UNIVERSITE JOSEPH KI-ZERBO

ECOLE DOCTORALE INFORMATIQUE ET CHANGEMENTS CLIMATIQUES

BURKINA FASO

Unité-Progrès-Justice

MASTER RESEARCH PROGRAM

- - - - - o - - - - -

SPECIALITY: INFORMATICS FOR CLIMATE CHANGE (ICC)

MASTER THESIS

Subject:

Applying Deep Learning to Sub-seasonal to Seasonal (S2S) prediction of rainfall over Burkina Faso

Defended on July 12, 2022 by:

Abderahim TOGUYENI

Major Supervisor:

Pr. Tizane DAHO

Co-Supervisor:

Dr. Ulrich Jacques DIASSO

Jury members:

Pr. Issa ZERBO	**President**
Dr. Seyni SALACK	**Reviewer**
Dr. Ulrich Jacques DIASSO	**Supervisor**

Academic year: 2021-2022

Dedication

All the glory to you, **Jesus Christ**, for the academic journey you designed for me!

Acknowledgements

This document, fruit of our two years of training, was worked out thanks to the contribution of several actors. Therefore, we take the opportunity offered by these lines to thank:

- the German Ministry of Education and Research (BMBF), through the West African Science Service on Climate Change and Adapted Land Use (WASCAL), for designing and funding this program;
- Prof. Tanga Pierre ZOUNGRANA, Director of the "Ecole Doctorale Informatique et Changement Climatique" (ED-ICC);
- Dr. Ousmane COULIBALY, Deputy Director of ED-ICC;
- Dr. Benewindé Jean-Bosco ZOUNGRANA, Scientific coordinator;
- all staff members of ED-ICC;
- Prof. Tizane DAHO, our major supervisor, for his invaluable advice and support;
- Dr. Ulrich Jacques DIASSO, our co-supervisor, for providing the subject and all the support needed to complete the thesis;
- all the lecturers of ED-ICC for instilling knowledge in us;
- Zourkalaini BOUBAKAR, my classmate, for all the support in programming skills on Deep Learning;
- the comrades of the MRP-Informatics for Climate Change, batch 2, for their team mind;
- my family, my colleagues, friends and acquaintances, for their multifaceted support.

Abstract

This study proposes to predict rainfall on a Sub-seasonal to seasonal (S2S) time scale over six (6) locations (Dori, Ouahigouya, Ouagadougou, Fada N'Gourma, Bobo-Dioulasso and Gaoua) in Burkina Faso, using a specific architecture of Deep Learning called LSTM. Historical monthly and daily climate parameters from different sources are used to calibrate the LSTM model. After data preprocessing, the model is set and run for each location. Afterwards, the model is evaluated using some statistical metrics such as R^2, NSE, RMSE and MAE. The performance evaluation of the model using these metrics shows that LSTM model is effective and performs well in predicting rainfall at monthly timescales. For instance, forecasting at monthly timescale exibits a R^2 ranging from 0.66 to 0.83, NSE ranging from 0.62 to 0.80, RMSE ranging from 32.9 to 59.9mm and MAE ranging from 21.1 to 39.7mm. Regarding the bimonthly rainfall prediction, R^2 ranges from 0.63 to 0.83, NSE ranges from 0.6 to 0.82, RMSE ranges from 34.0 to 62.8mm and MAE ranges from 21.8 to 42.8mm. These results allow to highlight the impact of climatic zones and topography. Indeed, the models have better results on slightly humid plateaus than on very rainy and elevated areas of the country. However, trying to bring these monthly forecasts down to daily scales, the models struggle to capture daily rainfall for all locations. This requires more investigations to be done as part of future studies.

Keywords: Sub-seasonal to seasonal (S2S), Burkina Faso, Deep learning, Long Short Term Memory (LSTM), model.

Résumé

Cette étude se propose de prédire les précipitations sur une échelle de temps sub-saisonnière à saisonnière (S2S) au niveau de six (6) localités (Dori, Ouahigouya, Ouagadougou, Fada N'Gourma, Bobo-Dioulasso et Gaoua) au Burkina Faso, en utilisant une architecture spécifique de Deep Learning appelée LSTM. Des paramètres climatiques historiques mensuelles et journalières obtenues de différentes sources sont utilisées pour calibrer le modèle LSTM. Après le prétraitement des données, les modèles sont calibrés et implémentés pour chaque localité. Ensuite, le modèle est évalué à l'aide de mesures statistiques telles que R^2, NSE, RMSE et MAE. L'évaluation des performances du modèle à l'aide de ces métriques montre qu'il est efficace pour prédire les précipitations à des échelles de temps mensuelles. Par exemple, les prévisions à l'échelle mensuelle présentent un R^2 variant de 0,66 à 0,83, NSE variant de 0,62 à 0,80, RMSE variant de 32,9 à 59,9mm et MAE variant de 21,1 à 39,7mm. En ce qui concerne la prévision bimensuelle des précipitations, les résultats montrent que R^2 varie de 0,63 à 0,83, NSE varie de 0,6 à 0,82, RMSE varie de 34,0 à 62,8 mm et MAE varie de 21,8 à 42,8 mm. Ces résultats permettent de mettre en évidence l'impact des zones climatiques et de la topographie. En effet, les modèles ont de meilleurs résultats sur les plateaux peu humides que sur les zones très pluvieuses et montagneuses du pays. Cependant, en essayant de ramener ces prévisions mensuelles à des échelles quotidiennes, les modèles ne parviennent pas à reproduire les précipitations quotidiennes pour tous les sites. Cela nécessite plus d'investigations à faire dans le cadre d'études futures.

Mots clés: Sub-saisonnier à saisonnier (S2S), Burkina Faso, Deep learning, Long Short Term Memory (LSTM), modèle.

Acronyms and abbreviations

AEJ : African Easterly Jet

AI : Artificial Intelligence

ANAM : Agence Nationale de la Météorologie

ANN : Artificial Neural Network

API : Application Programming Interface

ASECNA : Agence pour la Sécurité de la Navigation Aérienne en Afrique et à Madagascar

BMBF : German Ministry of Education and Research

CDO : Climate Data Operators

CDS : Climate Data Store

CHIRPS : Climate Hazards Group InfraRed Precipitation with Stations

DL : Deep Learning

DBN : Deep Belief Network

3D : Three-dimensional

CNN : Convolutional Neural Network

ECMWF : European Center for Medium-range Weather Forecast

ED-ICC : Ecole Doctorale Informatique pour le Changement Climatique

ENSO : El Niño/Southern Oscillation

ERA : European ReAnalysis

ESMs : Earth System Models

EVA : Evaporation

FIT : Front Inter-Tropical

hPa : hecto-Pascal

IFM : Institut Für Meteorologie

INPE : Instituto Nacional de Pesquisas Espaciais

IPCC : Intergovernmental Panel on Climate Change

ITCZ : Intertropical Convergence Zone

LDCs : Least Developed Countries

LSTM : Long Short-Term Memory

m : meter

MAE : Mean Absolute Error

MENA : Middle East and North Africa

ML : Machine Learning

mm : millimeter

MRP-ICC : Master Research Program Informatics for Climate Change

m/s : Meter per seconde

NINO3.4 : Anomaly SST

NMME : North American Multi-Model Ensemble

NOAA : National Oceanic and Atmospheric Administration

NSE : Nash-Sutcliffe Efficiency

NWP : Numerical Weather Prediction

PRESASS : Seasonal forecasts of Agro-climatic characteristics of the rainy season for the Sudanese and Sahelian zones

ReLU : Rectified Linear Unit

RMSE : Root Mean Squared Error

RNN : Recurrent Neural Networks

RR : Precipitation

R^2 : Coefficient of determination

SPEI : Standardized precipitation evapotranspiration index

SST : Sea Surface temperature

S2S : Subseasonal to Seasonal

TEJ : Tropical Easterly Jet

TMAX : Maximum air temperature

TMIN : Minimum air temperature

UBN : Upper Blue Nile basin

UDM : Relative humidity

USD : United State Dollar

WAM : West African Monsoon

WASCAL : West African Science Service Centre on Climate Change and Adapted Land Use

WFM : Wind speed

WJS : West Jet Subtropical

WMO : World Meteorological Organization

W700 : Wind speed at 700hPa

W850 : Wind speed at 850hPa

W925 : Wind speed at 925hPa

°C : Degree celsius

List of Tables

List of figures

Introduction

Context and justification

Climate change is already affecting every region on Earth, in multiple ways, increasing the need for reliable weather and climate forecasts with a view to cope, in particular, with the high frequency and intensity of extreme weather events (flood, drought, heatwaves, major hurricanes...) [1]. Even though there are various levels of certainty associated with the linkages between climate change and extreme weather events, there is no doubt that these climate related events are exacerbated by climate change [2].

Due to the low adaptive capacity of Sub-Saharan African countries, climate related events are undermining decision makers' efforts toward socioeconomic development. The case of the Sahel in West Africa is quite illustrative. Indeed, this region is known to be very sensitive to climatic fluctuations. Inter-annual variability of precipitation has immediate and strong consequences on agriculture, food security, water resources, biodiversity and energy. The region has faced series of extreme climatic events through the past. Recently, a long period of persistent drought jeopardized the ecosystems equilibrium. Indeed, from the 1970's to the mid 1990's, precipitations were strongly and repeatedly below average [3]. Conversely, during the last decade, several floods have severely hit the Sahel. Ouagadougou, in Burkina Faso, for instance, experienced its most severe floods on the 1^{st} of September, 2009. According to the World Bank, this flood affected more than 150,000 people, and resulted in estimated damages and losses to the economy of more than USD 130 million. In terms of rainfall projection, the region is not yet out of the rut because climate change is going to affect rainfall patterns. Precipitation is likely to decrease over large parts of the subtropics and changes to monsoon precipitation are expected [1].

From above, there is an urgent need to provide decision makers with accurate weather forecasts and climate predictions so that climate events don't annihilate their economical development, which is already weak.

Unfortunately, meteorological services in the Sahel are being issued forecasts for the timescales that range from short range (1–5 days), to medium range (7–15 days), and long range (3–12 months) [4]. Specifically, the two first are being issued as daily to

weekly outlooks of the weather while the last one is being provided and updated through Seasonal forecasts of Agro-climatic characteristics of the rainy season for the Sudanese and Sahelian zones (PRESASS). It means that, these meteorological services are not dealing so much with the gap between the medium range and the long range. To achieve this, it's critical to bridge the gap between daily weather forecasts and seasonal climate outlooks. Indeed, many management decisions in agriculture and food security, water, disaster risk reduction, and health fall into this gap in prediction capabilities [4]. Therefore, developing forecast capabilities for this time range would be of considerable societal value.

Problem statement

For the past several decades, forecasts of weather, ocean and other environmental phenomena made on short and medium range timescales have yielded invaluable information to improve decision making across many socio-economic sectors [5]. It worth pointing out that these improvements in weather and climate forecasts have to do with the advent of Numerical Weather Prediction (NWP) and computational facilities. Climate and weather information has gain in importance as a decision-making tool and for climate adaptation in the context of climate change. Unfortunateley, several fundamental issues are hindering the production and use of these forecasts. Firstly, the medium range timescales appear too short for any meaningful mitigating action to be taken. Secondly, there exists a gap, referred to as the "predictability desert", between the medium and long-range timescale forecasts [6]. This gap is also called Subseasonal to Seasonal (S2S) range.

In addition, the NWP models themselves have some inherent limits that lead to uncertainties in forecasts. These models can be referred to as bottom-up approaches because they forecast climate using physical boundary conditions. Although the performance of Earth System Models (ESMs) is improving, these models still suffer from significant forecast uncertainties. Such uncertainties in future climate may delay mitigation and adaptation to climate change [1]. Then, there is a need to pave the way toward building our own strategy to issues weather and climate predictions.

Therefore, in the Sahel, it appears necessary to explore others alternative to these NWP techniques which are highly computational costly. Moreover, with the discovery of nonlinearity in weather data, the focus has shifted towards the nonlinear prediction of the weather [7].

For the particular case of the parameter rainfall, several issues still prevent its accurate predictions, including the high variability of rainfall pattern in Sub-saharian Africa and the lack of mastery of the West African Monsoon (WAM), which is the main driver of rainfall in this region. To overcome all these challenges, several machine learning methods and models have been used to make reliable and timely rainfall predictions to overcome this ambiguity [19]. Deep Learning, in particular Long Short-Term Memory (LSTM), appears to be well suited to bring out the structural relationship between the various climate entities which are taken into account in the rainfall prediction. LSTMs are effective in solving many time series tasks which other feedforward networks could not solve [8].

Research questions

This work aims at answering this main question:

- **Q1:** How well does a LSTM model capture S2S rainfall over Burkina Faso?

Research hypotheses

The main hypothesis of this work is:

- **H1:** LSTM model used to predict S2S rainfall over Burkina Faso gives good scores.

Research objectives

The main objective of this research is to develop a LSTM model that is capable to predict S2S rainfall over Burkina Faso. Specifically, it is about to:

- **O1:** Determine the performances and efficiency of this LSTM model.

Thesis structure

This document is structured in 3 chapters:

- the first chapter highlights the knowledge acquired as well as related works and their limits in the field of the use of Artificial Intelligence in rainfall predictions;
- the second chapter is dedicated to the description of the data used and the methods adopted to obtain the results;

- these results will be analyzed and discussed in the third chapter.

The conclusion includes a reminder of the purpose of the study, a summary of the important results obtained, recommendations and some strong areas for further studies.

Chapitre 1

Literature review

In this study, a DL based Long Short-Term Memory (LSTM) will be used to predict rainfall from the timeseries weather data seen in section 2.2. First and foremost, we need to define clearly what we're talking about when we mention Deep Learning and LSTM, and more broadly the concepts of Artificial Intelligence and Machine Learning. Furthermore, the West African Monsoon (WAM) system is the main source of rainfall in the agriculturally based region of the Sahel [24]. Understanding transport across the WAM as well as the dynamics behind the WAM is of crucial importance in this study.

1.1 Artificial Intelligence, Machine Learning, and Deep Learning

The figure 1 suggests that Deep Learning (DL) is a subset of Machine Learning (ML) which in turn is a subset of Artificial Intelligence (AI). Let's have a look to what is behind these three (3) concepts.

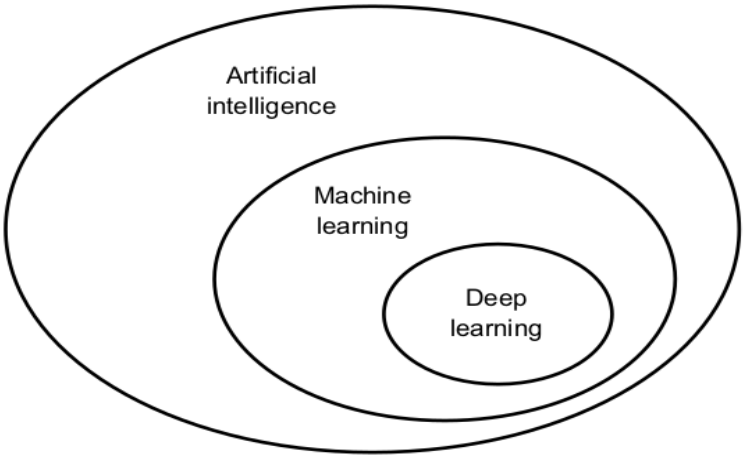


Figure 1: Subsets of Artificial Intelligence. *Source: Wikipedia*

- ***Artificial Intelligence (AI):***

  AI was born in the 1950s, when a handful of pioneers from the nascent field of computer science started asking whether computers could be made to "think"; a question whose ramifications we're still exploring today.

  Concisely, AI can be described as the effort to automate intellectual tasks normally performed by humans. As such, AI is a general field that encompasses ML and DL,

but that also includes many more approaches that may not involve any learning. Consider that until the 1980s, most AI textbooks didn't mention "learning" at all! Early chess programs, for instance, only involved hardcoded rules crafted by programmers, and didn't qualify as ML. In fact, for a fairly long time, most experts believed that human-level AI could be achieved by having programmers handcraft a sufficiently large set of explicit rules for manipulating knowledge stored in explicit databases. This approach is known as symbolic AI. It was the dominant paradigm in AI from the 1950s to the late 1980s [9], and it reached its peak popularity during the expert systems boom of the 1980s.

Although symbolic AI proved suitable to solve well-defined, logical problems, such as playing chess, it turned out to be intractable to figure out explicit rules for solving more complex, fuzzy problems, such as image classification, speech recognition, or natural language translation. A new approach arose to take symbolic AI's place: Machine Learning;

- ***Machine Learning (ML):***

  The usual way to make a computer do useful work is to have a human programmer write down rules -a computer program- to be followed to turn input data into appropriate answers, just like Lady Lovelace (who criticized the invention of the Analytical Engine in 1843 [9]) writing down step-by-step instructions for the Analytical Engine to perform. Machine learning turns this around: the machine looks at the input data and the corresponding answers, and figures out what the rules should be (see figure 2). A ML system is trained rather than explicitly programmed. It's presented with many examples relevant to a task, and it finds statistical structure in these examples that eventually allows the system to come up with rules for automating the task. For instance, if you wished to automate the task of tagging your vacation pictures, you could present a ML system with many examples of pictures already tagged by humans, and the system would learn statistical rules for associating specific pictures to specific tags.

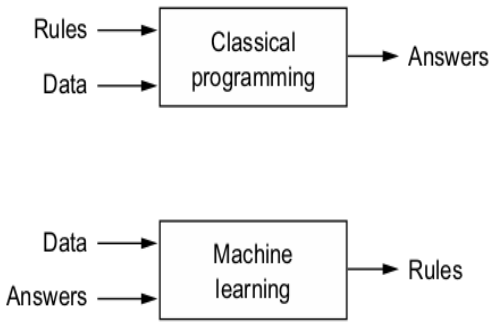


Figure 2: Machine Learning: a new programming paradigm.

Although ML only started to flourish in the 1990s, it has quickly become the most popular and most successful subfield of AI [9], a trend driven by the availability of faster hardware and larger datasets. ML is related to mathematical statistics, but it differs from statistics in several important ways, in the same sense that medicine is related to chemistry but cannot be reduced to chemistry, as medicine deals with its own distinct systems with their own distinct properties. Unlike statistics, ML tends to deal with large, complex datasets (such as a dataset of millions of images, each consisting of tens of thousands of pixels) for which classical statistical analysis such as Bayesian analysis would be impractical. As a result, ML, and especially deep learning, exhibits comparatively little mathematical theory - maybe too little -and is fundamentally an engineering discipline. Unlike theoretical physics or mathematics, ML is a very hands-on field driven by empirical findings and deeply reliant on advances in software and hardware.

- ***Deep Learning (DL):***

  DL is a specific subfield of ML: a new take on learning representations from data that puts an emphasis on learning successive layers of increasingly meaningful representations. The "deep" in "deep learning" isn't a reference to any kind of deeper understanding achieved by the approach; rather, it stands for this idea of successive layers of representations. How many layers contribute to a model of the data is called the *depth* of the model. Other appropriate names for the field could have been *layered representations learning* or *hierarchical representations learning* [9]. Modern deep learning often involves tens or even hundreds of successive layers of representations, and they're all learned automatically from exposure to training

data. Meanwhile, other approaches to ML tend to focus on learning only one or two layers of representations of the data (taking a pixel histogram and then applying a classification rule); hence, they're sometimes called *shallow learning*.

In DL, these layered representations are learned via models called neural networks, structured in literal layers stacked on top of each other. The term "neural network" refers to neurobiology, but although some of the central concepts in DL were developed in part by drawing inspiration from our understanding of the brain (in particular, the visual cortex), DL models are not models of the brain. There's no evidence that the brain implements anything like the learning mechanisms used in modern DL models. You may come across pop-science articles proclaiming that DL works like the brain or was modeled after the brain, but that isn't the case. It would be confusing and counterproductive for newcomers to the field to think of DL as being in any way related to neurobiology; you don't need that shroud of "just like our minds" mystique and mystery, and you may as well forget anything you may have read about hypothetical links between DL and biology. For our purposes, deep learning is a mathematical framework for learning representations from data.

What do the representations learned by a DL algorithm look like? Let's examine how a network several layers deep (see figure 3) transforms an image of a digit in order to recognize what digit it is.

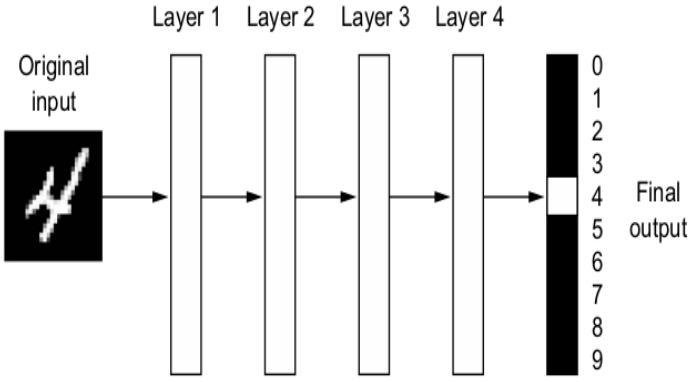


Figure 3: A deep neural network for digit classification. *Source: Chollet (2021).*

As you can see in figure 4, the network transforms the digit image into representations that are increasingly different from the original image and increasingly informative about the final result. You can think of a deep network as a multistage

information-distillation process, where information goes through successive filters and comes out increasingly *purified* (that is, useful with regard to some task).

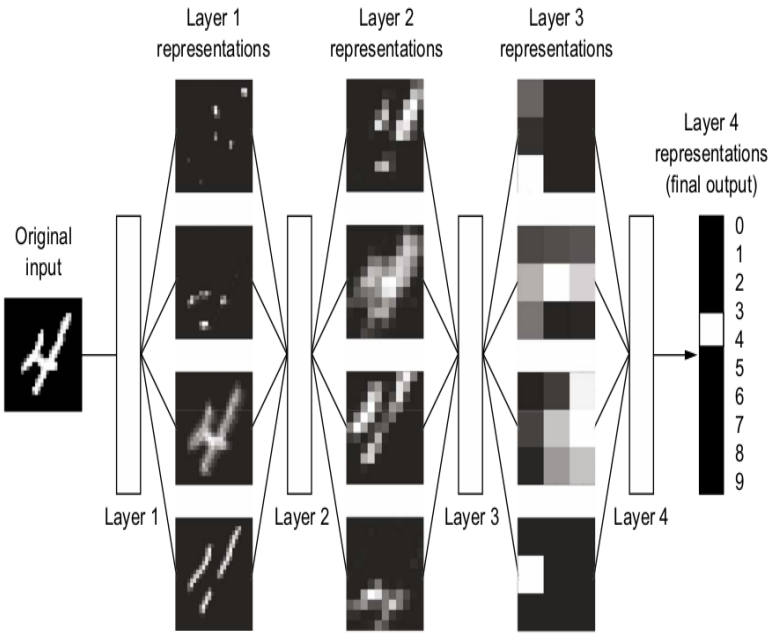


Figure 4: Data representations learned by a digit-classification model. *Source: Chollet (2021).*

So that's what deep learning is, technically: a multistage way to learn data representations. It's a simple idea—but, as it turns out, very simple mechanisms, sufficiently scaled, can end up looking like magic. The figure 5 shows how DL actually works.

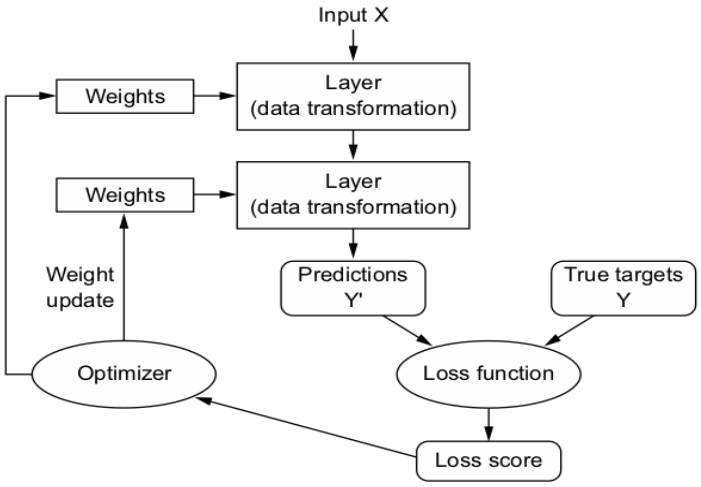


Figure 5: How Deep Learning works: the loss score as a feedback to adjust the weights

Initially, the weights of the network are assigned random values, so the network merely implements a series of random transformations. Naturally, its output is far

from what it should ideally be, and the loss score is accordingly very high. But with every example the network processes, the weights are adjusted a little in the correct direction, and the loss score decreases. This is the training loop, which, repeated a sufficient number of times (typically tens of iterations over thousands of examples), yields weight values that minimize the loss function. A network with a minimal loss is one for which the outputs are as close as they can be to the targets: a trained network. Once again, it's a simple mechanism that, once scaled, ends up looking like magic.

1.2 LSTM Model

There's a family of neural network architectures designed specifically for the type of data we are dealing with in this study (timeseries or sequence data, where causality and order matter): Recurrent Neural Networks (RNN). Among them, the Long Short-Term Memory (LSTM) algorithm, which is fundamental to deep learning for timeseries [9], was developed by Hochreiter and Schmidhuber in 1997 and has long been very popular. It was the culmination of their research on the vanishing gradient problem: allow past information to be reinjected at a later time, thus fighting the vanishing-gradient problem. This is essentially what LSTM does: it saves information for later, thus preventing older signals from gradually vanishing during processing. Let's index the W and U matrices in the cell, with the letter o (Wo and Uo) for output. The data flow that carries information across timesteps. Call its values at different timesteps c_t , where C stands for carry. This information will have the following impact on the cell: it will be combined with the input connection and the recurrent connection (via a dense transformation: a dot product with a weight matrix followed by a bias add and the application of an activation function), and it will affect the state being sent to the next timestep (via an activation function and a multiplication operation). Conceptually, the carry dataflow is a way to modulate the next output and the next state.

Now the subtlety, the way the next value of the carry dataflow is computed. It involves three (3) distinct transformations. All three have the form of a Simple RNN cell [9]:

$$y = activation(dot(state_t, U) + dot(input_t, W) + b) \tag{1}$$

But all three transformations have their own weight matrices, which we'll index with the letters i , f , and k .

- Pseudocode details of the LSTM architecture (1/2) [9]:

$$\begin{cases} output_t &= activation(dot(state_t, Uo) + dot(input_t, Wo) + dot(c_t, Vo) + bo) \\ i_t &= activation(dot(state_t, Ui) + dot(input_t, Wi) + bi) \\ f_t &= activation(dot(state_t, Uf) + dot(input_t, Wf) + bf) \\ k_t &= activation(dot(state_t, Uk) + dot(input_t, Wk) + bk) \end{cases} \tag{2}$$

- Pseudocode details of the LSTM architecture (2/2) [9]:

$$c_t + 1 = i_t * k_t + c_t * f_t \tag{3}$$

Add this as shown in figure 6, and that's it.

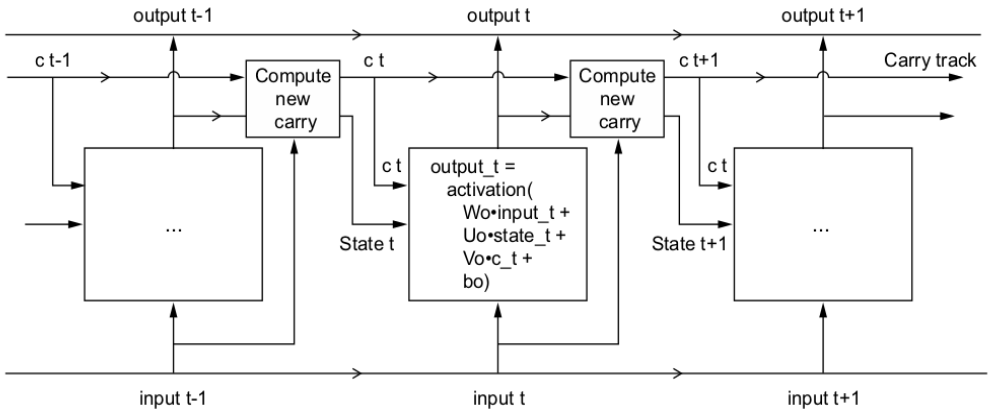


Figure 6: Anatomy of an LSTM architecture. *Source: Chollet (2021)*

1.3 Dynamics of the West African Monsoon (WAM)

The WAM is a coupled land-ocean-atmosphere system involving many spatial and temporal scales. It is a dynamic and hydrous response of the atmosphere to the differential between the ocean and the continent and therefore to the horizontal energy gradients in

the lower layers on a regional scale. Its sensitivity to surface conditions is accentuated by geography: low latitudes, absence of significant relief, quasi-zonal distribution of vegetation, soil humidity and albedos. The African monsoon circulation is organized on a regional scale, as shown in figure 7, around key elements of the zonal movement in the upper and middle troposphere (the African Easterly Jet (AEJ) and the Tropical Easterly Jet (TEJ)), as well as in the lower troposphere (the monsoon and harmattan flow). It also involves a convergence zone in the lower layers known as the Inter-Tropical Front (FIT) and two main convective structures: the ITCZ characterized by a maximum equivalent potential temperature θ_e and the Saharan thermal depression associated with a maximum potential temperature θ.

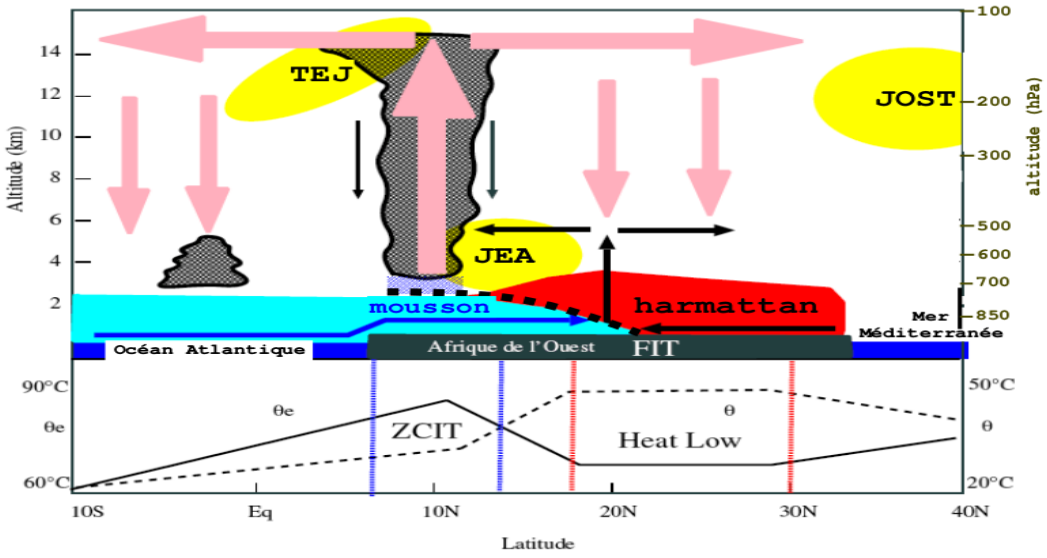


Figure 7: Conceptual diagram representing, in zonal average, the key elements of the African monsoon during the boreal summer. *Source: Beucher (2010)*

Its seasonal cycle arouses great interest for the scientific community because the date of arrival and the intensity of the monsoon constitute the first concerns of the West African populations. Indeed, the movement of the ZCIT follows the apparent position of the sun with a lag of 6 to 8 weeks. The first rainy season corresponds to the first jump of the ITCZ which migrates from 2°N to 5°N at the beginning of May: this is the pre-monsoon phase called "pre-onset" (figure 8). During this period, the FIT rises up to 15°N, the AEJ strengthens (10m/s monthly average) and migrates towards 10°N while the TEJ (7-8m/s) and the West Jet Subtropical (WJS) remain positioned at the equator [10].

The second rainy season marks the second jump of the ITCZ which goes from 5°N to

10°N: this is the monsoon phase called "onset" or Sahelian regime. During this phase, the sea surface temperature in the Gulf of Guinea drops sharply from 26°C to 24°C and the surface pressure increases by a few hPa [10]. Meanwhile, over the Sahara, the surface temperature continues to increase, which reinforces the thermal depression towards 25°N. Consequently, the horizontal gradients of pressure, temperature and humidity increase between the Sahara and the Gulf of Guinea, which favors the acceleration of the monsoon flow and its extension towards the North during this phase (20°N on average). Still, during this phase, the AEJ migrates from 10°N to 15°N while maintaining the same intensity and the TEJ rises towards 10°N while intensifying and the Tropical West Jet shifts towards 40°N [10].

Finally, from the end of August, the rain zone retreats towards the south: this is the period of withdrawal of the monsoon.

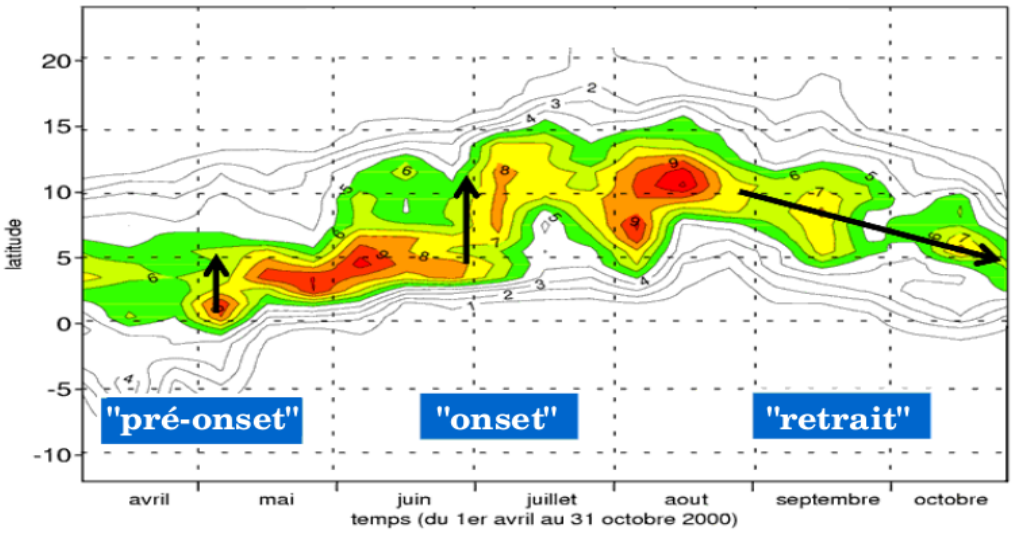


Figure 8: Average daily rainfall (in mm/day) between 10°E and 10°W. *Source: Beucher (2010).*

1.4 Related studies

This chapter presents a review of some of the recent works in rainfall prediction that are relevant to our work. Even though S2S rainfall prediction is at a relatively early stage of development, many studies focused on the performances of Machine Learning models in this timescale. However very few are related to Western African particularly to the Sahel region. Here are some conclusions from these studies.

Michael Scheuerer et al. used Artificial Neural Network (ANN) to establish relationships between NWP ensemble forecast and gridded observed 7-day precipitation accumulations, and to model the increase or decrease of the probabilities for different precipitation categories relative to their climatological frequencies over California. A Convolutional Neural Network (CNN) framework is proposed that extends the basic ANN and takes images of large-scale predictors as inputs that inform local increase or decrease of precipitation probabilities relative to climatology. The forecast skill relative to climatology is positive everywhere within the domain. The magnitude of skill, however, is low for week-3 and week-4, and suggests that additional sources of predictability need to be explored [11].

Guoxing Chen and Wei-Chyung Wang performed Short-term precipitation prediction using deep learning. They showed that a 3D convolutional neural network using a single frame of meteorology fields as input is capable of predicting the precipitation spatial distribution. The network is developed based on 39-years (1980–2018) data of meteorology and daily precipitation over the contiguous United States. The results bring fundamental advancements in weather prediction. First, the trained network alone outperforms the state-of-the-art weather models in predicting daily total precipitation, and the superiority of the network extends to forecast leads up to 5 days. Second, combining the network predictions with the weather-model forecasts significantly improves the accuracy of model forecasts, especially for heavy-precipitation events. Third, the millisecond-scale inference time of the network facilitates large ensemble predictions for further accuracy improvement. These findings strongly support the use of deep-learning in short-term weather predictions. Although the results above are based on next-day total precipitation, the scope of deep learning approach is much broader, applicable to other weather parameters (e.g., temperature and wind), at different regions, and for shorter or longer forecast leads, as long as sufficient data are readily available for training [12].

Sijie He et al. investigated ten (10) Machine Learning (ML) approaches to sub-seasonal temperature forecasting over the U.S. The results indicate that suitable ML models, e.g., XGBoost, to some extent, capture the predictability on sub-seasonal timescales and can outperform the climatological baselines, while DL models barely manage to match the best results with carefully designed architecture [13].

Hamada S. Badr et al. implement CNN using observed data and dynamical forecasts at S2S timescales for different countries in the Middle East and North Africa (MENA). Train-

ing is performed using predictors from historical observations and the North American Multi-Model Ensemble (NMME) forecasts along with gridded 5-km resolution CHIRPS precipitation as predictant. A custom loss function is defined to minimize the root mean squared error (RMSE) of the regional average for each defined climate region. They use CNN to generate high-resolution gridded precipitation predictions at monthly timescale. The proposed CNN workflow has been successfully applied to improve the skill and resolution of dynamically-based S2S precipitation forecasts for MENA [21].

Sanam Narejo et al. tried Multi-step rainfall forecasting using deep learning approach in a specific site in Italy. The focus of their work is direct prediction of multistep forecasting, where a separate time series model for each forecasting horizon is considered and forecasts are computed using observed data samples. Forecasting in this method is performed by proposing a deep learning approach, i.e, Temporal Deep Belief Network (DBN). The best model is selected from several baseline models on the basis of performance analysis metrics. The results suggest that the temporal DBN model outperforms the conventional Convolutional Neural Network (CNN) specifically on rainfall time series forecasting. However, it turned out that training DBN is more exhaustive and computationally intensive than other deep learning architectures [14].

Yashon O. Ouma et al. did rainfall and runoff time-series trend analysis using LSTM recurrent neural network and wavelet neural network in Nzoia hydrologic basin. They used long-term in situ observed data for 30 years (1980–2009) from ten (10) rain gauge stations and three discharge measurement stations, the rainfall and runoff trends in the Nzoia River basin are predicted through satellite-based meteorological data comprising of: precipitation, mean temperature, relative humidity, wind speed and solar radiation. Even though both models performed well, the study shows that in hydrologic basins with scarce meteorological and hydrological monitoring networks, the use satellite-based meteorological data in deep learning neural network models are suitable for spatial and temporal analysis of rainfall and runoff trends [15].

Chengcheng Chen et al. forecasted monthly rainfall distribution based on fixed sliding window long short-term memory. To this end, monthly rainfall data for a period of 41years (1980–2020) from two meteorological stations in Turkey, namely Rize and Konya, with different climatic conditions, were used. The results revealed that the LSTM model, as a more efficient tool, outperforms the RF model in forecasting rainfall at both stations. The

LSTM-based approach proposed could be adopted over any global climatic conditions to forecast the monthly rainfall with reasonable accuracy [16].

Ari YairBarrera-Animas et al. made rainfall prediction by making a comparative analysis of modern machine learning algorithms for time-series forecasting. Specifically, they presented a comparative analysis using simplified rainfall estimation models based on conventional Machine Learning algorithms and Deep Learning architectures that are efficient for these downstream applications. The evaluation metrics of Loss, Root Mean Squared Error, Mean Absolute Error, and Root Mean Squared Logarithmic Error were used to evaluate the models' performance. It ended up suggesting that models based on LSTM-Networks with fewer hidden layers perform better for this approach; denoting its ability to be applied as an approach for budget-wise rainfall forecast applications [17].

Demeke Endalie et al. developed a Deep learning model for daily rainfall prediction in Jimma, Ethiopia. They used daily records of weather parameters such as maximum temperature (tmax), minimum temperature (tmin), relative humidity, solar radiation, wind speed, and precipitation from 1985 to 2017. On this dataset, several experiments and comparisons with the existing machine-learning-based model are performed to validate the performance of the proposed predictive model. As a result, the proposed LSTM-based rainfall predictive model is suitable for use in a variety of applications requiring rainfall prediction, such as smart agriculture [18].

Juliana Aparecida Anochi et al. have evaluated different machine learning models for precipitation prediction over South America. They stressed that currently, NWP models are unable to precisely reproduce the precipitation patterns in South America due to many factors such as the lack of region-specific parametrizations and data availability. It turned out that machine learning models are able to produce predictions with errors under 2 mm in most of the continent in comparison to satellite-observed precipitation patterns for different climate seasons, and also outperform INPE's (Instituto Nacional de Pesquisas Espaciais) model for some regions (e.g., reduction of errors from 8 to 2 mm in central South America in winter). Another advantage is the computational performance from machine learning models, running faster with much lower computer resources than models based on differential equations currently used in operational centers [19].

Getachew Mehabie Mulualem et al. developed seven (7) ANN predictive models incorporating hydro-meteorological, climate, sea surface temperatures, and topographic at-

tributes to forecast the standardized precipitation evapotranspiration index (SPEI) for seven (7) stations in the Upper Blue Nile basin (UBN) of Ethiopia from 1986 to 2015. It was found that the coefficient of determination and the root-mean-square error of the best architecture ranged from 0.820 to 0.949 and 0.263 to 0.428, respectively. In terms of statistical achievement, they concluded that ANNs offer an alternative framework for forecasting the SPEI drought index [20].

Charity Oseiwah Adjei et al. used a LSTM Deep Learning Approach for rainfall forecasting in Sub-Sahara Africa-Ghana. They studied the predictive capacity of deep learning and built an Artificial Neural Network (ANN) model by using the Long Short-Term Memory (LSTM) algorithm and the Spearman coefficient based on selected parameters (precipitation, humidity, temperature, mean sea level pressure, wind speed, dew point, wind direction) reported by the weather station. They made hourly rainfall predictions about Axim, in the western region of Ghana. The model did very well because it ended up with a MSE result of 0.002 and a MAE result of 0.021. But some limitations have affected the quality of the study. Firstly, the authors didn't get historical data from the weather station located in the area of interest of the study. Secondly, the study focussed on a single site which did not allow to investigate the impact of climatic zones and topography of Ghana. It would be better to explore the performance of the model in several locations in the country to see how well it performs across the country [22].

Arsène Aizansi developed ANN models to predict monthly rainfall for six(6) geographically diverse weather stations across Benin Republic. He used data that cover January 1979 to December 2018. Eight (8) predictors have been selected including Minimum Relative Humidity, Maximum Relative Humidity, Minimum Air Temperature, Maximum Air Temperature, Evaporation, Zonal and Meridional wind at 850 hPa, Sea Surface temperature (SST). The ANN models performed well while highlighting that SST, zonal and meridional winds were relevant variables that contribute to rainfall. The importance of geography was stressed as well. Indeed the predictions were more accurate in higher altitudes. However, the study didn't embrace actually the entire country due to the limited number of weather stations; in addition there is a need to increase the predictors in order to covers the majority of climate parameters as well as relevant factors on which rainfall depends. Besides, it would be better to assess the ANN models over a large space scale such as Sahel, Savannah and Gulf of Guinea in west Africa [23].

It's clear that there is a need to take advantage of all those results to move forward by implementing a deep learning technique in Burkina Faso. By doing so, we will see how to improve the overall weather forecast in the country and in particular the S2S prediction of precipitation; while keeping in mind that issuing accurate rainfall forecasts is a perpetual process initiated since 1904 by Vilhelm Bjerknes (the precursor of the first deterministic models) and carried out so far by the scientific community. This study will attempt to make its modest contribution to this effort.

Chapitre 2

Materials and methods

2.1 Study area

Burkina Faso is the area of interest of this study. Specifically, the study covers six (6) meteorological stations of this country. These stations were choosen at the rate of two (2) stations per climatic zone.

A landlocked country in West Africa, Burkina Faso covers an area of $274200Km^2$. It is limited to the North and West by Mali, to the South by Côte d'Ivoire, Ghana, Togo, Benin and to the east by Niger. The country covers 625 km from North to South and $850km$ from East to West (figure 9).
Like other West African countries, the climate in Burkina Faso is under the influence of large-scale atmospheric circulation, controlled by the interaction of two air masses. Warm, dry continental air masses from the Sahara desert give rise to dusty winds called "Harmattan" which sweeps across West Africa from November to February. In summer, moist equatorial air masses from the Atlantic Ocean bring monsoon rains [25]. Thus, the variation of the influence of these air masses throughout the year is characterized by the north-south movement of the intertropical convergence zone (ITCZ). Due to the influence of these air masses, the precipitation regime in Burkina Faso makes it possible to distinguish three climatic zones [26]:

- Sahelian zone cumulating an average annual rainfall between 300 and 600 mm. In this area, the rainy season lasts on average of 3 to 4 months;
- Sudano-Sahelian zone recording annual rainfall from 600 to 900 mm for an average duration of rains of 4 to 5 months;
- Sudanian zone receiving a rainfall of between 900 and 1200 mm for a rainy season of about 5 to 6 months.

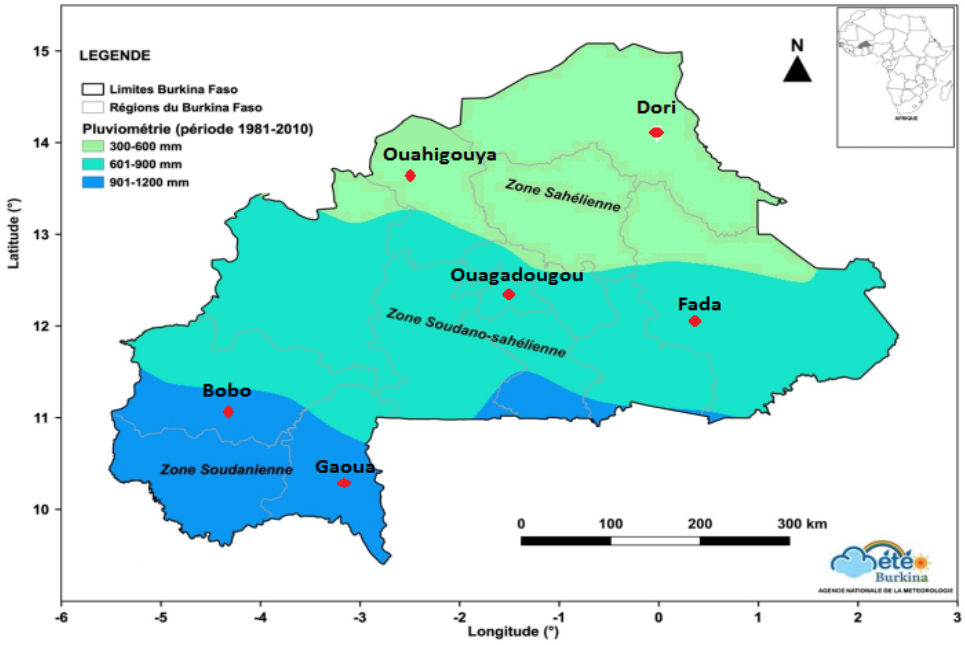


Figure 9: Map of Burkina Faso with the climatic zones (the stations in red). *Source: ANAM*

Table 1 summarises the main features of the six (6) weather stations.

Table 1: Features of the six (6) weather stations.

Station name	**WMO code**	**Latitude**	**Longitude**	**Elevation (m)**
Dori	65501	14°02'N	000°02'W	276
Ouahigouya	65502	13°34'N	002°25'W	337
Ouagadougou	65503	12°21'N	001°31'W	316
Fada N'Gourma	65507	12°02'N	000°22'E	308
Bobo-Dioulasso	65510	11°10'N	004°19'W	460
Gaoua	65522	10°20'N	003°11'W	333

2.2 Data

As part of this study, we only used historical meteorological data from three (3) sources. Most of the data was directly recorded in the six (6) meteorological stations. The data we got are on a daily and monthly scale and covers the period 1979-2020. More precisely, we got:

- daily weather data from National Meteorological Agency (ANAM) of Burkina Faso. These data are related to six (6) weather parameters namely Precipitation, Evaporation, Pressure, Minimum temperature, Maximum temperature and Wind speed. These data are available for each of the 6 weather stations. To have an overview of these data, **Appendix A** gives a visualization of all these weather parameters by station;

- daily reanalysis data, ERA5, on Sea Surface Temperature (SST) and wind speed at 700hPa, 850hPA and 925hPA. These data have been downloaded from Climate Data Store (CDS) https://cds.climate.copernicus.eu. The data cover the Earth on a 30km grid (0.25×0.25) and resolve the atmosphere using 137 levels from the surface up to a height of 80km;

- monthly SST anomalies (NINO3.4) data from the National Oceanic and Atmospheric Administration (NOAA). The NINO3.4 index is one of the several El Niño/Southern Oscillation (ENSO) indicators based on sea surface temperatures. NINO3.4 is the average sea surface temperature anomaly in the region bounded by 5°N to 5°S, from 170°W to 120°W with an impact on rainfall in the sahel [27]. This region has large variability on El Niño time scales, and is close to the region where changes in local sea-surface temperature are important for shifting the large region of rainfall typically located in the far western Pacific. The data are available through this link https://psl.noaa.gov/gcos_wgsp/Timeseries/Nino34/.

Table 2 summarizes essential characteristics of the data used in this study.

Table 2: Characteristics of the data used

N°	Parameters	Unit	Level	Step	Period	Source
1	Precipitation (RR)	mm	1.05m	daily	1979-2020	ANAM
2	Relative humidity (UDM)	%	1.5m			
3	Maximum air temperature (TMAX)	°C	1.5m			
4	Minimum air temperature (TMIN)	°C	1.5m			
5	Evaporation (EVA)	mm	surface			
6	Wind speed (WFM)	m/s	10m			
7	Pressure (PSEAD)	hPA	sea			
8	SST	°C	sea			ERA5
9	Wind speed at 700hPa (W700)	m/s	700hPA			
10	Wind speed at 850hPa (W850)	m/s	850hPA			
11	Wind speed at 925hPa (W925)	m/s	925hPA			
12	Anomaly SST (NINO3.4)	-	sea	monthly		NOAA

2.3 Processing tools

The processing of these data required the utilization of three (03) software that are run on command lines in a Linux environment:

- **Python** as a programming language on Ubuntu 20.04 64 bits operation system. Jupyter Notebook is the open document format for interactive computing that was used as a programming interface. Python supports several libraries designed for data analysis and machine learning in general. We used Keras module as a machine learning library. Keras is a deep learning API (created by François CHOLLET in March 2015) for Python, built on top of TensorFlow, that provides a convenient way to define and train any kind of Deep Learning model. Keras was initially developed for research, with the aim of enabling fast deep learning experimentation;

- **Ferret** is an interactive visualization and analysis tool developed by the National Oceanic and Atmospheric Administration (NOAA) to meet the needs of oceanographers and meteorologists who analyze large and complex data at grid points. It is

therefore a software well suited for reading data in NetCDF format. Ferret is free and can be downloaded from http://ferret.pmel.noaa.gov/Ferret/downloads;

- Climate Data Operators (**CDO**) is a software developed by IFM (Institute Für Meteorologie); it allows a standard processing of outputs from climate models and weather forecasts. It is a very powerful tool with more than 400 operators which include simple statistical and arithmetic functions, data selection, and tools of sub sampling and spatial interpolation. CDO is free and it can be downloaded on https://code.zmaw.de.

2.4 Methods

This section describes the approach adopted to achieve the objectives of this thesis. It consists of a first phase called data preprocessing in which raw data format is prepared to feed the model. Then comes the second phase, LSTM model calibration where the optimized set of hyperparameters were selected. Finally, the last phase deals with validation and testing of the model. The following figure (figure 10) summarizes the methods used in study.

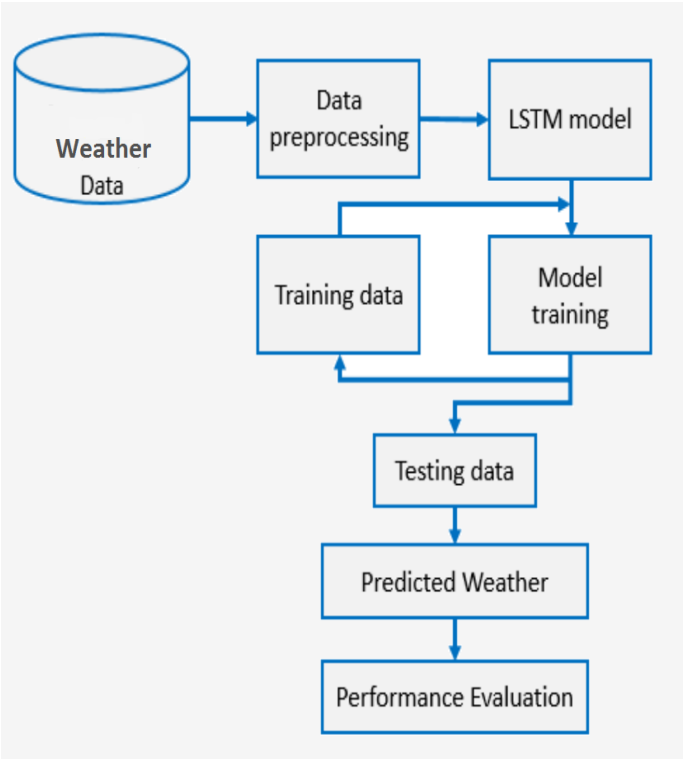


Figure 10: Summary of the methods.

2.4.1 Data preprocessing

Since this study deals with timeseries forecasting, we have to get a dataset (ideally in .csv file) which contains the predictant (precipitation) as well as the predictors (other weather parameters). To achieve this goal:

- we computed the correlation between the rainfall of each weather station and the SST in order to identify the relevant basins that have an influence on the corresponding rainfall patterns. Then we averaged the SST accordingly over the selected basin to get a single timeseries for this parameter;
- then we merged the daily data from all sources (ANAM, ERA5 and NOAA) to get a first dataset made of daily data (15341 rows and 12 columns);
- afterwards, we resampled this first dataset in order to get the second dataset made of monthly data (504 rows and 13 columns). The latter datased was used to perform **monthly (with a lead time of 0)** and **bimonthly (with a lead time of 1 month)** forecasts of rainfall over the six (6) weather stations. These monthly and bimonthly timescales correspond to the S2S timescale. As far as the first dataset is concerned, we used it to perform multisteps forecastings 30 days ahead. Indeed, even though we can get the accumulated rainfall through the monthly forecast it's also of crucial importance to have the daily distribution in the month; that's important to help, for instance, farmers to properly manage the development of their crops;
- finally, before feeding the model, we performed features selection to discard redundant predictors. This step was done through heat maps which displays all the correlation between different features (all the predictors).

2.4.2 LSTM model calibration

Before running the model, some steps are required:

- ***Data normalization:***

  Since every feature has values with varying ranges, we do normalization to confine feature values to a range of [0, 1] before training a neural network. We do this by using equation 4.

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (4)$$

With:

- X: the value to be scaled or normalized;
- X_{max}: the maximum of X;
- X_{min}: the minimum of X;
- X_n: the normalized value of X.

- ***Data splitting:***

  We are tracking weather data from past 42 years (1979-2020) corresponding to a dataset of 15341 rows, on a daily basis, and 504 rows, on a monthly basis. This data will be used to predict the rainfall after one (1) to two(2) months ahead.
  80% of the data will be used to train and 10% to validate the model, then 10% is reserved for the testing purpose [18].

- ***Model implementation:***

  We were inspired by the official site of Kera (**https://keras.io**) to set up our model. It consisted of:

  - a process called hyperparameter optimization. This is done through a minimization function called optimizer. The optimizer used here is the Adaptive Moment Estimation (Adam) optimizer. Adam optimization is a stochastic gradient descent method that is based on the adaptive estimation of first-order and second-order moments;
  - as far as the activation function is concerned, the Rectified Linear Unit (ReLU) was used. In Artificial Neural Networks (ANNs), the activation function is a mathematical "gate" in between the input feeding the current neuron and its output going to the next layer [9]. The ReLU is the most commonly used activation function in DL. The function returns 0 if the input is negative, but for any positive input, it returns that value back;

– the rests of parameters like, the learning rate, the numbers of epoch and the loss function (also called training and validation curves) were determined before compiling the model. Figure 11 displays the qualities of a training and validation curves.

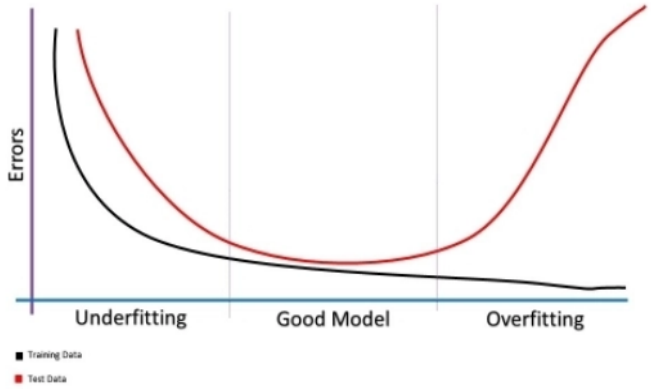


Figure 11: Description of validation and training loss curves. *Source: Baeldung (2020)*

The compilation was done for each weather station (**Appendix B** presents a summary of LSTM model per station) and the results are presented in the next chapter.

2.4.3 Performance measures of LSTM model

After training the model, it is ready to be used but before we have to test its performance. To achieve this, the test dataset was used in order to compare the observations and the model's predictions. Four (4) metrics were used to evaluate the models [28]: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Nash-Sutcliffe Efficiency (NSE) and coefficient of determination (R^2).

Let's assume Y_i the model prediction at time i, $\bar{Y_i}$ its average; X_i, the observation at the same time, $\bar{X_i}$ its average and n the number of observations. Then we have the following definitions:

- The MAE is the average of the absolute errors between model predictions and target values. It can be computed with the following equation 5 and ranges from 0 to ∞.

$$MAE = \frac{\sum_{i=1}^{n} |Y_i - X_i|}{n}, \qquad (5)$$

- The RMSE represents the error between model predictions and target values. It can be computed with the following equation 6 and ranges from 0 to ∞ too. The RMSE is frequently used to evaluate how closely the predicted values match the observed values, based on the relative range of the data.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i)^2}, \tag{6}$$

- The NSE was proposed by Nash and Sutcliffe in 1970. It can be defined as a measure of how the observed variance is simulated. It can be computed with the following equation 7. The NSE can range from $-\infty$ to 1. An efficiency $NSE = 1$ means a perfect prediction. An efficiency of 0 means that the model predictions are as reliable as the observed mean, while an efficiency $-\infty < NSE < 0$ indicates that the observed data mean is a better predictor than the model predictions.

$$NSE = 1 - \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}(X_i - \bar{X}_i)^2}, \tag{7}$$

- The R^2 is the proportion of the variance in the dependent variable that is predictable from the independent variables. It gives information about the length of the relationship between the models' predicted data and the observed data, and ranges from 0 to 1, with 1 being the best fit between predictions and actual data. It is computed with the equation 8.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}(X_i - \bar{Y}_i)^2}, \tag{8}$$

Chapitre 3

Results and discussion

This chapter presents the important findings of our study. These are, first of all, the results from the selection of relevant basins that are well correlated with SST. Then it will be about the result of the features selection which will make it possible to discard certain predictors that are redundant. Afterwards the results of the monthly, bimonthly and daily prediction of rainfall will be presented. Finally, the influence of topography and climatic zones on models' performances will be discussed.

3.1 Identification of relevant basins for SST selection

Since SST are oceanic data, if we want to use it as a predictor it worth identifying the different basins that influence the rainfall patterns over Burkina Faso. This identification was done through the correlation between each station rainfall and the SST. This made it possible to retain the basins which have a high correlation with the rainfall of the corresponding station (figure 12).

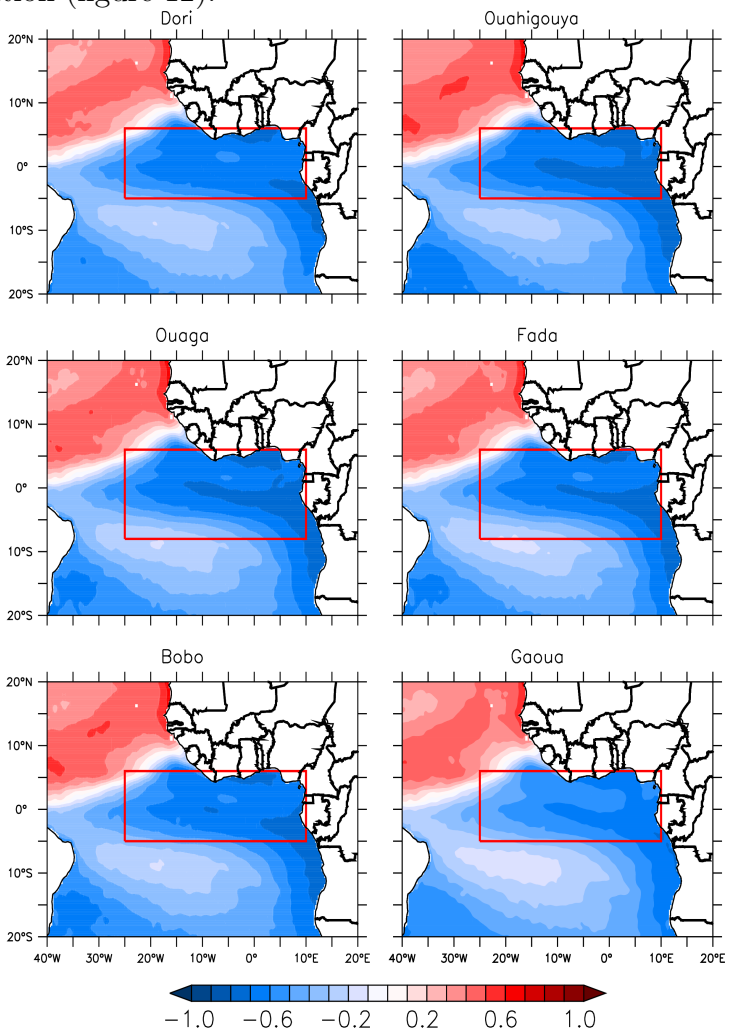


Figure 12: Correlation between rainfall of the six (6) weather stations and SST

The coordinates of the selected area for each stations are presented in table 3. Only high correlations of Golf of Guinea basins was considered [29] in order to take advantage of the southwesterly monsoon flux that convey moisture from the Atlantic Ocean to west Africa.

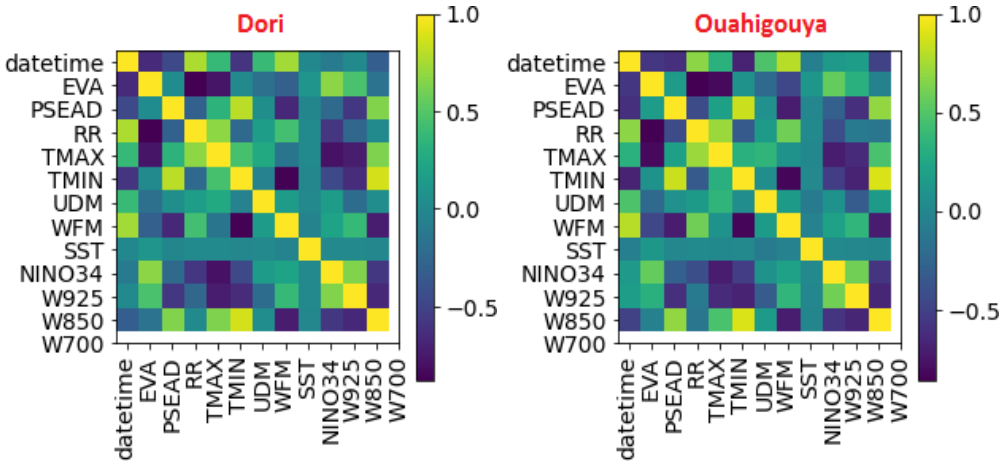
Table 3: Best correlated basins per station

Stations	**Selected basins**
Dori	5°S-6°N/025°W-010°E
Ouahigouya	5°S-6°N/025°W-010°E
Ouagadougou	8°S-6°N/025°W-010°E
Fada N'Gourma	8°S-6°N/025°W-010°E
Bobo-Dioulasso	5°S-6°N/025°W-010°E
Gaoua	5°S-6°N/025°W-010°E

Finally, the SST timeseries for each station were obtained by averaging the raw SST over these selected basins.

3.2 Features selection

We are trying to predict rainfall for each weather station based on historical (1979-2020) information of rainfall as well as eleven (11) others weather parameters. To avoid using redundant information, features selection is performed to discard some parameters that are highly correleted with the others. This is done through heat map (figure 13) which shows the correlation between all the features per station.



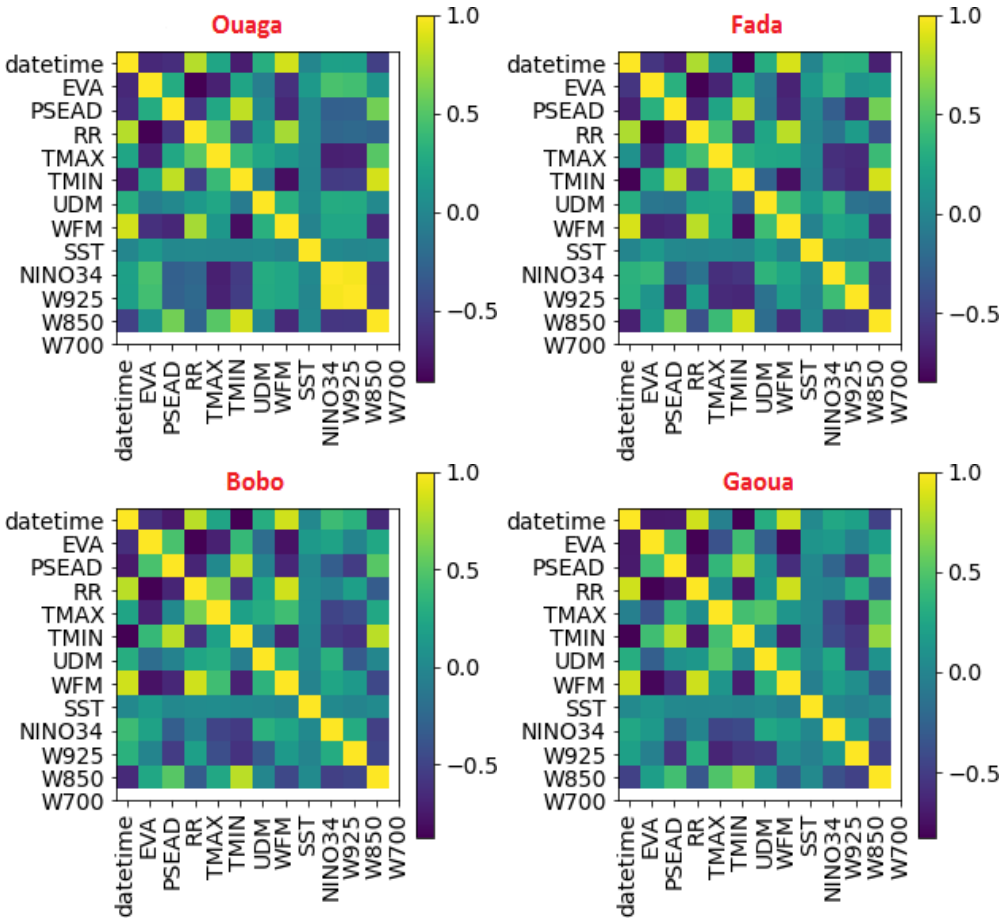


Figure 13: Feature correlation Heatmaps of the six (6) stations

We can see from the correlation heatmaps that few parameters like Minimum Temperature and wind speeds (surface and altitude) are redundant for almost all the stations. Since the parameter temperature is represented by the maximum temperature (Tmax) and the minimum temperature (Tmin), then we decided to do without the minimum temperature to avoid this redundancy revealed by this heat maps

3.3 Monthly rainfall prediction

The optimum numbers of hidden layers, numbers of epochs, and learning rates are determined by conducting a series of trials. Numerous combinations of varying numbers of these parameters are tested until a stable Training and Validation Loss (the loss function over the number of epochs as shown in figure 14) curve is obtained for each station. The training loss indicates how well the model is fitting the training data, while the validation loss indicates how well the model fits new data.

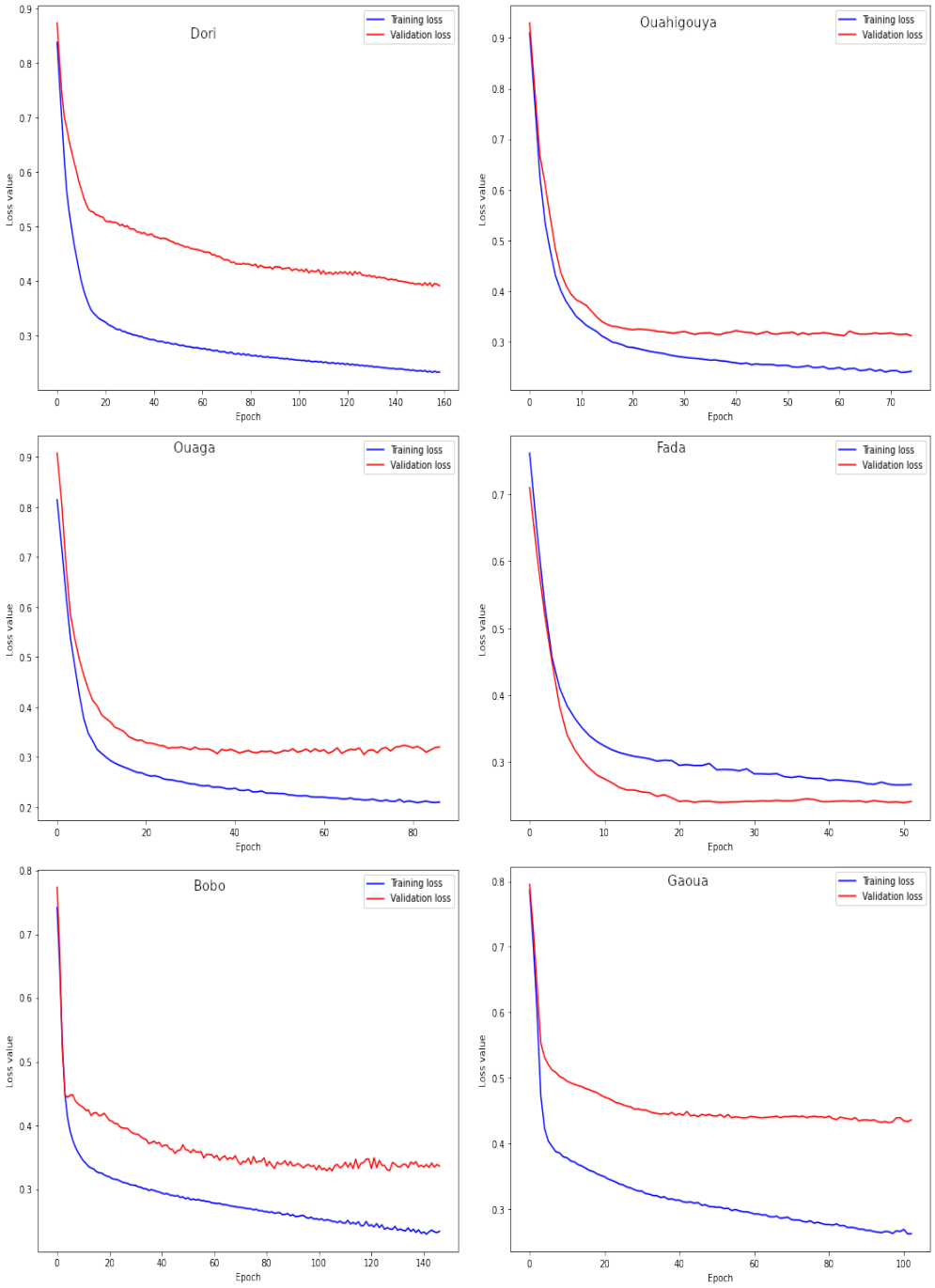
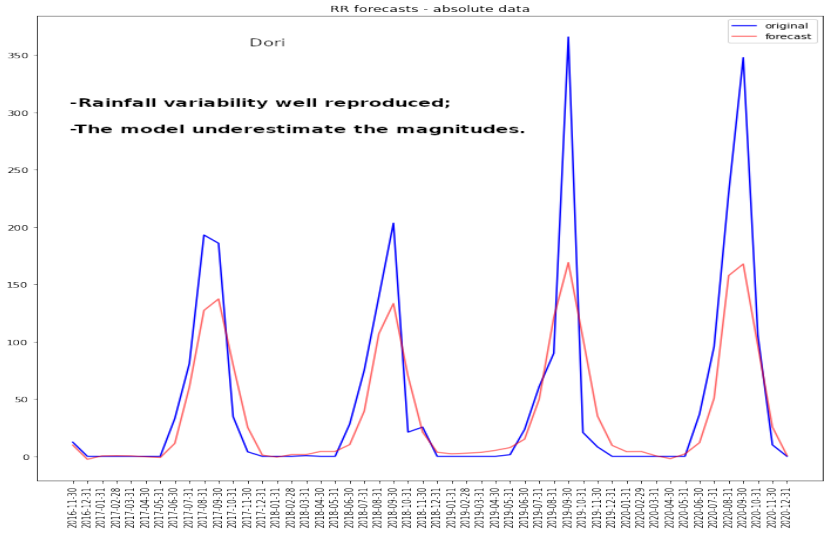
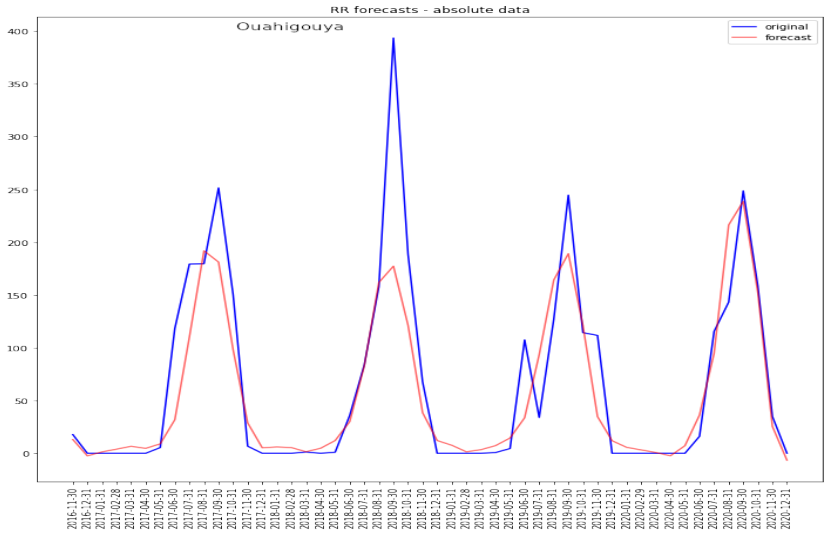


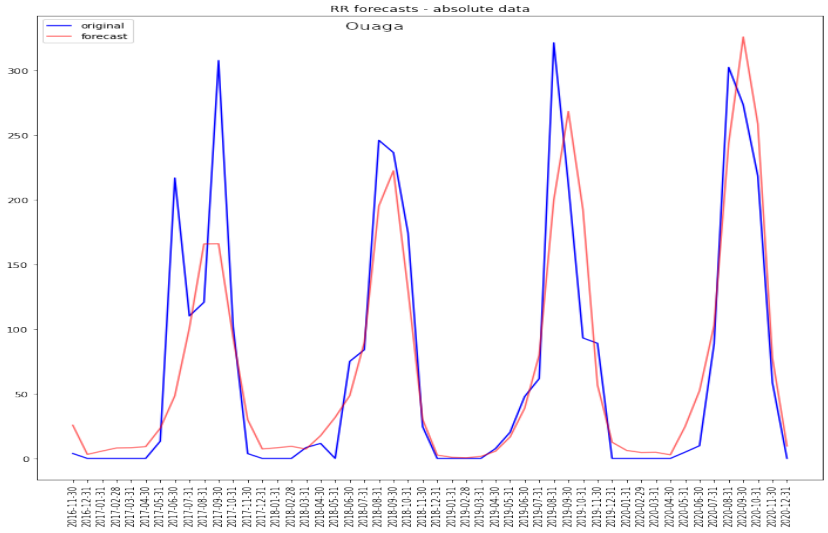
Figure 14: Training and validation loss of monthly rainfall prediction per station

The analysis of these training and validation loss curves reveals that the models are neither overfitting nor underfitting. However we wish we could have lessen the losses (errors) for example under 0.5 for all the stations.

Afterwards, the models are evaluated by comparing predictions from the models and observations over the test period (2016-2020) as featured in figure 15. We can see that the models predictions reasonably follow the series patterns at all locations even though the magnitudes are not well captured.


RR forecasts - absolute data
Dori
original
forecast
-Rainfall variability well reproduced;
-The model underestimate the magnitudes.


RR forecasts - absolute data
Ouahigouya
original
forecast


RR forecasts - absolute data
Ouaga
original
forecast

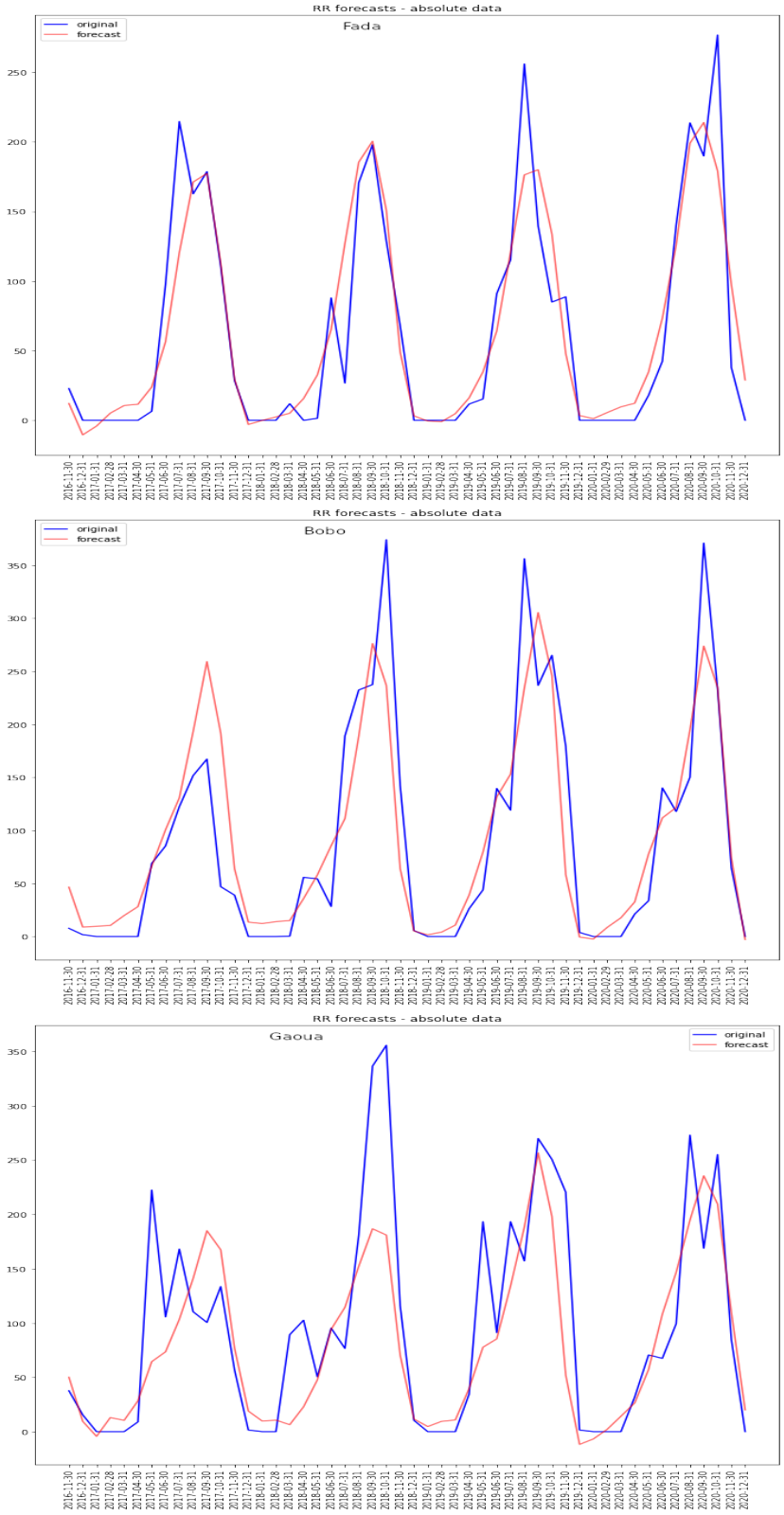


Figure 15: Observed rainfall vs. model prediction (in millimeter) per station

To go further in the evaluation, the regression plots or scattergrams (figure 16) were used to determine the correlation between the predicted rainfall and the observed rainfall, always based on the test dataset.

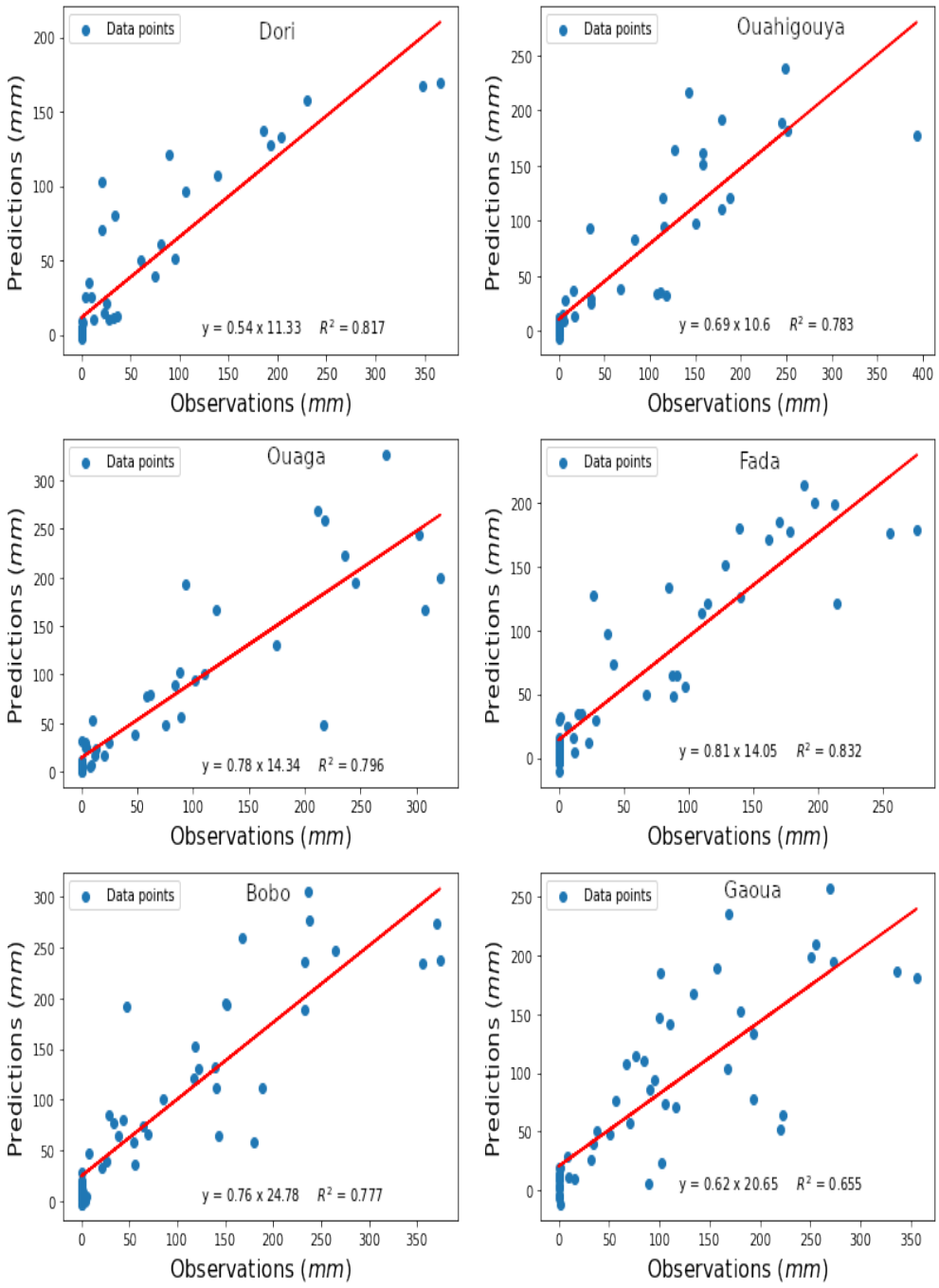


Figure 16: Scattergram per station

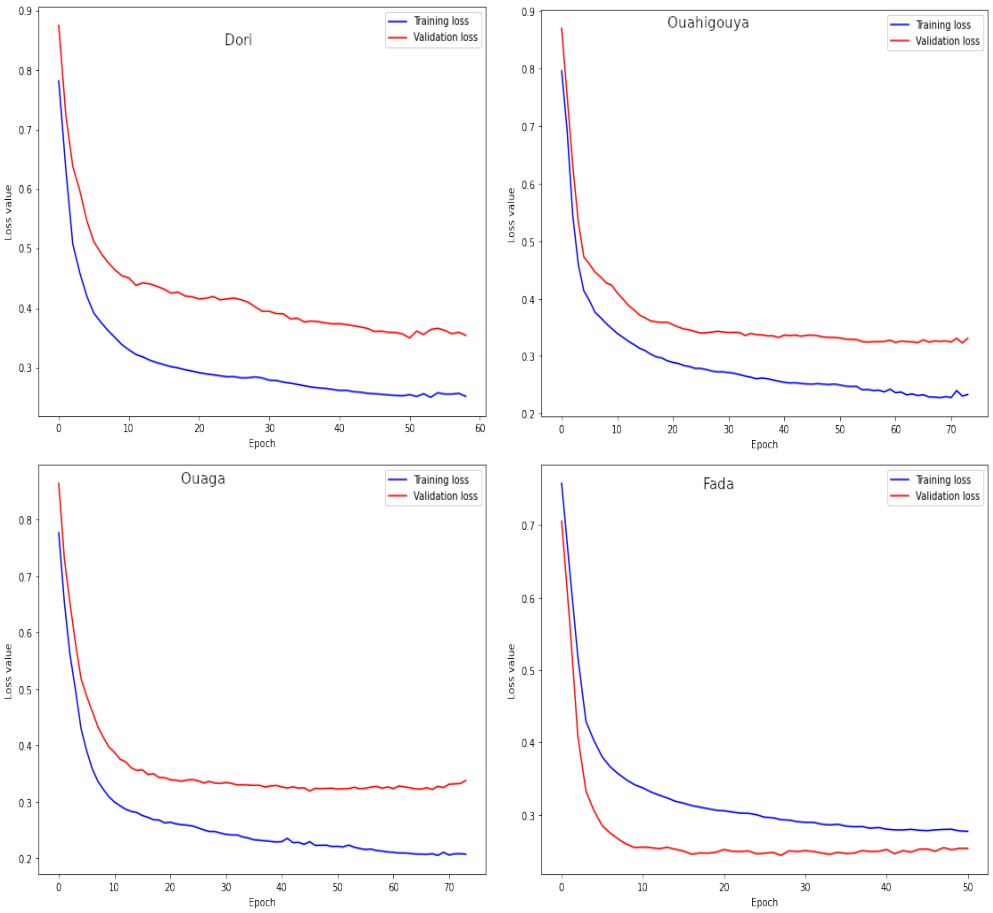
We can see that the coefficient of determination R^2 is greater than 0.65 for all the stations; which is not bad. For instance, Dori and Fada have the best coefficient of determination (more than 0.8). Table 5 summarizes all the prediction performances of the models.

Table 4: Models' prediction performances on test set (Period 2016-2020)

Stations	**NSE**	R^2	**RMSE (mm)**	**MAE (mm)**
Dori	0.710	0.817	46.5	24.1
Ouahigouya	0.758	0.783	44.6	24.5
Ouagadougou	0.800	0.809	43.9	24.2
Fada N'Gourma	0.831	0.832	32.9	21.1
Bobo-Dioulasso	0.776	0.777	50.2	33.4
Gaoua	0.624	0.655	59.9	39.7

3.4 Bimonthly rainfall prediction

After one (1) month ahead rainfall prediction, we performed a multi-step prediction to get rainfall two (2) months ahead. For this case again, several trials were conducted to get the optimized hyperparameters (numbers of hidden layers, numbers of epochs, and learning rates) that provide a stable training and validation loss curve for each station (as shown in figure 17).



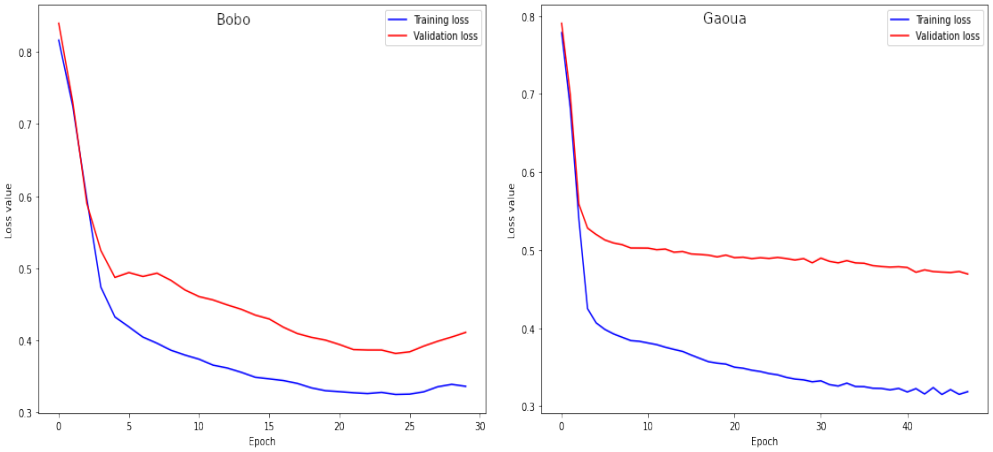
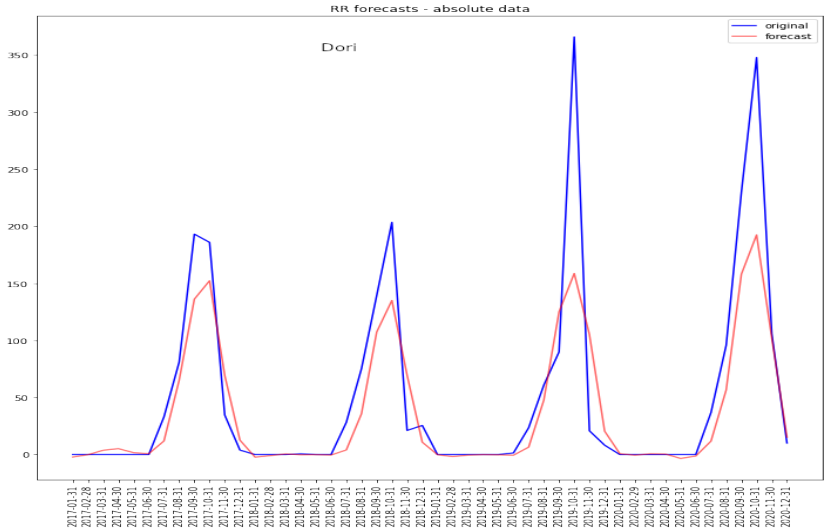
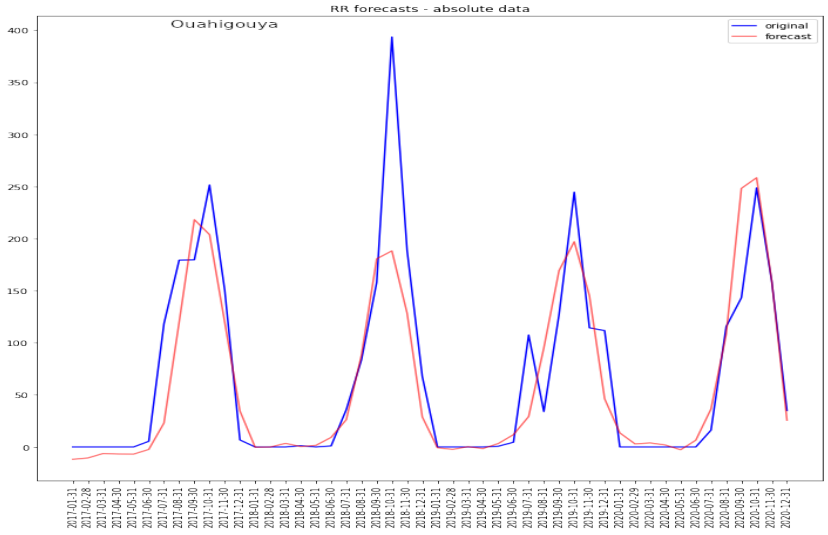


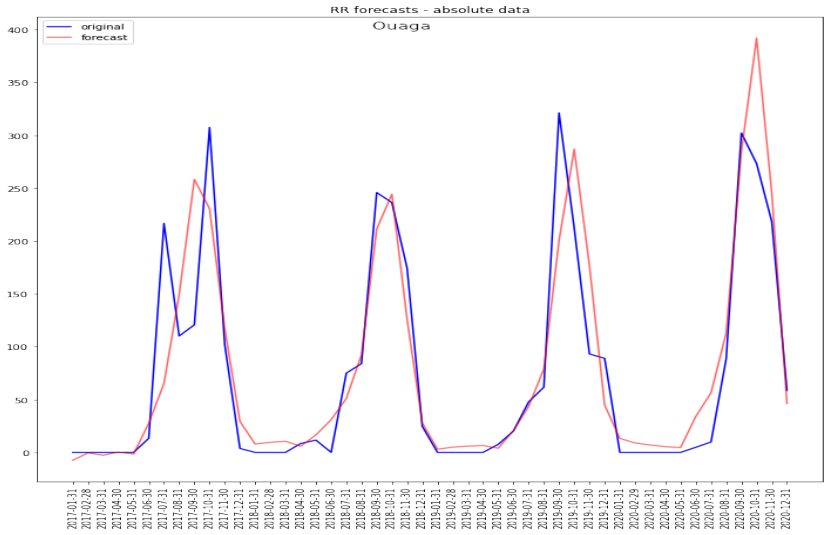
Figure 17: Training and validation loss of bimonthly rainfall prediction per station

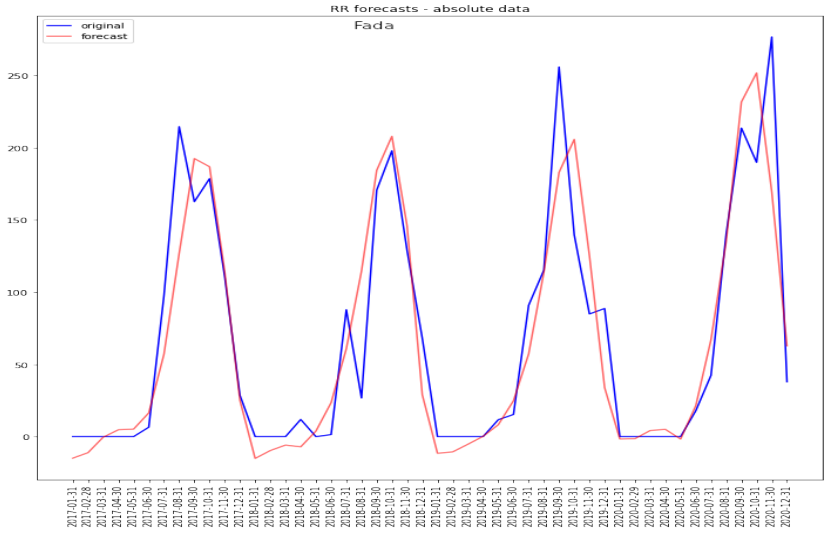
From these training and validation loss curves, one can notice that, for this case too, the models are neither overfitting nor underfitting.

The models were then evaluated by comparing predictions and observations over the test period (2016-2020) as shown in figure 18. Here again one can see the models predictions reasonably follow the series patterns; the problem lies in the magnitude that are not well captured as well.




RR forecasts - absolute data
Ouahigouya
original
forecast


RR forecasts - absolute data
Ouaga
original
forecast


RR forecasts - absolute data
Fada
original
forecast

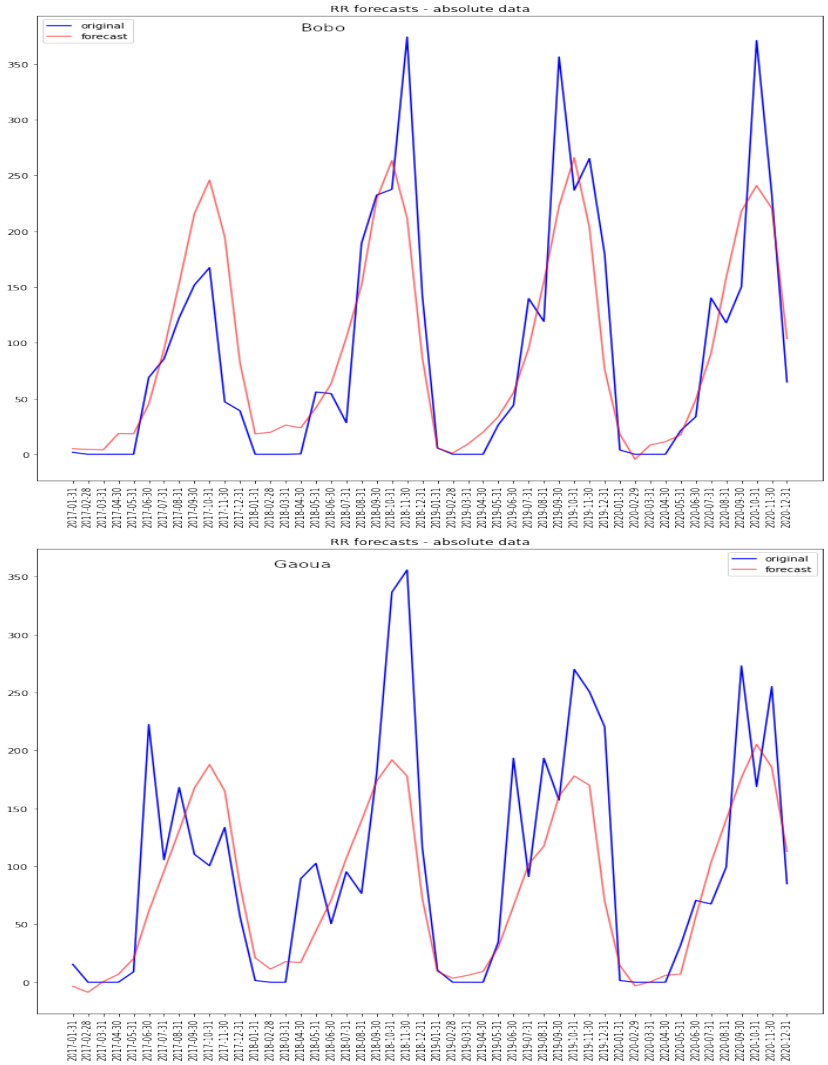
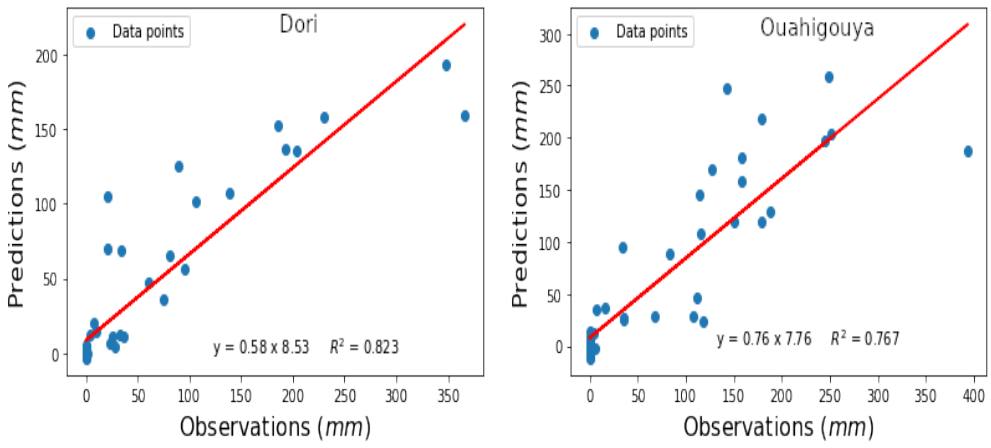


Figure 18: Observed rainfall vs. model prediction (in millimeter) per station (bimonthly prediction)

Here again, the regression plots (figure 19) were used to determine the correlation between the predicted rainfall and the observed rainfall, always based on the test dataset.



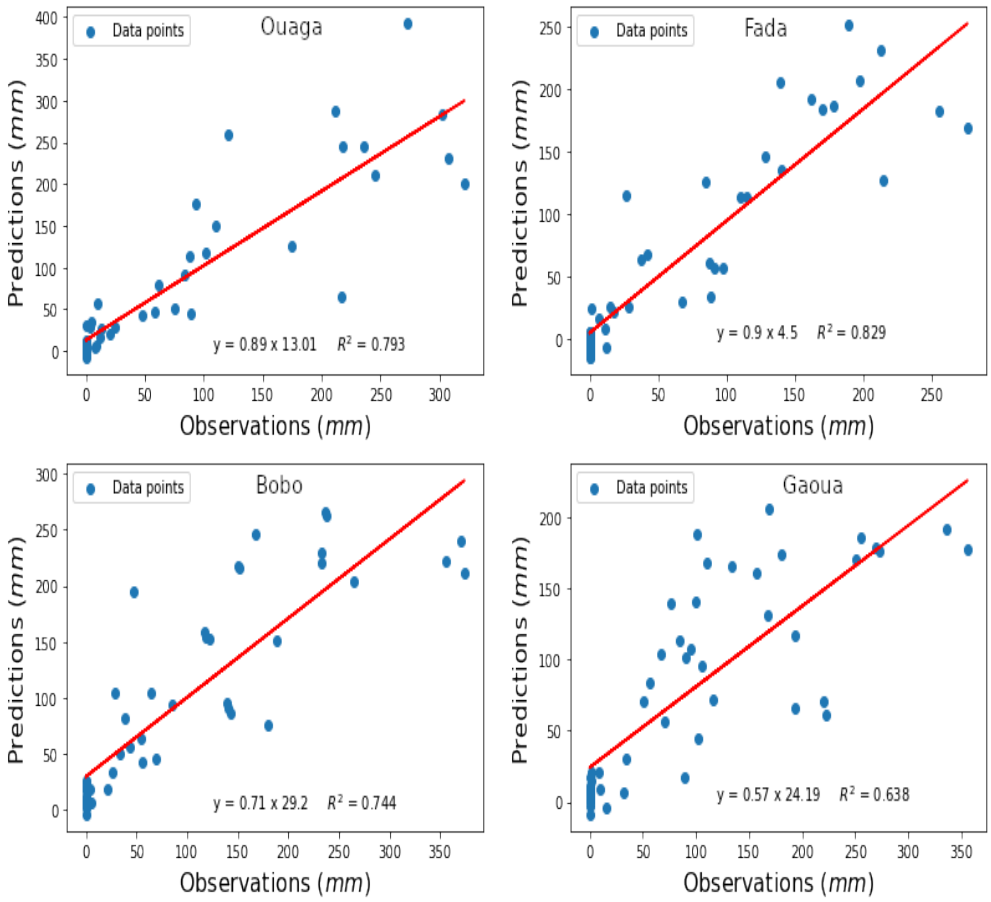


Figure 19: Scattergram per station (bimonthly prediction)

For this case again, one can notice that the coefficient of determination R^2 is greater than 0.6 for all the stations. For instance, Dori and Fada have the best coefficient of determination (more than 0.8) as well. Table 5 summarizes all the prediction performances of the models.

Table 5: Models' prediction performances on test set (Period 2016-2020)

Stations	**NSE**	R^2	**RMSE (mm)**	**MAE (mm)**
Dori	0.729	0.823	45.7	21.8
Ouahigouya	0.759	0.767	45.2	26.1
Ouagadougou	0.780	0.784	46.6	27.7
Fada N'Gourma	0.823	0.829	34.0	22.2
Bobo-Dioulasso	0.742	0.744	54.2	40.8
Gaoua	0.591	0.638	62.8	42.8

3.5 Daily rainfall prediction

In the previous sections, we saw how well the models were capable to predict rainfall one (1) month and two (2) months ahead. In this section, we will see how the models behave when it comes to discretize these monthly accumulations into 30 daily rainfalls. As said in the methods section, it is good to have the accumulated monthly prediction, but it is better to know the distribution of this monthly amount of rainfall in daily timescale. The figure 20 shows the models attempts to capture daily rainfalls within a specific month.

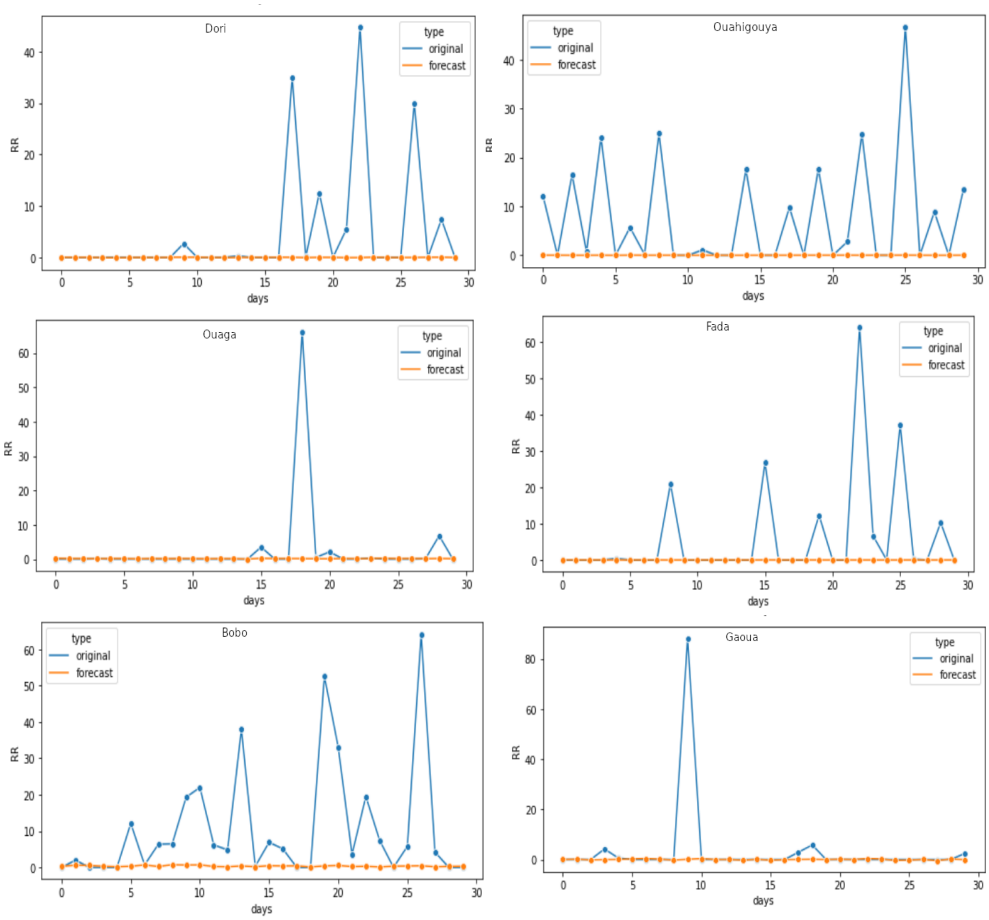


Figure 20: Thirty (30) days multi-step rainfall prediction per station

By analyzing the figure above, the models are not doing well. Unlike the monthly timescale, the models struggled to reproduce daily rainfall in all locations. This may reflect the limits between statistical models and dynamic models. To solve this problem, many authors [30],[31] suggest the combination of NWP and AI. This requires more time as well as more resources to investigate this matter as part of future studies.

3.6 Impact of climatic zones and topography on models' performances

By analyzing the performance of the models at the monthly and bimonthly scale, one can notice that the best results of the models were obtained in the Sudano-Sahelian zone (cf. table 6 obtained by averaging performance measures for the 2 stations of each climatic zone). Then comes, in terms of good results, the Sahelian zone and finally the Sudanian zone. It should be noted that this last zone, in terms of topography, is the highest of the country with several elevations including the highest peak, Mount Tenakourou. These results are in agreement with those obtained by Arsène Aizansi [23] who had shown that ANN had problems in very rainy and mountainous areas in Benin republic. However, this requires more investigation to find out why the models are good in the plateaus than in the elevated regions. But as we know, elevated areas are very rainy areas and we have seen that most models have problems capturing the magnitudes of rainfall.

Table 6: Models' performances regarding the climatic zones and topography

Forecast ranges	**Zones**	**NSE**	R^2	**RMSE (mm)**	**MAE (mm)**
Monthly	Sahelian	0.734	0.800	45.5	24.3
	Sudano-Sahelian	0.815	0.820	38.4	22.6
	Sudanian	0.700	0.716	55.0	36.5
Bimonthly	Sahelian	0.744	0.795	45.5	23.9
	Sudano-Sahelian	0.801	0.806	40.3	24.9
	Sudanian	0.666	0.691	58.5	41.8

Conclusion and perspectives

This study aimed to provide, at the sub-seasonal to seasonal (S2S) time scale, a prediction of rainfall over Burkina Faso using Deep Learning techniques. Specifically, it investigated the potential of using Long Short Term Memory (LSTM) networks to simulate rainfall at six (6) meteorological stations (Dori, Ouahigouya, Ouagadougou, Fada N'Gourma, Bobo-Dioulasso and Gaoua) at this time scale, which is not yet covered by traditional weather forecasts in the country. The LSTM model has been fed with historical data from the six (6) location, including precipitation, evaporation, pressure, wind speed, minimum and maximum temperatures, sea surface temperature.

To achieve this, the first step consisted in data preparation to meet the model requirements. This was done through "data pre-processing" starting by feature selections and ending up by data normalization. Afterwards, the optimal hyperparameters (learning rate, number of units, number of epochs and batch size) were selected for each location. Finally, the model was run and evaluated per location using some statistical metrics.

The results showed that LSTM model was effective and performed well with acceptable scores. For instance, forecasting at monthly timescale exibited a R^2 ranging from 0.66 to 0.83, NSE ranging from 0.62 to 0.80, RMSE ranging from 32.9 to 59.9mm and MAE ranging from 21.1 to 39.7mm. Regarding the bimonthly rainfall prediction, R^2 ranged from 0.63 to 0.83, NSE ranged from 0.6 to 0.82, RMSE ranged from 34.0 to 62.8mm and MAE ranged from 21.8 to 42.8mm.

Indeed, it turned out that the models reasonably followed the rainfall patterns at all locations. However, the model struggled to capture the magnitude of rainfall both monthly and bimonthly. Still at this monthly and bimonthly scale, we have seen the impact that the climatic zone and the topography have on the performance of the models. Indeed, the models had better results on slightly wet plateaus than on rainy and elevated areas.

We have also tried to bring these monthly forecasts down to the daily time scale; however the model was not able to capture daily rainfall pattern which turned out to be the most important information for some end users (farmers for instance). To solve this problem, it may require to investigate further, for instance, toward the combination between the state of the art of weather forecasting (Numerical Weather Prediction) and Deep Learning

techniques. However, this investigation goes beyond the scoop of this work.

Before finishing, Let us recall the main hypothesis of this work: **LSTM model used to predict S2S rainfall over Burkina Faso gives good scores**.

In the light of the results of this study, we could assert that these hypothesis have been confirmed although there is room for improvement. Nevertheless, we think we have paved the way allowing other students to continue this work (why not as part of a PhD), the usefulness of which is no longer to be demonstrated, given that the Artificial Intelligence, in general, and Deep Learning, in particular, is a topic of the hour backed currently by the World Meteorological Organization (WMO) [32]. This work could help the scientific community as well as the services competent in climate matters (ANAM, ASECNA) of Burkina Faso to gradually integrate this technique into their sovereign weather forecasting activities.

Bibliography

[1] **IPCC, (2022).** *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)].* Cambridge University Press. In Press.
https://www.ipcc.ch/report/ar6/wg2/about/how-to-cite-this-report/

[2] **Vellinga, P., Verseveld, W. J. (2000).** *Climate change and extreme weather events. WWF.*

[3] **Dardel et al (2013).** *The 25 years long drought in sahel and its impacts on ecosystems: long term vegetation monitoring from the sky and on the ground* ResearchGate.

[4] **Olaniyan et al. (2018).** *Evaluation of the ECMWF Sub-seasonal to Seasonal Precipitation Forecasts during the Peak of West Africa Monsoon in Nigeria*, Front. Environ. Sci.

[5] **Lazo et al. (2011).** *U.S. Economic sensitivity to weather variability.* Bull. A. Meteor. Soc. 92, 709–720. doi: 10.1175/2011BAMS2928.1.

[6] **Vitart et al. (2014).** *Sub-seasonal Predictions. European Centre for Medium Range Weather Forecasts* Technical Memorandum No., 738.

[7] **DE et al. (1986).** *Learning Internal Representations by Error Propagation*, pp. 318–362. Cambridge, MA: MIT Press. Google Scholar

[8] **Jason Brownlee (2017).** *The Promise of Recurrent Neural Networks for Time Series Forecasting*

[9] **Francois CHOLLET (2021).** *Deep Learning with Python Second edition*, pp.87-100

[10] **Beucher, F. (2010).** *Météorologie Tropicale: des alizés au cyclone.* Météo-France.

[11] **T. Kim and J. B. Valde (2020).** *Using Artificial Neural Networks for Generating Probabilistic Subseasonal Precipitation Forecasts over California* Monthly Weather Review, Volume 148: Issue 8, pp. 3489–3506.

[12] **Guoxing Chen and Wei-Chyung Wang (2021)** *Short-term precipitation prediction using deep learning* Cornell university. arXiv:2110.01843v1

[13] **Sijie et al. (2021).** *Sub-Seasonal Climate Forecasting via Machine Learning: Challenges, Analysis, and Advances* pp. 169.

[14] **Narejo S, Jawaid MM, Talpur S, Baloch R, Pasero EGA. (2021).** *Multi-step rainfall forecasting using deep learning approach.* PeerJ Comput. Sci. 7:e514 DOI 10.7717/peerj-cs.514

[15] **Ouma, Y.O., Cheruyot, R. Wachera, A.N. (2022).** *Rainfall and runoff time-series trend analysis using LSTM recurrent neural network and wavelet neural network with satellite-based meteorological data: case study of Nzoia hydrologic basin.* Complex Intell. Syst. 8, 213–236.
https://doi.org/10.1007/s40747-021-00365-2

[16] **Chen et al.(2022).** *Forecast of rainfall distribution based on fixed sliding window long short-term memory*, Engineering Applications of Computational Fluid Mechanics, 16:1, 248-261, DOI: 10.1080/19942060.2021.2009374

[17] **Barrera-Animas et al. (2022)**, *Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting*, Machine Learning with Applications, Volume 7, 100204, ISSN 2666-8270,
https://doi.org/10.1016/j.mlwa.2021.100204.

[18] **Endalie et al. (2022).** *Deep learning model for daily rainfall prediction: case study of Jimma, Ethiopia*, Water Supply (2022) 22 (3): 3448–3461.
https://doi.org/10.2166/ws.2021.391

[19] **Anochi et al. (2021).** *Machine Learning for Climate Precipitation Prediction Modeling over South America.* Remote Sens, 13, 2468.

[20] **Getachew Mehabie Mulualem and Yuei-An Liou (2020).** *Application of Artificial Neural Networks in Forecasting a Standardized Precipitation Evapotranspiration Index for the Upper Blue Nile Basin* Water.

[21] **Badr et al.(2020).** *Applications of Deep Learning to S2S Precipitation Prediction and Downscaling for the Middle East and North Africa* Americain Meteorological society.

[22] **Adjei et al. (2021).** *Rainfall Forecasting in Sub-Sahara Africa-Ghana using LSTM Deep Learning Approach* International Journal of Engineering Research Technology (IJERT)
http://www.ijert.org

[23] **Arsène Aizansi (2021).** *Monthly Rainfall Prediction Using Artificial Neural Network (Case Study: Republic of Benin).* Master Thesis, University Joseph Ki-ZERBO, pp. 90

[24] **Niang et al. (2020).** *Transport pathways across the West African Monsoon as revealed by Lagrangian Coherent Structures.* Sci Rep 10, 12543.
https://doi.org/10.1038/s41598-020-69159-9

[25] **Nicholson, S.E., (2013).** *The West African Sahel: A review of recent studies on the rainfall regime and its interannual variability.* International Scholarly Research Notices.

[26] **Troisième communication nationale du Burkina Faso: étude vulnérabilité/adaptation au changement climatique (2021).** *Rapport final du sous-goupe "étude climatique"*

[27] **Peter Molnar and Balaji Rajagopalan (2020).** *Mid-Holocene Sahara-Sahel Precipitation From the Vantage of Present-Day Climate*, Geophysical Research Letter.
https://doi.org/10.1029/2020GL088171

[28] **Samad et al. (2020).** *An Approach for Rainfall Prediction Using Long Short Term Memory Neural Network*, IEEE 5th International Conference on Computing Communication and Automation (ICCCA), pp. 190-195, doi: 10.1109/ICCCA49541.2020.9250809.

[29] **Teresa et al. (2012).** *Tropical SST and Sahel rainfall: A non-stationary relationship.* Geophysical Research Letters, American Geophysical Union, 39 (12), pp.L12705. 10.1029/2012gl052423. hal-01495119

[30] **Schultz M. G. et al. (2021).** *Can deep learning beat numerical weather prediction?* Phil. Trans. R. Soc. A.3792020009720200097
http://doi.org/10.1098/rsta.2020.0097

[31] **Philipp Hess and Niklas Boers (2021).** *Deep Learning for Improving Numerical Weather Prediction of Rainfall Extremes* 7 Convention Center - Poster Hall, D-
https://agu.confex.com/agu/fm21/meetingapp.cgi/Paper/885371 *Climate Change and Extreme Weather Events*, WWF, September 2000.

[32] **Mary Glackin (2019).** *Improving the Forecast: Value and Public-Private Collaboration in Data Driven Weather Insights* Bulletin n° : Vol 68 (1)

Appendix

Appendix A: Raw data visualization per station

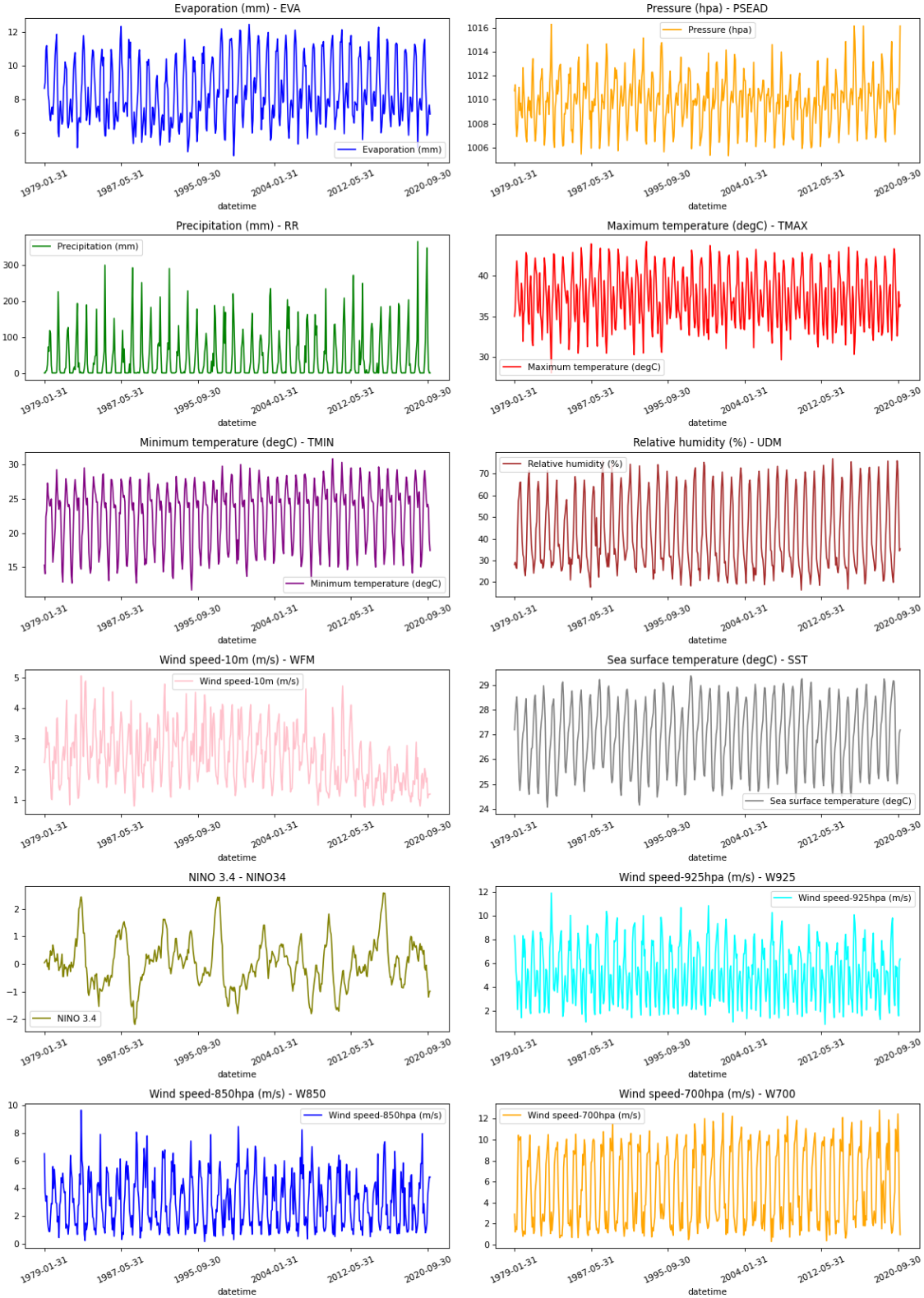


Figure 21: Raw data visualization of Dori

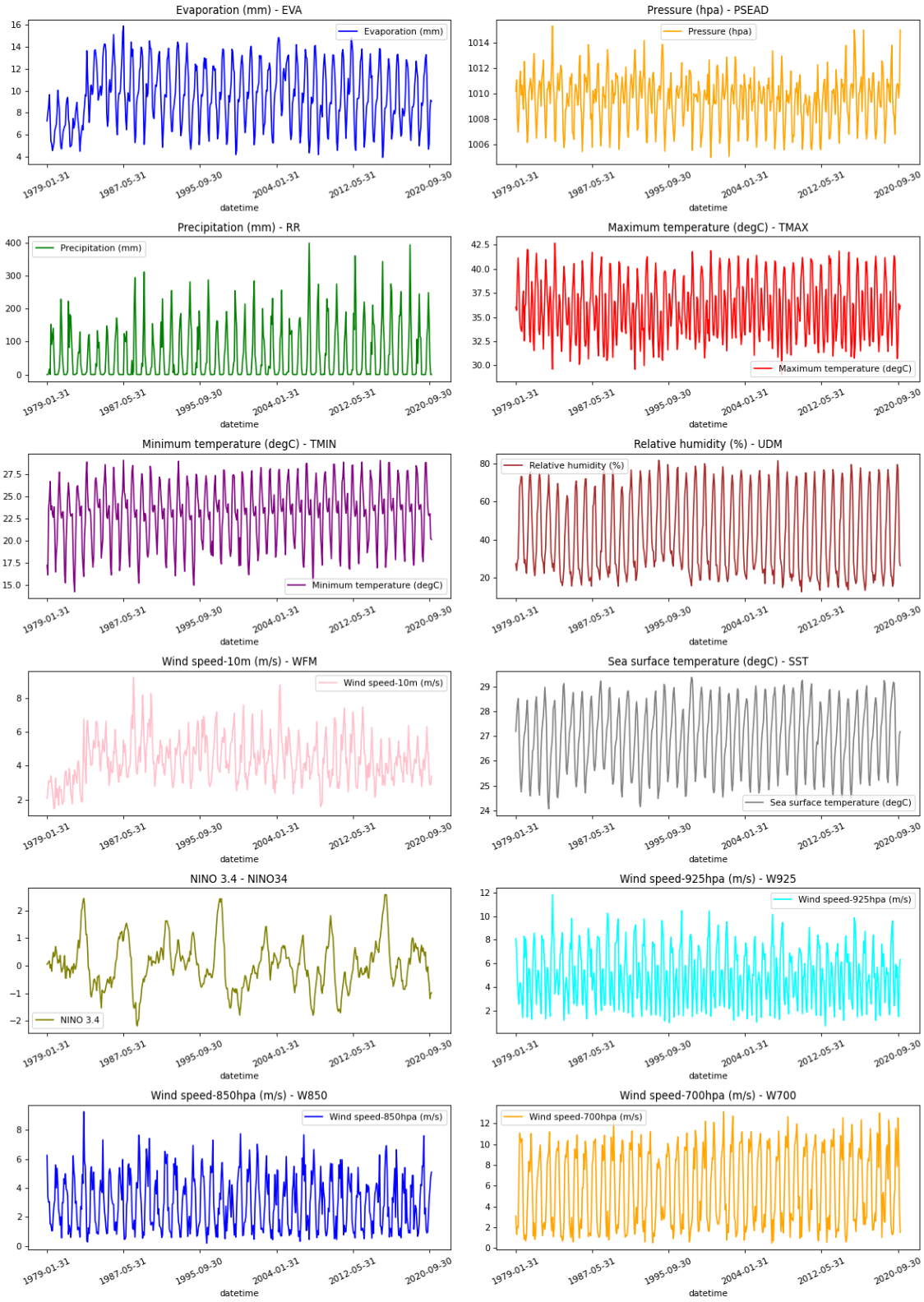


Figure 22: Raw data visualization of Ouahigouya

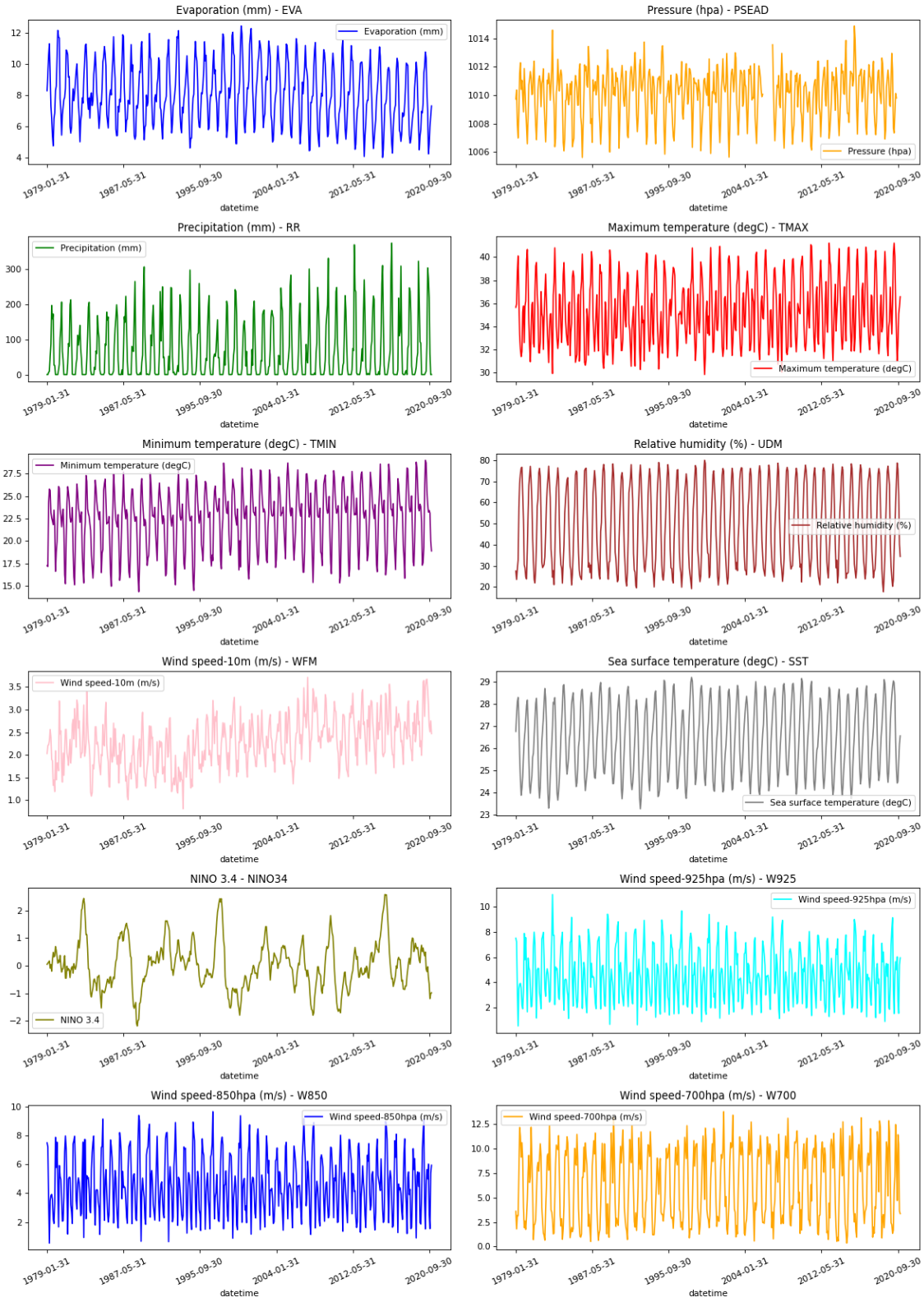


Figure 23: Raw data visualization of Ouagadougou

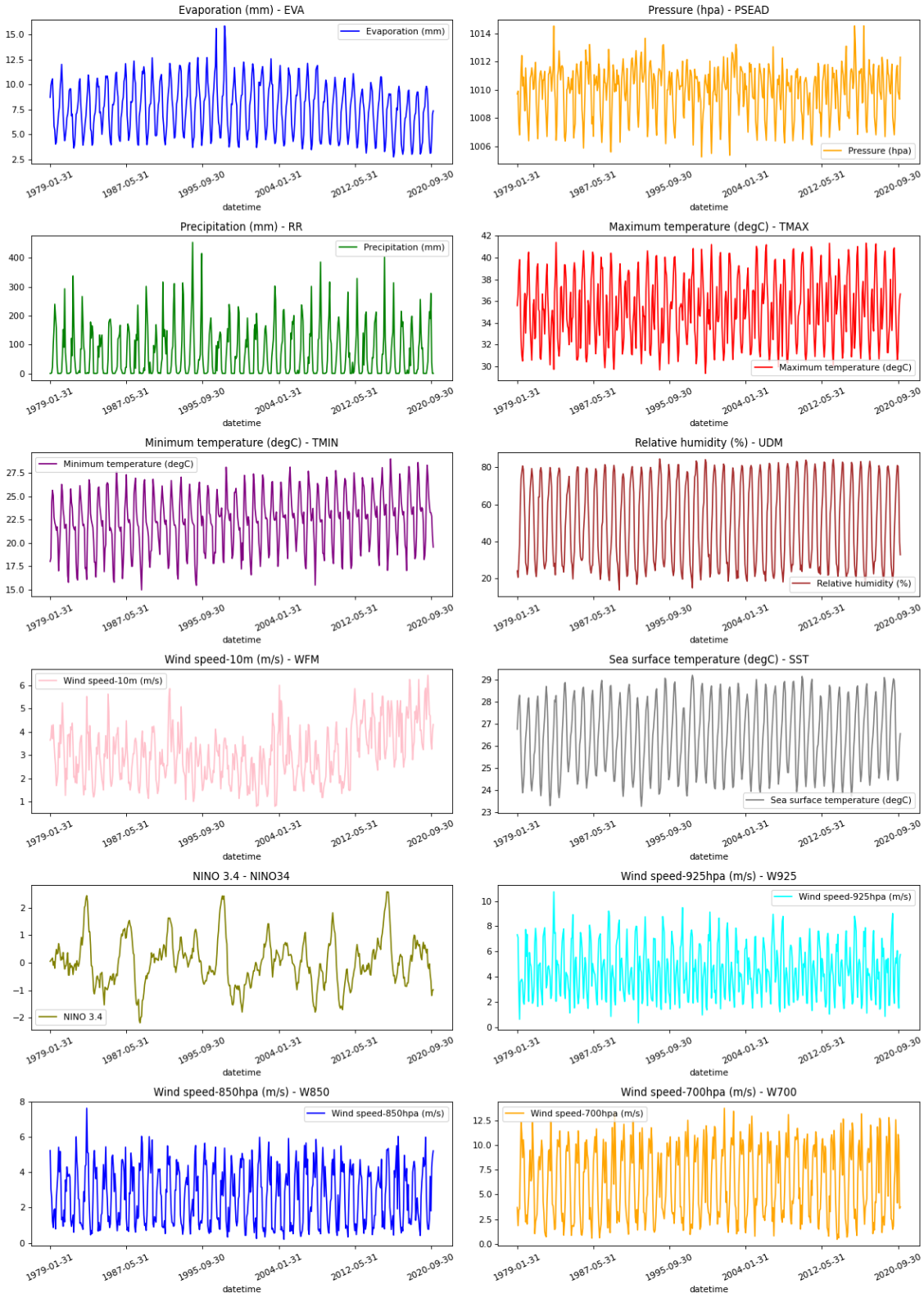


Figure 24: Raw data visualization of Fada

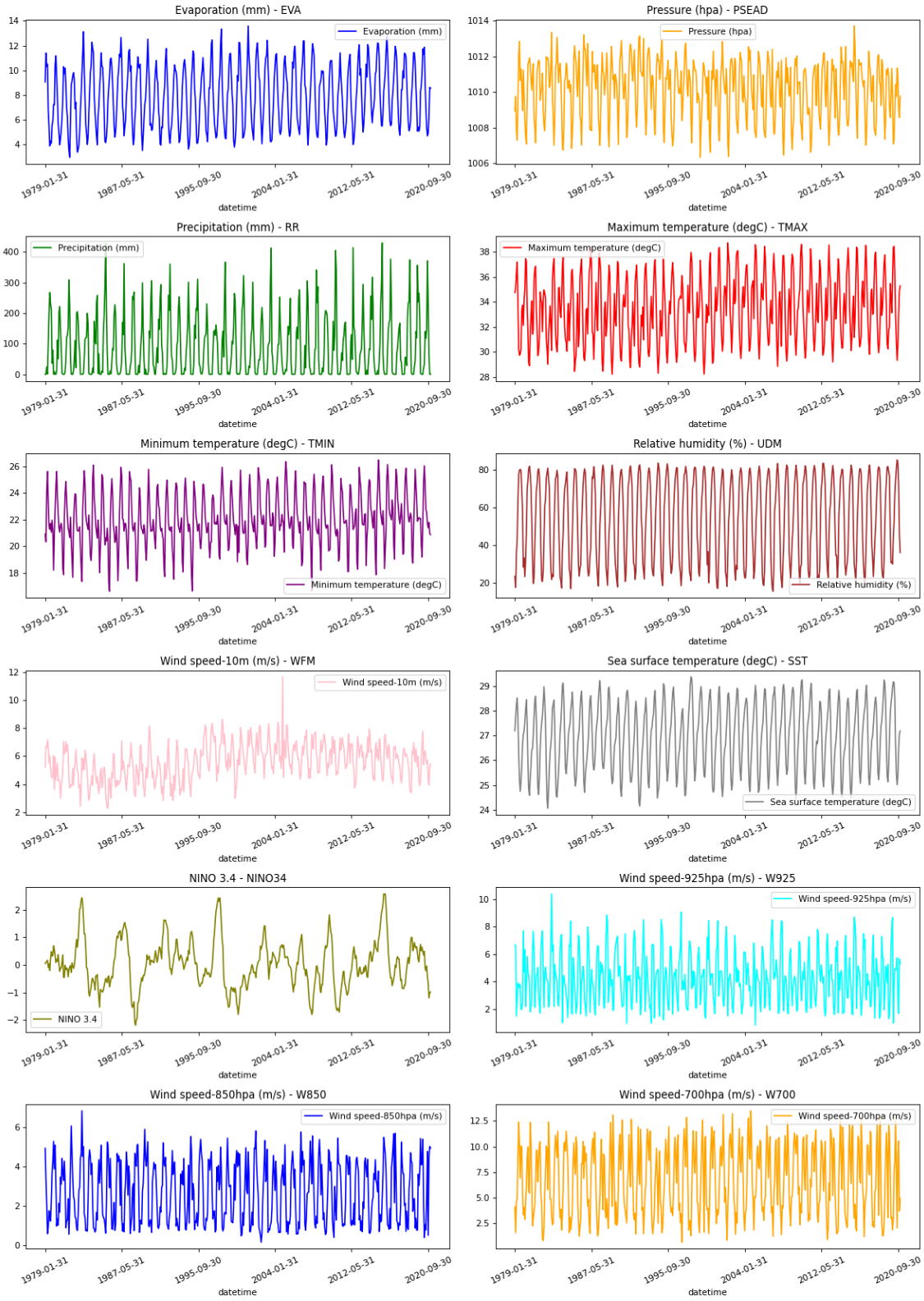


Figure 25: Raw data visualization of Bodo

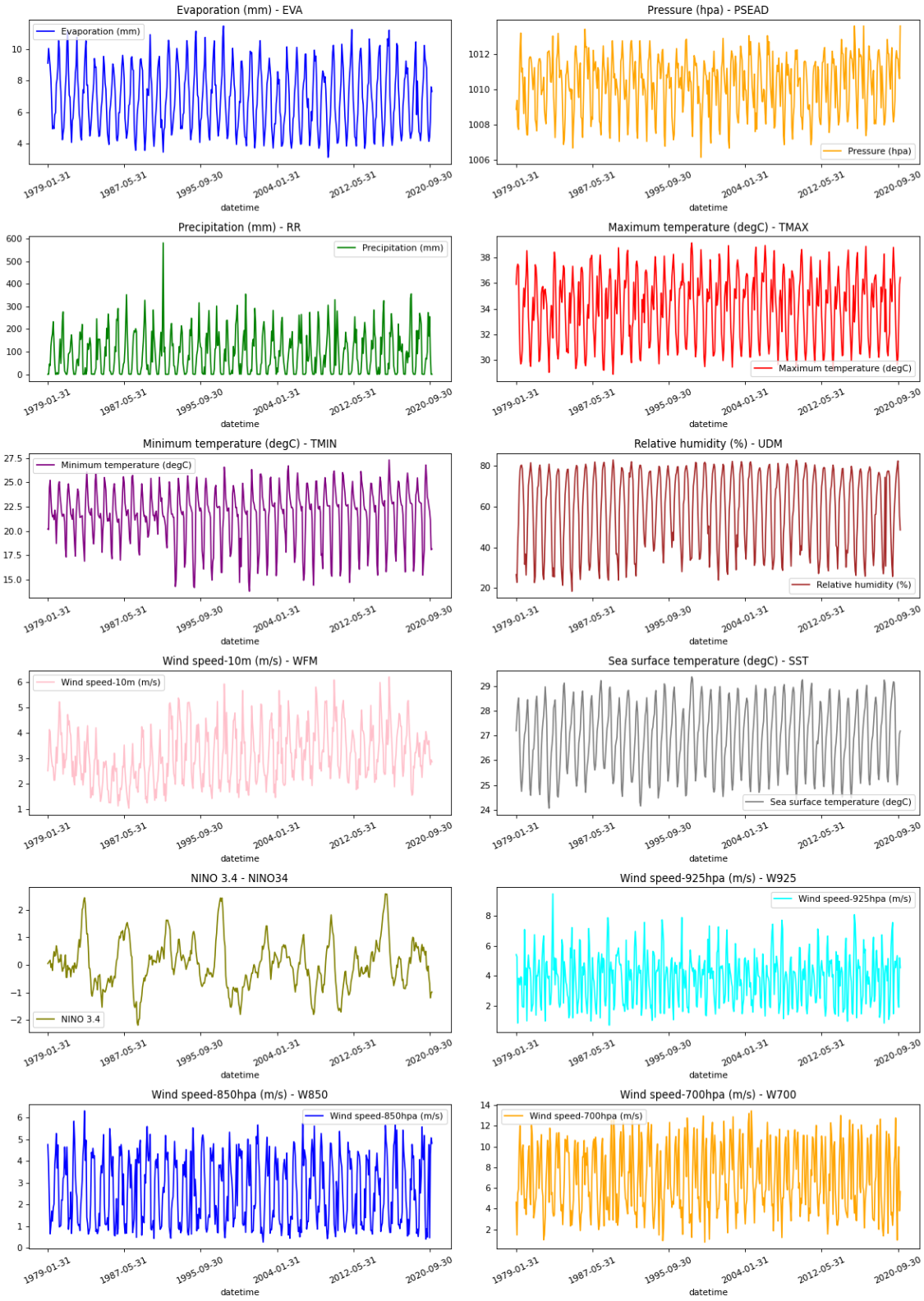


Figure 26: Raw data visualization of Gaoua

Appendix B: Model summary per station

Dori

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_17 (InputLayer)        [(None, 12, 12)]          0
_________________________________________________________________
lstm_16 (LSTM)               (None, 32)                5760
_________________________________________________________________
dense_16 (Dense)             (None, 1)                 33
=================================================================
Total params: 5,793
Trainable params: 5,793
Non-trainable params: 0
_________________________________________________________________
```

Ouahigouya

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_19 (InputLayer)        [(None, 12, 12)]          0
_________________________________________________________________
lstm_18 (LSTM)               (None, 70)                23240
_________________________________________________________________
dense_18 (Dense)             (None, 1)                 71
=================================================================
Total params: 23,311
Trainable params: 23,311
Non-trainable params: 0
_________________________________________________________________
```

Ouagadougou

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_2 (InputLayer)         [(None, 12, 12)]          0
_________________________________________________________________
lstm_2 (LSTM)                (None, 70)                23240
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 71
=================================================================
Total params: 23,311
Trainable params: 23,311
Non-trainable params: 0
_________________________________________________________________
```

Fada N'Gourma

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         [(None, 12, 12)]          0
_________________________________________________________________
lstm (LSTM)                  (None, 70)                23240
_________________________________________________________________
dense (Dense)                (None, 1)                 71
=================================================================
Total params: 23,311
Trainable params: 23,311
Non-trainable params: 0
_________________________________________________________________
```

Bobo Dioulasso

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_23 (InputLayer)        [(None, 12, 12)]          0
_________________________________________________________________
lstm_22 (LSTM)               (None, 70)                23240
_________________________________________________________________
dense_22 (Dense)             (None, 1)                 71
=================================================================
Total params: 23,311
Trainable params: 23,311
Non-trainable params: 0
_________________________________________________________________
```

Gaoua

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_27 (InputLayer)        [(None, 12, 12)]          0
_________________________________________________________________
lstm_26 (LSTM)               (None, 70)                23240
_________________________________________________________________
dense_26 (Dense)             (None, 1)                 71
=================================================================
Total params: 23,311
Trainable params: 23,311
Non-trainable params: 0
_________________________________________________________________
```

Figure 27: Model summary per station

Table of contents