. Scale from Red (High) to Yellow (Low).

As lightning strikes cannot be prevented or ameliorated, knowing the quantity of expected lightning will assist in relief preparation and risk management. Unfortunately though, the prediction of lightning is challenging. Lightning prediction is a non-linear, multifactorial, stochastic, cyclical (time-dependent) problem with multiple dependencies.

Recent developments in deep learning gave the rise to spatial LSTM (Long Short Term Memory) neural network model variants that not only considers temporal dimension but also spatial arrangement (Figure 1.2). The advantage of this is that these models consider its surrounding environment when interpreting information. Two of these models are the Convolutional Neural Network LSTM (CNN-LSTM) and the Convolutional LSTM (ConvLSTM). Both of these models have been used in weather prediction [14, 44, 5].

Most lightning prediction models focus on predicting the likelihood of lightning occurring or not occurring - and not the amount of expected lightning (Table 2.1 and Table 2.2). Further, current commercial models do not employ neural network deep

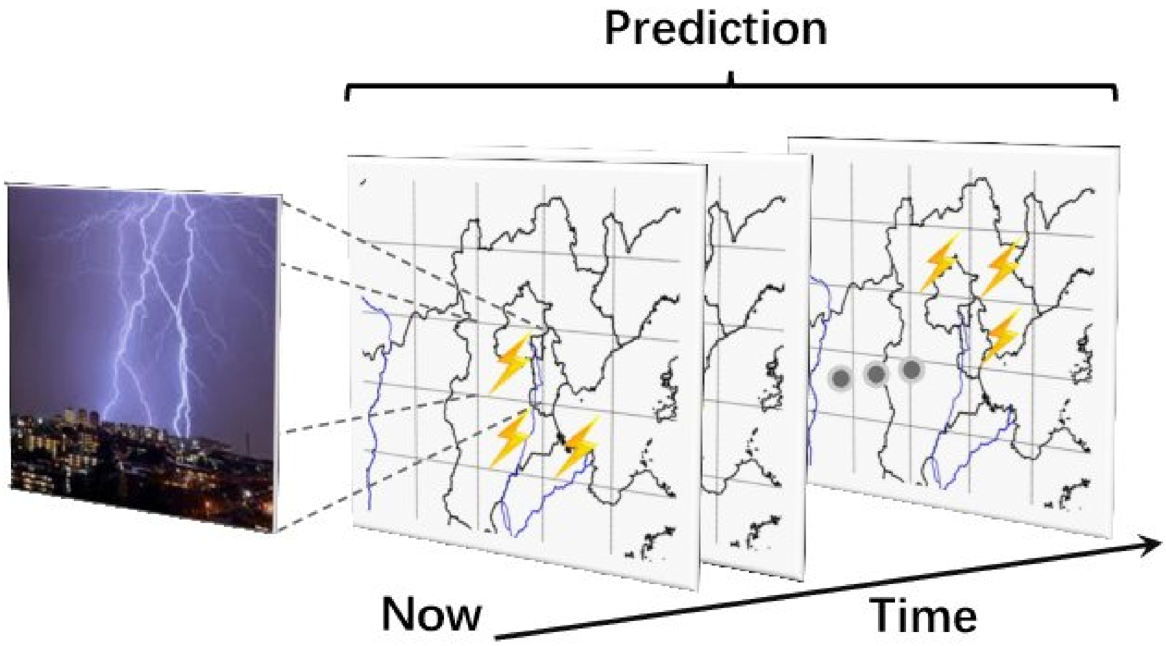

FIGURE 1.2: Spatio-Temporal input of CNN-LSTM and ConvLSTM Models. Data is viewed in context of its 2D spatial arrangement. Reproduced from [14].

learning which is an attractive tool for lighting forecasting.

1.2 Purpose Statement

The aim of this research was to develop quantitative thunderstorm severity prediction models that are accurate and predict in real time. Lightning frequency was used as an indicator of thunderstorm severity. The accuracy criterion was achieved with the employment of Long-Short-Term-Memory (LSTM) Recurrent Neural Networks (RNN) neural network variants as LSTM models are well-suited for time-series modelling [30]. For the model to be cost-effective, data from already deployed remote sensing instruments were used, *viz.*, the South African Lightning Detection Network and the Automatic Weather stations from the South African Weather Service (SAWS). The real-time prediction criterion was met by quick processing from the LSTM model variants after training.

1.3 Research Questions

To address the research problem and achieve the purpose of this research, the following research questions were investigated:

1. What is the prediction accuracy, as measured by Mean Absolute Error, of the following LSTM RNN Neural Network model variants to accurately predict short-term (1hr) cloud-to-ground lightning frequency in South Africa. These variants are:
   1.1 LSTM-Fully Connected network (LSTM-FC),
   1.2 Convolutional Neural Network LSTM (CNN-LSTM) and
   1.3 Convolutional LSTM network (ConvLSTM).
2. What is the prediction accuracy, as measured by Mean Absolute Error, of the LSTM RNN Neural Network model variants with various input dataset combinations for short-term (1hr) cloud-to-ground lightning frequency prediction in South Africa. These input datasets are:
   2.1 SALDN dataset only
   2.2 SAWS dataset only
   2.3 SALDN and SAWS datasets

1.4 Research Objectives

In order to address the research problem and achieve the purpose of this research, we sought the following research objectives:

1. Develop and train LSTM RNN Neural Network model variants for short-term (1hr) cloud-to-ground lightning frequency prediction in South Africa. These variants are:
   1.1 LSTM-Fully Connected network (LSTM-FC),
   1.2 Convolutional Neural Network LSTM (CNN-LSTM) and
   1.3 Convolutional LSTM network (ConvLSTM).

2. Evaluate (on Mean Absolute Error) and compare LSTM model variants on unseen datasets.

3. Develop and train LSTM RNN Neural Network models with various input datasets for short-term (1hr) cloud-to-ground lightning frequency prediction in South Africa. These input datasets are:

   3.1 SALDN dataset only

   3.2 SAWS dataset only

   3.3 SALDN and SAWS datasets

4. Evaluate (on Mean Absolute Error) and compare LSTM model variant performance based on input datasets.

1.5 Limitations

This research has several factors that limit generalization. Firstly, results from this study do not apply to other locations; for this, training will need to be repeated. Secondly, as lightning is stochastic, models need to be interpreted probabilistically and not deterministically. Third, models would have to be re-trained with larger datasets and higher resolutions for implementation purposes.

Fourth, due to the use of neural network models, this research will not provide insights into the explanatory variables, and how they interact to produce lightning density. Neural networks make interpretation of factors and parameters *impractical*. Fifth, the weather data used in this study are from Automatic Weather Stations only. Thus this data is discrete in nature and there will be no data for areas that have no weather stations. This lack of continuity is not optimal for the CNN-LSTM and ConvLSTM models.

1.6 Research Contribution

This study is first to evaluate the use of spatio-temporal neural networks for lightning prediction in South Africa. This study is also first to use remote sensing data from both automatic weather stations and lightning detection networks for lightning prediction within South Africa. Further, our study has laid the way for quantitative, as opposed to lightning-likelihood, lightning prediction analysis.

1.7 Overview

In this study, we evaluated the use of LSTM neural network model variants on their accuracy on short-term thunderstorm severity. We assess thunderstorm severity as the amount of lightning flashes per unit time within the predicted area. We also evaluated feature combinations.

Weather is non-linear, stochastic, cyclical (time-dependent) and has complex dependencies. LSTM RNN models are well suited to capture these complexities, and provide the possibility of infinite complexity scaling [34]. Further, CNN-LSTM and ConvLSTM models provide additional spatial recognition. The SAWS input data is used to construct lightning flash severity through an additive/multiplicative model. While the historical SALDN lightning data acts to regulate lightning severity. We hope that a pragmatic lightning severity prediction tool will assist in lightning risk management.

We found the CNN-LSTM model (MAE=51 flashes.hr^{-1}) performed best amongst LSTM model variants (LSTM-FC MAE=67; ConvLSTM MAE=86). When models were evaluated between input datasets, we found that SALDN only (MAE=59) outperformed SAWS only and SALDN+SAWS (SAWS MAE=74; SAWS+SALDN MAE=70). We conclude that CNN-LSTM models outperform prediction accuracy compared with ConvLSTM and FC-LSTM models but consideration on input data is required.

1.8 Dissertation Structure

This chapter discussed the usefulness of quantitative thunderstorm severity prediction models for risk mitigation. The research questions sought to evaluate prediction accuracy of three LSTM-RNN model variants with various input dataset combinations. The following chapter 2 gives background on smart disaster management, lightning physics and a literature review of lightning prediction. Chapter

3 details LSTM neural networks and discusses how this research was carried out. Results and a discussion of the results is found in chapter 4. Lastly, chapter 5 concludes and give recommendations for future work.

Chapter 2

Background and Related Work

In this chapter, we provide a conceptual framework for smart city disaster management. This is followed by a background of lightning physics and lightning attachment. And end with a literature review of lightning prediction research.

2.1 Conceptual Framework for Smart City Disaster Management

Ideally, a smart city should anticipate disasters (Figure 2.1). The prediction of thunderstorm severity will aid in the protection of both human life and infrastructure by responding to disasters before lightning damage occurs.

However, the standardized disaster response system in South Africa needs further advancement. Currently, lightning prediction is not quantitative; thus we are unable to predict the magnitude or extent of lightning damage.

We propose the use of artificial intelligence prediction algorithms within a lightning risk-management system. For example, a utility company may activate resources and personnel for expected number of electrical disruptions from a thunderstorm. This prediction will lead to quicker repairs, lower operational costs and increased economic activity. A schematic of this framework is shown in Figure 2.1.

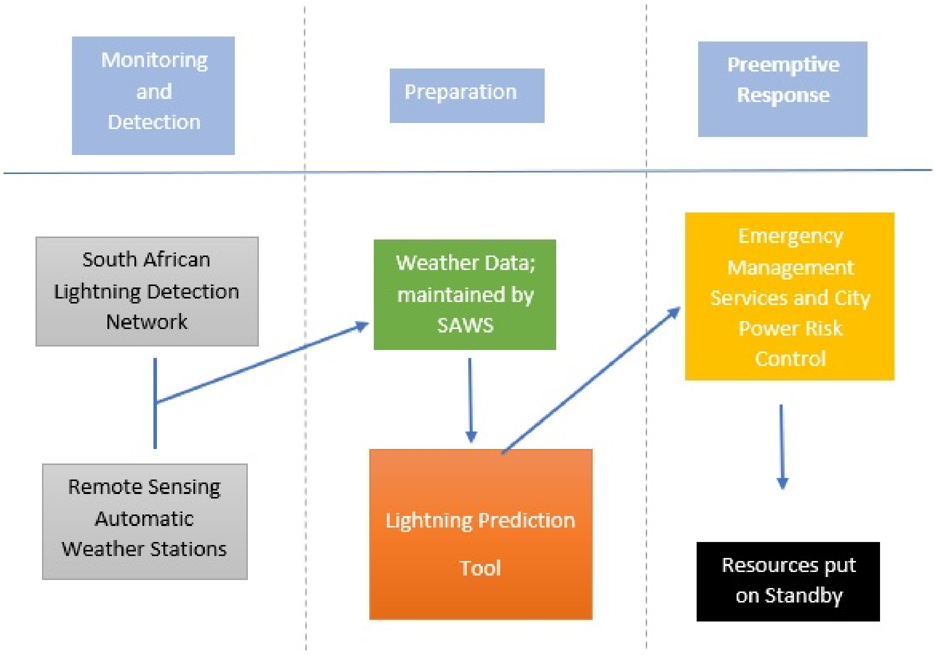

FIGURE 2.1: Conceptual framework of an intelligent decision warning system for smart city lightning risk management.

2.2 Thunderclouds and the Mechanics of a Lightning Flash

The prototypical charge structure of an individual thundercloud consist of tens of Coulombs of positive charge in its upper portion and a similar value of negative charge in its lower portion. Typically, a small positive charge is also found below the main negative charge (Figure 2.2). These electrostatically-charged tripole regions form due the collision - causing also a charge transfer - of particles by upward drafts within clouds between graupel (soft hail) and lighter ice crystals [1, 9, 33]. The positively-charged ice-crystals are light enough to be carried by updrafts to the top of the thundercloud, typically 10km above sea level, surrounded by super-cooled water droplets [9]. The negatively-charged graupel pellets are heavier and thus sink to bottom of the cloud, typically 6-8km above sea level.

As the electrostatic field within the thundercloud rapidly rise, an initial charge called the stepped leader breaks out from the cloud and seeks an ideal path to connect with an upward leader from the ground (Figure 2.2 at 1.2 ms) [6]. Around 90% of stepped leaders carry a negatively-charge [9].

Lightning attachment occurs when a downward leader from a cloud and an upward leader from the earth connect (Figure 2.2 at 20 ms to 20.20 ms). In a subsequent process termed, the first *return stroke*, several Coulombs of negative electric charge from the stepped leader then ferociously flows into the earth, resulting in the luminous stroke that we commonly observe (Figure 2.2 at 20.20 ms). The heat from stroke expands the surrounding air and propagates a shock wave that is called *thunder*.

Once the return stroke's current has stopped flowing, a lightning flash may either end or develop another stroke. If the lightning discharge ends, this is called a *single-stroke lightning flash*. In the other case, 80% of lightning flashes that lower negative cloud charge in temperate regions contain over one stroke, usually three to five strokes [9]. These additional lightning strokes occur if negative charge is available to the upper portion of the previous stroke channel. If a charge is available, a propagating dart leader travels down the previous return stroke channel. The dart leader thus sets up the subsequent return stroke (Figure 2.2 at 62.05 ms) [32], and again violently transfers negative charge to the earth.

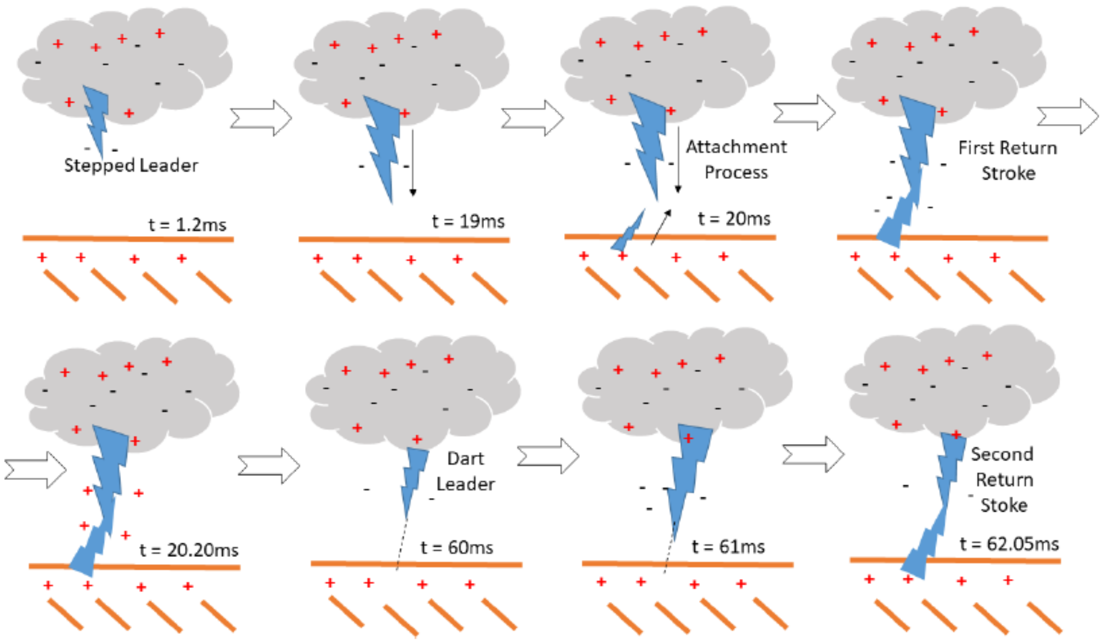

FIGURE 2.2: The lightning attachment process of a typical negative cloud to ground lightning flash. Timescale is in milliseconds. Redrawn and adapted from [9].

2.3 Review of Lightning Prediction

In this section, we start by discussing lightning prediction in general and then focus on lightning prediction for South Africa. We then will indicate gaps in knowledge.

2.3.1 Lightning Prediction Globally

Researchers have used various machine learning methods for lightning prediction over the last decade (Table 2.1). These tools include Support Vector Machine, Decision Trees, Regressions and Artificial Neural Networks. Datasets were typically historical lightning and/or weather parameters. In general, prediction accuracy remains high although accuracy is improving [11].

Of the machine learning methods evaluated, neural network models have performed best (Table 2.1). This is probably because neural networks are able to model complex non-linearity nature of lightning and weather events. A model used by Geng et al. (2019) called the Convolutional-LSTM neural network has shown especially good predictive power [14]. A similar model was used in this study.

2.3.2 Lightning Prediction for South Africa

In recent years, numerous studies have used machine learning tools for lightning prediction. Table 2.2 is a summary of Lightning prediction studies for South Africa. Booysens and Variri (2014) used decision tree and naive bayes methods to predict the likelihood of lightning from satellite data sources. The authors found a 97% spatial detection rate but only a 31% temporal detection rate. Gijben et al. (2017) also predicted the likelihood of lightning but used a stepwise logistic regression on numerical weather prediction data. The latest study by Essa et al. 2021 who have used a LSTM-FC Neural Network on an univariate dataset to predict lightning density using SALDN data only [11]. The authors predicted lightning with a mean absolute error accuracy of 2.87 flashes for an area of 770km^2 per hour.

TABLE 2.1: Summary of recent global lightning prediction studies. As adapted from [10].

Authors	Title	Dataset	Features	Method	Outcome
[14]	LightNet: A Dual Spatiotemporal Encoder Network Model for Lightning Prediction	Numerical Weather Prediction models and China National Lightning Detection Network	Numerical Weather Prediction variables and historical lightning data	Convolutional Long Short-Term Memory Neural Network with Spatio-Temporal dual encoders	Threefold improvement for six-hour predictions compared with three established forecast methods (PR92, LF1, LF2)
[15]	Numerical study of performance of two lightning prediction methods based on: Lightning Potential Index (LPI) and electric POTential difference (POT) over Tehran area	Lightning Imaging Sensor and Lightning Detection Network Data	Numerical Weather variables	Numerical Weather Prediction Model	LPI outperforms POT model
[25]	Forecasting lightning around the Korean Peninsula by postprocessing ECMWF data using SVMs and undersampling	Lightning Detection Networks and Weather station data	Weather variables: convective available potential energy, temperature, wind speed, relative humidity, K-index and the Showalter stability index	Support Vector Machines (outperforms) and Random Forests	Support Vector Machines and Random Forests had equitable threat scores of 0.0885 and 0.0828 respectively
[26]	Nowcasting lightning occurrence from commonly available meteorological parameters using machine learning techniques	Lightning Detection Networks and Weather station data	Surface air temperature, relative humidity, windspeed and surface air pressure	Decision Tree	80%+ Heidke Skill Score Score (Outperforms CAPE and Persistence Model)
[35]	NWP-based lightning prediction using flexible count data regression	NWP data and Lightning Detection Network	Numerical Weather variables and Historical Lightning data	Count-Data Regression	Outperforms a reference climatology method up to a forecast horizon of 5 days
[38]	Lightning Prediction Using Recurrent Neural Networks	Weather Station Data, Lightning Detection Network	-	Long Short-Term Memory (LSTM) Neural Networks	Max 84% accuracy

2.3.3 Significant Finding in the Related Work

We identify two gaps in knowledge. Firstly, studies have not yet used a spatio-temporal LSTM neural network for lightning prediction in South Africa. Secondly,

TABLE 2.2: Recent Southern African lightning prediction literature. As adapted from [10].

Authors	Title	Features	Method	Outcome
[4]	Detection of Lightning Pattern Changes Using Machine Learning Algorithms	Lightning data from Satellite Lightning Imaging Sensors	Various machine learning models	Naïve Bayes: a temporal 31% global pattern detection rate. Decision Tree: 97% spatial global pattern detection rate.
[10]	A LSTM Recurrent Neural Network for Lightning Flash Prediction within Southern Africa using Historical Time-series Data	Historical Lightning Data from SALDN	LSTM RNN Model	MAE of 2.87 flashes.hr^{-1} for one-hour predictions per ±770km^2
[11]	Short-term Prediction of Lightning in Southern Africa using Autoregressive Machine Learning Techniques	Historical Lightning Data from SALDN	ARIMA Model	AR and ARIMA, LSTM models MAPE of 15312, 15080 and 3705 flashes for three-hour predictions for South Africa
[17]	A statistical scheme to forecast the daily lightning threat over Southern Africa using the Unified Model	Relative humidity, Lifted index, Precipitable water, Equivalent temperature potential, Convective available potential energy, Average temperature	Stepwise Logistic Regression	Area Under Curve value of (AUC/ROC Curve) ±0.91

many studies did not quantitatively predict lightning frequency; most studies only predict the likelihood of lightning occurring.

2.4 Chapter Summary

This chapter began by describing how a quantitative lightning prediction model fits into a Smart City disaster management system. We then briefly discussed how thunderclouds form through the collision of graupel and ice crystals, and we also described the electrostatic interactions in a lightning flash. Next, we showed two knowledge gaps in current literature. These being that no previous study used spatio-temporal neural networks for lightning prediction in South Africa, and the lack of studies attempting to quantitatively predict lightning.

Chapter 3

Research Methodology

In this Chapter, we (1) give detail on LSTM neural network model variants used in this study, (2) discuss remote sending instruments and (3) give details on how this study was conducted.

3.1 LSTM Recurrent Neural Network Model Variants

RNN models are a category of artificial neural networks. Dissimilar to vanilla feed forward neural networks, RNN models have feedback connections between explanatory variables. This adaptation allows RNN models to be well suited for time series modeling. LSTM-RNN models additionally have an information filter system called a *cell* (Figure 3.1). A cell is composed of an input gate, an output gate and a forget gate. The cell functions to remember values over arbitrary time steps by regulating the flow of information by the three gates. LSTM RNN networks are thus well suited for classification, processing and prediction using time series data [30].

There are numerous varieties of LSTM RNN networks. Often, models are used in combination with other neural network architectures. Of interest to us are three models called the LSTM Fully-Connected (LSTM-FC) model, the Convolutional Neural Network LSTM (CNN-LSTM) model and the Convolutional-LSTM (ConvLSTM) model. The ConvLSTM model was chosen for its proven high accuracy in spatio-temporal weather prediction [14, 42]. The CNN-LSTM model was tested due to its suitability for spatio-temporal image prediction [5, 34]. And the LSTM-FC model was used to assess the effect that the convolutional operation specifically has on lightning prediction.

3.1.1 LSTM-FC

The LSTM-FC model has two neural network architectures placed one after the other. The data runs through a vanilla LSTM model and then is passed through a fully-connected artificial neural network. The latter dense network is required to process the LSTM data into the required output for analysis. LSTM-FC has the input, cell output and states as 1-dimensional vectors.

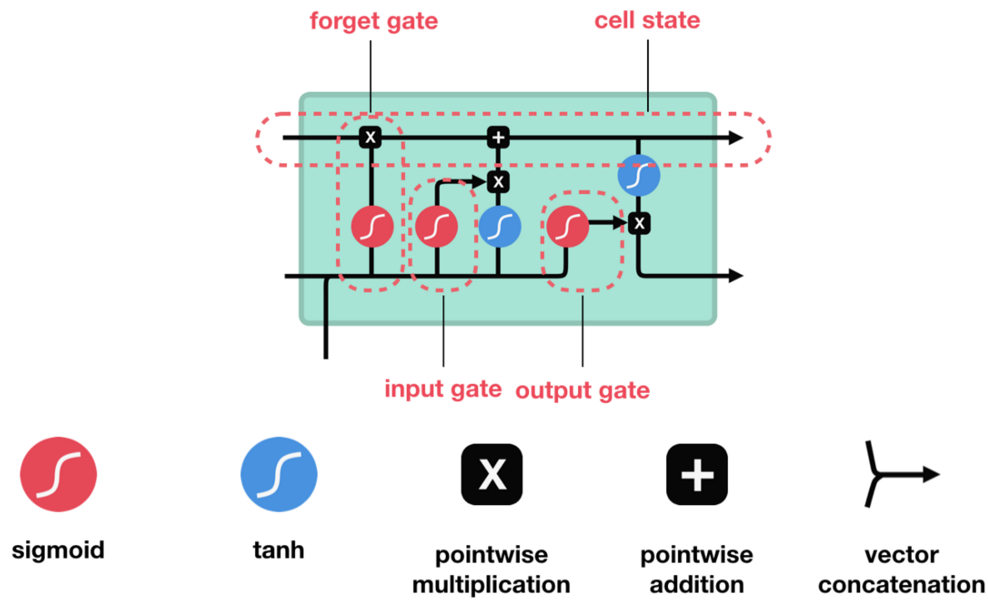

FIGURE 3.1: Schematic of LSTM-gates architecture. Data is from neighbouring time steps is filtered and processed by the forget, input and output gates. Reproduced from [31].

3.1.2 CNN-LSTM

The CNN-LSTM model is the combination of CNN and LSTM RNN models, placed one after the other. Convolutional Neural Networks (CNN) are well-adapted for image processing and recognition. It works by learning matrix-shaped filter(s) that convolve through an image to recognize features (Figure 3.2). For CNN-LSTM models, the data is then passed onto an LSTM neural network for further processing

(Figure 3.3). This makes CNN-LSTM especially well suited for image predictions in a series, e.g. video and time-series geomapping [40].

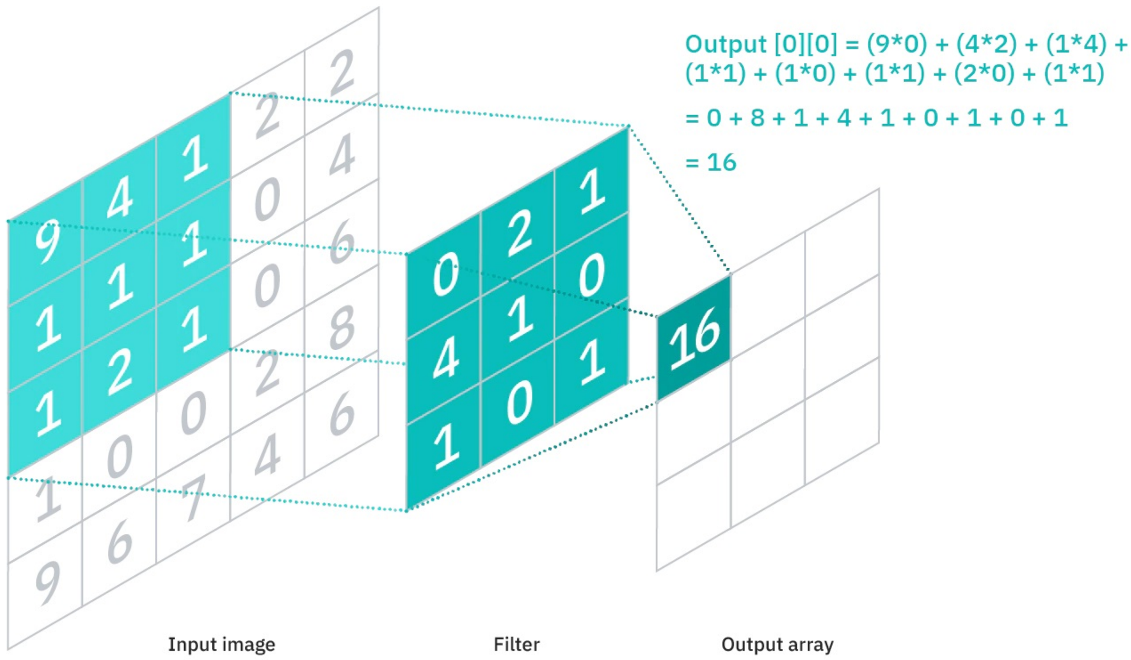

FIGURE 3.2: Schematic of CNN input structure and processing function. Reproduced from [21]

3.1.3 ConvLSTM

The ConvLSTM model follows on the same intention as the CNN-LSTM to predict spatio-temporal data. The difference, though, is in how the convolutional processing is implemented. The distinguishing feature of the ConvLSTM is that all cell inputs, cell outputs, input gates, output gates and forget gates of the ConvLSTM are 3-dimensional tensors (1-dimensional vectors for LSTM-FC) whose last two dimensions are spatial dimensions (Figure 3.2 and Figure 3.3) [42]. By doing so, it captures underlying spatial features by convolution operations in multiple-dimensions. ConvLSTM models have been used to forecast a variety of weather features, including precipitation and lightning [14, 39].

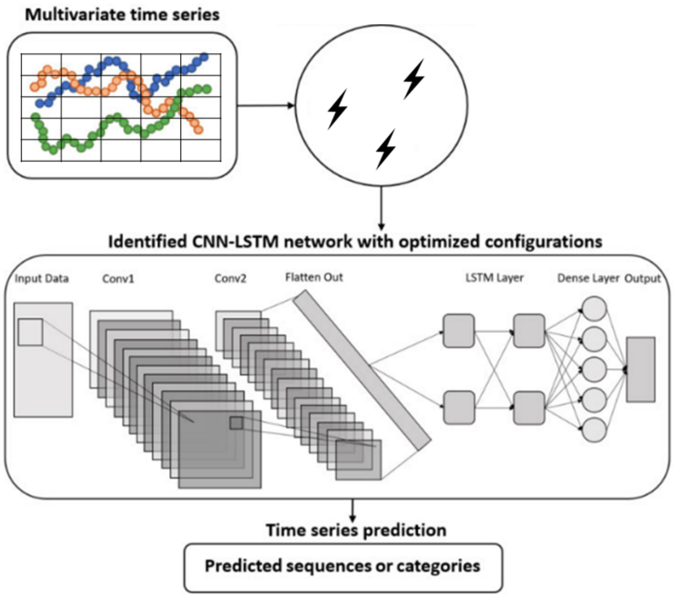

FIGURE 3.3: Schematic of the CNN-LSTM model architecture. Data is fed into a CNN network which is then passed onto a LSTM neural network. Adapted from [41].

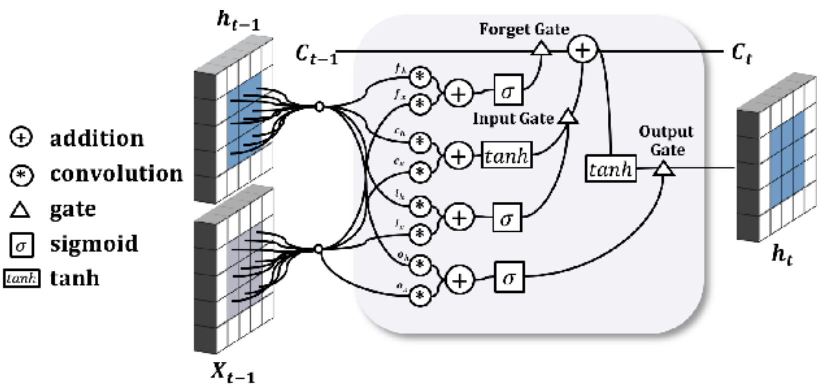

FIGURE 3.4: Schematic of ConvLSTM model architecture. Data is passed into a LSTM network that has convolutional operations for its input gate, output gate and forget gate processing. Reproduced from [43].

3.2 Data from Remote Sensing Instruments

Remote sensing instruments have the benefit of automation and location relevance. The remote sensing instruments used in this study consist of Automatic Weather Stations from the SAWS, and the South African Lightning Detection Network.

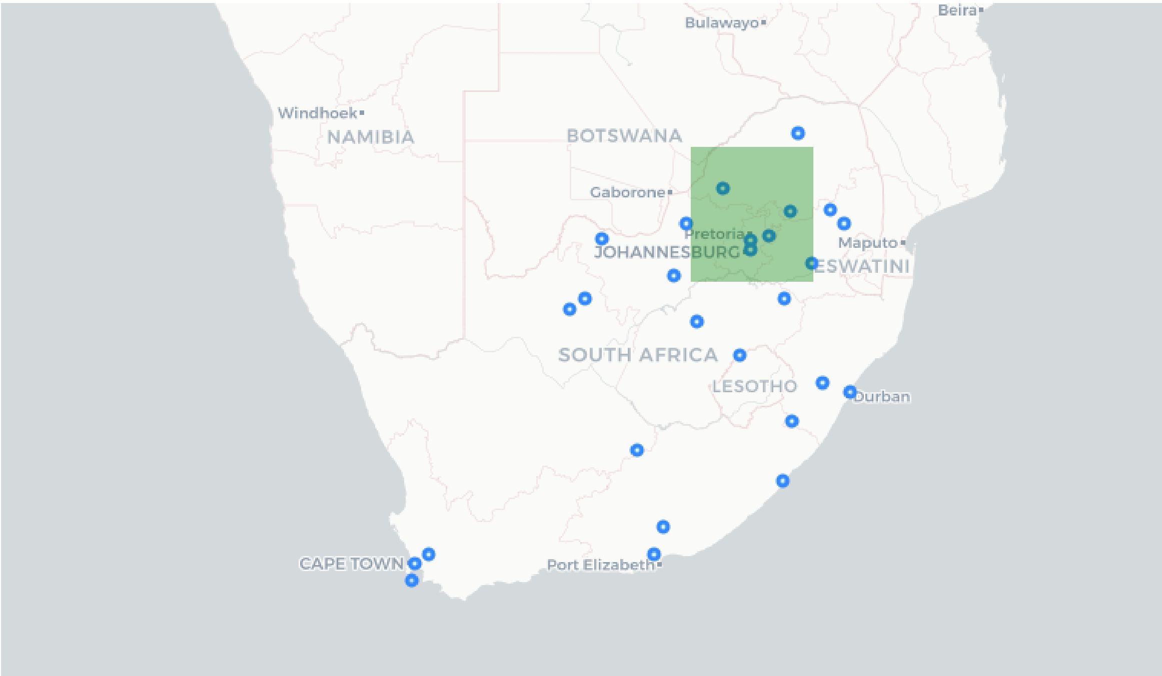

FIGURE 3.5: Location of certain SAWS stations in South Africa. The green shaded area represents the area that this study forecasted lightning within.

3.2.1 SAWS Automatic Weather Stations

The South African Weather Service (SAWS) is a public controlled entity overseen by the Ministry of Forestry, Fisheries and the Environment in South Africa [8]. One of SAWS's role is to maintain and run a quality control process meteorological and climatological data in South Africa. These variables include, *inter-alia*, rainfall intensity, surface temperature and humidity - all which contribute to the build up of a thundercloud. The variables are measured through a network of 231 Automatic Weather Stations around South Africa (Figure 3.5) [37]. These features are represented in a time series graph in Figure 3.6 and in a correlation matrix in Figure 3.7.

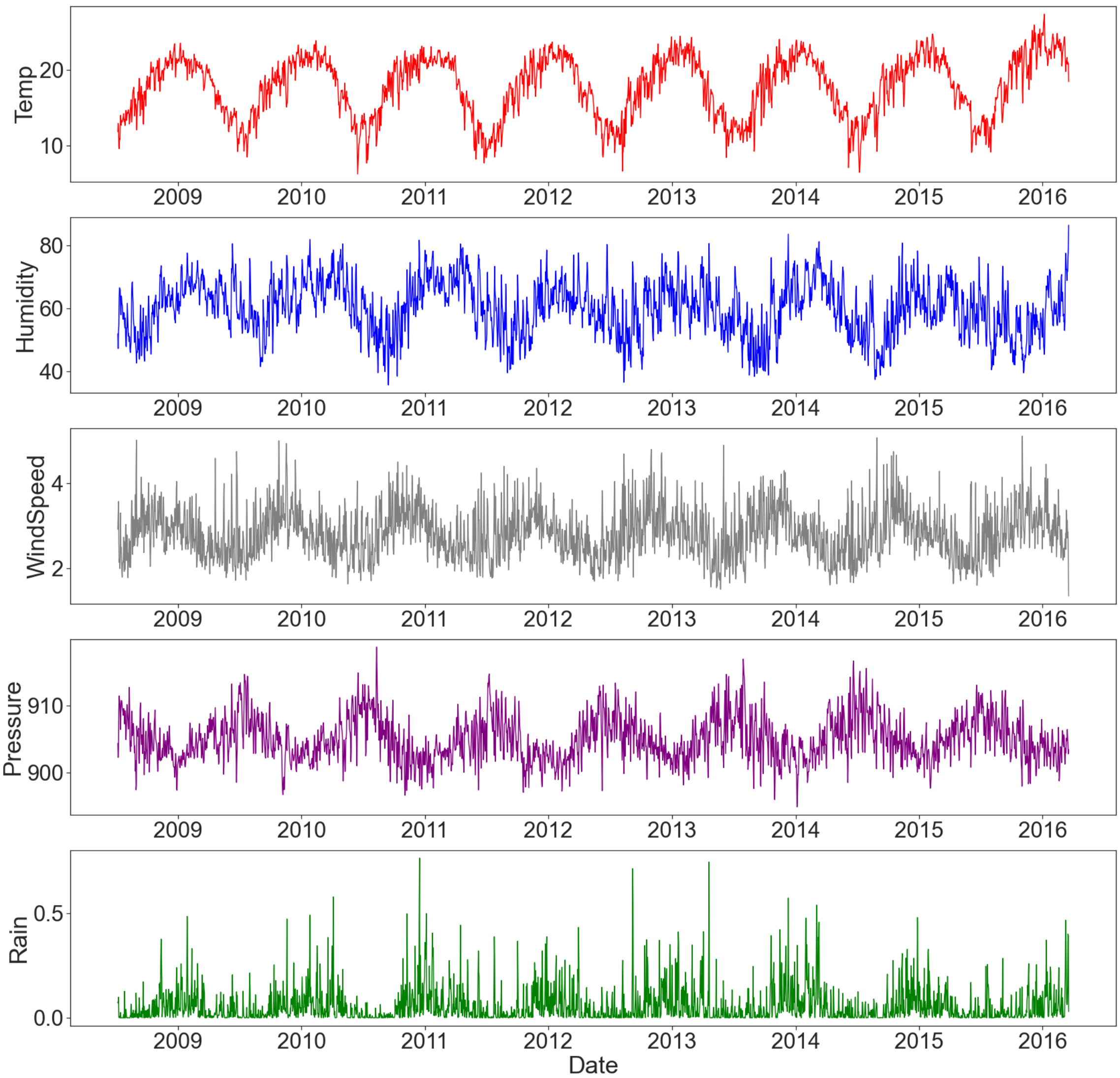

FIGURE 3.6: Time series of SAWS weather data. Average readings for all Automatic Weather Stations locations as shown in Figure 3.5.

We note seasonality for all variables, a negative correlation between pressure and temperature features and a positive correlation between humidity and rain features.

3.2.2 The South African Lightning Detection Network

The South African Lightning Detection Network (SALDN) is the most accurate instrument available for lightning detection and geolocation in South Africa [13]. As

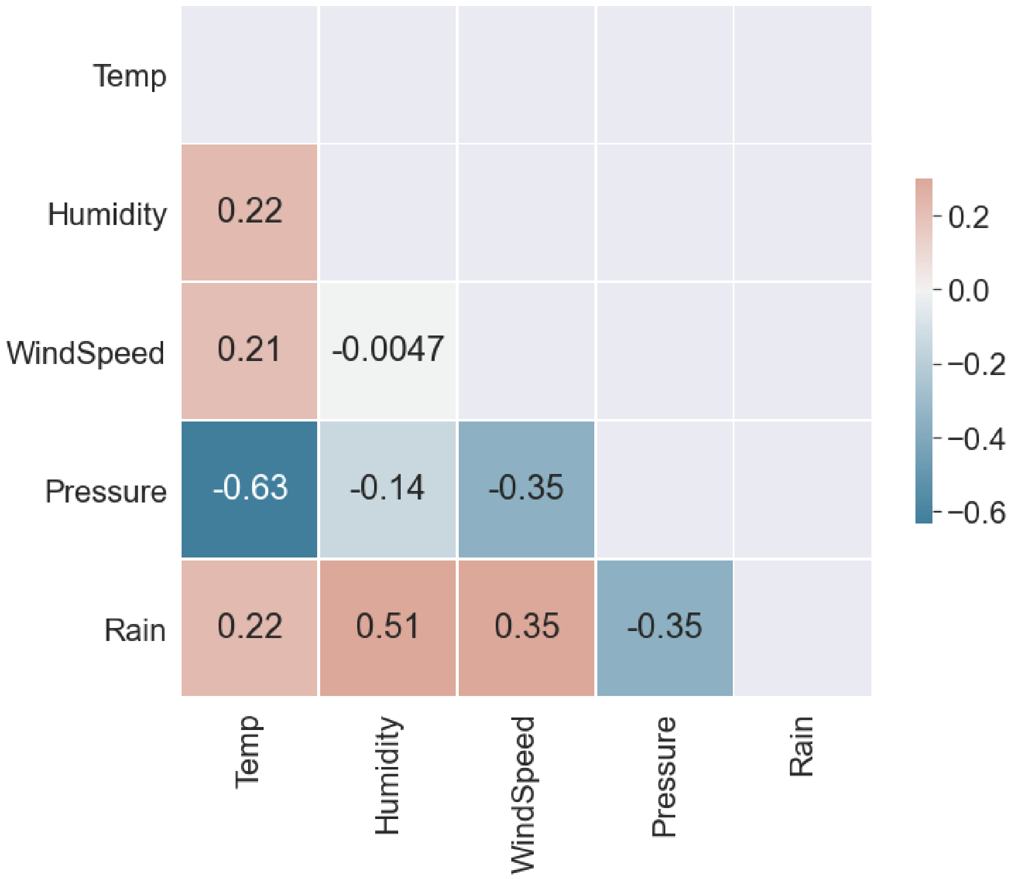

FIGURE 3.7: Correlogram (Pearson's correlation coefficient) between weather features from SAWS. Average readings for all Automatic Weather Stations locations as shown in Figure 3.5.

of 2015, it was composed of 24 Vaisala LS7000/01 series sensors that are spaced across South Africa [20, 12]. Currently, there are 25 Sensors with 18 LS7000 series, 4 LS7001 and 3 LS7002 sensors, as shown in Figure 3.8 [16]. Lightning detection networks work by detecting and geo-locating individual strokes which are then grouped into flashes [20]. The SALDN has an estimated detection efficiency of around 90% [17, 13]. Importantly, the SALDN network is designed to detect CG lightning events thus the network detects only a portion of total lightning [17]. Figure 1.1 shows a heatmap of lightning density for South Africa during 2015 using the SALDN dataset.

The lightning detection sensors relies on Magnetic Direction Finder (MDF) and/or

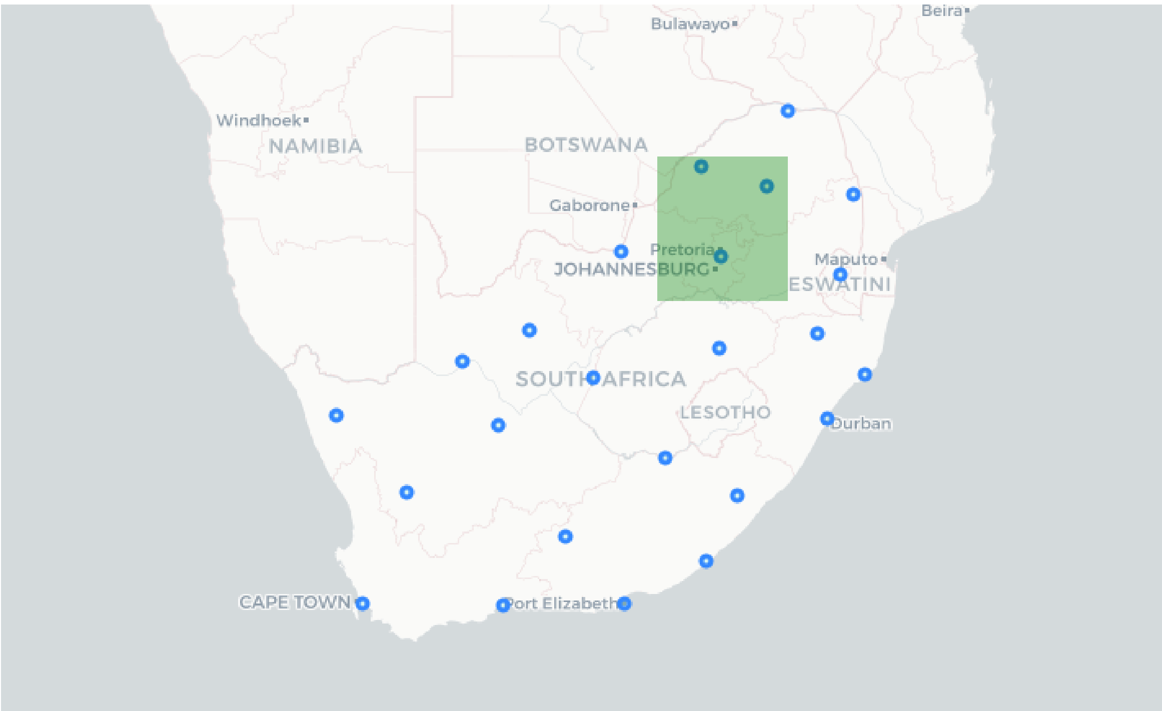

FIGURE 3.8: Locations of SALDN Sensors. The green shaded area represents the area that this study forecasted lightning within.

Time Of Arrival (TOA) methods for lightning detection [7]. The MDF method measures changes in magnetic field caused from a lightning stroke using two vertical, orthogonal wire loop antennas [28]. The plane of each loop are oriented at the North/South and East/West directions respectively [2]. The radiated field of a lightning flash induces an electric current in the loops and by comparing the voltage signals between the two loops, a direction to the flash may be determined. The TOA sensors records the time of arrival of radio-waves derived from lightning strokes. Together with other TOA sensors, the direction and location of the source may be triangulated [28].

3.3 Methods

This section will detail the protocol and analysis metrics used for this study.

3.3.1 Input Datasets and Pre-Processing

Lightning flashes. Data was obtained through the SALDN network for the period between 07 July 2008 to 17 March 2016, representing in excess of 140x10^6 lightning flashes. For each prediction, we input hourly lightning flash data for 240 previous time-steps, and also 240 previous time-steps for same period of the previous year. Spatially, we input lightning flashes in a grid-like format from the area with coordinates (23°S, 26°E) & (27.5°S, 26°E) & (23°S, 30.5E) & (27.5°S, 30.5°E).

Weather variables. Data was sourced from the SAWS for the period between 07 July 2008 to 17 March 2016. Variables included wind speed, humidity, surface temperature, pressure and rainfall intensity. We did input hourly weather data for 240 time-steps prior to the prediction time-step for automatic weather stations that fell within the same area as the lightning-flash data. The weather stations used were:

Scaling. All data was scaled. For the CNN-LSTM and LSTM-FC models, data was scaled according to the quartile ranges (1st and 3rd quartile, Robust Scalar SKLearn Version 0.23.2). This managers outliers. For the ConvLSTM model, we used a 0 to 1 scalar (MinMaxScaler, SKLearn Version 0.23.2). Missing values in the SAWS dataset were imputed with the median value for the respective area.

3.3.2 Model Implementation

Models were coded in Python language with Jupyter notebook server 6.1.4 (Python version 3.8.5., Tensorflow version 2.4.1. and Keras version 2.4.3). Models were run on AmazonSageMaker with notebook instance ml.m5.4xlarge. We used keras LSTM, keras ConvLSTM2D and keras Conv2D for LSTM-FC, ConvLSTM and CNN-LSTM models respectively. The neural networks and parameters were optimized with validation data using a gird search algorithm. Models were initially biased for higher complexity and complexity was then reduced during optimization. The hyperparamaters are noted below.

LSTM-FC. The LSTM-FC model error was minimized based on Mean Squared Error with Adam as the optimizer [22]. Epochs were 30. The model consisted of 3 x Tensorflow keras LSTM layers, 1 x dense layer:

- Layer 1: units=100, return sequences=True, dropout=0.2.
- Layer 2: units=50, return sequences=True.
- Layer 3: units=50.
- Layer 4: dense later, of 49 units, activation was Leaky-Rectified Linear Unit (alpha=0.15).

CNN-LSTM. The CNN-LSTM model error was minimized on Mean Squared Error, with Adam as the optimizer. Epochs were 3. The model consisted of 2 x Tensorflow keras Conv2D layers, 1 keras LSTM layer and 1 dense layer consisting of:

- Layer 1: filters=20, kernel size=3, padding=same.
- Layer 2: filters=20, kernel size=3, padding=same.
- Layer 3: LSTM layer, input=64.
- Layer 4: dense layer of 49 units, activation was Leaky-Rectified Linear Unit with (alpha=0.05).

ConvLSTM. The ConvLSTM model error was minimized on Mean Squared Error, with Adam as the optimizer. Epochs was between 3 to 5. The model consisted of 2 x Tensorflow keras ConvLSTM2D layers, 2 x dense layers consisting of:

- Layer 1: filters=20, kernel size=(3,3), return sequences=True.
- Layer 2: filters=20, kernel size=(3,3), return sequences=True.
- Layer 3: dense layer of 49 units, activation was Leaky-Rectified Linear Unit with (alpha=0.05).
- Layer 4: dense layer of 49 units, activation was Rectified Linear Unit.

3.3.3 Model Runs

All Models were run for input features as (1) SAWS data only, (2) SALDN data only and (3) SAWS and SALDN data as features; modelling were repeated 3 to 5 times. Models were trained on data between 07 July 2008 and 05 May 2013. Models were tested on data from 05 May 2013 to 18 March 2016. This represents a 70/30% train/test split.

3.3.4 Outputs

We predicted lightning flash frequency for an area between 26.5°E to 30°E longitude and 23.5°S to 27°S latitude every hour for test period. This is represented as the shaded area in Figure 3.5. This area was chosen due to the relatively high lightning frequency and relative abundance of automatic weather stations.

Negative predicted lightning flash values were adjusted to 0.

3.4 Performance Metric

For this study, model performance was evaluated based on Mean Absolute Error (MAE) metric (Equation 3.1) and the model loss function was trained to minimize Mean Squared Error (MSE) (Equation 3.2).

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{3.1}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2 \tag{3.2}$$

Let n = the number of data points, y_i is prediction value, and x_i is actual value

3.5 Python Code

We used Python coding language due to its versatility within the machine learning field. The LSTM code may be accessed at: https://github.com/YaseenZA/LightningPrediction

3.6 Chapter Summary

In Chapter 3, we explained how LSTM neural networks are suited for time-series forecasting and we also explained how the convolution operation extracts spatial features. We then introduced the remote sensing instruments that were used for data collection in this study. The final section detailed the methods used in this study to aid reproducibility.

Chapter 4

Results and Discussion

In this study, we evaluated the use of LSTM model variants with remote sensing data on lightning frequency prediction. In this chapter, we display results of this experiment and also discuss the context of these results.

4.1 Evaluation of Input Features

We assessed model performance based on input datasets and display the results in Table 4.1, Table 4.2 and Table 4.3. Here, we find that when input features are from the SALDN dataset only, the LSTM models had less predictive error then when input was SAWS only or SALDN+SAWS dataset. When all LSTM variant model predictions are combined, SALDN only as the input dataset produced the least error with a combined MAE of 59 flashes.hr^{-1} (Table 4.4). Input datasets as SAWS only and SALDN+SAWS had MAE (flashes.hr^{-1}) values of 74 an 70 respectively. Standard deviations were relatively high at 298, 333 and 318 for SALDN, SAWS and SALDN+SAWS datasets respectively (Table 4.4).

All of the LSTM model variants individually also predicted the least error with the SALDN only dataset, followed by either SAWS only or SALDN+SAWS datasets, as illustrated in Figure 4.1, Figure 4.2 and Figure 4.3. These figures show cumulative error for each input dataset.

Discussion: Evaluation between input features. Models with only the SALDN dataset as the input did outperform in terms of MAE compared to SAWS data and SALDN+SAWS datasets. We speculate that the SAWS input data did not increase predictive power because the introduced noise counteracted the possibility of formulating a more complex weather model. To clarify by using an example: although

TABLE 4.1: Lightning frequency prediction error values for LSTM-FC model based on input datasets. Units are flashes.hr^{-1}.

Statistic	SALDN only	SAWS only	SALDN+SAWS
Mean	60	73	67
Median	6	25	18
Standard Deviation	262	325	335

TABLE 4.2: Lightning frequency prediction error values for CNN-LSTM model based on input datasets. Units are flashes.hr^{-1}.

Statistic	SALDN only	SAWS only	SALDN+SAWS
Mean	45	58	50
Median	0	8	10
Standard Deviation	295	340	286

we may require hot temperatures as a prerequisite for thundercloud formation, only a small percentage of hot days will lead to thunderclouds.

Another explanation could be the lack of atmospheric weather data, as opposed to the surface weather data used in this study. Thundercloud formation requires three ingredients: atmospheric instability, atmospheric moisture and a trigger mechanism. Surface weather data is unlikely to capture the required atmospheric instability and moisture profiles required for thunderstorm formation. Heat on the other is only one of several possible trigger mechanisms required for thundercloud development, where other sources such as topography and moisture boundaries to name a few could also trigger convection. Atmospheric weather data may therefore be better suited for lightning prediction, and static inputs such as topography could also be considered in the future.

We have done further analysis to rule out some other possibilities (unpublished). We do not believe that the model(s) were not trained enough. We have trained

TABLE 4.3: Lightning frequency prediction error values for ConvLSTM model based on input datasets. Units are flashes.hr^{-1}.

Statistic	SALDN only	SAWS only	SALDN+SAWS
Mean	72	91	93
Median	23	46	58
Standard Deviation	336	332	331

TABLE 4.4: Average lightning frequency prediction error for all LSTM model variants for each input dataset.

Statistic	SALDN only	SAWS only	SALDN+SAWS
MAE (flashes.hr^{-1})	59	74	70
Standard Deviation	298	333	318

models with many more network layers to test for additional "intelligence."

The lack of continuity of SAWS weather data may explain the disparity between the CNN-LSTM and ConvLSTM models performance [36], and also provides guidance for future research. Our weather data was not available uniformly; thus making it difficult for features to become apparent. The CNN-LSTM model may be better suited as it relies less on the spatial-dependency of the model. Also, the smoother cumulative prediction may indicate an inability for the ConvLSTM to detect features. A possible solution would be to use satellite weather information instead, although there will be a trade-off of weather feature accuracy [29]. An additional benefit of using satellite data is that the neural network models may detect Cumulonimbus cloud structure.

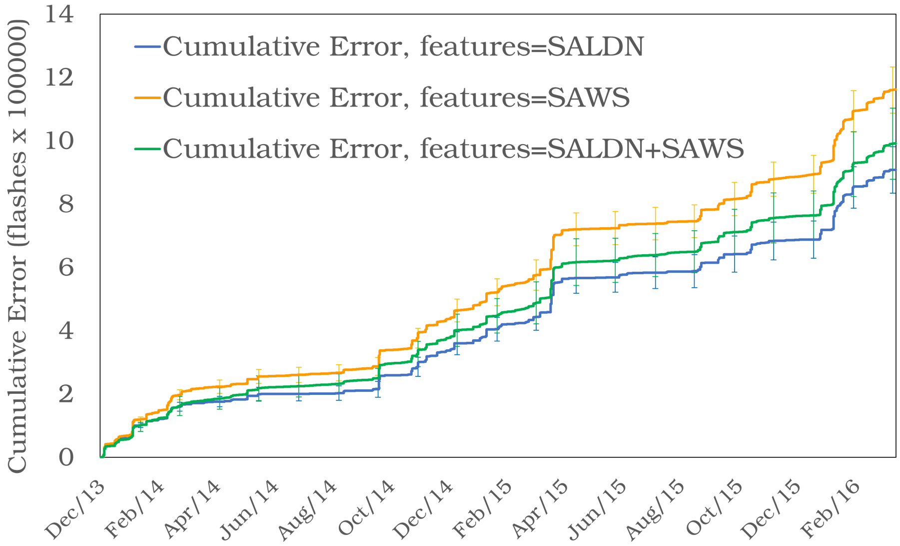

FIGURE 4.1: Cumulative lightning frequency prediction error (Mean Absolute Error) of CNN-LSTM model for each input dataset for test period. Error bars indicate standard error of the mean.

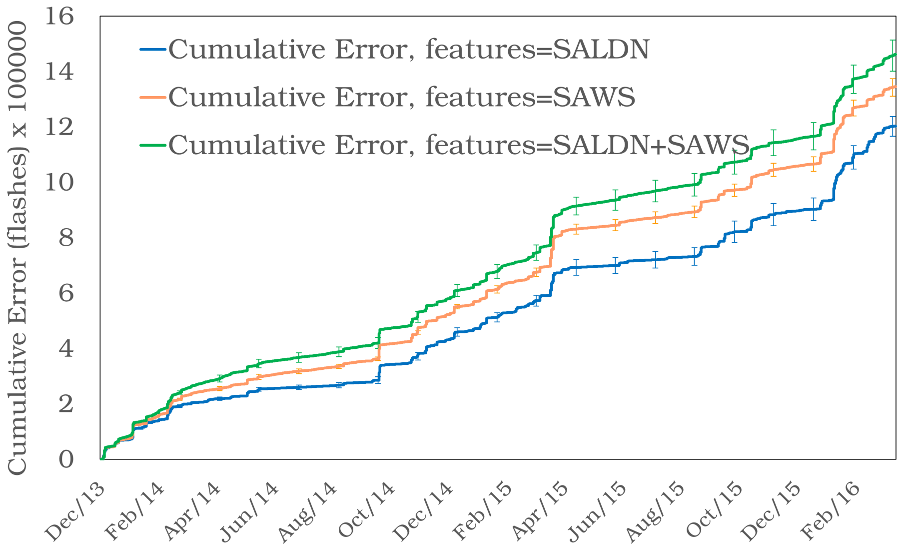

FIGURE 4.2: Cumulative lightning frequency prediction error (Mean Absolute Error) of LSTM-FC model for each input dataset for test period. Error bars indicate standard error of the mean.

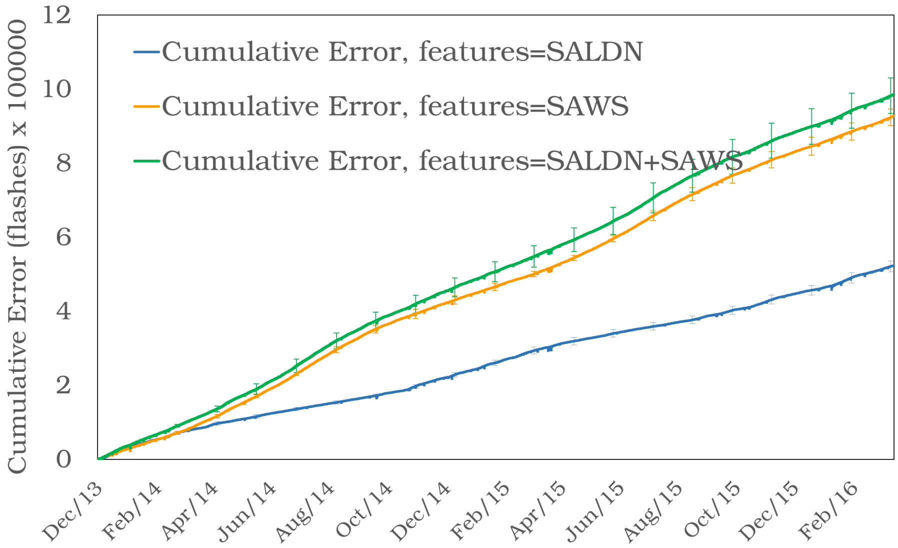

FIGURE 4.3: Cumulative lightning frequency prediction error (Mean Absolute Error) of ConvLSTM model for each input dataset for test period. Error bars indicate standard error of the mean.

4.2 Evaluation of LSTM Model Variants

The CNN-LSTM model outperformed LSTM-FC and ConvLSTM models for prediction accuracy. In Table 4.5, we see that that combined MAE was 51 flashes.hr^{-1} for the CNN-LSTM model, compared to 67 and 86 for the LSTM-FC and ConvLSTM models respectively. The CNN-LSTM model also outperformed LSTM-FC and ConvLSTM models for each input dataset (Table 4.1, Table 4.2, and Table 4.3). The CNN-LSTM model produced a MAE value of 45 flashes.hr^{-1} when features were SALDN dataset only.

Discussion: Evaluation between LSTM Models. The evidence that CNN-LSTM models outperforms LSTM-FC is substantial in weather predictions [44, 5]. The research on ConvLSTM performance is scant though. We note that the ConvLSTM model was less able to adjust for short-term weather changes, resulting in smoother graph predictions (Figure 4.4). Further, the kernel density estimation plot in Figure 4.7 for ConvLSTM indicates bias in error prediction as errors are not centered displays distribution of errors through a Kernel Density Estimation method.

TABLE 4.5: Average lightning frequency prediction error for all input datasets for each LSTM model variants.

Statistic	CNN-LSTM	LSTM-FC	ConvLSTM
MAE (flashes.hr^{-1})	51	67	86
Standard Deviation	307	308	333

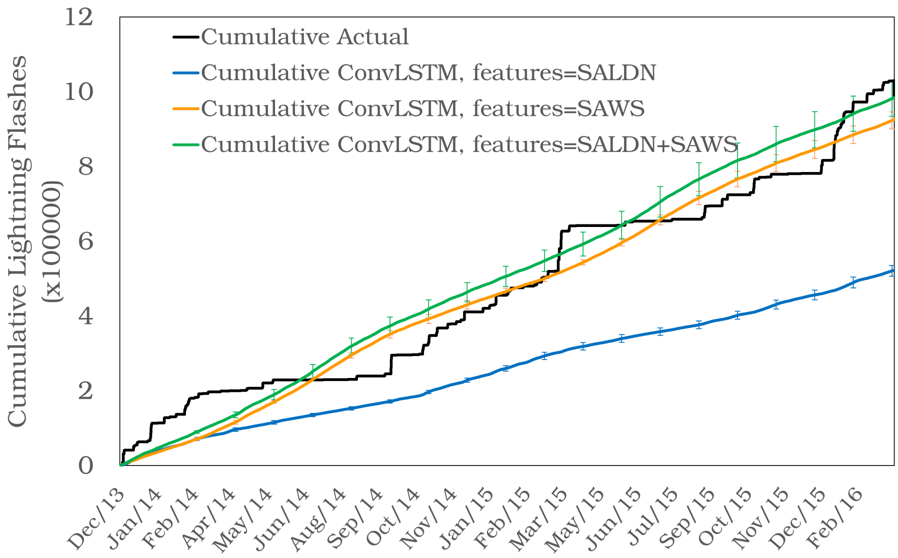

FIGURE 4.4: Cumulative lightning frequency prediction by ConvLSTM model during test period. Error bars indicate standard error of the mean.

4.3 Prediction Characteristics

We note that the ConvLSTM lightning frequency predictions were less responsive to input data compared to other LSTM model variants, as visualised in Figure 4.4. The the variability was smoother for the ConvLSTM model for all inputs compared to LSTM-FC and ConvLSTM Models (Figure 4.5 and Figure 4.6).

Distributions of prediction errors of all the LSTM model variants indicate that

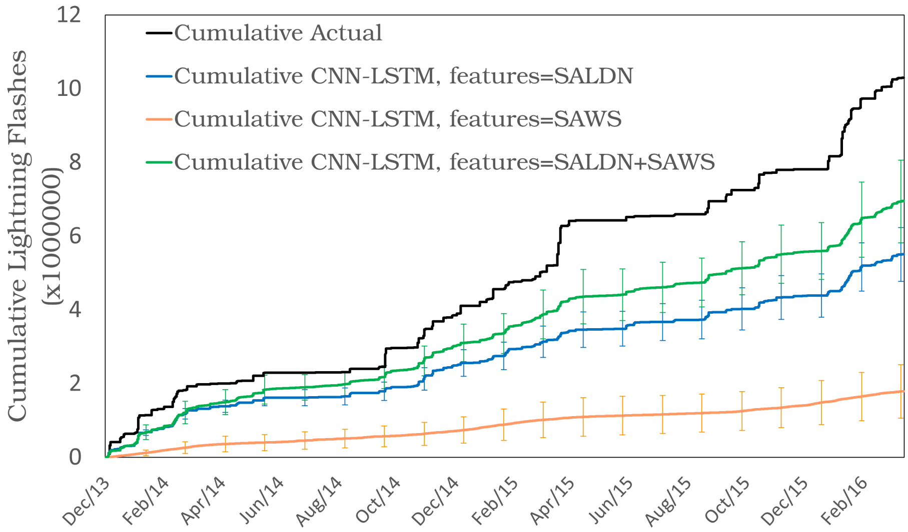

FIGURE 4.5: Cumulative lightning frequency prediction by CNN-LSTM model during test period. Error bars indicate standard error of the mean.

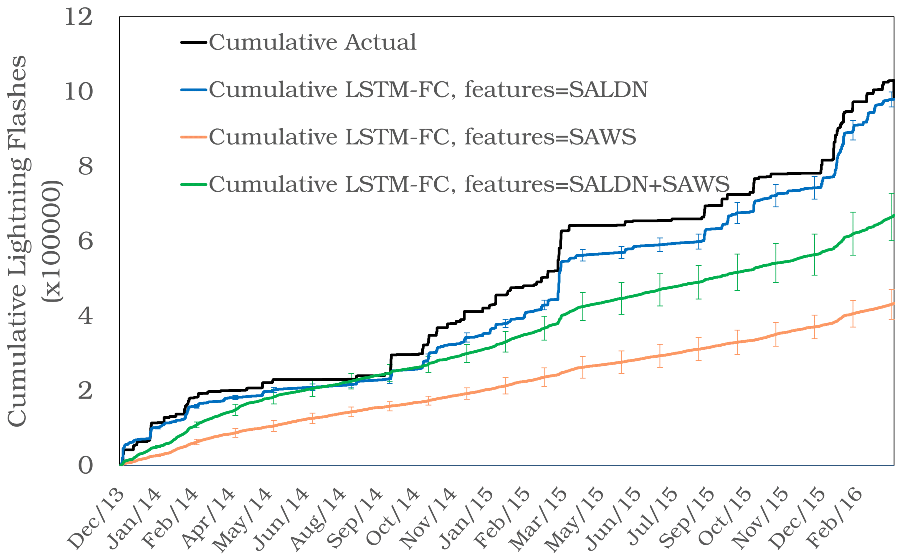

FIGURE 4.6: Cumulative lightning frequency prediction by LSTM-FC model during test period. Error bars indicate standard error of the mean.

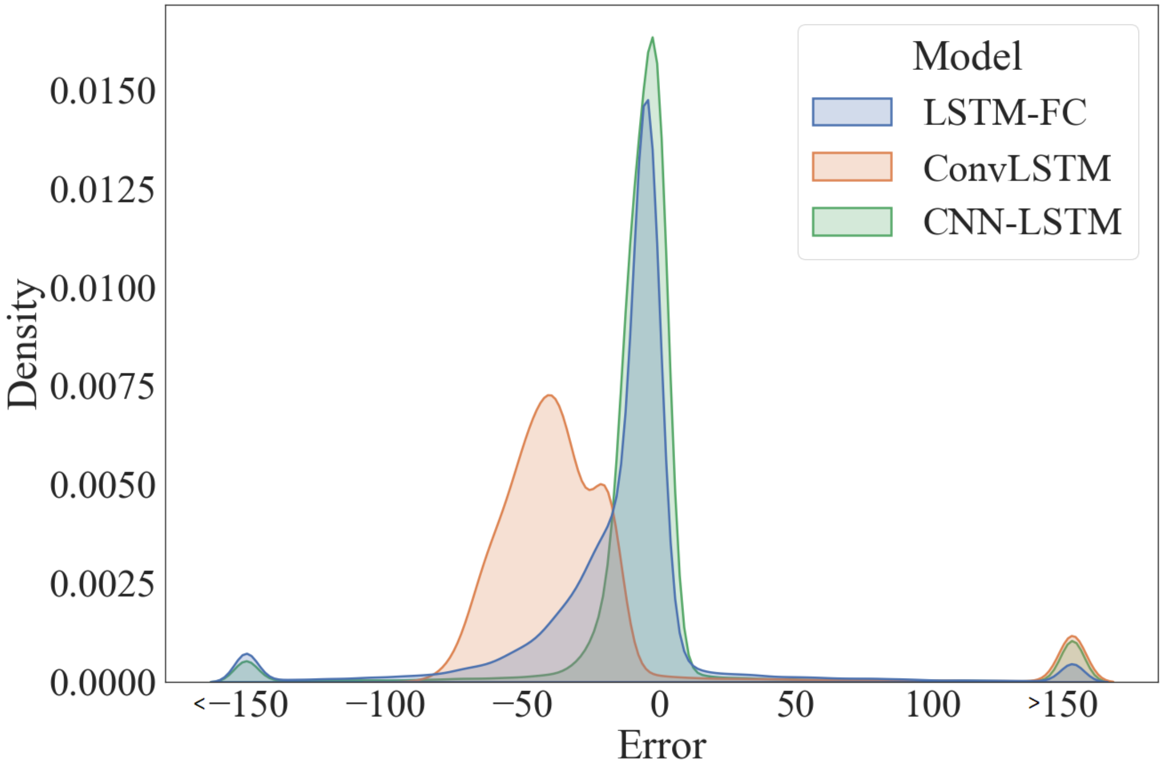

FIGURE 4.7: Kernel density estimate plot showing lightning frequency error distribution for each LSTM model variant. Error units are flashes.hr^{-1}.

the models predicted most points accurately; but there were a number of heavy weighted outliers, as indicated in the kernel density plot in Figure 4.7. ConvLSTM has a prediction bias while LSTM-FC and CNN-LSTM distribution was centered close to zero.

4.4 Chapter Summary

Our results indicate two main learning points. Firstly, we find that the CNN-LSTM model provides the best accuracy as assessed by MAE (Table 4.5, Figure 4.1). Secondly, models with SALDN only as input data outperformed SAWS only and SALDN+SAWS as input datasets (Table 4.2, Table 4.5).

Chapter 5

Conclusions and Future Work

This study sought to investigate the prediction accuracy of the LSTM-FC, CNN-LSTM and ConvLSTM model variants to predict short-term (1hr) cloud-to-ground lightning frequency within South Africa. Further, this study sought to assess prediction accuracy of these LSTM RNN model variants for different input datasets, namely the SALDN dataset only, SAWS dataset only and SALDN+SAWS datasets.

We have found that the CNN-LSTM (MAE=51 flashes.hr^{-1}) outperforms LSTM-FC and ConvLSTM for hourly lightning frequency predictions in South Africa (LSTM-FC MAE=67; ConvLSTM MAE=86). This is because the CNN-LSTM model is able to recognize spatio-temporal features which assist in prediction accuracy.

We also discovered that using SALDN only as the input dataset (MAE=59) did outperform using both SAWS only and SALDN+SAWS as input datasets (SAWS MAE=74; SALDN+SAWS MAE=70). This study has given direction for the implementation of spatio-temporal neural networks for the management of thunderstorm severity. We hope future work will build on this.

We advise future work should focus on data quality and continuity.

- Future research studies could add more feature variables, *inter alia,* atmospheric weather data. This may come from automatic weather stations, but can also come from space satellite sources. The advantage of satellite sources would be strong continuity of data, which would be more suitable for spatio-temporal models. Additionally, satellite data will be cost-effective to use.

- Satellite imagery may also be used with image recognition neural networks to identify cumulonimbus cloud structures. This may involve building a third CNN but the addition will most likely be beneficial.

Bibliography

[1] M. L. Akinyemi et al. "Lightning a Fundamental of Atmospheric Electricity". In: *IERI Procedia* 9 (2014). International Conference on Environment Systems Science and Engineering (ESSE 2014), pp. 47–52. ISSN: 2212-6678. DOI: https://doi.org/10.1016/j.ieri.2014.09.039. URL: http://www.sciencedirect.com/science/article/pii/S2212667814000872.

[2] A. Alammari et al. "Lightning mapping: Techniques, challenges, and opportunities". In: *IEEE Access* 8 (2020), pp. 190064–190082.

[3] T. B. Andersen and C. J. Dalgaard. "Power outages and economic growth in Africa". In: *Energy Economics* 38 (2013), pp. 19–23. ISSN: 0140-9883. DOI: https://doi.org/10.1016/j.eneco.2013.02.016. URL: http://www.sciencedirect.com/science/article/pii/S014098831300040.

[4] A. Booysens and S. Viriri. "Detection of Lightning Pattern Changes Using Machine Learning Algorithms". In: *Proceedings of the 2014 International Conference on Communications, Signal Processing and Computers*. 2014, pp. 78–84. URL: http://inase.org/library/2014/interlaken/bypaper/COSICO/COSICO-11.pdf.

[5] R. Chen et al. "A hybrid CNN-LSTM model for typhoon formation forecasting". In: *Geoinformatica* 23.3 (2019), pp. 375–396.

[6] G. V. Cooray. *The lightning flash*. 34. Iet, 2003.

[7] K. L. Cummins et al. "A combined TOA/MDF technology upgrade of the US National Lightning Detection Network". In: *Journal of Geophysical Research: Atmospheres* 103.D8 (1998), pp. 9035–9044.

[8] Department of Forestry, Fisheries and the Environment, South Africa. *South African Weather Service*, 2021. URL: https://www.environment.gov.za/statutorybodies/saws, (visited on 08/30/2021).

[9] J. R. Dwyer and M. A. Uman. "The physics of lightning". In: *Physics Reports* 534.4 (Jan. 2014), pp. 147–241. ISSN: 0370-1573. DOI: 10.1016/J.PHYSREP.2013.09.004. URL: https://www.sciencedirect.com/science/article/abs/pii/S037015731300375X.

[10] Y. Essa, R. Ajoodha, and H. G. P. Hunt. "A LSTM Recurrent Neural Network for Lightning Flash Prediction within Southern Africa using Historical Time-series Data". In: *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*. IEEE. 2020, pp. 1–6.

[11] Y. Essa, H. G. P. Hunt, and R. Ajoodha. "Short-term Prediction of Lightning in Southern Africa using Autoregressive Machine Learning Techniques". In: *2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*. IEEE. 2021, pp. 1–5.

[12] C. Evert and M. Gijben. "Official South African Lightning Ground Flash Density Map 2006 to 2017". In: *Inaugural Earthing Africa Symposium and Exhibition, Thaba Eco Hotel, Johannesburg*. 2017, pp. 5–9. URL: https://www.researchgate.net/profile/Richard_Evert/publication/321162358_Official_South_African_Lightning_Ground_Flash_Density_Map_2006_to_2017/links/5aa64f0aa6fdccd544bc90f9/Official-South-African-Lightning-Ground-Flash-Density-Map-2006-to-2017.pdf.

[13] H. G. Fensham et al. "Performance evaluation of the SALDN using highspeed camera footage of ground truth lightning events over Johannesburg, South Africa". In: *2018 34th International Conference on Lightning Protection (ICLP)*. 2018, pp. 1–5. DOI: 10.1109/ICLP.2018.8503415.

[14] Y. Geng et al. "LightNet: A Dual Spatiotemporal Encoder Network Model for Lightning Prediction". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. July 2019, pp. 2439–2447. ISBN: 978-1-4503-6201-6. DOI: 10.1145/3292500.3330717.

[15] M. Gharaylou et al. "Numerical study of performance of two lightning prediction methods based on: Lightning Potential Index (LPI) and electric POTential difference (POT) over Tehran area". In: *Journal of Atmospheric and Solar-Terrestrial Physics* 193 (2019), p. 105067. ISSN: 1364-6826. DOI: https://doi.org/10.1016/j.jastp.2019.105067. URL: http://www.sciencedirect.com/science/article/pii/S1364682619301026.

[16] M. Gijben. *MTG-LI Mission Advisory Group Meeting Presentation,* South African Weather Service. 2021. URL: https://www-cdn.eumetsat.int/files/2021-03/The\%20South\%20African\%20Lightning\%20Detection\%20Network.pdf (visited on 08/30/2021).

[17] M. Gijben, L. L. Dyson, and M. T. Loots. "A statistical scheme to forecast the daily lightning threat over southern Africa using the Unified Model". In: *Atmospheric Research* 194 (2017), pp. 78–88. ISSN: 0169-8095. DOI: 10.1016/J.ATMOSRES.2017.04.022. URL: https://www.sciencedirect.com/science/article/pii/S0169809516306354.

[18] R. L. Holle and M. A. Cooper. "Lightning Fatalities in Africa From 2010-2017". In: *2018 34th International Conference on Lightning Protection (ICLP)*. 2018, pp. 1–4. DOI: 10.1109/ICLP.2018.8503315.

[19] R. L. Holle et al. "Insured lightning-caused property damage in three western states". In: *Journal of Applied Meteorology and Climatology* 35.8 (1996), pp. 1344–1351.

[20] H. G. P. Hunt et al. "A multidisciplinary forensic analysis of two lightning deaths observed in South Africa". In: *International Journal of Disaster Risk Reduction* 51 (2020), p. 101814.

[21] IBM Cloud Education. *Convolutional Neural Networks*. 2020. URL: https://www.ibm.com/cloud/learn/convolutional-neural-networks (visited on 02/13/2022).

[22] D. Kingma and J. Ba. *Adam: A Method for Stochastic Optimization*. Dec. 2014. URL: https://arxiv.org/pdf/1412.6980.pdf.

[23] B. Maseko, G. Feig, and R. Burger. "Estimating lightning NOx production over South Africa". en. In: *South African Journal of Science* 117 (Oct. 2021), pp. 1 –11. ISSN: 0038-2353. URL: http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S0038-23532021000500022&nrm=iso.

[24] G. Medici et al. "The Intracloud Lightning Fraction in the Contiguous United States". In: *Monthly Weather Review* 145.11 (2017), pp. 4481 –4499. DOI: 10.1175/MWR-D-16-0426.1. URL: https://journals.ametsoc.org/view/journals/mwre/145/11/mwr-d-16-0426.1.xml.

[25] S. Moon and Y. Kim. “Forecasting lightning around the Korean Peninsula by postprocessing ECMWF data using SVMs and undersampling”. In: *Atmospheric Research* 243 (2020), p. 105026. ISSN: 0169-8095. DOI: https://doi.org/10.1016/j.atmosres.2020.105026. URL: http://www.sciencedirect.com/science/article/pii/S0169809520300995.

[26] A. Mostajabi et al. “Nowcasting lightning occurrence from commonly available meteorological parameters using machine learning techniques”. In: *npj Climate and Atmospheric Science* 2.1 (2019), p. 41. ISSN: 2397-3722. DOI: 10.1038/s41612-019-0098-0. URL: https://doi.org/10.1038/s41612-019-0098-0.

[27] V. Mucsi et al. “Lightning protection methods for wind turbine blades: An alternative approach”. In: *Applied sciences* 10.6 (2020), p. 2130.

[28] A. Nag et al. “Lightning locating systems: Insights on characteristics and validation techniques”. In: *Earth and Space Science* 2.4 (2015), pp. 65–93. DOI: https://doi.org/10.1002/2014EA000051. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014EA000051. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014EA000051.

[29] A. Nielsen, A. Iosifidis, and H. Karstoft. “CloudCast: A Satellite-Based Dataset and Baseline for Forecasting Clouds”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* PP (Mar. 2021), pp. 1–1. DOI: 10.1109/JSTARS.2021.3062936.

[30] G. Petneházi. *Recurrent neural networks for time series forecasting*. arXiv preprint arXiv:1901.00069. 2019. URL: https://arxiv.org/pdf/1901.00069.pdf.

[31] M. Phi. “Illustrated Guide to LSTM’s and GRU’s: A step by step explanation”. In: *Towards Data Science* (2018). URL: https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21 (visited on 02/14/2022).

[32] V. A. Rakov. *Fundamentals of Lightning*. Cambridge University Press, 2016. DOI: 10.1017/CBO9781139680370.

[33] V. A. Rakov and M. A. Uman. *Lightning: physics and effects*. Cambridge university press, 2003.

[34] M. G. Schultz et al. “Can deep learning beat numerical weather prediction?” In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379.2194 (2021), p. 20200097. DOI: 10.1098/rsta.2020.0097. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2020.0097. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2020.0097.

[35] T. Simon et al. “NWP-based lightning prediction using flexible count data regression”. In: *Advances in Statistical Climatology, Meteorology and Oceanography* 5 (Feb. 2019), pp. 1–16. DOI: 10.5194/ascmo-5-1-2019.

[36] H. Son and Y. Jang. “Partial Convolutional LSTM for Spatiotemporal Prediction of Incomplete Data”. In: *IEEE Access* 8 (2020), pp. 164762–164774.

[37] South African Weather Service. *Overview*. 2021. URL: https://www.weathersa.co.za/home/overview (visited on 08/27/2021).

[38] D. Speranza. “Lightning Prediction Using Recurrent Neural Networks Thesis”. PhD thesis. Air University, Air Force Institute of Technology, United States of America, 2019, p. 76. URL: https://scholar.afit.edu/etd/2317.

[39] A. Su et al. “A convection nowcasting method based on machine learning”. In: *Advances in Meteorology* 2020 (2020).

[40] G. Van Houdt, C. Mosquera, and G. Nápoles. “A review on the long short-term memory model”. In: *Artificial Intelligence Review* 53.8 (2020), pp. 5929–5955.

[41] H. Xie, Li Zhang, and C. P. Lim. “Evolving CNN-LSTM models for time series prediction using enhanced grey wolf optimizer”. In: *IEEE Access* 8 (2020), pp. 161519–161541.

[42] S. Xingjian et al. “Convolutional LSTM network: A machine learning approach for precipitation nowcasting”. In: *Advances in neural information processing systems*. 2015, pp. 802–810.

[43] Z. Yuan, X. Zhou, and T. Yang. “Hetero-ConvLSTM: A Deep Learning Approach to Traffic Accident Prediction on Heterogeneous Spatio-Temporal Data”. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (2018).

[44] H. Zang et al. “Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations”. In: *Renewable Energy* 160 (2020), pp. 26–41.