



**INTERNATIONAL INSTITUTE FOR WATER AND ENVIRONMENTAL
ENGINEERING (2iE)**

UNIVERSITY OF OTTAWA

THESIS UNDER JOINT SUPERVISION

For the award of:

**DOCTOR OF SCIENCE AND TECHNOLOGY IN WATER, ENERGY AND
ENVIRONMENT (2iE)**

Ph.D. IN CIVIL ENGINEERING (uOttawa)

Speciality: Water

by

Abdouramane Gado Djibo

June 28, 2016

Ref.: 2iE/2016-07

Title

**Exploration of Non-Linear and Non-Stationary Approaches to
Statistical Seasonal Forecasting in the Sahel**

JURY

M. Jean-Emmanuel PATUREL, IRD (Côte d'Ivoire)	President
M. Agnidé Emmanuel LAWIN, Université d'Abomey-Calavi (Benin)	Reporter
M. Mahaman Moustapha ADAMOU, Université Abdou Moumouni (Niger)	Reporter
M. Kazimierz ADAMOWSKI, University of Ottawa (Canada)	Examiner
M. Musandji FUAMBA, Polytechnique Montreal (Canada)	Examiner
M. Abdolmajid MOHAMMADIAN, University of Ottawa (Canada)	Examiner
Mrs. Hadiza MOUSSA SALEY, CESAG (Senegal)	Examiner
M. Ousmane SEIDOU, University of Ottawa (Canada)	Director of Thesis
M. Harouna KARAMBIRI, 2iE (Burkina Faso)	Director of Thesis

© Abdouramane Gado Djibo, Ottawa, Canada, 2016

**Laboratoire Hydrologie et Ressources en Eau (LEAH, 2iE)
Hydraulics Laboratory (Water Resources Group, uOttawa)**

Abstract

Water resources management in the Sahel region of West Africa is extremely difficult because of high inter-annual rainfall variability as well as a general reduction of water availability in the region. Observed changes in streamflow directly disturb key socioeconomic activities such as the agriculture sector, which constitutes one of the main survival pillars of the West African population. Seasonal rainfall forecasting is considered as one possible way to increase resilience to climate variability by providing information in advance about the amount of rainfall expected in each upcoming rainy season. Moreover, the availability of reliable information about streamflow magnitude a few months before a rainy season will immensely benefit water users who want to plan their activities. However, since the 90s, several studies have attempted to evaluate the predictability of Sahelian weather characteristics and develop seasonal rainfall and streamflow forecast models to help stakeholders take better decisions. Unfortunately, two decades later, forecasting is still difficult, and forecasts have a limited value for decision-making. It is believed that the low performance in seasonal forecasting is due to the limits of commonly used predictors and forecast approaches for this region. In this study, new seasonal forecasting approaches are developed and new predictors tested in an attempt to predict the seasonal rainfall over the Sirba watershed located in between Niger and Burkina Faso, in West Africa. Using combined statistical methods, a pool of 84 predictors with physical links with the West African monsoon and its dynamics were selected, with their optimal lag times. They were first reduced through screening using linear correlation with satellite rainfall over West Africa. Correlation analysis and principal component analysis were used to keep the most predictive principal components. Linear regression was used to get synthetic forecasts, and the model was assessed to rank the tested predictors. The three best predictors, air temperature (from Pacific Tropical North), sea level pressure (from Atlantic Tropical South) and relative humidity (from Mediterranean East) were retained and tested as inputs for seasonal

rainfall forecasting models. In this thesis it has been chosen to depart from the stationarity and linearity assumptions used in most seasonal forecasting methods:

1. Two probabilistic non-stationary methods based on change point detection were developed and tested. Each method uses one of the three best predictors. Model M1 allows for changes in model parameters according to annual rainfall magnitude, while M2 allows for changes in model parameters with time. M1 and M2 were compared to the classical linear model with constant parameters (M3) and to the linear model with climatology (M4). The model allowing changes in the predictand-predictor relationship according to rainfall amplitude (M1) and using AirTemp as a predictor was the best model for seasonal rainfall forecasting in the study area.
2. Non-linear models including regression trees, feed-forward neural networks and non-linear principal component analysis were implemented and tested to forecast seasonal rainfall using the same predictors. Forecast performances were compared using coefficients of determination, Nash-Sutcliffe coefficients and hit rate scores. Non-linear principal component analysis was the best non-linear model (R^2 : 0.46; Nash: 0.45; HIT: 60.7), while the feed-forward neural networks and regression tree models performed poorly.

All the developed rainfall forecasting methods were subsequently used to forecast seasonal annual mean streamflow and maximum monthly streamflow by introducing the rainfall forecasted in a SWAT model of the Sirba watershed, and the results are summarized as follows:

1. Non-stationary models: Models M1 and M2 were compared to models M3 and M4, and the results revealed that model M3 using RHUM as a predictor at a lag time of 8 months was the best method for seasonal annual mean streamflow forecasting, whereas model M1 using air temperature as a predictor at a lag time of 4 months was the best model to predict maximum monthly streamflow in the Sirba watershed. Moreover, the calibrated SWAT model achieved a NASH value of 0.83.
2. Non-linear models: The seasonal rainfall obtained from the non-linear principal component analysis model was disaggregated into daily rainfall using the method

of fragment, and then fed into the SWAT hydrological model to produce streamflow. This forecast was fairly acceptable, with a Nash value of 0.58. The evaluation of the level of risk associated with each seasonal forecast was carried out using a simple risk measure: the probability of overtopping of the flood protection dykes in Niamey, Niger. A HEC-RAS hydrodynamic model of the Niger River around Niamey was developed for the 1980-2014 period, and a copula analysis was used to model the dependence structure of streamflows and predict the distribution of streamflow in Niamey given the predicted streamflow on the Sirba watershed. Finally, the probabilities of overtopping of the flood protection dykes were estimated for each year in the 1980-2014 period. The findings of this study can be used as a guideline to improve the performance of seasonal forecasting in the Sahel. This research clearly confirmed the possibility of rainfall and streamflow forecasting in the Sirba watershed at a seasonal time scale using potential predictors other than sea surface temperature.

Keywords: West African monsoon, seasonal rainfall-streamflow forecasting, change point detection, seasonal flood forecast, probability risk assessment, Sahel, Sirba watershed

Résumé

La gestion des ressources en eau dans la région du Sahel, en Afrique de l'Ouest, est extrêmement difficile en raison de la forte variabilité interannuelle des précipitations ainsi que de la dégradation générale de la disponibilité de l'eau dans la région. Les changements observés dans les débits perturbent directement les activités socio-économiques clés telles que l'agriculture, une des principales sources de subsistance pour la population ouest-africaine. La prévision saisonnière des précipitations offre la possibilité d'accroître la résilience à la variabilité climatique en fournissant des informations à l'avance sur la quantité de pluie qu'on peut espérer pour la saison des pluies à venir. De plus, la disponibilité d'informations fiables sur la quantité des écoulements quelques mois avant la saison des pluies serait immensément bénéfiques pour les utilisateurs de l'eau qui pourront ainsi planifier leurs activités. Depuis les années 90, plusieurs études ont tenté d'évaluer la prévisibilité des caractéristiques de la pluviométrie sahélienne, et de développer des modèles de prévision saisonnière de précipitations et de débits pour aider les parties prenantes à prendre de meilleures décisions. Malheureusement, deux décennies plus tard, les méthodes de prévision saisonnière sont encore peu fiables et les prévisions ont une valeur pratique limitée pour la prise de décision. La faible performance dans la prévision saisonnière serait partiellement due aux limitations des prédicteurs et des approches de prévisions couramment utilisées pour cette région. Dans cette thèse, des nouvelles approches de prévision saisonnière ont été développées et des nouveaux prédicteurs sont testés avec pour objectif de prédire la quantité de précipitations saisonnières sur le bassin versant de la Sirba, un cours d'eau situé à cheval entre le Niger et le Burkina Faso, en Afrique occidentale. Un groupe de 84 prédicteurs ayant des liens physiques avec la mousson ouest-africaine et avec sa dynamique ont été sélectionnés. Le choix de ces prédicteurs est aussi basé sur le temps de latence optimal entre leur observation et la saison des pluies estimé en utilisant une combinaison de méthodes statistiques. La dimension des vecteurs

de prédictors a d'abord été réduite grâce à une analyse de corrélation linéaire avec des précipitations estimées par satellite sur la région ouest-africaine. L'analyse de corrélation et l'analyse en composantes principales ont été utilisées pour ne retenir que les composantes principales prédictives élevées. Un modèle de régression linéaire a été utilisé pour obtenir des prévisions saisonnières et le modèle a été évalué afin de classer les prédictors par ordre de performance. Les trois meilleurs prédictors, soient la température de l'air (du Pacifique tropical Nord), la pression du niveau de la mer (de l'Atlantique tropical sud) et l'humidité relative (à partir de la Méditerranée orientale) ont été retenus et utilisés ensemble comme entrées pour les modèles de prévision des précipitations saisonnières.

Dans cette thèse, des hypothèses de stationnarité et de linéarité ne sont pas émises comme c'est le cas dans la plupart des méthodes de prévision saisonnière:

1. Deux méthodes probabilistes non stationnaires basées sur la détection de ruptures ont été développées et ensuite testées. Chaque méthode utilise un des trois prédictors sélectionnés. Le modèle M1 permet un changement de valeur des paramètres du modèle linéaire en fonction de l'amplitude des précipitations annuelles, tandis que M2 permet des changements de valeurs des paramètres du modèle linéaire avec le temps. M1 et M2 ont été comparés au modèle linéaire classique avec paramètres constants (M3) et à la climatologie (M4). Le modèle qui permet un changement dans la relation pluie-prédictor selon l'amplitude des précipitations (M1), et utilisant AirTemp comme prédictor a obtenu le meilleur score pour la prévision saisonnière des précipitations dans la zone d'étude.
2. Des modèles non-linéaires, y compris les arbres de régression, le réseau de neurones, et l'analyse en composantes principales non-linéaire ont été mises en œuvre et testées pour prévoir les précipitations saisonnières en utilisant les mêmes prédictors. Les performances en prévision ont été comparées en utilisant le coefficient de détermination, le coefficient de Nash-Sutcliffe et le score de taux de succès. L'analyse en composantes principales non-linéaire était le meilleur modèle non-linéaire (R^2 : 0,46; Nash: 0,45; HIT: 60,7), tandis que le réseau de neurones et les modèles d'arbres de régression ont de résultats médiocres.

Toutes les méthodes de prévision des précipitations développées ont ensuite été utilisées pour la prévision des débits annuels moyens saisonniers et des débits maximum mensuels

en introduisant les précipitations prédites dans un modèle SWAT du bassin versant de la Sirba. Les résultats sont résumés comme suit:

1. Modèles non stationnaires: Les modèles M1 et M2 ont été comparés aux modèles M3 et M4. Les résultats obtenus ont révélé que le modèle M3 en utilisant RHUM comme prédicteur avec un temps de latence de 8 mois a obtenu le meilleur score pour les prévisions des débits annuels moyens saisonniers. Le modèle M1 en utilisant AirTemp comme prédicteur à un temps de latence de 4 mois est le meilleur modèle pour prédire les débits mensuels maximaux dans le bassin versant de la Sirba. En outre, le modèle SWAT calibré atteint une valeur NASH de 0,83.
2. Modèles non-linéaires: Les précipitations saisonnières obtenues à partir du modèle d'analyse en composantes principales non-linéaire ont été désagrégées au pas de temps journalier en utilisant la méthode du fragment, puis introduites dans le modèle hydrologique SWAT pour générer des débits. Cette prévision était assez acceptable avec une valeur Nash de 0,58.

L'évaluation du niveau de risque associé à chaque prévision saisonnière a été réalisée en utilisant une mesure simple de risque: la probabilité de débordement des digues de protection contre les inondations à Niamey, au Niger. Un modèle hydro-dynamique HEC-RAS du fleuve Niger de part et d'autre de Niamey a été développé en utilisant les données de débit de la période 1980-2014; une analyse de copules a été utilisée pour modéliser la structure de dépendance entre les débits et de prédire la répartition des débits à Niamey compte tenu de l'écoulement prévu sur le bassin versant de la Sirba. Finalement, les probabilités de déversement au-dessus des digues ont été estimées pour chaque année de la période 1980-2014. Les résultats de cette étude peuvent être utilisés comme un moyen d'évaluer et d'améliorer la performance de la prévision saisonnière au Sahel. Les travaux ont confirmé clairement la possibilité d'améliorer la prévision des précipitations et débits dans le bassin versant de la Sirba à l'échelle de temps saisonnière en utilisant des prédicteurs autres que les températures à la surface des mers.

Mots-clés: mousson ouest-africaine, prévision saisonnière des précipitations et débits, détection de ruptures, prévision des crues saisonnières, évaluation des probabilités de risques, Sahel, bassin versant Sirba

Acknowledgments

During this research, I received support and help of many people that I want to thank here.

First, my two directors of thesis, Harouna Karambiri, Director of the Doctoral School and Associate Professor at 2iE, and Ousmane Seidou, Associate Professor at the University of Ottawa (Canada), for their excellent guidance, enthusiasm, and intellectual support. I wish to thank Mr. Karambiri, my supervisor, for his patience, availability, understanding, and encouragement. You worked to establish the joint supervision agreement and greatly facilitated my mobility at the University of Ottawa. I also thank my co-director, Mr. Seidou for co-directing this thesis. You pushed me to my limits, gradually and always with a perfect dose. I am aware that if I am stronger and more confident today, you are largely responsible. My PhD experience has been so enriched by the presence of both of you, and I sincerely hope one day to return this favor.

I would also like to thank the jury, Agnidé Emmanuel Lawin, Mahaman Moustapha Adamou and Kazimierz Adamowski, for having accepted and taken the time to evaluate my thesis and providing comments that greatly improved the quality.

Also, I would like to thank Jean-Emmanuel Paturel, Abdolmajid Mohammadian, Hadiza Moussa Saley and Fuamba Musandji for agreeing to examine this thesis.

Special thanks to Fondation 2iE in collaboration with the African Development Bank (AfDB) for giving me a doctoral fellowship through the PARIASST project. I am also very grateful to the International Research Initiative on Adaptation to Climate Change (IRIACC) program, through the International Development Research Center (Canada), for funding a part of this research. Also, thanks to the National Directorate of Meteorology of Burkina Faso, the AGRHYMET Regional Centre, and the Resources Department of Water of Burkina Faso for providing me the meteorological and hydrological data.

I am also indebted to several people, including Nathalie Philippon, for her help and availability and who provided me with a lot of satellite rainfall and weather data that helped me to select good predictors. My deep appreciation to Hamma Yacouba, Niang Dial, Luc Descroix, Gil Mahe, Mamadou Koita, Malicki Zoromé, and Corentin Somé for their exchanges and advice on hydrological modeling, statistics and mapping issues.

My sincere gratitude to Hadiza Moussa Saley, who pushed me to continue my postgraduate studies and who has been an excellent mentor over the years since my Master's program. You helped me to acquire my first knowledge in MATLAB programming, and you aided me greatly with your availability and wise suggestions throughout this thesis.

Thanks to Geneviève Yameogo and Reine Dagbo for their support from the beginning of this thesis to the end.

I want to thank all the colleagues I met during my PhD studies in Canada. Their interest, questions and encouragement were energising. Thanks especially to Ketvara Sittichok, Zaid Aldafeeri, Oday Al-Heetimi, Saideh, Gohe, Mohammad Hydar, Mossi Idrissa AbdelKader, Mohammed Yatara and Moumouni Ali, who supported me a lot during my stay in Canada.

I am also very thankful to Professor Seidou's family, who kindly provided accommodations to me during the last 3 years; they have supported and encouraged me throughout my stay in Ottawa.

I am grateful to my colleagues in 2iE and elsewhere, Diafarou Moumouni, Bernadette, Soro Dimitri, Doto, Beteo, Hamidatu, Sangaré, Fowé, Amaré, Aida, Noellie, Bassirou Boubé, Hama Amadou, Awa Koita, Moussa Ibrahim, Laouali Sadissou, Ibrahim Abdou, Sanoussi Boureima, Boukari Tamboura, Abdelkader Issaka, Daouda Kouotou, Awa Gueye, Djibrilla Daouda Karimou, Daouda Salou, Ahmed Mamane and Nassirou Talatou for their empathy and encouragement.

Finally, I would like to thank my family, especially my mother for her unconditional support, love, patience, care, and all the support she has given me throughout my studies. As well, thank you to my brothers Boureima, Mahamadou, Ousseini, Hassane, AbdelKader, my wife Amsatou, and my sister Hadiza for your moral

support and advice in the most difficult periods and for your always being present to celebrate my successes.

This thesis is dedicated to the loving memory of my father.

Table of Contents

Abstract	ii
Résumé	v
Acknowledgments	viii
Table of Contents	xi
List of Figures.....	xv
List of Tables	xvii
List of Acronyms	xviii
Chapter 1. Introduction	1
1.1. Context and Problem Statement.....	1
1.2. Thesis Objectives	6
1.2.1 Main objective	7
1.2.2 Specific objectives	7
1.3. Research Questions and Hypothesis	7
1.3.1 Research questions.....	7
1.3.2 Hypothesis	8
1.4. Novelty	8
1.5. Scope of the study.....	9
1.6. Organisation of the Thesis	10
Chapter 2. West African Monsoon and Sahel Rainfall: A Review	12
2.1. West African Sahel Monsoon: General Circulation	12
2.2. West African Sahel Rainfall	15
2.2.1 Rainfall mechanism	15
2.2.2 Seasonal cycle and variability	18
2.3. Main Drivers of Sahel Rainfall	20
2.3.1 Teleconnections between Sahel rainfall and oceanic basins.....	20
2.3.2 Other drivers of Sahel rainfall	23
2.4. Synthesis and Partial Conclusion	25

Chapter 3. Background and Theory on Seasonal Forecasting and Hydrological Modelling	26
3.1. <i>Seasonal Forecasting</i>	26
3.1.1 Seasonal rainfall forecasting.....	27
3.1.2 Seasonal Streamflow forecasting.....	32
3.1.3 Seasonal flood forecasting.....	33
3.2. <i>Statistical Tools for Seasonal Forecasting</i>	34
3.2.1 Principal component analysis	34
3.2.2 Non-linear principal component analysis	35
3.2.3 Multiple linear regression	37
3.2.4 Stepwise regression	38
3.2.5 Artificial neural network.....	39
3.2.6 Regression trees	42
3.2.7 Non-stationary techniques	44
3.3. <i>Hydrological Modelling</i>	48
3.3.1 Classification of hydrological models	49
3.3.2 Criteria for adequate model selection	50
3.3.3 Calibration and validation of SWAT model	51
3.3.4 Model performance analysis.....	53
3.4. <i>Temporal Disaggregation</i>	53
3.4.1 Method of fragment	54
3.4.2 Modified method of fragment.....	56
3.5. <i>Forecast Assessment</i>	58
3.5.1 Coefficient of determination.....	59
3.5.2 Nash-Sutcliffe coefficient	60
3.5.3 Hit Rate Score.....	60
3.5.4 Bagging and boosting	61
3.5.5 Leave-one-out cross-validation	62
3.6. <i>Losses Associated with Seasonal Flood Forecasting</i>	64
3.6.1 Copula method.....	65
3.6.2 HEC-RAS model	65
3.7. <i>Synthesis and Partial Conclusion</i>	66
Chapter 4. Study Area and Data Presentation	69
4.1. <i>West Africa</i>	69
4.1.1 Relief	69
4.1.2 Vegetation.....	70
4.2. <i>Sirba Basin</i>	71
4.2.1 Geographical description	73
4.2.2 Climate.....	74
4.2.3 Hydrography	75
4.2.4 Geological and hydrogeological characteristics of the basin.....	76
4.2.5 Soils and land use	79
4.2.6 Demography	80
4.2.7 Environmental, social and economic activities.....	81

4.3.	<i>Data Presentation</i>	82
4.3.1	<i>Climate data</i>	82
4.3.2	<i>Hydrological data</i>	87
4.3.3	<i>Soil and land use data</i>	89
4.3.4	<i>Topographical data</i>	89
4.4.	<i>Synthesis and Partial Conclusion</i>	89
Chapter 5.	Seasonal Rainfall Forecasting	91
5.1.	<i>Selection of Predictors and Potential Lag Time</i>	91
5.2.	<i>Linear and Non-Linear Approaches for Statistical Seasonal Rainfall Forecast in the Sirba Watershed Region (Sahel)</i>	93
5.3.	<i>Seasonal Rainfall Forecasting with Changing and Constant Parameters</i>	125
5.4.	<i>Development and Assessment of Non-Linear and Non-Stationary Seasonal Rainfall Forecast Models for the Sirba Watershed, West Africa</i>	127
5.5.	<i>Regression Trees and Neural Network Seasonal Forecasts</i>	156
5.5.1	<i>Regression Tree for Seasonal Rainfall Forecasting</i>	156
5.5.2	<i>Neural Network Seasonal Rainfall Forecasting</i>	159
5.6.	<i>Synthesis and Partial Conclusion</i>	162
Chapter 6.	Seasonal Streamflow Forecasting	164
6.1.	<i>Sirba SWAT Hydrological Model</i>	164
6.1.1	<i>Justification and Method</i>	164
6.1.2	<i>Sirba Hydrological model</i>	167
6.2.	<i>Statistical Seasonal Streamflow Forecasting using Probabilistic Approach over West Africa</i>	170
6.3.	<i>Seasonal Streamflow Forecasting with Non-linear Methods</i>	203
6.4.	<i>Synthesis and Partial Conclusion</i>	204
Chapter 7.	Seasonal Flood Forecasting	206
7.1.	<i>Copula Approach</i>	206
7.2.	<i>Niger River Hydro-Dynamical Model at Niamey</i>	207
7.3.	<i>Probability of Overtopping</i>	208
7.4.	<i>A Copula-based Approach for Assessing Flood Protection Overtopping associated with a Seasonal Flood Forecast in Niamey, West Africa</i>	208
7.5.	<i>Synthesis and Partial Conclusion</i>	241
Chapter 8.	Conclusions and Perspectives	242
8.1.	<i>Conclusion</i>	242
8.2.	<i>Contributions</i>	245

8.3. <i>Future Work</i>	246
References	248

List of Figures

FIGURE 1.1 GEOGRAPHICAL POSITION OF THE SAHEL BETWEEN 12°N AND 20°N -----	3
FIGURE 2.1 ATMOSPHERIC CIRCULATION: H AND L MEAN LAYERS OF HIGH AND LOW PRESSURE RESPECTIVELY -----	13
FIGURE 2.2 ZONAL GLOBAL CIRCULATION TYPE WALKER -----	14
FIGURE 2.3 CONCEPTUAL SCHEME OF THE WEST AFRICAN CLIMATE MECHANISM-----	15
FIGURE 2.4 WEST AFRICAN MONSOON CYCLE -----	17
FIGURE 3.1 THE STEPS OF A STATISTICAL-NUMERICAL APPROACH -----	32
FIGURE 3.2 NON-LINEAR DIMENSIONALITY REDUCTION IN NLPCA -----	37
FIGURE 3.3 NEURAL NETWORK STRUCTURE WITH TWO HIDDEN LAYERS -----	40
FIGURE 3.4 SCHEMATIC REPRESENTATION OF A REGRESSION TREE -----	43
FIGURE 3.5 SCHEMATIC DIAGRAM OF THE RELATIONSHIP BETWEEN COMPLEXITY AND PERFORMANCE OF A HYDROLOGICAL MODEL AND DATA AVAILABILITY -----	51
FIGURE 3.6 FLOWCHART OF THE FRAGMENT DISAGGREGATION APPROACH (MF)-----	56
FIGURE 3.7 SUMMARY FLOW CHART OF THE MODIFIED FRAGMENT DISAGGREGATION APPROACH (MMF) -----	57
FIGURE 3.8 THE CROSS-VALIDATION PROCEDURE (BLUE BOXES ARE FOR VALIDATION AND WHITE BOXES ARE FOR TRAINING)-----	63
FIGURE 4.1 LAND USE IN AFRICA -----	71
FIGURE 4.2 LOCATION MAP OF SIRBA WATERSHED-----	72
FIGURE 4.3 ADMINISTRATIVE MAP OF SIRBA BASIN -----	74
FIGURE 4.4 HYDROGRAPHIC NETWORK OF THE SIRBA BASIN -----	76
FIGURE 4.5 GEOLOGICAL MAP OF THE SIRBA BASIN-----	78
FIGURE 4.6 OBSERVED RAINFALL AND STREAMFLOW MEASURE STATIONS -----	83
FIGURE 4.7 ILLUSTRATIVE IMAGE OF MEAN MONTHLY RELATIVE HUMIDITY -----	86
FIGURE 4.8 GAP RATIOS IN DAILY FLOW SERIES AT LIPTOUGOU STATION (1973-2011) ---	88

FIGURE 5.1 OBSERVED AND SIMULATED SEASONAL RAINFALL BASED ON RT USING RELATIVE HUMIDITY -----	158
FIGURE 5.2 OBSERVED AND SIMULATED SEASONAL RAINFALL BASED ON RT USING AIR TEMPERATURE -----	158
FIGURE 5.3 OBSERVED AND SIMULATED SEASONAL RAINFALL BASED ON FFNN USING AIRTEMP (UPPER PANEL) AND RHUM (LOWER PANEL) -----	161
FIGURE 6.1 SIRBA SUB-WATERSHEDS FROM SWAT-----	165
FIGURE 6.2 LAND USE CLASSES OF THE SIRBA SWAT HYDROLOGICAL MODEL -----	168
FIGURE 6.3 SLOPE CLASSES OF SIRBA SWAT HYDROLOGICAL MODEL -----	168
FIGURE 6.4 OBSERVED AND SIMULATED RAINFALL USING METHOD OF FRAGMENT (1989- 2003)-----	203
FIGURE 6.5 OBSERVED AND SIMULATED STREAMFLOW (1989-2002)-----	204

List of Tables

TABLE 3.1 COMPARATIVE SUMMARY OF BAGGING AND BOOSTING TECHNIQUES -----	61
TABLE 4.1 GEOLOGICAL FORMATION IN THE SIRBA WATERSHED (SOURCE: TAWEYE, 1995) -----	77
TABLE 4.2 SPECIFICS OF RAINFALL STATIONS -----	83
TABLE 4.3 DESCRIPTION OF ATMOSPHERIC DATA -----	86
TABLE 4.4 HYDROLOGICAL STATIONS OF THE SIRBA WATERSHED -----	88
TABLE 5.1 FFNN MODEL (SINGLE PREDICTOR) OUTPUT FOR SIRBA SEASONAL RAINFALL FORECAST -----	160
TABLE 6.1 RANKS OF SENSITIVE PARAMETERS OF SIRBA HYDROLOGICAL MODEL -----	169

List of Acronyms

Acronym	Definition
ACMAD	African center of meteorological application for development
AEJ	African easterly jet
AGRHMET	Agriculture, Hydrology and Meteorology Research Centre
AirTemp	Air temperature
AMMA	African Monsoon Multidisciplinary Analyses
ANN	Artificial neural network
ARMA	Autoregressive moving average
BADC	British atmospheric data centre
BGDM	Batch gradient descent with momentum
CART	Classification and regression trees
CC	Correlation coefficient
CPT	Climate predictability tool
CRU	Climatic Research Unit
CRV	Cross-validation method
CT	Classification tree
DEM	Digital elevation model
DGARH	Direction générale de l’agriculture et des ressources halieutiques
DRE	Direction des ressources en eau
ENSO	El Niño-Southern Oscillation
FFNN	Feedforward neural network
GCM	General circulation models
GDALR	Gradient descent with adaptive learning rate
GIS	Geographical information system
GLUE	Generalize likelihood uncertainty estimation
HEC-RAS	Hydrologic Engineering Center’s River Analysis System
HIT	Hit score
HRU	Hydrological response unit
IRI	International research institute for climate and society
ITCZ	Inter-Tropical Convergence Zone
KNN	K-nearest-neighbor
LHS	Latin hypercube sampling
LOOCV	Leave-one-out cross validation
MCMC	Monte-Carlo Markov Chain
MF	Method of fragment
MLR	Multiple linear regression
MMF	Modified method of fragment
MOS	Model output statistics

MSE	Moist static energy
NASH	Nash-Sutcliffe coefficient
NCEP	National centers for environment predictions
NLPCA	Non-linear principal component analysis
NOAA	National Oceanic and Atmospheric Administration
ParaSol	Parameter solution
PC	Principal component
PCA	Principal component analysis
PRESAO	PREvision Saisonnière en Afrique de l'Ouest
PSO	Particle swarm optimization
RCM	Regional climate models
RHUM	Relative humidity
ROC	Relative operating characteristic
RPSS	Rank probability skill score
RT	Regression tree
SCS	Soil conservation service
SLP	Sea level pressure
SR	Stepwise regression
SST	Sea surface temperature
STRM	Shuttle radar topography mission
SUFI2	Sequential uncertainty fitting
SWAT	Soil and Water Assessment Tool
SWAT-CUP	SWAT calibration and uncertainty programs
TD	Temporal disaggregation
TEJ	Tropical easterly jet
UWND	Zonal wind
VWND	Meridional wind
WAM	West African monsoon
WMO	World Meteorological Organization

Chapter 1. Introduction

This chapter gives a general introduction explaining the context under which the thesis work was carried out, before stating the problems that the present work has tried to assess and tackle. It also presents the objectives of the work, which include both the global objective and the specific objectives. The research questions and hypothesis are explained; afterward, the contributions, novelty and scope of the thesis are discussed. The last part of this chapter presents the organisation of the thesis.

1.1. Context and Problem Statement

The summer rainfall in semi-arid regions of the world is known for its unreliability, which has a large impact on the continental hydrological cycle, water resources and food security. The Sahel, extending across Africa from the Atlantic Ocean to 30°E and from 12° to 17°N, is the largest area of these regions, recording between 200 and 800mm/year from north to south, with ~80% of the rain being recorded in July–September, the rainy season (Nicholson, 2013).

Several studies have confirmed West Africa's degree of susceptibility to climate variability, and its constant exposure to the consequences of climate change (e.g., Christensen *et al.*, 2007; Gianini *et al.*, 2008). This has led some authors to question whether the persistent lengthy dry or rainy periods are truly historical aspects of the environment of the area (Ardoin *et al.*, 2003; Bouali, 2009). The Sahelian rainfall pattern is season-dependent and is directly related to the West African Monsoon (WAM), a dynamic yet to be fully understood by climatologists (Mohino *et al.*, 2011; Caminade & Terray, 2010; Biasutti *et al.*, 2008; Rowell, 2001, 2003; Janicot *et al.*, 2001; Palmer, 1986). This lack of a clear understanding of the WAM is, in part, why forecasts on all scales in the Sahel are problematic (see Figure 1.1). The uncertainty of the forecasts can directly affect local populations, since lack of awareness of the short- and medium-term evolution of rainfall and streamflows often results in inadequate or untimely measures to

cope with the increasingly frequent climate extremes. These can include shortage of precipitation as well as flooding, with their direct consequences of lower crop yields, loss of agricultural production, and the destruction of high value infrastructure such as roads and dams (Tarhule, 2005; Samimi *et al.*, 2012). Recurring droughts also affect agricultural production and streamflows (Samimi *et al.*, 2012; Brooks, 2004), and often catch local authorities and rural populations unawares, despite over a decade of publication of seasonal forecasts for West Africa (*PRESAO: Prévision Saisonnière en Afrique de l'Ouest*) (Hamatan, 2004; Ogallo *et al.*, 2000). In such unstable circumstances, any scientific information regarding the short (24 hours) and medium (6 months) terms of rainfall and streamflow trends is vital for decision-making and managing water resources. Agriculture, the primary socio-economic activity in the Sahelian zone, would be more efficient if reliable seasonal local information was available to help farmers make decisions (Hansen, 2002). Thus, the development of seasonal rainfall and streamflow forecasting models is anticipated and welcomed by all involved, particularly the rural population, as effective climatic information is akin to food security insurance. The models would increase resilience to climate variability by providing advance information about the expected amount of rain or runoff in the next rainy season (Hansen *et al.*, 2011).

Seasonal forecasting prediction attempts to assess what the climate will be over the next few months. It is currently the best means of addressing climate issues, and is also considered the most efficient approach to drought mitigation (Hayes *et al.*, 2005). METEO France defines seasonal forecasting as a probabilistic estimate of precipitation on a spatio-temporal scale of a season. Hence, it generally consists of searching for statistical relationships between indicators of the rainy season and parameters characterizing the state of the atmosphere, including oceans. A simple forecast is different than a seasonal forecast, as the latter is the result of an attempt to estimate the actual evolution of the climate in the future (e.g., seasonal, interannual or long-term scales).

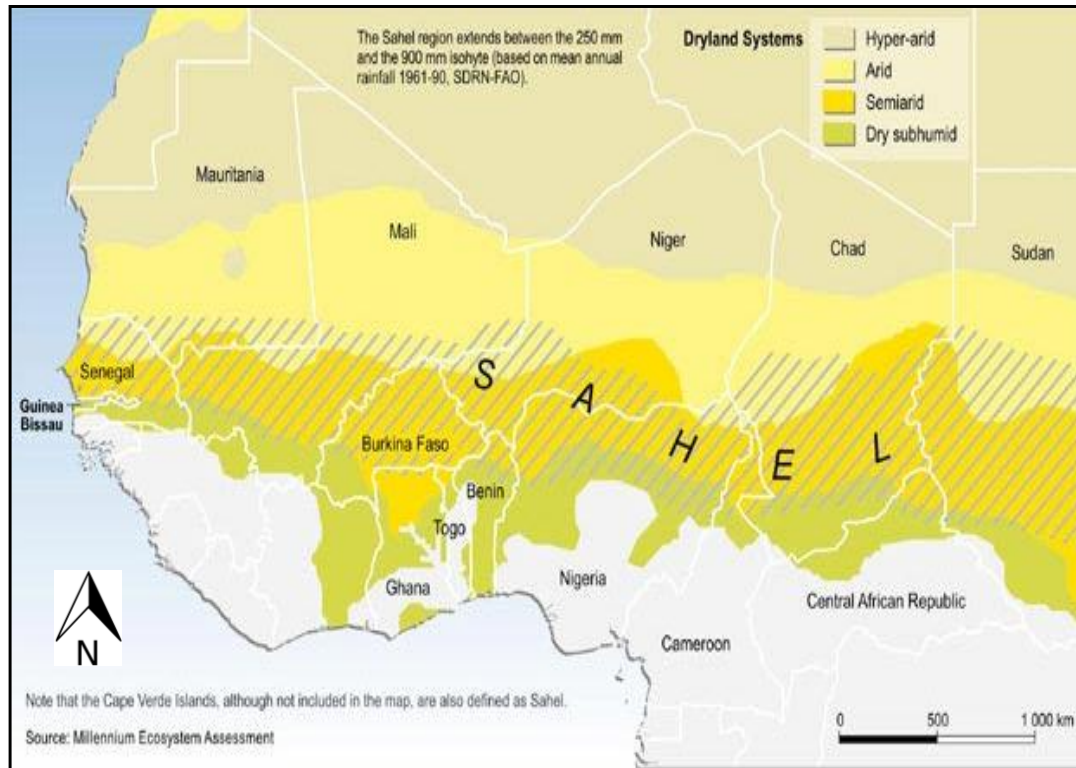


Figure 1.1 Geographical position of the Sahel between 12°N and 20°N (source: MEA, 2005)

Seasonal precipitation forecasts in West Africa have only been conducted in recent decades, and have been a major challenge for over 20 years. Between the late 1960s and the 1990s, the annual total rainfall in West Africa decreased. It began to increase during the early 2000s, though successive dry years led to a negative long-term rainfall trend which caused the most significant drought period on the planet (Fontaine & Janicot, 1996). Given the direct dependence on climate of agriculture and population lifestyles, these drought periods had significant impacts on the survival of the affected inhabitants.

In the Sahel, the African Center of Meteorological Application for Development (ACMAD) is the principal source of climate forecasts tailored to the needs of farmers and other socio-economic actors (Hamatan, 2002). In collaboration with African national meteorology services and AGRHYMET (Agriculture, Hydrology and Meteorology Research Center), ACMAD's primary mission is the development of climate information and the provision of climate forecasting tools for the short and medium term. These forecasts use a Climate Predictability Tool (CPT) to link sea surface temperature (SST) to

monthly or seasonal rainfall or streamflow data. CPT uses multiple linear regressions to link a grid predictor to a predictand. ACMAD makes two types of forecasts available to users:

- short-term forecasts of precipitation (24 hours), updated daily
- seasonal forecasts of precipitation, issued annually in the form of a bulletin

The seasonal forecast bulletins of precipitation are determined by the consensus of a panel of experts, who synthesize different forecasts provided by numerical models of global climate centers (Hamatan, 2002). Unfortunately, currently available models have several shortcomings with respect to representing processes and reproducing mechanisms, leading to a certain lack of practicality of the simulated climate. Moreover, by comparing their outputs to the observations, in terms of mean fields and teleconnections, it is clear that the improvements remain inadequate. This is particularly true over the extended area of the WAM, where the models still render badly reproduced cumulative and annual cycles of precipitation over West Africa, and more so in the Sahel zone (Philippon, 2002). For instance, it is still difficult to have confidence in the predictions of forecasting modes beyond the short term (i.e., a few days), particularly the precipitation fields. The models have difficulty simulating the tropics, where the convections playing on vertical instabilities are strong and the rains are intermittent (Janicot *et al.*, 1998).

Although somewhat useful, the forecasting information provided in PRESAO bulletins has limitations, e.g.:

- the forecasting is qualitative rather than quantitative and is issued in a categorical format in three scales: probability of above normal, normal, and below normal precipitation or streamflow
- the forecasting does not provide information on the temporal distribution of precipitation, which is essential for risk management

Thus, it is very difficult for organizational entities to effectively use the information from PRESAO bulletins to enhance their operations. The low predictive skill of the PRESAO was confirmed by a study by Konte (2011), who used the rank probability skill score (RPSS) and relative operating characteristic (ROC) to evaluate its performance from 1998 to 2010 in Sénégal. He found that the forecast reliability varied according to

geographical location in the region, and that overall the forecasts performed better than the climatology in only 54% of cases. Therefore, effectively streamlining the management of water resources in West Africa requires the development of a quantitative system of precipitation forecasting in the region, using a watershed approach.

Over the last decade, the field of seasonal forecasting has led to the emergence of many research projects worldwide, conducted by groups composed of national meteorology services and centers for climate and environmental research (e.g., the European projects DEMETER and ENSEMBLES). This escalation of interest has inspired the establishment of several models. Relevant efforts of the scientific community are based on three different but complementary approaches (Hastenrath, 1995): dynamical (based only on numerical models), statistical (based purely on statistics), and hybrid statistical-dynamical (a combination of numerical models and statistics).

Numerical models of seasonal forecasts use the outputs of atmospheric models similar to weather forecasting models. Some seasonal forecasts use only one numerical model, and others use several (multi-model forecasts). However, the confidence level of numerical forecasts varies strongly according to region. In a recent study, Batté and Déqué (2011) showed that the performance of numerical models (e.g., Model GCM, European Ensemble Model) for seasonal precipitation forecasting in West Africa is very limited. Raje and Mujumdar (2011) studied three downscaling methods for daily precipitation in the Panjab region, India, and observed that the output resolution of some types of general circulation models (GCMs) were inadequate for runoff simulation.

Moreover, several studies have proved that seasonal forecasting models based on a statistical approach, although imperfect, provide better performance than numerical models (Philippon, 2002; Garcia-Serrano *et al.*, 2013). Moron *et al.* (2004) also showed that numerical models for seasonal forecasts of precipitation in West Africa yield unrealistic results, and recommended using numerical multi-model outputs (e.g., DEMETER) as well as the development of statistical-dynamical models. Philippon (2002) provided a diagnosis of the methods used for seasonal forecasts of precipitation in eastern and western Africa from 1968 to 1998, and proposed a new statistical approach to generate more realistic forecasts. The seasonal forecasting models developed by Philippon (2002) are based on linear methods (i.e., multiple linear regressions and

canonical correlation analysis), and a non-linear method (i.e., an artificial neural network). A performance evaluation of these models revealed that the limitations of linear models are due to the linearity condition imposed on the predictors of precipitation; however, the use of neural networks could not overcome the linear constraints imposed on the predictors. Nevertheless, this study and others (e.g., Garric *et al.*, 2002) confirms the interest in adapting predictors, such as sea surface temperature and other climate indices, to the local or regional climate context. Since climate indices are averages of the climate variables of a particular region, it is possible that changing the region and/or the variables could help to achieve new indices that are better correlated with the observations of the study area.

The hybrid statistical-numerical approach, also known as model output statistics (MOS), is a combination method based on the principle of applying statistical methods to the output obtained from numerical models in order to perform further analysis.

Conversely, in West Africa, specifically Burkina Faso, a series of studies (Kirshen *et al.*, 2003; Ingram *et al.*, 2002; Roncoli *et al.*, 2001, 2002, 2004, 2009) showed that coping with the variability of the rainy season (July to September) requires farmers to use any relevant information on seasonal rainfall/streamflow forecasts. In this context, the challenge is to develop an operational seasonal forecasting system that delivers information that can be easily used by farmers to help them make optimal decisions. Consequently, satisfactory seasonal forecasting requires more than using model outputs without considering the local climate. In fact, what is needed is an effective methodology that takes the actual local climate characteristics into account, and which is based on purely statistical methods. This thesis addresses the aim of the scientific community to make seasonal forecasting a suitable tool to support correct decision-making.

1.2. Thesis Objectives

Goals were set in order to conduct this study, and they include global and specific objectives.

1.2.1 Main objective

The principal objective of this work is to develop a scientifically sound, seasonal forecasting system to improve the quality and usefulness of forecast outputs by systematically searching for the best predictors and lead times using both statistical (non-linear and non-stationary) and hydrological models..

1.2.2 Specific objectives

This research work aims to achieve the following specific objectives:

- (a) Develop and validate a method to generate synthetic series of precipitation
- (b) Develop and test non-linear and non-stationary approaches for seasonal rainfall and streamflow forecasting and compare their performance to the approaches already in use
- (c) Develop and calibrate a rainfall-runoff model for the Sirba watershed and use a temporal disaggregation method to forecast seasonal streamflow
- (d) Develop and calibrate a hydrodynamical model of the Niger River at Niamey to generate floodplain maps and quantify the probabilities of flood protection dykes being overtopped in Niamey

1.3. Research Questions and Hypothesis

1.3.1 Research questions

The following research questions are set out in four points:

Axis 1: How can the current performance of seasonal forecasts in West Africa using non-linear statistical approaches be improved on?

Axis 2: How can the seasonal forecast information at the operational level be made more useful, in a format suitable for the socio-economic actors (farmers and other producers)?

Axis 3: Can we reduce the uncertainties in forecasting using available climate data and other indicators of influence on the regional climate?

Axis 4: Is it possible to combine seasonal forecasting models for rainfall and streamflow to solve local flooding problems?

1.3.2 Hypothesis

In this thesis, the following hypotheses are tested:

Hypothesis 1: Numerical models of seasonal rainfall forecasts in West Africa (including multi-model approaches) give less realistic results.

Hypothesis 2: Statistical models of seasonal forecasts, although imperfect, provide comparable or higher performance than numerical approaches in terms of time and cost of implementation.

Hypothesis 3: The use of local climate characteristics coupled with global climatic indices is best suited for seasonal rainfall forecasting in West Africa.

Hypothesis 4: In operational mode, the information provided by the proposed statistical approach is effective, adapted to the socio-economic actors and help to deal with flooding issue.

1.4. Novelty

The novelties of this thesis are threefold:

1. The seasonal rainfall forecasts issued from ACMAD are mainly qualitative, and typically involve the probabilities of occurrence of three categorical events: above normal, normal, and below normal precipitation. Such forecasts are limited and are not so useful for the end users because large areas are considered, which does not take account of local climate conditions, as well as the information delivering format. Also, these forecasts do not give any information on the temporal distribution of the forecasted rainfall, which is very important in risk management. The forecasts from this thesis are probabilistic, and they are further used to manage inundation risk.
2. In the Sahel, in West Africa, the seasonal forecast models developed by researchers have mainly used linear methods where the relationship between rainfall and the predictor is considered to be constant, although in reality this cannot be. Also, the selection of predictors is done subjectively. In this thesis, the developed seasonal forecast approach is based on a sound method for selecting predictors with high predictive powers and optimal lag times. Moreover, the

developed models consider the rainfall/predictor relationship to change with time; and another set of models considers the rainfall/predictor relationship to change with rainfall magnitude.

3. The site application of seasonal forecasted rainfall and streamflow is a great innovation by this thesis, which uses them to solve local flooding problems from the Niger River (the third-longest African river).

Thus, this thesis contributes to highlighting the need to consider the selection of predictors as the most important step when forecasting West African seasonal rainfall. Also, it helps in understanding that the usual consideration of rainfall/predictor relationship parameters as constant is not really true.

1.5. Scope of the Study

In this thesis, statistical models are developed specifically for seasonal rainfall and streamflow forecasts. The study focuses only on statistical forecasting approaches, using predictors with a potential relationship with West African monsoon dynamics. These forecasting methods are explicitly based on non-linear approaches (e.g., regression trees, artificial neural networks) and non-stationary approaches, although some linear statistical methods are used for selecting potential predictors. In this study, rainfall is used as a predictand, while the predictors employed include air temperature (AirTemp), sea level pressure (SLP), relative humidity (RHUM), zonal wind (UWND), meridional wind (VWND), and sea surface temperature (SST). These predictors are used to forecast seasonal rainfall and streamflow. The developed runoff-rainfall model is based on the SWAT model, and the temporal disaggregation used in this work is focused on the fragment method. The hydrodynamical model was developed for the Niger River using the HEC-RAS model.

Therefore, the scope of this thesis is limited to those specified methods, predictors and models. The study is carried out on the Sirba watershed, a transboundary watershed shared by two countries, Burkina Faso and the Republic of Niger. It also involves the city of Niamey in the Republic of Niger.

1.6. Organisation of the Thesis

Based on the objectives, this thesis is presented in eight chapters.

Chapter 1: This chapter presents the thesis and its context, the problem statement, the objectives and scope of the study, and the organization of the thesis.

Chapter 2: This chapter presents the general framework of the study. It comprises an expanded explanation on the West African Sahel Monsoon, the Sahelian rainfall mechanism and variability. The potentials of seasonal forecasts in the West African Sahel are reviewed by presenting the teleconnections and the physiographic characteristics of the region.

Chapter 3: This part of the work provides a thorough background and theory on seasonal forecasting and hydrological modelling techniques. The different forecasting approaches for seasonal rainfall and streamflow used by the scientific community are reviewed, and a literature review is done on the statistical methods (e.g., neural network, principal component analysis, stepwise regression, regressions trees, non-linear principal component analysis, Bayesian method, simple/multiple linear regression) used to forecast Sahelian seasonal rainfall. This chapter also provides a review of the training techniques (e.g., boosting, bagging, leave-one-out cross-validation), on the methods used for the assessment of forecast results (coefficient of determination, Nash-Sutcliffe coefficient). The methods used for temporal disaggregation are reviewed and all the aspects regarding copula method, and HEC-RAS modelling. A comprehensive literature review on flood issues and related damages over the Sahel region is done in the last part of this chapter.

Chapter 4: The fourth chapter presents the study area with the characteristics of the West African Sahelian zone, and an inventory of the Sirba basin in terms of climate, geographic information, and environmental issues.

The various data used in this study are presented in this section, followed by their critical analysis, in order to assess their quality and representativeness across the study area. These data include climate data (rainfall and atmospheric data), hydrological data, land use data, soil data, and topographical data.

Chapter 5: This chapter presents the process developed and the results of seasonal rainfall forecasting. The developed approach for predictor selection and optimal lag time, the models allowing the relationship rainfall/predictor to change with time, the

models where the relationship rainfall/predictor changes with rainfall magnitude, and the models with constant parameters are all explained. The settled Bayesian model selection approach is also explained. The best predictors and their region in addition to the optimal lag time are presented. The assessments of the seasonal rainfall forecasting skills are detailed in this chapter regarding the neural network, regression trees and non-linear principal component analysis models. The performances of the developed forecast models are discussed and compared to those of the existing forecast models for the region.

Chapter 6: This chapter is mainly dedicated to seasonal streamflow forecasting. The developed approaches used for seasonal streamflow forecasting model development are all explained. This chapter also presents the methodology used to develop the Sirba SWAT hydrological model. This includes the calibration process and uncertainty analysis using the SWAT Calibration and Uncertainty Program (SWAT-CUP). The performances of the developed forecast models are discussed and compared to those of the existing forecast models for the region. The results obtained from the calibrated Sirba hydrological model are explained, and the forecasts of maximum monthly streamflows and annual mean streamflows are discussed. As well, the results from the non-linear models are presented.

Chapter 7: This chapter deals with all aspects of seasonal flood forecasting and the related overtopping. The copula method used to model the dependence structure of streamflows is explained. The development of the hydrodynamical model of the Niger River at Niamey is presented, and the HEC-RAS model is fully explained. The last part of this chapter details the concept behind the simple risk measure used to assess the probabilities of overtopping of the flood protection dykes at Niamey.

Chapter 8: Chapter Eight draws the global conclusions on all the work carried out in this study. This concerns the efficiency of the developed seasonal forecasting models and the probabilities of overtopping dykes, and the limits of the different models. The contributions of the present thesis to the knowledge of the scientific community are presented in this chapter. The perspectives resulting from the work done are presented in this section in order to allow the research on seasonal forecasting issues to progress.

Chapter 2. West African Monsoon and Sahel Rainfall: A Review

This chapter mainly focuses on the key element governing rainfall in West Africa. The West African Monsoon is fully described, including its general circulation and dynamics. The processes behind the West African Sahel rainfall are reviewed in order to provide a good understanding of the rainfall mechanism, the seasonal cycle and the different variabilities. Moreover, the potentials of seasonal rainfall forecasting are presented and detailed. The main drivers of the Sahelian rainfall are presented in this chapter, with a particular focus on the teleconnections, which change over periods of time. The different oceanic basins and other physiographic characteristics that impact Sahelian rainfall are presented, and the review in this chapter explains how the hydrological regimes and other activities depend mainly on the rainy season governed by the WAM.

2.1. West African Sahel Monsoon: General Circulation

The radiation balance across the whole system is found to be unbalanced. This balance is exceeded at the source regions (e.g., equator and lower layers), but is in deficit in the sink regions (e.g., the poles and higher layers). This uneven distribution is the origin of the atmospheric circulation at the global scale (Malardel, 2005; Bouali, 2009). Thus, to balance these energy gradients between sources and sinks, a meridian and vertical circulation is set up between the poles and the equator. However, the Coriolis force induced by the Earth's rotation prevents the direct transfer of energy between these two types of region and does not allow the creation of a single cell capable of transporting energy.

According to Palmen's (1951) model describing the general atmospheric circulation (Figure 2.1), there are three zones of wind circulation: Hadley cells found between the equator and 30° N and S, integrating the WAM and the circulation area of

WAM; Ferrel cells that lie at mid-latitudes (30° N to 60° N and 30° S to 60° S); and finally, the polar cells, north of 60° N and south of 60° S.

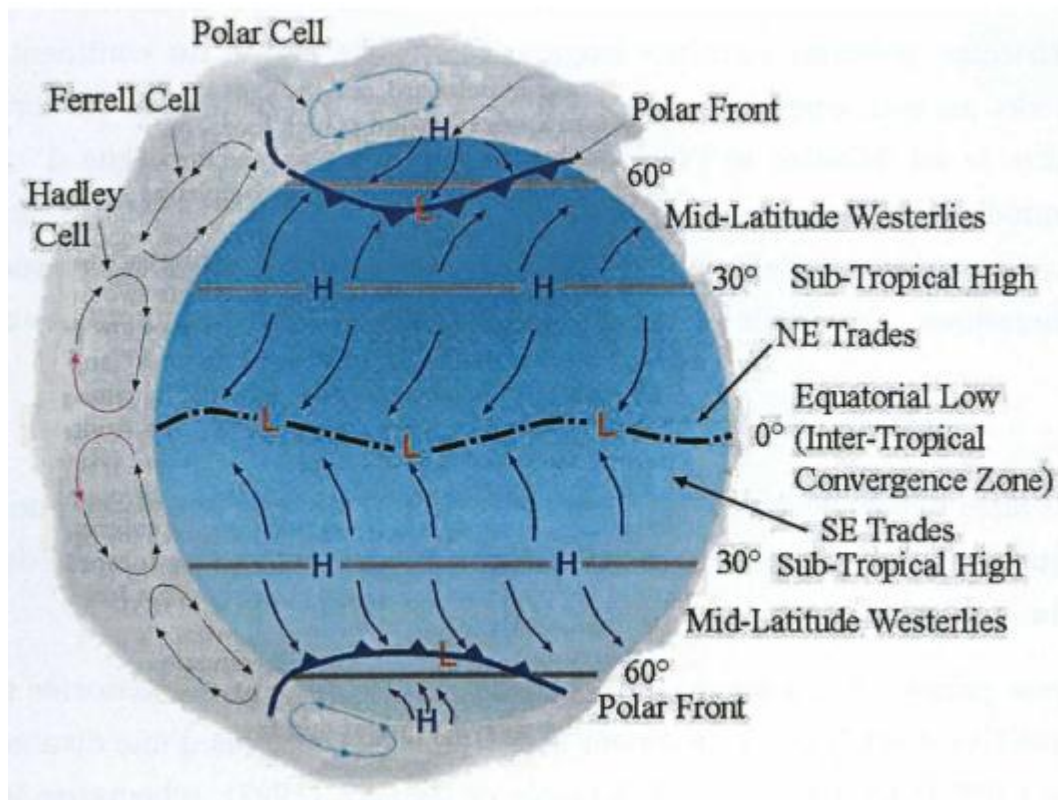


Figure 2.1 Atmospheric circulation: H and L mean layers of high and low pressure respectively (source: Palmen, 1951)

Unlike the cells of the Palmen model, the Hadley cells are asymmetrical and highly modulated by the seasonal cycle. The cells in the southern hemisphere (North) intensify in summer (winter) Boreal and lead to the displacement to the north (south) of the Inter-Tropical Convergence Zone (ITCZ). This area in between the cells of the northern and southern hemispheres is where the energy received by the atmosphere is at its maximum, thus characterizing a deep moist convection zone near the equator (Fontaine, 1993). The ground track of the winds of the two hemispheres of confluence zone is associated with a dry convergence zone called the ITCZ, less developed vertically and further north than the ITCZ. Hadley cells are characterized by two vertical branches: considering the movement of the air, a first ascending branch above the main energy source zones (equatorial), and a second downward one over the sink regions (subtropics to 30°N and 30°S), with air movement from sources to sinks in the lower layers and

sources to the sinks in the upper layers of the troposphere (15 km). Apart from the transfer of energy through the meridian and vertical Hadley circulation, another traffic type is west (zonal circulation: Walker type) is being established in response to the existing contrasts between firstly, the warm waters from western regions of the ocean basins and cold regions in the east, and also between hot continents, with a low heat capacity, and the coldest oceans, with a high thermal inertia (Figure 2.2). According to Caminade (2006), there are two groups of Walker-type cell: the first group is the Pacific / Atlantic, carrying the latent and sensible heat to the east and geopotential energy to the west, and the second group is the Africa / Indian Ocean set, characterized by a reverse circulation from the previous group. As for the Hadley circulation, the Walker types are characterized by Pedigree areas generating rainy systems and areas of subsidence conducting to the mitigation of cloudy upstream systems. However, they have the characteristic of being less forced by seasonality due to the low annual variation of ocean zonal thermal gradients and their location at low latitudes. Several studies, including that of Zhao and Moore (2007), showed that the decline in rainfall over West Africa during the second half of the twentieth century is associated with a lower Hadley ascending branch and a more intense Walker descending branch.

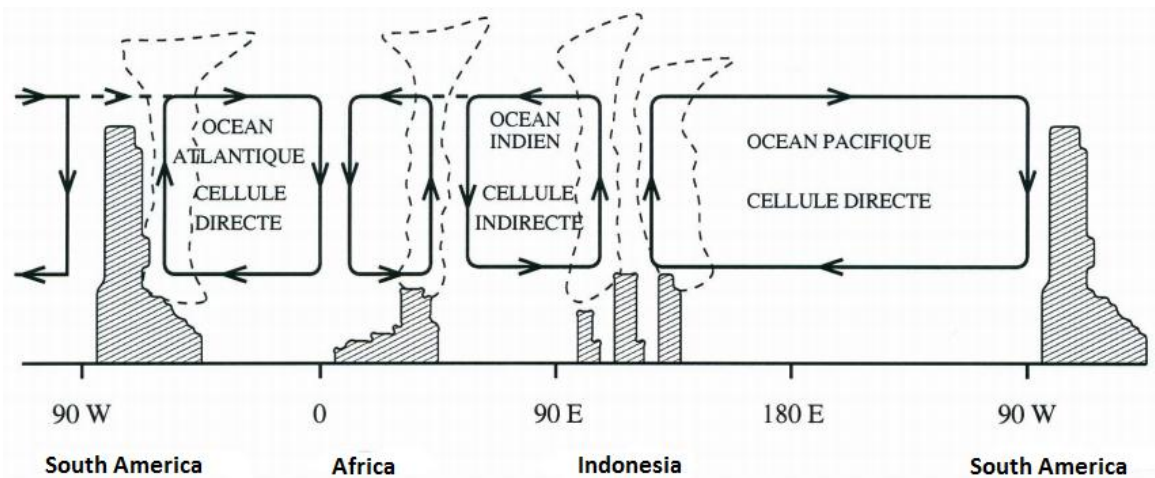


Figure 2.2 Zonal global circulation, Walker type (source: Caminade, 2006)

2.2. West African Sahel Rainfall

2.2.1 Rainfall mechanism

Soil-atmosphere-ocean interaction is the foundation of the West African Sahel climate mechanism (Lafore *et al.*, 2010). This relationship influences the dynamics of the regional ITCZ. The ITCZ is defined simply as the zone where two masses of air meet: a wet wind from the Atlantic Ocean to the south, and a hot dry wind from the Sahara desert to the north (Ramel, 2005). The high pressure Hadley cells and the depressed areas determine the dynamics of the two winds in the zone. The areas of high pressure, or anticyclones, are located within 30° North and 30° South (Figure 2.3). In this region, the wind dynamic is characterized by buoyancy from the Sainte-Helene anticyclone to the south and the Sahara anticyclone to the north. With respect to the circulation of these two air masses, they move in the lower atmospheric layers (less than 3 km altitude) and are known as trade winds.

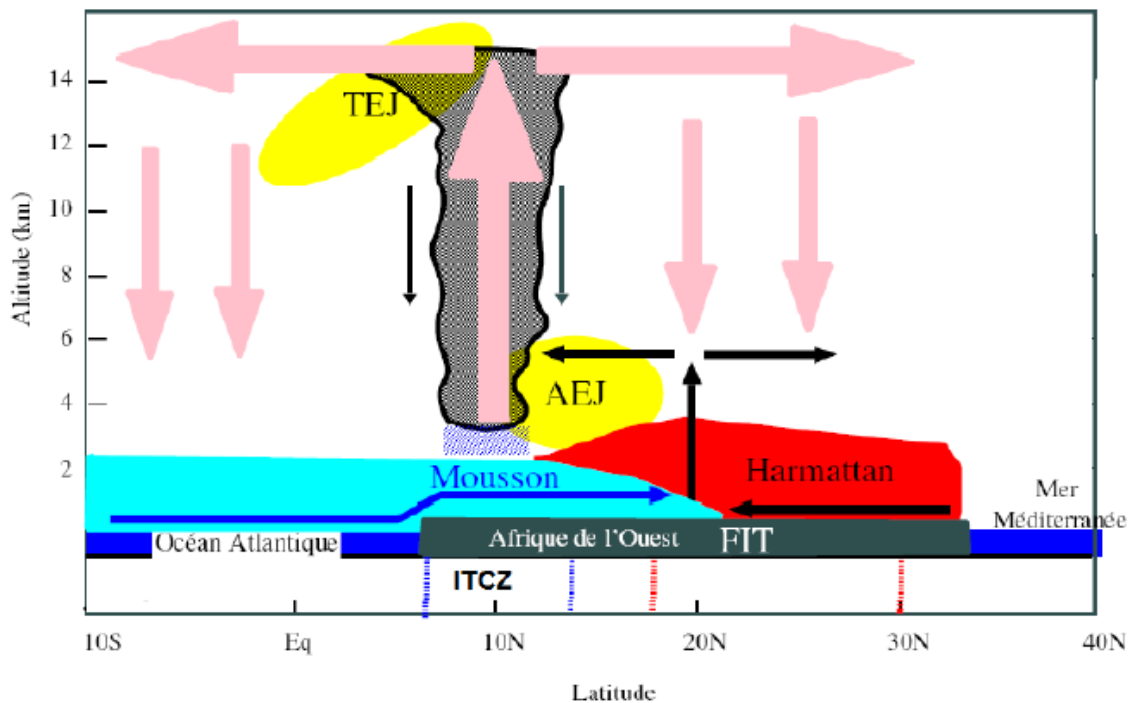


Figure 2.3 Conceptual scheme of the West African climate mechanism (adapted from Peyrillé, 2006)

The southern trade winds crossing the ocean and forest areas are full of water and are known as monsoon winds, while the trade winds from the north through the desert area are called Harmatan winds (warm and dry wind) (Figure 2.3). The temperature contrast between the continent and the Atlantic Ocean has a significant impact on the force of the southern trade winds (Fontaine & Bigot, 1993; Weldeab *et al.*, 2007), and the ITCZ delineates the northern limit of the monsoon front on the continent (Figure 2.3). Also, Gaye (2002) indicated that the ITCZ is the area of minimum pressure at ground level.

From June to September, circulation in the mid to upper atmospheric layers is largely dominated by two large air currents: the African Easterly Jet (AEJ) at 3 to 6 km elevation, and the Tropical Easterly Jet (TEJ) at 12 to 15 km elevation (Gaye, 2002; Gu & Adler, 2004; Parker *et al.*, 2005). It has been determined that these two air currents (Figure 2.1) have a significant impact on the monsoon rainfall, and that they act separately in different directions. A strong AEJ and a weak TEJ result in a rainy season with low rainfall, while rainy seasons with a low AEJ and high TEJ are characterized by heavy rainfall (Mahé & Citeau, 1993). However, Grist and Nicholson (2001) pointed out that a strong AEJ alone is not enough to generate seasonal precipitation; the latitudinal position of the AEJ further north also plays an important role and leads to higher rainfall.

Thus, depending on the intensity of the trade winds, the ITCZ moves from south to north (Figure 2.3) from February to August, and reaches its most northerly position at approximately 20° North. It then moves from north to south from September to January, and reaches its most southerly position at about 5° North (Penide, 2010). This movement of the ITCZ governs the repetition of dry and rainy seasons in West Africa, with dry seasons in areas located at north front of the two winds, as shown in Figure 2.4.

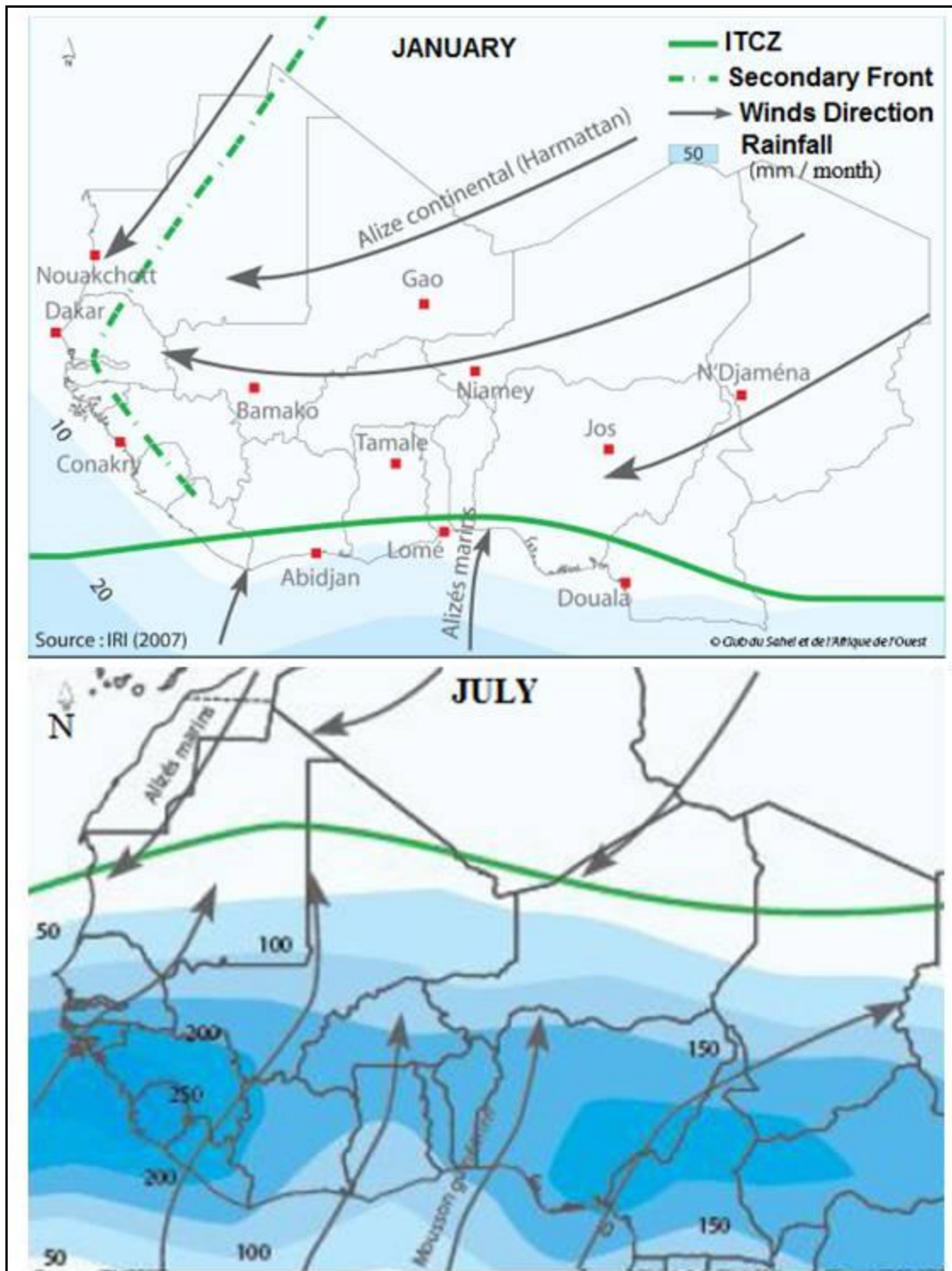


Figure 2.4 West African monsoon cycle (Adapted from CEDEAO & CSAO/OCDE, 2008)

The West African Sahel climate is mainly dominated by these two seasons (Figure 2.4), with the rainy season from April to October and the dry season from November to March (Sivakumar, 1988; Sultan & Janicot, 2003). The start of the rainy season is characterized by a rapid change in the quasi-stationary position of the ITCZ, from about 5° North in May and June to a second quasi-stationary position around 10° North in July and August (Sultan & Janicot, 2000).

Continental precipitation from the monsoon flux occurs in the form of rain or storm events which, on average, last less than 12 hours (LeBarbé & Lebel, 1997). These events may impact only a very limited area in the case of local convection, but with respect to organized convection or grain lines, they can cover a wide range. According to Gaye (2002), the grain lines generally move from east to west with very intense rainfall. Thus, the main element governing the atmospheric circulation over this region is the WAM, characterized by the opposite seasonal winds as described earlier. It is clear that the Sahelian rainfall pattern is season-dependent and directly related to the WAM.

2.2.2 Seasonal cycle and variability

The major element of the atmospheric circulation over West Africa is the monsoon circulation, which is characterized by a seasonal reversal of winds.

The word monsoon originates from an Arabic word pronounced as "*mawssim*" which simply means "season". In a general context, the monsoon is a seasonal reversal of wind direction, and on a global scale, the monsoon concerns the regions between 30° W-170° E / 25° S-35° N (Ramage, 1971).

According to the prevailing wind directions, the year can be divided into 4 periods:

- Boreal winter, corresponding to the northern regions during the period of the northeast trade winds
- The northern summer, which corresponds to the southwest monsoon
- Spring and boreal autumn, which correspond to transitions between the two previous types of circulation

Indeed, it is in boreal winter, especially in January, that the ITCZ is in its southernmost position in West Africa, around 2° N. Considering the atmospheric circulation area, this region is influenced by the northeast trade winds produced by thermal anticyclones from

Libyan area. The energy content, less than 335 kJ/kg over the northern parts of the region, illustrates the drought and stability of air masses associated with these flows from the northeast. On the West African coast, several energy gradients are observable, and they reveal the clash of two very different air masses: the flow of the Harmatan, dry and warm; and the southeast flow (from the Saint Helena anticyclone), warmer and very humid after going over the tropical Atlantic.

At a geopotential height of 200hPa, the circulation is dominated by westerly subtropical jets and zonal winds induced by temperature gradients between low and mid latitudes. These jets reach their maximum intensity during hemispheric winter. Regarding the SSTs from the Atlantic and Indian basins, they are characterized during such periods by maximum values at the position of the ITCZ. The Indian Ocean is however generally warmer and has more energy. In the northwest basin isotherms are oriented southwest / northeast either parallel to the northeast monsoon flow (Philippon, 2002).

From February and throughout the boreal spring, the ITCZ makes a slow recovery and jerkily towards the northern hemisphere. In the West African region, the increase in accruals in Guinea from April is not accompanied by a noticeable increase in isohyets towards the north. The evolution of the SST for the two basins is characterized by a maximum thermal expansion and movement of these maxima towards the north to the Indian Ocean. The most important changes between winter and spring in northern ocean concerns the currents.

The northern summer corresponds to the rainy season in the northern areas 7° N (Sahel), while the Guinean zone and the south equatorial area experiencing small and large dry season respectively. This distribution is associated with the northerly position of the ITCZ: in July it is localized around 10° N in West Africa after rising from latitude 5° N at the end of June (Sultan & Janicot 2000).

As for the transitional period of boreal autumn, this is the period of withdrawal of the ITCZ to the south. The ITCZ in October moves to a position comparable to that of April: around 5° N in West Africa. This is followed by a widespread and sharp decrease in rainfall over the Sahel and increased accruals over the lower Guinean coast. In the atmospheric circulation area, this decline is accompanied by a withdrawal of the

southwest flux. From the perspective of the SST, the northern autumn marks a return to warmer waters in the Gulf of Guinea and along the coast (Knox, 1987; Hastenrath *et al.*, 1993).

2.3. Main Drivers of Sahel Rainfall

2.3.1 Teleconnections between Sahel rainfall and oceanic basins

The Sahelian rainfall process has been studied for several years in order to understand the mechanisms behind each part of its process. Over the last two decades this has mobilized many laboratories and other research centers, both from Europe and Africa, to create a consortium called AMMA (African Monsoon Multidisciplinary Analyses) in order to help the scientific community obtain a better understanding of the WAM and the rainfall process.

Although the dynamics of the WAM are better understood nowadays, the main challenge with regard to its variability and predictability comes from its varying teleconnections. A teleconnection is the linkage of climate variables between two different areas, which may be close to or far away from each other (Kwon *et al.*, 2009; Liu & Alexander 2007). Teleconnections between Sahelian rainfall and the oceanic basins have changed greatly since the 1960s: the tropical Atlantic had the strongest influence during the years 1960–1970, and then the equatorial Pacific (i.e., El Nino Southern Oscillation) during the 1980–1990s, and the Mediterranean now (Gaetani & Fontaine, 2013; Jung *et al.*, 2006; Camberlin *et al.*, 2001; Rowell, 2001; Chase *et al.*, 2014, Fontaine *et al.*, 2011).

In West Africa, the rainfall pattern is firmly related to the seasonal movement of the ITCZ and, consequently, to the development of the WAM circulation (Olivry, 1993). Rainfall in the region presents extremely intermittent characteristics, particularly in the Sahelian zone where it is often associated with convective systems (Janicot, 1992). These convective systems interact with the AEJ and the equatorial Rossby waves that modulate the oceanic water flux on the continent. Lambergeon (1977) showed the inhibitory influence of the Azores and Sainte-Helene anticyclones on precipitation in the Sahelian region, and the reinforced effect of their coupling. He concluded that when the influence

of the Azores anticyclone is high and that of Sainte-Helene low, there is an increased deficit of rainfall in the Sahel, and also that the converse is true.

Mahé and Citeau (1993) proved that westward movement of the Sainte-Helene anticyclone in July to August weakens the monsoon circulation over the West African Sahel. This occurs when there is a dipolar structure of sea surface temperature (SST); that is, warmer in the South Atlantic and the equator, and colder in the North Atlantic. On the other hand, the impact of this dipolar structure on the sea surface pressure can also lead to the same phenomenon of lower values in the South Atlantic and the equator, and higher values in the North Atlantic (Janicot & Fontaine, 1993). These configurations follow the position of the ITCZ, which tends more southerly than the normal in the eastern Atlantic.

The SST has also been a key factor in the variability of the WAM, and consequently the Sahelian precipitation during the late 20th century, as shown in several studies (Garcia-Serrano *et al.*, 2013; Gaetani & Mohino, 2013; Mohino *et al.*, 2011; Solomon *et al.*, 2011; Folland *et al.*, 1986; Rowell *et al.*, 1995). Similarly, the tropical Atlantic Ocean is considered as the principal source of moisture for West Africa. Its impact on the WAM system has been known since the 1970s. This prompted Lamb (1978) to associate the Sahelian rainfall deficit to a colder SST in the north tropical Atlantic and to a warmer SST in the south Atlantic and at the equator, which promotes a southernmost ITCZ than the normal. These results have been confirmed and extended over longer time scales in many studies, including Hastenrath (1984, 1990), Druyan (1991), Lamb and Pepler (1991), and Janicot (1992). However, for Janicot *et al.* (1996), the relationship between Sahelian rainfall and Atlantic SSTs post-1970 substantially decreased to a point as to be statistically insignificant during the dry period. This teleconnection changed from Atlantic SST to the SST over eastern and center Equatorial Pacific, in agreement with the work of Moron (1994).

On a regional scale, Vizzy and Cook (2001, 2002) used an atmospheric GCM to show the sensitivity of West African precipitation to SST anomalies from the Gulf of Guinea. They linked positive SST anomalies to high evaporation in the Gulf of Guinea advected toward the Guinean coast. On the other hand, despite the distance between the African continent and the Pacific Ocean, there has been a significant statistical

relationship between the SST of the eastern equatorial Pacific and West African rainfall over the past 35 years (Janicot, 1997).

A rainy season with below average rainfall in West Africa is generally associated with a warm period of El Niño-Southern Oscillation (ENSO). Janicot *et al.* (1998) explained this relationship by a strengthening of the Walker circulation and a weakening of both the monsoon and the southern cell of the Hadley circulation. This state leads to an increase in trade winds over the northern tropical Atlantic and a reduction of water vapor in West Africa.

Clarke and Lebedev (1996) also argued that the warm El Niño event of the 1990s would have a greater impact than the same event at the beginning of the 1940s or 1970s. Janicot *et al.* (1996, 2001) also showed that teleconnections between the ENSO and Sahelian precipitation are not stationary. There are periods when the teleconnections are strong, and other periods when they are weak and dominated by another type of forcing. Their work highlighted the uncertainty between studies showing no significant signal of ENSO on Sahelian precipitation (Ropelewski & Halpert, 1987, 1989; Nicholson & Kim, 1997) and those that clearly pointed to the importance of ENSO forcing on the interannual fluctuations of WAM (Folland *et al.*, 1986; Hastenrath *et al.*, 1984; Ward, 1992; Palmer *et al.*, 1992). Shaman and Tziperman (2011) found that the interannual rainfall variability over the Mediterranean region is related to ENSO variability in the eastern Pacific via an eastward-propagating atmospheric stationary barotropic Rossby-wave train. Moreover, Lopez-Parages *et al.* (2012) explained how the teleconnection with the ENSO appears modulated by multidecadal oscillations of the SST over the Atlantic and Pacific basins.

Rowell *et al.* (1995) also showed a significant correlation between the equatorial Pacific SSTs and Sahelian precipitation on an interannual scale for periods less than eleven years. Ward (1992) noted that this correlation was even higher during years of cumulative rainfall deficit in West Africa. The work of Trzaska *et al.* (1996) on the effects of the Indian Ocean on Sahelian rainfall helped highlight a mode of an extratropical SST, representing a multi-decadal trend that describes a net reverse (i.e., hot in the south and cold in the north) of anomalies from the 1970s. Using filtering, Giannini *et al.* (2003) categorized the Sahelian precipitation time series as high- or low-frequency,

then applied a multi-decadal scale to associate the rainfall deficit in the Sahel with warm SSTs in the southern Atlantic, and particularly in the Indian Ocean. In a similar study, based on different atmospheric GCMs forced and fixed by observed positive monthly SST anomalies from 1950 to 1990, Hoerling *et al.* (2006) analyzed the role of the Indian Ocean on Sahelian drought. They noted that during the monsoon seasons (JAS) of 1950 to 1955, an increase in precipitation in the central and eastern Sahel between 15° and 20°N occurred simultaneously with a reduction in the western Sahel and central Africa. Hastenrath and Wolter (1992) explained that the rainfall deficit in these regions was due to warming of the Indian Ocean, while Bader and Latif (2003) regarded it as a direct response to the Indian Ocean anomaly.

In addition to the three major oceans, the Mediterranean Sea also appears to impact the WAM system. Raicich *et al.* (2003) and Rowell (2003) brought the Mediterranean Sea into the discussion due to its possible interactions with the WAM. Rowell (2003) found that the influence of temperature anomalies from the Mediterranean on the Sahel (for 1947-1996) is of similar magnitude with that of the Pacific. From numerical simulations, he showed that a warm Mediterranean promotes excess rainfall in the Sahel. Gaetani *et al.* (2010) and Polo *et al.* (2011) highlighted a teleconnection with the Mediterranean Sea. Additionally, Gaetani *et al.* (2010) pointed out that a positive precipitation response to warmer than average conditions in the Mediterranean Sea is found in the Sudano-Sahelian belt in August to September.

2.3.2 Other drivers of Sahel rainfall

The scientific community seems to have reached a consensus that the decadal variability of the West African climate is driven by variations in ocean surface temperatures. However, there are several other factors in play that could affect rainfall patterns, such as surface wind fields and altitude. For example, continental surface processes were shown to modulate or enhance ocean-climate signals (Zeng *et al.*, 1999). These processes often take place at regional levels gradually (i.e., day by day) due to atmospheric conditions. They can also occur at local levels of surface moisture and topography. Much further study on West Africa is needed to determine the degree of heterogeneity in surface properties (Ramel, 2005).

Regarding the role of the land surface on Sahelian rainfall variability, Webster *et al.* (1998) indicated that looking at moist static energy (MSE) (its three components: sensitive, latent and potential) could improve precipitation forecasts for the Sahel. They argued that the variation in temperature between the ocean and the continent is responsible for the monsoon circulation, and that the circulation is even better when the moisture gradient is considered. Eltahir (1996), Philippon & Fontaine (2002), and Hall and Peyrillé (2006) highlighted the role of these gradients on the dynamics of WAM and Sahelian rainfall variability, particularly at lower atmospheric layers. They explained that monsoon development would correspond to a dynamic response of the atmosphere to energy contrasts. Peyrillé (2006) emphasized the importance of MSE and the impact of these gradients at two different levels: the local level, where the MSE vertical gradient affects the rainy systems; and the regional level, where the MSE horizontal gradient affects the intensity of WAM circulation. According to Eltahir (1996), the distribution of vegetation and soil moisture on the continent is reflected by the MSE horizontal gradient effects. Zheng and Eltahir (1998) found that changes in vegetation (e.g., deforestation) on the Guinean coast have a direct substantial impact on atmospheric dynamics associated with the monsoon circulation through the MSE gradients. Thus, some authors, such as Wang and Eltahir (2000), have suggested the inclusion of vegetation dynamics in the modelling exercises, as it constitutes an important process for simulating Sahelian rainfall variability. In addition, while testing the impact of vegetation on rainfall variability simulation, Zeng *et al.* (1999) found that the decadal variability, which is an essential factor in Sahelian rainfall variability, is best reproduced when vegetation is added to the model.

The role of soil moisture in West African rainfall events is also addressed in some studies (e.g., Philippon & Fontaine, 2002; Cook, 1999). These authors emphasized that a positive soil moisture anomaly would strengthen the monsoon circulation through a modification of MSE gradients. It should be noted that their research was based on the theoretical work of Emanuel (1995). Douville *et al.* (2001) and Douville (2002; 2003) found that any reduction in soil moisture in ARPEGE is associated with low intensities of precipitation. Based on these results, they concluded that soil moisture contributes to the interannual variability of rainfall in the Sahel.

Several studies have pointed out the effect of geopotential heights on forecasting Sahelian precipitation. Bouali (2009) showed that winds at 850hPa have less impact on the monsoon flows over the continent, while those at 500hPa only partially describe the AEJ. However, in the reanalysis data, the author found a strong gradient throughout West Africa, based on correlations between Sahelian rainfall and zonal wind component at 500 hPa. Parker *et al.* (2005) confirmed this result, and suggested assessment of geopotential heights between 600 and 700hPa for high impacts on West African Sahelian rainfall. In addition, while analyzing the correlation between the Sahelian rainfall index and zonal wind reanalysis at 200hPa, Bouali (2009) found significant negative values for West Africa and the Atlantic Ocean.

2.4. Synthesis and partial conclusion

The atmospheric circulation over West Africa is mainly governed by the monsoon described by a seasonal reversal of winds. The WAM is in turn the key factor for precipitation occurrence over the Sahel. The rainfall pattern is firmly related to the seasonal movement of the ITCZ and, consequently, to the development of the WAM circulation. There are different climate factors influencing the Sahelian rainfall, known as teleconnections. Unfortunately, these teleconnections have changed over time, as illustrated by the change of oceanic basins impacting Sahelian rainfall from the 60s, 80s and 90s.

It therefore appears that the climate in West Africa (predominantly in the Sahel) is determined by interactions between global processes (e.g., sea surface temperatures) and regional processes (e.g., physiographic characteristic). The use of parameters related to these processes in seasonal rainfall forecasting models could generate more accurate forecasts for the Sahel.

Chapter 3. Background and Theory on Seasonal Forecasting and Hydrological Modelling

This chapter gives a thorough background and review on all modelling aspects involved in this thesis. This helps in better understanding the theory behind each concept. It gives details on the three different and complementary approaches used by the scientific community for seasonal forecasting, and the different statistical techniques for seasonal forecasting are fully presented, including both their advantages and weaknesses. The change point detection concept is explained, and promising works are entirely reviewed. The different steps in hydrological modelling are also presented, with a particular focus on the calibration and validation steps. The methods used to evaluate the performance of forecast models are described and their limits are depicted. Temporal disaggregation with its different approaches is presented. The theories on seasonal flood forecasting and the assessment of probabilities for flood protection dykes to be overtopped are presented.

3.1. Seasonal Forecasting

According to Udall and Hoerling (2005), seasonal forecasting prediction is the estimation of what the climate will be over the next few months. It is currently the best means of addressing climate issues, and is also considered the most efficient approach to drought mitigation (Hayes *et al.*, 2005). Its main objective is to predict the evolution of the climate system on time scales of months to seasons, at lead times beyond the limit of predictability of the instantaneous state of the atmosphere (Zwiers & Von Storch, 2004). METEO France defines seasonal forecasting as a probabilistic estimate of precipitation on a spatio-temporal scale of the season. Hence, it generally consists of searching for statistical relationships between indicators of the rainy season and parameters characterizing the state of the atmosphere, including oceans. Regardless of the approach,

the predictive ability is usually derived from teleconnections between the atmosphere and slowly varying boundary conditions, such as sea surface temperatures (Quan *et al.*, 2006) or soil moisture (Douville, 2003). Due to the long lead times and averaging of time scales involved in seasonal climate prediction, the number of observation samples available for training statistical models is usually small (George & Sutton, 2006).

Seasonal forecasting has several benefits. For instance, it helps decision-makers and stakeholders in various sectors such as water resources management, agriculture, risk management and disaster mitigation. Garric *et al.* (2002) indicated the importance of seasonal rainfall forecasting for northern Africa.

3.1.1 Seasonal rainfall forecasting

Over the last few decades, climatic fluctuations in the West African Sahelian countries have increased interest in rainfall forecasting due to the impacts of climate change / variability that frequently lead to severe disasters which endanger populations. These effects are not only due to the low rainfall in the annual monsoon season (Ali *et al.*, 2004), but also to the strong interannual and decadal variability of the monsoons (Sultan & Janicot 2004; Janicot *et al.*, 2001). The consequences have a strong impact on rural development sectors such as water resources, agriculture and health.

Diverse research in hydrological forecasting conducted by international weather centers (e.g., France, England, United States of America) has underscored the potential of seasonal precipitation forecasting in West Africa through the statistical relationship between oceanic basins and observed rainfall. Subsequently, some numerical and statistical forecasting models have been developed to predict seasonal rainfall and the streamflow of large rivers in the region.

This substantial interest in seasonal forecasting (Palmer & Anderson, 1994) has prompted the establishment of several models. Three different and complementary approaches are currently used (Hastenrath, 1995): the dynamical approach (based only on numerical models), the statistical approach (based only on statistics), and the hybrid statistical-dynamical approach (a combination of numerical models and statistics).

a) Dynamical Approach

The dynamical (or numerical) approach is based on equations of the physics and dynamics that describe the climate system (Kumar *et al.*, 1996; Brankovic & Palmer, 1997; Palmer *et al.*, 2000; 2004). This approach is most appropriate when time referencing is essential (e.g., the optimization of reservoir management rules). Unlike statistical methods, the numerical approach provides results in the form of time series, and is based on models representing the water cycle at various levels. There is a variety of numerical models, most of which are based on precipitation and temperature series. The latter is the common characteristic constraining modellers' choices, and it can facilitate and improve the comparison of different models. The first numerical forecast was introduced in 1922 by Britain's Lewis Fry Richardson, who attempted to numerically solve the forecasting equations of atmospheric behaviour developed by Vilhelm Bjerkness in 1904. Richardson's initial approach was progressively improved upon over the years, most notably in the 1970s with the development of the Global Climate Model (GCM), as well as coupled ocean-atmosphere GCMs and Regional Climate Models (RCMs). Much of the later numerical approach improvement is undoubtedly related to the progress in computing system tools (Coiffer, 2000) as well as consideration of physical, atmospheric and oceanic phenomena, which helped focus on the interactions between the components of the ocean-atmosphere system. This approach is based on boxes, and the atmosphere is schematically divided into a horizontal grid with different vertical heights. Each box is described by certain parameters, such as wind, humidity, pressure and temperature.

Although systematic improvement of forecasting scores using numerical models has been presented in several studies (Kharin & Zwiers, 2002; Krishnamurti *et al.*, 2000; Palmer *et al.*, 2004), other studies have showed that the advantages for seasonal forecasting were marginal compared to the best model (Doblas-Reyes *et al.*, 2000; Graham *et al.*, 2000; Peng *et al.*, 2002). Some of these studies showed that a multi-model approach that assigns different weights to each model (known as a super-ensemble) has better performance than an approach based on only one forecasting model, or on a simple average of several forecasting models (Krishnamurti *et al.*, 2000; Pavan & Doblas-Reyes, 2000; Yun *et al.*, 2003). However, several recent works have found that a super-ensemble

does not do better than a simple average approach when the period under study is short (Kharin & Zwiers, 2002; Peng *et al.*, 2002).

b) Statistical Approach

The statistical approach is mainly based on the characteristics of period of forecast and event occurrence. Thus, it can be defined as a technique to establish a direct relationship between the state of the atmosphere or ocean at the moment of the forecast and during event occurrences (e.g., precipitation), within a range of a few weeks or months. The existence of sufficiently strong and robust physical links between certain variables is regarded as a source of foreseeability, and is the basis of the statistical forecast. Predictions based purely on the statistical approach have contributed considerably to the discovery of relationships between precipitation and other climate variables (Schepen *et al.*, 2012; Lopez-Bustins *et al.*, 2008). Several statistical methods and tests have been proposed in the literature (Ibrahim *et al.*, 2014; Sittichok *et al.*, 2014; Bouali, 2009; Biasutti *et al.*, 2008; Hayes *et al.*, 2005, Philippon 2002; Fontaine *et al.*, 1999; Vautard *et al.*, 1996; Barnston, 1994), and they have become quite popular due to their ease of development and the limitations of the dynamical models. Philippon (2002) showed the importance of using statistical methods when searching for a relationship between seasonal precipitation and atmospheric variables. Establishing this relationship was based on statistical methods used individually and in combination.

However, it is important to check that the statistical relationships are simulating real physical processes of the climate system, which is obvious when dealing with atmospheric predictors. Statistical models require less computing resources than numerical models, and they can often make distinct contributions when the predictors associated with the target region (e.g., watersheds) are clearly established. In many circumstances, statistical seasonal forecasting models deliver comparable or better forecasting abilities than numerical models (Fontaine *et al.*, 1999). Due to the difficulty numerical models have in dealing with the complexity of natural systems, statistical models are still frequently used.

Statistical forecasting models are based on different types of statistical methods: linear and non-linear; and which of these is best depends on the constraints of the study. In West Africa, linear methods are mostly used for seasonal forecast studies, and non-

linear statistical methods are rarely used. Despite being little-used in modelling, these methods can improve seasonal forecasting compared to some other methods. According to French *et al.* (1992), rainfall is one of the most complex and difficult elements of the hydrological cycle to understand and model due to the great variation from one scale to another in both space and time. Quantitative rainfall forecasting is extremely difficult because of the complexity of the atmospheric processes, particularly the WAM mechanisms that generate rainfall. Many advances have been made in weather forecasting in recent decades, but accurate rainfall forecasting still remains one of the biggest challenges in operational hydrology (Gwangseob & Ana, 2001). This is due to the fact that not all atmospheric relationships are linear, though in several cases they have been assumed to be (e.g., the relationship between rainfall and streamflow). This was confirmed in an investigation by the UK Met Office Hadley Centre (2010), which found that the rainfall variability in the Sahel is the result of complex interactions between several processes. Thus, no one aspect seems to be able to explain Sahelian rainfall variability.

Non-linear methods are rarely applied in climate forecasting compared to linear methods due to the former's implementation difficulties related to programming languages and interpretation. Nevertheless, non-linear methods can improve the predictability of weather forecasting, since most climatic phenomena do not occur linearly. The main differences between linear and non-linear methods are as follows.

Mathematically, a function F is said to be non-linear only if it obeys Equation (3.1), which expresses that the response of the function is non-additive.

$$F(a_1X_1 + a_2X_2) \neq a_1F(X_1) + a_2F(X_2) \quad (3.1)$$

Non-linearity is fundamentally a characteristic of the physical climate system. A well-known example is the set of non-linear differential equations that govern the flow of the atmosphere and oceans.

Given the mathematical definition of non-linearity, a further distinction is made between non-linear statistical models that are linear in terms of the parameters to be estimated, and those that are non-linear based on those parameters. In the former case, the functional relationship described by the statistical model is non-linear (i.e., its response is non-additive), but derivatives from the function with respect to its parameters that do not

contain such parameters. In this case, the non-linear relationship can be made linear by applying a suitable transformation to the predictor (and/or predictand) variables. For example, the function in Equation (3.2) is non-linear by the same logic as in Equation (3.1), but linear in terms of parameter a .

$$F(X) = ae^X \quad (3.2)$$

The response can be made linear by taking the natural logarithm of X . However, the function (see Equation (3.3)) is non-linear with respect to both the meaning of Equation (3.1) and parameter a .

$$F(X) = 1/(1 + e^{aX}) \quad (3.3)$$

c) Hybrid statistical-Numerical Approach

The hybrid statistical-numerical approach, also known as model output statistics (MOS), is a mixed method based on the principle of applying statistical techniques on the output of numerical models (Wilks, 2006; Garric *et al.*, 2002). According to Hastenrath *et al.* (1995), the application of MOS involves five steps (see Figure 3.1), as summarized below:

- Conducting a thorough diagnosis of the numerical model output in order to understand the different interactions in the model in relation to the variable to be forecasted
- Testing and selecting the potential predictors
- Applying statistics to determine the relationship between the predictand and the selected potential predictors. This model can be set to use n couples of predictand-predictor values
- Predicting the observed value for a given year i , using data of that year that were not used during the model development phase
- Comparing the simulated data to the observed values, thereby validating the model's performance



Figure 3.1 The steps of a statistical-numerical approach (adapted from Hastenrath *et al.*, 1995)

3.1.2 Seasonal streamflow forecasting

There are two main approaches used in forecasting streamflow in a watershed: the direct method and the indirect method. The first approach consists of using some relationships (usually empirical) between streamflow at a given date and the observed predictors at a period prior to that date, while in the second method, the rainfall is first predicted using predictors and then fed into a rainfall-runoff model to get the streamflow.

In the first approach, statistical methods based on empirical relationships between historical streamflows (predictand) and hydrological or climate variables (predictor) are used. Most research on the Sahel has focused on the empirical relationships between climatic and oceanic variables with rainfall. Therefore, several studies have been carried out to find the appropriate statistical techniques and the best predictors to forecast streamflows for specific regions (Abudu *et al.* 2010; Kwon *et al.* 2009; Piechota *et al.* 1998). This direct method is rarely used, especially in the case of Sahelian streamflow forecast studies, as the majority of studies in the literature have used the second approach (i.e., the indirect method). Nevertheless, it is assumed in the few existing studies that streamflow is linked to the predictors by a multiple linear regression with parameters that are independent of time and predictor magnitude, though this fact is not really established.

Regarding the indirect method, statistical methods (Sittichok *et al.*, 2014; Garric *et al.* 2002; Barnston *et al.* 1996; Folland *et al.* 1991) or climate models (Thiaw & Mo 2005; Folland *et al.* 1991) are first used to forecast rainfall before using a runoff-rainfall model. However, it is noted that biases are usually introduced by climate models, but such biases are corrected using model output statistics techniques (Amadou *et al.* 2014; Ndiaye *et al.* 2011; Bouali 2009). In order to forecast seasonal streamflow, the predicted rainfall is used to run the rainfall-runoff model (Tuteja *et al.* 2011; Yossef *et al.* 2013;

Ndiaye *et al.* 2009). This approach has several advantages as it is a more realistic simulation of streamflow, since physical interactions between land, climate and river characteristics in the watershed are accounted for. Some authors (Wang *et al.* 2011; Yossef *et al.* 2013) have shown that rainfall-runoff models can produce consistent streamflow results once they are forced with accurate rainfall forecast and properly initialized using antecedent hydrological variables before the forecast is issued. One more advantage is that rainfall-runoff models are less sensitive to outliers in the input data compared to statistical models (i.e., the direct method). Nevertheless, the beneficial use of flow simulations apparently depends on the reliability of outputs and the lagged-time (in advance) flow that can be modeled. However, it should be noted that the number of studies dealing with seasonal streamflow forecasting is far less than those dealing with seasonal rainfall forecasting, despite the great need in the Sahelian region to have prior knowledge of the quantity of runoff in order to plan social and economic activities.

3.1.3 Seasonal flood forecasting

It is obvious that West Africa has been experiencing more intense flooding over the last few years. Even with flood management strategies in place, these events still continue. Various factors contribute to this, including the management of urban environments, the consequences of poverty, and the sensitivity to natural disasters (Awotwi *et al.*, 2015). It is now essential to assess and upgrade the capacities of stakeholders to manage and adapt to such circumstances. Strategies have been implemented to address the impacts on populations and cities, but they need more evaluation and tailoring before they can strengthen adaptation capacities enough to cope with the challenges. Some of these measures are simply not realistic, as they lack thorough preliminary study and do not take into account the characteristics of the area being considered (e.g., land use, agricultural practices, soil science).

Seasonal flood forecast models are usually based on a hydrological model that predicts the water level (Laganier *et al.*, 2004) and sometimes the extent of the flood area. This kind of forecasting is limited, as it cannot provide enough information to stakeholders in advance about impending floods in order to alert the population (O'Connor, 2006; Perrin *et al.*, 2001). Thus, a good seasonal flood forecasting should

involve a hydrodynamical model that helps simulate water movements (currents, waves, turbulence, etc.), compared to a simple hydrological model, which is limited. Additionally, a good seasonal flood forecast should be based on probabilistic forecasts over years (or seasons) that could further help to enhance the flood warning system (Gado Djibo *et al.*, 2015a). This enhancement, using a probabilistic approach to assess the overtopping of flood protection dykes (Asselman *et al.*, 2009), could give rise to a better estimation of the changes and effects on the seasonal flood patterns, or a better cost/benefit estimation of the seasonal flood (Ngoran *et al.*, 2015).

3.2. Statistical Tools for Seasonal Forecasting

The following paragraphs present a quick review of some statistical methods used for seasonal forecasting.

3.2.1 Principal component analysis

Principal component analysis (PCA) is increasingly being used in climate forecasting models, particularly for seasonal rainfall (Sittichok *et al.*, 2014; Ruiz *et al.*, 2005; Bader *et al.*, 2006; Bouali, 2009). Bouali (2009) used PCA to define the mode of Sahelian precipitation in GCMs and to extract coherent modes of spatial and temporal variability from the fields of precipitation simulated by digital models in West Africa. PCA has also been used to reduce dimensions or prevent collinearity at the level of the precipitation series, and to reduce the noise by choosing the first components to rebuild the precipitation series. The principle of PCA is to replace n correlated variables by new variables known as principal components (PC). PCs are uncorrelated linear combinations of maximum variance of the initial variables (Saporta, 2006). However, the procedure for variance maximization can create artificial components that do not represent true spatial and temporal structures, so in some cases a rotation of the factorial axes is required. This consists of redistributing the information in the first k components between new k components. The PCs from the PCA with rotation make it possible to better individualize and stabilize the space structures (Bouali *et al.*, 2008).

While searching for a relationship between the rough indices of rainfall anomalies and the natural mean flow of the Senegal River, PCA was conducted on the indices of the

anomalies at the level of each zone of the ARPEGE model (Bader *et al.*, 2006). This allowed expression of a major segment of the information in the set of indexes of the zone, starting from a number of independent variables (i.e., eigenvectors). Ruiz *et al.* (2005) also applied PCA to extract the major dominant modes from ocean surface temperatures while attempting to integrate El Nino into a statistical forecast analysis.

3.2.2 Non-linear principal component analysis

The chaotic evolution of weather, with its rapid and increasing differences between similar states, makes climate forecasting complicated because of the many non-linear processes. Identifying these non-linear aspects has become crucial for weather forecasters. Thus, the non-linear principal component analysis (NLPCA) algorithm developed by Scholz *et al.* (2007) has been improved and tested to forecast seasonal rainfall. Depending on the extensions used, NLPCA can be of different types, including hierarchical NLPCA, inverse NLPCA and circular NLPCA, all of which can be used individually or in combination. Hierarchical NLPCA forces the non-linear components to have the same hierarchical order as the linear components of a standard PCA. This hierarchical condition produces individual components of high significance. Inverse NLPCA considers non-linear PCA as an inverse problem with only the assumed data generation process being modeled, which provides the advantage that more complex curves can be identified. This makes NLPCA more useful for filling in missing data in incomplete datasets. Circular NLPCA helps non-linear PCA extract circular components that describe a closed curve, instead of the standard curve with an open interval. Circular NLPCA is very useful for analyzing data from cyclic or oscillatory phenomena.

NLPCA is considered to be a non-linear generalization of the standard linear PCA (Jolliffe, 1986; Diamantaras & Kung, 1996), with the principal components generalized from straight lines to curves. NLPCA is set using an auto-associative neural network (Kirby & Miranda, 1996; Kramer, 1991; Malthouse, 1998; DeMers & Cottrell, 1993; Hecht-Nielsen, 1995). With an auto-associative style, this network can be a replicator network, a bottleneck or sand-glass type network, or an autoencoder network. The auto-associative neural network (multilayer perceptron) helps match the input and output of a network, but with a middle layer that acts as a bottleneck and imposes a reduction in data

dimensions. This bottleneck layer provides the anticipated component values (i.e., scores).

NLPCA has been successfully applied in atmospheric and oceanic sciences (Hsieh, 2004; Monahan *et al.*, 2003), though it is new in the field of weather forecasting. In extracting components in either a linear or non-linear case, it is assumed that the data are determined by a number of factors; hence it can be considered as being generated by them. The data are located in a sub-space of the given data space because the number of varied factors is often less than the number of observed variables. The aim of NLPCA is to represent these factors by components which together describe the sub-space, as shown in Figure 3.2. The figure illustrates three-dimensional samples that are located in a one-dimensional sub-space, and can thus be described by a single variable (the component) without loss of information. The transformation is given by two functions (Φ_{extr} and Φ_{gen}). The extraction function (Φ_{extr}) maps each three-dimensional sample vector (on the left side) onto a one-dimensional component value (on the right side). The inverse mapping is given by the generation function (Φ_{gen}), which transforms any scalar component value back to the original data space (Scholz *et al.*, 2007).

The NLPCA approach used in this study is summarized in the following steps:

- First, the best predictors were selected based on their significant correlation with the seasonal rainfall data series.
- NLPCA was applied for dimensionality reduction and to screen and search for the best principal components (PCs), both linear and non-linear.
- Stepwise regression was applied to select the best PCs with high predictive power.
- Linear regression was used to predict the seasonal rainfall.

It should be noted that the number of PCs in the third step was fixed and monitored to be the first three best. This limitation of the PCs is due to time, as the higher the number of PCs, the more processing time required.

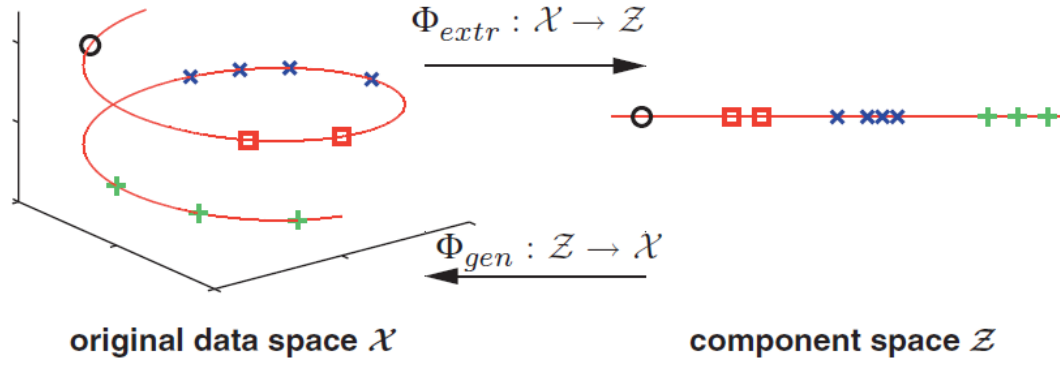


Figure 3.2 Non-linear dimensionality reduction in NLPKA (source: Scholz *et al.*, 2007)

3.2.3 Multiple linear regression

Multiple linear regression (MLR) is a mathematical method that helps explain a response variable y (predictand) and a series of explanatory variables X_i (predictors) by searching for a previous linear relationship between them.

If at a given time there are several observations of the couple (y, X_i) , the MLR is given in Equation (3.4).

$$y_j = b_0 + \sum_{i=1}^k b_i X_{ij} + \varepsilon(X)_j \quad (3.4)$$

where y_j is the j^{th} observation of the dependent variable, X_{ij} is the j^{th} observation of the i^{th} independent variables, k and $\varepsilon(X)_j$ are the number of explanatory variables and the residue of the observation respectively, j and b_i are the coefficients of regression, and b_0 is the coordinate at origin.

When building an MLR model, there should be more observations than explanatory variables, and $\varepsilon(x)_j$ is assumed to follow a normal distribution law. In this study, MLR is intended to predict seasonal rainfall/streamflow based on predictors as the last step of combined statistical methods.

Some previous works have been largely based on MLR models. For example, Bouali (2009) applied MLR to choose methods based on indices to forecast rainy seasons in West Africa with hybrid models; Roquelaure (2007) applied an MLR model to predict visibility when dealing with local overall forecasts of fogs and clouds; and Mason and Mimmack (2001) used MLR to generate probabilistic forecasts and their performance compared to numerical models. MLR models have also been developed for direct

forecasting of seasonal precipitation (June to September), and to choose predictors in ascending order (Korecha & Barnston, 2006). For improved flood forecasting for the Jhelum River in Pakistan, Archer and Fowler (2008) developed an MLR model using seasonal temperatures and precipitation as predictors.

3.2.4 Stepwise regression

Stepwise regression (SR) is a statistical method to select predictors through defined steps. It is a forward-selection screening technique that can address the problem of selecting the best set from a pool of potential predictors (Wilks, 2006). SR is an iterative method of choosing the most economical set of predictors that are effective in predicting the dependent variable (here, seasonal rainfall) on the basis of statistical criteria. Essentially, the technique determines which independent variable is the best predictor, which is second best, and so on. The emphasis is on finding the best predictor at each step.

When the predictors are highly correlated with each other and the rainfall dataset, one variable is often listed as a predictor and the other is not listed. This does not mean that the latter variable is not a predictor, but that it adds nothing that the first predictor has not already dealt with. Sometimes the best predictor is only marginally better than the second predictor, and minor variations in procedures (e.g., developed forecasting processes) could affect which of the two is chosen as the predictor. Thus, to find the best answer (i.e., best predictors) the SR process is repeated at all stages. In this study, the use of SR is justified by the large number of predictors (specifically, the number of PCs) that must be screened in order to achieve a pool of effective predictors according to their statistical contributions to determining the variance of seasonal rainfall. The detailed principles of SR are as follows (Jolliffe & Stephenson, 2012).

Suppose there are N potential predictors for a least-squares linear regression, and the forward selection process with the uninformative prediction equation is $\hat{y} = b_0$; that is, only the intercept term is ‘in the equation’, and thus it is the sample mean of the predictand.

Step1:

In the first step, all N potential predictors are examined to determine the strength of their linear relationships with the seasonal precipitation (predictand). In effect, all possible N simple linear regressions between the available predictors and seasonal precipitation are computed, and the predictor with the best linear regression among the candidate predictors is chosen as x_1 . At this stage of the screening procedure the prediction equation is $\hat{y} = b_0 + b_1x_1$. Note that typically the intercept b_0 will no longer be the average of the y values.

Step 2:

In the second step, more trial regressions are constructed using all the remaining $N-1$ predictors. However, these regressions also contain the variable x_1 ; that is, for any x_1 selected in step 1, the predictor variable yielding the best regression $\hat{y} = b_0 + b_1x_1 + b_2x_2$ is chosen as x_2 . This new x_2 will be recognized as the best because it produces the regression equation with $(P = 2)$ predictors that includes the previously chosen x_1 , with the highest R^2 and the smallest mean square error.

Step 3:

Successive steps in the selection procedure follow the same pattern; at each step the predictor in the potential predictor pool not yet in the regression is chosen. This produces the best regression in conjunction with the $(P - 1)$ predictors chosen in the previous steps.

Generally with SR, when the regression equations are recomputed, the regression coefficients for the intercept and the previously chosen predictors will change. The changes occur because the predictors are usually correlated to a greater or lesser degree, so information about the predictand is spread differently among the predictors as more are added to the equation (Wilks, 2006).

3.2.5 Artificial neural network

The artificial neural network (ANN) was derived from one of the first artificial neural models proposed in 1943 by Warren McCulloch and Walter Pitts.

A neural network, as the name implies, can be defined as an ensemble of neurons in a network that enables the output signals (i.e., *outputs*) of the neurons to become incoming signals (*inputs*) for other neurons. The general design of neural networks

represents neurons in successive layers. The first layer of the network is the *input layer*, the last layer is the *output layer*, and the intermediate layers are the *hidden layers*. They are called *hidden* because their operation cannot be clearly analyzed from outside the network; only the input and output signals of the network are known. The neurons of the input layer only standardize and distribute the input signals, and thus they cannot be considered as treating neurons (i.e., computing neurons). In this standardized architecture the neuron layers are completely interconnected; that is, all neurons in a layer are connected to all neurons of the adjacent layers. This standardization helps achieve a correct representation of the network, which can be used for more general training algorithms. In summary, the components involved in processing the main information are inputs, weights, weighted averages of input data, transformation functions, and outputs (see Figure 3.3).

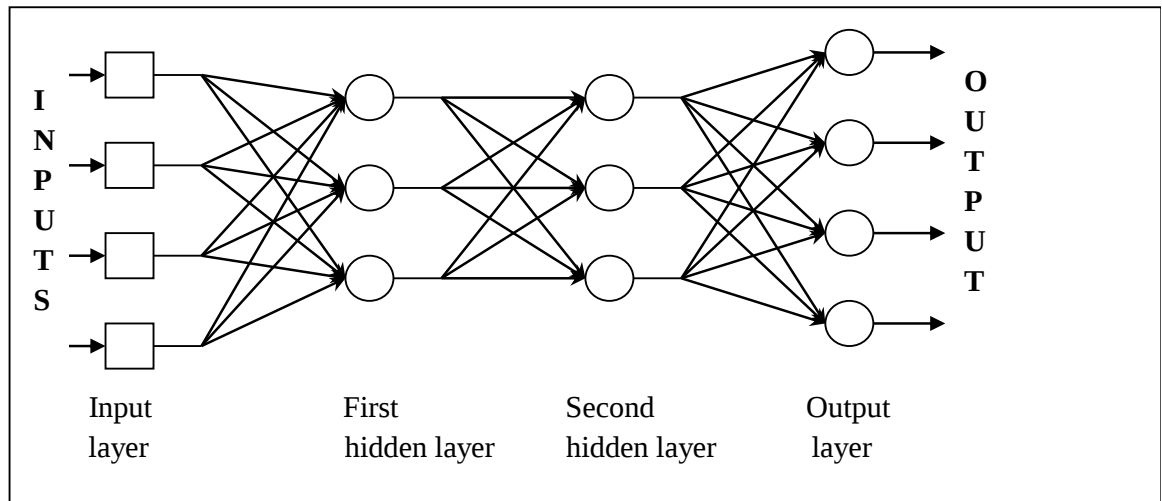


Figure 3.3 Neural network structure with two hidden layers

ANNs are widely used in various research fields, particularly for hydrologic processes such as rainfall-runoff (Hsu *et al.*, 1995; Shamseldin, 1997), streamflow (Zealand *et al.*, 1999; Campolo & Soldati, 1999; Abrahart & See, 2000), and weather (Hall *et al.*, 1999). Further work by Kumar *et al.* (1995) showed that this model can generate accurate forecasts of Indian summer monsoon rainfall. Goswami and Srividya (1996) developed a neural network-based forecasting system to predict monsoon rainfall from previous values of the rainfall series.

French *et al.* (1992) conducted rainfall predictions using ANNs, despite some limitations in the type of ANN used. Abraham *et al.* (2001) applied a fuzzy neural network and conjugate gradient algorithm for rainfall prediction in the south Indian Peninsula, but the predicted rainfall deviated from the actual rainfall due to the delay in monsoon and ENSO, which were used as predictors. Manusthiparom *et al.* (2003) used sea surface temperature and historical rainfall as input data for a standard back-propagation algorithm ANN to forecast rainfall one year in advance. While forecasting West and Central African seasonal rainfall, Philippon (2002) showed that ANN allows adaptation of the predictors to the local climate with the choice of appropriate training and transfer functions. Several works have developed ANNs for seasonal rainfall and streamflow forecasting in Africa and India (Wu *et al.*, 2010; Toth *et al.*, 2000; Coulibaly *et al.*, 2005; Chattopadhyay, 2007; Venkatesan *et al.*, 1997; Cannon & Mckendry, 1999; Cannon, 2000; Sahai *et al.*, 2000). Toth *et al.* (2000) compared three methods for short-term forecasting of precipitation: autoregressive moving average (ARMA), ANN, and K-nearest-neighbor (KNN), and found that ANN had the highest precision for forecasting flows. For more details regarding the application of ANNs in hydrology, refer to ASCE (2000a, 2000b).

In this work, the feed-forward neural network (FFNN) was used. FFNN is capable of representing non-linear processes between input and output variables, and non-linearity and other interactions between variables can be modeled without prior specification (Tangang *et al.*, 1998). Some descriptive variables were set before applying FFNN here, including network type, specification of connection weights, and training and transfer functions.

i. Specification of network size

In this step, the number of layers in the network and the nodes in each layer are defined. There is no specific rule to determine the exact number of hidden nodes. The only available method is trial and error, which means starting with a small number of nodes and gradually increasing the network size until the desired accuracy is achieved.

ii. Searching optimal values for connection weights

Two types of batch propagation algorithm are used for the training function: batch gradient descent with momentum (BGDM), and gradient descent with adaptive learning rate (GDALR), which is the more popular and suitable back-propagation algorithm and is used for 90% of ANN models developed in the field of hydrology (Philippon, 2002). The BGDM can provide faster convergence with momentum, which allows the FFNN to consider new tendencies in the error surface in addition to the local gradient. The momentum helps disregard minor features in the error surface in the same way a low-pass filter does. A network deprived of momentum can bog down in a local minimum, but with momentum it can overcome the minimum.

3.2.6 Regression trees

Classification and regression trees (CART) are non-parametric methods first developed by Breiman *et al.* (1984). CART recursively divides data into smaller strata in order to make the fit as close as possible. The trees are directed graphs (see Figure 3.4) that start with one node and branch out into many. They are fundamental in many fields, and using them for prediction has become popular over the last two decades since they are a viable option compared to other approaches (e.g., regression). CART consists of two methods: classification tree (used for categorical variables) and regression tree (used for continuous variables).

There are distinct differences between a classification tree (CT) and a regression tree (RT). CT searches for a set of rules that classify the data into predefined sets, while RT does not deal with predefined groups. When using RT, the groups are generated automatically during the classification process (Li & Sailor, 2000). Another difference is that CTs attempt to predict a discrete category (e.g., class) rather than a numerical value, while RT is continuous and allows the prediction of continuously dependent variables. In addition, RT does not assume normal distributed observations, or that the responses are related to predictors in a linear manner. Given the complex nature of climate systems, it is always best to avoid simplistic assumptions.

The major components of RT models are the selection and stopping rules. The selection rule determines which stratification to perform at each stage, and the stopping rule determines the final strata that are formed. Once the strata are created, the impurity

of each node is measured. There are three commonly used measures for node impurity in RT: misclassification error, Gini index, and cross-entropy or deviance. While all are similar, the latter two are differentiable and easier to optimize numerically, as it has been determined that the probabilities at nodes can easily affect the cross-entropy and Gini index. As well, in RT the least squares method is used to measure node impurity. For more details about the measurement of node impurity, refer to Hastie *et al.* (2001). Also, applying RT requires pruning, which is the process of reducing a tree by turning some branch nodes into leaf nodes and removing them from the original branch.

An example of RT used to forecast air data, illustrating RT splitting and structure, is presented in Figure 3.5.

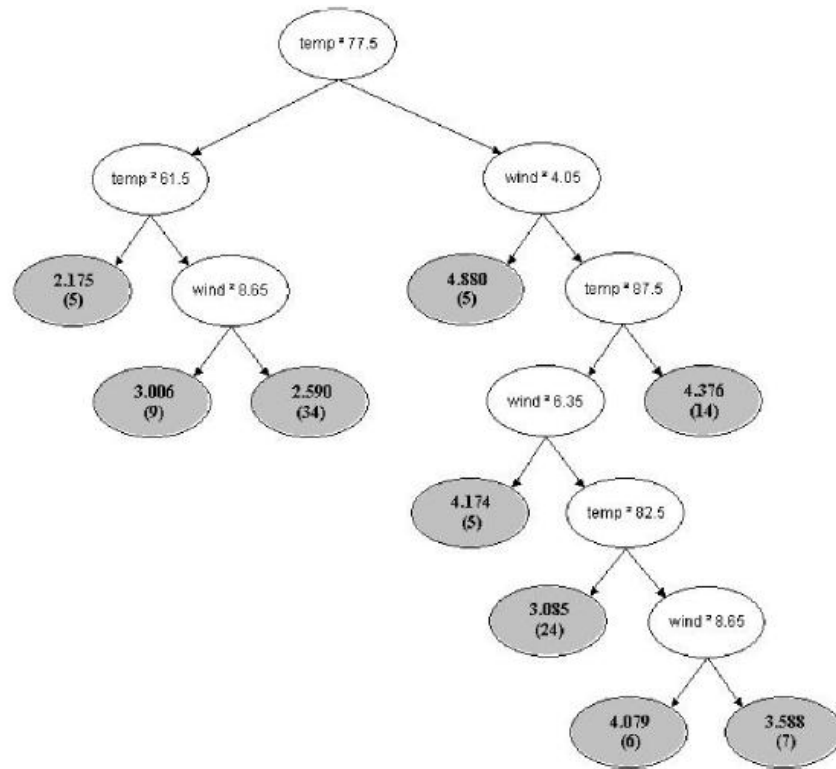


Figure 3.4 Schematic representation of a regression tree (source: Guangzhe & Brian, 2005)

Kannan and Ghosh (2010) found some advantages when using CART as a prediction method:

- there is no assumption of any type of statistical distribution on dependent or independent variables
- the predictor variables can take any format (e.g., continuous, categorical, interval)

- CART is not affected by collinearities, outliers, heteroscedasticity, or distribution errors that affect parametric procedures
- it can detect and reveal interactions in the dataset
- it is invariant under monotonic transformation of independent variables
- it efficiently manages modelling issues with huge dimensional data

However, despite these advantages, CART has limitations due to its novelty, as it is difficult to find help on the method. As explained in the preceding paragraph, the main issue when using RT is to determine the exact time to stop to prevent over-fitting; that is, the point when the fit of the training data is no longer likely to improve the predictive validity of the respective model. A general solution to this issue is to assess the quality of the fitted model by predicting observations using a test sample of data that have not yet been used in order to estimate the respective model. This helps determine the predictive accuracy, and detects when over-fitting begins. Two main techniques are used to deal with the over-fitting issue: bagging and boosting. Bagging is more useful for RT than partition methods, as it resamples observations by random sampling with replacement. Bagging and boosting methods are detailed in Section 3.5.5.

3.2.7 Non-stationary techniques

Data series are considered stationary if they remain invariant despite temporal changes. However, these conditions are nearly impossible, as there will inevitably be either abrupt changes or periodicities. These can lead a modeller to wrong predictions, as they could significantly impact the outcome of the forecast by inducing major uncertainties related to such changes. Therefore, it is very important to check the observed data carefully because change-point-free records are highly unlikely. Thus, the homogeneity of the observational data must be assessed prior to conducting forecast analysis, as recommended by the World Meteorological Organization (WMO). Homogenization involves the detection and correction of any changes in data series, and there are many methods in the literature for various fields of research (Vincent, 1998; Peterson *et al.*, 1998; Wang, 2003; Rodionov, 2004; Begert *et al.*, 2005; Beaulieu *et al.*, 2005, 2009; Fearnhead, 2006; Seidou *et al.*, 2007; Seidou & Ouarda, 2007; Villarini *et*

al., 2011). This review is focused on promising methods of change point detection (e.g., Bayesian inference approach) in order to allow their easy use when selecting seasonal rainfall/streamflow forecasting models.

a) Definition and sources of change points

Change points can be defined as discontinuities of time series that commonly exist in climate data (Reeves *et al.*, 2006). They can be either abrupt shifts in the mean value of a data series, or changes in trend or variance. According to Lund and Reeves (2002), a change point is a point when the series mean first undergoes a structural pattern change (e.g., slopes, trend) that could affect the discontinuity in series values. Thus, change points are points of discontinuity in a series, and can be due to several reasons, including movement of the observed station, changes in recording equipment, changes in measurement techniques, environmental changes, and climate change effects such as shifts in climate regimes (Lund & Reeves, 2002; Milly *et al.*, 2008; Daly *et al.*, 2007; Easterling *et al.*, 1996). The principal causes of inhomogeneity in a climate data series depend on the type of parameter measured. For example, changes in the time of observations or computation methods can cause inhomogeneity in the observed data series of average temperatures and moisture, but do not affect precipitation or pressure. Several cases of change points have been linked to anthropogenic effects such as dam construction, river regulation and land use changes. Inferences from climate data series frequently depend on the continuity of the measurement process; thus it is very difficult to achieve a change-point-free record.

b) Change point models: A review

Several change point models have been developed in the fields of hydrology and climate science, as well as engineering. The models were built using variables with statistical relationships based on known techniques (e.g., multiple linear regression, stepwise regression). These relationships in hydrological or climate models are often affected by sudden changes due to environmental, climatic (e.g., ENSO), and anthropogenic variations. In fact, many types of disturbance can cause apparent changes in a long-term climate time series and distort the true climate signal. These shifts (or changes) could impact the real physical characteristics of the variables in hydrological models (Woo & Thorne, 2003; Salinger, 2005), which is why several methods to detect

and correct these change points are proposed in the literature. Many climate variables are homogenized by change point detection, including temperature (Vincent, 1998; Begert *et al.*, 2005) and precipitation (Beaulieu *et al.*, 2005, 2009). Lund *et al.* (2001, as cited in Lund & Reeves, 2002) conducted research related to temperature change in the USA, and concluded that a single change point can affect the accuracy of linear temperature change estimation. Villarini *et al.* (2011) investigated the abrupt change of average daily discharge for 75 years over Central Europe, and detected some statistically significant change points in mean value and variance. While testing the non-stationarity of flood peak records in the USA's Delaware River basin, Smith *et al.* (2010) found that the presence of change points was associated with anthropogenic measures. Villarini *et al.* (2009) also investigated the validity of stationarity assumptions of annual maximum peak discharge data in the USA during the 20th century, and their results showed that almost half of the analyzed series exhibited statistically significant change points.

The methods in the literature for change point analysis can be grouped into two types: classical statistics methods, and Bayesian approach. The classical methods are usually only used for statistical tests to confirm or reject hypotheses based on the existence of a change. Likewise, the Pettitt test has been used in many previous studies (Tomozeiu *et al.*, 2005; Villarini *et al.*, 2011; Smith *et al.*, 2010; Villarini & Smith, 2010) to detect abrupt changes in the mean and variances in hydrology, due to its limits in cases of multiple change points requiring time series segmentation. Most authors (Solow, 1987; Easterling & Peterson, 1992; Vincent, 1998; Lund & Reeves, 2002; Wang, 2003) have focused on changes in the slope or intercept in linear regression models. This set of methods includes non-parametric methods (Servat *et al.*, 1997) and fuzzy methods (Yu *et al.*, 2001).

The second type is the Bayesian approach for change point analysis (Sarr *et al.*, 2013; Seidou *et al.*, 2007; Gelfand *et al.*, 1990; Seidou & Ouara, 2007; Barry & Hartigan, 1992, 1993; Perreault *et al.*, 2000; Xiong & Guo, 2004). This is one of the most popular techniques, as it helps obtain the statistical distribution for the dates of change as well as distributions for other parameters in the model. This method is associated with easy definition of priors, and determining the probability of any change point using its posterior probability distribution. The main advantage of the Bayesian approach

compared to classical methods is its effective description of parameter uncertainty. The Bayesian approach provides a comprehensive posterior probability distribution of change point position, whereas classical methods can only give the most probable change point position. However, this technique requires significant computational resources due to the involvement of the Monte-Carlo Markov Chain (MCMC) simulation. MCMC is often used in cases of inference when the change points are known, and if the change points are unknown, a reversible-jump MCMC can be applied to examine the joint space of the model and its parameters.

Fearnhead (2005, 2006) presented a recursion-based inference approach to deal with multiple change points problems, based on the theory of product-partition models from Barry and Hartigan (1992, 1993). In his recursion-based inference, Fearnhead (2006) used a set of recursive relationships to determine the posterior probabilities of different change points. This approach is specific, as it identifies only the number and position of changes occurring in the dataset. Gelfand *et al.* (1990) developed a Bayesian analysis model based on regression analysis of variances that are considered normal data where unequal variances are acceptable. Some authors (e.g., Perreault *et al.*, 2000) have dealt the case of a single change point using Gibbs sampling in a Bayesian model, without knowledge about the priors but with good knowledge of the occurrence of changes (i.e., their presence). Seidou *et al.* (2007) developed a Bayesian model that seems to be better than the others, confirmed by the fact their model can deal with either informative or non-informative priors, meaning that it can handle multivariate data and missing values. Furthermore, Seidou and Ouarda (2007) developed another Bayesian technique for multiple change point detection by proposing new classes of priors for the parameters of a multiple linear model, with straightforward computation of the posterior distribution of the change points.

With respect to the period of occurrence, there are two approaches to determine the presence of change points: retrospective (off-line) and sequential (on-line). The retrospective change point detection approach is used in most hydrological studies, though classical methods (change in slopes or intercepts) are also used, as explained previously. Some studies (Sagarin & Micheli, 2001; Bowman *et al.*, 2004) were based on curve-fitting methods to detect the changes. The sequential change point detection

approach can detect earlier changes (i.e., as soon as they occur) (Daumer & Falk, 1998; Gut & Steinebach, 2002; Moreno *et al.*, 2005). As explained, most current change point detection techniques rely on the Bayesian approach. Thus, this method is an important element in change point detection models, largely due to its ability to make inferences on the posterior distribution with respect to change point location.

c) Recursion-based change point inference approaches

This model is used to perform a finite number of recursive operations when dealing with a single change point in a normal distributed sample. This Bayesian inference approach was initially developed by Yao in 1984.

Other studies followed, such as that of Barry and Hartigan (1992, 1993), who were able to improve on Yao's method by using product-partition to handle a change point problem before generalizing it to multiple change points with prior assumptions. It should be noted that with product-partition models, all observations issued from a random partition of the data are considered to be independent. Progress was made toward improving this technique. Fearnhead (2006) regenerated the approach into a general inference procedure to detect the number and positions of change points, and Seidou and Ouarda (2007) greatly improved the approach by developing a fast recursion-based change point inference. They proposed a new class of priors for the parameters of their multiple linear regression model, and defined methods that allow direct computation of the posterior distribution of the change point. This latter model is not time-consuming because, unlike many Bayesian inference models, it does not involve MCMC simulation.

One drawback of the recursion-based inference change point model is the possibility of missing a change point during the detection process. This could be due to the relative position and sign, magnitude, or nature of the change point. Full details of the effects of these factors can be found in Beaulieu (2005).

3.3. Hydrological Modelling

A model can be described as a simplified representation of a physical system and the different processes of its operation. Models are created to simulate all or some aspects of a system's behavior. Hydrological models are used for the following reasons:

- Better assessment of the hydrological regime of rivers
- Optimization of networks of hydrologic observatories
- Plausible forecasting for more objective decision-making

Hydrological models are currently recognized as valid scientific tools for quantitative and qualitative investigation of the water resources of an area. Thus, they provide the basis for planning integrated management of these resources.

3.3.1 Classification of Hydrological Models

Hydrological models can be classified according to independent criteria. The following specific criteria are used to group models (Singh, 2005):

- the approach used in the layout: stochastic or deterministic
- degree of complexity of the layout: hydraulic, conceptual, or empirical
- characteristics of the input data or parameters: concentrated or distributed
- linearity of the differential equations: linear or non-linear
- time-dependence: stationary or non-stationary

Physical models are representations of a real system by another system with similar properties, but on a reduced scale which is much simpler to manage.

Hydraulic models (physically based models) describe all the phenomena in a watershed using physical equations.

Empirical models (black box models) do not describe the physical processes; they just report the link-runoff in order to provide a result that is as close as possible to the observed data. These models are also called synthetic models (Hingray *et al.*, 2009), due to their ability to consider the transformation as a global phenomenon.

Conceptual models (grey box models) apply the basic laws of physics. They are constructed in a framework that refers to processes similar to natural ones, such as the outflow from a tank, or motion in a regular channel.

Mathematical models are characterized by their ability to represent simple or complex systems. Though they can be linear or non-linear, the term 'linearity' has at least two meanings. A model is defined in terms of linear systems theory if it obeys the principle of superposition. A model is also linear in the sense of statistical regression; that is, if there is linearity in the parameters that must be estimated.

A model is defined as stationary or non-stationary if the relationship between the input and output, respectively, does not change or changes over time.

Concentrated or "lumped" models treat the entire water catchment as homogeneous, while semi-distributed models divide it into areas of sub-homogeneous zones, according to the contribution of the flow. In totally distributed models, the entire basin is divided into elementary units. Thus, distributed hydrological models consider the spatial variability of the characteristics of the basin and the main hydrological phenomena (e.g., infiltration, runoff, interception, erosion) that allow adequate simulation of the dynamics of the basin. Distributed models can study a rainfall-runoff phenomenon in detail, taking the spatial variability of the input data and the geomorphological and pedological characteristics of the basin into consideration.

Depending on the nature of the results, the models can be classified as either stochastic or deterministic. A stochastic model consists of a finite set of random variables with a given probability distribution that depend on time t and values they had in the past. A deterministic model considers all the variables, without random variation. These models attempt to numerically predict the evolution of a system (in space and time) through approximate solutions of mathematical equations that describe the physical laws.

3.3.2 Criteria for adequate model selection

A hydrological model is chosen according to the best use of the available information, and to some systematic procedures for selection and verification. The choice of the model that best fits the case study depends on many factors, as described in Hingray *et al.* (2009). These include:

- a. Simplicity: User-friendly models encourage user involvement. They are characterized by flexibility when entering and managing input data, and simple graphics options that facilitate and increase the learning rate, and the achievement and use of expected results.
- b. Quantity and quality of available data
- c. Accuracy of forecasting: use of the calibration process to define the degree of accuracy of a model's simulation

- d. Scale representation of the data and results: the ability of the model to represent an area in a comprehensible and accurate manner
- e. Ability to handle missing or incomplete data

Thus, as shown in Figure 3.5, combining the performance criteria, availability of data, and model complexity is recommended for effective model selection.

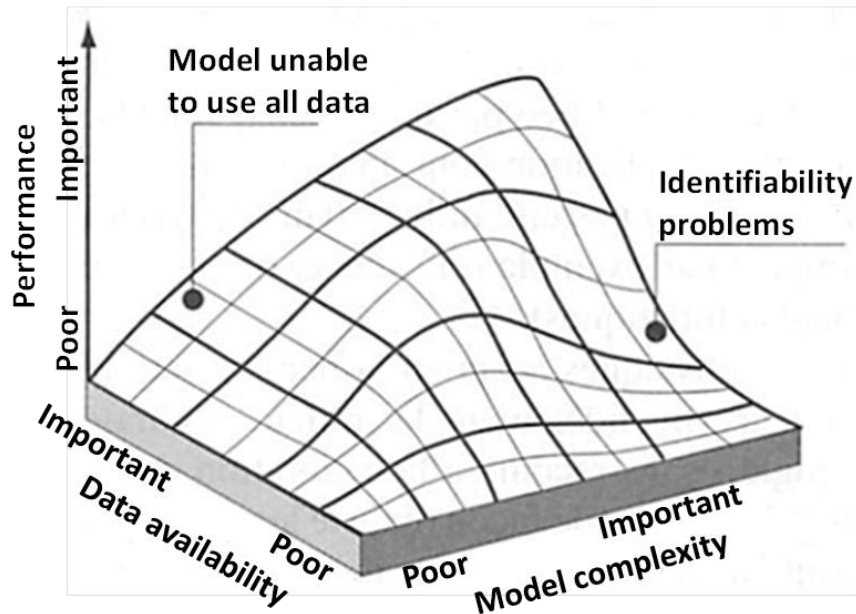


Figure 3.5 Schematic diagram of the relationship between complexity and performance of a hydrological model and data availability (adapted from Grayson & Blöschl, 2000)

For this study, a physical, semi-empirical conceptual distributed model known as SWAT (Soil and Water Assessment Tool) was chosen to manipulate and analyze hydrological and other climate data. Coupled with the GIS (Geographical Information System) to manage data raster, vector and alphanumeric, SWAT facilitates and automates the preparation of input data, making it more user-friendly in the integration phase, and managing the settings related to the simulation data.

3.3.3 Calibration and validation of the SWAT model

SWAT models, as with any conceptual model, commonly possess many parameters which are not directly measured from a site. Hence, there is a need to search for parameter values iteratively, which can significantly adjust the parameters, in order to

obtain the best match between the simulated values and corresponding observations. Such an approach is known as the calibration process. The main criteria used in SWAT to assess the best fit between observed and simulated parameters are called objective functions (Duan *et al.*, 1994). According to Yang *et al.* (2008), the last values of the parameters in the calibration step are used in the validation process.

The first step in carrying out prior calibration is the sensitivity analysis. Sensitivity analysis involves determining the response of a model to changes in parameter values, and identifies the parameters that have greatest effect on the model. It also helps in saving time during calibration, since only a few parameters of the SWAT model are optimized. With SWAT, sensitivity analysis is done based on changes of the mean of the output variable (flow). If observed data for the output variable are available, additional sensitivity analysis is done using the sum of the squared error between the observed and simulated values. It should be noted that there are two approaches in sensitivity analysis under SWAT: local and global methods. The local method changes parameter values only for one parameter at a time, while the global method adjusts all parameters at the same time. The latter method is used when there are a large number of simulations, while the former is suitable in the case that there are correct values for the fixed parameters (Arnold *et al.*, 2012). The set of significant parameters produced after completing the sensitivity analysis is further used in the calibration process. This calibration can be done either automatically, by using the auto calibration tools in SWAT, or through SWAT-CUP (SWAT Calibration and Uncertainty Programs).

The SWAT-CUP is the method mostly used in SWAT model calibration (Abbaspour, 2012), and has five different approaches to help carry out the model parameterization: sequential uncertainty fitting (SUFI2), particle swarm optimization (PSO), generalized likelihood uncertainty estimation (GLUE), parameter solution (ParaSol), and Markov Chain Monte Carlo (MCMC). Many studies have shown that SUFI2 presents better efficiency than others, with less computational cost (Ma *et al.*, 2014; Nkonge *et al.*, 2014; Luo *et al.*, 2011). SUFI2 uses Latin Hypercube Sampling (LHS) as a sampling method to define the values of particular parameters for the calibration period (Abbaspour, 2012). More details on the advantages and limits of each of these methods are provided in Yang *et al.* (2008).

According to Moriasi *et al.* (2007), there are no precise criteria to assess the model performance or accuracy. However, the most-used measures are statistical criteria: coefficient of determination (R^2), and the Nash-Sutcliffe coefficient (NASH) (Arnold *et al.*, 2012).

3.3.4 Model performance analysis

Since the uncertainty in hydrology is the result of the natural variability and complexity of systems and hydrological processes, it follows that the application of hydrological modeling is also subject to uncertainty. Therefore, performance analysis can be considered a solution to assessing the extent to which a hydrological model can reproduce the real characteristic of a watershed using observed data.

The performance assessment of a hydrological model consists of comparing the observed and simulated data to find the closeness between them. According to Zhang *et al.* (2009), the performance can be assessed in two ways:

- Qualitatively, by visual comparison, using the same plot for the observed and simulated values for a given variable. Several types of graphs can be considered (e.g., scatter plots).
- Quantitatively, using standard numerical performance criteria. The performance criterion is often a mathematical function of the errors obtained.

These mathematical functions are fully explained in Section 3.5.

3.4. Temporal Disaggregation

Many hydrological models often require precipitation data on a finer time scale (sub-daily, daily or monthly), but most hydrological data are on a coarser scale. These data enable simulating water resource systems to produce flows for the estimation of water yield in reservoirs (e.g., dams). The quality of precipitation data is also poor in several parts of the world, particularly Africa, due to an inadequate number of rain-gauge networks. Therefore, to get the data at the desired scale, one has to either set a weather generator to produce the data, or disaggregate the coarse data into a finer scale.

Mondal and Wasimi (2005) defined temporal disaggregation (TD) as a mathematical technique for downscaling coarser temporal or spatial levels into finer ones.

Thus, TD can be considered as a statistical simulation technique that allows movement from one scale to another (i.e., coarser to finer resolution). Nevertheless, TD should not be considered as exact downscaling, despite the fact that in both techniques, synthetic series are used to reproduce the important statistical characteristics in hydrological processes at a given time scale. According to Koutsoyiannis (2003), TD is not identical to downscaling because the latter produces finer scale time series with the required statistics that do not necessarily add up to any given coarse scale totals. With TD, there are a number of statistical characteristics from the original data (i.e., the data to be disaggregated) that should be preserved after disaggregation. Downscaling is particularly useful for hydrological applications when dealing with the outputs of GCM or RCM.

Most disaggregation models in the field of hydrology are general-purpose, and are not specific to certain hydrological processes or other applications. The general-purpose models cannot always be used for rainfall due to specific peculiarities, and this has raised interest in developing special disaggregation techniques for precipitation.

3.4.1 Method of fragment

The method of fragment (MF) was initially proposed by Svanidze (1977), and consists of generating annual datasets using a specific model (e.g., weather generators, transfer functions, and weather typing). The computation of monthly fragments is based on historical data, and disaggregating each value of the annual data requires a pool of fragments. The portions of annual precipitation that occur in each month of the year are known as fragments, and the sum of the fragments equals one. The steps of the MF approach are summarized in Equations (3.5) and (3.6).

$$W_i = \frac{d_i}{\sum_{i=1}^n d_i} \quad (3.5)$$

where W_i is the fragment to be calculated for month i , d_i is the data from the series chosen to produce the fragments (e.g., the chosen monthly data), and n is the number of data in that series.

Each fragment is then multiplied by the yearly data y being disaggregated, without altering the total precipitation in the year. This produces the new monthly values d_i' , as indicated in Equation (3.6).

$$d'_i = W_i \times y \quad (3.6)$$

This method was used in this work to increase the time scale of variables that use standardized historical data to generate the fragment series for each year. The biggest challenge with this method is selecting a set of fragments to multiply the forecasted rainfall total for increasing the temporal scale. Hence, each fragment W_i calculated for day d_i is multiplied by the seasonal data being disaggregated, such that the total precipitation in the season is not altered. After getting the historical average seasonal rainfall (July to September) for the 11 stations over the Sirba watershed, two main steps were applied on each seasonal simulated rainfall ($Y_{Se,Sim}$).

- a. The year of historical seasonal rainfall $Y_{Se,hist}$ with the average historical seasonal rainfall closest to the simulated seasonal rainfall was chosen.
- b. The daily rainfall of Y_{Sim} at station T_i was estimated using Equation (3.7). This method generated the rainfall time series for the 11 stations that keep the rainfall distribution over the whole basin.

$$Y_{daily,Sim,T_i} = Y_{daily,hist,T_i} \times \left[\frac{Y_{Se,Sim}}{Y_{Se,hist}} \right] \quad (3.7)$$

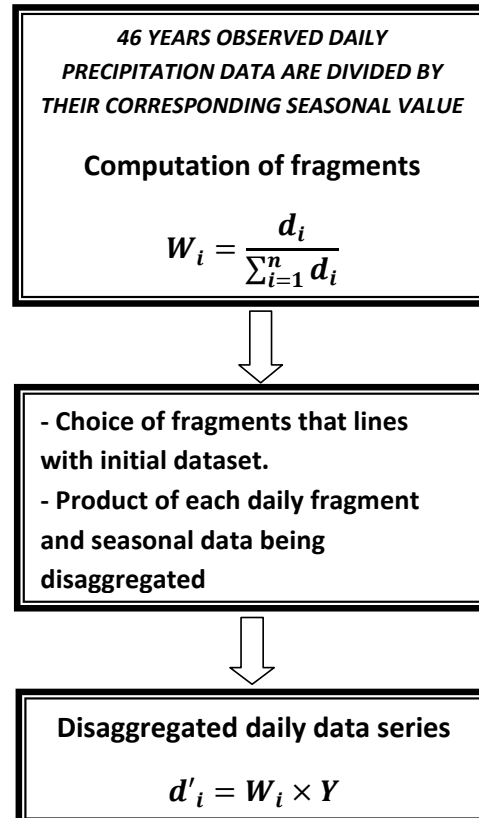


Figure 3.6 Flowchart describing the fragment disaggregation approach (*MF*)

3.4.2 Modified method of fragment

To address the limitations of MF and method of synthetic fragment, Maheepala and Perera (1996) proposed modifying Porter and Pink's (1991) method by changing how the monthly fragments for disaggregation of annual data set are chosen. This gave rise to a new approach called the modified method of fragment (MMF). MMF is similar to the first method except that generated data are used to compute the fragments rather than historical data (see Figure 3.8). A suitable generation model is applied to generate daily series and a seasonal dataset. The data generation model is a first-order Markov chain model (see Equation (3.8)), which is usually appropriate for such cases (Maheepala & Perera, 1996). Daily fragments are computed for each season, using the daily series.

$$X_t = \rho X_{t-1} + \varepsilon_t(1 - \rho^2)^{1/2} \quad (3.8)$$

where:

X_t = Standardized variable at time t

ε_t = Normally distributed random number with zero mean and a unit standard deviation at time t

ρ = Lag-one serial correlation of the variable

For the last step, each value of the generated seasonal precipitation is disaggregated using daily fragments, and a smoothing factor ($\alpha_j + \beta_j$) for selecting the series to be used as fragments is introduced. This factor, with minimum value, should be considered as described in Equations (3.9) and (3.10) and detailed by Maheepala and Perera (1996). The steps are illustrated in Figure 3.8.

$$\alpha_j = \left(\frac{S_k - \sum_{j=1}^{92} D_j}{\sigma_S} \right)^2 \quad (3.9)$$

$$\beta_j = \left(\frac{d_{k-1} - d_{j-1}}{\sigma_d} \right)^2 \quad (3.10)$$

where:

S = generated seasonal data

D = historical daily data

d = last day of daily data

j = season (month) of daily data

k = season (month) of seasonal data

σ = standard deviation

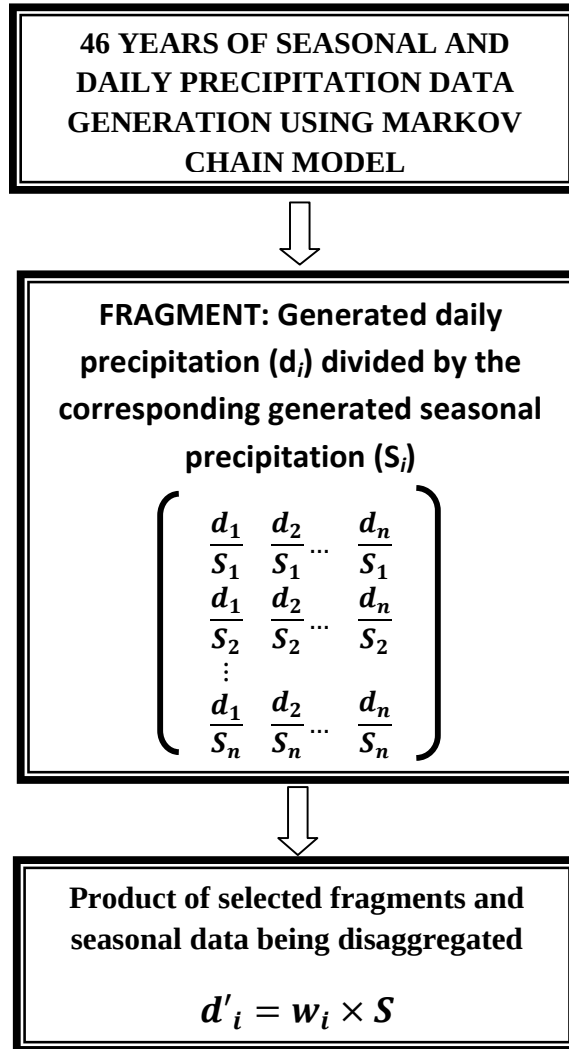


Figure 3.7 Summary flow chart of the modified fragment disaggregation approach (MMF)

3.5. Forecast Assessment

One of the essential steps in developing a forecasting system is assessing the quality of the models. Weather and climate forecasting is internationally monitored, and the WMO has published a comprehensive survey of forecast verification methods since the late 1980s (Wilks, 2006; Jolliffe & Stephenson, 2012). The model assessment field is separate from the development of prediction systems, since the increasing complexity of forecasts has made it an active area of scientific research (Casati *et al.*, 2008). It constitutes the main step in judging a forecasting model's performance.

Forecast assessment includes examining and reviewing the relationships between observed and predicted datasets, as well as comparing the performance of forecasting systems to that of the reference states. Jolliffe and Stephenson (2012) defined forecast verification as the exploration and assessment of the quality of a forecasting system, based on samples of previous forecasts and corresponding observations.

However, there is a clear difference between verification and validation of a forecasting model. Verification determines if a model gives an accurate representation of reality, whereas validation only establishes that the model provides a reasonable representation of reality, and is adapted to the target.

The measures used in verification often assess the accuracy or association of simulated and actual observations. Accuracy is the extent of the link between individual pairs of forecasts and observations, and association is the overall strength of the relationship between these individual pairs. Thus, the correlation coefficient is a measure of the linear relationship, while the coefficient of determination or mean squared error is a measure of accuracy. The two types of measurement are often distinct, as follows.

- Measurement of model accuracy: as there are many methods to measure accuracy, Murphy (1988) recommended using several of them simultaneously to assess well-defined abilities of the model.
- Measurement of model performance: this compares a model's accuracy to that of a standard model. A forecasting model cannot be considered useful if its performance is lower than the performance of standard models (Livezy, 1995).

The following sections detail some tests commonly used in assessing forecasting models, including those selected for this work.

3.5.1 Coefficient of Determination

The coefficient of determination (R^2) is the square value of the correlation coefficient (CC), according to Bravais-Pearson. The R^2 can also be expressed as the square ratio between the covariance and the multiplied standard deviations of the observed and simulated values. It estimates a single dispersion against the combined dispersion of the simulated and observed data. The R^2 can be considered as being the proportion of the variation of the predictand that is described or accounted for by the regression. This concept is sometimes expressed as the "proportion of variation explained". However, this assertion is distorted, and the confusion is explained by the fact that a regression analysis can only quantify the nature and strength of a relationship between two variables; it cannot determine which variable causes the other (Wilks, 2011).

One of the major drawbacks of R^2 , if it is considered alone, is that only the dispersion is quantified. Hence, a model that always over- or under-predicts can still result in good R^2 values close to 1, even if all the predictions are wrong (Krausse *et al.*, 2005).

As shown in Equation (3.11), CC is dimensionless and is within the limits of -1 and 1.

$$CC = \frac{\sum_{i=1}^n (OB_i - \overline{OB})(PR_i - \overline{PR})}{\sqrt{\sum_{i=1}^n (OB_i - \overline{OB})^2} \times \sqrt{\sum_{i=1}^n (PR_i - \overline{PR})^2}} \quad (3.11)$$

where OB and PR are the observed and forecasted values respectively, and n is the amount of data.

A correlation of +1 or -1 denotes a perfect linear association between the observed and simulated data, while a correlation of zero means that there is no linear relationship between the variables. It should be noted that if there are two uncorrelated variables they are not necessarily independent, as there may be a non-linear relationship between them.

3.5.2 Nash-Sutcliffe coefficient

The Nash-Sutcliffe statistic (NASH) can be defined as a measure of how the observed variance is simulated. The efficiency proposed by Nash and Sutcliffe (1970) is computed as one minus the sum of the absolute squared differences between the predicted and observed values, normalized by the variance of the observed values during the period under investigation (see Equation (3.12)).

$$\text{NASH} = 1 - \frac{\sum_{i=1}^n (\text{OB}_i - \text{PR}_i)^2}{\sum_{i=1}^n (\text{OB}_i - \overline{\text{OB}})^2} \quad (3.12)$$

The range of NASH is between 1 and $-\infty$. A value of 1 indicates a perfect match of modeled to observed data, and $\text{NASH} = 0$ means that the model's predictions are as accurate as the mean of the observed data. An efficiency $\text{NASH} < 0$ indicates that the mean value of the observed time series would have been a better predictor than the model.

Legates and McCabe (1999) pointed out some limits of using NASH in forecasting models. The greatest drawback is the fact that the differences between the observed and predicted values are calculated as squared values. As a result, larger values in a time series are strongly overestimated and lower values are neglected. In addition, NASH is not very sensitive to systematic models' over- or under-prediction.

3.5.3 Hit rate score

This criterion is different from the previous, as it is suitable for categorical data analysis. However, the model performance is assessed by converting the continuous data to categorical data, and then uses the hit score (HIT) as an estimator.

This is done by considering each value of a continuous variable particularly falling into a categorical set of data. The contingency table is 2 x 2, explaining four combinations of events for each pair of forecasted and observed data. If the simulated and measured data are considered in three categories (for example, normal, below-normal, and above-normal) this contingency table will consist of nine combinations of events (Wilks, 2011). Equations (3.13) and (3.14) are used to calculate HIT for 2 x 2 and 3 x 3 contingency tables respectively. Using this type of measurement, the properties of both

prediction and observation data (for example, mean and variance) are not responsible for assessment. It measures only how many data fall into the same section. If HIT shows a high percentage, it means that predictions correspond well to the observations, whereas a low value indicates a lower model performance. HIT is within the interval 0 to 100%.

$$HIT = \left(\frac{a+d}{n} \right) * 100 \quad (3.13)$$

$$HIT = \left(\frac{r+v+z}{n} \right) * 100 \quad (3.14)$$

3.5.4 Bagging and boosting

The bagging method was proposed by Breiman (1996) in order to help improve the prediction process by reducing the variance (not the bias) related to a forecast. This technique is particularly useful for fixing the lack of robustness of unstable classifiers such as RTs and ANNs. The method is called either bootstrap or bagging, and it is based on a relatively simple process. For bagging, a number of bootstrap samples are drawn from the available data, a prediction method is applied to each sample, and then the results are combined by averaging to obtain the overall forecast; the variance is reduced due to the averaging process.

In contrast to bagging, boosting uses the weighted average of the results obtained from applying a forecasting method to various samples. However, the samples are not drawn from the population in the same way at each step. Boosting is often applied to weak learners such as RTs with two nodes.

Though bagging and boosting are both ensemble algorithms, they differ in that bagging is a purely random process and boosting is an adaptive process. A comparative summary of these two methods is presented in Table 3.1.

Table 3.1 Comparative summary of bagging and boosting techniques

BAGGING	BOOSTING
<i>CHARACTERISTICS</i>	
Bagging is a random mechanism	Boosting is an adaptive mechanism and is generally deterministic
Learning takes place on a different	Generally, learning takes place on the

bootstrap sample with each iteration	whole initial sample with each iteration
At each iteration, the resulting model must perform well over all the observations	At each iteration, the resulting model must perform well on certain observations. A model performing well on certain outliers will perform less well on other observations
In the final aggregation, all the models have the same weight	In the final aggregation, the models are generally weighted according to their error rate
ADVANTAGES AND DISADVANTAGES	
A method for reducing variance by averaging models	Can reduce the variance and bias of the base classifier, but the variance can increase with a stable base classifier.
Loss of readability if the base classifier is a decision tree	Loss of readability if the base classifier is a decision tree
Ineffective on stumps unless double randomization is provided, as in random forest	Effective on stumps
Faster convergence	Slower convergence
The algorithm can be parallelized	It is a sequential algorithm and cannot be parallelized
No over-fitting, thus better than boosting in the presence of noise	Risk of over-fitting, but better than bagging overall on non-noisy data
Bagging is effective more often than boosting	When boosting is effective, it is better than bagging

3.5.5 Leave-one-out cross-validation

While searching for model errors using discriminant analysis, Lachenbruch and Mickey (1968) developed the cross-validation method (CRV), a statistical technique that provides an estimate of forecasting ability that is less biased than the usual hindcast ability estimates. It works by steadily deleting one or more datums in a dataset, and developing the forecasting model from the remaining data before testing it on the deleted cases. The process is a nonparametric approach, and can be applied in building any model (Michaelsen, 1987). Also, in a forecast using several predictors, CRV can be used to screen a set of potential predictors to find the best subset; this is why it was used in many forecasting applications (Breiman *et al.*, 1984). Several studies (Barsnton & Van Den Dool, 1992; Von Storch & Ziwers, 1999) showed that precautions should be taken when applying this method in order to avoid distortion of the results. To be precautionary, the

information used to develop the model should be quite independent of the information applied during the phase test. ACMAD and CLIPS (1998) used CRV to evaluate the performance of all the national models of seasonal rainfall forecasting used annually. Figure 3.9 summarizes the steps involved in carrying out a CRV.

- Observations are first divided into k samples of the same size in order to conduct k adjustments (training) by removing a sample each time to estimate the performance of the model with the iteration k (Bishop, 1997).
- Then, the parameters and performances of the model are obtained by computing their respective averages through the use of all iterations.

The natural limit of CRV corresponds to when the number of available observations is equal to the number of k samples.

Though larger sample sizes can be managed by omitting a subset each time, with small numbers of observations (as in this work) it is advisable to omit each observation one at a time. This type of CRV is known as leave-one-out cross-validation (LOOCV), and is another type of CRV used with small datasets which ignores each observation individually until the end of the data span. This method of partitioning is one of the most-applied when the number of observations is restricted (Michaelsen, 1987; Elsner & Schmertmann, 1994).

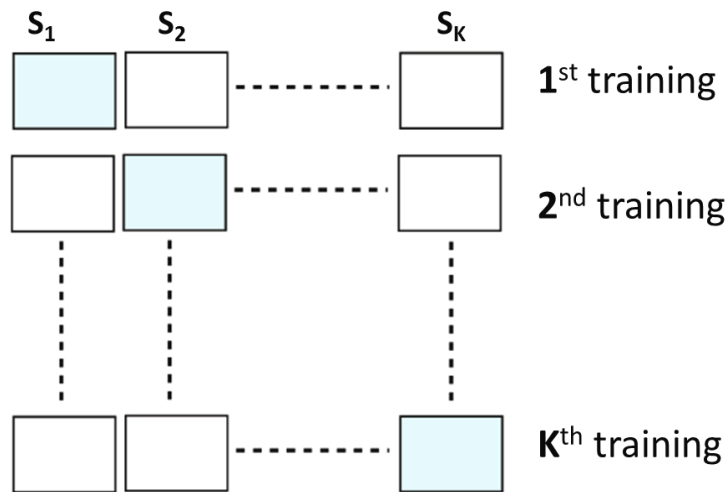


Figure 3.8 The cross-validation procedure (blue boxes are for validation and white boxes are for training)

3.6. Losses Associated with Seasonal Flood Forecasting

The growth of populations, rapid and uncontrolled urbanization of risk areas (informal settlements), instability of health protection, significant prevalence of disease, endemic poverty, poor preparedness and disaster response, inadequacy of climate information, and weak institutional capacity all increase the vulnerability of Africa to climate disasters (Gaye, 2009). If distinctions exist, West African regions show similar characteristics for their urbanization and socio-economic and political development (Ouédraogo, 2008).

It is apparent that West Africa has been experiencing more intense flooding over recent years. Even with flood management strategies in place, these events persist. It is now essential to assess and upgrade the capacities of stakeholders to manage and adapt to such circumstances. Flooding is one of the most serious natural disasters, causing huge loss of life and property every year in this area. Moreover, the International Federation of the Red Cross and Red Crescent Societies pointed out that floods were by far the greatest cause of homelessness. According to WMO (2011) statistics on water-related natural disasters during the last two decades, the issue of floods is more frequent and has more economic impact on populations compared to other extreme events.

Therefore, due to these flood-induced problems, there is a great need to forecast floods in order to provide advance information about impending floods to stakeholders in order to reduce the probable losses. Seasonal flood forecasting is very important for the city of Niamey due to its position on the Niger River. Seasonal flooding experienced by Niamey causes displacement of the local population living near the river, economic losses, and psychological impact on the victims. It should be noted that Niamey is a city that lacks resources, so the funds disbursed to aid the victims should be invested in development in order to provide better living conditions for the population. Hence, seasonal flood forecasts could help Niamey to avoid or lessen the overall impacts of a flood and also save lives, livestock and property.

The development of a seasonal flood forecast model for Niamey involves the development of a hydrodynamical model of the Niger River at Niamey using the HEC-RAS model, and applying probability distributions to generate probabilistic forecasts

(Gado Djibo *et al.*, 2015b) for each year. Moreover, the probabilities of flood protection dykes being overtopped were estimated.

3.6.1 Copula method

Let X and Y be random variables with distribution function F_X and F_Y . A copula (2-dimensional) is known as any distribution function C whose marginal bivariate is a uniform distribution on $[0, 1]$. In other words, C satisfies the four properties of a bivariate distribution function in addition to Equations (3.15) to (3.18):

$$C(x, y) = 0 \quad \text{if } x \leq 0 \text{ or } y \leq 0 \quad (3.15)$$

$$C(x, y) = 1 \quad \text{if } x \geq 1 \text{ and } y \geq 1 \quad (3.16)$$

$$C(x, y) = x \quad \text{if } y \geq 1 \quad (3.17)$$

$$C(x, y) = y \quad \text{if } x \geq 1 \quad (3.18)$$

Copula theory allows modelling the dependence structure between variables independently of marginals (Sklar, 1959). It has been widely applied in various fields such as biology (Frees & Valdez, 1998) and Insurance and Reinsurance (Klugman & Parsa, 1999; Belguise, 2001; Venter, 2002, 2003; Charpentier, 2003). In hydrology, Favre *et al.* (2004) presented two applications: the first concerns the modelling of flow rates upstream and downstream of a river with a tributary, and the second concerns the joint modelling of annual maximum flows based on the volume of runoff. DeMichele *et al.* (2005) used the Gumbel copula to model the positive dependence between the peak flow and flood volume. Zhang and Singh (2006) used bivariate copulas for frequency analysis of the variables: peak flow and volume of runoff and flood volume and its duration.

3.6.2 HEC-RAS model

The hydraulic model HEC-RAS (Hydrologic Engineering Center's River Analysis System) and the HEC-GeoRAS model associated with ArcGIS were used to study and evaluate flood flows and flood zone mapping in a range of the Niger River at Niamey (160 km either side of Niamey city). The HEC-RAS program has been under development by the US Army Corps of Engineers since its inception in 1964, and different sub-programs are continuously being established. It was designed specifically for application in floodplain management and flood-insurance studies to evaluate

floodway encroachment and to simulate estimated flood inundation (U.S. Army Corps of Engineers, 2009; Soleymani *et al.*, 2014).

This model is used worldwide. Johnson and Dominique (1988) used the HEC-RAS model to predict and determine the desired land boundaries, 10 km along the Wyoming River - Gary Bull in the USA. Using the model, the water surface profile of the river was plotted. Using HEC-RAS software, David and Smith (2000) evaluated the hydraulic behavior of the flood. David *et al.* (2002) studied floods for a period of 5 years in the United States and prepared flood zone maps. Knebl *et al.* (2005), using the hydrological model HEC-HMS, hydraulic model HEC-RAS, and radar precipitation estimates (NEXRAD) in the basin of San Antonio, Central Texas, United States, suggested a logic model for flooding and compared the model with the flood of the summer of 2002. The results showed the model's efficiency in regional-scale flood forecasting. Napradean and Chira (2006) did flood-hazard mapping for small watersheds near Baya Sea in Astore valley, and for this purpose, a combination of Wet Spa and HEC-RAS was used.

3.7. Synthesis and Partial Conclusion

Seasonal forecasting is found to be of paramount importance for the West African Sahel. It appeared that for seasonal rainfall forecasting, the statistical and dynamical approaches were the most used in the Sahel region compared to the MOS. GCMs were especially widely used for seasonal rainfall forecasting in this region, and they resulted in acceptable to good forecasting abilities which were assessed frequently using correlation. But due to their low resolution and high computational cost, as well as time, GCMs are limited for seasonal rainfall forecasting in the Sahel. This fact has caused the statistical approach to be widely used to forecast seasonal Sahelian rainfall. The forecast abilities were moderate to high, and the lead time was between 0 and 1 month. However, some lead times of up to 12 months were found. In contrast to seasonal rainfall forecasts, there was no study on seasonal streamflow forecasts in the Sahel. Most of the forecasts were issued from simple hydrological model simulations, while in this work the seasonal streamflow was forecast a few months ahead in order to enable stakeholders and users to plan their activities.

Regarding the seasonal flood forecast, there is a lack of forecasting studies that can be obviously found in the Sahel region. Most of the models are hydrological models, which were limited compared to hydro-dynamical models. Moreover, there was no modelling study that tried to assess flood risk in the Sahel region with the related probabilities.

Hydrological modelling was also reviewed before starting to develop the model, including the different types of models and the criteria needed to make a good selection. It was particularly mention the limits of each model. In this study, the SWAT model was selected and used to forecast streamflow. It is a physical-based model which is used for hydrologic variable simulations such as streamflow, sediment yield, and so on. The most tedious part when dealing with the SWAT model is the calibration process needed to have accurate outputs. SWAT-CUP is commonly used for the calibration, with its five options for SWAT users. The advantages and limits of these methods have been explained in this review, and it was depicted that most of studies found that the SUFI-2 seemed to be widely used despite that it is responsible for all uncertainties as well as GLUE. The performance of the calibrated model is mostly assessed using R^2 or NASH.

Statistical tools used for seasonal forecasting were also reviewed. They help to directly link rainfall to climatic indices or oceanic basins. There have been many studies done worldwide and particularly in the Sahel; however, no study was found regarding the connection between oceanic indices and streamflow in the Sahel. Additionally, it was found that most studies in the Sahel region have used linear statistical methods rather than non-linear approaches for forecasting rainfall and streamflow.

The techniques used for temporal disaggregation were also reviewed, and it was found that the method of fragments was mostly used, despite that it has many challenges such as the lack of maintaining correlation between the disaggregated data of a previous year and a given year.

There were several tools used to measure forecast abilities, and it was found that the Hit rate score was not widely used in seasonal forecasting in the West African Sahel, despite that it gives a good understanding of how predictions are well corresponding to observations. Moreover, the copula approach was explained and reviewed, with specific cases regarding its use.

Therefore, it is found that the Sahel region needs an adapted seasonal forecasting model that can consider its specific aspects and characteristics, as this approach could help in improving forecast abilities. Additionally, there are some statistical tools which could particularly help in combining seasonal forecasting models. Thus, combining a seasonal rainfall forecasting model to that of a seasonal streamflow forecasting and a seasonal flood forecasting models is of paramount importance for the Sahel region. Finally, coupling all these models with a simple risk measure regarding the probabilities of overtopping of flood protection dykes is very important in order for stakeholders to make correct decisions.

Chapter 4. Study Area and Data Presentation

This chapter presents the study area by describing West Africa through the most important aspects such as the relief and vegetation, which could have an impact on the rainfall of the Sahel region. The Sirba basin, on which the thesis work is focused, is fully described in all aspects in order to provide a good understanding of the local characteristics in terms of climate, physiography, and social and economic activities. The data used in this thesis work are presented and their quality discussed.

4.1. West Africa

West Africa abuts the Atlantic Ocean and extends from Nigeria to the coast of Senegal in the south and west. It is bounded to the east by Cameroon and Chad, and to the north by the Sahara desert. It includes eight French-speaking countries whose neighbors are English-speaking countries such as Nigeria, Ghana, Gambia, Sierra Leone and Liberia, and also Portuguese-speaking, such as Cape Verde and Guinea-Bissau. Its population of about 330 million with a growth rate over 3% per year. This population is mainly rural. This part of Africa is divided into two climatic zones: (1) the south equatorial, with a hot and humid climate, abundant rainfall and forests; and (2) the Northern Sahel, with savannas, steppes and desert regions (Philippon, 2002). The most important rivers flowing through this part of Africa are the Niger, the Gambia and the Senegal. More details on the West African relief and vegetation are provided in the following.

4.1.1 Relief

The West African relief is relatively simple compared to other parts of the continent. It is mainly characterized by large tabular spaces, plateaus or low-lying plains about 200m on average. The plains occupy the coastal border and are larger in scope in Senegal, Gambia and the Niger valley. Aside from the Senegal River, Volta River, and

Niger River, these plains are drained by small watercourses; however, they have much higher than regular flows and reaches irrigating the coastal plains of East Africa. The two major water features associated with these plateaus are the Niger basin and lake Tchad. Lake Tchad, the inner Niger delta, and the southern marshes provide large amounts of water vapor to the atmosphere (Carrington *et al.*, 2001).

Medium altitude plateaus and mountains characterize the southern margins and the West African northeast. The uplands of "l'Air", Tibesti and Cameroon are the three major topographical items. Semazzi and Sun (1997, as cited by Philippon, 2002) showed the important role played by the Atlas-Hoggar in Sahelian precipitation during northern summer; it causes the appearance of a quasi-stationary wave in the northeast flow, associated with anticyclonic circulation northeast to the uplands and cyclonic circulation southeast. A second group of plateaus and uplands includes the mountains of Fouta Djallon and the Guinean Mountains (Nimba Mountains) to the west, and the Bauchi plateaus to the east.

4.1.2 Vegetation

In West Africa, vegetation formations are irregularly arranged by zonal bands. The homogeneity of West African vegetation is only relative because at fine scales, most vegetation is being moth-eaten by cultivated areas, or composed of different facies. This vegetation mosaic at the fine scale results from the contrasting soil conditions, which is in connection with the overlapping of reliefs, drainage and different soils as well as climate changes in the long and short term (Cole, 1982).

Three main types of vegetation of regional importance are identified in West Africa: dense forest, savannah and shrubland (see Figure 4.1).

The dense forest is mainly localized along the coast of the Guinea Gulf. As for the savannah, it can be wooded, bushy or grassy. It occupies the Sudano-Sahelian area located north of 7°N. Shrublands are without herbaceous stratum. The shrub savannah and shrubland occupy respectively 38 and 26% of the terrain in West Africa (Fuller & Ottke, 2002).

The role of vegetation (and soil moisture) on the variability of the atmospheric system was long neglected compared to the role of SSTs. But according to Hutjes *et al.*

(1998), the vegetation is a key element of the coupling between mainland and atmospheric part of the water cycle. Its interactions with the atmosphere result in exchanges that are controlled by albedo, roughness and the flow of sensible and latent heat.

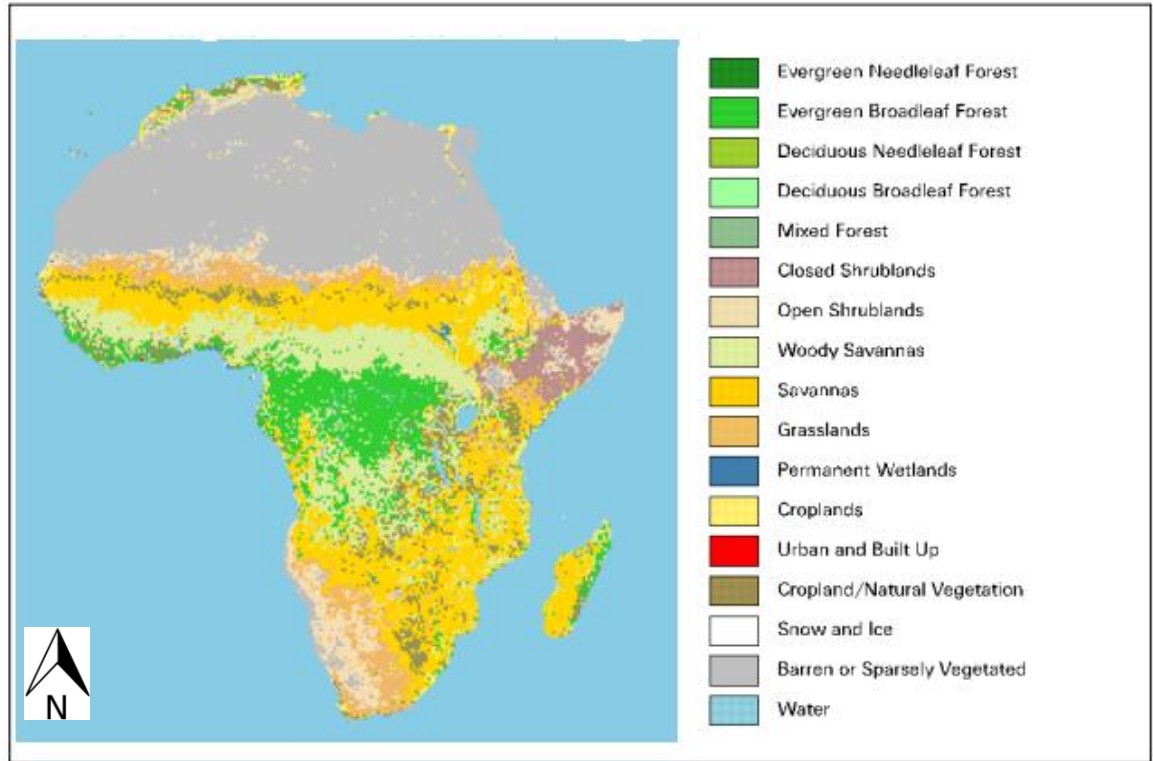


Figure 4.1 Land use in Africa (source: Philippon, 2002)

4.2. Sirba Basin

The study area in this work is the Sirba basin in northwestern Africa, which is a transboundary watershed shared by two countries: Burkina Faso and the Republic of Niger (see Figure 4.2). Several studies were conducted in this basin to better characterize and understand Sahelian climate dynamics and the impact of climate variability and changes.

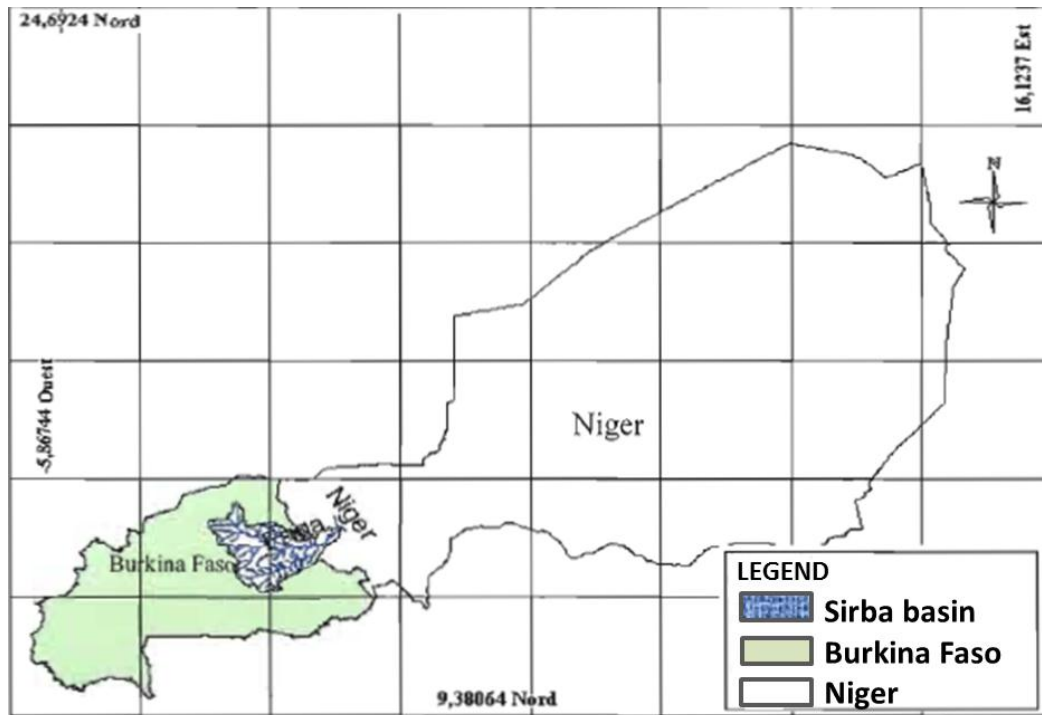


Figure 4.2 Location map of the Sirba watershed (source: AGHRYMET, 2004)

This choice of location was motivated by the fact that the Sirba basin is central in the Sahel region, so it is influenced by the climate characteristics of both the northern Sahel and the Sahara desert, and the southern Sahel and the Sudanian savanna. Moreover, there are many climate stations inside and around the basin that have been collecting climate data daily for more than 40 years. Another reason is that, locally, the Sirba tributary plays an important role in the hydrological regime of the Niger River at Niamey, as it contributes to the Sudanian flood in September (Sighomnou *et al.*, 2013).

The Sirba watershed is largely characterized by the rural sectors of agriculture, livestock breeding, and exploitation of forest resources. These are the main production activities of the weak local economy that is adversely affected by the depletion of the natural resource, poverty, recurring droughts, and high population growth, all of which are exacerbated by the effects of climate change. Thus, all these activities depend on streamflow, as rainfall is erratic in the area. Therefore, the present work can benefit the rural population (e.g., farmers) to plan the type of crops to sow according to the seasonal forecast information received a few months ahead.

4.2.1 Geographical description

This transboundary watershed, shared by Burkina Faso and Niger, is located between latitudes 12°55'54"-14°23'30"N and longitudes 1°27'W-1°23'42"E, with an area of 38,750 km² (Mara, 2010). The Burkina segment is more than 80% of the total area. The Sirba River is one of the most important tributaries of the Niger River in the Liptako Gourma region (i.e., southwestern Niger, eastern Burkina Faso, and the adjacent border of Mali) (Taweye, 1995).

The relief of the watershed is somewhat rough, while the geomorphology is dominated by vast alluvial plains with armored hillocks with rocky outcrops in some places. The hillocks are connected to plains and rivers by glacis whose slope varies between 2 and 3%. These are where runoff, solids transport and soil development (changes) occur. The average altitude is 280m, with some peaks reaching 320m to 450m due to corresponding hills unaffected by erosion (Mara, 2010).

An analysis of the administrative and topographic maps of Burkina Faso and the Niger Republic shows that the Sirba basin encompasses 376 villages and 25 departments in eight provinces (Ganzourgou, Gnagna, Gourma, Kourittenga, Namentenga, Sanmatenga, Seno and Soum) in Burkina Faso, and three departments (Kollo, Say and Tera), five rural entities, and 44 administrative villages in Niger (Sawadogo, 2002). Figure 4.3 shows the administrative zones in the Sirba watershed.

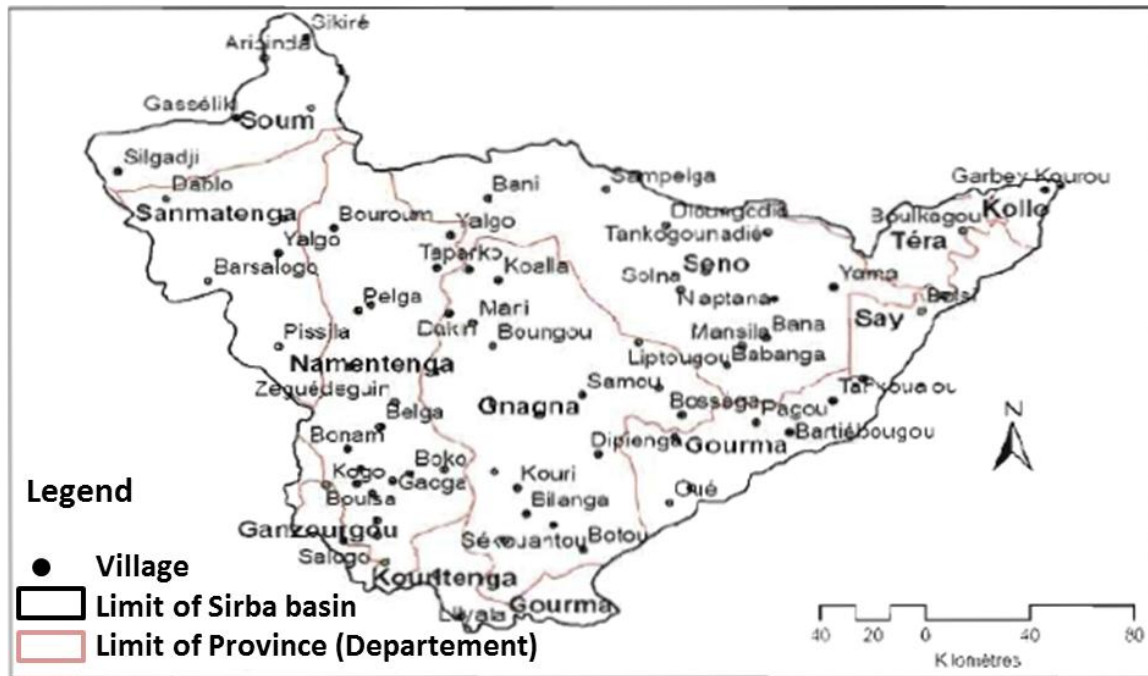


Figure 4.3 Administrative map of the Sirba basin (source: AGRHYMET, 2004)

4.2.2 Climate

The climate of the Sirba basin is tropical and Sudano-Sahelian. In general, the Sirba is influenced by three sub-climate zones based on the amount of precipitation, which decreases from south to north: the southern Soudanian zone with mean annual rainfall between 700 and 800mm, the northern Soudanian zone with mean annual rainfall ranging from 550 to 650mm, and the Sahelian zone with mean annual rainfall of 300-500mm (Taweye, 1995). Thus, throughout the watershed, the mean annual rainfall exceeds 500mm near latitude 14°N (Dori, Tera) and 800mm near latitude 12°N (Koupela). The largest quantities of rainfall are observed during the months of July to September (JAS), regardless of the climate zone. The climate is generally characterized by two seasons: a dry season (October to April) due to the harmattan (dry wind), and a rainy season (May to September) influenced by the WAM (cold wind) (Descroix *et al.*, 2009). In the watershed, the duration of the dry season increases gradually toward the north, and varies from six to seven months in the south to over eight months to the east and north (Fada, Zorgho). Evaporation is very high, exceeding the mean annual amount of precipitation, and the relative humidity is erratic, decreasing during the harmattan

period and increasing with the return of the WAM. Temperatures in the basin are high throughout the year, between 18°C and 40°C, and increase from south to north. Finally, the mean value of the wind speed is between 1.6 and 2.8m/s (Taweye, 1995).

4.2.3 Hydrography

The hydrographic network of the Sirba basin is relatively dense, as it consists of three main tributaries (Sirba, Faga, and Yeli), as well as a few dammed water reservoirs (Mara, 2010) (Figure 4.4). Based on the rainfall pattern, the hydrological regime in the watershed is a Sahelian type, as it is characterized by non-sustainable flows with an exoreic pattern (Descroix *et al.*, 2009). The streamflows produced on the northeastern slopes drain to the Niger River (Savadogo, 2004).

At the upper bed of the Sirba there is a series of depressions with intermittent flow. However, during the wettest years, some sections of the Sirba have water constantly. Locally, the Sirba plays an important role in the hydrological regime of the Niger River at Niamey, as it is involved in the Sudanian flood in September (Taweye, 1995). The two major tributaries, the Faga and the Yeli, join the Sirba near the border between Niger and Burkina Faso. With a length of 360km, the Faga is the longest tributary of the basin. The outlet of the Sirba basin is located at Garbé-Kourou in Niger, where the streamflows drain into the Niger River.

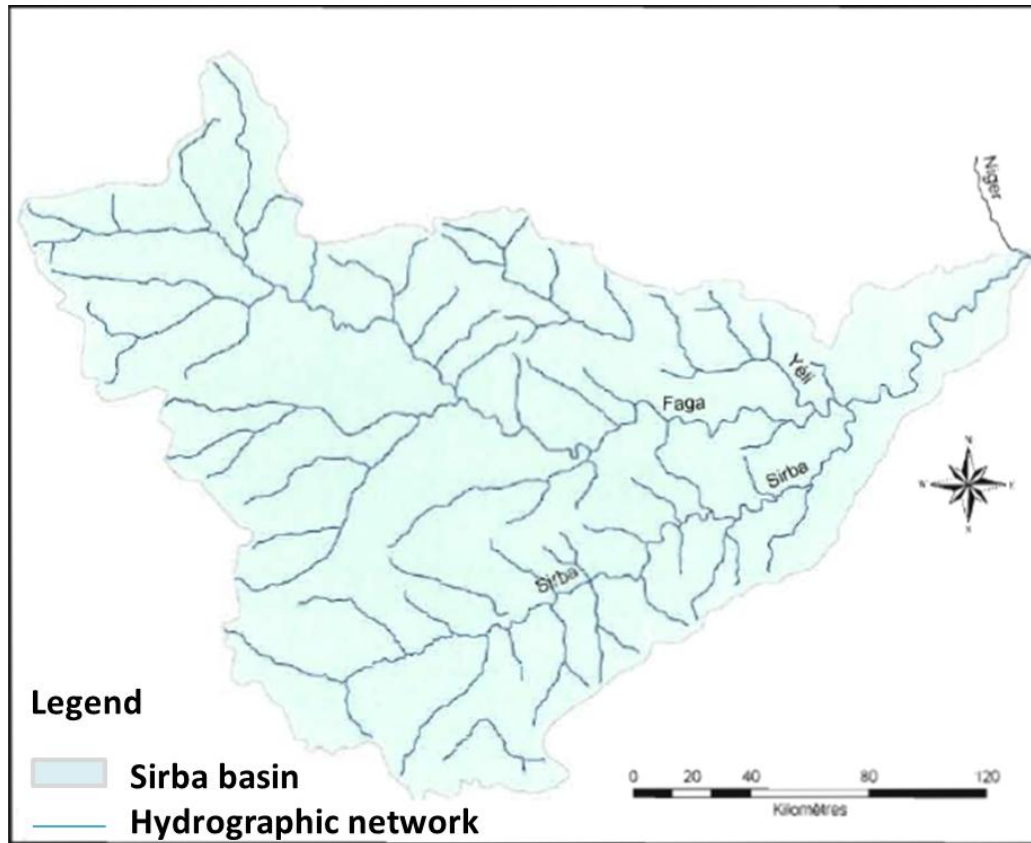


Figure 4.4 Hydrographic network of the Sirba basin (source: AGHRYMET, 2004)

4.2.4 Geological and hydrogeological characteristics of the basin

With respect to the geological structure, in Burkina Faso the Sirba basin occupies the eastern component of a large Birimian furrow elongation that runs northeast to southwest from Sebba to Manga, through Bogandé, Bilanga, Koupéla and Garango. This path includes greenstone, amphibolite (purplish red shale), quartzite, and a set of volcanic-sedimentary rocks extruded by various plutonic rocks (Figure 4.5). In the watershed, these formations are found near the border with Niger, where they run from Tiabongou to Songori until they reach the Niger River. At the center of the basin, granites run through Bassiéri, Bilanga, and Yanga.

From a socio-economic perspective, these formations generate very heavy clay, soil and regolith, and are not manned because of the lack of water facilities (e.g., water boreholes). In Niger there are two main formations: birrimian and granites (syntectonic

granites) from Liptako. Table 4.1 shows the ratio of the area of each geological formation in the Sirba basin. To summarize, referring to the study by Savadogo (2004), the geology of the basin is dominated by two major formations: granites and birrimian.

Table 4.1 Geological formations in the Sirba watershed (source: Taweye, 1995)

Geological nature	Corresponding area (km²)	Percentage area (%)
Syntectonic granites	29 270	75.54
Birrimian (shale, greywacke, metamorphic, greenstone)	9 200	23.74
Calco-alkaline granites post-tectonic	280	0.72

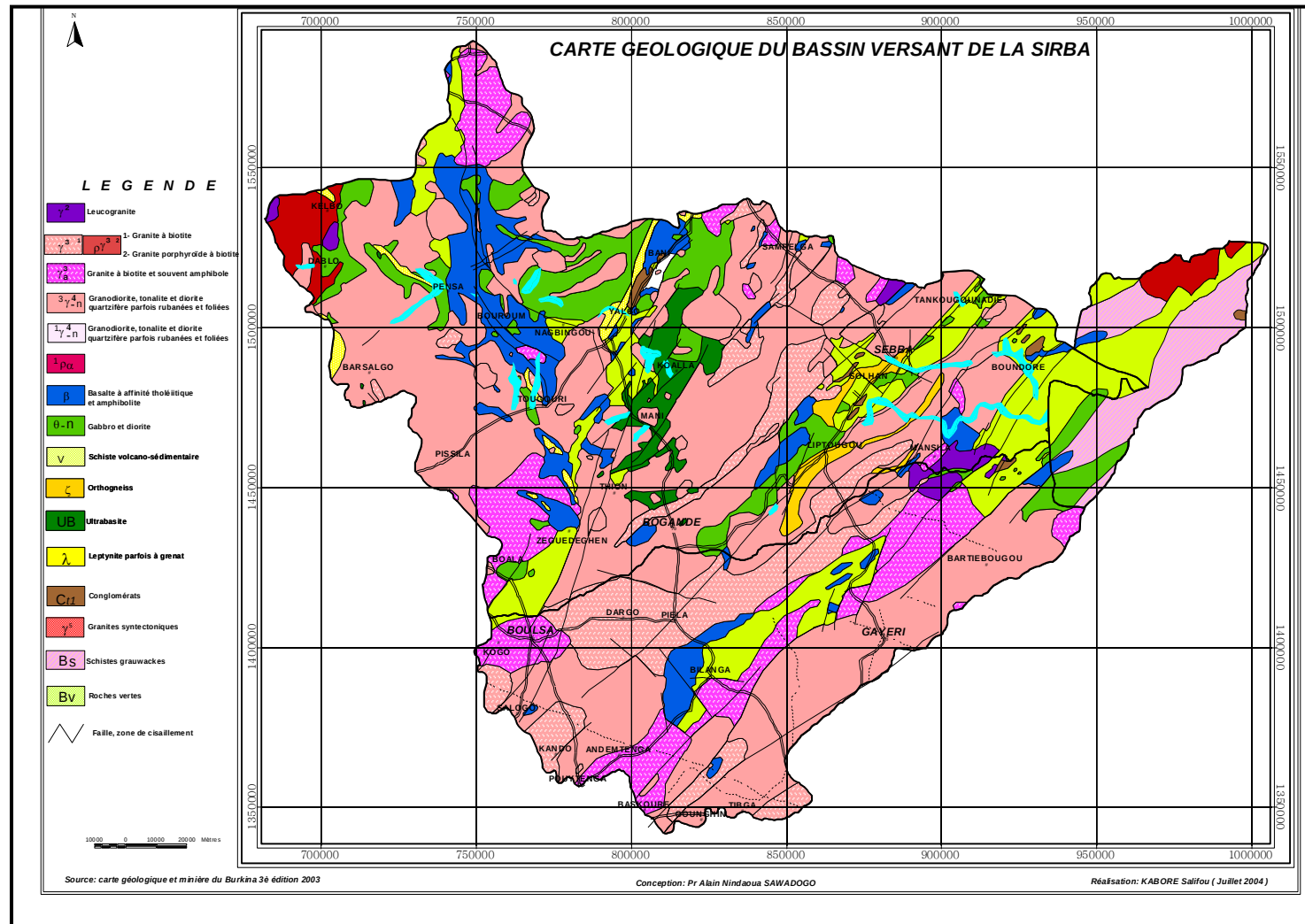


Figure 4.5 Geological map of the Sirba basin (source: AGHRYMET, 2004)

4.2.5 Soils and land use

The pedology of the Sirba watershed is related to its geomorphological evolution and the climate of the region. The soil type is similar to that of northeastern Burkina Faso, and most soils in the watershed are highly weathered, friable and of low structure and high density. They are tropical ferruginous soils with sandy surface textures and surface crusts, and they generally lack nutrients, exhibit poor structure, and have low organic matter content (Dembele & Somé, 1991). The soils are poorly developed due to the nature of the parent rock and their thinness. Thus, the geomorphology is general because the Niger River is a sequence of:

- flood basins favorable to rice development
- different levels of mostly sandy alluvial terraces with discontinuous bands of clay depression areas
- colluvium-alluvial glacis with low ramps (less than 3%) and long slopes
- glacis of denudation with steep ramps
- continental shelf battleship

To summarize, two types of soil are predominant in the basin: crystalline rock composed of former kaolinitic materials, and soil from the alteration of montmorillonite materials. The former has low fertility and the second has relatively high fertility, and both have significant drainage. These hydromorphic and ferruginous soils are found mostly in the farmlands.

However, despite the thinness of these soils, the vegetation still grows due to the high influence of rainfall. The vegetation evolves from south to north: savannah dotted with clear sparse forest, to steppe shrubs and bushes. The vegetation in the watershed is thorny and lightly wooded savannah (Andersen *et al.*, 2005). Thorny types are generally stunted due to the dry climate and overgrazing (Devineau & Serpantié, 1991). They are either widely dispersed, or form thickets in more or less parallel bands alternating with bare patches (striped bush). The following species are found in different areas of the basin:

- Gum trees (*dacryodes hexandra*) that form real wooded belts around permanent and semi-permanent ponds and along the thalwegs

- Balanitis (*balanites aegyptiaca*) that are scattered throughout the basin, as well as jujube trees (*zizyphus vulgaris*)
- Baobab trees (*Adansonia digita*) that are sometimes near important settlements, particularly in the Aribinda, Bouroum, and Liptougou regions
- Grass (*Andropogon gayanus*, *penissetum*) that generally forms a discontinuous carpet, except in depressions with clay soils, where meadows of wild fonio (*Parricum Lactum*) are dominant
- Shea trees (*Vitellaria paradox*) and “nééré” trees (*Parkia biglobosa*) in the extreme south of the watershed

4.2.6 Demography

The population of the Sirba watershed is mainly situated in the provinces of Ganzourgou, Gnagna, Gourma, Kourittenga, Namentenga, Sanmatenga, Seno, and Sourn in Burkina Faso, and the departments of Kollo, Say, and Tera in Niger Republic.

The watershed has an estimated population of 1,105,209 inhabitants, and an average density of 28.6 inhabitants per km² (Mara, 2010). The area had low population density until the 1980s, when it began receiving more immigrants due to its relatively high average rainfall compared to that of the central plateau, and the irrigation schemes in place. Since then the population has increased exponentially due to the traditional gold mining activities in the area, leading to the establishment of several gold exploitation sites.

The population of the region is largely composed of four social groups. In the Niger section, they are mainly *Zarma-Songhai* and *Fulani*, who were there before the colonial penetration and occupy all the villages along the Sirba River. *Touareg*, who are somewhat nomadic and engage in farming and gardening, can also be found, as well as other ethnic groups such as the *Hausa* and *Gourmantché*, who are there because of the gold mines in the area. In the Burkina Faso section of the basin, the population is composed of three major groups: the *Gourmantché* and the indigenous majority, the *Mossi* and the *Fulani*. The first two groups are sedentary farmers, though now they increasingly farm livestock. The *Fulani*, breeders by tradition, now also practice agriculture.

4.2.7 Environmental, social, and economic activities

The Sirba watershed is largely characterized by the rural sectors of agriculture, livestock breeding, and exploitation of forest resources. These are the main production activities of a weak local economy that is adversely affected by the depletion of the natural resource potential, poverty, recurring droughts, and high population growth, all of which are exacerbated by the effects of climate change. The dominant socio-economic activities in the basin are practiced by the local population. The main economic and environmental issues are:

- Increasing demographic and pastoral pressure around irrigation schemes. The schemes are drowning under water, and thus limit the potential cropland (see the case of the construction of the Bilanga dam, which led to farmers retreating to the highlands). They are also attractive places for agricultural immigrants.
- The integration of agriculture (except in the Niger area) and livestock breeding is still low. However, the practice of agriculture and livestock breeding by farmers is increasing, and the high livestock numbers could be involved in the densification of agriculture (i.e., the use of ruminant manure, and cattle, donkeys, and horses for animal traction). However, the improvements are still modest, as animal traction is undeveloped and only households near farms can benefit from a regular supply of organic manure. Raising awareness, disseminating information, and conducting training could accelerate the process.
- Random and decreasing rainfall dictates the actions that could contribute to better management of water availability. The supervision of agricultural and pastoral sectors should be intensified, and the mutual benefits of integration better understood. In addition, raising awareness of the vulnerability of ecological inheritance and the need to include conservation, protection, and improvement in production must be initiated and/or strengthened to motivate producers to practice eco-development.

4.3. Data Presentation

The data used in this study include climate, hydrological, land use, soil and topographical data. The climate data were used to develop the seasonal forecasting model, while the hydrological data (along with the climate data), land use, soil and topographical data are used to develop the Sirba hydrological model and the Niger River hydrodynamical model at Niamey.

4.3.1 Climate data

The climate data include rainfall (*In situ* daily data and satellite data) and atmospheric data. The *in situ* daily rainfall data (i.e., from observed station) were used to develop the seasonal rainfall and streamflow models. The satellite precipitation data and atmospheric data were used in the selection of potential predictors, before finalizing the development of the rainfall forecasting model.

a) Rainfall Data

Two sets of rainfall data were used: observed data from rain gauges, and precipitation data from the Climate Research Unit (CRU) database.

i) In-situ rainfall

The *in situ* daily rainfall data are from a network of 11 rain gauge stations in Burkina Faso and Niger. These daily rainfall were obtained from the national meteorological offices of Burkina Faso and Niger, and span the period 1960-2008. Five of these stations are located within the watershed, while the remaining six stations at most of 25km from the watershed boundary (see Figure 4.6). The Thiessen polygon method is a standard method widely used to calculate average areal precipitation (Ball & Luk, 1998). It was implemented to estimate average rainfall over the watershed from the 11 rainfall time series (Table 4.2). Though there is a lack of good climate data throughout Africa, only a few gaps were found in these records, and they do not alter the quality. The ratio of the missing data was less than 10% and varies from 0 to 7% over the observation period.

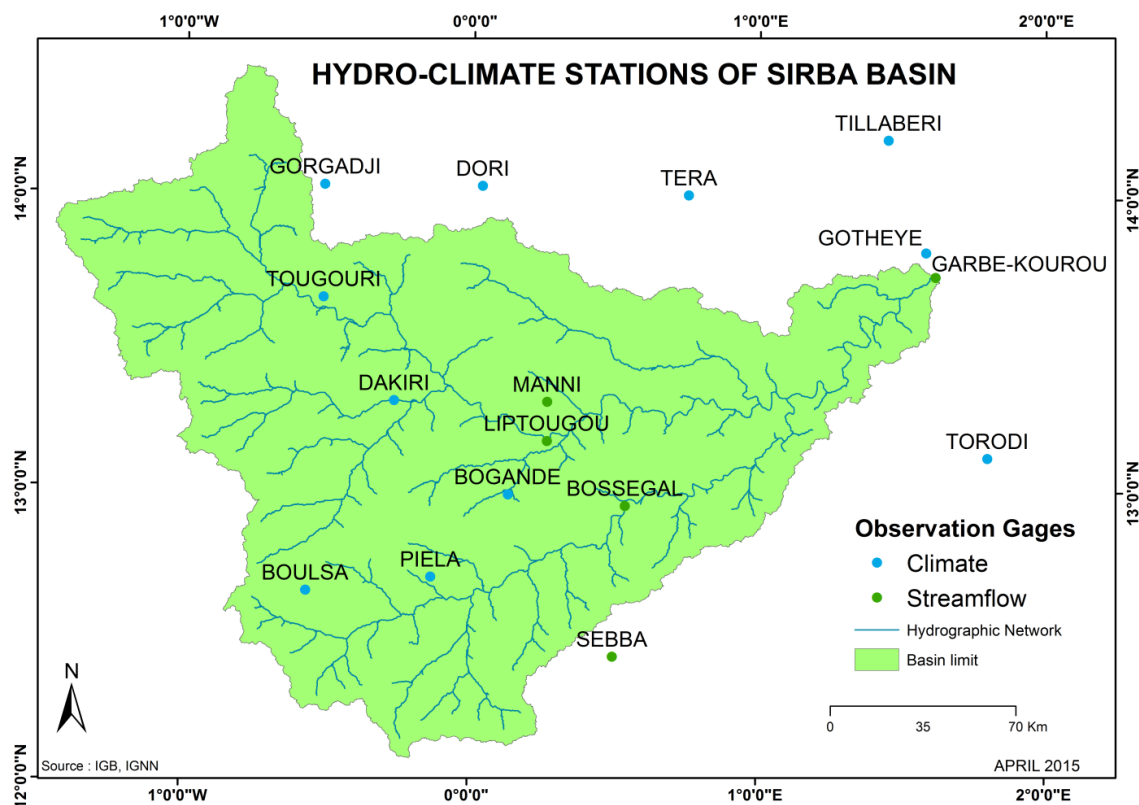


Figure 4.6 Observed rainfall and streamflow measurement stations

Table 4.2 Specifics of rainfall stations

Station number (code)	Station name	Longitude (degrees: °)	Latitude (degrees: °)	Country
320006	Torodi	1.8	13.12	Niger
320002	Tera	0.82	14.03	Niger
320004	Tillaberi	1.45	14.20	Niger
320005	Gotheye	1.58	13.82	Niger
200082	Boulsa	-0.57	12.65	Burkina Faso
200026	Dori	0.03	14.03	Burkina Faso
200085	Bogande	0.13	12.98	Burkina Faso
200048	Dakiri	-0.27	13.30	Burkina Faso

Station number (code)	Station name	Longitude (degrees: °)	Latitude (degrees: °)	Country
200024	Gorgadji	-0.52	14.03	Burkina Faso
200086	Piela	-0.13	12.70	Burkina Faso
200047	Tougouri	-0.52	13.65	Burkina Faso

ii) Climatic research unit rainfall

The second type of rainfall data is derived from observations generated by spatial procedures (New *et al.*, 2002). This dataset are monthly precipitation time series of Climatic Research Unit (CRU TS 3.21 0.5° global) with a spatial resolution of 0.5° × 0.5° and defined on latitudes 10°N–15°N, and longitudes 2°W–2°E. They are high-resolution, gridded precipitation time series, which is more than the extent of the Sirba watershed. They are sourced from the British Atmospheric Data Centre (BADC, <http://www.cru.uea.ac.uk/cru/data/hrg/>) and span the period 1901–2012 (BADC, 2015). In this study, the series are used as predictand to select the best potential predictors for the seasonal rainfall forecasting model. In general, CRU data are monthly data covering all land areas (excluding Antarctica) on a square grid of 0.5° latitude and longitude. While there are other global climate time series (Willmott & Matsuura, 2001). The CRU dataset provides high spatial resolution and long-term coverage, as well as the most comprehensive compilation of surface climate variables (New *et al.*, 2000). There is free access to the different versions of CRU data, which is updated regularly. Because of these advantages, CRU data are used in many global studies on climate variability and for hydrological modeling in West and Central Africa (Ardoin-Bardin, 2004; Held *et al.*, 2005; Diello, 2007; Mahé *et al.*, 2008; Doll *et al.*, 2003). They are chosen for their quality, as well as the ability to compare variations in climate with variations in other phenomena. However, satellite observations are limited in that they locate good rainfall but underestimate considerably. This is because they are based on the assumption that the clouds whose peaks reach high altitudes give more precipitation. Moreover, the estimation of precipitation in some regions with the use of microwave data from existing satellites is

limited by the small steps repetition rate over a given area. The description of these data is presented in Table 4.3.

b) Atmospheric Data

The third type of climate data is atmospheric, and includes Sea Level Pressure (SLP), Relative Humidity (RHUM), Air Temperature (AirTemp), Zonal Wind (UWND), Meridional Wind (VWND), and Sea Surface Temperature (SST). These variables are monthly NCEP-DOE Reanalysis data sourced from the National Oceanic and Atmospheric Administration (NOAA: <http://www.esrl.noaa.gov>), except the SST data series (NOAA NCDC ERSST version3b sst) which are obtained from the IRI data library (International Research Institute for Climate and Society: <http://iridl.ldeo.columbia.edu>) (IRI, 2015). These data span from January 1979 to August 2013, except for the SST data which cover the period January 1960 to December 2013. The spatial and temporal resolutions depend on the individual variables (i.e., the provided source), and Table 4.3 presents the description of these variables. Figure 4.7 highlights the monthly averages of the NCEP Reanalysis 2 of relative humidity (90°N to 90°S and 0 to 357.5°E).

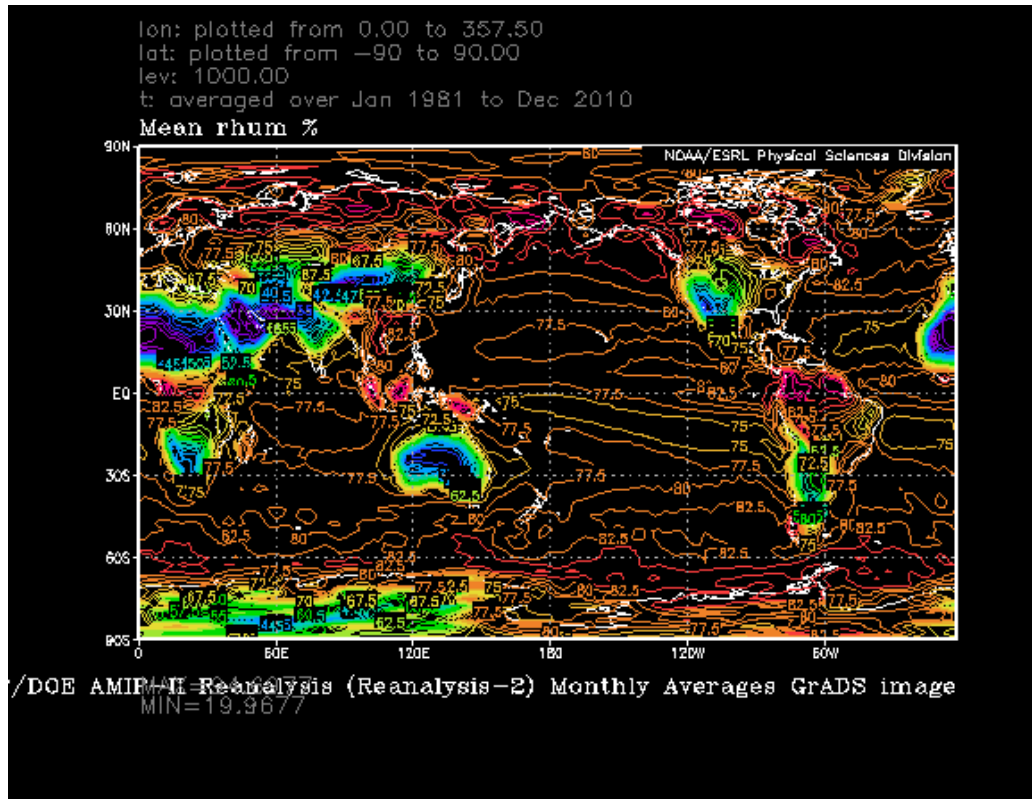


Figure 4.7 Illustrative image of mean monthly relative humidity (source: <http://www.esrl.noaa.gov>)

Table 4.3 Description of atmospheric data

<i>Parameter</i>	<i>Units</i>	<i>Level</i>	<i>Reference data</i>	<i>Spatial coverage</i>	<i>Temporal coverage</i>
Sea level pressures (SLP)	Pa/s	1000 hPa	NCEP 2	2.5°x2.5° grid 15N - 45S, 60W - 10E (Atlantic Ocean)	1979/01/01 to 2013/08/31
Air temperature (AirTemp)	°K	1000 hPa	NCEP 2	2.5°x2.5° grid 20N - 15S, 120W - 70W (Pacific Ocean)	1979/01/01 to 2013/08/31
Meridional wind (VWND)	m/s	1000 hPa	NCEP 2	2.5°x2.5° grid 90N - 90S, 0E - 357.5E	1979/01/01 to 2013/08/31

<i>Parameter</i>	<i>Units</i>	<i>Level</i>	<i>Reference data</i>	<i>Spatial coverage</i>	<i>Temporal coverage</i>
Zonal wind (UWND)	m/s	1000 hPa	NCEP 2	2.5°x2.5° grid 90N - 90S, 0E - 357.5E	1979/01/01 to 2013/08/31
Relative humidity (RHUM)	%	1000 hPa	NCEP 2	2.5°x2.5° grid 40N - 30N, 20E - 35E (Mediterranean Basin)	1979/01/01 to 2013/08/31
Sea surface temperature (SST)	°C	Surface	NOAA NCDC ERSST version3b	2°x2° grid 39N - 15S, 60W - 15E (Atlantic Ocean)	1854/01/01 to 2013/08/31
Climatic research unit rainfall (CRU)	mm	Surface	CRU	0.5°x0.5° grid 2°W-2°E, 10°N-15°N	Jan. 1901 to Dec. 2012

4.3.2 Hydrological data

In addition to climate data, hydrological data are also used in this study to develop seasonal forecast models and the Sirba hydrological model. The data are daily mean streamflows collected from five stations (see Figure 4.6). The first data, from four stations (Table 4.4), are collected from the Directorate General for Agriculture and Fisheries (*Direction Générale de l'Agriculture et des Ressources Halieutiques*: DGARH) in Burkina Faso. The daily streamflow data collected from the flow measure station at Garbé-Kourou (outlet of the Sirba basin) are sourced from the Directorate General of Water Resources (*Direction Générale des Ressources en Eau*: DGRE) in Niger.

These data are characterized by missing data ratios of 0 to 47% over the period. Figure 4.8 shows that the annual missing data ratios are significant (>10%) at the beginning of the series at the Liptougou flow station until 1978, and more pronounced from July to October, and there were missing measures in 1977. In the more than 39 years the dataset has been used, only the periods of 1984 to 1986, 1992 to 1994, and 1999 to 2002 are without gaps. A summary of the characteristics of these data is provided in Table 4.4. Also, some data related to dams (e.g., physical characteristics, operational modes) were supplied by the mentioned national services. They explain the flow patterns,

and provide useful information about dams and other waters structures in the Sirba watershed. These data are needed to develop the Sirba SWAT hydrological model.

Table 4.4 Hydrological stations of the Sirba watershed

Station Name	Longitude	Latitude	Observation period	Ratio of missing data (%)	Country
Liptougou	0.26	13.16	1973-2011	16.32	Burkina Faso
Manni	0.26	13.30	1973-2003	29.18	Burkina Faso
Bossegal	0.5405	12.94	1973-2004	47.43	Burkina Faso
Sebba	0.50	12.43	1981-2010	19.80	Burkina Faso
Garbé-kourou	1.89	13.72	1989-2002	9.7	Niger

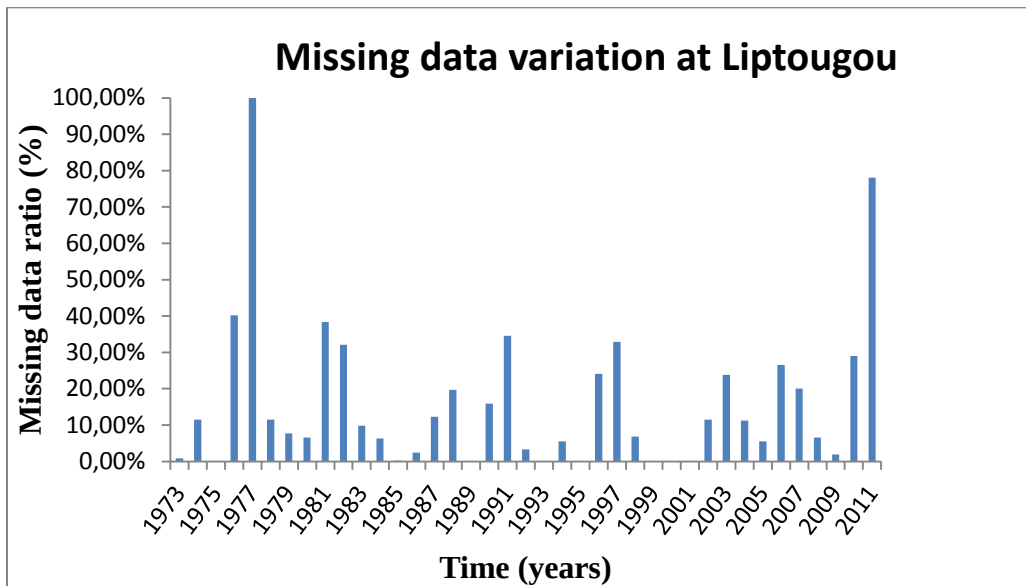


Figure 4.8 Gap ratios in daily flow series at Liptougou station (1973-2011)

4.3.3 Soil and land use data

Soil and land use coverages are downloaded from the waterbase database (George & Leon, 2007). The land coverages in the database have a resolution of 500m, and come from the Global Land Cover Facility (Hansen *et al.*, 1998). The soil map has a resolution of 1km, and was generated by the Food Aid Organization and the United Nations Organization for Education, Science and Culture (FAO/UNESCO, 2003).

The land use and soil type maps are later reclassified and overlaid to represent the study area conditions, and the threshold values for land use, soil and slope maps are set.

4.3.4 Topographical Data

The Sirba watershed and sub-watersheds are delineated using a 90m resolution digital elevation of West Africa, downloaded from the Shuttle Radar Topography Mission (SRTM, 2004). After the delineation (using the automatic DEM-based method), the watershed is divided into nine sub-basins according to the topography in order to perform the spatial parameterization of the SWAT model.

4.4. Synthesis and Partial Conclusion

The geographical description of West Africa showed that the relief of this region is also a source for rainfall occurrence, as the different mountains and the vegetation in Guinea Gulf play an important role in West African Sahel rainfall. The choice of the Sirba basin is also wise because this area has three sub-climate zones (southern Sudanian zone, northern Sudanian zone, Sahelian zone) with different rainfall regimes. The study area, being central in the Sahel region, is a rural area in which most of the land is utilized for agriculture and livestock farming. Additionally, many studies have been conducted in this area due to the data availability over a long period and at a fine scale. The daily rainfall data and hydrological data are of good quality despite the presence of some gaps. The remaining atmospheric data are easily accessible from international research center websites, and they do not have missing data that could alter the forecasting process. It is particularly found that the hydrological data have more missing data compared to other

climate and atmospheric data. All these data combined to the good choice of the area would make easier the different model development.

Chapter 5. Seasonal Rainfall Forecasting

The selection of predictors constitutes the most important step in seasonal rainfall forecasting, particularly in the Sahel, where rainfall is dependent on the WAM, which in turn is influenced by different factors. The approach developed to select the predictors, as well as the resulting products, is presented in this chapter. Also, the selection of a good lag time is important in dealing with seasonal rainfall forecasting, and this aspect is presented and fully detailed. Furthermore, in this chapter the models developed for seasonal rainfall forecasting are described, with a particular focus on models with changing parameters and the multiple change point detection algorithm. The Bayesian model selection method is well explained, and the criteria used to find the best seasonal forecasting model are clearly clarified. Finally, the chapter closes with seasonal rainfall forecasting using three types of regression trees and a feed-forward neural network.

5.1. Selection of Predictors and Potential Lag Time

This is a summary of the selection of predictors and the lag time, as is fully explained in Section 5.2 in the study by Gado *et al.* (2015a).

This process consisted of searching for good predictors and using them to improve the seasonal forecasting abilities. A preliminary screening is conducted using a combination of simple correlation analysis and the physical relationship between the WAM and each predictor.

First, a pool of predictors was formed based on information related to the WAM. It was used to test variables with a known relationship with the WAM, and to avoid those with no effect on the monsoon dynamics (e.g., aerosols) or the consequences of monsoon velocity. Another criterion was to consider predictors with the best ratio between spatial resolution (i.e., as fine as possible) and temporal length.

The next step was to define a large CRU grid of rainfall that covers more than the extent of the Sirba watershed. The precipitation was averaged over July, August and

September (JAS), which is the period of highest rainfall in the Sahel. The averaging was done for various lengths of time according to the length of each predictor.

As a preliminary test, the rainfall for the periods 1960 to 2010, 1970 to 2010, 1980 to 2010, 1990 to 2010, and 2000 to 2010 was analyzed to determine the most favorable periods; that is, those with significant signals with the pool of predictors. Based on the defined periods, a group of predictors was considered for the preliminary test to check the presence of significant correlations ($R > 0.5$) between each predictor and CRU rainfall. After this first selection of the pool of predictors using the monthly precipitation time series of Climatic Research Unit (CRU TS 3.21 0.5° global), they were further screened by using the observed precipitation from the rain gauge stations to detect the best predictors among the group. To do this, the method developed by Sittichok *et al.* (2014) was employed to link the observed rainfall and each predictor.

After searching for periods with significant signals with the set of predictors, the period 1970 to 2010 was selected as the best for the rainfall dataset (predictand), as it gave the best signals in the different correlation maps. The establishment of synchronous correlations between predictors and predictand (i.e., the correlations between rainfall and predictors averaged over the JAS period) revealed good signals during the defined period 1970 to 2010.

The applied methods were able to make a clear contribution to the selection of a pool of predictors for developing rainfall forecasting models over the Sahelian region. The period that yielded the best NASH (i.e., the optimal lag time) was selected for each tested predictor, namely AirTemp, RHUM, SLP, UWND, VWND and SST. The results of this process showed that AirTemp, RHUM and SLP are the best predictors with respect to NASH values of 0.53, 0.52 and 0.46 respectively, and are also the best coefficients of determination and hit rate scores.

5.2. Linear and Non-Linear Approaches for Statistical Seasonal Rainfall Forecast in the Sirba Watershed Region (SAHEL)

Climate **2015**, *3*, 727-752; doi:10.3390/cli3030727

OPEN ACCESS

climate

ISSN 2225-1154

www.mdpi.com/journal/climate

Article

Linear and Non-Linear Approaches for Statistical Seasonal Rainfall Forecast in the Sirba Watershed Region (SAHEL)

Abdouramane Gado Djibo ^{1,2,*}, Harouna Karambiri ¹, Ousmane Seidou ², Ketvara Sittichok ², Nathalie Philippon ³, Jean Emmanuel Paturel ⁴ and Hadiza Moussa Saley ⁵

¹ International Institute for Water and Environmental Engineering (2iE), 01 BP 594, Ouagadougou 01, Burkina Faso; E-Mail: harouna.karambiri@2ie-edu.org

² Department of Civil Engineering, University of Ottawa, Ottawa, ON K1N 6N5, Canada;

E-Mails: Ousmane.Seidou@uottawa.ca (O.S.); ksitt028@uottawa.ca (K.S.)

³ Centre de Recherches de Climatologie, UMR6282 Biogéosciences CNRS, Université de Bourgogne, Dijon 21000, France; E-Mail: Nathalie.Philippon@u-bourgogne.fr

⁴ Institut de Recherche pour le Développement (IRD), Abidjan 08 BP 3800, Côte d'Ivoire;

E-Mail: manupaturel@gmail.com

⁵ Centre Africain d'Études Supérieures en Gestion (CESAG), Dakar BP 3802, Sénégal; E-Mail: hmsaley@gmail.com

* Author to whom correspondence should be addressed.

Received: 19 June 2015 / Accepted: 1 September 2015 / Published:

Abstract: Since the 90s, several studies were conducted to evaluate the predictability of the Sahelian rainy season and propose seasonal rainfall forecasts to help stakeholders to take the adequate decisions to adapt with the predicted situation. Unfortunately, two decades later, the forecasting skills

remains low and forecasts have a limited value for decision making while the population is still suffering from rainfall interannual variability: this shows the limit of commonly used predictors and forecast approaches for this region. Thus, this paper developed and tested new predictors and new approaches to predict the upcoming seasonal rainfall amount over the Sirba watershed. Predictors selected through a linear correlation analysis were further processed using combined linear methods to identify those having high predictive power. Seasonal rainfall was forecasted using a set of linear and non-linear models. An average lag time up to eight months was obtained for all models. It is found that the combined linear methods performed better than non-linear, possibly because non-linear models require larger and better datasets for calibration. The R^2 , Nash and Hit rate score are respectively 0.53, 0.52, and 68% for the combined linear approach; and 0.46, 0.45, 61% for non-linear principal component analysis.

Keywords: rainfall forecasting; neural network; non-linear principal component analysis; Sirba basin; West African monsoon; air temperature

1. Introduction

The summer rainfall of semi-arid regions of the world is known for its unreliability, which has a large impact on the continental hydrological cycle, water resources and food security. The Sahel, extending across Africa from the Atlantic Ocean to 30°E and from 12 to 17°N is the largest area of these regions, recording between 200 and 800mm/year from north to south, ~80% of the rain being recorded in July–September the cool rainy season. Temperature in this region ranges from approximately 18 to 36 °C. The recurrent droughts and subsequent famines that struck the Sahel in the 1970s (1972–1974), and the 1980s (1983–1985) and make it unique at the global scale led the scientific community to investigate possible mechanisms responsible for these dramatic events and to develop forecasting models to help coping with such phenomena. This area experienced severe droughts almost every two to three years. The devastate drought in 2012 affected more than 18 million people in nine countries with food insecurity, high grain prices and environmental degradation. This 2012 crisis came after the severe drought in 2010 [1–18].

The Sahelian rainfall pattern is season dependent and is directly related to the West African Monsoon (WAM).

Although, the WAM's dynamic is better understood nowadays, the main challenge with regards to its variability and predictability comes from its varying teleconnections. A

teleconnection is the linkage of climate variables between two different areas, which may be close or far to each other [19,20]. Teleconnections between Sahelian rainfall and the oceanic basins have changed quickly much since the 60's: the tropical Atlantic had the strongest influence during the years 60–70, then the equatorial Pacific (El Nino/Southern Oscillation) during the 80–90's, and the Mediterranean now [21–25]. Thus, the absence of well-established predictors that can be used to predict seasonal rainfall as well as streamflow in the Sahel partly explains why forecasts at all scales in the Sahel are tricky. Many studies attempted to forecast Sahelian seasonal rainfall and streamflow for the purpose to overcome the droughts impacts [4,7,26–33]. Unfortunately, most of these works focused on only sea surface temperatures (SST) over years [24,34–36]. Nevertheless, few studies attempted to use the atmospheric dynamics for carrying out seasonal forecasts. This prediction relies on the explicit simulation of major atmospheric processes [37–44]. Garric *et al.* [42] used the ARPEGE atmospheric model forced by SST anomalies and multivariate linear regression using SST and rainfall predictors observed before the monsoon season. They showed that the ARPEGE model did not give better seasonal rainfall predictions than simple regression systems.

Thus, the objective of this paper is to develop a method that would identify new skillful predictors for seasonal rainfall in the Sahel, and to compare a set of linear models to non-linear ones for forecasting JAS (July to September) rainfall amounts. This would provide the community with actionable seasonal information that would constitutes a major tool for farmers, decision makers and water resources managers in this region.

For such purpose, a pool of predictors is built by analyzing the physical influence on the WAM. Each predictor is tested as an input to a linear rainfall forecasting model as in [4]. At the end of the process, an optimal lag time and an optimal season are obtained to extract the predictor. Retained predictors are afterward also tested in new developed non-linear models. Finally, the forecast skills of the two groups of models are discussed.

2. Review of the Main Drivers of the Sahelian Rainfall Variability

In West Africa, the rainfall pattern is firmly related to the seasonal movement of the inter-tropical convergence zone (ITCZ) and consequently to the development of the WAM circulation [45].

The SST constitutes a key factor in the variability of the WAM and therefore the Sahelian precipitations as shown in several studies [12,34,37,46–48]. The tropical Atlantic Ocean is considered as the principal source of moisture for West Africa. Its impact on the WAM system was shown since 1970s when the Sahelian rainfall deficit was associated to colder SST in the north tropical Atlantic and warmer SST in the south and at the equator which promotes a southernmost ITCZ than the normal. These results were confirmed, and then

extended to longer time scales by many studies [13,49–52]. However, Janicot *et al.* [53] noticed that the relationship between Sahelian rainfall and Atlantic SSTs considerably decreased to a point to be statistically not significant during the dry period (*i.e.*, post 1970). This teleconnection changed from Atlantic SST to the SST over eastern and center Equatorial Pacific, in agreement with the work of [54].

A rainy season with below average rainfall in West Africa is usually associated to a warm period of El Niño-Southern Oscillation (ENSO). Janicot *et al.* [55] explained this link by a strengthening of the Walker circulation and a weakening of both the monsoon and the southern cell of the Hadley circulation. This situation leads to an increase in trade winds over the northern tropical Atlantic and a reduction in the water vapor in West Africa. Recently, a teleconnection with the Mediterranean Sea has been highlighted [56–58]. It seems to impact the WAM system in addition to the Atlantic and Pacific [58]. Rowell [56] found that the influence of temperature anomalies in the Mediterranean on the Sahel (for 1947–1996) is of a similar magnitude with that in the Pacific. From numerical simulations, he showed that a warm Mediterranean sea promotes excess rainfall in the Sahel. Additionally, Gaetani *et al.* [57] pointed out that a positive precipitation response to warmer than average conditions in the Mediterranean Sea is found in the Sudano-Sahelian belt in August to September.

Several recent papers analyzed the interactions of the different oceanic basins and the resulting impact on the Sahelian rainfall variability. Shaman and Tziperman [59] found that the interannual rainfall variability over the Mediterranean region is related to ENSO variability in the eastern Pacific *via* an eastward-propagating atmospheric stationary barotropic Rossby-wave train. Moreover, Lopez-Parages *et al.* [60] explained how the teleconnection with the ENSO appears modulated by multidecadal oscillations of the SST over the Atlantic and Pacific basins.

With regards to the role of the land surface on the Sahelian rainfall variability, Webster *et al.* [61] indicated that the use of the moist static energy (MSE) (its three components: sensitive, latent and potential) could improve the rainfall forecasts in the Sahel. They argued that the variation in temperature between the ocean and the continent is responsible for the monsoon circulation; and this circulation is even better explained when the moisture gradient is considered. Eltahir [62], Philippon and Fontaine [63], Hall and Peyrillé [64] highlighted the role of these gradients on the dynamics of WAM and the Sahelian rainfall variability. Zheng and Eltahir [65] found that a change in vegetation (*e.g.*, deforestation) on the Guinean coast has a direct substantial impact on atmospheric dynamics associated with the monsoon circulation through MSE gradients. Thus, some authors such as Wang and Eltahir [66] suggested the inclusion of vegetation dynamics in the modeling exercises as it constitutes an important process for simulating Sahelian rainfall variability. In addition, while testing the impact of vegetation in rainfall

variability simulation, Zeng *et al.* [67] found that the decadal variability is best reproduced when interactive vegetation is added to the model. The role of soil moisture on West African rainfall event is also addressed in some studies [62,68]. These authors emphasized that a positive anomaly of soil moisture would strengthen the monsoon circulation through a modification of MSE gradients Douville *et al.* [69], Douville [70], and Douville [71] found that any reduction in soil moisture in ARPEGE is associated with low intensities of precipitations. Thus, based on these results they concluded that soil moisture contributes to the interannual variability of rainfall in the Sahel.

It therefore appears that the climate in West Africa (predominantly in the Sahel) is determined by interactions between global processes (e.g., sea surface temperatures) and regional processes (e.g., physiographic characteristics). The use of parameters related to these processes in seasonal rainfall forecasting models would generate more skillful forecasts for the Sahel.

3. Materials and Methods

3.1. Study Area

The study area considered in this work is the Sirba watershed. This watershed, shared by Burkina Faso and Niger, is situated between latitudes 12°55'54"S–14°23'30"N and longitudes 1°27'W–1°23'42"E with an area of 38,750 km² (Figure 1) [72]. This choice is motivated by the fact that the Sirba basin is central in the Sahel region, and there are many climate stations inside and around the basin that have been collecting climate data daily for more than 40 years. Another reason is that, locally, the Sirba tributary plays an important role in the hydrological regime of the Niger river at Niamey, as it participates in its Sudanian flood in September. The Sirba extends over three sub-climate zones based on the amount of rainfall decreasing from south to north: a southern Sudanian zone with mean annual rainfall between 700 and 800 mm, a northern Sudanian zone with mean annual rainfall ranging from 550 to 650 mm and a Sahelian zone with mean annual rainfall of 300 to 500 mm [73]. The largest quantities of rainfall are observed during the months of July to September (JAS), regardless of the climate zone. The climate is generally characterized by the presence of two seasons: a dry season (October to April) due to the Harmattan (dry wind) and a rainy season (May to September) influenced by the WAM (wet wind). The hydrographic network is relatively dense and consists of three main tributaries (Sirba, Faga and Yeli) as well as some water reservoirs from dams [71]. Based on the description of the rainfall pattern, the hydrological regime in the Sirba watershed is of Sahelian type, as it is characterized by non-sustainable flows with an exoreic operation pattern. At the upper bed of the Sirba, there is a series of depressions

with intermittent flow. However some sections of the Sirba reaches have water constantly during the wettest years. Its vegetation formation is thorny, lightly-wooded savannah.

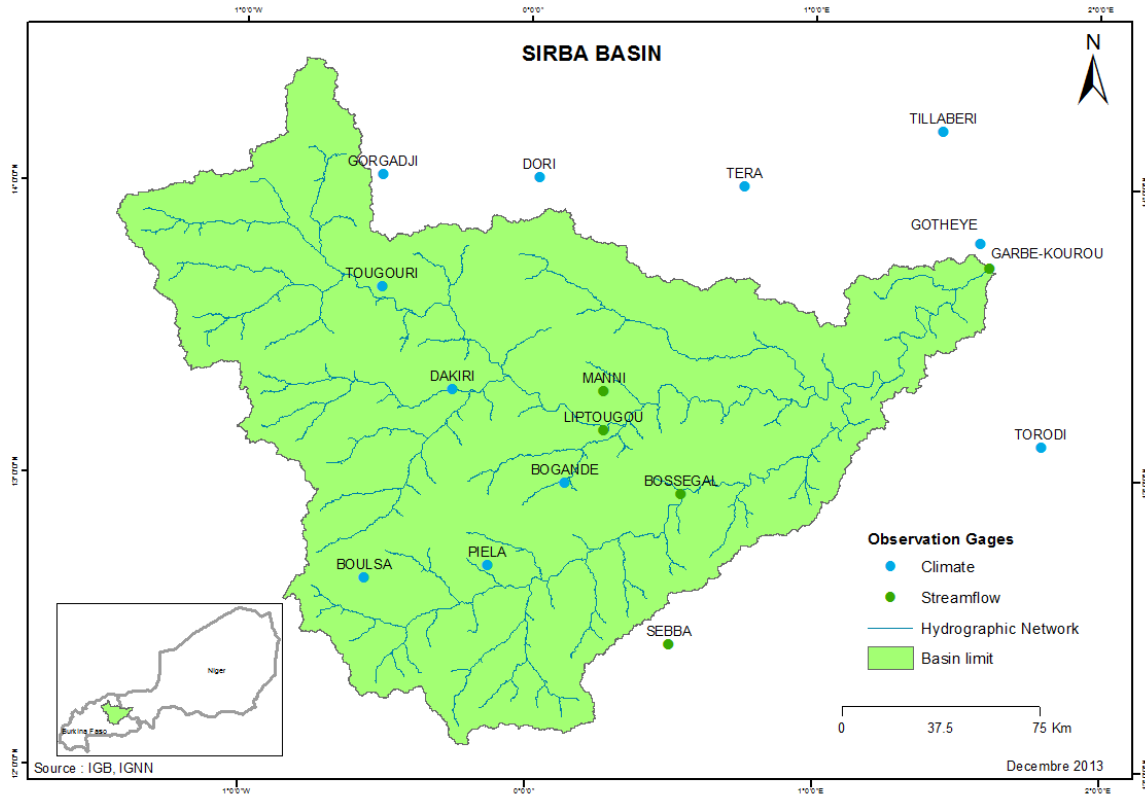


Figure 1. Sirba watershed and observation stations.

3.2. Climate and Atmospheric Data

Climate data used in this study include rainfall and atmospheric data. *In situ* daily rainfall data span the period 1960–2008, and are obtained from national meteorological offices of Burkina Faso and Niger. Five of these stations are located within the watershed while the remaining six stations are located at most 25 km from the watershed boundary (Figure 1). The Thiessen polygon method is a standard method widely used to calculate average areal precipitation [74]. It was implemented to estimate average rainfall over the watershed from the 11 rainfall time series (Table 1). Though there is a lack of good climate data throughout African Sahel, only a few gaps were found in these records, and they do not alter the quality. The *in situ* rainfall time series have less than 10% missing data as this ratio varies from 0 to 7% for all 11 stations over the period. Moreover, monthly precipitation time series of Climatic Research Unit (CRU TS 3.21 0.5° global) with a spatial resolution of $0.5^\circ \times 0.5^\circ$ and defined on latitudes 10°N – 15°N , longitudes 2°W – 2°E are used. They are sourced from the British Atmospheric Data Centre (BADC, <http://www.cru.uea.ac.uk/cru/data/hrg/>) and span the period 1901–2012 [75].

The atmospheric data considered are Sea Level Pressure (SLP), Sea Surface Temperature (SST), Relative Humidity (RHUM), Air Temperature (AirTemp), Meridional Wind (VWND), and Zonal Wind (UWND). These variables are monthly NCEP-DOE Reanalysis data sourced from the National Oceanic and Atmospheric Administration (NOAA: <http://www.esrl.noaa.gov>) except the SST data series (NOAA NCDC ERSST version3b sst) obtained from the IRI data library (International Research Institute for Climate and Society: <http://iridl.ldeo.columbia.edu>) [76]. These data span from January 1979 to August 2013, except for SST data which cover the period January 1960 to December 2013. The relationship of these predictors with the WAM is briefly described in the following paragraphs.

Table 1. Specifics of rainfall stations.

Station number (code)	Station name	Longitude (degrees: °)	Latitude (degrees: °)	Country
320006	Torodi	1.80	13.12	Niger
320002	Tera	0.82	14.03	Niger
320004	Tillaberi	1.45	14.20	Niger
320005	Gotheye	1.58	13.82	Niger
200082	Boulsa	-0.57	12.65	Burkina Faso
200026	Dori	0.03	14.03	Burkina Faso
200085	Bogande	0.13	12.98	Burkina Faso
200048	Dakiri	-0.27	13.30	Burkina Faso
200024	Gorgadji	-0.52	14.03	Burkina Faso
200086	Piela	-0.13	12.70	Burkina Faso
200047	Tougouri	-0.52	13.65	Burkina Faso

(a) Zonal Wind and Meridional Wind

Gallée *et al.* [77] have highlighted the importance of the meridian gradient of MSE on the movement of the ITCZ northward. According to their studies, MSE led the WAM and creates an environment favorable for deep convection over the Sahel. Thus, the spread of WAM to north is associated with strong gradient of MSE in the lower layers, the convergence result in the triggering of convection over the Sahel. These authors emphasized the relationship between the WAM, precipitation, the MSE, the meridian wind and sensible and latent heat in the Sahel. The strong north-south gradient (weak) of the meridian wind seems to be the result of a strong south-north gradient (weak) of MSE. The first maximum MSE gradient develops, followed by the maximum meridian wind. Then the MSE meridian gradient leads meridian wind speed and precipitation to a secondary maximum. In general, changes in precipitation appear to be a consequence of

MSE with a major role of the meridian wind.

The zonal and Meridian winds play a very important role in the flow in the middle and upper troposphere. Indeed, in the middle and upper troposphere, the zonal wind profile is dominated by three jets: AEJ (African East Jet), TEJ (Tropical East Jet) and WSJ (West South Jet). AEJ is at the level of 600 hPa and TEJ at 200 hPa. These two jets, which are located above the Sahel and Guinea, are important for the atmospheric dynamics in West Africa. The variation of these winds regulates the position of the AEJ, which, in turn, explains why during wet years (dry) for the Sahel [68,78]. Grist and Nicholson [79] showed evidence of AEJ above the Sahel (10°N to 15°N). JET looks very bound to the African monsoon. Indeed, through the Walker circulation, the intensity of the jet effect of the monsoon which forms the lower part of the cell [80]. Sultan [81] shows that the jet's installation date is a good indicator of the development of the monsoon. Finally, Nicholson and Grist [82] suggested that the jet is a response to precipitation but is not a cause of the variability of rainfall.

(b) Air Temperature

The impact of Air temperature on the WAM occurs through a link with atmospheric dynamics over West Africa. According to Fontaine *et al.* [83], the air temperature at two meters presents its maximum before the wet period (during the month of May) because of the maximum exposure at the top of the atmosphere and is followed by a maximum equivalent potential temperature (Θ_e) in August because of the maximum of the zenith angle of the sun. Θ_e in the lower troposphere is equivalent to MSE whose transport is due to the circulation of large-scale (the Hadley). But Flaounas *et al.* [84] and Guiavarch *et al.* [85] like other studies have clearly shown the relationship between this surface energy and WAM. Thus, it appears that the relationship between air temperature and the WAM is not a direct link but rather a role on the dynamics of WAM in terms of anomalies reflections in atmospheric circulation like Hadley types.

(c) SST

Several studies have analyzed the relationship between WAM and SSTs. Coëtlogon *et al.* [86] studied the link between WAM and the contrast between SST and the temperature at the coast of Guinea. According to their results, from spring to summer, a band of cold water settles between Ecuador and the coast of Guinea and enhances the temperature gradient at the surface meridian. This causes the acceleration of the WAM which then moves further north. The appearance of this band of cold water is attributed to the process of recovery in deep and cold ocean masses “upwelling”, mostly due to surface winds. Thus, the acceleration of WAM resulted from a positive feedback system since the

acceleration intensifies upwelling increasing itself from WAM that spreads further north. Moreover, Peyrillé and Lafore [87] developed a two-dimensional idealized model to reproduce the monsoon system in West Africa. Their study reveals the importance of SST in the Mediterranean. The lack of moisture transport or transport by zonal eddies above the African continent requires forcing external moisture advection in the Mediterranean to get realistic monsoon in West Africa. Thus, they show that warm SST in the Mediterranean entails strengthening moisture advection in the lower layers. SST of West Mediterranean seems to have a stronger impact on the variability of precipitation in the Gulf of Guinea (Sahel) [88]. Several studies suggested some links between the variability of SST during the season of coastal rain and precipitation, especially through the installation of the equatorial upwelling. Gu and Adler [89] describe, using satellite observations and reanalysis, the seasonal evolution of the tropical Atlantic. They show that, in the Gulf of Guinea, convection is modulated by seasonal forcing of the ocean and the SST gradient meridian. In addition, it was shown that the Pacific Ocean [34], the Atlantic Ocean [90], the Mediterranean [88] as well as the phenomenon El Niño-Southern Oscillation [10] generate atmospheric disturbances and, in this way, affect the African monsoon.

(d) SLP

A climatological analysis by Baldi *et al.* [91] suggests that the West Africa monsoon influence the central-western [92, 93]. Using NCEP/NCAR global reanalysis [91] analyzed in detail the events characterizing summer 2002 over Mediterranean, Europe and North Atlantic, in particular the anomalous SST and Sea Level Pressure (SLP) fields relatively to the mean climate patterns Mediterranean summer, and specifically the SLP (weakly), the temperature and the rainfall. They also found that the overall pattern of the WAM changed in July, when a lower pressure developed from Iceland to central Mediterranean along a northwest to south-east axis, with anomalously high pressures in the south-west and north-east. Moreover, the summer average SLP field was similar to the pattern observed in July. The surface air temperature field over Mediterranean closely follows the sea level pressure patterns in summer (e.g., Maheras and Kutiel, [94]). They also run sensitivity analysis which shows how the SST anomalies can produce quantitatively significant anomalies in the sea level pressure patterns over North Sahel (positive).

(e) RHUM

The interaction between the flows of heat from the surface of ground water and in the atmosphere has been studied by Lafore [95] and Fontaine *et al.* [88] over the period 1979–2001. Four phases of the ITCZ (inter tropical convergence zone) were identified: early March, mid-April, May and late June (establishment of WAM). These phases appear to be sensitive to the relative humidity of last year. The interaction between these phases is as follows: positive anomalies in this humidity entail an increase in the humidity of the atmosphere, and the convergence of humidity flux- and a decrease in surface albedo. Therefore, the net solar radiation is strengthened on the surface but the air temperature in the lower layers decreases. Net radiation on the surface increases as well as the flow of heat to the atmosphere. This process results in the strengthening of MSE in the lower layers and the strengthening of the circulation of WAM. On the other hand, Fontaine *et al.* [83] showed that in the three regions in West Africa (Guinea, 6°N–10°N, Sudan 10°N–15°N, the Sahel 15°N–20°N), the convergence integrated moisture flux in the entire atmospheric column is significantly correlated with the precipitation at different scales. It is interesting that in the Sahel, unlike the rest of West Africa, the relationship between precipitation and soil evaporation (consequently the RUM) in the Sahel is almost linear. Moreover, Broman *et al.* [96] performed A K-means cluster analysis to identify spatially coherent regions of relative humidity variability during the two periods over West African Sahel. They found that correlating the cluster indices with large-scale circulation and SSTs indicates that the land–ocean temperature gradient and the corresponding circulation, tropical Atlantic sea surface temperatures (SSTs), and to a somewhat lesser extent tropical Pacific SSTs all play a role in modulating the timing of the monsoon season relative humidity onset and retreat.

Thus, it is clear that the RUM is a link because it is connected to the mainland that can exert forcing on the atmospheric dynamics through which it is a reflection of abnormalities in the circulation of WAM.

From the above explanation, it is clear that despite that the atmosphere has no inertia the predictors (SLP, AirTemp, and RUM) are connected either to the ocean or continent that can have a forcing on atmospheric dynamics and they are the reflections of anomalies in the atmospheric circulation types Hadley or Walker. Other predictors are related to the Pacific or the Mediterranean to the tropical Atlantic. Table 2 summarizes the atmospheric data and their geographical locations.

3.3. Selection of Predictors and Optimal Lag Time

Monthly CRU precipitation time series cover an area larger than the extent of the Sirba basin. They were initially used as predictand for selecting a pool of potential predictors

having a known relationship with the WAM, and to avoid those with no effect on the monsoon dynamics.

Table 2. Description of atmospheric data.

Parameter	Units	Level	Reference Data	Spatial coverage	Regions of the Predictors	Temporal coverage
Sea level pressures (SLP)	Pa/s	1000 hPa	NCEP 2	2.5° × 2.5° grid 15N–45S, 60W–10E	Atlantic ocean	1979/01/01 to 2013/08/31
Air temperature (AirTemp)	°K	1000 hPa	NCEP 2	2.5° × 2.5° grid 20N–15S, 120E–70W	Pacific ocean	1979/01/01 to 2013/08/31
Meridional wind (VWND)	m/s	1000 hPa	NCEP 2	2.5° × 2.5° grid 90N–90S, 0–180W	Sahel (Easterly jet)	1979/01/01 to 2013/08/31
Zonal wind (UWND)	m/s	1000 hPa	NCEP 2	2.5° × 2.5° grid 30N–25S 10W–10E	Sahel (Easterly jet)	1979/01/01 to 2013/08/31
Relative humidity (RHUM)	%	1000 hPa	NCEP 2	2.5° × 2.5° grid 40N–30N, 20E–35E	Mediterranean basin	1979/01/01 to 2013/08/31
Sea surface temperature (SST)	°C	Surface	NOAA NCDC ERSST version3b	2° × 2° grid 39N–15S, 60W–15E	Atlantic ocean	1854/01/01 to 2013/08/31
Climatic research unit rainfall (CRU)	mm	Surface	CRU	0.5° × 0.5° grid 2°W–2°E, 10°N–15°N		January 1901 to December 2012

These monthly precipitations time series were averaged over the season July–September (JAS), which is the core of the rainy season in the Sahel. As a preliminary test, different time periods 1960–2010, 1970–2010, 1980–2010, 1990–2010, 2000–2010 are considered to check the most favorable periods in terms of signals within the pool of potential predictors. The reason for testing these sub-periods resulted from the previously mentioned studies (Section 2) which showed that the teleconnections of the Sahelian rainfall have evolved since the 60's. Based on these defined periods, 13 groups (with 89 sub-components) of predictors were considered for the preliminary test to check the presence of significant correlations ($R > 0.5$) between each predictor and CRU rainfall.

After this first selection which relies on the CRU precipitations dataset as predictand, predictors were selected using the *in situ* rainfall (from rain gauge stations) as predictand.

This selection is done using the method developed in [4]. This method was employed to link the observed rainfall and each predictor through some statistical techniques. The candidate predictor was aggregated over all possible time windows (where a time window's length in months is an integer) during the 18 months prior to the rainy season onset and each of the obtained time series was used as an explanatory variable in a linear model having the seasonal rainfall on the Sirba watershed as explained variable.

The choice of the period over which the predictor is averaged will impact the performance of the forecast. Since the best period is not known *a priori*, predictor data sets were aggregated over various periods with different lengths and different start dates. The periods were restricted to start at the beginning of a calendar month and finish at the end of a calendar month. The beginning of a period has to be later or equal to January 1st of the previous year (year Y-1, where Y is the year containing the rainy season for which the forecast is issued).

The end of the period must be prior or equal to June 30th of year Y. Figure 2 shows how time windows were systematically generated. The upper bar indicates all months starting from January of the previous year (year Y-1) to June of the year the forecast is issued (year Y). In the first run, for example, only the predictor of January (Y-1) was selected to use as a predictor. Predictor averaged over January-February (Y-1) was used as a predictor in the second run. This process was iterated at one-month increments until June (Y) was reached as the end of the period. The process was repeated until the beginning and the end of the periods were June (Y).

For each time window, a linear model linking the predictor averaged over that time window and seasonal rainfall on the Sirba was built as follows:

- (a) For each year Y that the predictor was available,
 - (i) The predictor of year Y-1 was removed from the predictor grid;
 - (ii) The rainfall of year Y was removed from the rainfall data set;
 - (iii) A coefficient of correlation (R) is used to screen the remaining predictor data: a correlation analysis between the predictor at each grid point and the rainfall was computed and its level of significance (P-value <0.05) was assessed. Once the correlation was not significant, the grid point was discarded. The remaining grid points were then ordered decreasingly;
 - (iv) Afterward, a principal component analysis (PCA) was applied on the retained predictor gridded data from the previous step to reduce the number of predictors;

- (v) Since PCA gave rise to more sets of new predictor data, a stepwise regression (5% confidence interval) was used to keep only grid points with high predictive power;
 - (vi) A linear regression was fitted between the predictors and precipitation time series;
 - (vii) The fitted linear regression was used to simulate the rainfall of year Y. If predictor and rainfall were in the same year (Year Y), only predictor and rainfall time series for that year were removed in the first step.
- (b) Then, the coefficient of determination (R^2), Nash-Sutcliffe coefficient (Nash), and Hit-Rate scores (HIT) were computed to estimate the model's performance.

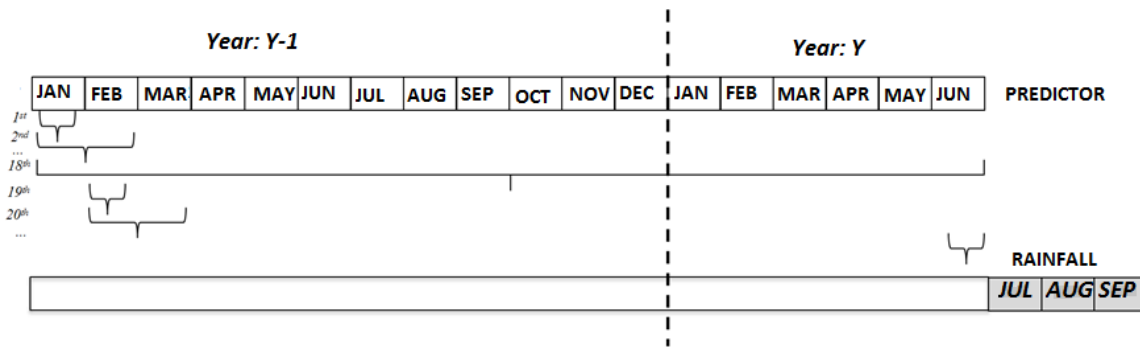


Figure 2. Predictor averaging periods (Adapted from Sittichok *et al.* [4]).

In summary, all predictors were selected according to their physical link with the WAM based on the hypothesis that only predictors having dynamical link with the WAM seem to give good forecast skills. Each of these predictors was screened through simple correlation test to find its link with the CRU rainfall data defined over a region covering more than the Sirba watershed. The main criteria used to find either the predictor should be retained or rejected is the correlation coefficient (R) that has to be greater or equal to 0.5 (*i.e.*, $p\text{-value} < 0.05$). The CRU rainfall was used at this stage for the purpose to have a good assessment of all possible predictors having impact on the WAM. The retained predictors were further screened based on *in situ* rainfall using the approach explained previously and summarized in Figure 2. All these steps were summarized in the new flowchart of Figure 3. It should be noted that the figures provided were based on leave-one out cross validation (LOOV) which was also used in selecting the predictor. The LOOV was used due to the small length of data used.

The first three (3) best predictors in terms of high Nash values were used to test the performance of the two sets of methods (linear approach and non-linear approach).

3.4. Linear Approach

Several linear methods are applied successively for selecting the predictors and developing the seasonal rainfall forecast models. They include: correlation analysis, principal component analysis (PCA), stepwise regression, linear regression, and cross validation.

They were used to perform three successive tasks: predictor selection; predictor dimension reduction and linear regression. The first application discarded meaningless predictors from the original data set using the coefficient of correlation as criteria of selection. At this stage, the correlation coefficient between the predictor at each grid point and the rainfall on the Sirba watershed was calculated and its level of significance (p -value < 0.05) was tested. When the correlation was not significant, the grid point was discarded. The remaining grid points were then decreasingly ordered according to the correlation

p -value. Only the best grid points were included in the analysis. Afterward, PCA was applied on the retained predictors to reduce their number. A forward stepwise regression method (5% confidence interval threshold) was then applied to keep predictors having only a significant predictive power.

It should be noted that a leave-one-out cross validation was used in the model application to avoid the bias which might occur during the development of empirical equations using statistical models.

3.5. Non-Linear Approach

Two non-linear methods were tested for each of the three best predictors selected based on the correlation analysis. The description of these methods and how they are applied is detailed in the next paragraphs. The R^2 , Nash, and HIT were calculated to estimate the model's performance.

3.5.1. Non-Linear Principal Component Analysis

The non-linear principal component analysis (NLPCA) algorithm developed by [97] is generally considered as a non-linear generalization of standard linear PCA and was successfully applied in atmospheric and oceanic sciences [98–102]. The principal components (PCs) are generalized from straight lines to curves, thus the NLPCA helps to extract PCs either linear or not. This could improve seasonal rainfall forecast skills because it is well known that most of atmospheric/climate relationships are not linear as always assumed. Each predictor was first screened using R^2 before being fed into the NLPCA. However, due to the high computational time of NLPCA the number of PCs is

narrowed in considering only the three best PCs in the process. Figure 4 presents the entire process of the NLPCA seasonal forecast model. More details on the way NLPCA model works and its difference with the ordinary PCA can be found in Scholz *et al.* [97].

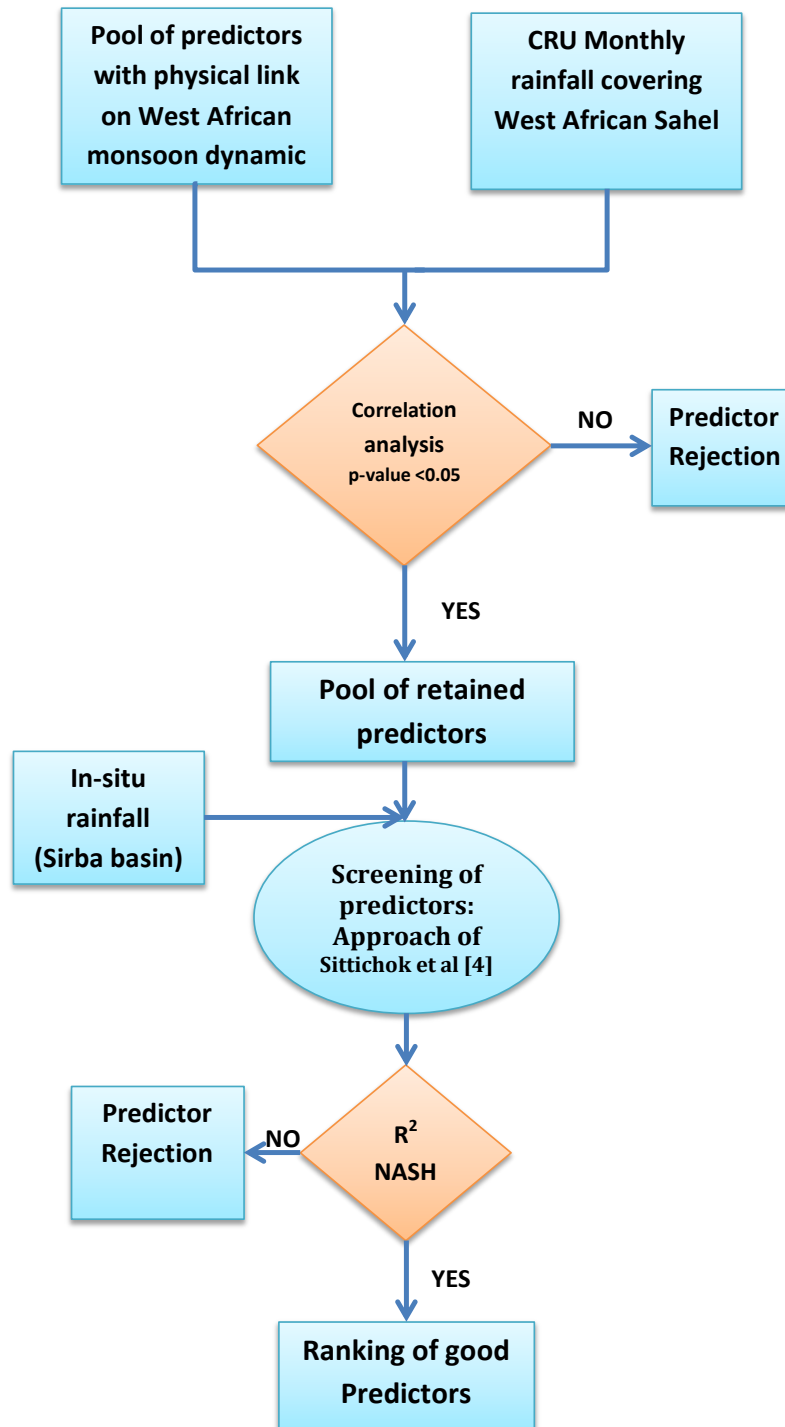


Figure 3. Steps for the selection of predictors.

3.5.2. Feedforward Neural Network

The feedforward neural network (FFNN) is just tested in this work because to show its performance which is somehow poor. For more details on this method, the reader can refer to [103–109].

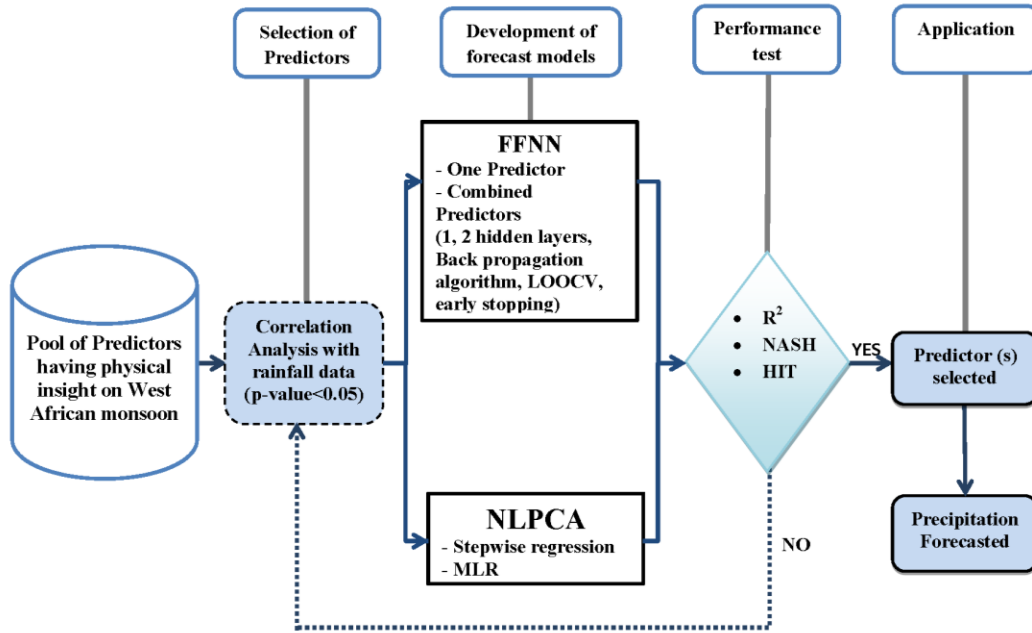


Figure 4. Developed non-linear statistical approach for rainfall forecast.

4. Results and Discussions

4.1. Selected Predictors and Lag Time Period

In the preliminary setting of a pool of predictors, the synchronous correlations between predictors-predictand (*i.e.*, correlations between rainfall and predictors averaged over JAS period) have revealed the presence of the largest correlations for the period 1970–2010. Thus, it was used as the reference study period. For instance, Figure 5 presents the correlation between rainfall and SLP over this time period with a lag time of three months between the two variables.

During the second step dedicated to selecting the predictors, the time window (or season) which yielded the best Nash coefficient (and therefore the optimal lag time) was determined. Tables 3, 4 and 5 summarize the final selected predictors used to forecast seasonal rainfall using combined linear methods, NLPCA and FFNN, respectively. It is obvious that for all models the best predictors according to the forecast skills are AirTemp, RHUM and SLP. In addition, the lag time (eight months lead time on average)

obtained for most predictors is large enough to develop early warning systems for the decision makers and socio-economic actors about the issue of the forthcoming rainy season.

4.2. Seasonal Rainfall Forecast

The performance of the combined linear models showed that AirTemp (from Pacific Tropical North), RHUM (from Mediterranean East) and SLP (from Atlantic tropical South) are the best predictors with respective Nash coefficients of 0.53, 0.52 and 0.46 (see Table 3). They have also the best coefficients of determination (53%, 58%, 48%) and Hit rate scores (67.9%, 64.3%, 60.7%), respectively. While SST (from Atlantic Ocean) obtained 0.34, 43%, and 58.5% as Nash, R^2 and Hit rate scores, respectively. These results from the predictors AirTemp, RHUM, and SLP seem to be better than those obtained

in [4] who used linear methods to forecast seasonal rainfall in the same area based on Pacific and Atlantic SSTs as predictor. They obtained 0.45, 0.38 and 66.67% respectively for R^2 , Nash, and Hit rate score. Figures 6 and 7 present the observed and simulated seasonal rainfall for combined linear models.

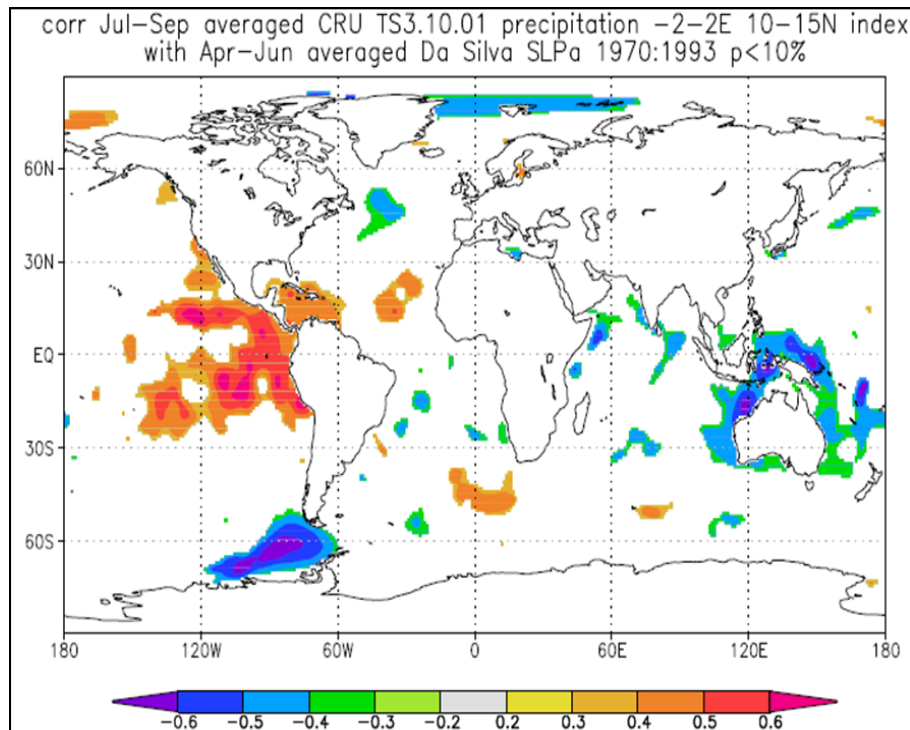


Figure 5. Correlation map between seasonal precipitation and sea level pressure (Da Silva analysis, sea).

Table 3. Combined linear methods for seasonal rainfall forecast.

PREDICTOR	NMAX*	R ²	Nash coef.	HIT Score	Best period M1-M2**	Lag period
Sea Level Pressure (SLP) at 1000hPa	50	0.48	0.46	60.71	17-18	0
Relative Humidity (RHUM) at 1000hPa	80	0.58	0.52	64.29	10-10	8 months
Air Temperature (AirTemp) at 1000hPa	10	0.530	0.527	67.86	1-4	7 months
Meridional Wind (VWND) at 1000hPa	170	0.31	0.28	53.57	5-5	8 months
Zonal Wind (UWND) at 1000hPa	190	0.33	0.324	71.43	11-11	7 months
Sea surface temperature (SST)	30	0.43	0.34	58.54	3-6	12 months

(**) M1=1:12 (January to December); M2=M1:18 (considered month of M1 to the next coming June)

(*) NMAX: number of best grid points retained after screening the predictor grid based on R²

Table 4. Seasonal rainfall forecast model skills using non-linear principal component analysis (NLPCA).

PREDICTOR	R ²	NASH	HIT score	Lag time Period
Sea Level Pressure (SLP) at 1000hPa	0.32	0.31	53.57	9 months
Relative Humidity (RHUM) at 1000hPa	0.36	0.36	53.57	7 months
Air Temperature (AirTemp) at 1000hPa	0.46	0.45	60.71	8 months

Table 5. Feedforward neural network (FFNN) model (single predictor) output for Sirba seasonal rainfall forecast.

Predictors	R ²	Nash	HIT score (%)	Lag time (months)
AirTemp	0.26	0.20	48.24	4
RHUM	0.18	0.10	29.12	4
SLP	0.21	0.09	18.03	2
SST	0.18	0.044	11.49	5

For the NLPCA model, the issued seasonal forecast skills can be judged satisfactory regarding the short study period considered, because non-linear models need longer study periods to over perform the linear ones. Results showed that the predictor AirTemp (R^2 : 0.46; Nash: 0.45; HIT: 60.7%) was the best, and then followed by RHUM and SLP, respectively. It was also found that this method provides a to larger lag time compared to combined non-linear methods despite of its relative low forecast skills. Table 4 presents the model performance and the lag time, while Figure 8 illustrates some of the rainfall forecast obtained from the NLCPA model using, respectively, the predictors AirTemp and RHUM.

Overall, the set of linear models performs better than the non-linear ones. This suggests that there is less benefit using non-linear methods when dealing with small samples, as found in previous studies [27,96,102].

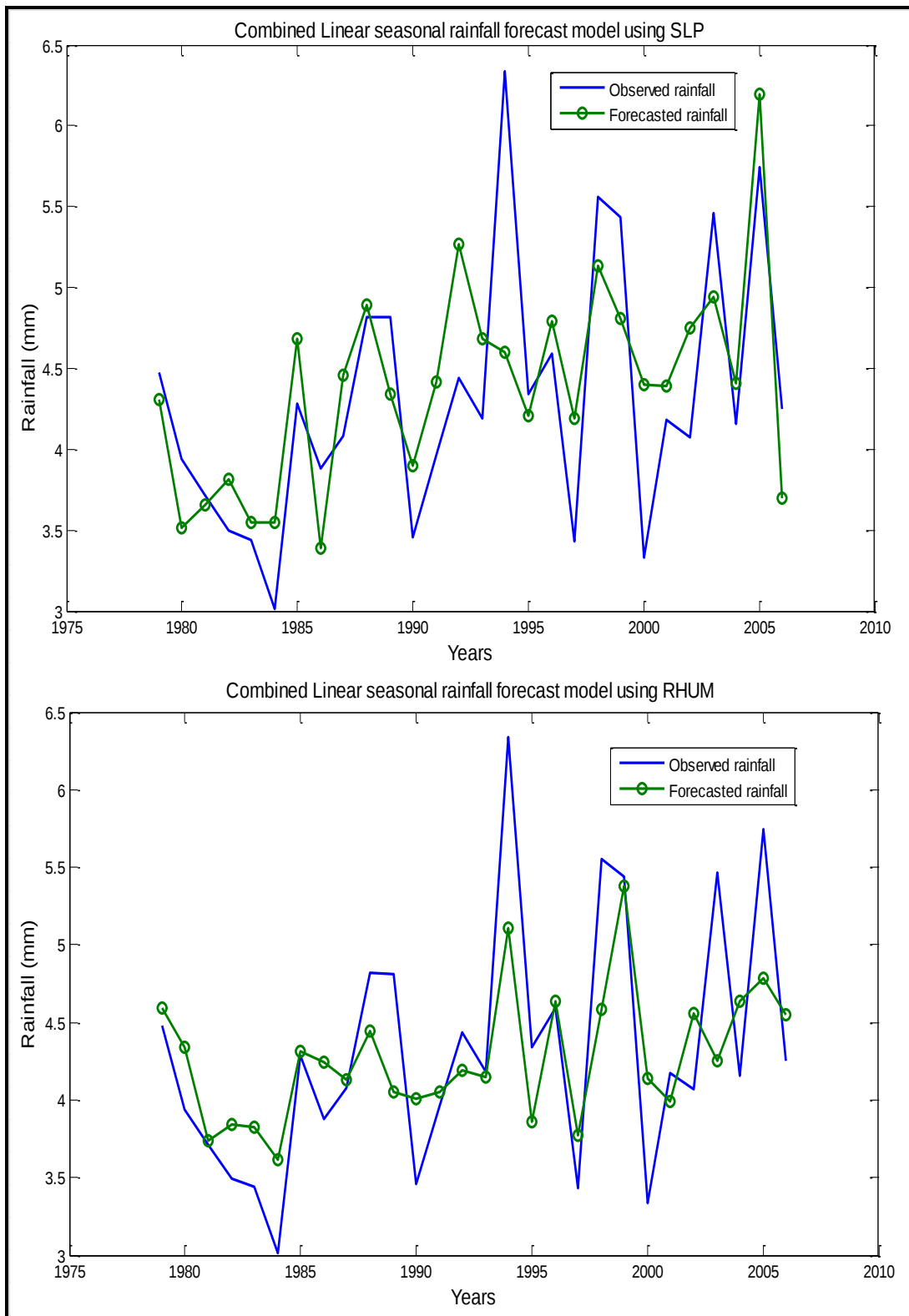


Figure 6. Combined linear model for seasonal rainfall forecast using SLP (upper panel) and RHUM (lower panel).

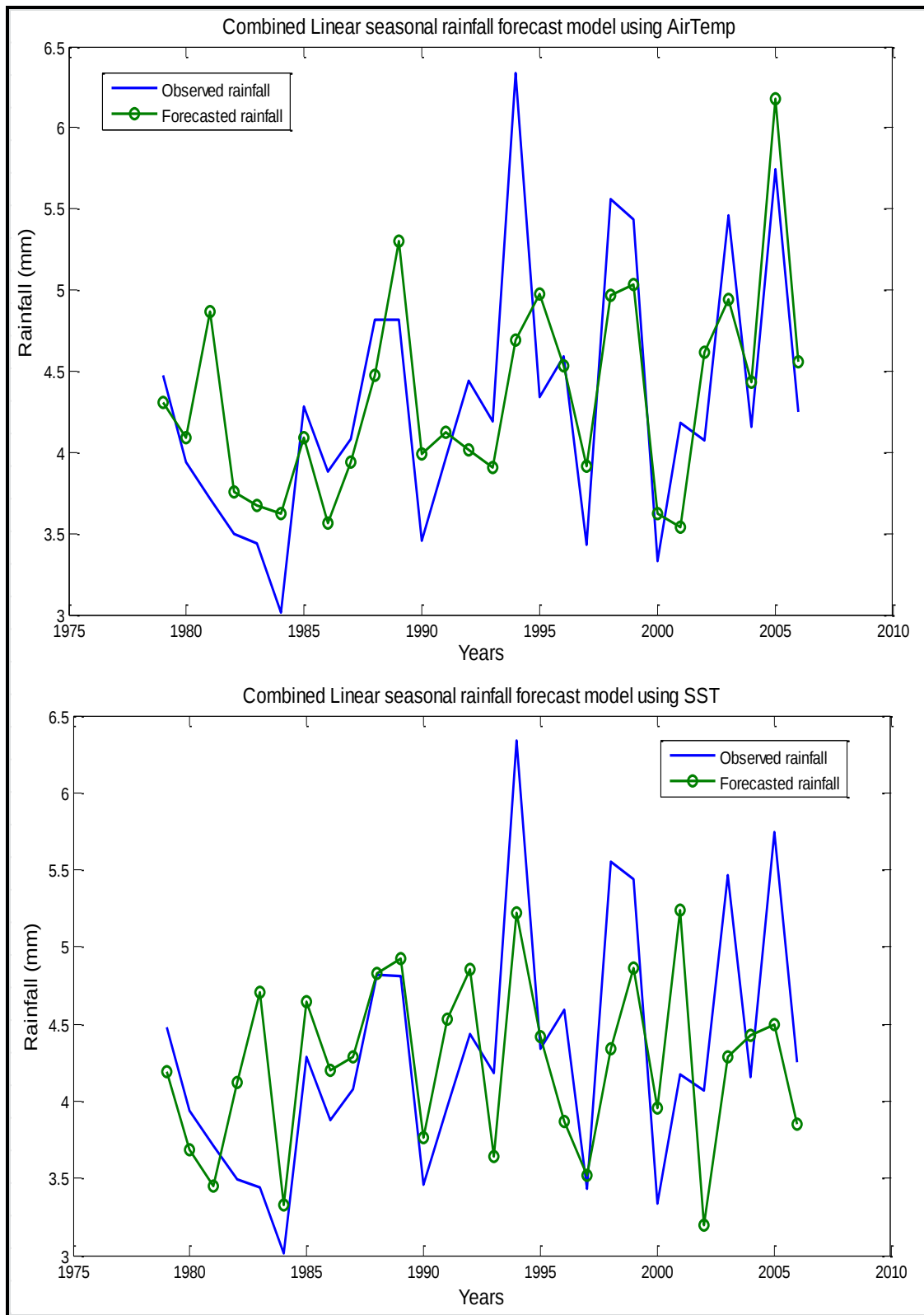


Figure 7. Combined linear model for seasonal rainfall forecast using AirTemp (upper panel) and SST (lower panel).

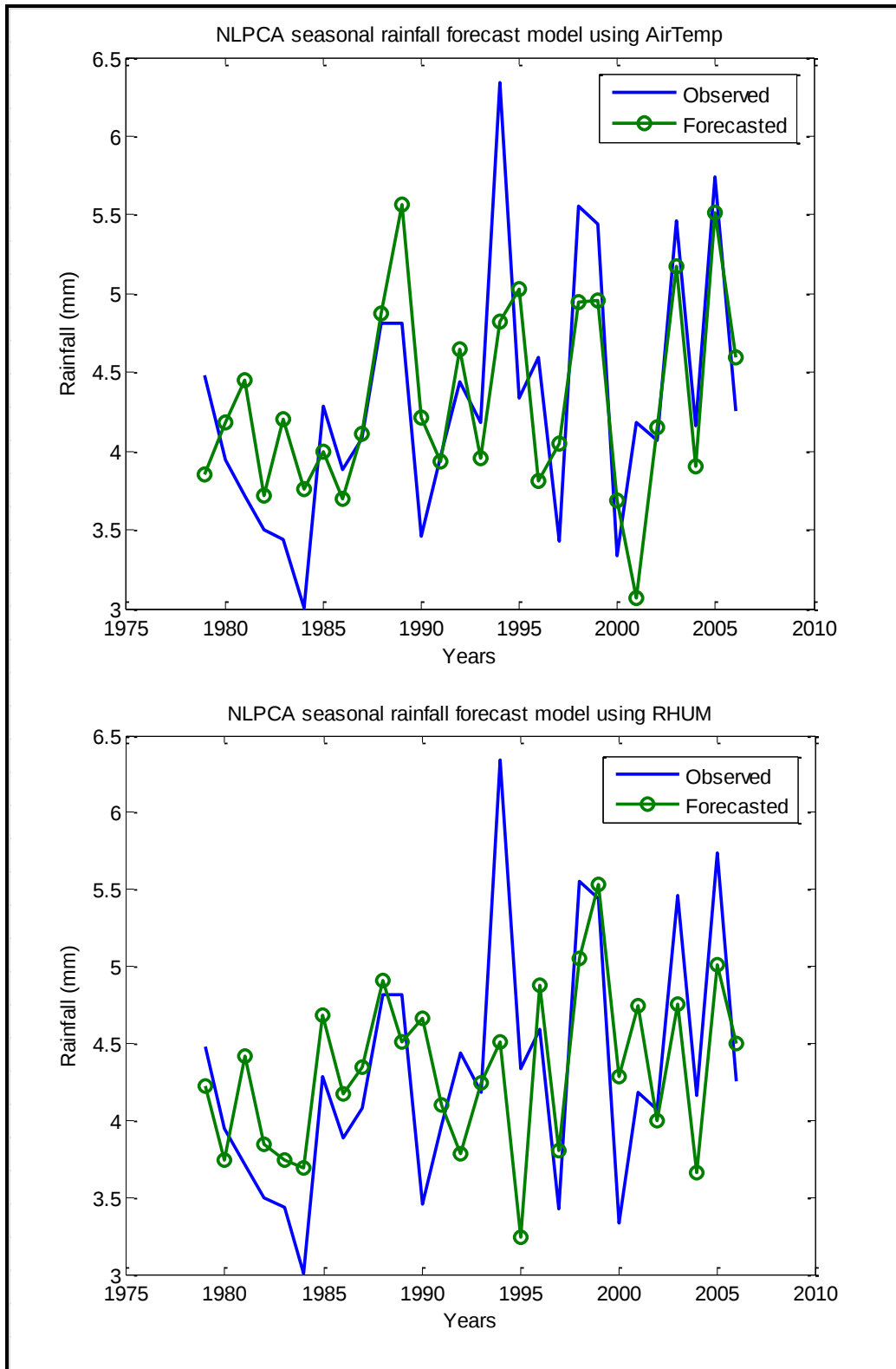


Figure 8. NLPCA seasonal rainfall forecast model using AirTemp (upper panel) and RHUM (lower panel).

5. Conclusions

Two non-linear methods and a combined linear approach were used to forecast JAS (July to September) rainfall on the Sirba watershed, West Africa. Predictors were first screened using a series of steps to isolate those having the highest predictive power. At the end of the process three predictors, air temperature (from Pacific Tropical North), sea level pressure (from Atlantic Tropical South) and relative humidity (from Mediterranean East) were retained and tested as inputs for seasonal rainfall forecasting models. Forecast performances were compared using R^2 , Nash and HIT. Results showed that the combined linear approach performed better than the non-linear models. The best forecasts were obtained using air temperature as predictor ($R^2 = 53\%$; Nash = 0.53; HIT = 67.9%; Lead-time = 7 months). The next best model uses relative humidity as predictor ($R^2 = 58\%$; Nash = 0.52; HIT = 64.3%, Lead-time = eight months). Nonlinear Principal Component Analysis (NLPCA) was the best non-linear method while FFNN performed poorly. These new predictors found in this study could lead to better forecasts of seasonal rainfall over West Africa, an issue which has challenged forecasters over many years. This paper also helped understanding that non-linear methods could also be used instead of the usual linear methods. The specificity of this work is the use of other predictors rather than the SST which gave acceptable results than the SST. However, the limit of the approach resides on the small length of data used. Therefore, to generalize the results for other scientists, it is recommended that during the forecast a Sahelian global index must be constructed using CRU data while examining its correlation with the index of the watershed and the skill of models. As a future step of this work, a multi-model approach will be used to compare the resulting skills to that of the best model.

Acknowledgements

The authors would like to thank the International Research Initiative on Adaptation to Climate Change (IRIACC) program through the International Development Research Center (Canada) for funding this research.

Author Contributions

Abdouramane Gado Djibo developed the models, performed analyses and wrote the paper. Nathalie Philippon, Ousmane Seidou, Harouna Karambiri, Hadiza Moussa Saley, Ketvara Sittichok and Jean Emmanuel Paturel contributed to analysis and interpretation of results. Nathalie Philippon, Ousmane Seidou, and Ketvara Sittichok proofread the manuscript and contributed to answer reviewers' comments.

Conflicts of Interest

The authors declare no conflict of interest.

References

- Gado Djibo, A.; Seidou, O.; Karambiri, H.; Sittichock, K.; Paturel, J.E.; Saley, H.M. Development and assessment of non-linear and non-stationary seasonal rainfall forecast models for the Sirba watershed, West Africa. *J. Hydrol. Reg. Stud.* **2015**, *4*, 134–152.
- Sarr, M.A.; Gachon, P.; Seidou, O.; Bryant, C.R.; Ndione, J.A.; Comby, J. Inconsistent linear trends in Senegalese rainfall data. *Hydrol. Sci. J.* **2014**, doi:10.1080/02626667.2014.926364.
- Sarr, M.A.; Zoromé, M.; Seidou, O.; Bryant, C.R.; Gachon, P. Recent trends in selected extreme precipitation indices in Senegal-A changepoint approach. *J. Hydrol.* **2013**, *505*, 326–334.
- Sittichok, K.; Gado Djibo, A.; Seidou, O.; Saley, H.M.; Karambiri, H.; Paturel, J. Statistical seasonal rainfall and streamflow forecasting for the Sirba watershed, using sea surface temperature. *Hydrol. Sci. J.* **2014**, doi:10.1080/02626667.2014.944526.
- Samimi, C.; Fink, A.H.; Paeth, H. The 2007 flood in the Sahel: causes, characteristics and its presentation in the media and FEWS NET. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 313–325.
- Brooks, N. *Drought in the African Sahel: Long term perspectives and future prospects*; Tyndall Centre Working Paper; University of East Anglia: Norwich, UK, 2004.
- Amadou, A.; Gado Djibo, A.; Seidou, O.; Sittichok, K.; Seidou Sanda, I. Changes to flow regime on the Niger River at Koulikoro under a changing climate. *J. Hydrol. Sci.* **2014**, doi: 10.1080/02626667.2014.916407.
- Ibrahim, B.; Karambiri, H.; Polcher, J.; Yacouba, H.; Ribstein, P. Changes in rainfall regime over Burkina Faso under the climate change conditions simulated by 5 regional climate models. *Clim. Dyn.* **2014**, *42*, 1363–1381.
- Le Barbe, L.; Lebel, T.; Tapsoba, D. Rainfall variability in West Africa during the years 1950–90. *J. Clim.* **2002**, *15*, 187–202.
- Janicot, S.; Trzaska, S.; Poccarr, I. Summer Sahel-ENSO teleconnection and decadal time scale SST variations. *Clim. Dyn.* **2001**, *18*, 303–320.
- Fontaine, B.; Janicot, S. Sea surface temperature fields associated with West African rainfall anomaly types. *J. Clim.* **1996**, *9*, 2935–2940.
- Folland, C.K.; Palmer, T.N.; Parker, D.E. Sahel rainfall and worldwide sea temperature 1901–1985. *Nature* **1986**, *320*, 602–607.

- Hastenrath, S. Decadal scale changes of the circulation in the tropical Atlantic sector associated with Sahel drought. *Int. J. Climatol.* **1990**, *20*, 459–472.
- Lamb, P.J. West African water variations between recent contrasting Subsaharan droughts. *Tellus* **1983**, *35*, 198–212.
- Lamb, P.J.; Pepler, R.A. Further case studies of tropical Atlantic surface atmospheric and oceanic patterns associated with sub-Saharan drought. *J. Clim.* **1992**, *5*, 476–488.
- Nicholson, S.E. Rainfall and atmospheric circulation during drought periods and wetter years in West Africa. *Mon. Wea. Rev.* **1981**, *109*, 2191–2208.
- Nicholson, S.E. Land surface processes and Sahel climate. *Rev. Geophys.* **2000**, *38*, 117–139.
- Nicholson, S.E. The West African Sahel: A review of recent studies on the rainfall regime and its interannual variability. *ISRN Meteor.* **2013**, *2013*, 453–521.
- Kwon, H.H.; Brown, C.; Xu, K.; Lall, U. Seasonal and annual maximum streamflow forecasting using climate information: application to the Three Gorges Dam in the Yangtze River basin, China. *Hydrol. Sci. J.* **2009**, *54*, 582–595.
- Liu, Z.; Alexander, M. Atmospheric bridge, oceanic tunnel, and global climatic teleconnections. *Rev. Geograph.* **2007**, *45*, 1–34.
- Gaetani, M.; Fontaine, B. Interaction between the West African Monsoon and the summer Mediterranean climate: An overview. *Fisica de la Tierra* **2013**, *25*, 41–55.
- Jung, T.; Ferranti, L.; Tompkins A.M. Response to the summer 2003 Mediterranean SST anomalies over Europe and Africa. *J. Climate*, **2006**, *19*, 5439–5454..
- Camberlin, P.; Janicot, S.; Poccarrd, I. Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature: Atlantic VS. ENSO. *Int. J. Clim.* **2001**, *21*, 973–1005.
- Rowell, D.P. Teleconnections between the tropical Pacific and the Sahel. *Q. J. R. Meteorol. Soc.* **2001**, *127*, 1683–1706.
- Chase, T.N.; Pielke S.R.; Avissar, R. Teleconnections in the earth system. Encyclopedia of Hydrological Sciences. Available online: <http://onlinelibrary.wiley.com/doi/10.1002/0470848944.hsa190/pdf> (accessed on 12 August 2014).
- Ndiaye, O.; Ward, M.N.; Thiaw, W.M. Predictability of seasonal Sahel rainfall using GCMs and lead-time improvement through the use of a couple model. *J. Clim.* **2011**, *24*, 1931–1949.
- Bouali, L. Prévisibilité et Prévision Statistico-Dynamique des Saisons des Pluies Associées à la Mousson Ouest Africaine à Partir d'Ensembles Multi-modèles. Ph.D. Thesis, Université de Bourgogne, Bourgogne, France, 2009.

- Garbrecht, J.D.; Schneider, J.M.; Van Liew, M.W. Monthly runoff predictions based on rainfall forecasts in a small Oklahoma Watershed. *J. Am. Water Resour. Assoc.* **2007**, *42*, 1285–1295.
- Tuteja, N.K.; Shin D.; Laugesen, R.; Khan, U.; Shao, Q.; Wang, E.; Li, M.; Zheng, H.; Kuczera, G.; Kavetski, D.; *et al.* *Experiment evaluation of the dynamic seasonal streamflow forecasting approach*; Bureau of Meteorology: Melbourne, Australia, 2011.
- Chiew, F.H.S.; McMahon, T.A. Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability. *Hydrol. Sci. J.* **2002**, *47*, 505–522.
- Ndiaye, O.; Goddard, L.; Ward, M.N. Using regional wind fields to improve general circulation model forecasts of July–September Sahel rainfall. *Int. J. Clim.* **2009**, *29*, 1262–1275.
- Wang, E.; Zhang, Y.; Luo, J.; Chiew, F.; Wang, Q.J. Monthly and seasonal streamflow forecasts using rainfall-runoff modeling and historical weather data. *Water Resour. Res.* **2011**, *47*, 1–13.
- Yossef, N.C.; Winsemius, H.; Weerts, A.; Beek, R.V.; Bierkens, F.P. Skill of global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. *Water Resour. Res.* **2013**, *49*, 4687–4699.
- Mohino, E.; Rodriguez-Fonseca, B.; Mechoso, C.R.; Gervois, S.; Ruti, P.; Chauvin, F. Impacts of the tropical Pacific/Indian Oceans on the seasonal cycle of the West African monsoon. *J. Clim.* **2011**, *24*, 3878–3891.
- Rodriguez-Fonseca, B.; Janicot, S.; Mohino, E.; Losada, T.; Bader, J.; Caminade, C.; Chauvin, F.; Fontaine, B.; Garcia-Serrano, J.; Gervois, S.; *et al.* Interannual and decadal SST-forced responses of the West African monsoon. *Atmos. Sci. Lett.* **2011**, *12*, 67–74.
- Singh, O.P. Cause-effect relationships between sea surface temperature, precipitation and sea level along the Bangladesh coast. *Theor. Appl. Clim.* **2001**, *68*, 233–243.
- Garcia-Serrano, J.; Doblas-Reyes, F.J.; Haarsma, R.J.; Polo, I. Decadal prediction of the dominant West African monsoon rainfall modes. *J. Geophys. Res. Atmos.* **2013**, *118*, 5260–5279.
- Philippon, N.; Doblas-Reyes, F.J.; Ruti, P.M. Skill, reproducibility and potential predictability of the West African monsoon in coupled GCMs. *Clim. Dyn.* **2010**, *35*, 53–74.
- Rodrigues L.R.L.; Garcia-Serrano, J.; Doblas-Reyes, F. Seasonal prediction of the intraseasonal variability of the West African monsoon precipitation. *Fisica de la Tierra* **2013**, *25*, 73–87.
- Sultan, B.; Janicot, S.; Diedhiou, A. The West African monsoon dynamics. Part I: Documentation of intraseasonal variability. *J. Clim.* **2003**, *16*, 3389–3406.

- Marteau, R. Cohérence spatiale et prévisibilité potentielle des descripteurs intrasaisonniers de la saison des pluies en Afrique Soudano-Sahélienne : Application à la culture du mil dans la région de Niamey. Ph.D. Thesis, Université de Bourgogne, Bourgogne, France, 2010.
- Garric, G.; Douville, H.; Déqué, M. Prospects for improved seasonal predictions of monsoon precipitation over Sahel. *Int J. Climatol.* **2002**, *22*, 331–345.
- Tippet, M.K.; Giannini, A. Potentially predictable components of African summer rainfall in an SST-forced GCM simulation. *J. Clim.* **2006**, *19*, 3133–3144.
- Batté, L.; Déqué, M. Seasonal predictions of precipitation over Africa using coupled ocean-atmosphere general circulation models: Skill of the ENSEMBLES project multimodel ensemble forecasts. *Tellus A* **2011**, *63*, 283–299.
- Olivry, J.C. Evolution récente des régimes hydrologiques en Afrique Intertropicale. L'Eau la terre et les hommes, Presses Universitaires de Nancy : Nancy, France, 1993.
- Rowell, D.P.; Folland, C.K.; Maskell, K.; Ward, M.N. Variability of summer rainfall over Tropical North Africa (1906–1992): Observations and modelling. *Q. J. R. Meteorol. Soc.* **1995**, *121*, 669–704.
- Solomon, A.; Goddard, L.; **Kumar, A.; Carton, J.; Deser, C.; Fukumori, I.; Greene, A.M.; Hegerl, G.; Kirtman, B.; Kushnir, Y.; et al.** Distinguishing the roles of natural and anthropogenically forced decadal climate variability. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 141–156.
- Gaetani, M.; Mohino, E. Decadal prediction of the Sahelian precipitation in CMIP5 simulations. *J. Clim.* **2013**, *26*, 7708–7719.
- Hastenrath, S. Interannual variability and annual cycle: Mechanisms of circulation and climate in the Tropical Atlantic Sector. *Mon. Weather. Rev.* **1984**, *112*, 1097–1107.
- Druyan, L.M. The sensitivity of sub-Saharan precipitation to Atlantic SST. *Clim. Chang.* **1991**, *18*, 17–36.
- Lamb, P.J.; Pepler, R.A. West Africa. In *Teleconnections Linking Worldwide Climate Anomalies*, Glantz, M.H., Katz, R.W., Nicholls, N., Eds.; Cambridge University Press: Cambridge, England, 1991; pp. 121–189.
- Janicot, S. Spatio-temporal variability of West African rainfall. Part II: associated surface and air mass characteristics. *J. Clim.* **1992**, *5*, 499–511.
- Janicot, S.; Moron, V.; Fontaine, B. Sahel droughts and ENSO dynamics. *Geophys. Res. Lett.* **1996**, *23*, 515–518.
- Moron, V. Guinean and Sahelian rainfall anomaly indices at annual and monthly scales (1933–1990). *Int. J. Clim.* **1994**, *14*, 325–341.

- Janicot, S.; Harzallah, A.; Fontaine, B.; Moron, V. West African monsoon dynamics and Eastern Equatorial Atlantic and Pacific SST anomalies (1970–88). *J. Clim.* **1998**, *11*, 1874–1882.
- Rowell, D.P. The impact of Mediterranean SSTs on the Sahelian rainfall season. *J. Clim.* **2003**, *16*, 849–862.
- Gaetani, M.; Fontaine, B.; Roucou, P.; Baldi, M. Influence of the Mediterranean Sea on the West African monsoon: intraseasonal variability in numerical simulations. *J. Geophys. Res.* **2010**, *115*, doi: 10.1029/2010JD014436.
- Polo, I.; Ullmann, A.; Roucou, P.; Fontaine, B. Weather regimes in the Euro-Atlantic and Mediterranean sector and relationship with West African rainfall over the period 1989–2008 from a self-organizing maps approach. *J. Clim.* **2011**, *24*, 3423–3432.
- Shaman, J.; Tziperman, E. An atmospheric teleconnection linking ENSO and Southwestern European precipitation. *J. Clim.* **2011**, *24*, 124–139.
- Lopez-Parages, J.; Rodriguez-Fonseca, B. Multidecadal modulation of El Niño influence on the Euro-Mediterranean rainfall. *Geophys. Res. Lett.* **2012**, *39*, doi: 10.1029/2011GL050049.
- Webster, P.J.; Magana, V.O.; Palmer, T.N.; Shukla, J.; Tomas, R.A.; Yanai, M.; Yasunari, T. The monsoon: Processes, predictability and prediction. *J. Geophys. Res.* **1998**, *103*, 14451–14510.
- Eltahir, E.A.B. Role of vegetation in sustaining large-scale atmospheric circulations in the tropics. *J. Geophys. Res.* **1996**, *101*, 4255–4268.
- Philippon, N.; Fontaine, B. The relationship between the Sahelian and previous 2nd Guinean rainy seasons: A monsoon regulation by soil wetness. *Ann. Geophys.* **2002**, *20*, 575–582.
- Hall, N.M.J.; Peyrillé, P. Dynamics of the West African Monsoon. *J. Phys. IV France* **2006**, *139*, 81–99.
- Zheng, X.; Eltahir, E.A.B. The role of vegetation in the dynamics of West African monsoons. *J. Clim.* **1998**, *11*, 2078–2096.
- Wang, G.; Eltahir, E.A.B. Role of vegetation dynamics in enhancing the low frequency variability of the Sahel rainfall. *Water Resour.* **2000**, *36*, 1013–1021.
- Zeng, N.; Neelin, J.D.; Lau, K.M.; Tucker, C.J. Enhancement of interdecadal climate variability in the Sahel by vegetation interaction. *Science* **1999**, *286*, 1537–1540.
- Cook, K.H. Generation of the African easterly jet and its role in determining West African precipitation. *J. Clim.* **1999**, *12*, 1165–1184.
- Douville, H.; Chauvin, F.; Broqua, H. Influence of soil moisture on the Asian and African monsoons. Part I: Mean monsoon and daily precipitation. *J. Clim.* **2001**, *14*, 2381–2403.

- Douville, H. Influence of soil moisture on the Asian and African monsoons. Part II: Interannual variability. *J. Clim.* **2002**, *15*, 701–720.
- Douville, H. Assessing the influence of soil moisture on seasonal climate variability with AGCMs. *J. hydrometeorol.* **2003**, *4*, 1044–1066.
- Mara, F. Développement et analyse des critères de vulnérabilité des populations sahéliennes face à la variabilité du climat: le cas de la ressource en eau dans la vallée de la Sirba au Burkina Faso. Ph.D. Thesis. Université du Québec à Montréal, Montréal, QC, Canada, 2010.
- Taweye, A. *Contribution à l'étude hydrologique du bassin versant de la Sirba à Garbé-Kourou*. Centre Régional AGRHYMET : Ouagadougou, Burkina Faso, 1995.
- Ball, J.E.; Luk, K.C. Modeling spatial variability of rainfall over a catchment. *J. Hydrol. Eng.* **1998**, *3*, 122–130.
- British Atmospheric Data Centre (BADC), <http://www.cru.uea.ac.uk/cru/data/hrg/>. (Consulted on 08/12/2013)
- International Research Institute for Climate and Society (IRI) data library. <http://iridl.ldeo.columbia.edu>. (consulted on 18/11/2013)
- Gallée, H.; Moufouma-Okia, W.; Bechtold, P.; Brasseur, O.; Dupays, I.; Marbaix, P.; Messenger, C.; Ramel, R.; Lebel, T. A high-resolution simulation of a West African rainy season using a regional climate model. *J. Geophys. Res.* **2004**, *109*, doi: 10.1029/2003JD004020.
- Thorncroft, C.; Blackburn, M. Maintenance of the African easterly jet. *Q. J. R. Meteorol. Soc.* **1999**, *125*, 763–786.
- Grist, J.P.; Nicholson, S. A Study of the Dynamic Factors Influencing the Rainfall Variability in the West African Sahel. *J. Clim.* **2001**, *14*, 1337–1359.
- Janicot, S.; Mounier, F.; Hall, N.M.J.; Leroux, S.; Sultan, B.; Kiladis, G.N. Dynamics of the West African monsoon. Part IV: Analysis of 25-90-Day variability of convection and the role of the Indian monsoon. *J. Clim.* **2009**, *22*, 1541–1565.
- Sultan, B. Etude de la Mise en Place de la Mousson en Afrique de l'Ouest et de la Variabilité Intra Saisonnière de la Convection. Applications à la Sensibilité des Rendements Agricoles. Ph.D. Thesis, Université Denis Diderot, Paris, France, 2002.
- Nicholson, S.; Grist, J. The seasonal evolution of the atmospheric circulation over West Africa and equatorial Africa. *J. Clim.* **2003**, *16*, 1013–1030.
- Fontaine, B.; Louvet, S.; Roucou, P. Definition and predictability of an OLR based West African monsoon onset. *Int. J. Climatol.* **2008**, *28*, 1097–1088.
- Flaounas, E.; Janicot, S.; Bastin, S.; Roca, R.; Mohino, E. The role of the Indian monsoon onset in the West African monsoon onset: observations and AGCM nudged simulations. *Clim. Dyn.* **2012**, *38*, 965–983.

- Guiavarch, C.; Tréguier, A.M.; Vangriesheim, A. Deep currents in the Gulf of Guinea: Along slope propagation of intraseasonal waves. *Ocean Sci.* **2009**, *5*, 141–153.
- Coëtlogon, G.; Janicot, S.; Lazar, A. Intraseasonal variability of the ocean – atmosphere coupling in the Gulf of Guinea during boreal spring and summer. *Q. J. R. Meteorol. Soc.* **2010**, *136*, 426–441.
- Peyrillé, P.; Lafore, J.P. An idealized Two-Dimensional framework to study the West African monsoon. Part II: Large-scale advection and the diurnal cycle. *J. Atmos. Sci.*, **2007**, *64*, 2783–2803.
- Fontaine, B.; Garcia-Serrano, J.; Roucou, P.; Rodriguez-Fonseca, B.; Losada, T.; Chauvin, F.; Gervois, S.; Sijikumar, S.; Ruti, P.; Janicot, S. Impacts of warm and cold situations in the Mediterranean basins on the West African monsoon: observed connection patterns (1979–2006) and climate simulations. *Clim. Dyn.* **2010**, *35*, 95–114.
- Gu, G.; Adler, R. Seasonal evolution and variability associated with the West African monsoon system. *J. Clim.*, **2004**, *17*, 3364–3377.
- Giannini, A.; Saravanan, R.; Chang, P. Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. *Science* **2003**, *302*, 1027–1030.
- Baldi, M.; Dalu, G.; Maracchi, G.; Pasqui, M.; Cesarone, F. Heat waves in the Mediterranean: A local feature or a larger-scale effect? *Int. J. Climatol.* **2006**, *26*, 1477–1487.
- Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; *et al.* The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Am. Met. Soc.* **1996**, *77*, 437–471.
- Kistler, R.; Kalnay, E.; Collins, W.; Saha, S.; White, G.; Woollen, J.; Chelliah, M.; Ebisuzaki, W.; Kanamitsu, M.; Kousky, V.; *et al.* The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Am. Met. Soc.* **2001**, *82*, 247–268.
- Maheras, P.; Kutiel, H. Spatial and temporal variations in the temperature regime in the Mediterranean and their relationship with circulation during the last century. *Int. J. Climatol.* **1999**, *19*, 745–764.
- Lafore, J.; Flamant, C.; Giraud, V.; Guichard, F.; Knippertz, P.; Mahfouf, J.; Mascart, P.; Williams, E. Introduction to the AMMA special issue on SAdvances in understanding atmospheric processes over West Africa through the AMMA field campaign’. *Q. J. R. Meteorol. Soc.* **2010**, *136*, 2–7.
- Broman, D.; Rajagopalan, B.; Hopson, T. Spatiotemporal variability and predictability of Relative Humidity over West African Monsoon Region. *J. Clim.* **2014**, *27*, 5346–5363.

- Scholz, M.; Martin, F.; Joachim, S. Nonlinear principal component analysis: neural network models and applications. In *principal manifolds for Data visualization and dimension reduction*; Gorban, A.N., Kegl, B., Wunsch, D.C., Zinovyev, A., Eds; Springer: Berlin, Germany, 2007; pp. 44–67.
- Wu, A.; Hsieh, W.W. The nonlinear Northern Hemisphere atmospheric response to ENSO. *Geophys. Res. Lett.* **2004**, *31*, doi: 10.1029/2003GL018885.
- Hsieh, W.W. Nonlinear multivariate and time series analysis by neural network methods. *Rev. Geophys.* **2004**, *42*, doi: 10.1029/2002RG000112.
- Monahan, A.H.; Fyfe, J.C.; Pandolfo, L. The vertical structure of winter time climate regimes of the northern hemisphere extratropical atmosphere. *J. Clim.* **2003**, *16*, 2005–2021.
- Diamantaras, K.; Kung, S. *Principal Component Neural Networks*. Wiley: Hoboken, NJ, USA, 1996.
- Jolliffe, I.T.; Stephenson, D.B. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 2nd edition; John Wiley & Sons: Hoboken, NJ, USA, 2012.
- Tangang, F.T.; Hsieh, W.W.; Tang, B. Forecasting regional sea surface temperatures in the tropical Pacific by neural network models, with wind stress and sea level pressure as predictors. *J. Geogr. Res.* **1998**, *103*, 7511–7522.
- Canon, A.J.; Mckendry, I. Forecasting all-India summer monsoon rainfall using regional circulation principal components: A comparison between neural network and multiple regression models. *Int. J. Clim.* **1999**, *19*, 1561–1578.
- Badr, H.S.; Zaitchik, B.F.; Guikema, S.D. Application of statistical models to the prediction of seasonal rainfall anomalies over the Sahel. *J. Appl. Meteor. Climatol.* **2014**, *53*, 614–636.
- El-Shafie, A.H.; El-Shafie, A.; El-Magzoghi, H.G.; Shehata, A.; Taha, M.R. Artificial neural network technique for rainfall forecasting applied to Alexandria, Egypt. *Int. J. Phys. Sci.* **2011**, *6*, 1306–1316.
- Ansari, H. Forecasting Seasonal and Annual Rainfall Based on Nonlinear Modeling with Gamma Test in North of Iran. *Int. J. Eng. Pract. Res.* **2013**, *2*, 16–29.
- Hung, N.Q.; Babel, M.S.; Weesakul, S.; Tripathi, N.K. An artificial neural network for rainfall forecasting in Bangkok, Thailand. *Hydrol. Earth Sys. Sci. Discuss.* **2008**, *5*, 183–218.
- Philippon, N. Une nouvelle approche pour la prévision statistique des précipitations saisonnières en Afrique de l'Ouest et de l'Est : méthodes, diagnostics (1968–1998) et applications (2000–2001). Ph.D. Thesis, CRC—Université de Bourgogne, Dijon, France, 2002.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

5.3. Seasonal Rainfall Forecasting with Changing and Constant Parameters

The different seasonal rainfall forecasting models developed and the achieved results were published in Gado *et al.* (2015b) in section 5.4. However, a summary is provided to make the entire approach more comprehensive.

Rainfall forecasting models often arbitrarily assume that rainfall is linked to predictors by a multiple linear regression, with parameters that are independent of time and predictor magnitude. Two probabilistic methods, based on change point detection that allows the relationship to change according to the amount of time or rainfall, were developed in this work using normalized Bayes factors. Each method uses one of the following predictors: SLP, AirTemp or RHUM. Method 1 (M1) allows for changes in model parameters according to annual rainfall magnitude, while Method 2 (M2) allows for changes in model parameters with time. M1 and M2 were compared to the classical linear model with constant parameters (M3) and to the model with climatology (M4). The Bayesian approach is used here to select the best seasonal rainfall forecasting model from the developed models with changing parameters and those with constant parameters. In the Bayesian model selection, the posterior and prior probabilities of the models and observed data were first computed (see Equation (5.1)).

$$Posterior Prob_{model} = \frac{(Prior Prob_{model} \times Likelihood_{model})}{Prob_{observations}} \quad (5.1)$$

where $Likelihood_{model} = \prod_{i=1}^n likelihood_i$, i is the year of the forecasted rainfall, and n is the number of years.

All twelve models were considered when applying Equation (5.1). The ratio of a model's posterior probability to that of the observations constitutes a comparative criterion after normalization. For this purpose, the normalized Bayes factor B_f (see Equation (5.2)) was used for all competing models in order to provide a weighted comparison of the likelihood of each model given the observed data. B_f compares the posterior likelihood of the data d of a given model M_i to that of the reference model M_r . For details on the Bayes factor, refer to Min *et al.* (2007).

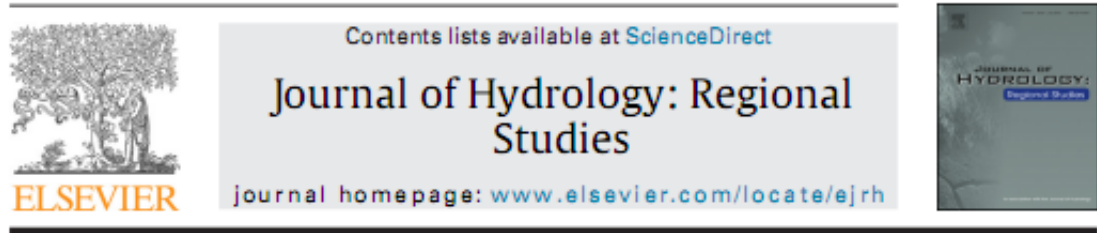
$$B_f = \frac{Likelihood(d/M_i)}{Likelihood(d/M_r)} \quad (5.2)$$

Selection of the best seasonal rainfall forecasting model required analyzing the Bayes factors for all 12 models, with the model having a consistently favorable B_f being the best. In addition, the evolution of likelihoods of forecasted seasonal precipitation (JAS) for each model was checked with graphical representation. The graph shows the likelihood of each forecasted rainfall value on a colored scale where red and blue tend toward a probability of 1 and 0 respectively.

The Bayesian change point detection method was based on a multiple change point detection algorithm. This method was adopted because it can handle an unknown number of changes and it displays the complete probability distribution of the dates of the changes. It can also evaluate abrupt changes in the relationship between the principal and the number of relevant explanatory variables. In these cases, the estimated trend for each segment of the time series is performed based on proxies. A full description of the model algorithm is given in Section 5.4.

5.4. Development and Assessment of Non-Linear and Non-Stationary Seasonal Rainfall Forecast Models for the Sirba Watershed, West Africa

Journal of Hydrology: Regional Studies 4 (2015) 134–152



DEVELOPMENT AND ASSESSMENT OF NON-LINEAR AND NON-STATIONARY SEASONAL RAINFALL FORECAST MODELS FOR THE SIRBA WATERSHED, WEST AFRICA

GADO DJIBO Abdouramane^{a,b*}, SEIDOU Ousmane^b, KARAMBIRI Harouna^a, SITTICHOK Ketevera^b, PATUREL Jean Emmanuel^c, MOUSSA SALEY Hadiza^d

^aInternational Institute for Water and Environmental Engineering (2iE), 01 BP 594 Ouagadougou, Burkina Faso.

^bDepartment of Civil Engineering, University of Ottawa, ON, Canada.

^cInstitut de Recherche pour le Développement (IRD), Abidjan, Côte d'Ivoire.

^dCentre Africain d'Études Supérieures en Gestion, Dakar, Sénégal.

* Corresponding author at: Department of Civil Engineering, University of Ottawa, ON, Canada.

Article History:

Received 21 January 2015

Received in revised form 2 May 2015

Accepted 9 May 2015

ABSTRACT

Study region: The Sirba watershed, Republic of Niger and Burkina Faso, West Africa

Study focus: Water Resources Management in the Sahel region, West Africa, is extremely difficult because of high inter-annual rainfall variability. Unexpected floods and droughts often compromise economic activities and sometime lead to severe humanitarian crises. Seasonal rainfall forecasting is one possible way to increase resilience to climate variability by providing information in advance about the amount of rainfall expected in each upcoming rainy season. Rainfall forecasting models often arbitrarily assume that rainfall is linked to predictors by a multiple linear regression with parameters that are independent of time and of predictor magnitude. Two probabilistic methods based on change point detection that allow the relationship to change according

to time or rainfall magnitude were developed in this paper using Normalized Bayes Factors. Each method uses one of the following predictors: Sea Level Pressure (SLP), Air Temperature (AirTemp) and Relative Humidity (RHUM). Method M1 allows for change in model parameters according to annual rainfall magnitude, while model M2 allows for changes in model parameters with time. M1 and M2 were compared to the classical linear model with constant parameters (M3) and to the climatology (M4).

New hydrological insights for the region: The model that allows a change in the predictor-predictand relationship according to rainfall amplitude (M1) and using AirTemp as predictor is the best model for seasonal rainfall forecasting in the study area.

Keywords: change point detection; seasonal rainfall forecast; Bayes factor; Sirba watershed; West Africa

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

<http://dx.doi.org/10.1016/j.ejrh.2015.05.001>

2214-5818/© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

1. Introduction

Several studies show the degree to which West Africa is vulnerable to climate variability, including those by Gianini et al. (2008) and Christensen et al. (2007). The Sahelian rainfall pattern is season dependent and is directly related to the West African Monsoon (WAM) which dynamic is yet to be fully understood by climatologists (Mohino et al., 2011; Caminade and Terray, 2010; Biasutti et al., 2008; Camberlin et al., 2001; Rowell, 2001, 2003; Janicot et al., 2001; Palmer, 1986). This lack of knowledge about the WAM dynamic is part of the reason for which forecasts in the Sahel at all scales are problematic. The uncertainty in the forecasts directly affects local populations (Hayes et al., 2005). Indeed, the lack of awareness of the short and medium term evolution of rainfall and streamflows often results in populations being poorly prepared to cope with increasingly frequent climate extremes, including lack of precipitation, and floods and their direct corollaries such as lower crop yields, total loss of agricultural production or the destruction of economically valuable infrastructure, such as roads and dams (Tarhule, 2005; Samimi et al., 2012). Recurrent droughts also regularly affect agricultural production, streamflows often take authorities and largely rural local populations by surprise, despite over a decade of publication of seasonal forecasts in West Africa (*PRESAO: Prévision Saisonnière en Afrique de l'Ouest*. Hamatan, 2002; Ogallo et al., 2000). In such an unstable situation, any scientific information regarding the short (24 hours) and medium (6 months) terms of rainfall and streamflow trends becomes a crucial tool for decision-making and water resources management. Agriculture, the primary socio-economic activity in the Sahelian zone, could be more efficient if, local and reliable seasonal information was available to help farmers make critical agricultural decisions

(Hansen, 2002). Thus, the development of seasonal rainfall and streamflow forecast models is highly anticipated by all concerned, particularly the rural population, as it would enable effective use of climatic information that would help ensure food security. The models would increase resilience to climate variability by providing advance information about the expected amount of rain or runoff in the next rainy season (Hansen et al., 2011).

Relevant efforts of the scientific community are based on three different but complementary approaches (Hastenrath, 1995): dynamical (based purely on numerical models), statistical (based purely on statistics) and hybrid statistical-dynamical (a combination of statistics and numerical models).

The dynamical approach is based on numerical models of physics and dynamics equations that describe the climate system (Kumar et al., 1996; Brankovic and Palmer, 1997; Palmer et al., 2000; 2004). The statistical approach consists of establishing a direct relationship between the state of the atmosphere or ocean at the moment of the forecast and during event occurrences (e.g. precipitation) within the period of a few months or weeks (Schepen et al., 2012; Lopez-Bustins et al., 2008). The existence of sufficiently strong and robust physical links between certain variables is regarded as foreseeable, and is the basis of the statistical forecast. The hybrid statistical-numerical approach also known as model output statistics (MOS), is a combination method based on the principle of applying statistical methods to the output obtained from numerical models, in order to perform further analysis.

Statistical models are quite popular, given their ease of development and the limitations of dynamical models (Sittichok et al., 2014; Ibrahim et al., 2014; Mara, 2010; Bouali, 2009; Biasutti et al., 2008; Hayes et al., 2005; Philippon and Fontaine, 2002; Janicot et al., 2001; Hunt, 2000; Thiaw et al., 1999). However, it is notable that all models developed from these studies arbitrarily consider the relationship between the predictors and the predictand (rainfall in the Sahel) to be independent of time and rainfall magnitude.

The objective of this paper is to depart from that hypothesis to develop statistical seasonal rainfall forecasting models with changing parameters, and to instead compare the new models to the classical linear model with constant parameters and to the climatology.

First, a linear rainfall forecasting model is developed for each of the predictors under consideration, as in Sittichok et al. (2014). At the end of the process, an optimal lag time and optimal season are obtained to average the predictor. Using the latter lag time and the predictor time series, new models are developed that allow the linear regression parameter to change according to time or rainfall amplitude. The performance of the new models is then compared to that of the original linear model, and to a model representing the rainfall climatology in the study area.

2. Materials and methods

2.1 Study area

The area under study considered in this work is the Sirba watershed, a transboundary watershed, shared by Burkina Faso and Niger, located between latitudes 12°55'54"-14°23'30"N and longitudes 1°27'W-1°23'42"E with an area of 38750 km² (Mara, 2010). Fig. 1 depicts the geographical situation and characteristics of the area, which is influenced by three sub-climate zones based on the decrease of rainfall from south to north: the southern Soudanian zone with mean annual rainfall of 700 to 800mm, the northern Soudanian zone with mean annual rainfall of 550 to 650mm and the Sahelian zone with mean annual rainfall of 300 to 500mm (Taweye, 1995). Most of rainfall occurs from July to September (JAS), regardless of the sub-climate zone. The climate is characterized in part by having only two seasons: a dry season (October to April) due to the harmattan (dry wind) and a rainy season (May to September) influenced by the WAM (cold wind) (Descroix et al., 2009). The hydrographic network is relatively dense, and consists of three main tributaries (Sirba, Faga, and Yeli) plus a few dam water reservoirs (Mara, 2010). Based on descriptions of the rainfall pattern, the hydrological regime in the Sirba watershed is the Sahelian type, and its vegetation formation is thorny, lightly-wooded savannah (Andersen, 2005; Descroix et al., 2009). The reason for choosing the Sirba basin is threefold. First, it is located approximately in the middle of the Sahel region, so, it is influenced by the climate characteristics of both northern Sahel and the Sahara desert, and southern Sahel and the Sudanian savanna. Second, there are many climate stations inside and around the basin that collect climate data daily. And third, there is more than 40 years of precipitation data available.

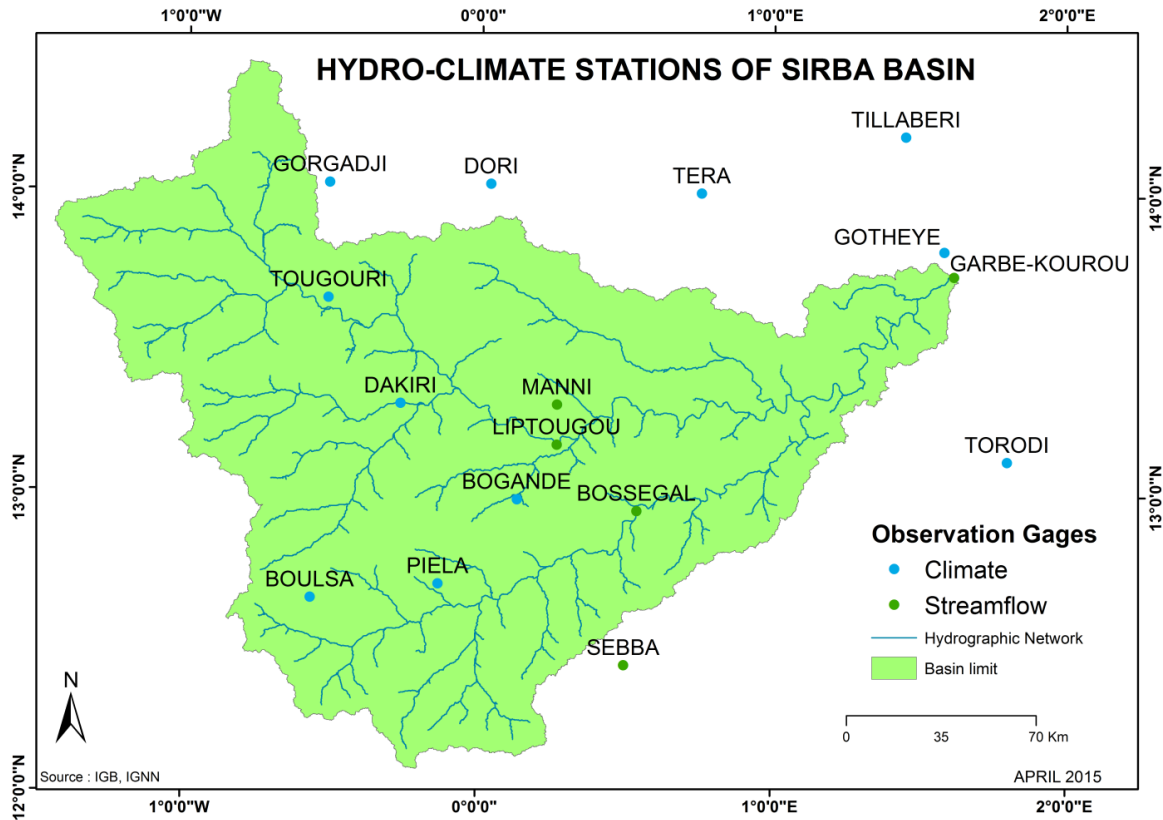


Fig. 1. Location map of the Sirba watershed and its hydro-meteorological stations

2.2 Climatic Data

The predictand in this study is the average seasonal precipitation in the Sirba watershed. It was calculated using daily rainfall data that was recorded by a network of 11 rain gage stations in Burkina Faso and Niger, from 1960-2008. The rainfall time series were provided by the national meteorological offices of Burkina Faso and Niger. Five of the stations are located within the watershed, and the remaining six are a maximum of 25km from the watershed boundary (see Fig. 1). Using the 11 rainfall time series, the Thiessen polygon method was applied to estimate the average rainfall in the watershed.

The atmospheric data are Sea Level Pressure (SLP), Relative Humidity (RHUM), Air Temperature (AirTemp), Zonal Wind (UWND) and Meridional Wind (VWND). The variables are monthly *NCEP-DOE Reanalysis* data obtained from the National Oceanic and Atmospheric Administration (NOAA: <http://www.esrl.noaa.gov>). They relate to the grid 90° N-90° S latitudes and 0° E-357.5° E longitudes, and span the period from January 1979 to August 2013. Tables 1 and 2 present the rainfall and atmospheric data, respectively.

Table 1
Details of rainfall stations

Station number (code)	Station name	Longitude	Latitude	Country
320006	Torodi	1.8	13.12	Niger
320002	Tera	0.82	14.03	Niger
320004	Tillaberi	1.45	14.20	Niger
320005	Gotheye	1.58	13.82	Niger
200082	Boulssa	-0.57	12.65	Burkina Faso
200026	Dori	0.033	14.03	Burkina Faso
200085	Bogande	0.13	12.98	Burkina Faso
200048	Dakiri	-0.27	13.30	Burkina Faso
200024	Gorgadji	-0.52	14.03	Burkina Faso
200086	Piela	-0.13	12.70	Burkina Faso
200047	Tougouri	-0.52	13.65	Burkina Faso

Table 2
Description of climate variables

Parameter	Units	Level	Reference data	Spatial coverage	Temporal coverage
Sea level pressures (SLP)	Pa/s	1000 hPa	NCEP 2	2.5°x2.5° grid 90N - 90S, 0E - 357.5E	1979/01/01 to 2013/08/31
Air temperature (AirTemp)	°K	1000 hPa	NCEP 2	2.5°x2.5° grid 90N - 90S, 0E - 357.5E	1979/01/01 to 2013/08/31
Relative humidity (RHUM)	%	1000 hPa	NCEP 2	2.5°x2.5° grid 90N - 90S, 0E - 357.5E	1979/01/01 to 2013/08/31
Meridional wind (VWND)	m/s	1000 hPa	NCEP 2	2.5°x2.5° grid 90N - 90S, 0E - 357.5E	1979/01/01 to 2013/08/31
Zonal wind (UWND)	m/s	1000 hPa	NCEP 2	2.5°x2.5° grid 90N - 90S, 0E - 357.5E	1979/01/01 to 2013/08/31

2.3 Selection of the optimal lag time for each predictor

Monthly precipitation time series from the Climatic Research Unit (CRU TS 3.21 0.5° global), with a spatial resolution 0.5° x 0.5° defined on 2° W-2° E longitude, and 10° N-15° N latitude (covering more than the area of Sirba basin) were initially used as predictand for selecting a pool of potential predictors for seasonal rainfall forecasting.

After establishing the pool of predictors, the observed precipitation from rain gage stations was used to determine the best predictors in the group. The method developed by Sittichok et al. (2014) was used to link the observed rainfall with each predictor, and the candidate predictor was aggregated over all possible time windows (with a time window length in months is an integer) in the 18 months prior to the rainy season. Each of the obtained time series was used as input to a linear model linking it to the seasonal rainfall on the Sirba watershed. How the periods are generated is shown in Fig. 2. For each period, a linear model linking the predictor averaged over that period and seasonal rainfall on the Sirba is built as follows:

1. For each year Y that the predictor was available.
 - i) the predictor of year $Y-1$ was removed from the predictor grid;
 - ii) the rainfall of year Y was removed from the rainfall data set; and,
 - iii) the dimension of the remaining predictor data set was reduced using the coefficient of determination (R^2) to screen predictor gridded points and obtain a small number of predictors. Principal component analysis (PCA) was then applied on the remaining predictor gridded data from the previous step to reduce the number of predictors
2. A linear regression was fitted between the predictor and precipitation time series.
3. The fitted linear regression was used to simulate the rainfall of year Y . If predictor and rainfall were in the same year (Year Y), only predictor and rainfall time series for that year were removed in the first step.
4. When the simulated rainfall was available for every year in the historical period, the objective functions R^2 , Nash, and Hit-Rate score were calculated to estimate the model's performance.

The period that yielded the best Nash coefficient (i.e. the optimal lag time) is then selected. Table 3 summarizes the final selected predictors used in this study to forecast seasonal rainfall.

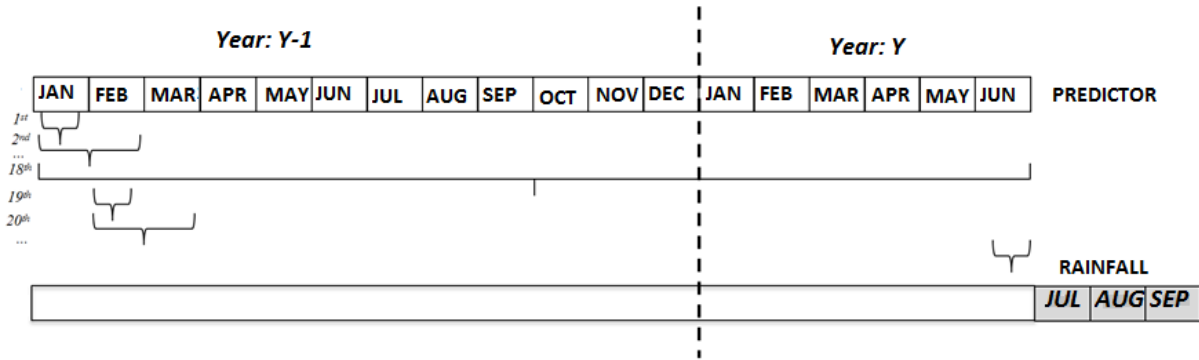


Fig. 2. Predictor averaging periods

Table 3

Selected predictors with their lag time for seasonal forecast

Predictors	NMAX ^b	R ²	Nash	HIT Rate Score	Best period M1-M2 ^a	Lagged Period
Sea Level Pressure at 1000hPa	50	0.48	0.46	60.71	17 – 18	0
Relative Humidity at 1000hPa	80	0.58	0.52	64.29	10 – 10	8 months
Air Temperature at 1000hPa	10	0.530	0.527	67.86	1 – 4	14 months

(a) M1=1:12 (January to December); M2=M1:18 (considered month of M1 to the next coming June)

(b) NMAX: number of best grid points retained after screening the predictor grid based on R²

2.4 Seasonal forecasting models with changing parameters

The adapted algorithm of the Bayesian change point detection method is presented before describing the developed models with changing parameters.

2.4.1 Multiple change point detection algorithm

Change points can be defined as discontinuities of time series that normally exist in climate data (Reeves et al., 2006). They can occur for many reasons, including, observed station movement, changes in recording equipment, changes in measurement techniques, environmental changes and climate change effects such as shifts in climate regimes (Lund and Reeves, 2002). There are many methods in the literature to detect and correct change points in various fields of research (Vincent, 1998; Begert et al., 2005; Beaulieu et al., 2005, 2009; Fearnhead, 2006; Seidou et al., 2007; Seidou and Ouarda, 2007; Villarini et al., 2011). Indeed, the Bayesian method for change point analysis is one of the most popular techniques, as it helps obtain the statistical distribution for the dates of

change as well as the distribution for the other parameters in the model (Sarr et al., 2013; Seidou et al., 2007; Seidou and Ouarda, 2007; Xiong and Guo, 2004; Perreault et al., 2000; Gelfand et al., 1990; Barry and Hartigan, 1993). In this study, the Bayesian change point method proposed by Seidou and Ouarda (2007) is employed to evaluate abrupt changes in mean or direction of trends for climatic variables. This method was adopted because it handles an unknown number of changes and displays the complete probability distribution of the dates of the changes. The Bayesian change point detection model used in this present case can also evaluate abrupt changes in the relationship between the principal and a number of relevant explanatory variables. In these cases, the estimated trend for each segment of the time series is performed based on proxies.

A brief description of the model algorithm follows. Readers can refer to Seidou and Ouarda (2007) and Ehsanzadeh & Adamowski (2010) for further details.

Let $Y = (y_1, y_2, \dots, y_n)$ be the n -sample of observations representing the response variable, m be the unknown number of change points and $\tau_0 = 0, \tau_1, \dots, \tau_{m+1} = n$. Let $Y_{t:s}$ be observations from time t to time s ; $Y_{t:s} = (y_t, y_{t+1}, y_{t+2}, \dots, y_s)$ ($t \leq s$). Then, for $k = 1, \dots, m + 1$, the k^{th} segment is the set of data observed between $\tau_{k-1} + 1$ and τ_k . A parameter ϕ_k is associated with the k^{th} segment and $\pi(\phi_k)$ denotes the prior distribution of ϕ_k . As established by Fearnhead (2006), the posterior probability of change points is given by:

$$\begin{cases} Pr(\tau_1/Y_{1:n}) = P(1, \tau_1)Q(\tau_1 + 1)g_0(\tau_1)/Q(1) \\ Pr(\tau_k/\tau_{k-1}, Y_{1:n}) = P(\tau_{k-1} + 1, \tau_k)Q(\tau_k + 1)g(\tau_k - \tau_{k-1})/Q(\tau_{k-1} + 1), \\ \text{for } k = 2, \dots, m \end{cases} \quad (1)$$

where (g) is the probability distribution of the time interval between two consecutive change points, and g_0 is the probability distribution of the first change point. For $s \geq t$ and $y_i \in Y_{t:s}$;

$P(t, s) = \int \prod_{i=t}^s f(y_i/\theta)\pi(\theta)d\theta$ is the probability of t and s belonging to the same segment. $Q(t)$ is the likelihood of segment $Y_{t:n}$ given a change point at $t - 1$, and is derived from a recursive relation using $P(t, s)$ and both g and g_0 (see theorem 1, Fearnhead, 2006).

Now, let $X = (x_{1j}, x_{2j}, \dots, x_{nj})$, and $j = 1, \dots, d^*$ denote the set of d^* explanatory vectors including any intercepts. Thus, the multiple linear regression can be written as:

$$y_i = \sum_{j=1}^{d^*} \theta_j x_{ij} + \varepsilon_i, i = 1, \dots, n \text{ or } Y = X\theta + \varepsilon \quad (2)$$

where $\theta = (\theta_1, \theta_2, \dots, \theta_{d^*})$ is the vector of the regression parameters and $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{d^*})$ is the Gaussian vector of residuals with mean zero and variance σ^2 . Note that relation (1) changes after each change point and is recomputed for each segment. In a given segment, the parameter vector ϕ is defined as:

$$\phi = (\theta_1, \theta_2, \dots, \theta_{d^*}, \sigma) \quad (3)$$

and it follows that:

$$f(y_i/\emptyset) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-0.5\left(\frac{y_i - \sum_{j=1}^{d^*} \theta_j x_{ij}}{\sigma}\right)^2\right) \quad (4)$$

In this study, the prior distribution to be used depends only on the scale parameter σ and as such:

$$\pi_1(\emptyset) = \pi_1(\sigma) = p(\sigma/a, c) = \frac{\sigma^{-a} \exp\left(-\frac{c}{2\sigma^2}\right)}{2^{\frac{a-3}{2}} c^{\frac{a-1}{2}} \Gamma\left(\frac{a-1}{2}\right)} \quad a > 1, c > 0 \quad (5)$$

where a and c are the hyper parameters. Hence, as shown in Seidou & Ouarda (2007), in this setting the posterior probability of the change point displayed in Equation (1) is given by:

$$P(t, s) = (2\pi)^{\frac{d^*}{2}} \frac{(\pi(\varepsilon_{t:s}^T \varepsilon_{t:s} + c))^{\frac{(s-t+a-1)}{2}} \Gamma\left(\frac{s-t+a-d^*}{2}\right)}{(c\pi)^{\frac{a-1}{2}} |X_{t:s}^T X_{t:s}|^{1/2} \Gamma\left(\frac{a-1}{2}\right)} \quad \text{for } s \geq t \quad (6)$$

In this study, parameter a in Eq. (6) is fixed at 2, so that the prior distribution is non-informative.

The Bayesian change point detection model first estimates the posterior distribution of probability of the number of changes. The most probable number of detected changes (associated with the highest probability of occurrence) is then selected as the number of change points observed in the data series. Conditional on this number, the Bayesian inference then provides the time position of detected changes and their respective (posterior) distribution of probability of occurrence. Finally, the magnitude of the detected changes is determined. The identified changes could represent shifts in the mean, changes in the direction of a trend, or a combination of both.

2.4.2 Model M1

Model M1 was developed to detect potential changes in the relation between predictor and predictand and assumes that the relationship changes with precipitation amplitude. To obtain the forecast for a given year i ($i=1979 - 2002$) the model is applied as follows:

1. Year i is removed from the predictor and predictand time series.
2. Stepwise regression is used to fit a linear relation between the predictor and the predictand in the remainder of the series, and an initial forecast assuming a single equation for all points is issued. The equation is also used to issue an initial forecast for year i .
3. The data is sorted in increasing order of the forecasted predictand from step 2, and a new position i_l is assigned to year i . The initial forecast for year i is between the initial forecast for year i_l and the initial forecast for year $i_l + 1$.

4. The change point detection method by Seidou et al. (2007) is applied to the remaining data. The method generates 1,000 time series of length $N-1$, with a random number of change points at random locations. The density of the change points in a given time interval is proportional to the probability of change in that interval (Seidou et al., 2007).
5. For each of the 1,000 generated sequences of change points, stepwise regression is applied to fit a linear relation between the predictor and predictand on any segments delineated by the change points, and, both the optimal (least square) forecast and the standard deviation of the residual are calculated. If m is the order of the current generated sequence of change points, i is in the k^{th} segment, $x_1, x_2, \dots, x_n$ the values of the predictors for year i and $\alpha_1^{k,m}, \alpha_2^{k,m}, \dots, \alpha_n^{k,m}$ are the coefficients of the equation for the segment k , then the least square estimate is $\hat{Y}_i = \alpha_0^{k,m} + \alpha_1^{k,m} \times x_1 + \alpha_2^{k,m} \times x_2 + \dots + \alpha_n^{k,m} \times x_n$
6. Ten probabilistic forecasts are generated by sampling ten values from a normal distribution. The mean of the normal distribution is the least square forecast and its standard deviation is the standard deviation of the forecasts.
7. The 10,000 forecasts for year i are used to calculate the empirical probability density of the forecast. The estimate of the distribution is nonparametric, uses a normal kernel function, and is evaluated at 1,000 equally spaced points that cover the range of the data set.

At the end of the process, a probability density function is obtained for the forecast in year i .

2.4.3 Model M2

Model M2 is similar to M1, except that it assumes that the predictand-predictor relationship changes with time (i.e., the regression parameters change over time). The same approach as in M1 was followed, but there was no ordering of the data set after each exclusion of year i ($i=1979 - 2002$).

1. Year i is removed from the predictor and predictand time series.
2. The Seidou et al. (2007) change point detection method is applied to the remaining data. The method generates 1,000 time series of length $N-1$, with a random number of change points at random locations. The density of change points in a given time interval is proportional to the probability of change in that interval (Seidou et al., 2007).
3. For each of the 1,000 generated sequences of change points, stepwise regression is applied to fit a linear relation between the predictor and predictand on any segments defined by the change points. Both the optimal (least square) forecast and the standard deviation of the residual are calculated. If m is the order of the current generated sequence of change points, i is in the k^{th} segment, $x_1, x_2, \dots, x_n$,

and the values of the predictors for year i and $\alpha_1^{k,m}, \alpha_2^{k,m}, \dots, \alpha_n^{k,m}$ are the coefficients of the equation for segment k , then the least square estimate is $\hat{Y}_i = \alpha_0^{k,m} + \alpha_1^{k,m} \times x_1 + \alpha_2^{k,m} \times x_2 + \dots + \alpha_n^{k,m} \times x_n$

4. Ten probabilistic forecasts are generated by sampling ten values from a normal distribution. The mean of the normal distribution is the least square forecast and its standard deviation is the standard deviation of the forecasts.
5. The 10,000 forecasts for year i are used to calculate the empirical probability density of the forecast. The estimate of the distribution is nonparametric, uses a normal kernel function, and is evaluated at 1,000 equally spaced points that cover the range of the data set.

At the end of the process, a probability density function is obtained for the forecast in year i ($i=1979-2002$).

Fig. 3 recapitulates the steps involved in the models M1 and M2.

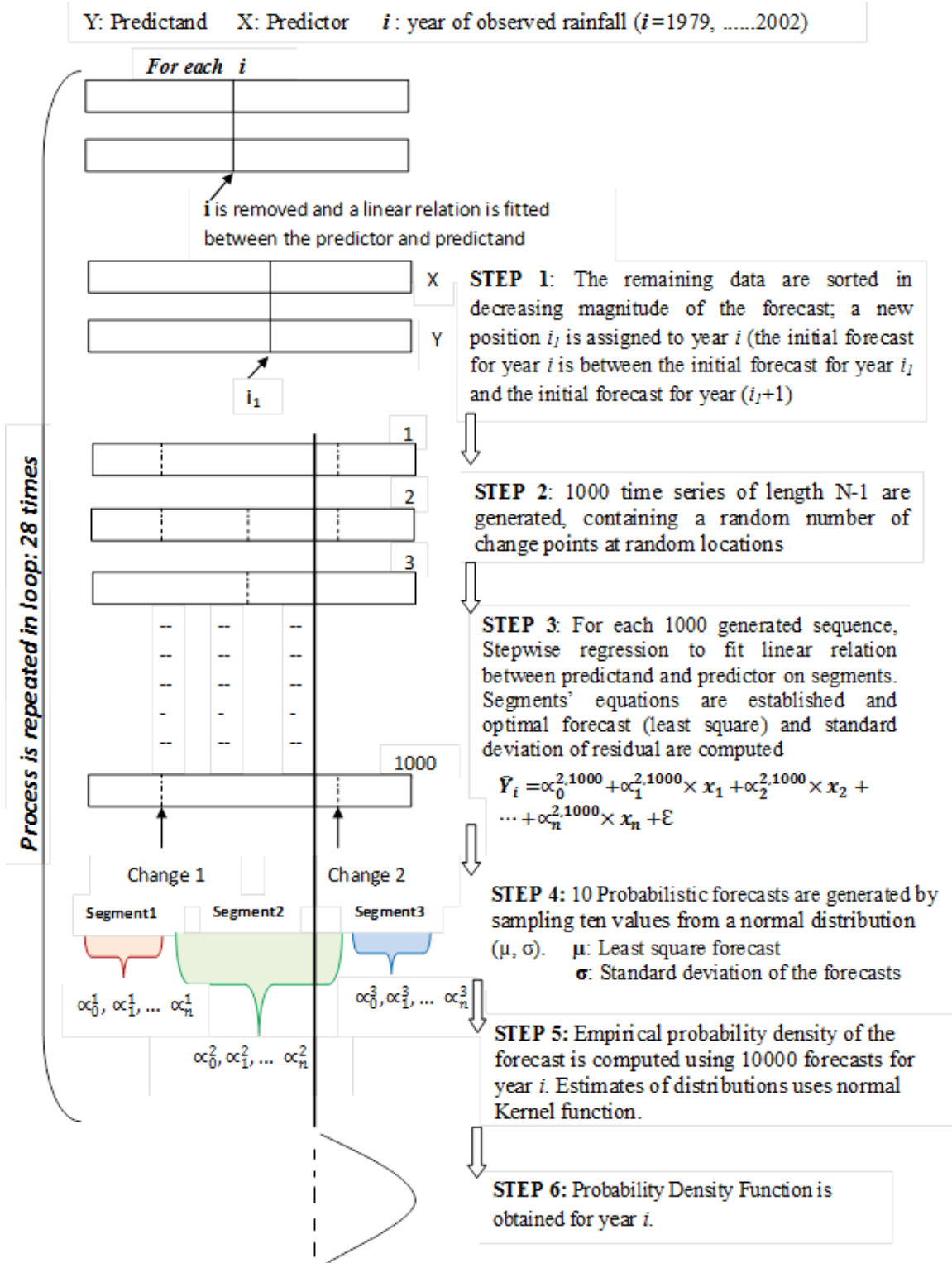


Fig. 3. Steps in seasonal rainfall forecasting models with changing parameters M1 and M2 (All steps are followed except Step 1 which is not included while using model M2)

2.5 Seasonal forecasting models with constant parameters

Two models with constant parameters were developed and tested in order to find the best seasonal rainfall forecast model. The first method (M3) is the classical linear model with constant parameters, and the second (M4) is based on the climatology.

2.5.1 Model M3

In model M3, no change points are assumed in the linear regression between predictand and predictors. The model M3 is applied as follows (see Fig. 4):

1. Year i is removed from the predictor and predictand time series.
2. Stepwise regression is used to fit a linear relation between the predictor and predictand. Both the optimal (least square) forecast and the standard deviation of the residual are calculated. If i is in the k^{th} segment, $x_1, x_2, \dots, x_n$, and the values of the predictors for year i and $\alpha_1^k, \alpha_2^k, \dots, \alpha_n^k$ are the coefficients of the equation for the segment containing i , then the least square estimate for year i is $\hat{Y}_i = \alpha_0^k + \alpha_1^k \times x_1 + \alpha_2^k \times x_2 + \dots + \alpha_n^k \times x_n$
3. Ten probabilistic forecasts are generated by sampling ten values from a normal distribution. The mean of the normal distribution is the least square forecast and its standard deviation is the standard deviation of the forecasts.
4. The 10000 forecasts for year i are used to calculate the empirical probability density of the forecast. The estimate of the distribution is non-parametric, uses a normal Kernel function, and is evaluated at 1000 equally spaced points that cover the range of the data set.

At the end of the process, a probability density function is obtained for the forecast in year i ($i=1979$ to 2002).

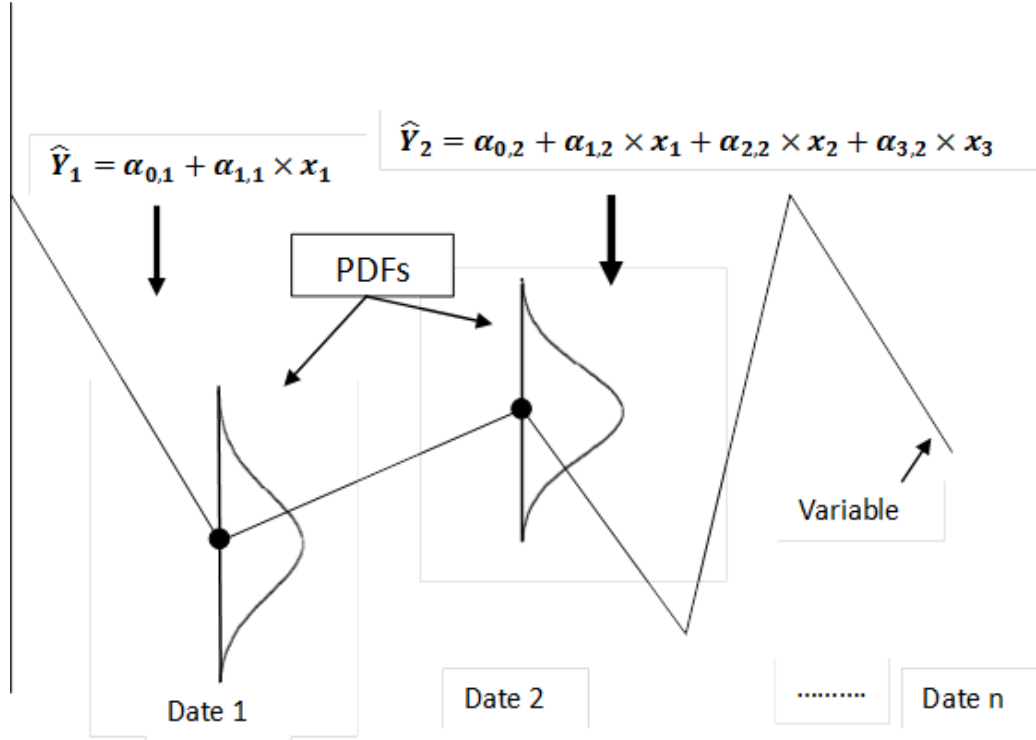


Fig. 4. Graphical description of model M3

2.5.2 Model M4

Under model M4, the climatology is used to estimate the seasonal rainfall. The probability density of the forecast is a normal distribution in which the average observed precipitation is the mean and the standard deviation is the standard deviation of the observed precipitation (see Fig. 5). Model M4 is applied as follow:

1. Year i is removed from the predictor and predictand time series.
2. The average and the standard deviation of the observed precipitation are calculated on the remainder of the data.
3. The probability distribution of the forecast is generated at 1,000 points over the range of the data, using a normal distribution. The mean of the normal distribution is the average observed precipitation, and the standard deviation is the standard deviation of the observed precipitations.
4. The 10,000 forecasts for year i are used to calculate the empirical probability density of the forecast. The estimate of the distribution is nonparametric, uses a normal kernel function, and is evaluated at 1,000 equally spaced points over the range of the data set.

At the end of the process, a probability density function is obtained for the forecast in year i .

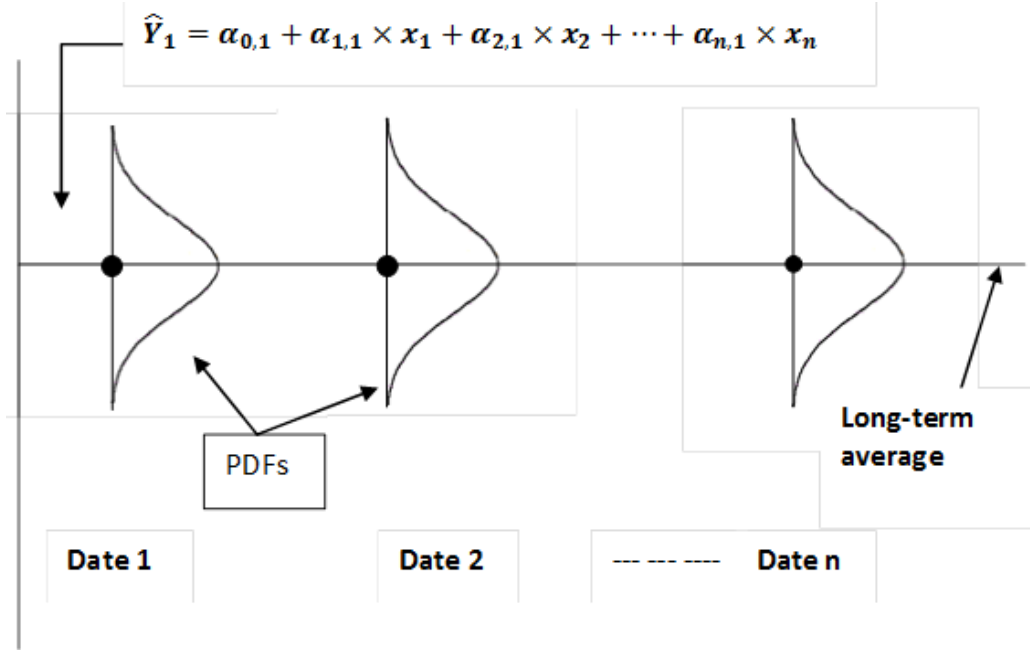


Fig. 5. Schematically presentation of model M4 using climatology

2.6 Bayesian model selection

In this paper, the Bayesian approach is used to select the best seasonal rainfall forecast model from developed models with changing parameters and those with constant parameters. The posterior and prior probabilities of the models and the observed data were computed first (see Eq. (7)).

$$Posterior Prob_{model} = \frac{(Prior Prob_{model} \times Likelihood_{model})}{Prob_{observations}} \quad (7)$$

where, $Likelihood_{model} = \prod_{i=1}^n likelihood_i$

i is the year of forecasted rainfall and, n is the number of years.

To apply Eq. (7), all twelve models (i.e. M1, M2, M3, and M4 used with each of the three predictors) were considered. The ratio of a model's posterior probability to observations constitutes a comparative criterion after normalization. Normalized Bayes factors B_f (see Eq. (8)) were calculated for each model to facilitate comparison between the models' results, and to provide a weighted comparison of the likelihood of each model given the observed data. B_f compares the posterior likelihood of data d of a given model M_i to that of the reference model M_r . For more details about Bayes factors, refer to Min et al. (2007).

$$B_f = \frac{Likelihood(d/M_i)}{Likelihood(d/M_r)} \quad (8)$$

Selection of the best seasonal rainfall forecast model required analysis of the Bayes factors for all 12 models. Table 5 demonstrates how to interpret Bayes factor. There is strong evidence favorably supporting model M_r (a reference model) and M_i (a given

model), when B_f is less than 1/10 and higher than 10, respectively. In contrast, B_f between 1/3 and 3, meant that both the M_r and M_i models are weak models. Hence, the best forecast model is the one for which B_f is always favorable in terms of value. In addition, the evolution of likelihoods of forecasted seasonal precipitation (JAS) was confirmed via a graphical representation of each model. The graphs show the likelihood of each forecasted rainfall value on a coloured scale, where red and blue represent a probability of 1 and 0, respectively.

The entire methodology used in this work is summarized by the flowchart presented in Fig. 6.

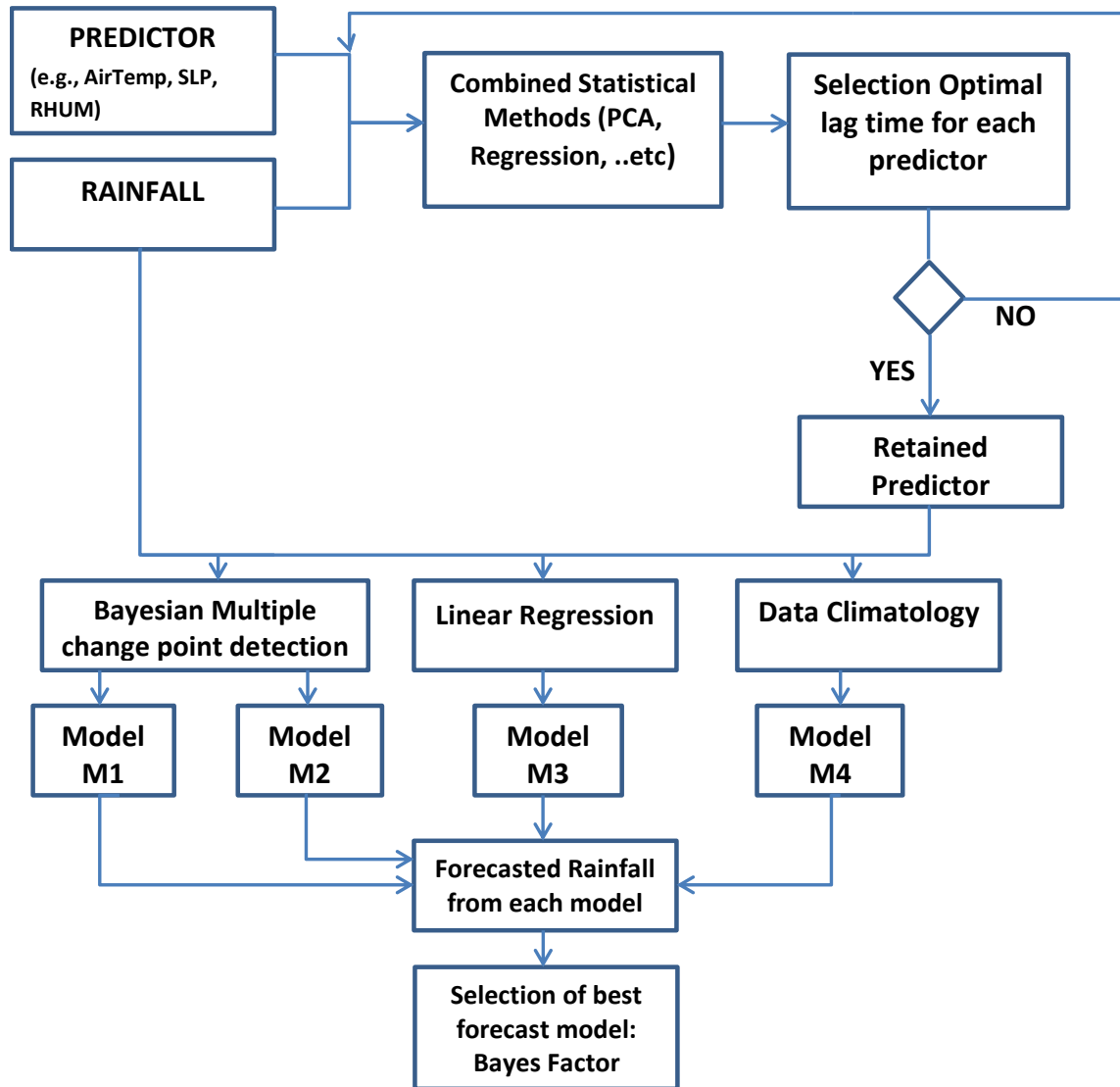


Fig. 6. Selection process of best seasonal rainfall forecast model

2.7 Performance measures

In this study, the relative performances of the four rainfall seasonal forecasting models (M1, M2, M3 and M4) were compared quantitatively using the Nash-Sutcliffe criterion (Nash). This criterion was chosen because it can present the differences in magnitude between observed and simulated data during the entire time period. The best Nash value equates to the best performance.

The performance of each model (under each predictor) was further evaluated based on two other criteria: i) the number of forecasted values per model with high likelihoods (e.g. 80-100%); and ii) the model's performance if the data support it favorably, based on Bayes factors; if so it is deemed to have credible high performance.

Therefore, the model is considered to perform better if almost 100% of its forecasted values have high likelihoods, which is clearly indicated by the red plot on the graph. The opposite (i.e., low likelihoods) is plotted in blue.

3. Results and discussions

3.1 Changes in relationship

It was found that the linear relation between the predictor and predictand systematically displayed the presence of one or more change points. The probability of change, as well as the conditional probability of the change points, was calculated according to the work of Seidou and Ouarda (2007). Table 4 summarizes the number of change points and their respective locations for each model. It was observed that the number of change points varies between models. Fig. 7 shows the output of model M1 as a histogram representing the probability of occurrence of the first change in the data, in the case of AirTemp. The weight at the first date (position) is the probability of no change. Clearly, the effectively localized histogram indicates that the position of the change is well identified. It was assumed that no change occurs when the weight is above 0.5, and, when the probability of change is above 0.5, the position of the change was assumed to be beyond the first year that had the highest probability. Thus, in Fig. 7 at year 1990 (lower panel at position 12) it was observed that the probability of change is 55%, and the probability of no change is 45%. The conditional probability of the existing change point is also given, and shows there is a 60% chance of change at position 12.

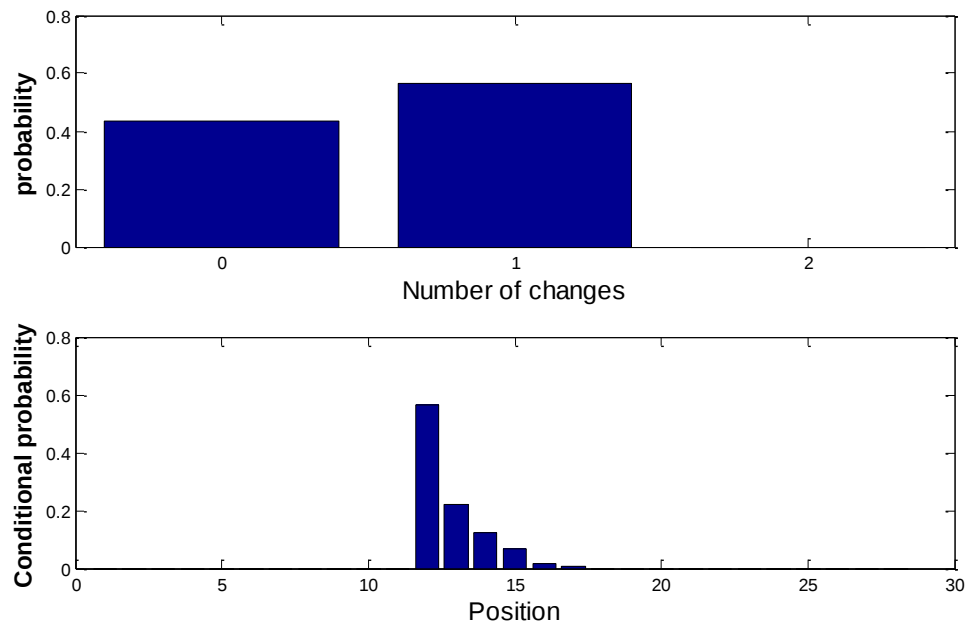


Fig. 7. Histogram of change point detection result for model M1

Table 4

Number of change points and their most probable locations for each model

Models	Number of change points	Position of changes
M1 _a	1	1992
M1 _{rh}	1	1990
M1 _s	1	1990
M2 _a	1	1990
M2 _{rh}	1	1990
M2 _s	1	1992

3.2 Performance of forecasting models

The relative performance of the four rainfall seasonal forecasting models described in Sections 2.4 and 2.5 was obtained using the observed and forecasted seasonal time series. The results showed Nash values of 0.76, 0.52, 0.46 and 0.58 for models M1, M2, M3 and M4 respectively. Since the objective function used to present the forecast skill is Nash, the best performance equates to the best Nash, which indicated that model M1 outperformed the others, followed by model M4.

The limitations of model M3 could be because, unlike the Nash criteria, regression is not suitable for measuring the difference in magnitude of both the observed and simulated data. As for model M2, its limitations are the result of the imposed condition that makes the rainfall-predictor relationship change over time. Models M1 and M4 seems to perform acceptably.

Considering the performance of the models under each of the three predictors, it is interesting that for the models using AirTemp as the predictor, 92%, 38%, 81%, and 54% of the forecasted values have high likelihoods for models M1, M2, M3 and M4, respectively. Thus, the strongest model is M1_a, followed by M3_a, M4_a and M2_a. For models using RHUM as the predictor, the performance of the models in decreasing order is M1_{rh}, M4_{rh}, M2_{rh}, and M3_{rh}, as they have 86%, 79%, 29%, and 28% of forecasted values with high likelihoods, respectively. For the last predictor SLP, the performance of each model is shown by the inequalities M4_s>M1_s>M2_s>M3_s, as models M1, M2, M3, and M4 have 86%, 59%, 36%, and 89% of forecasted values have high likelihoods, respectively.

3.3 Selection of the best model

Selection of the best seasonal rainfall forecast model involved analyzing the Bayes factors for all model combinations, and visually examining the location of seasonal (JAS) precipitation on a graphical representation of each model's posterior likelihood. If the observations are largely in areas of high likelihood according to a given model, that model is deemed credible. Figs. 8-10 present the graphs for models M1, M2 and M3. On each graph, some rainfall values were in the red range (high probability) and others were in the blue range (low probability). Analysis of the evolving likelihoods (Fig. 8) found that 22 of the 28 forecasted rainfall values under model M1 using AirTemp are in the red range, which indicates high probability.

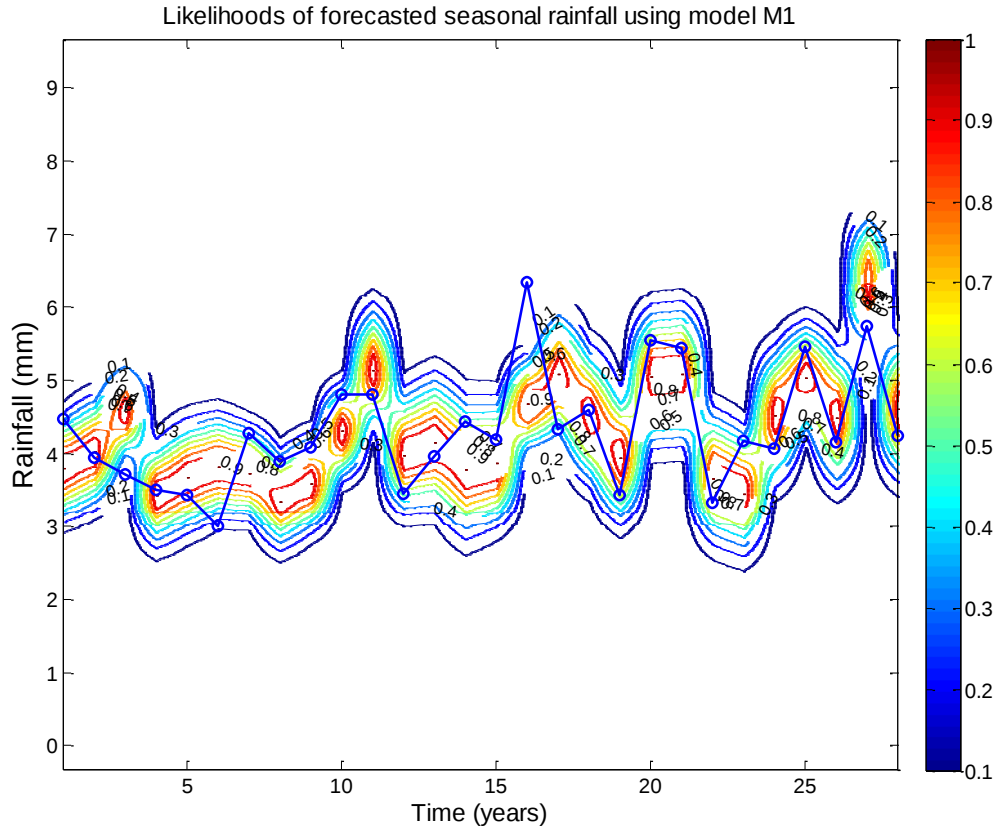


Fig. 8. Evolving probabilities of forecasted seasonal rainfall from M1 using Air Temperature

Figs. 9 and 10 show the evolving likelihoods of forecasted seasonal precipitation for models M2 and M3, respectively. On these graphs, most of the forecasted seasonal rainfalls tend toward blue (i.e. low probability). In models M2 and M3, only 29% and 36% of the observations fall in areas of high likelihood respectively, so the models are deemed not credible.

Table 5

Scale for Bayes factor interpretation (source: Min et al., 2007)

Bayes Factor	Interpretation
$B_f < 1/10$	Strong evidence for M_r
$1/10 \leq B_f < 1/3$	Moderate evidence for M_r
$1/3 \leq B_f < 1$	Weak evidence for M_r
$1 \leq B_f < 3$	Weak evidence for M_i
$3 \leq B_f < 10$	Moderate evidence for M_i
$B_f \geq 10$	Strong evidence for M_i

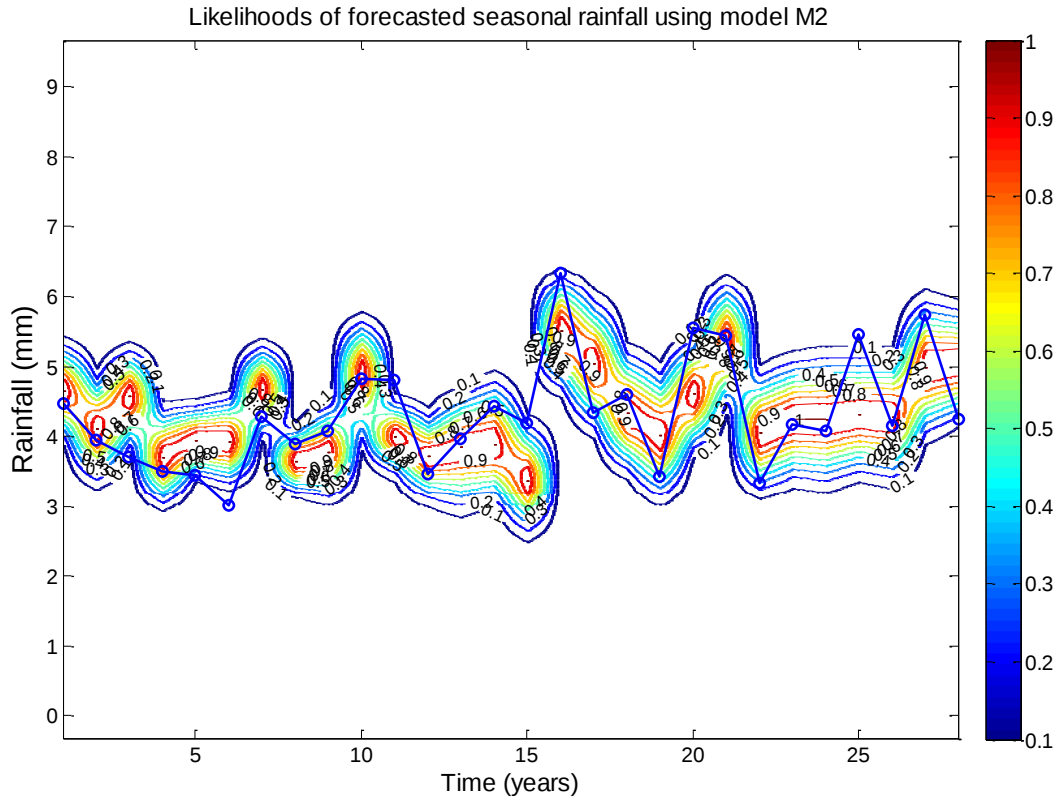


Fig. 9. Evolving probabilities of forecasted seasonal precipitations from M2 using RHUM

Table 6 displays the normalized Bayes factors for all models. Table 5 shows how to interpret the magnitude of the Bayes factors, and concludes that there is weak, moderate or strong evidence to support the competing models. Table 7 shows that there is a strong evidence for Model M1, using AirTemp as predictor. Bayes factors favorably supported model M1 (AirTemp) because it had strong evidence (St.E) either as reference or given model, compared to the others which had moderate or weak evidence as in Table 7. Thus, for seasonal rainfall forecasts, evaluating changes in the relationship of predictand-predictor with the rainfall amplitude seems to be the best approach for the Sahelian region.

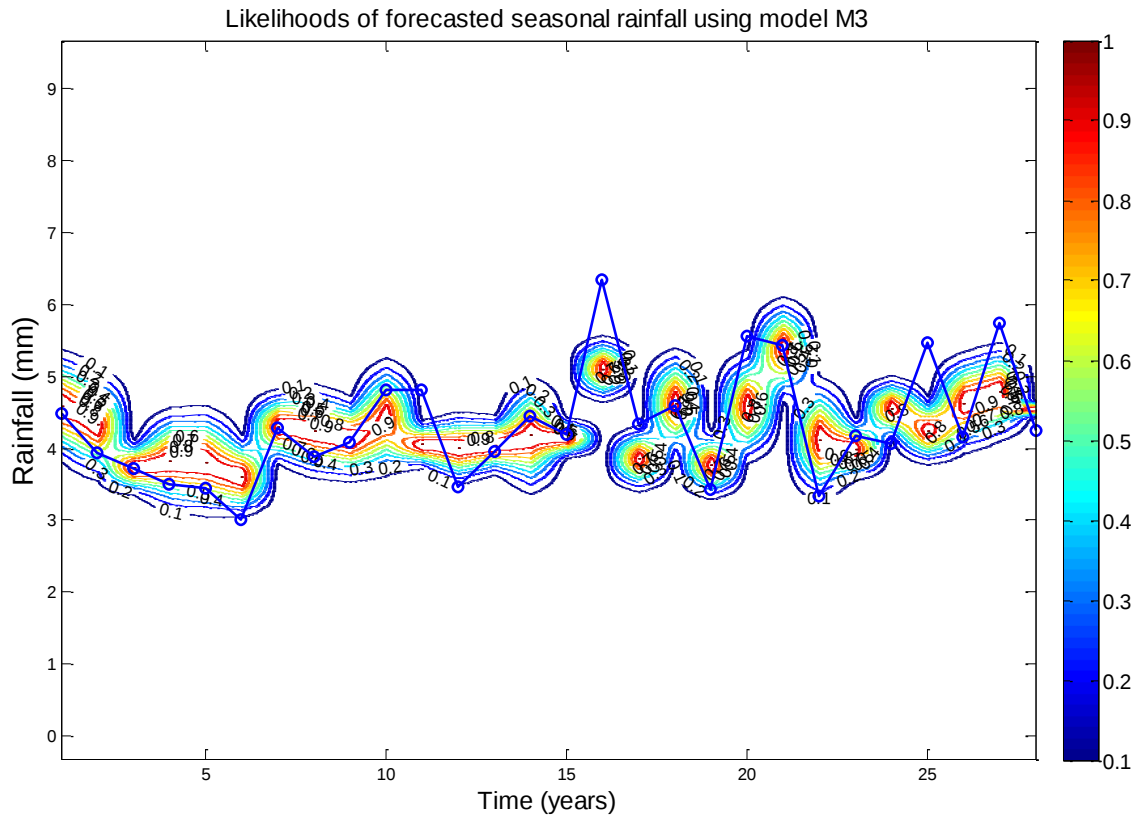


Fig. 10. Evolving probabilities of predicted seasonal precipitations from M3 using SLP

Thus, seasonal forecast models with parameters that change according to rainfall magnitude could be considered optimal for seasonal rainfall forecast over the Sirba watershed, rather than classical models where parameters are constant. Combining this changing parameter model with the Bayesian change point detection procedure, and using the normalized Bayes factor, constitutes an acceptable means of forecasting seasonal rainfall over West Africa, and address an issue that has challenged forecasters for more than two decades.

Table 6

Normalized Bayes factors of twelve seasonal rainfall forecast models

$\begin{matrix} M_i \\ M_r \end{matrix}$		AirTemp				RHUM				SLP			
		M1 _a	M2 _a	M3 _a	M4 _a	M1 _{rh}	M2 _{rh}	M3 _{rh}	M4 _{rh}	M1 _s	M2 _s	M3 _s	M4 _s
AirTemp	M1 _a	1	8.62E-05	2.53E+00	5.53E-04	9.53E-02	1.19E-14	4.63E-20	5.53E-04	6.09E-06	8.82E-12	5.91E-23	5.53E-04
	M2 _a	1.16E+04	1	2.94E+04	6.41E+00	1.11E+03	1.38E-10	5.37E-16	6.41E+00	7.07E-02	1.02E-07	1.74E-18	6.41E+00
	M3 _a	3.95E-01	3.41E-05	1	2.18E-04	3.77E-02	4.69E-15	1.83E-20	2.18E-04	2.41E-06	3.48E-12	1.50E-22	2.18E-04
	M4 _a	1.81E+03	1.56E-01	4.58E+03	1	1.72E+02	2.15E-11	8.37E-17	1.00E+00	1.10E-02	1.60E-08	2.71E-19	1.00E+00
RHUM	M1 _{rh}	1.05E+01	9.04E-04	2.66E+01	5.80E-03	1	1.24E-13	4.85E-19	5.80E-03	6.40E-05	9.25E-11	1.57E-21	5.80E-03
	M2 _{rh}	8.43E+13	7.27E+09	2.13E+14	4.66E+10	8.03E+12	1	3.90E-06	4.66E+10	5.14E+08	7.43E+02	1.26E-08	4.66E+10
	M3 _{rh}	2.16E+19	1.86E+15	5.47E+19	1.19E+16	2.06E+18	2.56E+05	1	1.19E+16	1.32E+14	1.91E+08	3.23E-03	1.19E+16
	M4 _{rh}	2.09E+03	1.80E-01	5.29E+03	1.00E+00	1.99E+02	2.48E-11	9.67E-17	1	1.27E-02	1.84E-08	3.13E-19	1.00E+00
SLP	M1 _s	1.64E+05	1.41E+01	4.15E+05	9.07E+01	1.56E+04	1.95E-09	7.59E-15	9.07E+01	1	1.45E-06	2.45E-17	9.07E+01
	M2 _s	1.13E+11	9.77E+06	2.87E+11	6.27E+07	1.08E+10	1.35E-03	5.25E-09	6.27E+07	6.91E+05	1	1.70E-11	6.27E+07
	M3 _s	1.69E+22	5.76E+17	6.69E+21	3.69E+18	6.37E+20	7.93E+07	3.09E+02	3.69E+18	4.08E+16	5.89E+10	1	3.69E+18
	M4 _s	1.40E+03	1.21E-01	3.54E+03	1.00E+00	1.33E+02	1.66E-11	6.47E-17	1.00E+00	8.53E-03	1.23E-08	2.09E-19	1

*a: subscript for model developed using AirTemp as predictor**rh: subscript for model developed using RHUM as predictor**s: subscript for model developed using SLP as predictor*

Table 7
Comparison of twelve seasonal rainfall forecast models

$\begin{matrix} M_i \\ M_r \end{matrix}$		AirTemp				RHUM				SLP			
		M1 _a	M2 _a	M3 _a	M4 _a	M1 _{rh}	M2 _{rh}	M3 _{rh}	M4 _{rh}	M1 _s	M2 _s	M3 _s	M4 _s
AirTemp	M1 _a	Wk.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a
	M2 _a	St.E. M1 _a	Wk.E. M2 _a	St.E. M3 _a	Md.E. M4 _a	St.E. M1 _{rh}	St.E. M2 _a	St.E. M3 _{rh}	Md.E. M4 _{rh}	St.E. M2 _a	St.E. M2 _a	St.E. M2 _a	Md.E. M4 _s
	M3 _a	St.E. M1 _a	St.E. M3 _a	Wk.E. M3 _a	St.E. M3 _a	St.E. M1 _{rh}	St.E. M2 _{rh}	St.E. M3 _a	St.E. M3 _a	St.E. M1 _s	St.E. M3 _a	St.E. M3 _s	St.E. M3 _a
	M4 _a	St.E. M1 _a	Md.E. M4 _a	St.E. M3 _a	Wk.E. M4 _a	St.E. M1 _{rh}	St.E. M4 _a	St.E. M4 _a	Wk.E. M4 _a	St.E. M4 _a	St.E. M4 _a	St.E. M3 _s	Wk.E. M4 _a
RHUM	M1 _{rh}	St.E. M1 _a	St.E. M1 _{rh}	St.E. M1 _{rh}	St.E. M1 _{rh}	Wk.E. M1 _{rh}	St.E. M1 _{rh}	St.E. M1 _{rh}	St.E. M1 _{rh}	St.E. M1 _s	St.E. M1 _{rh}	St.E. M3 _s	St.E. M1 _{rh}
	M2 _{rh}	St.E. M1 _a	St.E. M2 _a	St.E. M2 _{rh}	St.E. M4 _a	St.E. M1 _{rh}	Wk.E. M2 _{rh}	St.E. M2 _{rh}	St.E. M4 _{rh}	St.E. M1 _s	St.E. M2 _{rh}	St.E. M3 _s	St.E. M4 _s
	M3 _{rh}	St.E. M1 _a	St.E. M3 _{rh}	St.E. M3 _a	St.E. M4 _a	St.E. M1 _{rh}	St.E. M2 _{rh}	Wk.E. M3 _{rh}	St.E. M3 _{rh}	St.E. M1 _s	St.E. M2 _s	St.E. M3 _{rh}	St.E. M4 _s
	M4 _{rh}	St.E. M1 _a	Md.E. M4 _{rh}	St.E. M3 _a	Wk.E. M4 _a	St.E. M1 _{rh}	St.E. M4 _{rh}	St.E. M3 _{rh}	Wk.E. M4 _{rh}	St.E. M4 _{rh}	St.E. M4 _{rh}	St.E. M3 _s	Wk.E. M4 _{rh}
SLP	M1 _s	St.E. M1 _a	St.E. M2 _a	St.E. M1 _s	St.E. M4 _a	St.E. M1 _s	St.E. M1 _s	St.E. M1 _s	St.E. M4 _{rh}	Wk.E. M1 _s	St.E. M1 _s	St.E. M1 _s	St.E. M4 _s
	M2 _s	St.E. M1 _a	St.E. M2 _a	St.E. M3 _a	St.E. M4 _a	St.E. M1 _{rh}	St.E. M2 _{rh}	St.E. M2 _s	St.E. M4 _{rh}	St.E. M1 _s	Wk.E. M2 _s	St.E. M2 _s	St.E. M4 _s
	M3 _s	St.E. M1 _a	St.E. M2 _a	St.E. M3 _s	St.E. M3 _s	St.E. M3 _s	St.E. M3 _s	St.E. M3 _{rh}	St.E. M3 _s	St.E. M1 _s	St.E. M2 _s	Wk.E. M3 _s	St.E. M4 _s
	M4 _s	St.E. M1 _a	Md.E. M4 _s	St.E. M3 _a	Wk.E. M4 _a	St.E. M1 _{rh}	St.E. M4 _s	St.E. M4 _s	Wk.E. M4 _{rh}	St.E. M4 _s	St.E. M4 _s	St.E. M4 _s	Wk.E. M4 _s

St.E.: Strong Evidence (for example, *St.E. M1_a*: strong evidence for *M1_a*) *Md.E.:* Moderate Evidence *Wk.E.:* Weak Evidence

a: subscript for model developed using AirTemp as predictor

rh: subscript for model developed using RHUM as predictor

s: subscript for model developed using SLP as predictor

4. Conclusion

Seasonal forecast models, with either changing parameters or constant parameters, were developed and tested in this study, using three predictors (air temperature, sea level pressure and relative humidity). Normalized Bayes factors, and graphs of the likelihood of forecasted rainfall under each model, were compared. It was found that the best seasonal rainfall forecast model uses air temperature as the predictor and allows parameter changes according to rainfall magnitude. Thus, seasonal forecast models with changing parameters could be the best for seasonal rainfall forecasting in the Sirba watershed. Indeed, changes in the predictand-predictor relationship according to rainfall amplitude, combined with the Bayesian model selection procedure, appear to be the best technique for forecasting seasonal rainfall in the Sahel.

5. References

- Andersen, I., Dione, O., Jarosewich-Holder, M., Olivry, J.C., 2005. The Niger River Basin: a vision for sustainable management. In: Golitzen, K.G. (Ed.), *Directions in Development. The World Bank*, Washington DC, USA. p. 145.
- Barry, D., Hartigan, J.A., 1993. A Bayesian analysis for change point problems. *J. Am. Stat. Assoc.*, 88, 309-319.
- Beaulieu, C., Ouarda, T.B.M.J., Seidou, O., 2005. Comparative study of homogenization techniques for precipitation data series (in French). Progress report no. 3 (Project on the homogenization of precipitation data), *Ouranos Consortium*, Montreal.
- Beaulieu, C., Seidou, O., Ouarda, T.M.B.J., Zang, X., 2009. Intercomparison of homogenization techniques for precipitation data continued: comparison of two recent Bayesian change point models. *Water Resources Research* Vol. 45. Doi: 10.1029/2008WR007501
- Begert, M., Thomas, S., Walter, K., 2005. Homogenous temperature and precipitation series of Switzerland from 1864 to 2000. *International Journal of Climatology* 25 (1), 65-80
- Biasutti, M., Held, I.M., Sobel, A.H., Giannini, A., 2008. SST forcings and Sahel rainfall variability in simulations of the twentieth and twenty-first centuries. *Journal of Climate*, 21, 3471-3486.
- Bouali, L., 2009. Prévisibilité et prévision statistico-dynamique des saisons des pluies associées à la mousson ouest africaine à partir d'ensembles multi-modèles. Thèse de doctorat, *Université de Bourgogne*, France. 159 pp
- Brankovic, C., Palmer, T.N., 1997. Atmospheric seasonal predictability and estimates of ensemble size. *Monthly Weather Review*, 125, 859-874.
- Caminade, C., Terray, L., 2010. Twentieth century Sahel rainfall variability as simulated by the ARPEGE AGCM, and future changes. *Climate Dynamics*, 35, 75-94.

- Christensen, J.H. et al., 2007. Regional climate projections, in climate change 2007: the physical science basis. In: Solomon S. et al. (eds) Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. *Cambridge Univ. Press*, New York, pp. 847-940
- Descroix, L., Mahe, G., Lebel, T., Favreau, G., Galle, S., Gautier, E., Olivry, J-C, Albergel, J., Amogu, O., Cappelaere, B., Dessouassi, R., Diedhiou, A., Breton, E.L., Mamadou, I. and Sighomnou, D., 2009. Spatio-temporal variability of hydrological regimes around the boundaries between Sahelian and Sudanian areas of West Africa: a synthesis. *Journal of Hydrology*, 375, 90-102.
- Ehsanzadeh, E., Adamowski, K., 2010. Trends in timing of low stream flows in Canada: Impact of autocorrelation and long term persistence. *Hydrological Processes*, 24, 970-980.
- Fearnhead, P., 2006. Exact and efficient Bayesian inference for multiple changepoint problems. *Stat. Comput.*, 16, 203-213.
- Gelfand, A.E., Hills, S.E., Racine-Poon, A., Smith, A.F.M., 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. *J. Am. Stat. Assoc.*, 85, 972-985
- Giannini, A., Biasutti, M., Held, I.M., Sobel, A.H., 2008. A global perspective on African climate. *Climatic Change* 90, 359-383
- Hamatan, M., 2002. Synthèse et évaluation des prévisions saisonnières en Afrique de l'Ouest. DEA Sciences de l'Eau dans l'Environnement Continental. *Université Montpellier 2*. 115pp.
- Hansen, J.W., 2002. Realizing the potential benefits of climate prediction to agriculture: issues, approaches, challenges. *Agricultural Systems*, 74(3), pp.309-330.
- Hansen, J.W. et al., 2011. Review of Seasonal Climate Forecasting for Agriculture in Sub-Saharan Africa. *Experimental Agriculture*, 47(2), pp.205-240.
- Hastenrath, S., 1995. Recent advances in tropical climate prediction. *J. Climate*: 8, 1519-1532
- Hayes, M., Svoboda, M., LeComte, D., Redmond, K., Pasteris, P., 2005. Drought monitoring: New tools for the 21st century, in: Drought and Water Crises: Science, Technology, and Management Issues, edited by: Wihite, D.A., *Taylor Francis*, BocaRaton (LA), 53-69
- Hunt, B.G., 2000. Natural climatic variability and Sahelian rainfall trends. *Global and Planetary Change*, 24, 107-131.
- Ibrahim, B., Karambiri, H., Polcher, J., Yacouba, H., Ribstein, P., 2014. Changes in rainfall regime over Burkina Faso under the climate change conditions simulated by 5 regional climate models. *Clim Dyn* 42:1363-1381
- Janicot, S., Trzaska, S., Poccarr, I., 2001. Summer Sahel-ENSO teleconnection and decadal time scale SST variations. *Climate Dynamics*, 18, 303-320.

- Kumar, A., Hoerling, M., Ji, M., Leetmaa, A., Sardeshmukh, P., 1996. Assessing a GCM's Suitability for Making Seasonal Predictions. *Journal of Climate*, 9, 115-129.
- Lopez-Bustins, J.A., Martin-Vide, J., Sanchez-Lorenzo, A., 2008. Iberia winter rainfall trends based upon changes in teleconnection and circulation patterns. *Global and Planetary Change*, 63, 171-176.
- Lund, R., Reeves, J., 2002. Detection of undocumented changepoints: a revision of the two-phase regression model. *Journal of Climate*, 15, 2547-2554
- Mara, F., 2010. Développement et analyse des critères de vulnérabilité des populations sahéliennes face à la variabilité du climat: le cas de la ressource en eau dans la vallée de la Sirba au Burkina Faso. Thèse de doctorat. *Université du Québec à Montréal*. 273pp
- Min, S., Simonis, D., Hense, A., 2007. Probabilistic climate change predictions applying Bayesian model averaging. *Phil. Trans. R. Soc. A*, 365: 2103-2116
- Mohino, E., Janicot, S., Bader, J., 2011. Sahel rainfall and decadal to multi-decadal sea surface temperature variability. *Climate Dynamic*, 37: 419-440.
- Ogallo, L.A., Boulahya, M.S., Keane, T., 2000. Applications of seasonal to interannual climate prediction in agricultural planning and operations. *Agricultural and Forest Meteorology*, 103, pp.159-166.
- Palmer T.N., 1986. Influence of the Atlantic, Pacific and Indian Oceans on Sahel rainfall. *Nature*, 322, 251-253.
- Palmer, T.N., Alessandri, U., Andersen, P., Cantelaube, M., Davey, P., Délécluse, Déqué, M., Diez, E., Doblas-Reyes, J. F., Feddersen, H., Graham, R., Gualdi, S., J. F. Guérémy, J. F., Hagedorn, R., Moshen, M., Keenlyside, N., Latif, M., Lazar, A., Maisonnave, E., Marletto, V., Morse, A.P., Orfila, B., Rogel, P., Terres, J.M., Thomson, M.C., 2004. Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bulletin of the American Meteorological Society*, 85, 853-872.
- Palmer, T.N., Brankovic, C., Richardson, D.S., 2000. A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations. *Quarterly Journal of the Royal Meteorological Society*, 126, 2013-2034.
- Perreault, L., Bernier, J., Bobee, B., Parent, E., 2000. Bayesian change-point analysis in hydrometeorological time series 2. Part 2. Comparison of change-point models and forecasting. *J. Hydrol.*, 235, 242-263
- Philippon, N., Fontaine, B., 2002. The relationship between the Sahelian and previous 2nd Guinean rainy seasons: a monsoon regulation by soil wetness. *Annales Geophysicae*, 20, 575-582.
- Reeves, J., Chen, J., Wang, X.L., Lund, R., Lu, Q., 2006. A review and comparison of change point detection techniques for climate data. *Journal of Applied Meteorology and Climatology*, 46, 900-914.
- Rowell, D.P., 2001. Teleconnections between the tropical Pacific and the Sahel. *Quarterly Journal of the Royal Meteorological Society*, 127, 1683-1706.

- Rowell, D.P., 2003. The impact of Mediterranean SSTs on the Sahelian rainfall season. *Journal of Climate*, 16, 849-862.
- Samimi, C., Fink, A.H., Paeth, H., 2012. The 2007 flood in the Sahel: causes, characteristics and its presentation in the media and FEWS NET. *Natural Hazards and Earth System Sciences*, 12: 313-325.
- Sarr, M.A., Zoromé, M., Seidou, O., Bryant, C.R., Gachon, P., 2013. Recent trends in selected extreme precipitation indices in Senegal-A changepoint approach. *Journal of Hydrology*, doi: <http://dx.doi.org/10.1016/j.jhydrol.2013.09.032>
- Schepen, A., Wang, Q.J., Robertson, D.E., 2012. Combining the strengths of statistical and dynamical modeling approaches for forecasting Australian seasonal rainfall. *Journal of Geophysical Research*, 117, 148-227.
- Seidou, O., Asselin, J.J., Ouarda, T.B.M.J., 2007. Bayesian multivariate linear regression with application to change-point models in hydrometeorological variables. *Water Resour. Res.*, doi: 10.1029/2005WR004835
- Seidou, O., Ouarda, T.B.M.J., 2007. Recursion-based multiple changepoint detection in multiple linear regression and application to river streamflows. *Water Resources Research*, Vol. 43, W07404, doi: 10.1029/2006WR005021
- Sittichok, K., Gado Djibo, A., Seidou, O., Saley, H.M., Karambiri, H., Paturel, J., 2014. Statistical seasonal rainfall and streamflow forecasting for the Sirba watershed, using sea surface temperature. *Hydrological Sciences Journal*. DOI: 10.1080/02626667.2014.944526
- Tarhule, A., 2005. Damaging rainfall and flooding: the other Sahel Hazards. *Climatic Change*, 72(3), pp.355-377.
- Taweye, A., 1995. Contribution à l'étude hydrologique du bassin versant de la Sirba à Garbé-Kourou. Dissertation, Centre Régional AGRHYMET, 96pp.
- Thiaw, W., Barnston, A.G., Kumar, V., 1999. Predictions of African rainfall on the seasonal time scale. *J. Geophys. Res.* 104, 31589-31597
- Villarini, G., Smith, J.A., Serinaldi, F., Ntelekos, A.A., 2011. Analyses of seasonal annual and maximum daily discharge records for central Europe. *Journal of hydrology* 399, 299-312
- Vincent, L.A., 1998. A technique for the identification of inhomogeneities in Canadian temperature series. *J. Clim.*, 11, 1094-1105.
- Xiong, L., Guo, S., 2004. Trend test and change-point detection for the annual discharge series of the Yangtze River at the Yichang hydrological station. *Hydrological Sciences Journal* 49 (1), 99-112

5.5. Regression Trees and Neural Network Seasonal Forecasts

5.5.1 Regression tree for seasonal rainfall forecasting

Regression tree (RT) models were used to improve the seasonal forecasting skills, as such models have achieved the best results in several climate forecasting studies (Kannan & Ghosh, 2010; Razi & Athappilly, 2005). Three types of RTs are tested to develop the seasonal rainfall forecasting model: extremely randomized tree, random forest, and bagging tree.

- ***Extremely randomized tree***

With this method, the RT algorithm is applied with the default conditions (i.e., without constraints). During model development a single regression tree is grown, then the default parameters are set to grow an ensemble of extremely randomized trees. Despite rigorous parameterization, this technique explains the training pattern of the data, and easily obtains the predicted values without a high computational time. The tree is pruned, and the number of nodes and leaves are continuously adjusted until an acceptable mean square error value is reached. If tree pruning is disabled, only filling with models (use of average values) is done. For each model, sub-set selection is performed using backward selection. However, due to the short dataset, LOOCV is applied while constructing the forecasting model.

- ***Random forest***

This method can help construct an ensemble of trees, with the number depending on the size of the dataset; 100 trees are used in this work. The predictand is forecast under a defined algorithm, and the Breiman random forest algorithm is adapted. The algorithm attempts to improve the forecasting ability by de-correlating the trees, with each tree having the same probability. At each tree split, a random sample of m features is drawn ($m = \sqrt{p}$, where p is the number of features), and only the selected features are considered for splitting. There is an error in each bootstrap sample because an individual tree is built on the sample, so this error rate (also called out-of-bag error rate) is checked. LOOCV is also applied in this study due to the span of available data.

- **Bagging tree**

The bagging technique is used in this approach when developing the RT. It consists of using average trees over many samples to reduce the variance. The common criterion used for splitting is: *“a node is not furthermore divided once the standard deviation of the output variable data at the node is inferior to the split threshold of the standard deviation of the output variable data of the entire initial data series”*. The default values defined in the algorithm by Wang and Witten (1997) are used to prune the tree. In applying the LOOCV, the number of k -fold cross-validations is set to equal the length of the dataset of the predictand.

The three techniques of the RT method gave poor results. The best R^2 value was 0.19 (for RHUM), while all ANSH values were negative (e.g., -0.40 for UWND). These values highlight that the three RT techniques were unable to deal with the variance. Figures 5.1 and 5.2 show the observed and forecasted rainfall using RHUM and AirTemp respectively. Analyzing the graphs, it is identified a persistent over-fitting in all RT models, even when the bagging training technique and LOOCV were used to overcome the problem. This could be due to the fact that RT methods are more suitable for long data series that enable k -fold cross-validation, while the data series used in this work spans a small period. Forecasting models built with RT are often subject to overprediction unless long datasets or good training techniques are used. Another factor could be the use of time-averaging (i.e., daily rainfall data to seasonal data), which can mask underlying non-linear relationships. Yuval and Hsieh (2002) showed that time-averaging can have a significant impact on the detectability of non-linear relationships.

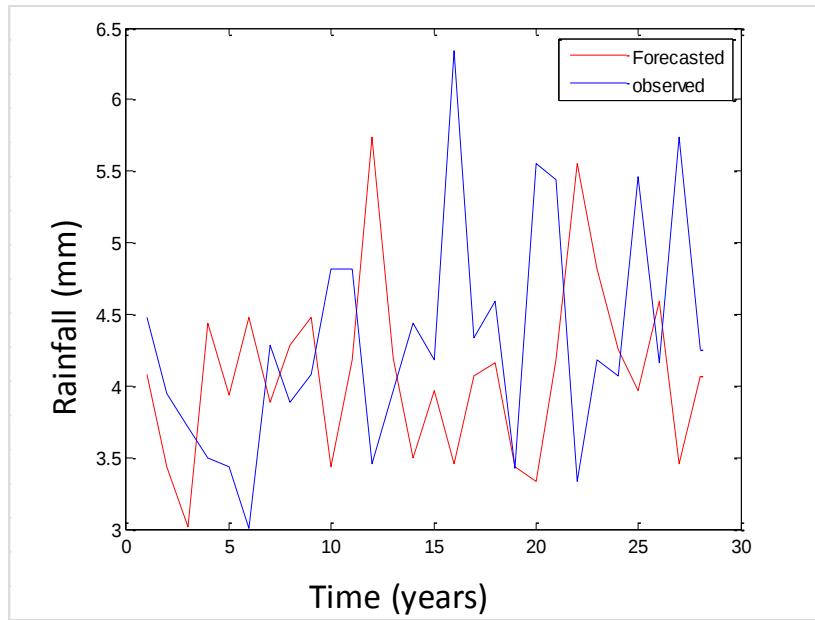


Figure 5.1 Observed and simulated seasonal rainfall based on RT using Relative Humidity

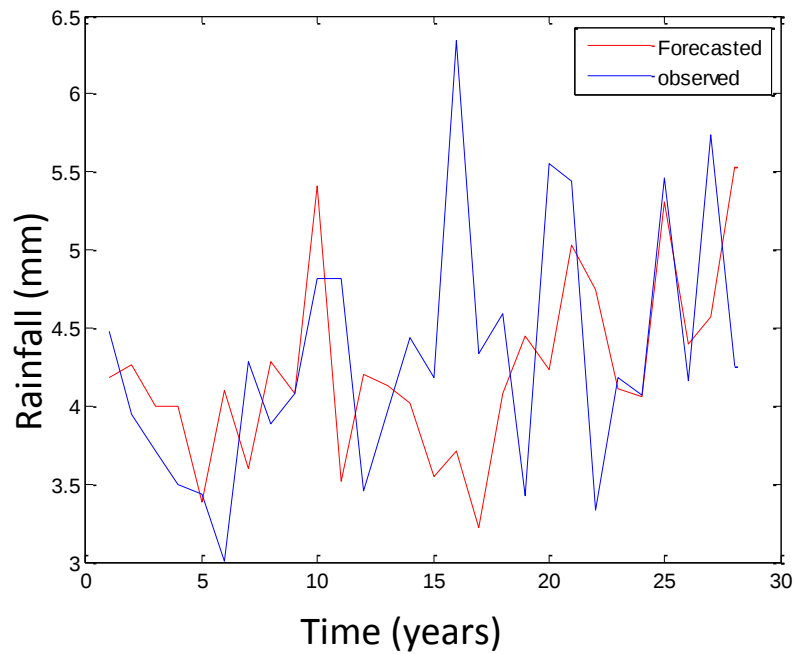


Figure 5.2 Observed and simulated seasonal rainfall based on RT using Air Temperature

5.5.2 Neural network seasonal rainfall forecasting

The feedforward neural network (FFNN) was also tested in this work because it is capable of representing non-linear processes between input and output variables, and also, non-linearity and other interactions between variables can be modeled without prior specification (Tangang *et al.*, 1998; Canon & Mckendry, 1999). After screening the predictor, it was introduced into an FFNN to predict seasonal rainfall. In applying the FFNN, some descriptive variables were first set, and the batch gradient descent with momentum (BGDM) algorithm was used as a training function. This back-propagation algorithm is more suitable and is used in 90% of artificial neural network models in the field of hydrology and climate forecasting (Badr *et al.*, 2014; El-Shafie *et al.*, 2011; Ansari, 2013; Hung *et al.*, 2008). A sensitivity analysis was carried out prior to the forecast process, because it allows to understand which input has the most significant impact on the outputs after removing each input element one by one. This sensitivity analysis was based on the architecture of the neural network. Two sets of FFNN models were developed and tested: one set using only one predictor, and a second type using a combination of predictors. The characteristics of the first-developed FFNN model are as follows:

- 1 input layer containing 1 entry cell composed of one predictor
- 1 hidden layer C_1 with a hyperbolic tangent sigmoid (tansig) as transfer function. The cells are connected to those at the input as well as to the bias through the weights
- 1 output layer C_2 having only 1 cell due to the use of only one predictand

In the second FFNN model, where predictors were combined, the configuration of the hidden neurons was made supplier due to the nature of the forecast. The network architecture was:

- 1 input layer with two neurons representing two predictors
- 2 hidden layers K_1 and K_2 with hyperbolic tangent sigmoid (tansig) and Log-sigmoid (logsig) as transfer functions, respectively
- 1 output layer K_3 with one neuron representing the forecasted rainfall

The results obtained for the FFNN model were disappointing given the reported high ability of FFNN to deal with non-linearities (see Table 5.1). The FFNN architecture

defined gave the best potential output. For the first set of FFNN models, it was found that the highest R^2 and NASH values were 0.26 and 0.20, respectively. A similar result was obtained by Philippon (2002), who obtained an R^2 value of 0.45 after applying a multi-perceptron neural network on Central and West African seasonal rainfall. In the second FFNN model, the best combination was RHUM and SLP together, with an $R^2 = 0.37$ and NASH = 0.28, respectively. This result was achieved after 1000 epochs and a mean square error of 10^{-4} . Figure 5.3 shows the observed and simulated rainfall. A possible reason for these low coefficients is that the number of observations is limited (31 years). One more factor is that non-linear models have more parameters than linear models, and therefore require more and better information in the calibration process.

Table 5.1 FFNN model (single predictor) output for Sirba seasonal rainfall forecast

Predictors	R^2	Nash	HIT score (%)	Lag time (months)
AirTemp	0.26	0.20	48.24	4
RHUM	0.18	0.10	29.12	4
SLP	0.21	0.09	18.03	2
SST	0.18	0.044	11.49	5

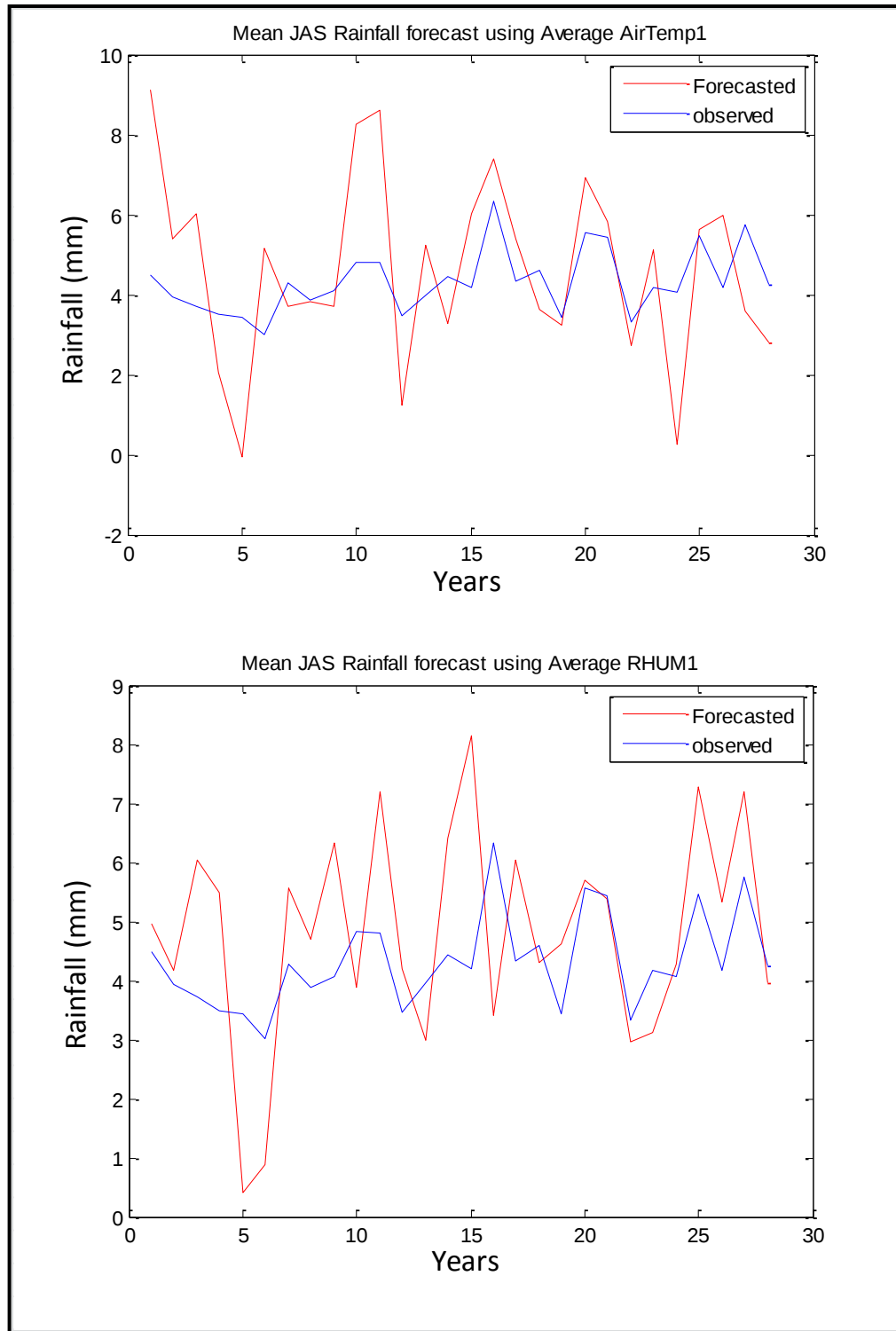


Figure 5.3 Observed and simulated seasonal rainfall based on FFNN using AirTemp (upper panel) and RHUM (lower panel)

5.6. Synthesis and Partial Conclusion

Seasonal rainfall forecasting was carried out, and was primarily based on the identification of the best predictors and the optimal lag time. The best predictors were selected through a process based on combined linear statistical methods. These techniques helped in screening a pool of predictors having physical insight with the WAM. It was determined that AirTemp, RHUM and SLP were the best predictors among the group with lag time varying between 5 to 8 months. Two groups of models were developed and tested using the obtained predictors.

The first group is composed of seasonal forecast models, with either changing parameters or constant parameters, which were developed and tested using three predictors (air temperature, sea level pressure and relative humidity). The use of normalized Bayes factors and graphs of the likelihood of forecast rainfall enabled the finding that the seasonal rainfall forecast model using AirTemp as a predictor and allowing parameter changes according to rainfall magnitude performs better than the others. Thus, seasonal forecast models with changing parameters could be the best for seasonal rainfall forecasting in the Sirba watershed. Indeed, allowing changes in the predictand–predictor relationship according to rainfall amplitude, combined with the Bayesian model selection procedure, appear to be a good technique for forecasting seasonal rainfall in the Sahel. Regarding the performance of the models under each of the three predictors, it is noted that for the models using AirTemp as predictor, the strongest model is $M1_a$, followed by $M3_a$, $M4_a$ and $M2_a$. For models using RHUM as predictor, the performance of the models in decreasing order is $M1_{rh}$, $M4_{rh}$, $M2_{rh}$, and $M3_{rh}$. Lastly, for SLP, the performance of each model is shown by the inequalities $M4_s > M1_s > M2_s > M3_s$.

The second group of models is a set of non-linear statistical methods, namely, NLPCA, FFNN and RT. The NLPCA, compared to the two other methods, performed satisfactorily with regard to the short study period considered. The NLPCA with AirTemp (R^2 : 0.46; NASH: 0.45; HIT: 60.7%) was the best, followed by RHUM and SLP, respectively. It was also found that this method provides a too larger lag time. The FFNN model performed poorly based on the obtained results, and this was mainly due to the short length of the data, which was not enough to train the model. The performance of

the regression tree model was the worst, and this kind of non-linear model needs more data for the pruning process in order to achieve better results.

In conclusion, these seasonal forecasting models helped in understanding that non-linear methods can be used instead of the usual linear methods. The specificity of this work is the use of other predictors rather than the SSTs which gave acceptable results than the SSTs.

However, the limit of the approach resides on the short length of data used. Therefore, in order to generalize the results for other work, it is recommended that during the forecast a Sahelian global index be constructed using CRU data (satellite rainfall) while examining its correlation with the index of the watershed and the ability of models.

Chapter 6. Seasonal Streamflow Forecasting

The development of seasonal streamflow forecasting models based on changing and constant parameters and the Sirba SWAT hydrological model are presented in this chapter. The chapter also details the different seasonal streamflow forecasting models using non-linear model.

6.1. Sirba SWAT Hydrological Model

6.1.1 Justification and method

The SWAT model was used in this study to develop the Sirba hydrological model. This model was chosen due to its high diversity of hydrological processes. The selection is justified by its several advantages, as follows. SWAT:

- is one of the most broadly applied watershed-scale simulation tools
- is used around the world to address watershed questions
- is used in hundreds of scientific studies published worldwide
- has the capability to predict the effect of soil, land use and management on water
- can represent changes in management as actual processes (i.e., physically based)
- is designed to use a realistic number of inputs (readily-available input)
- is open source, well documented and computationally efficient

SWAT is a physically based model developed by the USDA Agriculture Research Service (Arnold *et al.*, 1998), and is used to simulate river flows and other processes (e.g., sediment transport, nutrients and pesticides cycle, bacterial growth). It helps deal with environmental issues, which are the most critical challenging aspects of hydrological model development, particularly in the West African Sahel. Hence, SWAT is one of the most comprehensive models and can take almost all the components required for hydrological model development into account. It enables the prediction of the influence of land management practices related to sediment, water and agricultural chemical yields

in substantial basins. Thus, it is a great tool to ensure food security in the Sahelian regions with respect to how well land management practices can be modelled.

The model was set to use the SCS curve number method (Soil Conservation Service Engineering Division 1972) to calculate surface runoff, and the Penman–Monteith method for potential evapo-transpiration estimation. The Arc-SWAT interface (Olivera *et al.* 2006) was used for project development. The Sirba watershed and sub-watersheds are delineated using a 90m resolution digital elevation of West Africa, downloaded from the Shuttle Radar Topography Mission (SRTM, 2004). After the delineation (using the automatic DEM-based method), the watershed is divided into nine sub-basins (see Figure 6.1) according to the topography in order to perform the SWAT model spatial parameterization.

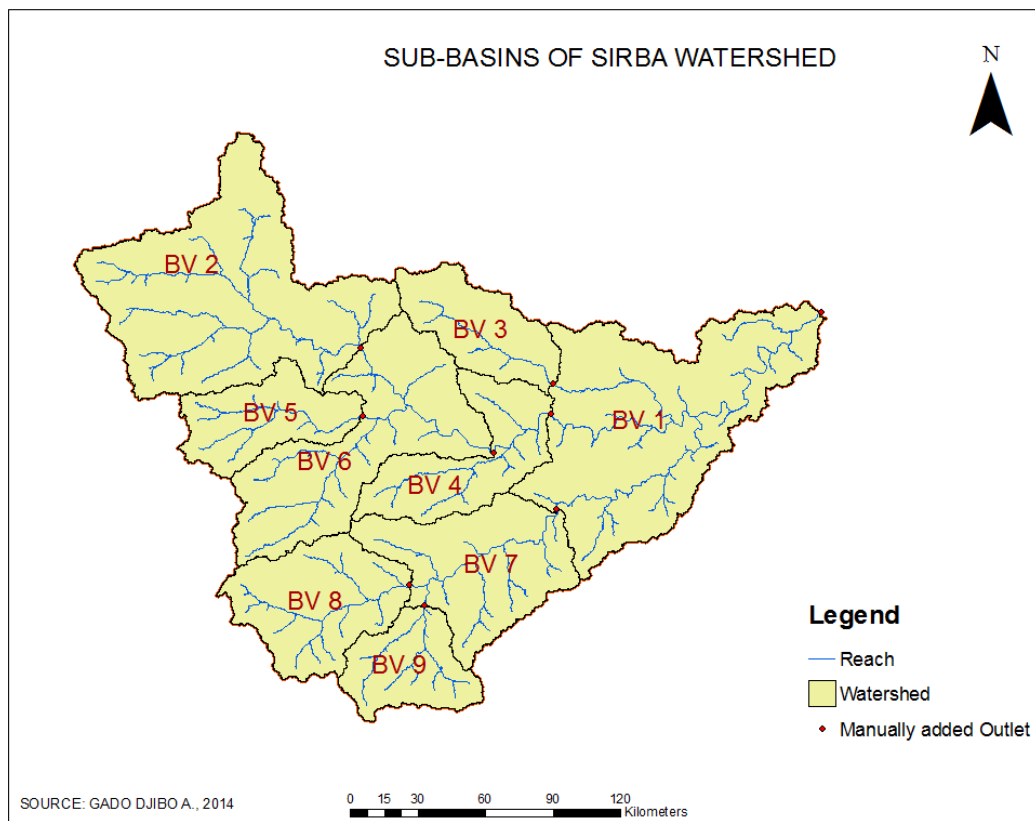


Figure 6.1 Sirba sub-watersheds from SWAT

Before the SWAT model is set up, the distribution of HRUs in the sub-watersheds is determined, based on the land use and soil data layers specified in the previous step. The HRUs can be used to assess the varying hydrologic conditions between sub-watersheds. By subdividing the watershed into areas with unique land use and soil combinations, the model can reflect differences in evapo-transpiration and other hydrological conditions for various soils and crops. With the SWAT model, the runoff is first forecast per HRU, then routed to obtain the total runoff for the basin. This technique increases the accuracy and provides a higher quality physical description of the water balance. The Sirba watershed is set to have five reservoirs, 64 HRUs, and one outlet at Garbé-Kourou. The characteristics of the dams (i.e., surface and area of the lake at various levels) are specified in the model reservoir routing parameters.

Rainfall and streamflow data, as described in Section 4.3.1, were used to develop the Sirba hydrological model. For other climate parameters, NCEP (National Centers for Environment Predictions) reanalysis is used to find a dataset of wind speed, relative humidity, and average temperature on a daily scale. The data for a given watershed is taken at the grid-point of the NCEP data, which is closest to the watershed centre. Since maximum and minimum temperatures are not available in the NCEP database, they were set as equal to the mean temperature in all SWAT simulations. The daily flow data at the hydrological station of Garbé-Kourou, the area-volume relationship, and water levels at different dams are acquired from the DGARH and DGRE.

Due to the relatively short span of the observation data and the large number of parameters of the model, only the sensitivity analysis and calibration are performed over the entire period. With SWAT, sensitivity analysis is done based on changes of the mean of the output variable (flow). If the observed data for the output variable are available, additional sensitivity analysis is done using the sum of the squared error between the observed and simulated values. Using the available observed flow at Garbé-Kourou station, additional sensitivity analysis is performed using the SWAT-CUP software (Abbaspour, 2007) by including all parameters that are likely to influence the objective function. The model was run 270 times during sensitivity analysis. The objective function in the sensitivity analysis is the Nash coefficient, calculated with the simulated and observed stream flow at Garbé-Kourou.

For the calibration step, the objective function selected was the Nash coefficient. The calibration was performed using the SWAT-CUP interface, as well as the parameters that have the most effect on the streamflow. The SUIF2 algorithm was applied to calibrate the discharge at the outlet of the watershed.

The parameters were adjusted after each simulation, and the model output was compared to the observed values. The sub-watershed and HRUs were then chosen wherever a change was applied. This process can be done for many parameters in one simulation. After each calibration process, the NASH "goodness-of-fit" was applied to the corresponding assessment. The parameters with the most influence on the streamflow at Garbé-Kourou were used repeatedly during the calibration of the model, and a graphical analysis of the observed and forecasted streamflows was evaluated. The calibration was ended once the NASH coefficient was judged as satisfactory (i.e., when the curves of the observed and simulated flows match more closely on the hydrograph).

6.1.2 Sirba hydrological model

Land use and soil coverages were developed and classified during the development of the Sirba hydrological model, resulting in the identification of six classes of land use: pasture (23.5%), cropland/woodland mosaic (24%), shrubland (0.1%), savanna (51%), barren or sparsely vegetated (0.12%), and water bodies (1%). Three types of slopes and one soil class (sandy loam: Lc13-1a-127) were also identified. These results are described in Figures 6.2 and 6.3.

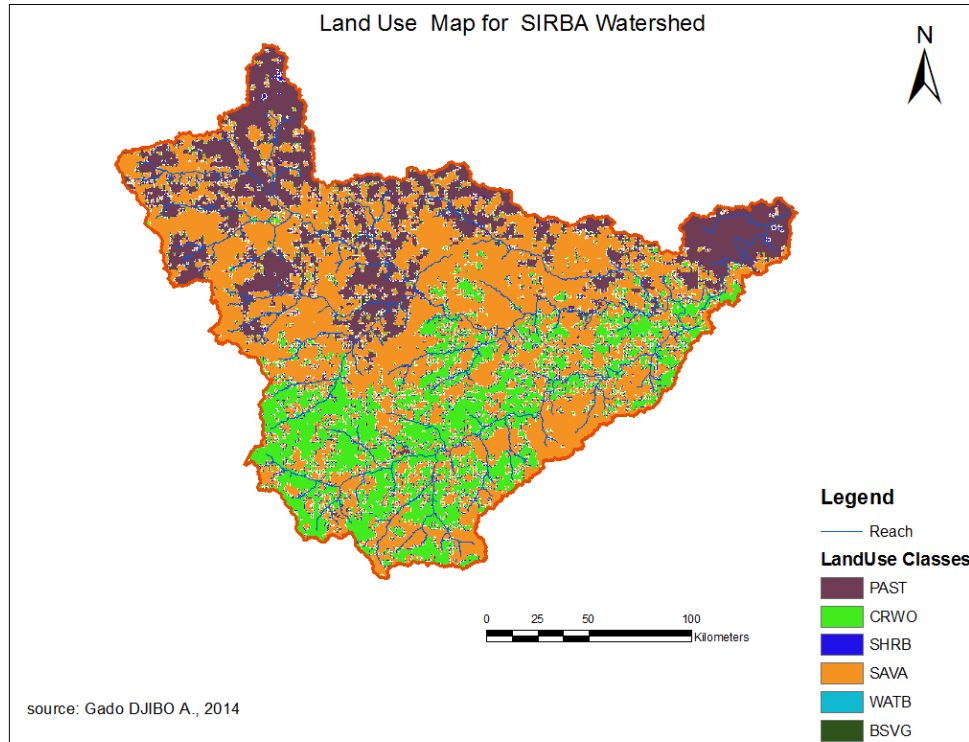


Figure 6.2 Land use classes of the Sirba SWAT hydrological model

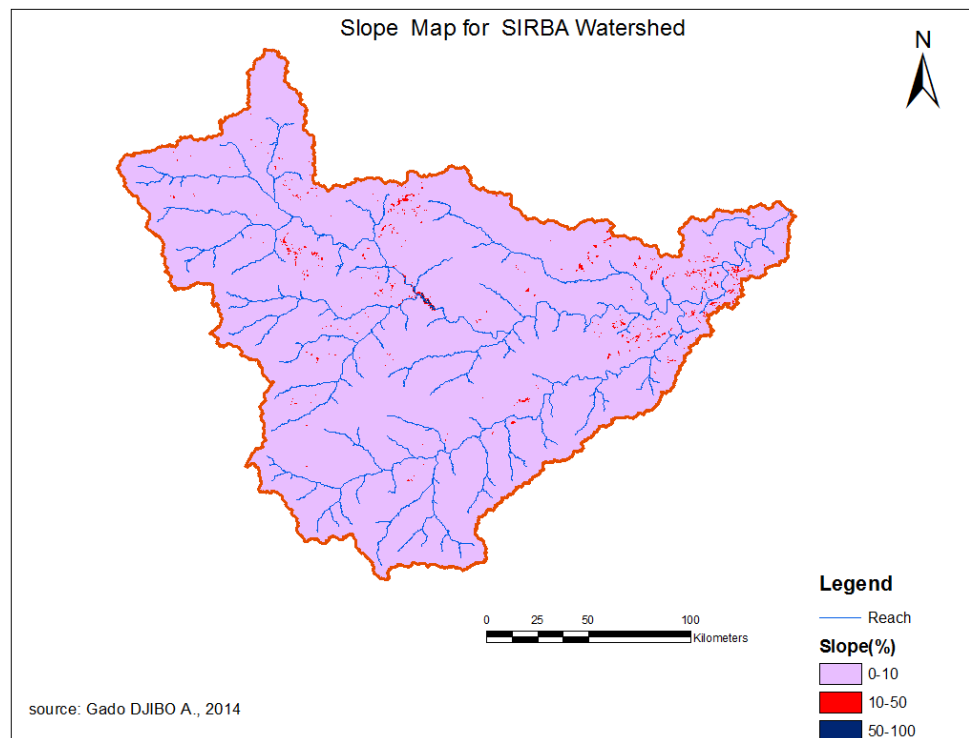


Figure 6.3 Slope classes of Sirba SWAT hydrological model

The sensitivity of all SWAT parameters was obtained using the SUFI2 algorithm implemented in SWAT-CUP. This algorithm was founded on a t-test, and allows the removal of less sensitive parameters using their p-value as criteria. A parameter is less sensitive (i.e., not likely to affect streamflow) if its p-value is greater than 0.5. The number of parameters from the sensitivity analysis and their respective ranks are presented in Table 6.1.

Table 6.1 Ranks of sensitive parameters of the Sirba hydrological model

Rank	Parameter	Rank	Parameter	Rank	Parameter
1	Cn2	10	Surlag	19	Biomix
2	Esco	11	Epco	20	Ch_N2
3	Sol_Awc	12	Gw_Delay	27	Sftmp
4	Sol_Z	13	Slope	27	Smfmn
5	Gwqmn	14	Sol_K	27	Smfmx
6	Ch_K2	15	Alpha_Bf	27	Smtmp
7	Canmx	16	Ssubbsn	27	Tlaps
8	Revapmn	17	Sol_Alb	27	Timp
9	Blai	18	Gw_Revap		

These results show that the most sensitive parameters are Cn2, Esco and Sol_Awc. The first important parameter is the SCS curve number for moisture condition II, which controls how precipitation is split in runoff and infiltration. The second is the soil evaporation compensation factor which affects soil prone to evaporation, measured in evapo-transpiration; the higher the ESCO, the lower the evaporation. The last parameter is the water capacity of soil layers: the higher the Sol_AWC, the less water available.

However, these are outputs of mathematical analysis, and may not be representative of the true physical state of the watershed being modeled.

Calibration of models on a watershed scale is challenging due to potential uncertainties in process simplification, processes not accounted for by the model, and processes in the watershed unknown to the modeler.

After sensitivity analysis, only the parameters ranked very sensitive were selected for calibration, based on their p-values. The ranges and description of parameters in the final calibrated SWAT model are given in Paper 1.

After calibration, the simulated monthly hydrograph reproduced the observations well, with a NASH value of 0.83.

The uncertainties in this model could be due to inputs such as rainfall and temperature. Rainfall and temperature data are measured at local stations, and regionalization of these data can introduce significant errors; if an anomalous site is used, runoff results could be skewed high or low. With SWAT, climate data for every sub-basin are provided by the station nearest to the centre of the sub-basin. The direct accounting of rainfall or temperature distribution errors is quite difficult, as information from many stations is required. However, the "elevation band" option in SWAT can decrease this error by adjusting the temperature and rainfall to account for the orographic effects of a sub-basin.

Given these possible sources of error, and the NASH values from similar studies in the West African Sahel (see Paper 1), the calibration results of the Sirba SWAT hydrological model in this study can be considered acceptable.

6.2. Statistical seasonal streamflow forecasting using probabilistic approach over West Africa

The seasonal streamflow forecasting was based on two set of models: models with changing parameters and models with constant parameters. The models' development and the achieved results were published in Gado *et al.* (2015c) as detailed below.

Statistical seasonal streamflow forecasting using probabilistic approach over West African Sahel

Abdouramane Gado Djibo^{1,2} · Harouna Karambiri¹ ·
Ousmane Seidou² · Ketevera Sittichok² ·
Jean Emmanuel Paturel³ · Hadiza Moussa Saley⁴

Received: 3 April 2015 / Accepted: 6 June 2015
© Springer Science+Business Media Dordrecht 2015

Abstract

Runoff changes are tightly connected to precipitation in West African Sahel in such a way that any impact on precipitation would result in potential changes in runoff. Unfortunately, climate change and variability impacts induced changes in streamflow which directly disturb water availability for socioeconomic activities particularly agricultural sector which constitutes the main survival issue of West African population. Thus, available streamflow information a few months in advance prior to a rainy season with an acceptable forecasting skill will immensely benefit for water users to make an operational planning for water management decision. Streamflow is usually either forecasted directly by linking streamflow to predictors through a multiple linear regression or indirectly using a rainfall-runoff model to transform predicted rainfall into streamflow. Seasonal annual mean streamflow and maximum monthly streamflow were forecasted in this study by using two statistical methods based on change point detection using Normalized Bayes Factors. Each method uses one of the following predictors: Sea Level Pressure (SLP), Air Temperature (AirTemp) and Relative Humidity (RHUM). Models M1 and M2 respectively allow for change in model parameters according to rainfall amplitude (M1), or along time (M2). They were compared to forecasting models where precipitation is obtained using the classical linear model with constant parameters (M3) and the climatology (M4). The obtained results revealed that model M3 using RHUM as predictor at a lag time of 8 months was the best method for seasonal annual streamflow forecast. Whereas, model M1 using air temperature as predictor at a lag time of 4 months is the best model to predict maximum monthly streamflow in the Sirba watershed.

Keywords: Streamflow forecast, Bayes factor, SWAT, Posterior probability, Sirba watershed, Sahel

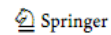
¹ International Institute for Water and Environmental Engineering (2iE), 01 BP 594 Ouagadougou, Burkina Faso

² Department of Civil Engineering, University of Ottawa, Ottawa, ON, Canada

³ Institut de Recherche pour le Développement (IRD), Abidjan, Côte d'Ivoire

⁴ Centre Africain d'Études Supérieures en Gestion, Dakar, Sénégal

Published online: 18 June 2015



1. Introduction

Water availability in West African Sahel region is important for ensuring human health, empowering economic activities and developing the ecosystem. But, climate change and variability had potential pressures on water resources, making water availability and accessibility very difficult (IPCC 2007). Rowe et al. (1994) showed that an increase of 2° in temperature associated to a 10% decrease in precipitation would reduce runoff by 40%. A similar study carried on the Black Volta basin indicated a much higher decrease in river flows due to climate changes and variability (Wolfram et al., 2008). Also, Roudier et al. (2014) revealed that runoff changes are tightly correlated to precipitation in West Africa and any impact on precipitation would result in potential changes in runoff. Therefore, the Sahelian streamflow is rainfall dependent which is directly based on West African Monsoon (WAM). Unfortunately, the WAM is not completely understood by climatologists and this fact makes forecasts at all scales difficult due to uncertainties. The uncertainty in the forecasts directly affects local populations. Indeed, lack of awareness of the evolution of streamflows often disturb water availability for socioeconomic activities (e.g., human consumptive, industrial, recreational uses) and particularly agricultural sector which constitutes the main survival issue of West African population. This is exacerbated by droughts affecting streamflows (Samimi et al. 2012; Brooks 2004) and sometimes leading to inundations, destruction of roads and dams, loss of agricultural production and so on. All these adverse effects surprise the decision makers and the population despite the existence of seasonal streamflow forecast bulletin for major rivers in West Africa over years (Hamatan, 2004). This forecast bulletin is provided by African Centre of Meteorological Application for Development (ACMAD) and Agriculture, Hydrology and Meteorology Research Center (AGRHYMET). In such conditions, any forecast information on streamflow trend would be a good decision tool for authorities and socio-economic actors. This local climate forecast information could improve

agriculture and other economic activities which help ensuring food security for West Africa Sahelian population. Thus, the development of seasonal streamflow forecast models is greatly looked forward and cheered by all actors, to enable an efficient use of climatic information for food security. Such streamflow forecast model could increase resilience to climate variability by providing advanced information about the expected amount of runoff in the next rainy season.

The scientific community made several efforts in that direction, which are based on two approaches in forecasting streamflow on a watershed: direct method and indirect method. The first approach consists to use some relationships (usually empirical) between streamflow at a given date and observed predictors at a period prior to that date. While, in the second method, the rainfall is first predicted using predictors and then fed into a rainfall-runoff model to get streamflow.

In the first approach, statistical methods based on empirical relationships between historical streamflows (predictand) and hydrological or climate variables (predictor) are used. Most research in the Sahel focused on empirical relationships used between climatic and oceanic variables with rainfall. Therefore, several studies were carried out to find the appropriate statistical techniques and the best predictors to forecast streamflows for specific regions (Abudu et al. 2010; Kwon et al. 2009; Piechota et al. 1998; Cayan and Riddle 1993). This first method is rarely found especially in the case of Sahelian streamflow forecast studies, as majority of studies in the literature used the second approach (i.e., indirect method). Nevertheless, it is assumed in the existing few studies that streamflow is linked to predictors by a multiple linear regression with parameters that are independent of time and of predictor magnitude, though this fact is not really established.

Regarding the indirect method, statistical methods (Sittihok et al. 2014; Garric et al. 2002; Barnston et al. 1996; Washington and Downing 1999; Folland et al. 1991) or climate models (Thiaw and Mo 2005; Rowel et al. 1992; Folland et al. 1991) are first used to forecast rainfall before using a runoff-rainfall model. However, it is noted that biases are usually introduced by climate models. But such biases are corrected using model output statistics techniques (Amadou et al. 2014; Ndiaye et al. 2011; Bouali 2009). In order to forecast seasonal streamflow, the predicted rainfall is used to run the rainfall-

runoff model (Garbrecht et al. 2007; Tuteja et al. 2011; Yossef et al. 2013; Ndiaye et al. 2009). This approach has several advantages as it is a more realistic simulation of streamflow since physical interactions between land, climate and river characteristics in the watershed are accounted for. Some authors (Wang et al. 2011; Yossef et al. 2013) show that rainfall-runoff models can produce consistent streamflow once they are forced with accurate rainfall forecast and properly initialized using antecedent hydrological variables before the forecast issued. One more advantage is that, rainfall-runoff models are less sensitive to outliers in the input data compared to statistical models (i.e., direct method). Nevertheless, the beneficial use of flow simulations apparently depends on the reliability of outputs and the lagged-time (in advance) flow can be modeled. However, it should be noted that the number of studies dealing with seasonal streamflow forecasting is far less than those dealing with seasonal rainfall forecasting despite the great needs in this Sahelian region to have a prior knowledge of the quantity of runoff in order to plan socio economic activities against adverse effects.

As explained previously, most studies arbitrarily considered the relationship between the predictors and the predictand (streamflow in the Sahel) to be independent of time and streamflow magnitude. The aim of this paper is to test statistical seasonal streamflow forecasting models with changing parameters, and to compare the new models to the classical linear model with constant parameters and to the climatology. This helps to check either the usual seasonal forecast skills can be improved or not, and to find the best model to predict annual mean streamflow and maximum monthly streamflow. The performance of the new models is compared with that of the original linear model and that of a model representing the streamflow climatology in the study area.

The Sirba is a significant and essential tributary of the Niger river, which drains water from the Burkina Faso zone up to Garbé-Kourou (in Niger Republic), where it flows into the Niger river. The hydrological behavior of the Sirba has promoted several studies in different areas. The Sirba basin is characterized by a predominance of mostly rural sector activities (e.g. agriculture, forest resource exploitation, artisanal gold mining), and is a watershed with diverse socio-economic entities that can be representative samples for questions concerning climate issues (Mara, 2010). The basin is also among the regions most affected by fluctuations in rainfall regimes. Indeed, the area is naturally

characterized by scarce and intermittent spatial precipitation distribution, and a very high evaporation rate. It is prone to recurrent droughts, and continuous wind erosion, as well as heavy rainfall events during the monsoon season (Taweye, 1995). These factors, along with over-exploitation of the fragile natural ecological systems, could lead to irreversible changes and damage potential agriculture production (Gachon et al., 2007; Janicot et al., 1996). It has also been observed that recurring drought in the region in recent decades has led to severe limitations in socio-economic development, particularly in rural areas where the economy depends on rainfed agriculture. The Burkina Faso IWRM (Integrated Water Resources Management) program of 2001 shows that the Sirba basin is predominately characterized by successive droughts, rainfall deficits, violent winds and extreme temperatures. All these facts make a completing case for the development of skillful seasonal forecasting models for the in the Sirba.

2. Materials and methods

2.1 Study area

The study area considered in this work is the Sirba watershed. This transboundary watershed, which is shared by Burkina Faso and Niger, is located between latitudes 12°55'54" - 14°23'30" N and longitudes 1°27' W - 1°23'42" E with an area of 38750 km² (Fig. 1). The Sirba has three sub-climate zones based on the amount of rainfall decreasing from south to north: a southern Sudanian zone with mean annual rainfall between 700 and 800mm, a northern Sudanian zone with mean annual rainfall ranging from 550 to 650mm and a Sahelian zone with mean annual rainfall of 300 to 500mm. The largest quantities of rainfall are observed during the months of July to September (JAS), regardless of the climate zone. The climate is generally characterized by the presence of two seasons: a dry season (October to April) due to the harmattan (dry wind) and a rainy season (May to September) influenced by the WAM (cold wind). The hydrographic network is relatively dense and consists of three main tributaries (Sirba, Faga and Yeli) as well as some small reservoirs from five dams with non-operated spillways. Based on the description of the rainfall pattern, the hydrological regime in the Sirba watershed is a Sahelian type, while its vegetation formation is thorny and of a lightly wooded savannah. The Sirba watershed is largely characterized by the rural sectors of agriculture, livestock breeding and exploitation of forest resources. These are the main production activities of

a weak local economy that is adversely affected by the depletion of the natural resource potential, poverty, recurring droughts and high population growth, all of which are exacerbated by the effects of climate change. Thus, all these activities depend on streamflow as rainfall is erratic in the area. Moreover, locally, the Sirba tributary plays an important role in the hydrological regime of the Niger river at Niamey, as it participates in its Sudanian flood in September. Therefore, the present study can benefit the rural population (e.g. farmers) to plan on the type of crops to sow according the streamflow forecast information received few months ahead. Also, they can avoid areas that usually prone to inundation.



Fig. 1 Sirba watershed and hydro-meteorological stations

2.2 Data presentation

Climate data are used in this study to select best predictors and develop seasonal forecast models. These climate data include rainfall and atmospheric data. The atmospheric data are used as predictors in the present study and they include: Sea Level Pressure (SLP), Relative Humidity (RHUM), and Air Temperature (AirTemp). These variables are monthly NCEP-DOE Reanalysis data obtained from the National Oceanic and

Atmospheric Administration (NOAA: <http://www.esrl.noaa.gov>). These variables are within the grid 0°E-357.5°E longitude and 90°N-90°S latitude, and span from January 1979 to August 2013 (Gado et al. 2015).

The observed rainfall data come from a network of 11 rain gage stations located in Burkina Faso and Niger (Fig.e 1). These daily databases span the period 1960-2008, and are obtained from national meteorological offices of Burkina Faso and Niger. Five of these stations are located within the watershed while the remaining six are located at most 25km from the watershed boundary (see Fig.1). The Thiessen polygon method was implemented to estimate average rainfall over the watershed from the 11 rainfall time series. In addition to climate data, hydrological data are also used in this study to develop seasonal forecast models and Sirba hydrological model. They are mean daily streamflow flow data collected from the flow measure station at Garbé-Kourou. They cover the period 1989-2002, and are sourced from the Directorate of Water Resources (Direction des Ressources en Eau: DRE) in Niger. Tables 1 and 2 present the rainfall and atmospheric data respectively.

Table 1 Presentation of rainfall stations

Station number (code)	Station name	Longitude	Latitude	Country
320006	Torodi	1.8	13.12	Niger
320002	Tera	0.82	14.03	Niger
320004	Tillaberi	1.45	14.20	Niger
320005	Gotheye	1.58	13.82	Niger
200082	Boulsa	-0.57	12.65	Burkina Faso
200026	Dori	0.033	14.03	Burkina Faso
200085	Bogande	0.13	12.98	Burkina Faso
200048	Dakiri	-0.27	13.30	Burkina Faso
200024	Gorgadji	-0.52	14.03	Burkina Faso
200086	Piela	-0.13	12.70	Burkina Faso
200047	Tougouri	-0.52	13.65	Burkina Faso

Source Gado et al. (2015)

Table 2 Description of atmospheric variables

Parameter	Units	Level	Reference data	Spatial coverage	Temporal coverage
Sea level pressures (SLP)	Pa/s	1000 hPa	NCEP 2	2.5°x2.5° grid 90N - 90S, 0E - 357.5E	1979/01/01 to 2013/08/31
Air temperature (AirTemp)	°K	1000 hPa	NCEP 2	2.5°x2.5° grid 90N - 90S, 0E - 357.5E	1979/01/01 to 2013/08/31
Relative humidity (RHUM)	%	1000 hPa	NCEP 2	2.5°x2.5° grid 90N - 90S, 0E - 357.5E	1979/01/01 to 2013/08/31

Source Gado et al. (2015)

2.3 Rainfall forecast results

Seasonal rainfall was forecasted over the Sirba watershed by Gado et al. (2015) using the climate and atmospheric data detailed above. Indeed, they first selected the optimal lag time for each of the three best predictors (AirTemp, RHUM, SLP) using the method developed in Sittichok et al. (2014). Indeed, the candidate predictor was aggregated all possible time windows in the 18 months prior to the rainy season and each the obtained time series is used as an input to a linear model linking it to the seasonal rainfall on the Sirba watershed. For each period, a linear model linking the predictor averaged over that period and seasonal rainfall on the Sirba is built. For more details in building up the linear model, refer to Sittichok et al. (2014) and Gado et al. (2015).

After getting the best predictors (Table 3) and their corresponding optimal lag time, two types of models associated to a Bayesian change point detection method were used to forecast the seasonal rainfall. The first sets of models (M1, M2) were models with changing parameters. Model M1 helped detecting potential changes in the relation between predictor and rainfall, and it assumed that the relationship changes with precipitation amplitude. While M2 was similar to M1, except that it considered that the rainfall-predictor relationship changes with time (i.e., the regression parameters change with time).

Table 3 Predictors used for seasonal forecast

PREDICTORS	NMAX**	R ²	NASH	HIT Rate Score	Best period M1-M2*	Lag time
Sea Level Pressure at 1000hPa	50	0.48	0.46	60.71	17 – 18	0
Relative Humidity at 1000hPa	80	0.58	0.52	64.29	10 – 10	8 months
Air Temperature at 1000hPa	10	0.530	0.527	67.86	1 – 4	4 months

Source Gado et al. (2015)

(*) M1=1:12 (January to December); M2=M1:18 (considered month of M1 to the next coming June)

(**) NMAX: number of best grid points retained after screening the predictor grid based on R²

The approach used in M1 can be summarized as follows:

- In the forecasting process, the year i was removed from the predictor and predictand time series.
- Then, a stepwise regression was used to fit a linear relation between the predictor and the predictand on the remainder of the series, and an initial forecast assuming a single equation for all points is issued.
- Based on the forecasted predictand, the data were sorted to give a new position to year i .
- Seidou et al. (2007) change point detection method was used to generate 1000 time series of length $n-1$, containing a random number of change points at random locations.
- A stepwise regression was used to fit linear relation between the predictor and predictand on the segments delineated by the change points.
- Probabilistic forecasts were generated which helped to compute empirical probability density of the forecast. At the end of the process, a probability density function was obtained for the forecast in year i .

For model M2, the same approach was applied as explained in M1 except that there was no ordering of data set after each removal of year i .

The second set of models (M3, M4) consisted with model M3 a classical linear model with constant parameters and a second model M4 which was based on the climatology.

After developing these 12 models using the three predictors, a normalized Bayes factor was used for all 12 competing models in order to provide a weighted comparison of the likelihood of each model given the observed data. For more details on these models, the reader can refer to Gado et al. (2015).

Indeed, for a given model, the probabilistic forecasts generated (i.e., those used to get the probability densities) for each year are further used in the Sirba hydrological model. The entire steps for this process are explained in the next sections.

2.4 Sirba hydrological model development

Since the rainfall forecasts estimates (Gado et al. 2015) are to be introduced into a hydrological model to obtain corresponding streamflow values, there is necessity to develop a Sirba hydrological model. Therefore, SWAT (Soil and Water Assessment Tool) model is used in this work to build-up the Sirba hydrological model. It is a physically based model (Arnold et al. 1998) and used to simulate river flows and other processes in various regions (Demirel et al. 2009; Huang et al. 2009; Tripathi et al. 2004; Xu et al. 2009).

Daily Rainfall and streamflow data (as described previously) were used in developing the Sirba SWAT model. For other climate parameters, the NCEP (National Centers for Environment Predictions) reanalysis was used to find dataset of wind speed, relative humidity and average temperature at the daily scale.

A digital elevation model (90 m resolution) of West Africa was downloaded from the Shuttle Radar Topography Mission (SRTM 2004) and used to delineate the Sirba watershed, as well as its sub-watersheds. Soil and landuse coverages were downloaded from the Waterbase database (George and Leon 2007). The land coverage in the database has a resolution of 500m and comes from the Global Land Cover Facility (Hansen et al. 1998). The soil map has a resolution of 1km and was generated by the Food Aid Organization and the United Nations Organization for Education, Science and Culture (FAO/UNESCO 2003).

The model was set to use the SCS curve number method (Soil Conservation Service Engineering Division 1972) to calculate surface runoff, and the Penman-Monteith method for potential evapotranspiration estimation. The Arc-SWAT interface (Olivera et al. 2006) was used for project development. The watershed was divided into 9 sub-basins

according to the topography in order to perform the spatial parameterization of the SWAT model. Also, the watershed was set to have 5 reservoirs, 64 Hydrological Response Units (HRUs) and 1 outlet at Garbé-Kourou. The characteristics of the dams were specified in the model reservoir routing parameters. With observed flow at Garbé-Kourou station, sensitivity analysis was performed using the SWAT-CUP software (Abbaspour 2007) and including all parameters that are likely to influence the chosen objective function: Nash-Sutcliffe coefficient. The calibration was performed using the SWAT-CUP interface and involved parameters that have the most influence on the streamflow. The SUIF2 algorithm was selected to calibrate the discharge at the outlet of the watershed. The calibration was stopped once the Nash coefficient was judged suitable.

2.5 Seasonal streamflow forecasting approach

The seasonal streamflow is forecasted using the probabilistic forecasts which were used to generate the probability density in forecasting seasonal rainfall (see Gado et al. 2015). For a given model, probabilistic forecasts for year Y are introduced into the Sirba hydrological model to obtain similar values for streamflow. Further, the outputs were used to construct probability density of the forecast and finally to get an estimate. The steps are schematically represented in Fig. 2. The developed approach is detailed as follow to obtain the forecast for each year:

1. From the set of seasonal rainfall probabilistic forecasts a pool of 500 values were chosen for each year (year Y). This gave rise to 14000 values for the entire span considered.
2. For year Y , the 500 probabilistic forecast estimates were introduced into the Sirba SWAT model using the ".pcp" file located in the "TxtInOut" folder. It is important to note that these forecasts have different probabilities.
3. The model was run to obtain corresponding values for streamflow without changing the model initial parameters. It is noted that after simulation the obtained values have different probabilities.
4. The obtained set of forecasts for year Y is used to compute the empirical probability density of the forecast. This is done under relatively free assumptions,

in such a way that the estimate of the distribution is non-parametric and it uses a normal kernel function.

5. This estimate based on normal kernel function, uses a window parameter (or width) which is a function of the number of the set of forecasts. The resulted vector of density values is then assessed at 100 equally spaced points (i.e., default value) on the entire range of dataset. It should be noted that the width represents the bandwidth of the Kernel-smoothing window and it was set to be optimal for estimating normal densities.
6. At the end of the process, a probability density function is obtained for the forecast in *year Y*.
7. The same process is repeated for each year till the end of the 28th year of observed streamflow. This helps to cover the entire span (1989-2002) under each of the 12 models.

It should be noted that this process was first applied for mean annual streamflows over July-August-September (JAS). Additionally, the same procedure was used for maximum monthly streamflows. This helps to compare the best models for these different kinds of streamflows in the Sirba basin.

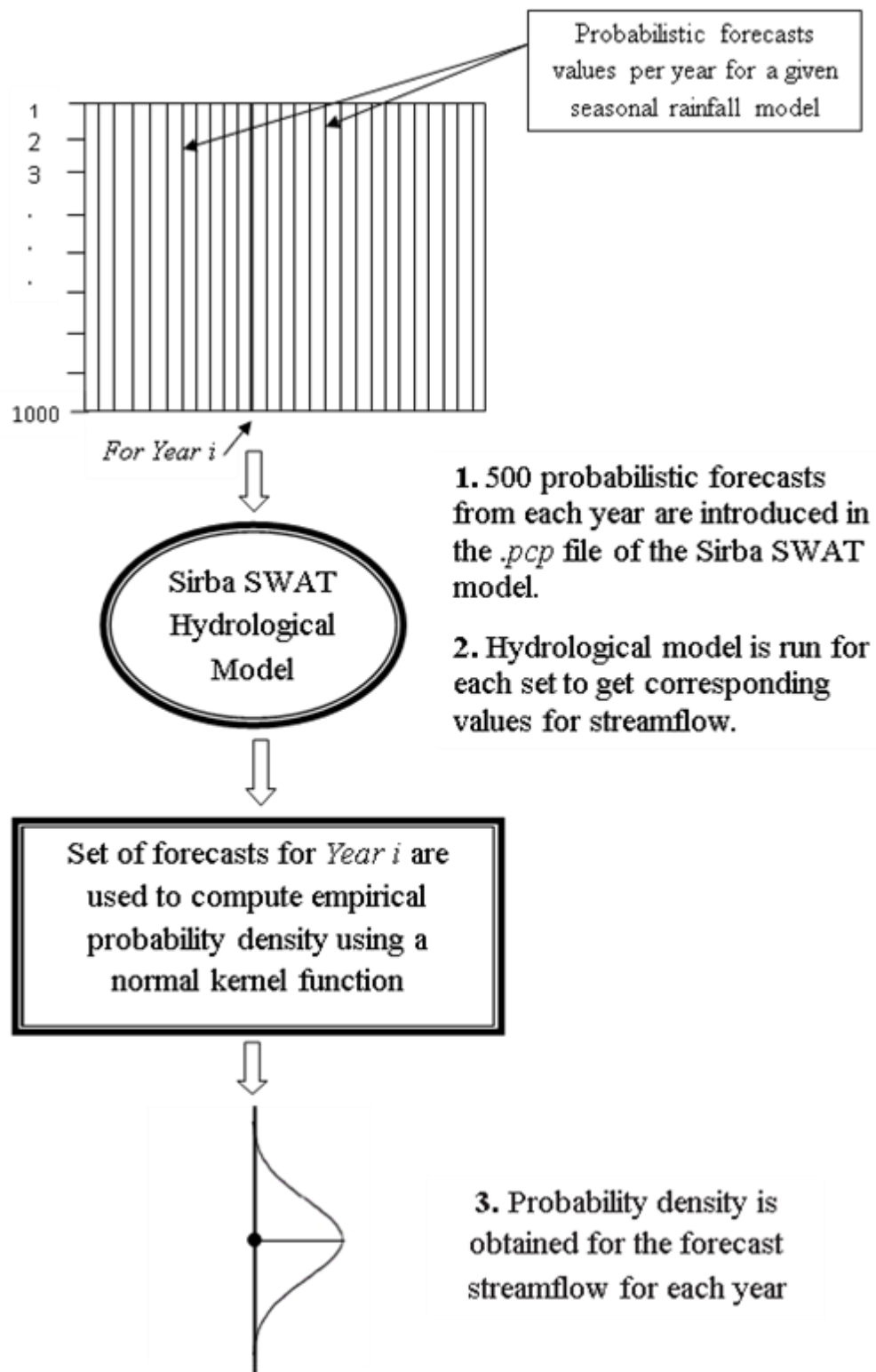


Fig. 2 Schematic representation of probabilistic forecasts generation for streamflow estimate

2.6 Bayesian model selection

Bayesian approach is used to select the best seasonal streamflow forecast model among the developed 12 models. For such purpose, prior and posterior probabilities of all models and observed streamflow data were first computed (see Eq. 1).

$$Posterior Prob_{model} = \frac{(Prior Prob_{model} \times Likelihood_{model})}{Prob_{observations}} \quad (1)$$

Where, $Likelihood_{model} = \prod_{i=1}^n likelihood_i$, i : year of forecasted streamflow, n : number of years (1-28)

A ratio of a models' posterior probability to that of observations constitutes a comparative criterion after normalization. This led to employ a normalized Bayes factor, B_f , (Eq. 2) which gives a weighted comparison of the likelihood of each model given the observed streamflow. This helped comparing the posterior likelihood of the data, d , of a given model, M_i , to that of the reference model M_r .

$$B_f = \frac{Likelihood(d/M_i)}{Likelihood(d/M_r)} \quad (2)$$

The selection of best seasonal streamflow forecast model consisted of analyzing the Bayes factors for all 12 models in order to find for which model B_f is always encouraging. Large values of B_f indicate that the data support M_i over M_r . It should be noted that the Bayes factor (Min et al. 2007) values for each model were computed using the observed streamflow data. Moreover, a graphical representation is used for each model to examine the evolution of likelihoods of forecasted seasonal streamflow. This scatter diagram presents the likelihood of each forecasted streamflow value on a coloured scale, where the reddish and bluish colours respectively tend toward a probability of 1 and 0.

3. Results and discussions

3.1 Sensitivity analysis and calibration of SWAT hydrological model

The model calibration was preceded by the sensitivity analysis. This sensitivity analysis showed that the most sensitive parameters were $Cn2$, $Esco$ and Sol_Awc . The first important parameter was the SCS curve number for moisture condition II, which controls how precipitation is split in runoff and infiltration. The second parameter was the soil evaporation compensation factor which affects portion of the soil prone for evaporation, counted in evapotranspiration. The higher the $ESCO$, lower is the evaporation. The last

third parameter was the available water capacity of soil layer (*Sol_Awc*); the higher the *Sol_Awc*, the less water to the reach. In total, 10 sensitive parameters having higher effects on streamflow were considered (Table 4). Therefore, using the most sensitive parameters for calibration, the ranges of fitted parameters in the final calibrated SWAT model were obtained and presented in Table 5. After calibration, the simulated monthly hydrograph fairly reproduced the observations (Fig. 3), with a Nash of 0.83. This coefficient is similar to those found in many studies in the same region. Indeed, Amadou et al. (2014) obtained a comparable Nash value (0.89) for a SWAT hydrological model of the Niger River at Koulikoro. Schuol and Abbaspour (2007) got a Nash in the interval 0-0.7 for most stations in the Upper Niger Basin while developing a SWAT model for Niger River. Their negative Nash values in the middle and lower basin are linked to data scarcity and for not taken into account dam operation. Likewise, Laurent and Ruelland (2010) obtained 0.88 to 0.91 (calibration/validation) as Nash coefficient for the Bani SWAT hydrological model, located in south-west in Mali.

Table 4 Description of sensitive parameters in Sirba SWAT hydrological model

Rank	Acronym	Definition	Range
1	Cn2	Moisture condition II curve number	20-90
2	Esco	Soil evaporation compensation coefficient	0.01-1
3	Sol_Awc	Available water capacity of the soil layer	0-1
4	Sol_Z	Depth from soil surface to bottom of layer (mm)	0-3500
5	Gwqmn	Threshold water level in shallow aquifer for base flow (mm H ₂ O)	0-5000
6	Ch_K(2)	Effective hydraulic conductivity (mm/hr)	0-150
7	Canmx	Maximum canopy storage (mm)	0-100
8	Revapmn	Threshold depth of water in the shallow aquifer for “revap” to occur (mm)	0-500
9	Blai	Potential maximum leaf area index for the plant	0-12
10	Surlag	Surface runoff lag coefficient	1-24

Table 5 Fitted parameters after calibration of Sirba SWAT hydrological model

Parameter	Fitted Value	Description	Effect	Unit
v_GW_DELAY.g w	403.59	Groundwater delay	Affects the groundwater movement, when increased. It is the lag between the times that water exits the soil profile and enters the shallow aquifer.	day
r_CN2.mgt	-0.18	Initial SCS runoff curve number for wetting condition	Affects the direct runoff	%
v_EPCO.hru	0.29	Plant uptake	For high value (1.0), the model	-

Parameter	Fitted Value	Description	Effect	Unit
		compensation factor	allows more of the water uptake demand to be met by lower layers in the soil. For low values (0.0) the model allows less variation from the depth distribution.	
r_ALPHA_BF.gw	1.10	Base flow Alpha factor	Shows the direct index of ground water flow response to changes in recharge	days
v_GWQMN.gw	329.01	Threshold depth of water required in the shallow aquifer required for base flow for base flow to occur	The lower the GWqmn the higher the stream flow	mm
r_GW_REVAP.gw	0.08	Groundwater “Revap” coefficient	Affects groundwater prone to evaporation, a low value of Groundwater “Revap” leads to high base flow	-
v_CH_K2.rte	5.97	Channel effective hydraulic conductivity	Allows interaction between groundwater flow and river flow.	mm/h
v_REVAPMN.gw	31.33	Threshold Depth of water in shallow aquifer for “Revap”	When increased, the base flow will increase	mm
v_CANMX.hru	64.21	Maximum canopy storage	Impacts on runoff	mm
v_ESCO.hru	0.98	Soil Evaporation compensation factor	Affect portion of the soil prone for evaporation, counted in evapotranspiration. The higher the ESCO, the lower the evaporation	-
v_SURLAG.bsn	14.67	Surface lag coefficient	It controls the fraction of the total water that is allowed to enter the stream on any specific day. When surlag increases, less water is held in storage	day
v_SLSUBBSN.hru	9.95	Slope sub-basin	Contribute directly to surface runoff generation	%
v_CH_N2.rte	0.081	Manning’s “n” value for the main channel	Affects the surface runoff	-
r_HRU_SLP.hru	0.63	Mean slope of HRU	Impacts on the sediment yield	%
v_BIOMIX.mgt	0.25	Biological mixing efficiency	Defines the activity of soil organisms which influence soil porosity and water fluxes	
r_SOL_AWC (1-2).sol	0.96	available water capacity of soil layer	The higher the Sol_AWC, the less water to the reach	%
r_SOL_Z (1-2).sol	-0.40	Depth from soil surface to bottom of layer	Acts on soil properties	mm
r_SOL_K (1-2).sol	0.88	Saturated hydraulic conductivity	Relates soil water flow rate to the hydraulic conductivity	%
r_SOL_ALB (1-	0.21	Moist soil albedo	Impacts on the moisture content	-

Parameter	Fitted Value	Description	Effect	Unit
2).sol				

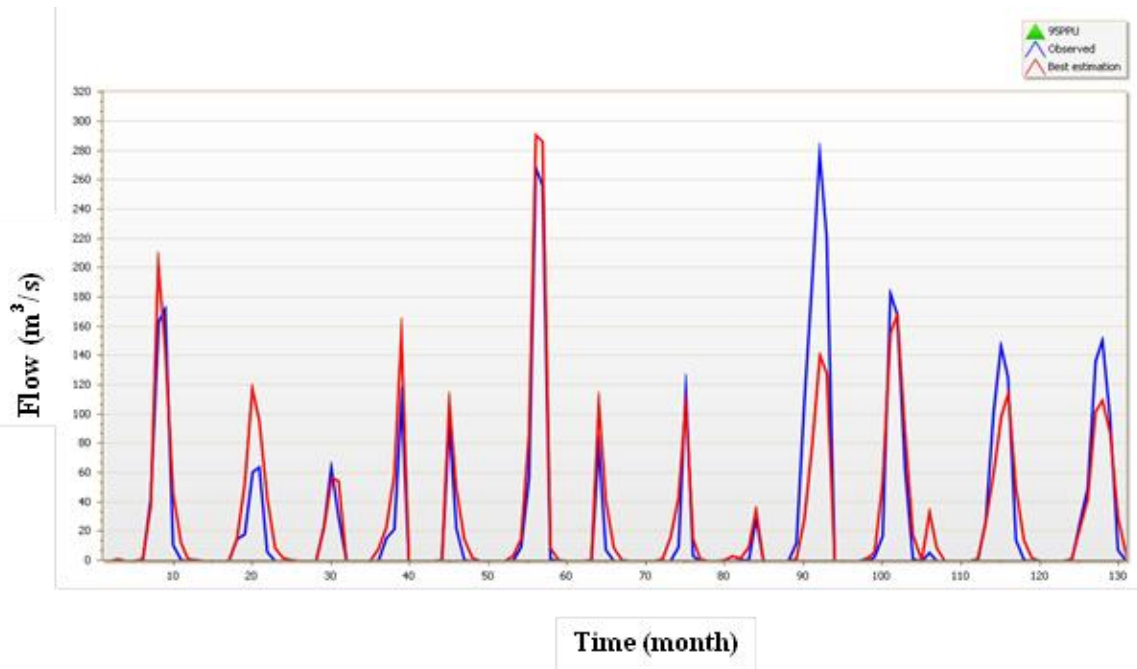


Fig. 3 Observed (blue) and simulated (red) hydrograph at Garbé-kourou (1989-2002)

3.2 Selection of best streamflow forecast model

After testing different models for finding the best one for the annual mean streamflows forecast over July-August-September (JAS) some analysis were done.

Analysis of Bayes factor and graphical representation of models' posterior likelihoods enables to select the best seasonal annual mean streamflow forecast model. Once, observations range in reddish colour band (high likelihood) according to a given model, that model is deemed reliable. On each of the graphs in Figures 4 to 6, some observed streamflow values were inside the reddish (high probability) area while others fell within the bluish range (low probability). Figure 4 showed that 91% and 100% of the forecasted streamflow values under model M3 using respectively AirTemp and RHUM are within the reddish area. Therefore, since this area designates the range of high probabilities, model M3 using RHUM can be considered as a credible model. Model M4 was found to perform well for all predictors, but this fact seemed that the models are not sensitive to the climatology which was used.

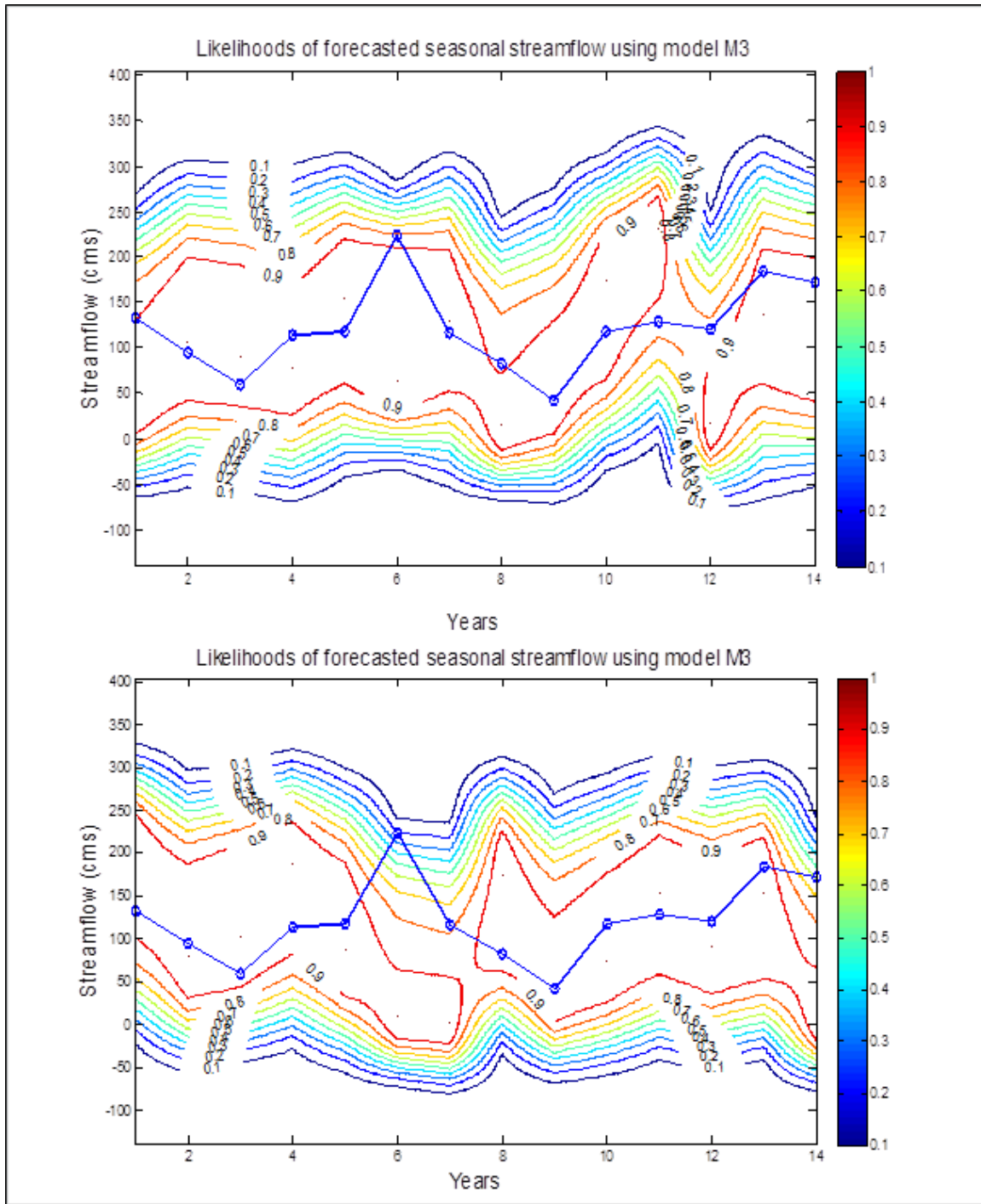


Fig. 4 Evolving likelihoods of forecasted seasonal annual mean streamflow from model M3 using RHUM (upper panel) and AirTemp (lower panel)

It is observed on Figs. 5 and 6 that most of the forecasted seasonal annual mean streamflows tend to reddish colour (i.e., high probabilities). For instance models M2

(using RHUM), M1 (using SLP) had more than 70% of forecasted values within the interval of high probabilities. However, the evolving likelihoods of forecasted streamflows for models M1 and M2 are low compared to those of M3. The normalized Bayes factors for all models are given in Table 6, while Table 7 provides the means to interpret the magnitude of Bayes factor. It helps concluding if there is strong, moderate, or weak evidence in favor of one of the competing models. Bayes factors were favorable for model M3 (using RHUM) and relatively low for other models, thereby proving that there is strong evidence for this model (Table 8). Also, Table 8 showed the presence of strong evidence for models using RHUM. A comparison of these models under each predictor showed that for SLP, the performance of models is found in decreasing order as $M3_s$, $M4_s$, $M2_s$, and $M1_s$; while for AirTemp the inequality $M3_a > M4_a > M2_a > M1_a$ was found. Furthermore, RHUM revealed $M3_{rh}$ as the strongest model followed respectively by $M4_{rh}$, $M2_{rh}$, and $M1_{rh}$. Thus, seasonal annual mean streamflow forecast, considering constant parameters seem to be the best approaches for this Sahelian region among the tested models. However, it should be noted that model M2 seemed to perform relatively well; and this performance could be enhanced if there were enough streamflow dataset. Also, forecasted values tend to reddish area for almost models.

Model M3 could be regarded as the most favorable for forecasting seasonal annual mean streamflow over the Sirba watershed with a lag time of 8 months. Thus, a combination of models with constant parameters and Bayesian model selection approach using normalized Bayes factor, forms as a whole a reliable forecast of seasonal annual mean streamflow over West Africa. Comparatively, the results obtained for the case of maximal monthly streamflows (JAS) showed that 100% and 92% of the forecasted maximum monthly streamflows under model M1 using respectively AirTemp and RHUM are within the area of high probabilities (Fig. 7). Thereby, model M1 seems to be the best model for predicting the maximum monthly streamflow as the likelihoods for other models are mostly in bluish range (low probabilities). This model M1 using AirTemp help predicting such flows 4 months ahead.

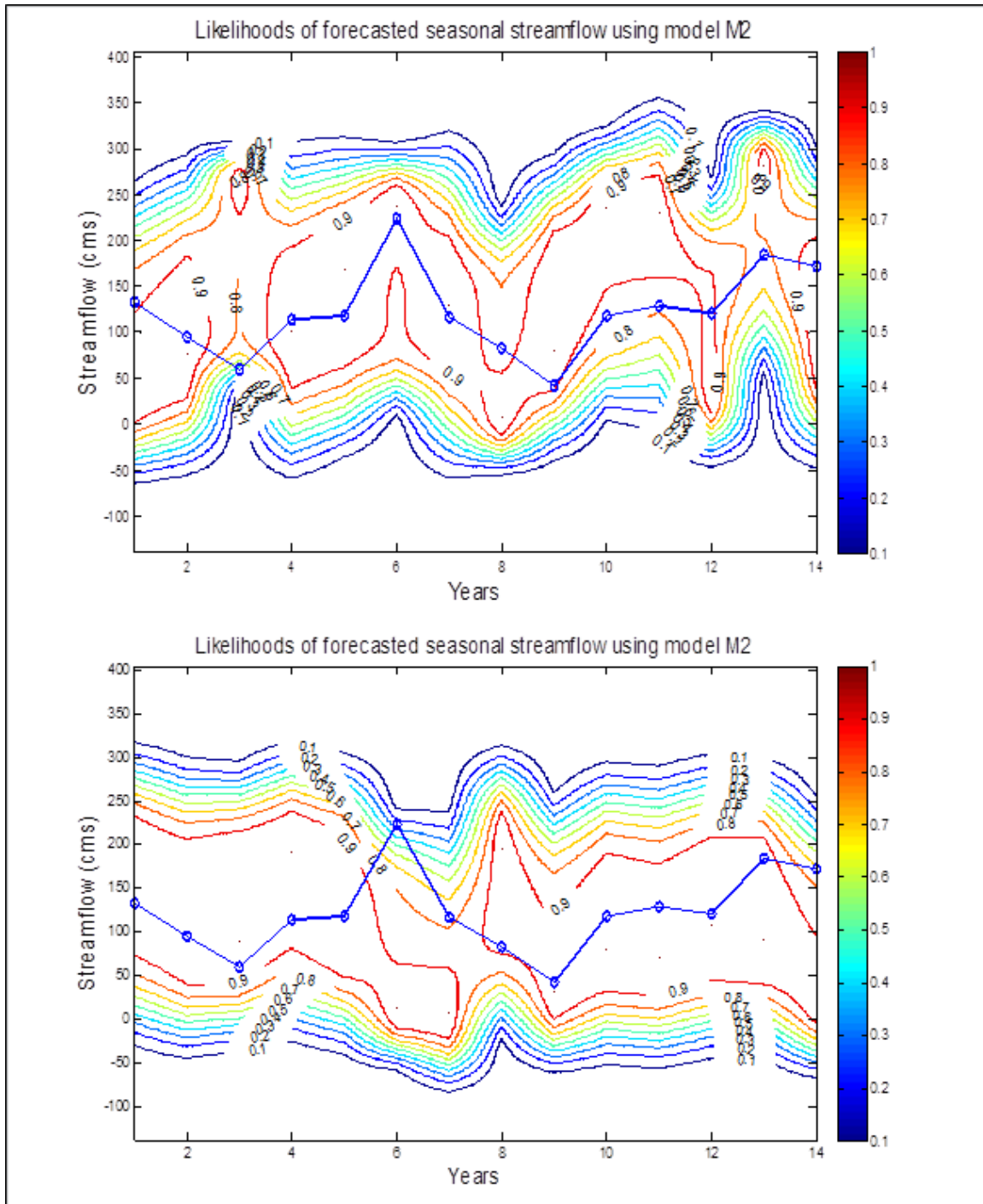


Fig. 5 Evolving likelihoods of forecasted seasonal annual mean streamflow from model M2 using RHUM (upper panel) and AirTemp (lower panel)

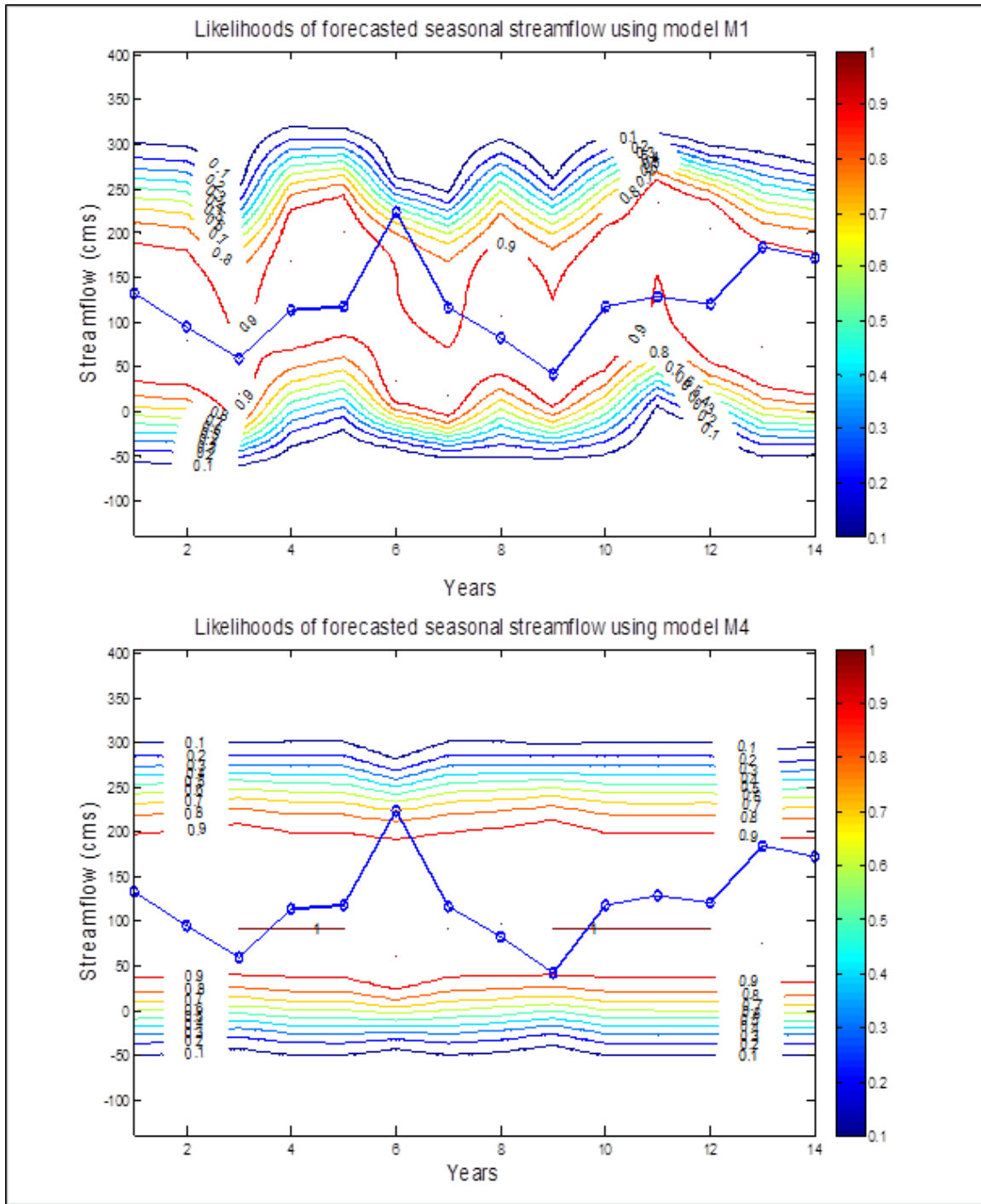


Fig. 6 Evolving likelihoods of predicted seasonal annual mean streamflow from models M1 (upper panel) and M4 (lower panel) using SLP

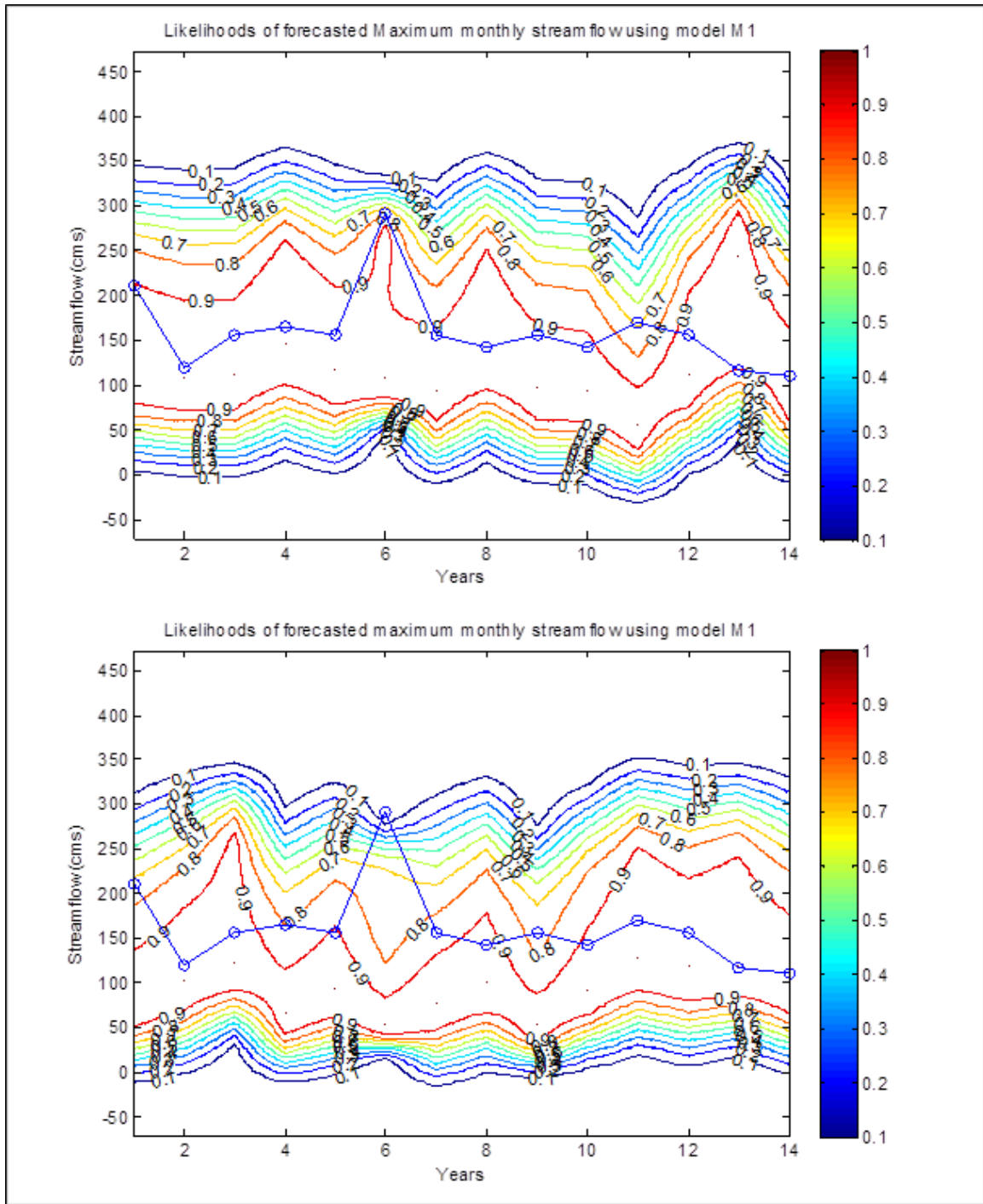


Fig. 7 Evolving likelihoods of forecasted maximum monthly streamflow from model M1 using AirTemp (upper panel) and RHUM (lower panel)

Additionally, a thorough analysis of the different Bayes factors in Table 9 and 10 exhibited large values of B_f to the benefit of model M1 using AirTemp which means a strong evidence for such model among the competing ones. Moreover, an appraisal of all models under each predictor pointed the following relations: $M1a > M3a > M2a > M4a$, $M1s > M4s > M3s > M2s$, and $M4rh > M1rh > M3rh > M2rh$ respectively for AirTemp, SLP and RHUM. Therefore, models with changing parameters could be reflected as the preeminent to forecast the maximum monthly streamflow in this region.

Regarding, the uncertainties, our model predicts seasonal streamflows and then statistically distributes it over the year using the fragment method. It is not possible in our study to look at different periods of the year separately, and the uncertainty will simply be proportional to the average streamflow in a given month.

One interesting finding in this paper is that different characteristics of the streamflow hydrograph are best forecasted by different models, suggesting that a multimodel approach to forecasting may be an interesting solution to climate forecasting in the Sahel. The study also highlighted that there are strong non-linearities between climate based predictors and streamflow in the study area. Therefore the current practice in the region of directly linking predictors such as sea surface temperature to streamflow in the Sahel is far from being optimal.

Table 6 Normalized Bayes factors for seasonal annual mean streamflow forecast models

$M_i \backslash M_r$		AirTemp				RHUM				SLP			
		$M1_a$	$M2_a$	$M3_a$	$M4_a$	$M1_{rh}$	$M2_{rh}$	$M3_{rh}$	$M4_{rh}$	$M1_s$	$M2_s$	$M3_s$	$M4_s$
AirTemp	$M1_a$	1	2.90E+01	8.15E+09	6.08E+09	3.18E+03	1.93E+05	4.95E+14	6.08E+09	5.60E-02	3.34E+00	1.65E+12	6.08E+09
	$M2_a$	3.45E-02	1	2.81E+08	2.10E+08	1.10E+02	6.67E+03	1.71E+13	2.10E+08	1.93E-03	1.15E-01	5.68E+10	2.10E+08
	$M3_a$	1.23E-10	3.56E-09	1	6.07E+04	7.45E-01	3.90E-07	8.15E+04	7.45E-01	6.87E-12	4.10E-10	2.02E+02	7.45E-01
	$M4_a$	1.65E-10	4.77E-09	2.37E-05	1	5.23E-07	3.18E-05	1.56E+11	1.00E+00	9.21E-12	5.50E-10	2.71E+02	1.00E+00
RHUM	$M1_{rh}$	3.15E-04	9.13E-03	1.34E+00	1.91E+06	1	6.08E+01	2.56E+09	1.91E+06	1.76E-05	1.05E-03	5.19E+08	1.91E+06
	$M2_{rh}$	5.17E-06	1.50E-04	2.57E+06	3.14E+04	1.64E-02	1	1.23E+04	3.14E+04	2.90E-07	1.73E-05	8.52E+06	3.14E+04
	$M3_{rh}$	2.02E-15	5.86E-14	4.22E+04	6.42E-12	3.90E-10	1.65E-05	1	1.23E-05	1.13E-16	6.76E-15	3.33E-03	1.23E-05
	$M4_{rh}$	2.83E-07	8.19E-06	2.30E+03	1.72E+03	8.98E-04	5.46E-02	1.40E+08	1	1.58E-08	9.45E-07	4.66E+05	1.72E+03
SLP	$M1_s$	1.79E+01	5.18E+02	1.46E+11	1.09E+11	5.68E+04	3.45E+06	8.84E+15	1.09E+11	1	5.98E+01	2.94E+13	1.09E+11
	$M2_s$	2.99E-01	8.67E+00	2.44E+09	1.82E+09	9.50E+02	5.78E+04	1.48E+14	1.82E+09	1.67E-02	1	4.92E+11	1.82E+09
	$M3_s$	6.07E-13	1.76E-11	4.95E-03	3.69E-03	1.93E-09	1.17E-07	3.01E+02	3.69E-03	3.40E-14	2.03E-12	1	3.69E+01
	$M4_s$	2.44E-07	7.08E-06	1.99E+03	1.48E+03	7.76E-04	4.72E-02	1.21E+08	1.48E+03	1.37E-08	8.17E-07	4.02E+05	1

a: subscript for model developed using AirTemp as predictor

rh: subscript for model developed using RHUM as predictor

s: subscript for model developed using SLP as predictor

Table 7 Scale for Bayes factor interpretation (source: Min et al. 2007)

Bayes Factor	Interpretation
$B_f < 1/10$	Strong evidence for M_r
$1/10 \leq B_f < 1/3$	Moderate evidence for M_r
$1/3 \leq B_f < 1$	Weak evidence for M_r
$1 \leq B_f < 3$	Weak evidence for M_i
$3 \leq B_f < 10$	Moderate evidence for M_i
$B_f \geq 10$	Strong evidence for M_i

Table 8 Comparison of twelve seasonal annual mean streamflow forecast models

M _i M _r		AirTemp				RHUM				SLP			
		M1 _a	M2 _a	M3 _a	M4 _a	M1 _{rh}	M2 _{rh}	M3 _{rh}	M4 _{rh}	M1 _s	M2 _s	M3 _s	M4 _s
AirTemp	M1 _a	Wk.E.M1 _a	St.E.M2 _a	St.E.M3 _a	St.E.M4 _a	St.E.M1 _{rh}	St.E.M2 _{rh}	St.E.M3 _{rh}	St.E.M4 _{rh}	St.E.M1 _a	St.E.M2 _s	St.E.M3 _s	St.E.M4 _s
	M2 _a	St.E.M2 _a	Wk.E.M2 _a	St.E.M3 _a	St.E.M4 _a	St.E.M1 _{rh}	St.E.M2 _{rh}	St.E.M3 _{rh}	St.E.M4 _{rh}	St.E.M2 _a	St.E.M2 _a	St.E.M3 _s	St.E.M4 _s
	M3 _a	St.E.M3 _a	St.E.M3 _a	Wk.E.M3 _a	St.E.M3 _a	St.E.M3 _a	St.E.M3 _a	St.E.M3 _{rh}	St.E.M3 _a	St.E.M3 _a	St.E.M3 _a	St.E.M3 _a	St.E.M3 _a
	M4 _a	St.E.M4 _a	St.E.M4 _a	St.E.M3 _a	Wk.E.M4 _a	St.E.M4 _a	St.E.M4 _a	St.E.M3 _{rh}	St.E.M4 _{rh}	St.E.M4 _a	St.E.M4 _a	St.E.M3 _s	St.E.M4 _s
RHUM	M1 _{rh}	St.E.M1 _{rh}	St.E.M2 _a	St.E.M3 _a	St.E.M4 _a	Wk.E.M1 _{rh}	St.E.M2 _{rh}	St.E.M3 _{rh}	St.E.M4 _{rh}	St.E.M1 _{rh}	St.E.M1 _{rh}	St.E.M3 _s	St.E.M4 _s
	M2 _{rh}	St.E.M2 _{rh}	St.E.M2 _{rh}	St.E.M3 _a	St.E.M4 _a	St.E.M2 _{rh}	Wk.E.M2 _{rh}	St.E.M3 _{rh}	St.E.M4 _{rh}	St.E.M2 _{rh}	St.E.M2 _{rh}	St.E.M3 _s	St.E.M4 _s
	M3 _{rh}	St.E.M3 _{rh}	St.E.M3 _{rh}	St.E.M3 _{rh}	St.E.M3 _{rh}	St.E.M3 _{rh}	St.E.M3 _{rh}	Wk.E.M3 _{rh}	St.E.M3 _{rh}	St.E.M3 _{rh}	St.E.M3 _{rh}	St.E.M3 _{rh}	St.E.M3 _{rh}
	M4 _{rh}	St.E.M4 _{rh}	St.E.M4 _{rh}	St.E.M3 _a	St.E.M4 _{rh}	St.E.M4 _{rh}	St.E.M4 _{rh}	St.E.M3 _{rh}	Wk.E.M4 _{rh}	St.E.M4 _{rh}	St.E.M4 _{rh}	St.E.M3 _s	St.E.M4 _{rh}
SLP	M1 _s	St.E.M1 _a	St.E.M2 _a	St.E.M3 _a	St.E.M4 _a	St.E.M1 _{rh}	St.E.M2 _{rh}	St.E.M3 _{rh}	St.E.M4 _{rh}	Wk.E.M1 _s	St.E.M1 _s	St.E.M3 _s	St.E.M4 _s
	M2 _s	St.E.M2 _s	St.E.M2 _s	St.E.M3 _a	St.E.M4 _a	St.E.M1 _{rh}	St.E.M2 _{rh}	St.E.M3 _{rh}	St.E.M4 _{rh}	St.E.M1 _s	Wk.E.M2 _s	St.E.M3 _s	St.E.M4 _s
	M3 _s	St.E.M3 _s	St.E.M3 _s	St.E.M3 _a	St.E.M3 _s	St.E.M3 _s	St.E.M3 _s	St.E.M3 _{rh}	St.E.M3 _s	St.E.M3 _s	St.E.M3 _s	Wk.E.M3 _s	St.E.M3 _s
	M4 _s	St.E.M4 _s	St.E.M4 _s	St.E.M3 _a	St.E.M4 _a	St.E.M4 _s	St.E.M4 _s	St.E.M3 _{rh}	St.E.M4 _{rh}	St.E.M4 _s	St.E.M4 _s	St.E.M3 _s	Wk.E.M4 _s

St.E.: Strong Evidence (for example, St.E. M1_a: strong evidence for M1_a)

Md.E.: Moderate Evidence

Wk.E.: Weak Evidence

a: subscript for model developed using AirTemp as predictor

rh: subscript for model developed using RHUM as predictor

s: subscript for model developed using SLP as predictor

Table 9 Normalized Bayes factors for maximum monthly streamflow forecast models

$\begin{matrix} M_i \\ M_r \end{matrix}$		AirTemp				RHUM				SLP			
		M1 _a	M2 _a	M3 _a	M4 _a	M1 _{rh}	M2 _{rh}	M3 _{rh}	M4 _{rh}	M1 _s	M2 _s	M3 _s	M4 _s
AirTemp	M1 _a	1	2.33E+04	3.95E+01	2.81E+03	1.25E+01	3.43E+13	3.15E+19	2.43E+04	2.06E+05	2.45E+11	1.88E+22	1.57E+03
	M2 _a	1.24E+03	1	5.40E-05	2.67E-01	5.06E-04	7.73E+09	2.31E+15	8.01E-01	1.82E+01	9.87E+06	7.65E+17	6.55E-01
	M3 _a	9.53E+00	1.41E-03	1	6.58E+02	3.47E+01	2.79E+14	6.55E+19	5.88E+03	5.12E+05	2.77E+11	5.46E+21	5.65E+03
	M4 _a	2.21E+04	3.88E-01	6.58E+03	1	6.60E-03	5.47E+10	2.12E+16	1.00E+00	8.97E+01	6.86E+07	6.57E+18	1.00E+00
RHUM	M1 _{rh}	1.95E+01	2.53E-04	6.64E+01	6.70E-03	1	9.09E+12	3.91E-19	1.89E+02	3.97E-05	1.98E+10	4.35E+20	5.67E+02
	M2 _{rh}	6.43E+14	8.61E+09	2.13E+14	5.61E+10	9.01E+12	1	3.13E+05	3.39E-11	6.04E+08	7.76E+02	1.34E-08	5.68E+10
	M3 _{rh}	1.16E+19	1.25E+15	6.55E+19	1.09E+16	5.79E-19	2.89E+05	1	1.89E+16	2.35E+14	3.09E+08	2.99E+02	1.88E+16
	M4 _{rh}	2.19E+03	2.09E-01	8.53E+03	1.00E+00	4.19E+02	1.54E-11	2.18E+16	1	8.08E+01	1.98E-08	6.54E+18	1.00E+00
SLP	M1 _s	2.64E+05	1.61E+02	5.14E+05	9.18E+01	7.90E-05	7.72E+08	2.32E+14	9.10E+01	1	7.89E+05	2.54E-17	9.22E+01
	M2 _s	2.13E+12	8.89E+07	3.14E+11	5.64E+07	3.08E+10	7.89E+02	2.12E+08	3.12E-08	7.12E+05	1	1.88E-11	5.68E+07
	M3 _s	3.67E+22	6.60E+16	5.87E+21	4.68E+18	7.12E+20	2.36E-08	3.90E+02	3.90E+18	3.51E-17	1.77E-11	1	7.65E+18
	M4 _s	3.60E+03	2.11E-01	2.66E+03	1.00E+00	1.67E+02	5.65E+10	2.21E+16	1.00E+00	9.19E+01	4.33E+07	3.69E+18	1

Table 10 Comparison of twelve maximum monthly streamflow forecast models

$\begin{matrix} M_i \\ M_r \end{matrix}$		AirTemp				RHUM				SLP			
		M1 _a	M2 _a	M3 _a	M4 _a	M1 _{rh}	M2 _{rh}	M3 _{rh}	M4 _{rh}	M1 _s	M2 _s	M3 _s	M4 _s
AirTemp	M1 _a	Wk.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a	St.E. M1 _a
	M2 _a	St.E. M1 _a	Wk.E. M2 _a	St.E. M3 _a	Md.E. M2 _a	St.E. M1 _{rh}	St.E. M2 _a	St.E. M3 _{rh}	Md.E. M4 _{rh}	St.E. M2 _a	St.E. M2 _a	St.E. M2 _a	Md.E. M4 _s
	M3 _a	St.E. M1 _a	St.E. M3 _a	Wk.E. M3 _a	St.E. M3 _a	St.E. M1 _{rh}	St.E. M2 _{rh}	St.E. M3 _a	St.E. M3 _a	St.E. M1 _s	St.E. M3 _a	St.E. M3 _s	St.E. M3 _a
	M4 _a	St.E. M1 _a	Md.E. M2 _a	St.E. M3 _a	Wk.E. M4 _a	St.E. M1 _{rh}	St.E. M4 _a	St.E. M4 _a	Wk.E. M4 _a	St.E. M4 _a	St.E. M4 _a	St.E. M3 _s	Wk.E. M4 _a
RHUM	M1 _{rh}	St.E. M1 _a	St.E. M1 _{rh}	St.E. M1 _{rh}	St.E. M1 _{rh}	Wk.E. M1 _{rh}	St.E. M1 _{rh}	St.E. M1 _{rh}	St.E. M4 _{rh}	St.E. M1 _s	St.E. M1 _{rh}	St.E. M3 _s	St.E. M1 _{rh}
	M2 _{rh}	St.E. M1 _a	St.E. M2 _a	St.E. M2 _{rh}	St.E. M4 _a	St.E. M1 _{rh}	Wk.E. M2 _{rh}	St.E. M3 _{rh}	St.E. M4 _{rh}	St.E. M1 _s	St.E. M2 _{rh}	St.E. M3 _s	St.E. M4 _s
	M3 _{rh}	St.E. M1 _a	St.E. M3 _{rh}	St.E. M3 _a	St.E. M4 _a	St.E. M1 _{rh}	St.E. M3 _{rh}	Wk.E. M3 _{rh}	St.E. M4 _{rh}	St.E. M1 _s	St.E. M2 _s	St.E. M3 _{rh}	St.E. M4 _s
	M4 _{rh}	St.E. M1 _a	Md.E. M4 _{rh}	St.E. M3 _a	Wk.E. M4 _a	St.E. M4 _{rh}	St.E. M4 _{rh}	St.E. M4 _{rh}	Wk.E. M4 _{rh}	St.E. M4 _{rh}	St.E. M4 _{rh}	St.E. M3 _s	Wk.E. M4 _{rh}
SLP	M1 _s	St.E. M1 _a	St.E. M2 _a	St.E. M1 _s	St.E. M4 _a	St.E. M1 _s	St.E. M1 _s	St.E. M1 _s	St.E. M4 _{rh}	Wk.E. M1 _s	St.E. M1 _s	St.E. M1 _s	St.E. M1 _s
	M2 _s	St.E. M1 _a	St.E. M2 _a	St.E. M3 _a	St.E. M4 _a	St.E. M1 _{rh}	St.E. M2 _{rh}	St.E. M2 _s	St.E. M4 _{rh}	St.E. M1 _s	Wk.E. M2 _s	St.E. M3 _s	St.E. M4 _s
	M3 _s	St.E. M1 _a	St.E. M2 _a	St.E. M3 _s	St.E. M3 _s	St.E. M3 _s	St.E. M3 _s	St.E. M3 _{rh}	St.E. M3 _s	St.E. M1 _s	St.E. M3 _s	Wk.E. M3 _s	St.E. M4 _s
	M4 _s	St.E. M1 _a	Md.E. M4 _s	St.E. M3 _a	Wk.E. M4 _a	St.E. M1 _{rh}	St.E. M4 _s	St.E. M4 _s	Wk.E. M4 _{rh}	St.E. M4 _s	St.E. M4 _s	St.E. M4 _s	Wk.E. M4 _s

4. Conclusion

Comparative study is carried out using two set of models having respectively changing and constant parameters. These models were applied on three predictors, namely, sea level pressure, relative humidity, and air temperature. Seasonal rainfall probability forecasts derived from such models were used in combination with Sirba hydrological model to estimate seasonal streamflow. Graphs of likelihoods in addition to Bayesian model selection using normalized Bayes factor helped to confront rivaling models. The results showed that model M3 using RHUM as predictor is the best seasonal annual mean streamflow forecast model followed by model M4 among forecast models. Also, model M1 using AirTemp is deemed the preeminent to forecast the maximum monthly streamflow. Thereby, models with constant parameters and those allowing for parameters changes according to rainfall magnitude could be respectively regarded as the best for seasonal annual mean (resp. maximum monthly) streamflows forecast over the Sirba watershed. Therefore, this type of model associated with the Bayesian model selection approach appeared to be a good way to forecast seasonal annual mean (resp. maximum monthly) streamflow in this Sahelian region.

5. References

- Abbaspour KC (2007) User Manual for SWAT-CUP, SWAT Calibration and Uncertainty Analysis Programs, Swiss Federal Institute of Aquatic Science and Technology, EAWAG, Dübendorf, Switzerland, 2007.
- Abudu S, King JP, Pagano TC (2010) Application of partial least-square regression in seasonal streamflow forecasting. *Journal of Hydrologic Engineering*, 15, 612-623
- Amadou A, Gado Djibo A, Seidou O, Sittichok K, Seidou Sanda I (2014) Changes to flow regime on the Niger River at Koulikoro under a changing climate. *Journal of hydrological sciences*. doi: 10.1080/02626667.2014.916407
- Arnold JG, Srinivasan R, Muttiah RS, Williams JR (1998) Large area hydrologic modeling and assessment - Part 1: Model development. *Journal of American Water Resources Association* 34 (1), 73-89
- Barnston AG, Thiao W, Kumar V (1996) Long-lead forecasts of seasonal precipitation in Africa using CCA. *Weather and Forecasting*, 11: 506-520.
- Bouali L (2009) Prévisibilité et prévision statistico-dynamique des saisons des pluies associées à la mousson ouest africaine à partir d'ensembles multi-modèles. Dissertation, Université de Bourgogne, France. 159 pp

- Brooks N (2004) Drought in the African Sahel: long term perspectives and future prospects. Tyndall Centre Working Paper, vol 61. Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich.
- Cayan DR, Riddle LG (1993) The influence of precipitation and temperature on seasonal streamflow in California. *Water Resources Research*, 29: 1127-1140.
- Demirel MC, Venancio A, Kahya E (2009) Flow forecasts by SWAT model and ANN in Pracana basin, Portugal. *Advanced in Engineering Software*, 40: 467-473.
- FAO/UNESCO (2003) Digital Soil Map of the World and Derived Soil Properties.Rev.1.(CD Rom), 2003. Available from http://www.fao.org/catalog/what_new-e.htm. Accessed 30 May 2014.
- Folland C, Owen J, Ward MN, Colman A (1991) Prediction of seasonal rainfall in the Sahel region using empirical and dynamic models. *Journal of Forecasting*, 10: 21-56.
- Gachon P, Gauthier N, Bokoye A, Pauishkura D, Cotnoir A, Trembley Y, Vigeant G (2007) Variabilité, extrême et changement climatique au Sahel: de l'observation à la modélisation. Rapport de projet -Entente de coopération entre le Centre AGRHYMET et Environnement Canada. Version mars 2007. p. 10-64.
- Gado Djibo A, Seidou O, Karambiri H, Sittichock K, Paturel JE, Saley HM (2015) Development and assessment of non-linear and non-stationary seasonal rainfall forecast models for the Sirba watershed, West Africa. *Journal of hydrology. Regional Studies*. DOI: 10.1016/j.ejrh.2015.05.001
- Garbrecht JD, Schneider JM, Van Liew MW (2007) Monthly runoff predictions based on rainfall forecasts in a small Oklahoma Watershed. *Journal of the American Water Resources Association* 42: 1285-1295.
- Garric G, Douville H, Deque M (2002) Prospect for improved seasonal predictions of monsoon precipitations over the Sahel. *International Journal of Climatology*, 22: 331-345.
- George C, Leon LF (2007) WaterBase: SWAT in an open source GIS. *The Open Hydrology Journal* 1: 19-24. Bentham Science Publisher Ltd. 1874-3781/07 2007.
- Hamatan M (2004) Synthèse et évaluation des prévisions saisonnières en Afrique de l'Ouest. *Sécheresse*, 15 (3) : 279-286
- Hansen M, DeFries R, Townshend J, Sohlberg R (1998) 1 Km Land Cover Classification Derived from AVHRR. <http://glcf.umiacs.umd.edu/data/landcover>. Accessed 30 May 2014.
- Huang Z, Xue B, Pang Y (2009) Simulation on stream flow and nutrient loadings in Gucheng Lake, Low Yangtze River Basin, based on SWAT model. *Quaternary International* 208: 109-115.
- IPCC (2007) Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the IPCC. In M.L. Parry,

- O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson, eds. Cambridge University Press, Cambridge, UK, 976pp.
- Janicot S, Moron V, Fontaine B (1996) Sahel droughts and ENSO dynamics. *Geophysical Research Letters*, vol. 23, p. 515-518.
- Kwon H, Brown C, Xu K, Lall U (2009) Seasonal and annual maximum streamflow forecasting using climate information: application to the Three Gorges Dam in the Yangtze River basin, China. *Hydrological Science Journal* 54: 582-595.
- Laurent F, Ruelland D (2010) Modélisation à base physique de la variabilité hydroclimatique à l'échelle d'un grand bassin versant tropical. Proc. of 6th world FRIEND Int. Conference, Fez, Morocco, 25-29 october 2010. IAHS Publ.
- MARA F (2010) Développement et analyse des critères de vulnérabilité des populations sahéniennes face à la variabilité du climat: le cas de la ressource en eau dans la vallée de la Sirba au Burkina Faso. Thèse de doctorat. Université du Québec à Montréal. 273pp
- Min S, Simonis D, Hense A (2007) Probabilistic climate change predictions applying Bayesian model averaging. *Phil. Trans. R. Soc. A* 365: 2103-2116
- Ndiaye O, Goddard L, Ward MN (2009) Using regional wind fields to improve general circulation model forecasts of July-September Sahel rainfall. *International Journal of Climatology* 29: 1262-1275.
- Ndiaye O, Ward MN, Thiaw WM (2011) Predictability of seasonal Sahel rainfall using GCMs and lead-time improvement through the use of a couple model. *Journal of Climate* 24: 1931-1949.
- Olivera F, Valenzuela M, Srinivasan R, Choi J, Cho HD, Koka S, Agrawal A (2006) ArcGIS-SWAT: A geodata model and GIS interface for SWAT. *Journal of the American Water Resources Association* 42: 295-309.
- Piechota TC, Chiew FHS, Dracup JA, McMahon TA (1998) Seasonal streamflow forecasting in eastern Australia and the El Nino-Southern Oscillation. *Water Resources Research* 34: 3035-3044.
- Roudier P, Ducharne A, Feyen L (2014) Climate change impacts on runoff in West Africa: a review. *Hydrol. Earth Syst. Sci.* 18: 2789-2801
- Rowe CM, Kuivinen KC, Flores-Mendoza F (1994) Sensitivity of Streamflow to Climate Change: A Case Study for Nebraska. *Great Plains Research: A Journal of Natural and Social Sciences*. Paper 154.
- Rowell DP, Folland CK, Maskell K, Owen JA, Ward MN (1992) Modelling the influence of global sea surface temperatures on the variability and predictability of seasonal Sahel rainfall. *Geophysical Research Letters*. 19: 905-908.
- Samimi C, Fink AH, Paeth H (2012) The 2007 flood in the Sahel: causes, characteristics and its presentation in the media and FEWS NET. *Natural Hazards and Earth System Science*. 12: 313-325.

- Schuol J, Abbaspour KC (2007) Using monthly weather statistics to generate daily data in a SWAT model application to West Africa. *Ecological modelling* 201: 301-311
- SRTM (2004) DEM data from International Centre for Tropical Agriculture (CIAT), available from the CGIAR CSI SRTM 90m Database: <http://srtm.csi.cgiar.org>. Accessed 30 May 2014
- Seidou O, Asselin JJ, Ouarda TBMJ (2007) Bayesian multivariate linear regression with application to change-point models in hydrometeorological variables. *Water Resour. Res.* doi: 10.1029/2005WR004835
- Sittichok K, Gado Djibo A, Seidou O, Saley HM, Karambiri H, Paturel J (2014) Statistical seasonal rainfall and streamflow forecasting for the Sirba watershed, using sea surface temperature. *Hydrological Sciences Journal*. doi:10.1080/02626667.2014.944526
- Taweye A (1995) Contribution à l'étude hydrologique du bassin versant de la Sirba à Garbé-Kourou. Dissertation, Centre Régional AGRHYMET, 96pp.
- Thaiw W, Mo KC (2005) Impact of sea surface temperature and soil moisture on seasonal rainfall prediction over the Sahel. *Journal of Climate* 18: 5330-5343.
- Tripathi MP, Panda RK, Raghuwanshi NS, Singh R (2004) Hydrological modelling of a small watershed using generated rainfall in the soil and water assessment tool model. *Hydrological Process* 18: 1811-1821.
- Tuteja NK, ShinD, Laugesen R, Khan U, Shao Q, Wang E, Li M, Zheng H, Kuczera G, Kavetski D, Evin G, Thyer M, MacDonald A, Chia T, Le B (2011) Experiment evaluation of the dynamic seasonal streamflow forecasting approach. Technical Report, Bureau of Meteorology, Melbourne.
- Wang E, Zhang Y, Luo J, Chiew F, Wang QJ (2011) Monthly and seasonal streamflow forecasts using rainfall-runoff modeling and historical weather data. *Water Resources Research* 47, W05516 (1-13)
- Washington, R., Downing, T.E. (1999). Seasonal forecasting of African rainfall: prediction, responses and household food security. *The Geographical Journal*, 165, 255-274.
- Wolfram L, Constanze L, Barnabas A (2008) Impact of Climate Change on the Black Volta Basin and the Bui Dam. GLOWA Volta Policy Brief, *Research report* march 2008. (<http://ghanadamsdialogue.iwmi.org/Data/Sites/2/media/glowavoltapolicybriefuidam170308.pdf>).
- Xu ZX, Pang JP, Liu CM, Li JY (2009) Assessment of runoff and sediment yield in the Miyun reservoir catchment by using SWAT model. *Hydrological Process* 23: 3619-3630
- Yossef NC, Winsemius H, Weerts A, Beek RV, Bierkens FP (2013) Skill of global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. *Water Resources Research* 49: 4687-4699.

6.3. Seasonal Streamflow Forecasting with Non-linear Methods

The seasonal rainfall predicted by the NLPCA model using AirTemp is considered for predicting the seasonal streamflow. The forecasted seasonal rainfall was disaggregated from seasonal to daily scale using the method of fragment (MF). To estimate the performance of MF, the simulated monthly rainfall (after disaggregation) was compared to the observed rainfall using NASH, R^2 and HIT criteria. The results showed that the simulated rainfall closely agreed with the observed data, with Nash, R^2 and HIT values of 0.81, 0.87 and 85.82% respectively. Thus, MF performed quite well with respect to the obtained model skill values for this temporal disaggregation model. The simulated and observed monthly rainfall are compared in Figure 6.4.

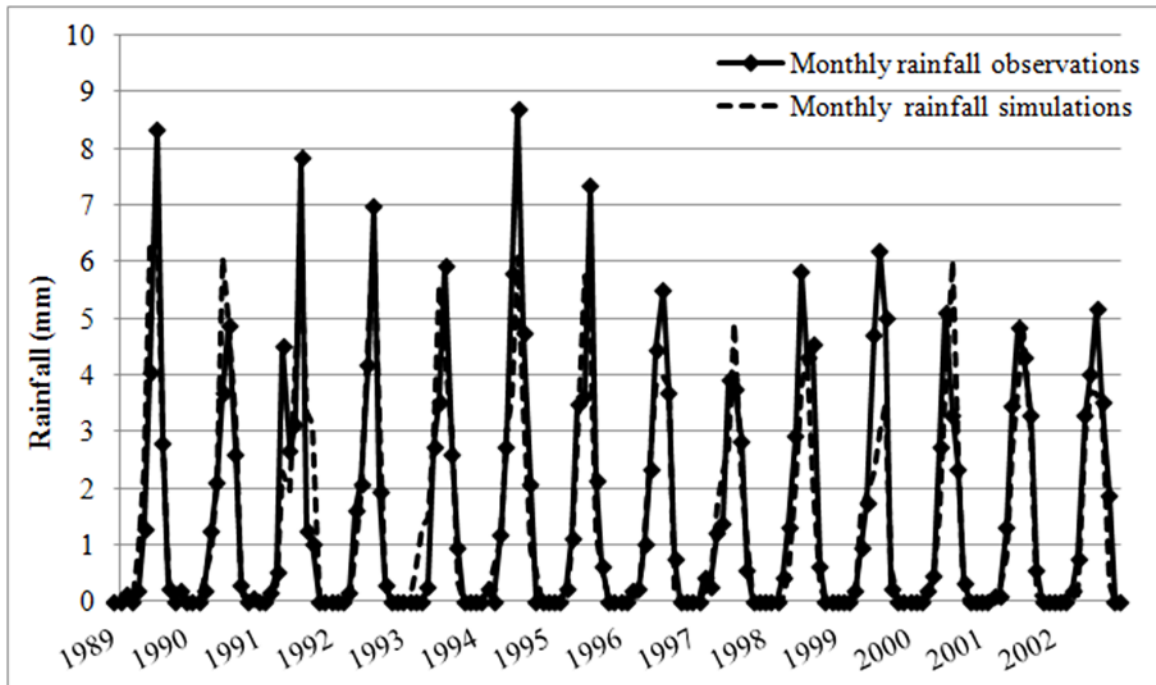


Figure 6.4 Observed and simulated rainfall using method of fragment (1989-2003)

After acquiring the disaggregated rainfall, it was fed into the Sirba SWAT model to force the hydrological model to produce monthly streamflows, which were later aggregated into seasonal streamflows. Figure 6.5 presents the observed and simulated streamflows, and it is obvious that the result was fairly acceptable (NASH value of 0.58) knowing all the uncertainties involved in the seasonal rainfall forecasting and in a

hydrological model development. However, this seasonal streamflow forecast can be improved by working the temporal disaggregation model as well the calibration of the Sirba hydrological model.

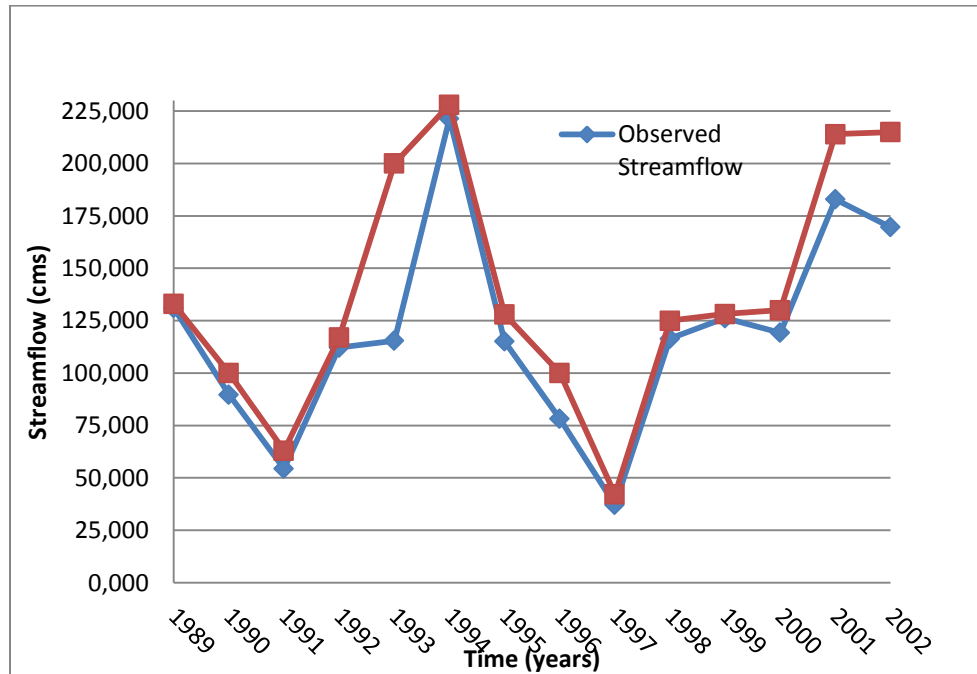


Figure 6.5 Observed and simulated streamflows (1989-2002)

6.4. Synthesis and partial conclusion

A SWAT hydrological model was developed for the Sirba watershed, and the model was calibrated after a sensitivity analysis. Evaluation of the model showed that the simulated monthly hydrograph reproduced the observations well, with a NASH value of 0.83. This value was found to be within the range of those obtained by several studies in the same region. Thus, the results of the Sirba SWAT hydrological model in this study can be considered acceptable. The calibrated Sirba SWAT hydrological model was further used in the study for a seasonal streamflow forecasting. For such purpose, two sets of models with changing parameters and models with constant parameters were developed and tested over the Sirba watershed to forecast seasonal streamflow using three predictors, namely, SLP, RHUM, and AirTemp. The simulation results showed that the models with constant parameters were deemed preminent for seasonal annual mean

streamflow forecasting, while the models allowing for parameter changes according to rainfall magnitude can be regarded as the best to forecast seasonal maximum monthly streamflows over the Sirba watershed.

The other developed seasonal streamflow forecasting model was based on the NLPCA method. This forecasting was done using the forecast rainfall obtained from the NLPCA seasonal rainfall forecast model, which were disaggregated to a daily scale and fed into the Sirba SWAT hydrological model. The result was fairly acceptable (NASH = 0.58) regarding the different kind of uncertainties that usually occurred in a forecasting and hydrological model development.

Based on the different forecasting models developed and tested, it can be concluded that models with changing parameters and models with constant parameters once associated with the Bayesian model selection approach can be a good way to forecast seasonal annual mean (respectively maximum monthly) streamflows in the Sahelian region.

Chapter 7. Seasonal Flood Forecasting

In this chapter, the development a seasonal flood forecasting model and the evaluation of the level of risk associated with seasonal forecasts are presented. A simple risk measure was used: the probability of overtopping of the flood protection dykes in Niamey.

In developing the seasonal forecasting model, many steps were taken and they are presented in this chapter. Among these steps, there is the development of a hydro-dynamical model of Niger River at Niamey, the copula method used to find the adequate copula that better models the observed streamflows from Garbé-Kourou and Kandadji. Also, the probabilities of overtopping of the flood protection dykes in Niamey are assessed.

7.1. Copula Approach

Formally, a d -dimensional copula is a distribution function defined on the hypercube $[0; 1]^d$ and whose margins are uniform. Sklar's (1959) theorem allows the connection of the notion of copula to the multivariate distribution function.

The concept of copula is a powerful and flexible tool, since it allows modeling the dependence regardless of the effect of the behavior of margins; that is to say, the $F_1; F_2; \dots; F_d$ distribution functions of variables $X_1; X_2; \dots; X_d$ considered alone.

In the modern theory of stochastic dependency, copulas are the tool of choice. Indeed, a copula contains all the pertinent information about the dependency structure of a vector of continuous random variables, and is used to isolate the effects of marginals.

Full details on how the copula was used in this work to assess the dependence between streamflows (at Garbé-Kourou and Kandadji) and the results from the analysis were presented in Paper 4 (section 7.4). Also, the copula that best fits the dependency structure was presented in that paper, as well as the criteria used for the selection.

7.2. Niger River Hydrodynamical Model at Niamey

In setting up a flood forecasting model, the type of model to be used is driven by the purpose of the model. Some criteria defined by WMO (2011) help in making a judicious choice of a real-time flood forecasting model. These are:

- Relationship between lead time / time of concentration
- Robustness of the model
- Necessary computational time

However, the last criterion should not prevent the testing of accurate and advanced (or sophisticated) models, despite their high computational time. In addition, there are limits and advantages of each set of models (physical-process models, data-driven models, conceptual models) that help to select an appropriate model for flood forecasting. The objective is to get a model that gives less uncertainty in forecasting floods.

Based on these criteria, the hydraulic model HEC-RAS and HEC-GeoRAS associated with ArcGIS were selected to develop the Niger River hydrodynamical model. They helped in studying and evaluating flood flows and flood-zone mapping in a range of the Niger River at Niamey (160 km sides Niamey). Peak flows before August spanning the period 1980-2013 were used, and the data are from two flow-measure stations: Kandadji (on the Niger River) and Garbé-Kourou (outlet of the Sirba watershed located on the Niger River). The model was calibrated to reproduce the last rating curve over the period 2009-2014 using the observed elevations data for Niamey and the bathymetry data (over 160 km sides Niamey) of the Niger River. A floodplain map of the city of Niamey was generated.

It was found that the calibrated model over the period 2009-2014 almost perfectly reproduced the last rating curve over the indicated period. The simulations showed the overtopping of the 6 levees during specific years, and this confirmed the periods that Niamey underwent flooding.

The remaining processes and results from the model development were presented in Paper 4 (section 7.4).

7.3. Probability of Overtopping

The most basic and important steps in floodplain management are the determination of the exact boundaries of the floodplain and the estimation of flood damage. Unfortunately, most models only predict the flooding without assessing the resulting damage. Thus, this process is very important for decision-makers in order to make preparations for the forecasted flood.

Flood risk assessment involves identification of the hazard, exposed assets, vulnerability, damages, and the losses, both economic and human. In this work, after developing the seasonal flood forecasting model, a simple risk measure was used to compute the probability of overtopping of the flood protection dykes at Niamey.

All details on computing the overtopping probabilities and the seasonal flood forecast for Niamey are presented in Gado *et al.* (2016) (see Section 7.4).

7.4. A Copula-based Approach for Assessing Flood Protection Overtopping associated with a Seasonal Flood Forecast in Niamey, West Africa

The model development and the achieved results were published in Gado *et al.* (2016), as detailed in the following.

A copula-based approach for assessing flood protection overtopping associated with a seasonal flood forecast in Niamey, West Africa

Abdouramane Gado Djibo¹, Ousmane Seidou², Harouna Karambiri¹, Hadiza Moussa Saley³, Ketevera Sittichok², Jean Emmanuel Paturel⁴, Nathalie Philippon⁵

1 International Institute for Water and Environmental Engineering (2iE), 01 BP 594 Ouagadougou, Burkina Faso

2 Department of Civil Engineering, University of Ottawa, Ottawa, ON, Canada

3 Centre Africain d'Etudes Supérieures en Gestion, Dakar, Senegal

4 Institut de Recherche pour le Développement (IRD), Abidjan, Côte d'Ivoire

5 Centre de Recherches de Climatologie, UMR6282 Biogéosciences CNRS, Université de Bourgogne, Dijon 21000, France; E-Mail: Nathalie.Philippon@u-bourgogne.fr

Abstract

Flood is one of the most important natural disasters that causes huge loss of life and properties every year around the world. Moreover, the International Federation of the Red Cross and Red Crescent Societies pointed out that floods were by far the greatest cause of homelessness. In West Africa, many countries are damaged from flooding almost each season. Thus, this study aimed to set a seasonal flood forecast model and carried out an evaluation of the level of risk associated with each seasonal forecast. Hydrologic Engineering Center's River Analysis System model was used to develop the hydro-dynamical model of Niger river at Niamey, then a simple risk measure was used to calculate the probability of overtopping the flood protection dykes in Niamey. The results show that the hydro-dynamical model reproduced well the rating curve over the period 2009-2014. Also, the Gumbel copula method was found to be the best among the tested 5 copulas with a Ks-value of 71.61%. The copula analysis pointed to a dependency between the tributaries contributing to the seasonal flood at Niamey. It is found that for the six dykes the probabilities of being overtopped by the flood are very high to low, varying from 100% to 16.67 %.

Keywords: Seasonal flood forecast; HEC-RAS model; Dykes; Copula; West Africa

1. Introduction

Climate variability is known as a short-term deviation from the normal of the climate at a given place. Deviations affect parameters such as temperature, wind and precipitation. In different parts of the world, the climate is also distinguished by extreme weather events such as droughts, flooding, and so on. These extreme weather events cause recurrently

worldwide significant human loss and economic disasters. This is the case, for example, of floods in the Sahel in West Africa. Climate change is thought to be the reason behind the increase in magnitude and frequency of extreme weather events such as droughts, intensive precipitations, floods, strong winds, storm surges that were reported in several recent publications (UNDP, 2010; IPCC, 2007). The extreme events have severe consequences for societies in underdeveloped countries where the populations are more vulnerable due to their low adaptive capacity (Scheraga & Grambsch, 1998); this fact was specifically detailed in the study by Heather (2014) who particularly emphasized on the lack of sustainable solutions for communities to fight against these calamitous phenomena.

Flooding can be defined as the overflowing by water of the normal confines of a stream or other body of water, or accumulation of water by drainage over areas that are not normally submerged (WMO, 2011). Different types of floods are mentioned in the literature and they can be categorized into 10 groups: Flash floods, fluvial (riverine) floods, single event floods, multiple event floods, seasonal floods, coastal floods, estuarine floods, urban floods, snowmelt floods, and ice-and debris-jam floods. For more details on these floods, the reader can refer to WMO (2011); Antoine et al. (2008); Dasylva (2009); Laganier et al. (2004) and Ngoran et al. (2015).

Flood is one of the most important natural disasters which cause extensive loss of life and properties every year around the world. According to WMO statistics on the types of water-related natural disasters in the world (WMO 2011), floods have been more frequent and had more economic impact on the population as compared to other extreme events during the last two decades. Indeed, the United Nations Educational, Scientific and Cultural Organization (UNESCO) showed in its water assessment program that 15% of deaths issued from natural disasters are due to flooding. For instance the Asian continent was affected by 44% of flood disasters during the period 1987-1997, that led to the loss of 228 000 lives (Awotwi et al., 2015). Moreover, the International Federation of the Red Cross and Red Crescent Societies (IFRC) pointed out that floods were by far the greatest cause of homelessness. In West Africa, many countries have been recently damaged by flooding: Burkina Faso and Niger have been hardly hit by heavy flooding in 2009 and faced many emergency situations. At least 30 people were killed and 350,000 displaced

when torrential rains soaked much of West Africa in September 2009. In 2007, 300 people died and 800,000 were affected by the inundations (OCHA, 2009). It is particularly found that the Niger River and its tributaries (at upstream) always contribute to the flooding issue in Niger especially at Niamey city. Many dead and displaced populations were registered during such flooding events. In 2013, the economic losses induced by flooding in Niger were estimated to 32 billion FCFA (OCHA, 2013).

Therefore, there is a great need to forecast flood in order to provide early-warning about impending flood to stakeholders to alert the population in order to reduce the probable losses. Seasonal flood forecasts are of prime importance for Niamey city due to its location at the banks of the Niger River. Seasonal floodings experienced annually by the city generate the displacement of the local population living near the river, economic loss and psychological impact on victims. It should be noted that Niamey is a city that lacks of adequate resources, so the funds disbursed to aid the victims should have been invested in the development of the city in order to provide better living conditions to the population. Hence, the seasonal flood forecast should help Niamey city to avoid or to lessen the overall impacts of a flood and also by saving lives, livestock and property.

In a broad sense, flood-risk assessment combines a frequency analysis of extreme hydrological phenomena and an evaluation of the flood damages.

This article discusses the hydrological variability by assessing the potential damages of induced flood events over seasons. The development of a seasonal flood forecast model for Niamey city, involves the development of a hydro-dynamical model of Niger at Niamey using HEC-RAS model, and applying probability distributions to generate probabilistic forecast for each year. Thus, the aim of this study is to develop a seasonal flood forecast model in Niamey and to quantify the probabilities of flood protection dykes to be overtopped. In attaining this objective a copula-based approach is chosen.

2. Seasonal flood forecasting approach

The performance of an engineered flood protection system must be evaluated over the system's lifetime. Consequently, the hydrologic conditions can be described as a function of time, particularly if the river environment changes over time due to the impact of land-use change or because of climate perturbations. In order to tackle such issues, flood

forecasting is needed to help taking the necessary and adequate measures to avoid such flood or to reduce its impacts. The critical part of any flood forecasting system is the hydrological model (i.e., rainfall-runoff model). It is for this reason that Serban & Askew (1991) considered the hydrological model as the element that ensures the success and efficiency of a flood forecasting system. Thus, the choice of the rainfall-runoff model is fundamental for flood forecasting attainment in order to produce the exact flood projections based on input data (hydro-meteorological and other data). Flood forecasters usually rely heavily on real-time data about rainfall and river water levels as well as rainfall forecasts (Najafi et al., 2011). But there are uncertainties in long-range forecasts, noticeably biases in precipitation frequency (i.e., too many rainy days), and intensity (i.e., too small precipitation values). Thus bias correction of seasonal rainfall forecasts is essential before using them as inputs for hydrological models. To circumvent such process, we developed a new approach for seasonal flood forecast for Niamey city. The seasonal flood forecast model for Niamey city was developed based on the probabilistic forecasted streamflows. It involves the development of a hydro-dynamical model of Niger at Niamey using HEC-RAS model, and applying probability distributions to generate probabilistic forecast for each year. The developed seasonal flood forecast model is a tool that could help stakeholders to take adequate decision to reduce the resulting damages. Also, the quantification of probabilities of flood protection structures to be overtopped could help a lot to avoid high induced economic losses. Therefore, the main idea behind this seasonal flood forecast model is that any flood forecasting model or method has to consider the will of decision makers (Klemes, 2002; O'Connor, 2006; Perrin et al., 2001).

3. Materials and methods

3.1 Study Area

Two areas are considered in this study, namely, the Niger River at Niamey city for flood damages and the Sirba watershed (Sirba tributary) for flow prediction. Located in western Niger between longitudes 2° 01' - 2° 14' East and latitudes 13° 25' - 13° 36' North, Niamey is landlocked in the department of Kollo. It is bounded to the north-east by the rural town of Hamdallaye, at east and south by the rural town of Liboré and at west by the

canton of Lamordé and the rural town of Karma (Tinni, 2011). Niamey, capital of Niger, is built on both banks of the Niger River. The left bank houses the districts I, II, III, and IV; while the fifth district is on the right bank. These five districts give rise to a set of 96 quarters occupying a surface of 239.263 km² (Tinni, 2011). The Figure 1 presents the location and characteristics of Niamey city.

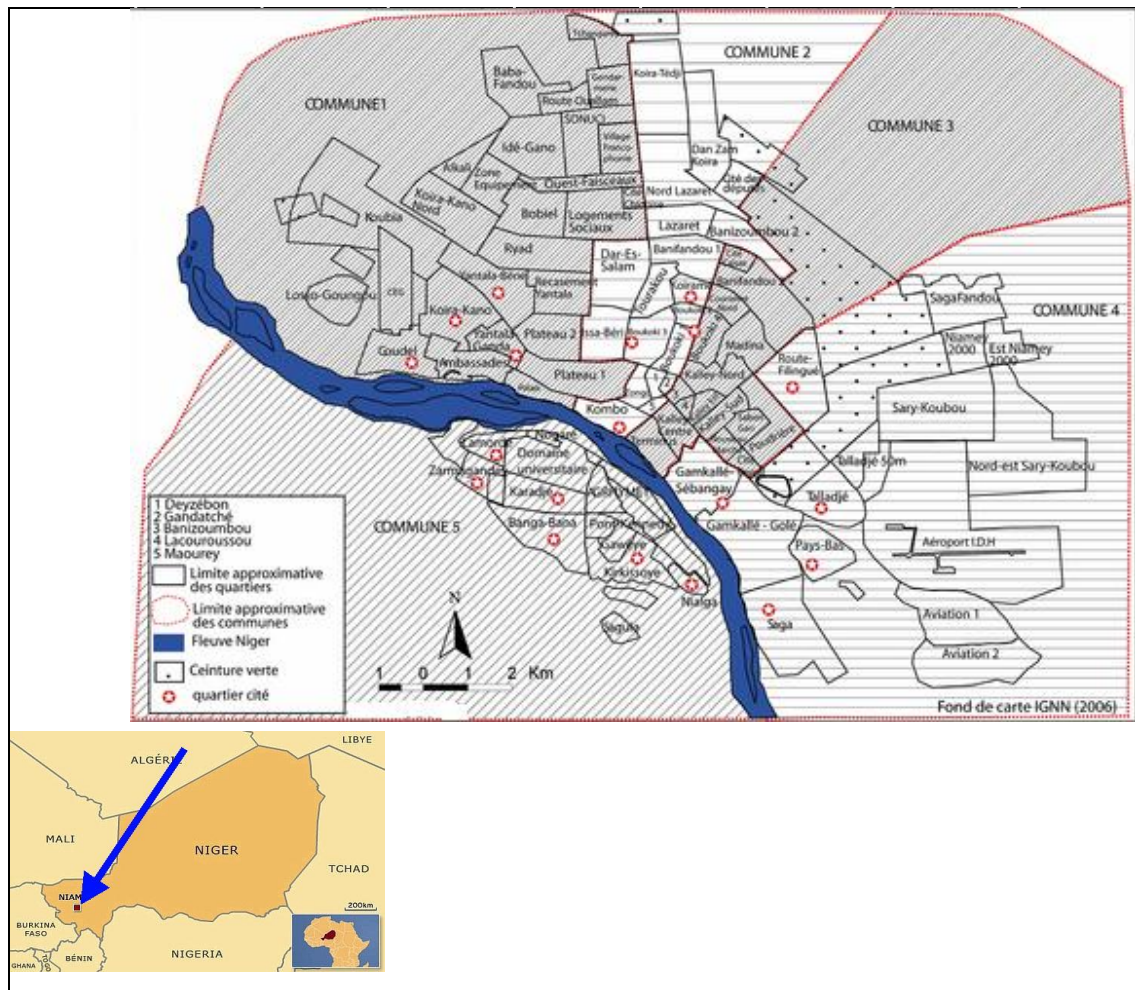


Figure 1: Description of Niamey city (Adapted from Sanda, 2010)

The development of Niamey falls within the administrative needs of the colonial authorities and subsequently of the Niger governments. The city spread diffusely without proper planning developments. Currently, the urban area at the right bank is experiencing unparalleled growth driven by ever-increasing rural exodus (Diaz Olvera et al., 2002; Gilliard, 2005). This spatial expansion meets the growing demographic needs. Its population increased from about 30,000 inhabitants in 1960 to 1,222,066 people in 2010,

that is, from 1 to 8% of the total population in fifty years. This growing of the demographic weight confirms the importance of rural-urban migration and high annual population growth rate (3.3%) (Motcho, 2006; INS, 2010). This high demographic and spatial pressure also generates land issues involving competition between housing, industry and agriculture, and the use of vulnerable space, for instance the Niger River banks are without any prior planning. Thus, this work is carried out on the portion of the Niger River at Niamey (along 160 km upstream and downstream) in order to seek ways on how to tackle this flooding issue. The choice of Niamey city is motivated by the fact that the city experienced almost each year seasonal flooding which usually led to the displacement of the local population living near the river, and huge economic losses. Moreover, this area offers the opportunity to carry out such kind of study because the city undergoes an anarchic occupation and the population has low fiscal revenue to cope with the issued consequences.

Concerning the Sirba watershed, it is a transboundary watershed in between Burkina Faso and the Niger Republic. The Sirba watershed lies within the latitudes 12°55'54''–14°23'30''N and longitudes 1°27'W–1°23'42''E with an area of 38,750 km² (Mara, 2010). The Sirba extends over three sub-climate zones based on the mean annual amount of rainfall decreasing from south to north: a southern Sudanian zone with mean annual rainfall between 700 and 800 mm, a northern Sudanian zone with mean annual rainfall ranging from 550 to 650 mm and a Sahelian zone with mean annual rainfall of 300 to 500 mm. The highest quantities of rainfall are observed during the months of July to September (JAS), regardless of the climate zone. The climate is generally characterized by the presence of two seasons: a dry season (October–April) due to the harmattan (dry hot wind) and a rainy season (May–September) due to the presence of the West African Monsoon (cold wet wind). The hydrographic network is relatively dense and consists of three main tributaries (Sirba, Faga and Yeli) as well as some small reservoirs from five dams with non-operated spillways (Figure 2). Based on the description of the rainfall pattern, the hydrological regime in the Sirba watershed is a Sahelian type, while its vegetation formation is thorny and of a lightly wooded savannah (Gado Djibo et al., 2015). The reason for choosing the Sirba basin is due to the fact that the Sirba tributary contributes to the local flood of Niger River. It is also one of the most important affluent

of Niger River; and it is specifically chosen to predict the seasonal flows in order to develop the seasonal flood forecast model.

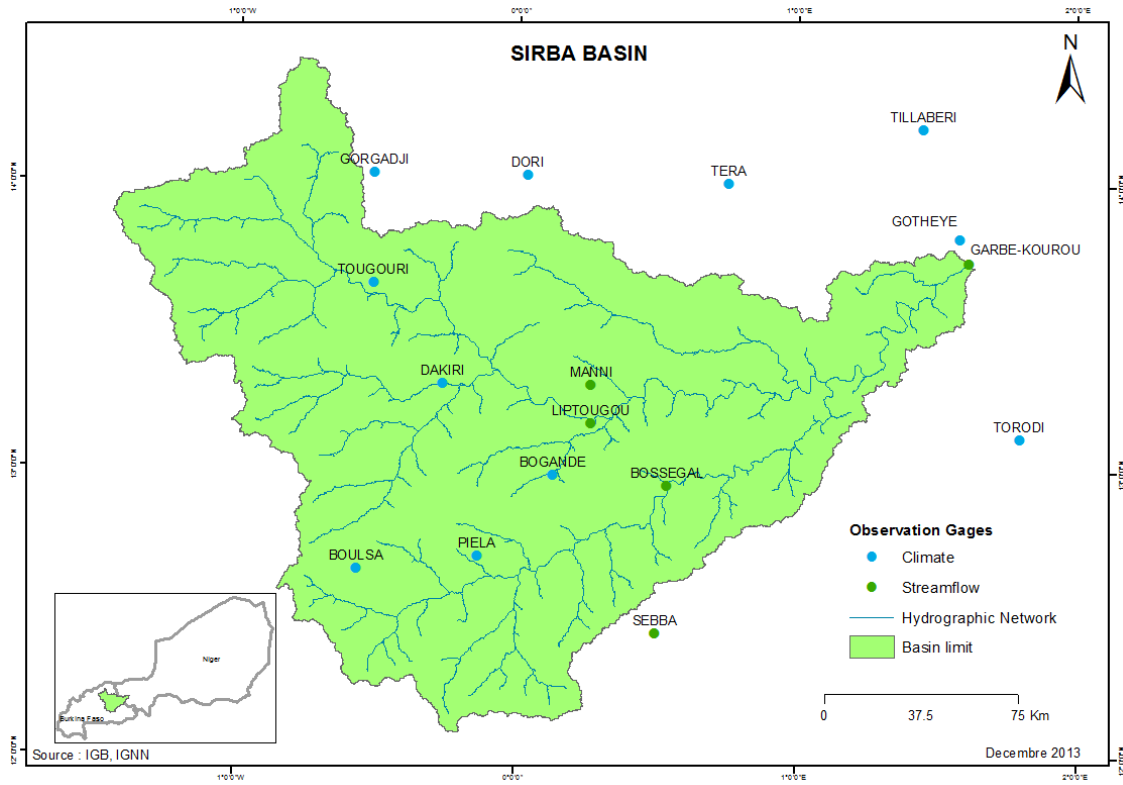


Figure 2: Location of Sirba watershed (source: Gado Djibo et al., 2015)

3.2 Modeling the dependence of flood peak at the Kandadji and Garbe - Kourou station using Copulas

Let X and Y be random variables with distribution function F_X and F_Y . We called copula (2-dimensional) any distribution function C whose marginal bivariate is a uniform distribution on $[0, 1]$. In other words, C satisfies the four properties of a bivariate distribution function in addition to the following 4 equations:

$$C(x, y) = 0 \text{ if } x \leq 0 \text{ or } y \leq 0 \dots\dots\dots (1)$$

$$C(x, y) = 1 \text{ if } x \geq 1 \text{ and } y \geq 1 \dots\dots\dots (2)$$

$$C(x, y) = x \text{ if } y \geq 1 \dots\dots\dots (3)$$

$$C(x, y) = y \text{ if } x \geq 1 \dots\dots\dots (4)$$

Copula theory allows to model the dependence structure between variables independently of marginals (Sklar, 1959). It has been widely applied in various fields, such as biology (Frees & Valdez, 1998) and Insurance and Reinsurance (Klugman & Parsa, 1999;

Belguise, 2001; Venter, 2002, 2003; Charpentier, 2003). In hydrology, Favre et al. (2004) present two applications of copula: the first concerns the modeling of flow rates upstream and downstream of a river with tributary, and the second concerns the joint modeling of annual maximum flows based on the volume of runoff. De Michele et al. (2005) used the Gumbel copula to model the positive dependence between the peak flow and flood volume. Zhang & Singh (2006) used bivariate copulas for frequency analysis of the variables: peak flow and volume of runoff, and flood volume and its duration.

In this study, the copula method is applied for the analysis of the dependence structure for the couple of streamflows at Kandadji and at Garbé-Kourou. This helps to check which flow station contributes more to the flooding occurring at Niamey city along the Niger River. This checking was made in 3 main steps detailed in the next paragraphs.

a. Identification and adequacy of marginals

Eleven statistical distributions were applied and tested over the streamflow data series of Kandadji and Garbé-Kourou. The distributions are individually tested using the maximum likelihood method to compute the parameters of each distribution. The statistical criterion used to assess the goodness of fit is the Kolmogorov-Smirnov (Ks) test (Wilks, 2006). All statistical distributions were ranked based on the goodness of fit values, where the lesser value denotes a high rank. Additionally, the p-value is also used because the higher the p-value the lower the statistical value of the goodness of fit. This ranking helps classifying the probability distributions and eases the identification of the best distribution. The probability densities of the best fit were computed. More details on these tests are provided in Wilks (2006) and Solaiman (2011).

b. Detecting dependency

To characterize and measure the dependence between flow at Kandadji and Garbé-Kourou, two criteria are used: the Spearman (ρ_s) and Kendall (τ) rank correlation coefficients. These two correlation coefficients are the most suitable as they only depend on the rank of each observation compared to Pearson correlation coefficient which depends on variables values. Furthermore ρ_s and τ are invariant under strictly increasing transformations of random variables X and Y.

In addition, distributions of ρ_s and τ are near normal distribution of average zero and respective variances $\frac{1}{1-n}$ and $\frac{2(2n+5)}{9n(n-1)}$. So random variables X and Y can be considered as dependent on α threshold, if: $\sqrt{(n-1)}|\rho_s| > z_{\alpha/2}$ and $\sqrt{\frac{9n(n-1)}{2(2n+5)}}|\tau| > z_{\alpha/2}$ respectively (Emna & Assia, 2008). Therefore, the dependence parameter was estimated using the above conditions.

c. Selection and adequacy of the copula

Five copulas were used to find the dependence structure. They are t-copula, Gaussian copula, Frank copula, Gumbel copula, and Clayton copula. It consisted to compare the empirical values and the estimated values for the different copula. This comparison is made using the Kolmogorov-Smirnov test (Marsaglia et al., 2003) to a threshold α . This test is recommended by several studies such as Genest et al. (2006) and Genest & Rivest (1993). Also, the quantile-quantile (QQ) plot is performed after resampling because it helps checking ultimately if the convergences between the distribution of series being tested and those of reference are very obvious.

3.3 Hydro-dynamical model development for Niger river at Niamey

In this study, the hydraulic model HEC-RAS (Hydrologic Engineering Center's River Analysis System) and HEC-GeoRAS associated to ArcGIS were used to study and evaluate flood flows and map flood zone along 160km upstream Niamey.

HEC-RAS, a one-dimensional, hydraulic-flow model was used to model the Niger River at Niamey. The HEC-RAS program has been developed by the US Army corps of Engineers since its inception in 1964, and different sub-programs have been continuously established. It was designed specifically for application in floodplain management and flood-insurance studies to evaluate floodway encroachment and to simulate estimated flood inundation (U.S. Army Corps of Engineers, 2009; Soleymani et al., 2014). This model is used worldwide. Indeed, Johnson et al. (1988) used HEC-RAS model to predict and determine the desired land boundaries, 10 km along the Wyoming River - Gary Bull in America. Using the model, the water surface profile of the river was plotted. Using HEC-RAS software, David and Smith (2000), evaluated the hydraulic behavior of the

flood. David, et al. (2002) studying flood for a period of 5 years in the United States, prepared flood zone maps. Knebl et al. (2005) using the hydrological model, HEC-HMS, and hydraulic model, HEC-RAS, and radar precipitation estimate (NEXRAD) in the basin of San Antonio, Central Texas, United States suggested logic model for flood, and compared the model with the summer 2002 flood. Results showed the model efficiency for regional-scale flood forecasting. Napradin et al. (2006) made flood hazard mapping for small watersheds near Baya Sea in Astore valley. For this purpose combination of Wet Spa and HEC-RAS was used.

The HEC-RAS model was used to compute water-surface elevations and develop flood-inundation areas along the study area for the flood discharges. Peak flows before august were used (1980-2013), and they are from two flow measurement stations, Kandadji (on Niger River) and Garbé-Kourou (exit of the Sirba watershed located at Niger River).

The steps in building the hydro-dynamical model, and the required data at each step, are briefly explained in the next paragraphs as they are standard steps. It consisted first of building the import file (digital terrain model (DTM) of the Niger River system) in the ArcInfo TIN (Triangular Interpolation Networks) format under HEC-GeoRAS environment. This helps setting-up analysis environment for HEC-GeoRAS. The geometry file containing the RAS layers and line themes were then created. It should be noted that two nodes were specified in this model, and also adequate number of cross-sections cut lines. 6 levees were set in this model; these levees protect the city of Niamey against flooding from Niger River. The geometry data were then imported into HEC-RAS then the model was run for some simulations after entering flow data and boundary conditions. These simulations helped to obtain water surface profile and velocity data after processing.

The developed HEC-RAS model was calibrated to reproduce the last rating curve over the period 2009-2014 using the observed elevations data at Niamey and bathymetry data of Niger River. A flood inundation map of Niamey city was generated after exporting the acquired data from the HEC-RAS simulations to HEC-GeoRAS, for GIS analysis of floodplain mapping.

3.4 Assessment of potential damages related to the seasonal flood forecasts

Economic loss assessment based on established vulnerability maps are crucial for West African Sahel especially the city of Niamey, as the recurrent flood events can lead to the loss of human lives and valuable properties. Also, another potential benefit resides in the thought of some specialists (Laganier et al., 2004; Dasylyva, 2009) who believe that remedial approaches and quick technical solutions are less effective for dealing with flood events than projected measures. With increased accessibility to digital information and the effectiveness of computer-based analysis, GIS has played an important role in hydrologic and hydraulic modeling. The main advantage of using GIS in modeling is the considerable potential for extracting digital information from a digital elevation model (DEM). Thus, the probabilistic forecast streamflows obtained from Gado et al. (2015) are used in the hydro-dynamical model of Niger River at Niamey to find the probabilities of overtopping the dykes protecting the city of Niamey against flood. These probabilities were obtained after running the hydro-dynamical model and considering the number of dykes (among the six) that each peak flow can overtop. These probabilities of overtopping the dykes explain the extent of the danger of floods at Niamey.

4. Results and discussions

4.1 Niger River HEC-RAS model

The Niger River hydro-dynamical model was developed by combining HEC-GeoRAS and HEC-RAS computer programs to extract floodplain cross sections from the elevation contour data of the City of Niamey. A total of 988 cross sections were extracted along the study reach of 160 km. All cross sectional profiles were drawn and the longitudinal profile for each of the 34 years (1980-2012). Typical HEC-RAS cross sectional and longitudinal profiles obtained for the year 2012 are shown in Figures 3 and 4, respectively. The developed model contains 6 levees which were completely over flooded in 2007, 2009, 2010 and 2012, the years when Niamey actually underwent several. The floods were computed at all cross sections from the HEC-RAS model. An analysis of all those profiles showed that the year 2010 was the worst in terms of negative impacts on the City of Niamey, followed by the year 2012. This finding is justified by the records of

the government through its early warning center. In addition, in 2012, the flood led to the collapse of more than 3,000 houses and 23,000 people affected (Sighomnou, 2013).

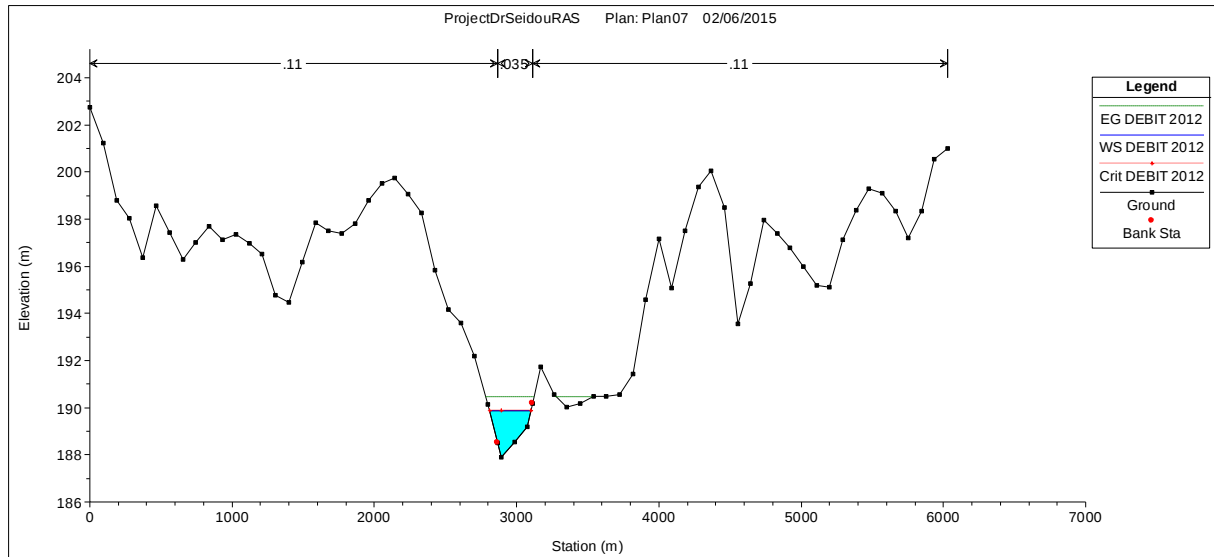


Figure 3: Example of a cross sectional profile for the year 2012 streamflow

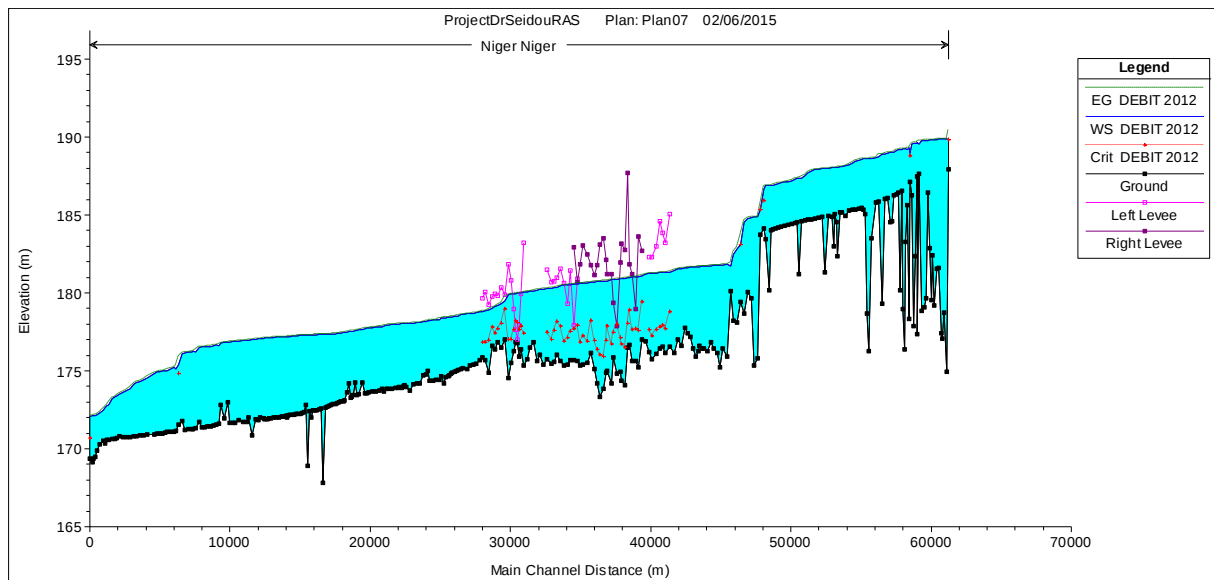


Figure 4: A longitudinal profile of Niger River at Niamey for 2012 flows

Figure 5 depicts the calibration of the seasonal flood forecast model. It can be easily seen that the model calibrated over the period 2009-2014 reproduces almost perfectly the last

rating curve over the indicated period. The two curves match well confirming how satisfactory the calibration is.

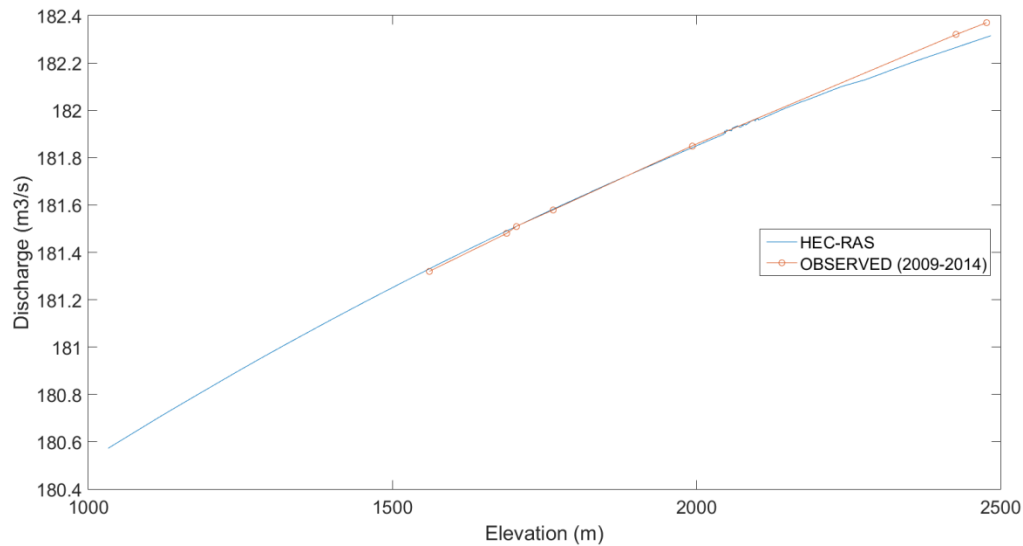


Figure 5: Calibrated seasonal flood forecast model: simulated (blue) and observed (red)

This is the first model for Niger River developed using the HEC-RAS model. However, the limitations and errors associated with developing flood-inundation areas are dependent on the topographic data, hydrologic data, and one-dimensional hydraulic modeling used in a floodplain study.

4.2 Kandadji and Sirba streamflows dependency on flooding at Niamey

a. Goodness of fit test

The streamflow peaks at Garbé-Kourou and Kandadji were analyzed over the period 1980-2013. The best fit was obtained with the generalized Pareto (GPAR) distribution for flows at Garbé-Kourou (Figure 6) while the Log-Pearson III distribution (LP-III) was found as the best fit for flows at Kandadji (Figure 7). These distributions were the best at 95% confidence interval. It can be seen on these figures that all the observations align with the fit and fall within the confidence range.

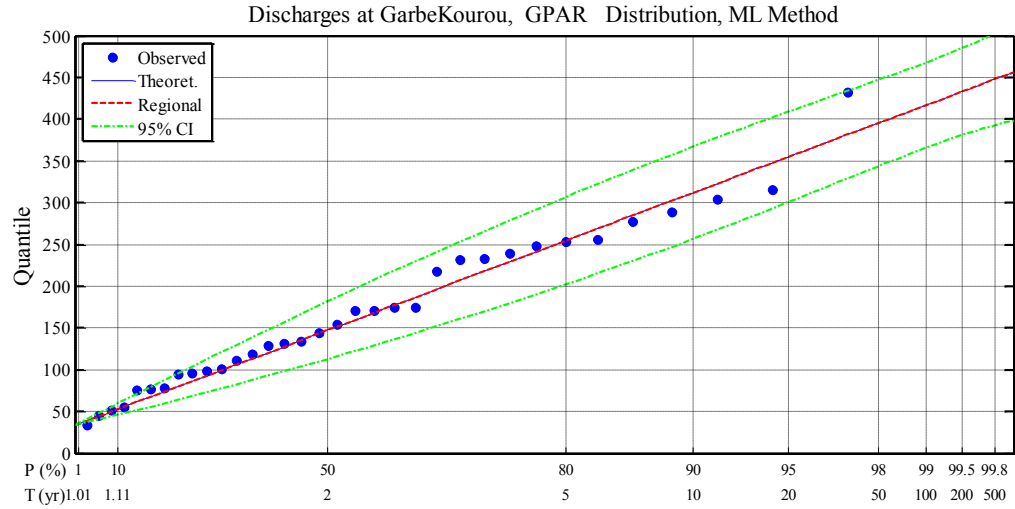


Figure 6: Best probability distribution fit (GPAR) for streamflow peaks at Garbé-Kourou

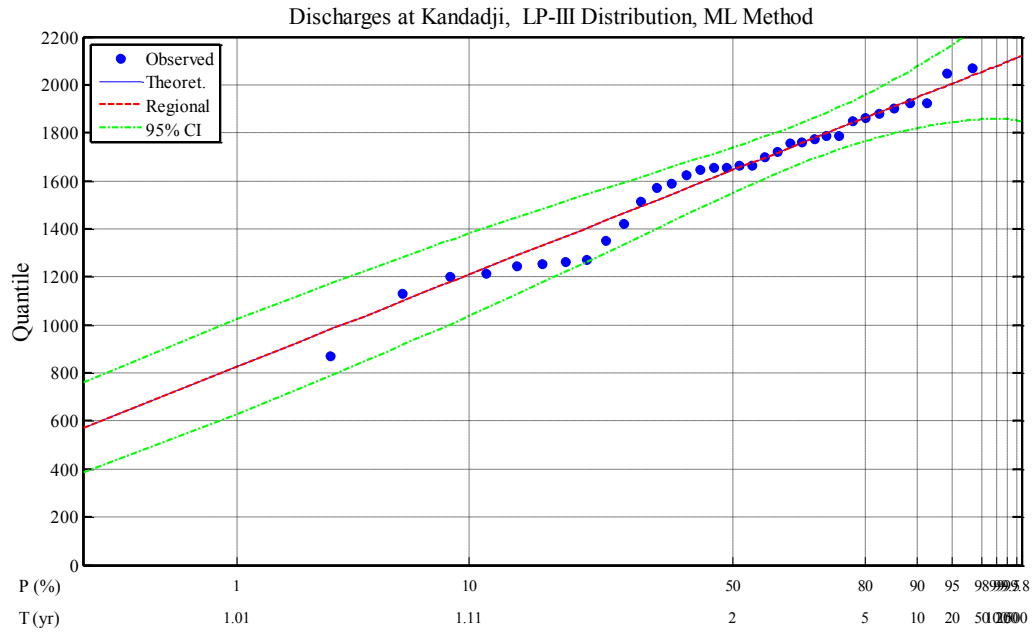


Figure 7: Best probability distribution fit (LP-III) for streamflows at Kandadji

The Figures 8 and 9 present the tested distribution for comparing the 11 probability distributions for flows at Kandadji and Garbé-Kourou, respectively. In addition, Table 1 presents the results for the statistical distributions that gives the best fit.

Table 1: Parameters of the best fitted distributions for streamflows peaks at Kandadji and Garbé- Kourou based on Ks test.

Distribution	Ks test		Parameters
	Statistic test	p-value	
Log-Pearson III (LP-III)	0.10	0.846	$\gamma=7.66$; $\alpha=-0.13$; $\beta=2.22$
Generalized Pareto (GPAR)	0.07	0.913	$e=33.22$; $\alpha=189.88$; $k=0.43$

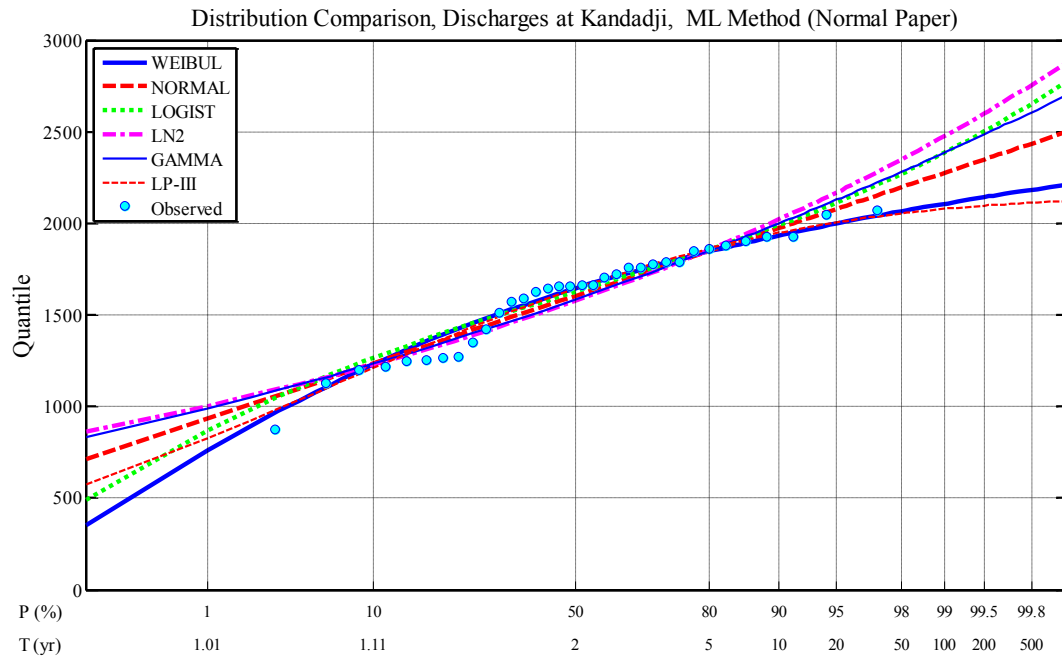


Figure 8: Distribution comparison of 6 probability distributions for streamflow peaks at Kandadji

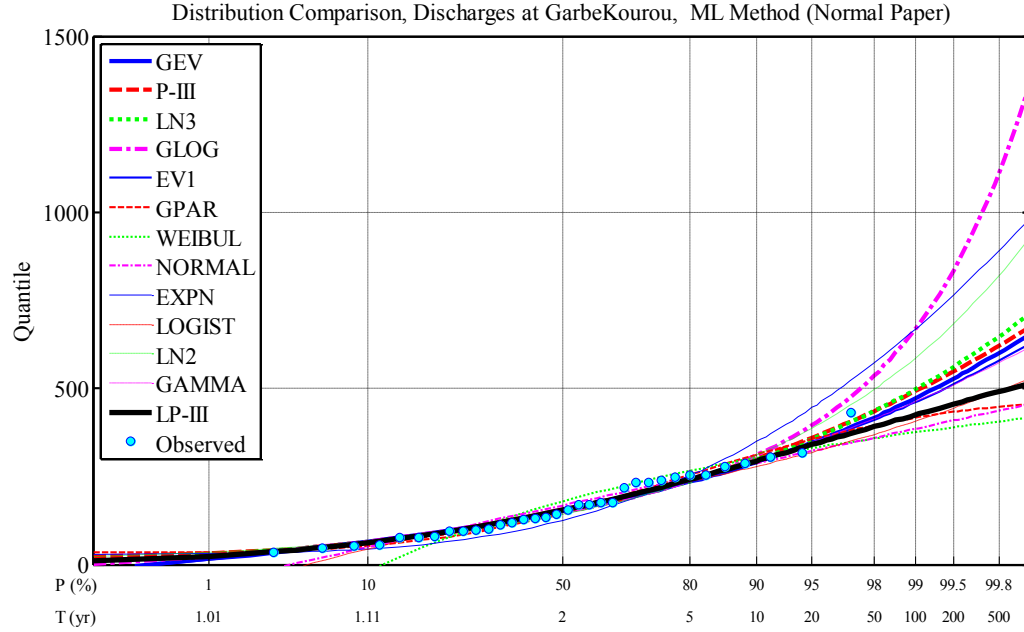


Figure 9: Distribution comparison of 11 probability distributions for streamflow peaks at Garbé-Kourou

b. Copula analysis

The dependence between streamflows at Kandadji and Garbé-kourou was measured by calculating the rank correlation coefficients of Spearman and Kendall. These coefficients were 0.447 and 0.385 respectively.

To estimate the dependence parameter, the hypothesis of independence between the streamflows at Kandadji and Garbé-Kourou was tested using the Spearman and Kendall rank coefficients obtained. It was found on one hand for the Spearman coefficient that

$\sqrt{(n-1)}|\rho_s| = 2.52 > 1.96 (z_{\alpha/2})$, and on the other hand for the Kendall coefficient

$\sqrt{\frac{9n(n-1)}{2(2n+5)}}|\tau| = 151.78 > 1.96 (z_{\alpha/2})$. These values correspond to $z_{\alpha/2}$ at $\alpha=5\%$

threshold. $z_{\alpha/2}$ was obtained from the table of Z-test at $\alpha=5\%$.

These values indicate that the hypothesis of independence between these two variables is to be rejected, thus there is dependence between these streamflows.

About the suitability of the copula, the data were transformed to the copula scale (unit square) using appropriate cumulative distribution function (CDF) to transform X (streamflows at Garbé-Kourou) to U and Y (streamflows at Kandadji) to V, so that U and

V have values between 0 and 1. The best distribution for X is the generalized Pareto distribution, while Log-Pearson III distribution is the best distribution for Y. Figure 10 shows a scatter-plot of the initial streamflows data of Kandadji and Garbé-Kourou, and Figure 11 displays the transformed streamflows to unit in order to be fitted.

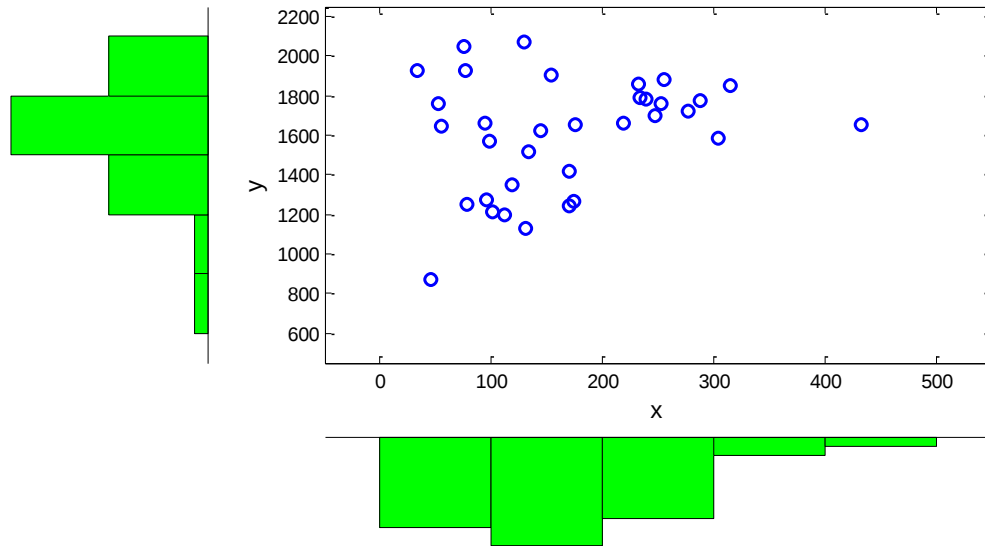


Figure 10: Scatter plot of observed streamflows at Garbé-kourou (X) and Kandadji (Y)

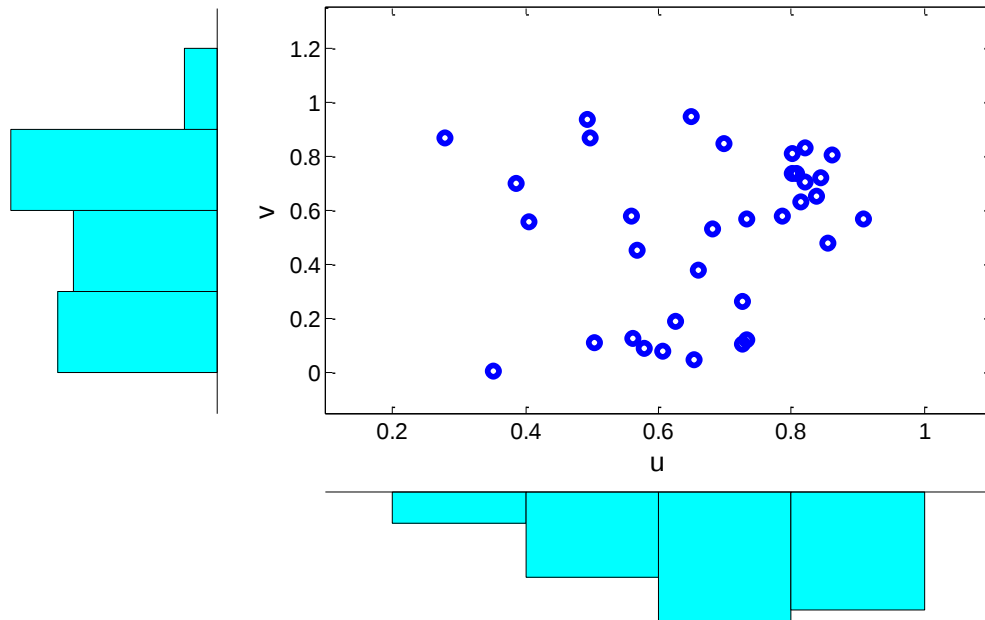


Figure 11: Cumulative distributions of streamflows at Garbé-kourou (U) and Kandadji (V)

After transforming these data to unit, each of the five copulas was fitted to the obtained couple (U, V) by generating first a random sample from the corresponding copula (U_{sim}, V_{sim}) . Thereafter, using the appropriate inverse cumulative distribution function, the previous couple (U_{sim}, V_{sim}) was transformed to X_{sim}, Y_{sim} , so that to transform the random sample back to the original scale of data.

Figures 12 to 16 present the fitted copula and the transformed copula back to the initial scale for t , Gaussian, Gumbel, Frank and Clayton copulas, respectively. It is observed on the lower panels of those figures that there is dependence between the streamflows. They also showed that the scatter plots are mostly located on the positive parts of the axes.

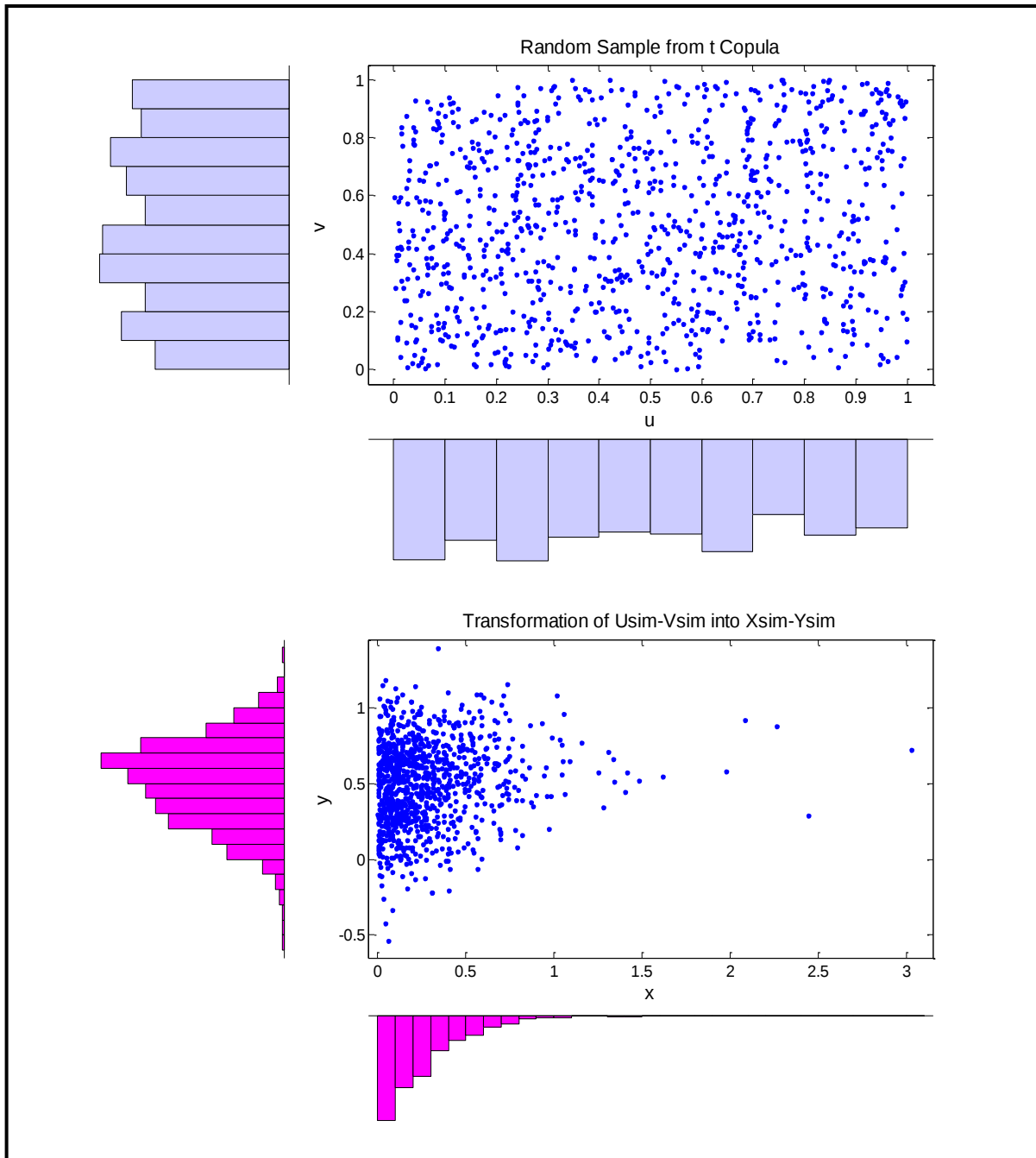


Figure 12: Fitted t copula using random sample (upper panel) and transformed copula back to the initial scale (lower panel).

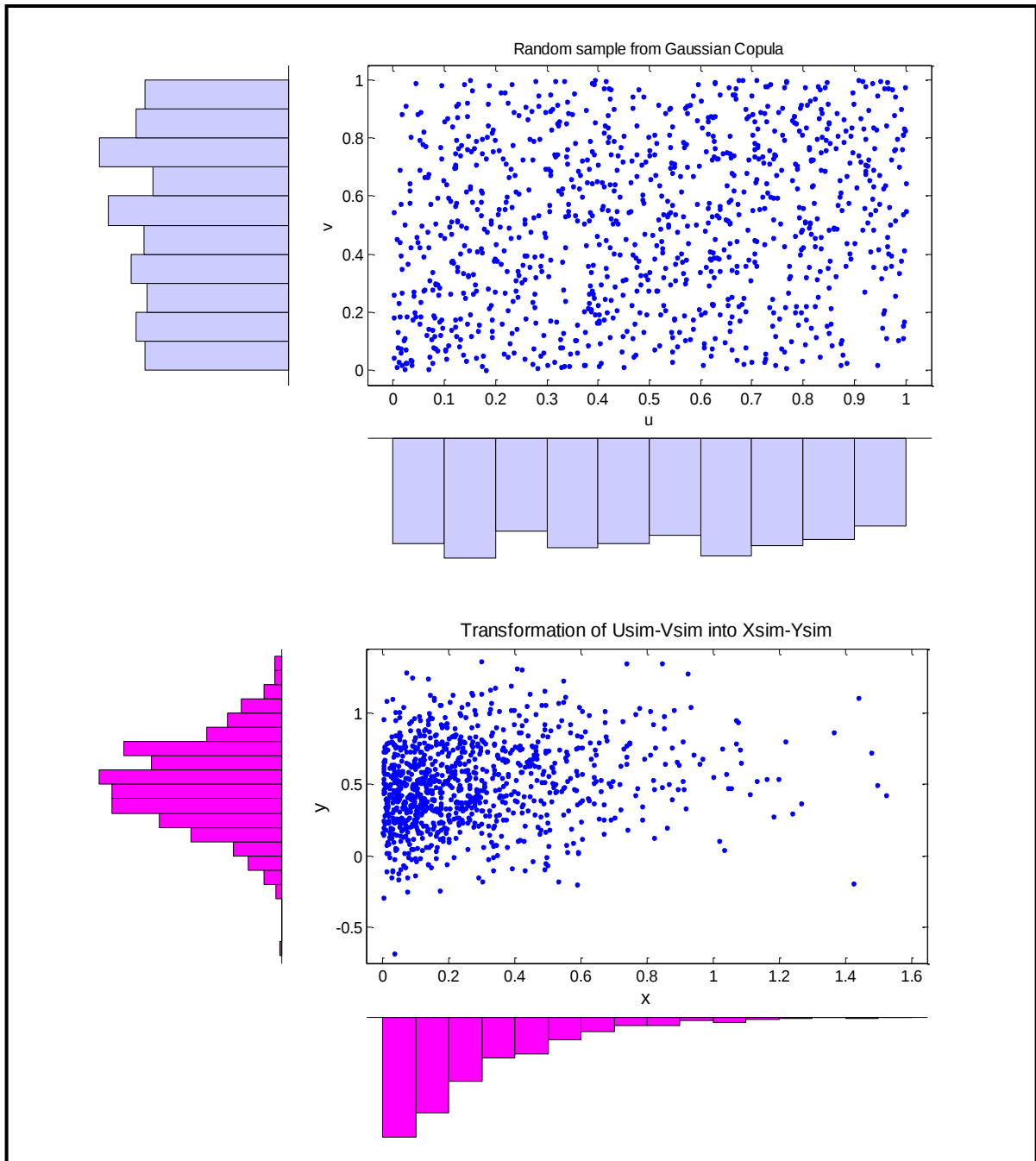


Figure 13: Fitted Gaussian copula using random sample (upper panel) and transformed copula back to the initial scale (lower panel)

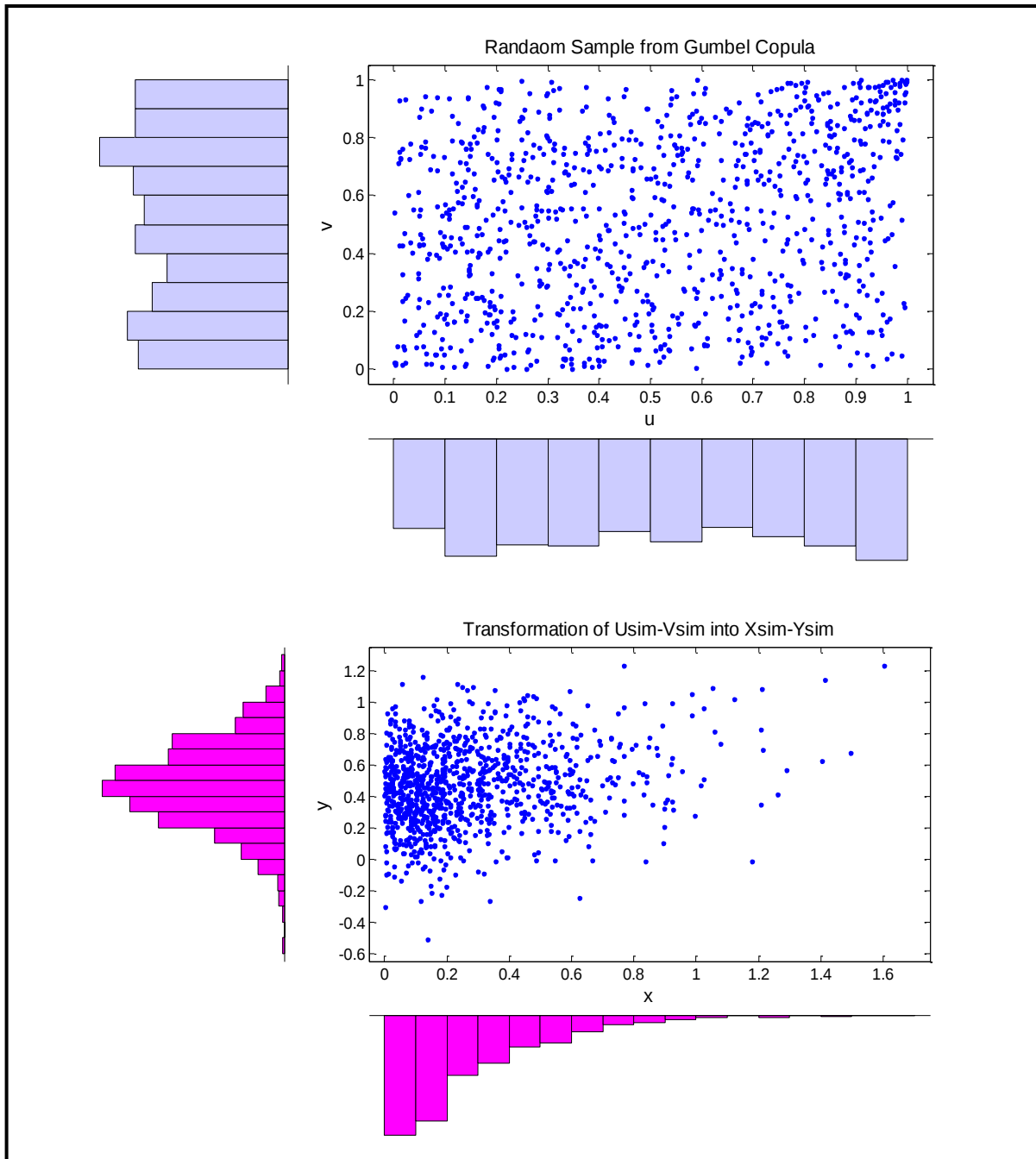


Figure 14: Fitted Gumbel copula using random sample (upper panel) and transformed copula back to the initial scale (lower panel)

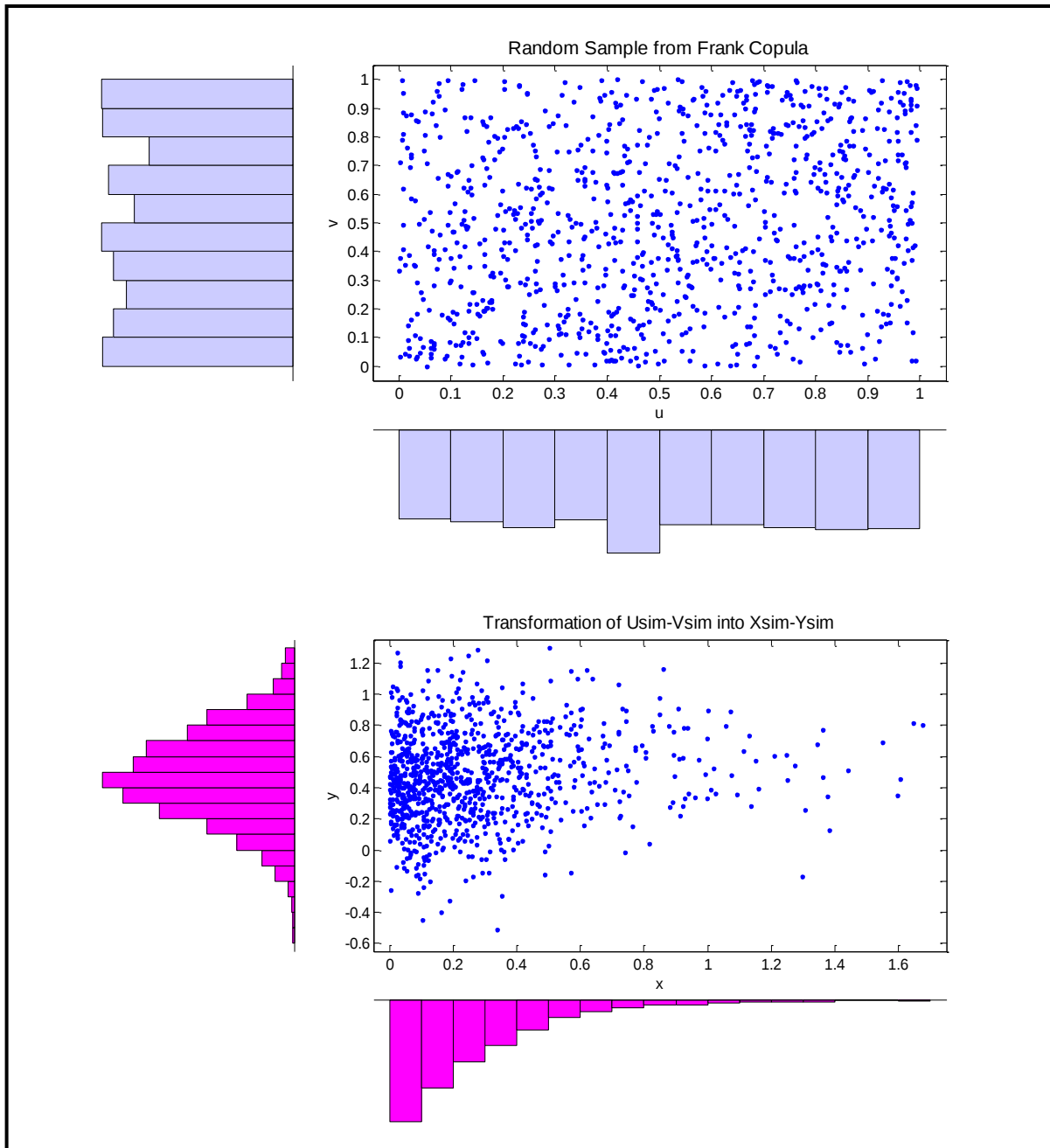


Figure 15: Fitted Frank copula using random sample (upper panel) and transformed copula back to the initial scale (lower panel)

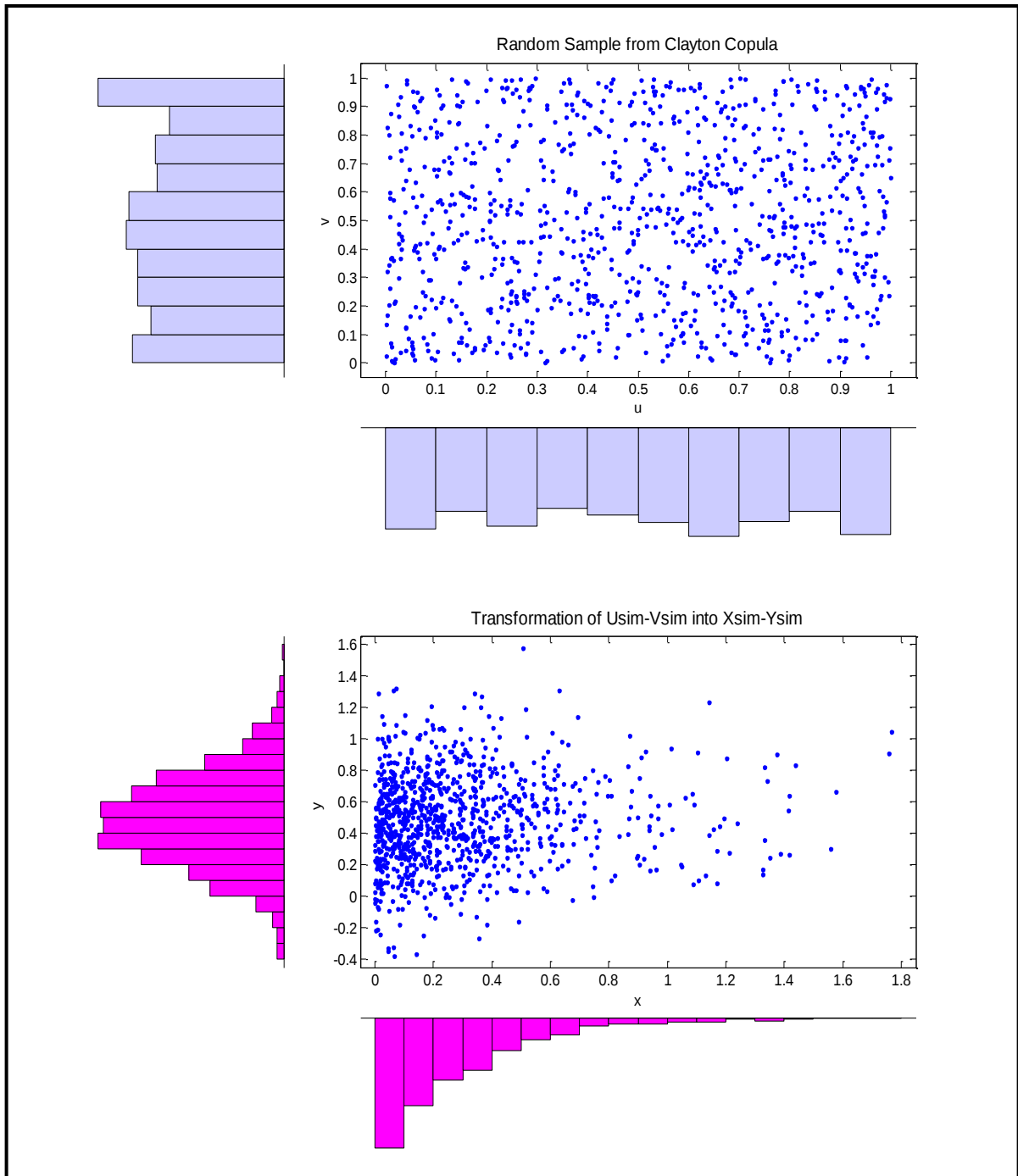


Figure 16: Fitted Clayton copula using random sample (upper panel) and transformed copula back to the initial scale (lower panel)

The great difficulty is to choose the copula that best capture the dependency between data structures. To choose the copula that models the best observations, a comparison is made

by a Kolmogorov-Smirnov. Therefore, the analysis of the 5 copulas based on the Kolmogorov-Smirnov test show that the Gumbel copula is the best as it has the lowest statistic value (0.031) and highest p-value (0.716) as shown in Table 2. Moreover, regarding the value of Spearman and Kendall rank coefficients, the Gumbel copula was found as the best with 0.71 and 0.80, respectively. However, the final selection was based on the QQ-plots (Figure 17) as after resampling they allow ultimately to decide. Examining these figures, it is obvious that the Gumbel copula gives the best fitting compared to the remaining 4 copulas.

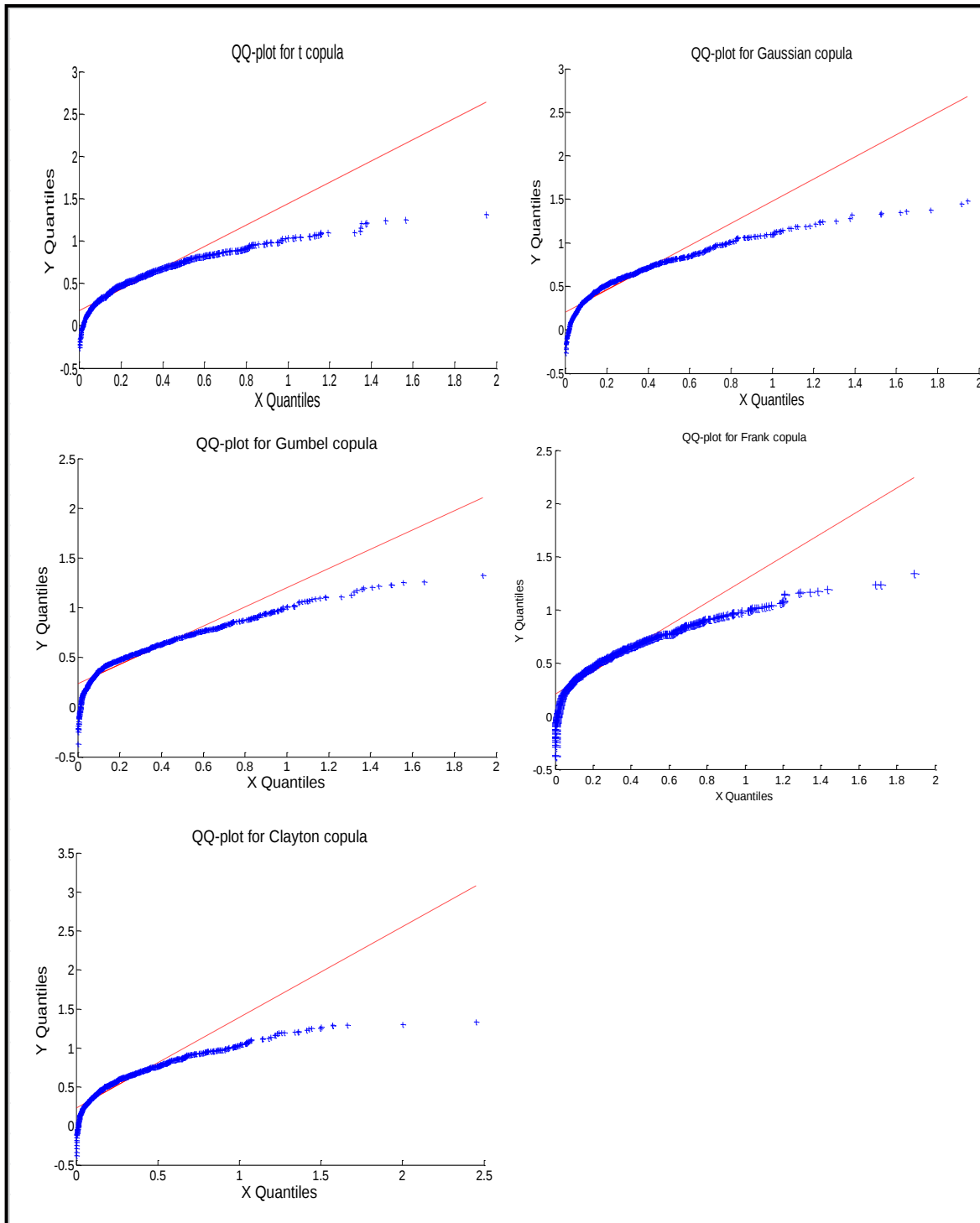


Figure 17: QQ-plot for the five copulas

Table 2: Criteria table for the choice of copula

Copulas	Ks (test statistic)	Ks (p-value)	Spearman rank coefficient	Kendall rank coefficient	Dependence hypothesis (H)
t	0.0370	0.4931	0.15	0.64	H ₀ accepted (H=0)
Gaussian	0.0420	0.3344	0.22	0.18	H ₀ accepted (H=0)
Gumbel	0.0310	0.7161	0.71	0.80	H ₀ accepted (H=0)
Frank	0.0620	0.0410	0.57	0.46	H ₀ rejected (H=1)
Clayton	0.0390	0.4253	0.14	0.71	H ₀ accepted (H=0)

4.3 Flood risk evaluation

After analyzing the floodplain, model results also indicate that the present protection system set in the Niger River at Niamey did not retain the floods throughout the sections of the city. It was observed that most of the dykes protecting the city of Niamey were overtopped by the different floods that occurred. The results also indicate that most of floods overtopped the levees reach on the upstream. However, it is noteworthy that in 1984 no dyke was endangered to be exceeded by the mean flow. For the six dykes, it is found that the probabilities of being overtopped by flows are very high (100% to 16.67 %) as shown in Table 3. For instance, the likelihood of exceeding these dikes show that the years 2010, and 2012 that experienced inundation events have flooding probabilities (66.7%, 33.3%) higher compared to other years. This is in line with different works explaining two sets of periods: the years (1984) of extreme drought in the region (no flood, no overtopping) and period with the increase in rainfall with high intensities in the region leading to flooding with high impacts (i.e. high probabilities of overtopping).

These probabilities were obtained after running the hydro-dynamical model and considering the number of dykes (among the six) that each peak flow can overtop. These probabilities of overtopping the dykes explain the extent of the danger of floods at Niamey. However, these results may be improved by using more detailed topographic data, such as LIDAR (Light Detection and Ranging), in the future (Bales et al., 2007).

Figure 18 shows an example of a floodplain generated from the Niger River HEC-RAS model based on the 2012 flows with the different cross sections. In Figure 19 one can observe some agricultural areas and houses within the floodplain after its overlapping t over a GoogleEarth image of Niamey.

Table 3: Probabilities of overtopping dykes for some specific yearly peak flows before August.

Peak flows	Probability of levees to be overtopped (%)
1981	16.67
1984	0
1998	83.33
1999	66.67
2003	83.33
2010	100
2012	100

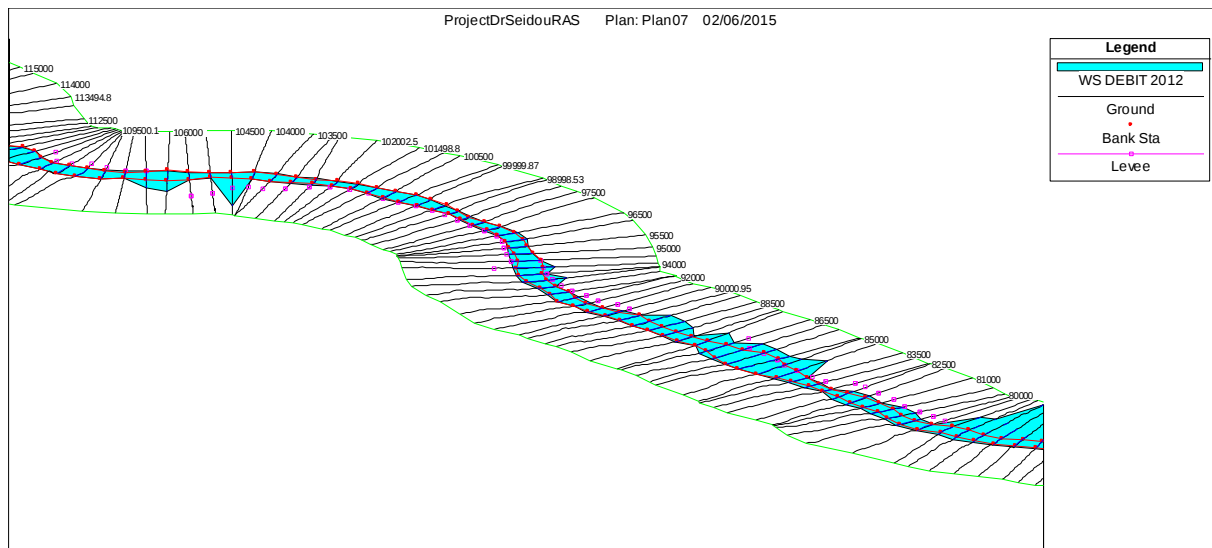


Figure 18: Simulation of levees overtopping from the hydro-dynamical model of Niger River at Niamey.



Figure 19: Impact of flooding by overlaying the floodplain on Google Earth image of Niamey

5. Conclusion

A seasonal flood forecast model was developed in this study for the city of Niamey. It involved the building up of a hydro-dynamical model of Niger River at Niamey based on HEC-RAS model and the use of a simple risk measure to calculate the probability of overtopping the flood protection dykes in Niamey. The results show that the seasonal forecast model reproduced well the rating curve over the period 2009-2014. The Gumbel copula was found to be the best among the 5 copulas after comparing them based on the Kolmogorov-Smirnov test and the QQ-plot. It showed that a dependency between streamflows at Kandadji and Garbé-Kourou based on the rank correlation coefficients of Spearman and Kendall coefficients. It is found that for the six dykes the probabilities of being overtopped by the flood are very high and varying from 100% to 16.67 %. The next steps of this work in addition to the seasonal flood forecast model coupled to copula analysis is to carry out a probability risk assessment. This could constitute a good tool for decision makers in Niamey, and also be used to tackle the flood issue faced by several West African countries.

6. References

- Antoine, J.M., Desailly, B., Galtié, J.F., Gazelle, F., Peltier, A., Valette, P. (2008). Les mots des risques naturels. *Presses universitaires du Mirail*, 127 pp.
- Awotwi, A., Kumi, M., Jansson, P.E., Yeboah, F., Nti, I.K. (2015). Predicting Hydrological Response to Climate Change in the White Volta Catchment, West Africa. *J. Earth Sci Clim Change* 6: 249. doi:10.4172/2157-7617.1000249
- Bales, J.D., Wagner, C.R., Tighe, K.C., Terziotti, S. (2007). LiDAR-derived flood-inundation maps for real-time flood-mapping applications, Tar River basin, North Carolina: U.S. *Geological Survey Scientific Investigations Report* 2007–5032, 42 p.
- Belguise, O. (2001). Tempêtes: étude des dépendances entre les branches auto et incendie avec la théorie des copulas. Mémoire DUAS troisième année. *Actuariat*, Univ. Louis Pasteur, Strasbourg, France.
- Charpentier, A. (2003). Tail distribution and dependence measures. Proceedings ASTIN Actuarial Studies in Non-life Insurance. Berlin. <http://www.actuaries.org/ASTIN/Colloquia/Berlin/Charpentier.pdf>.
- David, A., Smith, A. (2000). HEC-RAS 2.2 for backwater and Scour analysis-phase one, University of Kansas. Department of Civil and Environmental Engineering, University of Kansas Lawrence, Kansas, 88P.
- David, L.K., Mastin, M.C., Olsen, T.D. (2002). Fifty-year flood inundation maps for catacamas. Honduras, U.S. Department of the Interior U.S. Geological Survey, 9 pp.
- Dasylyva, S. (2009). Inondations à Dakar et au Sahel, Gestion durable des eaux de pluie. Etudes et Recherches, n° 267-268-269, *enda Editions*, Dakar, 259 pp.
- De Michele, C., Salvadori, G., Canossi, M., Petaccia, A. & Rosso, R. (2005). Bivariate statistical approach to spillway design flood. *J. Hydrol. Engng* ASCE 10(1), 50-57.
- Diaz Olvera, L., Plat, D., Pochet, P. (2002). Etalement urbain, situation de pauvreté et accès à la ville en Afrique subsaharienne. L'exemple de Niamey. in Bussiere Y., Madre J-L., (Eds), Démographie et transport : Villes du Nord et villes du Sud. Paris: *Ed. L'Harmattan*. pp 147-175
- Emna G., Assia C. (2008). Modélisation de la structure de dépendance hauteur-durée d'évènements pluvieux par la copule de Gumbel. *Hydrological sciences journal*, 53 :4, 802-817, DOI : 10.1623/hysj.53.4.802
- Favre, A. C., El Adlouni, S., Perreault, T. L., Thiemonge, N. & Bobée, B. (2004). Multivariate hydrological frequency analysis using copulas. *Water Resour. Res.* 40, W01101, 1-12.
- Frees, E.W., Valdez, E.A. (1998). Understanding relationships using copulas. *N. Am. Actuarial J.* 2(1), 1-25.

Gado Djibo, A., Karambiri, H., Seidou, O., Sittichok, K., Paturel, J.E., Saley, H.M. (2015). Statistical seasonal streamflow forecasting using probabilistic approach over West African Sahel. *Nat Hazards*, DOI 10.1007/s11069-015-1866-8

Gado Djibo, A., Seidou, O., Karambiri, H., Sittichok, K., Paturel, J.E., Saley, H.M. (2015). Development and assessment of non-linear and non-stationary seasonal rainfall forecast models for the Sirba watershed, West Africa. *Journal of Hydrology: Regional Studies* 4: 134–152

Genest, C., Quessy, J.F., Rémillard, B. (2006). Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scand. J. Statist.* 33(2), 337-366

Genest, C., Rivest, L. (1993). Statistical inference procedures for bivariate Archimedean copulas. *J. Am. Statist. Assoc.* 88, 1034-1043

Gilliard, P. (2005). L'extrême pauvreté au Niger mendier ou mourir. Paris: *Ed. L'Harmattan*. 280 p.

Heather J.H. (2014). Developing the Rivers of East and West Africa: An Environmental History. *Environmental History* (2014) doi: 10.1093/envhis/emu116

Institut National de la Statistique - INS, (2010). Le Niger en chiffres. République du Niger, Niamey: INS-Niger. 2 p. Institut National de la Statistique, 2002. Recensement de la population au Niger. République du Niger, Niamey : INS, Bureau Central du Recensement. 137 p.

IPCC (2007) Climate change 2007: impacts, adaptation and vulnerability. Contribution of working group II to the Fourth assessment report of the IPCC. In: Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, Hanson CE (eds) Cambridge University Press, Cambridge

Johnson, S.K., Dominique, J.O. (1988). Extracting topographic structure from digital elevation data for Geographic Information Systems analysis. *Photogrammetric Engineering and Remote Sensing*. 54(11), 1593-1600.

Klemeš, V. (2002). Risk analysis: the unbearable cleverness of bluffing. In: Risk, Reliability, Uncertainty and Robustness of Water Resources Systems (J.J. Bogardi and Z.W. Kundzewicz, eds). International Hydrology Series (UNESCO). Cambridge, Cambridge University Press.

Knebl, M.R., Yang, Z.L., Hutchison, K., Maidment, D.R. (2005). Regional Scale Flood Modeling using (NEXRAD). Rainfall GIS, and HEC-HMS/RAS: A Case study for the San Antonio River Basin Summer 2002 storm Event, *Journal of Environment Management*, 75, 325-336.

Klugman, A. S., Parsa, R. (1999). Fitting bivariate loss distribution with copulas. *Inst. Math. Econ.* 24, 139-148.

Laganier, R., Scrawell, H.J. (2004). Risque d'inondation et aménagement durable des territoires. *Presses Universitaires du Septentrion*, 226 pp.

Mara, F. (2010) Développement et analyse des critères de vulnérabilité des populations sahéliennes face à la variabilité du climat: le cas de la ressource en eau dans la vallée de la Sirba au Burkina Faso. Thèse de doctorat. Université du Québec à Montréal

Marsaglia, G., Tsang, W., Wang, J. (2003). Evaluating Kolmogorov's Distribution. *Journal of Statistical Software*. Vol. 8, Issue 18.

Motcho K.H. (2006). La réforme de la Communauté Urbaine de Niamey. Italie, Turino: *Working Paper*, no.16. 19 pp.

Najafi, M., Moradkhani, H., Wherry, S. (2011). Statistical Downscaling of Precipitation Using Machine Learning with Optimal Predictor Selection. *J. Hydrol. Eng.*, 16(8), 650–664. Olarinde, L.O.; Manyong, V.M.; and Akintola, J.O, 2010. Economic Perspectives of the Diversity of Risks among Crop Farmers in the Northern Guinea Savanna of Nigeria, Joint 3rd African Association of Agricultural Economists (AAAE) and 48th Agricultural Economists Association of South Africa (AEASA) Conference, Cape Town, South Africa, September 19-23, 2010

Napradean, I and Chira, R. (2006). The hydrological modeling of the Usturoi Valley - Using two modeling programs - WetSpa and HecRas. *Carpathian Journal of Earth And Environmental Sciences*.v1.issn (1842-4090.53-62).

Ngoran, S.D., Dogah, K.E., Xue, X.Z. (2015). Assessing the Impacts of Climate Change on Water Resources: The Sub-Saharan Africa Perspective. *Journal of Economics and Sustainable Development*. (6). 185-193

O'Connor, K.M. (2006). River flow forecasting. In: River Basin Modelling for Flood Risk Mitigation (D.W. Knight and A.Y. Shamseldin, eds). London, *Taylor and Francis Group*

Perrin, C., Michel, C., Andréassian, V. (2001). Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *J. Hydrol.*, 242:275-301.

Sanda G.H. (2010). Cartographie de la dynamique de l'occupation des sols et de l'érosion dans la ville de Niamey et sa périphérie. Mémoire Maîtrise en géographie, *Université Abdou Moumouni de Niamey*. 152pp

Scheraga, J.D., Grambsch, A.E. (1998). Risks, opportunities, and adaptation to climate change. *Climate Research*, 10, 85-95.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist.* 8, 229-231. Université Paris VIII, France.

Solaiman, T.A., (2011). Uncertainty Estimation of Extreme Precipitations under Climatic Change: A Non-Parametric Approach. *PhD Thesis*, Department of Civil and Environmental Engineering, The University of Western Ontario

Soleymani M., Hossein S., Amir K., Fereydoon K., Hossein B. (2014). Determining Flood Zones Using HEC-RAS Model (Case study: Gale Hassan River situated in Atrak watershed). *Bull. Env. Pharmacol. Life Sci.*, 3 (7), 102-107

Sighomnou, D., Descroix, L., Genthon, P., Mahé, G., Moussa, I. B., Gautier, E., Coulibaly, B. (2013). La crue de 2012 à Niamey: un paroxysme du paradoxe du Sahel? *Science et Changements planétaires/Sécheresse*, 24(1), 3-13.

Tinni B. A. (2011). Mobilités résidentielles et habitats à Niamey. Mémoire de maitrise. *Université Abdou Moumouni*, Niamey Niger. 171pp

United Nations Development Programme (UNDP). (2010). Human development report 2010. The real wealth of nations: Pathways to human development. 20th anniversary edition. New York: Palgrave Macmillan for the United Nations Development Programme (UNDP). <http://hdr.undp.org/en/reports/global/hdr2010/>. Accessed 10 June 2015.

United Nations Office for the Coordination of Humanitarian Affairs – OCHA (2009). Afrique de l'Ouest: Inondations 2009. Rapport de situation # 3, 3 pp

United Nations Office for the Coordination of Humanitarian Affairs – OCHA (2013). Le processus de planification humanitaire. *Appel Global Niger 2013*. 108pp.

US Army Corps of Engineers-USACE, (2009). HEC-GeoRAS, GIS Tools for Support of HECRAS using ArcGIS. User's Manual, Version 4.2, U.S. Army Corps of Engineers, Davis, California, 246 p.

Venter, G. (2002). Tails of copulas. *Proc. Casualty Actuarial Society*. 89, 68-113.
Venter, G. (2003). Quantifying correlated reinsurance exposures with copulas. *Casualty Actuarial Society Forum, Spring*, 215-229.

Wilks, D. (2006). Statistical Methods in the Atmospheric Sciences. *Academic Press*; 3rd edition. New York, U.S.A. 704pp

World Meteorological Organization - WMO (2011). Manual on Flood Forecasting and Warning. *Chairperson, Publications Board*. 142 pp

Zhang, L. & Singh, V. P. (2006). Bivariate flood frequency analysis using the copula method. *J. Hydrol. Engng* 11(2), 150-164.

7.5. Synthesis and partial conclusion

A seasonal flood forecasting model was developed using the probabilistic forecasted streamflows obtained from the seasonal streamflow forecasting model.

Eleven probability distributions were tested for streamflows at Kandadji and Garbé-Kourou, based on the Ks-test. For streamflows at Garbé-Kourou, the best fit was obtained with the generalized Pareto (GPAR) distribution, while for streamflows at Kandadji, the Log-Pearson III distribution (LP-III) was found to be the best. These results helped to determine the best type of distribution to use for the copula analysis.

The copula approach was then used to model the dependence of the streamflows at Garbé-Kourou and Kandadji, regardless of the effect of the behavior of distribution margins. This method is a powerful and flexible tool, as it allowed finding that the Gumbel copula was the best among the 5 copulas tested, according to its lowest statistic value (0.031) and highest p-value (0.716) from the Kolmogorov-Smirnov test. Moreover, the Gumbel copula gave the best fit compared to the remaining copulas, based on the quantile-quantile plots after resampling.

The hydraulic HEC-RAS model was used to develop the hydrodynamical model of the Niger River at Niamey. This model was calibrated in such a way that it produced well the last rating curve over the period 2009-2014. This result was achieved due to the use of observed elevation data for Niamey and bathymetry data for the Niger River. The model simulations showed that most of the levees protecting the city of Niamey were overtopped by the different floods that occurred, especially the same historical years during which floods has occurred.

Moreover, the assessment of submersion probabilities showed that for the six dykes, the probabilities of being overtopped by the flood are very high and varying from 100% to 16.67 %.

Therefore, it appears that coupling a seasonal flood forecasting model to a copula-based approach can help in carrying out risk assessments to assess the probabilities of overtopping protection dykes. This can help stakeholders, national authorities, and the population to take protective measures before flooding occurs.

Chapter 8. Conclusions and Perspectives

This chapter summarizes the problem and the contributions, along with their limitations. It also refers to the previous chapters where details are provided. The future works are identified in terms of perspectives resulting from this thesis work. They are presented in order to make more comprehensive the field of seasonal forecasting.

8.1. Conclusion

The development and assessment of seasonal forecasting models will help local people and rural populations who mainly rely on agriculture for their survival by changing their farming systems through a good management system. This thesis research has endeavored to identify a seasonal rainfall forecasting system which can improve forecasting abilities using statistical approaches. The developed models were tested on the Sirba watershed.

This forecasting system was first based on the development of a model to find the best predictors and optimal lag times which are *sine qua non* conditions for obtaining good seasonal forecasting abilities. Eighty-four predictors having physical insight on the West African monsoon were considered, based on the AMMA research program. Predictors were first reduced by screening through linear correlation with satellite rainfall over the West African region. The retained predictors were then screened through a series of steps in order to isolate those having the highest predictive power. The series of steps included the use of: (i) correlation analysis to screen the predictors' grid points due to their huge number. In this step, only predictors having a P-value of less than 0.05 were retained for further processing. (ii) Then, principal component analysis was applied to reduce the predictor vector dimension before using stepwise regression to keep only the principal component having high predictive power. (iii) In the last step, linear regression was used to get synthetic forecasts. The model's performance was assessed using the coefficient of determination, the Nash-Sutcliffe coefficient, and hit rate score. At the end of the process,

three predictors, air temperature (from Pacific Tropical North), sea level pressure (from Atlantic Tropical South), and relative humidity (from Mediterranean East) were retained and tested as inputs for seasonal rainfall forecasting models. Then, seasonal rainfall forecasting models, with either changing parameters or constant parameters, were developed and tested using each of the three predictors.

The forecasting models with changing parameters were two probabilistic methods based on change point detection that allow the relationship (rainfall/predictor) to change according to time or rainfall magnitude: model M1 allows for change in model parameters according to annual rainfall magnitude, while model M2 allows for changes in model parameters with time. The forecasting models with constant parameters were those that often arbitrarily assume that rainfall is linked to predictors by a multiple linear regression with parameters: model M3 is the classical linear model with constant parameters, and model M4 is based on the climatology. A Bayesian multiple change point detection algorithm was developed and used with a Bayesian model selection approach. A normalized Bayes factors and graphs of the likelihood of forecasted rainfall under each model, were compared. It was found that the best seasonal rainfall forecasting model uses air temperature as a predictor and allows parameter changes according to rainfall magnitude. Thus, seasonal forecasting models with changing parameters could be the best for use in the Sirba watershed. Indeed, changes in the predictand-predictor relationship according to rainfall amplitude, combined with the Bayesian model selection procedure, appear to be a good technique for forecasting seasonal rainfall in the Sahel.

Non-linear models including regression trees, feedforward neural network and non-linear principal component analysis were implemented and tested to forecast seasonal rainfall using the predictors. Forecast performances were compared using the coefficient of determination, the Nash-Sutcliffe coefficient, and hit rate score. The results showed that non-linear principal component analysis was the best non-linear method, while the feed-forward neural network and the regression tree models performed poorly. The limits of these models may result from the short length of data used, because such models need long lengths of data to ensure their training.

Many factors such as climate change and variability impacts induced changes in streamflow, which directly affect water availability for all activities, particularly

agriculture, which is the main survival means of the West African population. Hence, streamflow information a few months prior to a rainy season with an acceptable forecasting ability would immensely benefit water users in performing operational planning for water management decisions. This research attempts to develop seasonal streamflow forecasting models that can help such populations change their cropping systems to improve their agricultural yields. Two sets of models for seasonal streamflow forecasting were developed and tested using the three predictors. These models have both changing and constant parameters, as previously described. The seasonal rainfall probability forecasts obtained from such models were used in addition to the developed Sirba SWAT hydrological model to estimate the seasonal streamflow. Using a Bayesian model selection based on normalized Bayes factors combined with graphs of forecast likelihoods, the models' performances were assessed. Analysis of the results showed that model M3 using RHUM as a predictor is the best seasonal annual mean streamflow forecast model, followed by model M4, while model M1 using AirTemp is deemed the best to forecast maximum monthly streamflows. From this, it can be concluded that that models with changing parameters and models with constant parameters once associated with the Bayesian model selection approach could be a good way to forecast seasonal annual mean (respectively maximum monthly) streamflows in the Sahel region.

Long periods of aridity with sudden flood events usually prove to be disastrous for the local communities in the West African Sahel region. This thesis work has developed a seasonal flood forecasting model with a copula-based approach in order to assess the probabilities of flood protection structures in Niamey being overtopped. This work was done in three steps: copula fitting, seasonal flood model development, and probabilistic assessment of flood protection dykes being overtopped.

Copula methods were used to model the dependence of streamflows at Garbé-Kourou and Kandadji, regardless of the effect of the behavior of distribution margins. For the distributions, eleven probability functions were tested over streamflows at Kandadji and Garbé-Kourou. The best fit was obtained with the generalized Pareto (GPAR) distribution for streamflows at Garbé-Kourou, while the Log-Pearson III distribution (LP-III) was found the best for streamflows at Kandadji. The test of the five copulas proved that the Gumbel copula was the best-fitting one due to its lowest statistical value and

highest p-value from the Kolmogorov-Smirnov test, and also based on the QQ-plot. The seasonal flood forecasting model was established using the HEC-RAS model which helped to develop the hydrodynamical model of Niger River at Niamey in addition to bathymetry data and the probabilistic streamflow forecasts obtained from the seasonal streamflow forecasting model. The calibrated seasonal flood forecast model reproduced almost perfectly the rating curve over the period 2009-2014. After developing the seasonal flood forecasting model, the six dykes protecting the city of Niamey were found to have high probabilities of being overtopped.

In conclusion, this work has shown the importance of developing a good method for selecting predictors of Sahelian rainfall with optimal lag times rather than assigning them subjectively as has been done in many studies in the region. Moreover, the results have pointed out that other potential predictors such as air temperature, sea level pressure, and relative humidity could be used instead of just using the sea surface temperatures, as has been traditional. The study has helped in understanding that for seasonal forecasting it is advantageous to use many criteria to estimate a model's performance, because the use of one estimator could cause biased estimates, as it may be sensitive to outliers. Therefore, any seasonal forecasting model for the West African Sahel should consider the local climate of the region in order to make accurate forecasts.

8.2. Contributions

This thesis has contributed to the understanding of seasonal rainfall and streamflow forecasting in the Sahel. This was done through the development of new seasonal forecasting systems using models with changing and constant parameters: first new seasonal forecasting models considering the relationship between rainfall and predictor to change according to rainfall magnitude or time.

One interesting finding in this work is that different characteristics of streamflow hydrographs were best forecasted by different models, suggesting that a multi-model approach to forecasting may be a rewarding solution to climate forecasting in the Sahel. The study also highlighted that there were strong non-linearities between climate-based predictors and streamflow in the study area. Therefore, the current practice in the region,

of directly linking predictors such as sea surface temperature to streamflow, is far from optimal.

A large contribution of this work is the quantitative prediction of seasonal rainfalls and their probabilities. This kind of forecasting allows policy-makers and local populations to better plan their activities and also to take appropriate measures ahead of time in order to mitigate impacts. The quantitative probabilistic forecasting in this work also helps to better manage the risks recurrent in the West African region.

Another contribution of this research is that it helps in understanding that non-linear models can also be used instead of the usual linear methods. Thus, the detection of the limits related to non-linear models will help in improving their forecasting abilities in the Sahel area.

A major contribution is the development of a seasonal rainfall forecasting model combines with a seasonal streamflow forecasting model and a seasonal flood forecasting model. This will serve policy-makers and the West African population to better adapt to the recurrent floods which affect most cities in the region.

The novelty of this work is in the use of other predictors rather than the usual SSTs which gave acceptable results than SSTs. The new predictors identified in this study could lead to better forecasting of seasonal rainfall and streamflow over West Africa, an issue which has challenged forecasters for many years.

8.3. Future work

This work has developed and tested these models on the Sirba watershed. Therefore, from this perspective, there is a need to continue developing and testing these seasonal forecasting models over the large basins of West Africa, and even the entire African continent.

Further to this, the developed Sirba SWAT hydrological model should be carried out for other neighboring basins.

From the seasonal streamflow forecasting results, a multi-model approach to forecasting is an interesting solution to climate forecasting in the Sahel. Thus, this work will continue to test the multi-model approach in order to confirm its abilities.

Regarding the obtained results from non-linear models, there is need to continue their development by adopting more specific training techniques and gathering more data in order to achieve better forecasts.

Another step is to generalize the models for other works on the Sahel by constructing a Sahelian global index using CRU data and examining its correlation with the index of the watershed and the abilities of various models.

As a future step in this work, a multi-model approach will be used to compare the resulting abilities to those of the best seasonal rainfall model.

The last perspective of this work is to carry out a probabilistic risk assessment to quantify the damages associated to the seasonal flood forecasts in Niamey.

References

- Abbaspour, K.C. (2012). SWAT-CUP 2012: SWAT calibration and uncertainty 7 programs. Eawag: Swiss Federal Institute of Aquatic Science and 8 Technology. Retrieved from <http://www.neprashtechology.ca>.
- Abraham, A., Steinberg, D., Philip, S.N. (2001). Rainfall forecasting using soft computing model sand multivariate adaptive regression splines, available at: <http://meghnad.iucaa.ernet.in/~nspp/ieee.pdf>.
- Abrahart, R.J., See, L. (2000). Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecast in two contrasting catchments, *Hydrol. Process*, 14, 2157-2172
- Abudu, S., King, J.P., Pagano, T.C. (2010). Application of partial least-square regression in seasonal streamflow forecasting. *Journal of Hydrologic Engineering*, 15, 612-623.
- ACMAD, CLIPS (1998). Prévision climatique en Afrique, World Meteorological Organization, WMO/TD, 927, 210 pp.
- Ali, A., Amani, A., Lebel, T. (2004). Estimation des pluies au Sahel: utilisation d'un modèle d'erreur pour évaluer les réseaux sol et produits satellitaires, *Sécheresse*, 3, 15, p. 271-278.
- Amadou, A., Gado Djibo, A., Seidou, O., Sittichok, K., Seidou Sanda, I. (2014). Changes to flow regime on the Niger River at Koulikoro under a changing climate. *Journal of hydrological sciences*. DOI: 10.1080/02626667.2014.916407
- Andersen, I., Dione, O., Jarosewich-Holder, M., Olivry, J.C. (2005). The Niger River Basin: a vision for sustainable management. In: Golitzen, K.G. (Ed.), *Directions in Development*. The World Bank, Washington, DC, USA, p. 145.
- Ansari, H. (2013). Forecasting Seasonal and Annual Rainfall Based on Nonlinear Modeling with Gamma Test in North of Iran. *International Journal of Engineering Practical Research*, Vol. 2 Issue 1, 16-29.
- Archer, D.R., Fowler, H.J. (2008). Using meteorological data to forecast seasonal runoff on the river Jhelum, Paksitan. *Journal of Hydrology*, 361, 10-23
- Ardoïn-Bardin, S. (2004). Variabilité hydroclimatique et impacts sur les ressources en eau de grands bassins hydrographiques en zone soudano-sahélienne. Thèse de doctorat, Université Montpellier II, France.
- Ardoïn, S., Lubès-Niel, E., Servat, E., Dezetter, A., Boyer, J.F. (2003). Analyse de la persistance de la sécheresse en Afrique de l'ouest : caractérisation de la situation de la décennie 1990. In : *Hydrology of Mediterranean and Semiarid Regions* (ed. E. Servat, W. Najem, C. Leduc, A.Shakeel), 223-228. *IAHS Pub.* 278

- Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., Harmel, R.D., Griensven, A.V., Van Liew, M.W., Kannan, N., Jha, M.K. (2012). SWAT: model use, calibration, and validation. *American Society of Agriculture and Biological Engineers*, 55, 1491-1508.
- ASCE, (2000a). Task committee on Application of Artificial Neural Networks in Hydrology, Part I, *J. Hydrol. Eng.*, 5(2), 115-123
- ASCE, (2000b). Task committee on Application of Artificial Neural Networks in Hydrology, Part 2, *J. Hydrol. Eng.*, 5(2), 124-137
- Asselman, N., Bates, P., Woodhead, S., Fewtrell, T., Soares-Frazão, S., Zech, Y., Velickovic, M., De Wit, A., Ter Maat, J., Verhoeven, G., Lhomme, J. (2009). Flood inundation modelling - model choice and proper application. *Floodsite report T08-09-03*, 142 pp.
- Awotwi, A., Kumi, M., Jansson, P.E., Yeboah, F., Nti, I.K. (2015). Predicting Hydrological Response to Climate Change in the White Volta Catchment, West Africa. *J. Earth Sci Clim Change* 6: 249. doi:10.4172/2157-7617.1000249
- Bader, J., Latif, M. (2003). The impact of decadal-scale Indian Ocean sea surface temperature anomalies on Sahelian rainfall and the North Atlantic Oscillation. *Geophysical Research Letters*, Vol. 30, Issue 22.
- Bader, J.C., Piedelievre, J.P., Lamagat, J.P. (2006). Prévision saisonnière du volume de crue du Fleuve Sénégal: utilisation des résultats du modèle ARPEGE Climat / Seasonal forecasting of the flood volume of the Senegal River, based on results of the ARPEGE Climate model, *Hydrological Sciences Journal*, 51: 3, 406-417
- Badr, H.S., Zaitchik, B.F., Guikema, S.D. (2014). Application of Statistical Models to the Prediction of Seasonal Rainfall Anomalies over the Sahel. *J. Appl. Meteor. Climatol.*, 53, 614-636. DOI: <http://dx.doi.org/10.1175/JAMC-D-13-0181.1>
- Ball, J.E., Luk, K.C. (1998). Modeling spatial variability of rainfall over a catchment. *J. Hydrol. Eng.* 3, 122–130.
- Barnston, A.G. (1994). Linear Statistical Short-Term Climate Predictive Skill in the Northern Hemisphere, *Journal of Climate*, 7, 1513-1564.
- Barnston, A.G., Thiao, W., Kumar, V. (1996). Long-lead forecasts of seasonal precipitation in Africa using CCA, *Weather and Forecasting*, 11, 506-520.
- Barnston, A.G., Van Den Dool, H.M. (1992). A degeneracy in cross-validated skill in regression-based forecasts. *Journal of Climate*: 6, 963-977
- Barry, D., Hartigan, J. A. (1992). Product partition models for change point models, *Ann. Stat.*, 20, 260-279.
- Barry, D., Hartigan, J.A. (1993). A Bayesian analysis for change point problems, *J. Am. Stat. Assoc.*, 88, 309-319.
- Batté, L., M. Déqué, (2011). Seasonal predictions of precipitation over Africa using coupled ocean-atmosphere general circulation models: Skill of the ENSEMBLES

- project multimodel ensemble forecasts. *Tellus*, 63A, 283-299, doi:10.1111/j.1600-0870.2010.00493.x.
- Beaulieu, C., Ouarda, T.B.M.J., Seidou, O. (2005). Comparative study of homogenization techniques for precipitation data series (in French), *Progress report* no. 3 (Project on the homogenization of precipitation data), Ouranos Consortium, Montreal.
- Beaulieu, C., Seidou, O., Ouarda, T.M.B.J., Zang, X. (2009). Intercomparison of homogenization techniques for precipitation data continued: comparison of two recent Bayesian change point models. *Water Resources Research* Vol. 45. Doi: 10.1029/2008WR007501
- Belguise, O. (2001). Tempêtes: étude des dépendances entre les branches auto et incendie avec la théorie des copulas. Mémoire DUAS troisième année. Actuariat, Univ. Louis Pasteur, Strasbourg, France.
- Begert, M., Thomas, S., Walter, K. (2005). Homogenous temperature and precipitation series of switzerland from 1864 to 2000. *International Journal of Climatology* 25 (1), 65-80
- Bishop, C.M. (1997). Neural Networks for Pattern Recognition, third edition, *Clarendon Press*, Oxford. 145pp.
- Bouali, L. (2009). Prévisibilité et prévision statistico-dynamique des saisons des pluies associées à la mousson ouest africaine à partir d'ensembles multi-modèles, *thèse de doctorat*, Université de Bourgogne, France. 159 pp
- Bouali, L., Philippon, N., Fontaine, B., Lemond, J. (2008). Performance of DEMETER calibration for rainfall forecasting purposes: Application to the July–August Sahelian rainfall, *Journal of Geophysical Research*, 113, D15111, doi:10.1029/2007JD009403.
- Bowman, A.W., Pope, A., Ismail, B. (2004). Detecting discontinuities in nonparametric regression curves and surfaces, Research Report, Department of Statistics, University of Glasgow. [<http://www.stats.gla.ac.uk/%7Eadrian/research-reports/discontinuity-paper.ps>]
- Biasutti, M., Held, I.M., Sobel, A.H., Giannini, A. (2008). SST forcings and Sahel rainfall variability in simulations of the twentieth and twenty-first centuries. *Journal of Climate*, 21, 3471-3486
- Brankovic, C., Palmer, T.N. (1997). Atmospheric seasonal predictability and estimates of ensemble size, *Monthly Weather Review*, 125, 859-874.
- Breiman, L. (1996). Bagging predictors. *Machine Learning* 24, 123-140
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984). Classification and regression tree. *Wadsworth*, 358pp.
- British Atmospheric Data Centre (BADC). High-Resolution Gridded Datasets (and Derived Products). Available Online: <http://www.cru.uea.ac.uk/cru/data/hrg/> (accessed on 9 September 2015).

- Brooks, N. (2004). Drought in the African Sahel: long term perspectives and future prospects. Tyndall Centre Working Paper, vol 61. *Tyndall Centre for Climate Change Research, School of Environmental Sciences*, University of East Anglia, Norwich.
- Camberlin, P., Janicot, S., Pocard, I. (2001). Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature: Atlantic VS. ENSO. *Int. J. Clim.*, 21, 973–1005.
- Caminade, C. (2006). Rôle de l'océan et influence des émissions d'origine anthropique sur la variabilité climatique en Afrique, thèse de doctorat, *Université de Toulouse III*.
- Caminade, C., Terray, L., (2010). Twentieth century Sahel rainfall variability as simulated by the ARPEGE AGCM, and future changes. *Climate Dynamics*, 35, 75-94.
- Campolo, M., Soldati, A. (1999). Forecasting river flow rate during low-flow periods using neural networks, *Water Resour. Res.*, 35 (11), 3547-3552
- Canon, A.J. (2000). Forecasting Indian rainfall using regional circulation fields as predictors: an ensemble neural network approach. Master of Science thesis, Vancouver, 87pp
- Canon, A.J., Mckendry, I. (1999). Forecasting all-India summer monsoon rainfall using regional circulation principal components: a comparison between neural network and multiple regression models. *Int. J. Clim.*: 19, 1561-1578
- Casati, B., Wilson, L.J., Stephenson, D.B. (2008). Forecast verification: Current status and future directions. *Meteorol. Appl.*, 15, 3-18.
- CEDEAO and CSAO/OCDE (2008). Le climat et les changements climatiques. Atlas de l'Intégration Sous-régionale en Afrique de l'Ouest, *série environnement*. 24 p.[En ligne]. http://www.fao.org/nr/clim/docs/clim_080502_fr.pdf (Page consultée le 19 juillet 2014)
- Centre Régional AGRHYMET. (2004). Système TBASE de données numériques. Niamey, Niger.
- Charpentier, A. (2003). Tail distribution and dependence measures. Proceedings ASTIN Actuarial Studies in Non-life Insurance. Berlin. <http://www.actuaries.org/ASTIN/Colloquia/Berlin/Charpentier.pdf>.
- Chase, T.N., Pielke S.R., Avissar, R. Teleconnections in the Earth System. Encyclopedia of Hydrological Sciences. Available online: <http://onlinelibrary.wiley.com/doi/10.1002/0470848944.hsa190/pdf> (accessed on 12 August 2014).
- Chattopadhyay, S., Chattopadhyay, G. (2007). Identification of the best hidden layer size for three-layered neural net in predicting monsoon rainfall in India. *Journal of Hydroinformatics* 10 (2), 181-188.
- Christensen, J.H. et al (2007). Regional climate projections, in climate change 2007: the physical science basis. In: Solomon S. et al. (eds) Contribution of working group I

- to the fourth assessment report of the intergovernmental panel on climate change, pp 847-940. *Cambridge Univ. Press*, New York
- Clarke, A.J., Lebedev, A. (1996). Long term changes in the equatorial Pacific trade winds. *Journal of climate*, 9, 1020-1029.
- Coiffier, J. (2000). Un demi-siècle de prévision numérique du temps, *La Météorologie*, 8, 30.
- Cole, M.M. (1982). The influence of soils, geomorphology and geology on the distribution of plant communities in savanna ecosystems, in "Ecology of tropical savannas", Huntley and Walker, 145-174
- Coulibaly, P., Hache, M., Fortin, V., Bobee, B. (2005). Improving daily reservoir inflow forecasts with model combination. *Journal of Hydrologic Engineering* 10 (2), 91-99.
- Cook, K.H. (1999). Generation of the African easterly jet and its role in determining West African precipitation. *J. Climate*, 12, 1165-1184.
- Daly, C., Gibson, W.P., Taylor, G.H., Doggett, M.K., Smith, J.I. (2007). Observer bias in daily precipitation measurements at United States cooperative network stations. *Bulletin of the American Meteorological Society* 88 (6), 899-912.
- Daumer, M., Falk, M. (1998). On-line change-point detection for state space models using multi-process Kalman filters, *Linear Algebra Appl.*, 284, 125-135.
- David, L.K., Mastin, M.C., Olsen, T.D. (2002). Fifty-year flood inundation maps for catacamas. Honduras, U.S. Department of the Interior U.S. Geological Survey, 9 pp.
- David, A., Smith, A. (2000). HEC-RAS 2.2 for backwater and Scour analysis-phase one, University of Kansas. Department of Civil and Environmental Engineering, University of Kansas Lawrence, Kansas, 88P.
- Dembele, Y., Somé, L. (1991). Propriétés hydrodynamiques des principaux types de sol du Burkina Faso, (Proceedings of the Niamey Workshop, February 1991) *IAHS Publ.*199 (357), 217-227.
- DeMers, D., Cottrell, G.W. (1993). Nonlinear dimensionality reduction. In Hanson, D., Cowan, J. and Giles, L., (eds.) *Advances in Neural Information Processing Systems* 5. Morgan Kaufmann, San Mateo, CA, pp. 580-587
- DeMichele, C., Salvadori, G., Canossi, M., Petaccia, A. & Rosso, R. (2005). Bivariate statistical approach to spillway design flood. *J. Hydrol. Engng ASCE* 10(1), 50-57.
- Devineau, L., Serpantié, G. (1991). Paysages végétaux et systèmes agraires au Burkina Faso. Caractérisation et suivi des milieux terrestres en région arides et tropicales. Deuxièmes journées de télédétection. *Colloques et Séminaires. Bondy ORSTOM* 373-385 p.
- Descroix, L., Mahe, G., Lebel, T., Favreau, G., Galle, S., Gautier, E., Olivry, J.-C., Albergel, J., Amogu, O., Cappelaere, B., Dessouassi, R., Diedhiou, A., Breton, E.L., Mamadou, I., Sighomnou, D. (2009). Spatio-temporal variability

- of hydrological regimes around the boundaries between Sahelian and Sudanian areas of West Africa: a synthesis. *J. Hydrol.* 375, 90–102.
- Diamantaras, K., Kung, S. (1996). *Principal Component Neural Networks*. Wiley, New York.
- Diello, P. (2007). *Interrelations Climat - Homme - Environnement dans le Sahel Burkinabé : impacts sur les états de surface et la modélisation hydrologique*. Thèse de doctorat, Université de Montpellier II, Sciences et techniques du Languedoc.
- Doblas-Reyes, F.J., Déqué, M., Piedelièvre, P. (2000). Multi-model spread and probabilistic seasonal forecasts in PROVOST, *Quarterly Journal of Royal Meteorological Society*, 126, 2069–2088.
- Doll, P., Kaspar, F., Lehner, B. (2003). A global hydrological model for deriving water availability indicators: model tuning and validation. *J. Hydrol.* 270, 105–134.
- Douville, H., Chauvin, F., Broqua, H. (2001). Influence of soil moisture on the Asian and African monsoons. Part I: Mean monsoon and daily precipitation. *J. Clim.*, 14, 2381–2403.
- Douville, H. (2002). Influence of soil moisture on the Asian and African monsoons. Part II: Interannual variability. *J. Clim.*, 15, 701–720.
- Douville, H. (2003). Assessing the influence of soil moisture on seasonal climate variability with AGCMs. *J. hydrometeorol.*, 4, 1044–1066.
- Duan, Q., Sorooshian, S., Gupta, V.K. (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology*, 158, 265–284.
- Druyan, L.M. (1991). The sensitivity of sub-Saharan precipitation to Atlantic SST, *Climate Change*, 18, 17–36.
- Easterling, D.R., Peterson, T.C. (1992). Techniques for detecting and adjusting for artificial discontinuities in climatological time series: A review, *Proc. of the 5th International Meeting on Statistical Climatology*, 22–26 June 1996, Toronto, Ontario, Canada.
- Easterling, D.R., Peterson, T.C., Karl, T.R. (1996). On the development and use of homogenized climate datasets. *Journal of Climate* 9 (6), 1429–1434.
- El-Shafie, A.H., El-Shafie, A., El-Magzoghi, H.G., Shehata, A., Taha, M.R. (2011). Artificial neural network technique for rainfall forecasting applied to Alexandria, Egypt. *International Journal of the physical sciences*, 6, 1306–1316
- Elsner, J., Schmertmann, C. (1994). Assessing Forecast Skill through Cross Validation, *Weather. Forecasting*, 9, 619–624.
- Eltahir, E.A.B. (1996). Role of vegetation in sustaining large-scale atmospheric circulations in the tropics. *Journal of Geophysical Research*, 101 (D2), 4255–4268.
- Emanuel, K.A. (1995). On thermally direct circulations in moist atmospheres. *Journal of Atmospheric Sciences*, 52, 1529–1534.

- FAO/UNESCO (2003). Digital Soil Map of the World and Derived Soil Properties. Rev.1. (CD Rom), 2003. Available from http://www.fao.org/catalog/what_new-e.htm
- Favre, A. C., El Adlouni, S., Perreault, T. L., Thiemonge, N. & Bobée, B. (2004). Multivariate hydrological frequency analysis using copulas. *Water Resour. Res.* 40, W01101, 1-12.
- Fearnhead, P. (2005). Exact Bayesian curve fitting and signal segmentation, *IEEE Trans. Signal Process.*, 53(6), 2160-2166.
- Fearnhead, P. (2006). Exact and efficient Bayesian inference for multiple changepoint problems, *Stat. Comput.*, 16, 203-213.
- Folland, C.K., Owen, J.A., Ward, M.N., Colman, A.W. (1991). Prediction of seasonal rainfall in the Sahel region of Africa using empirical and dynamical methods, *Journal of Forecasting*, 10, 21-56.
- Folland, C.K., Palmer, T.N., Pather, D.E. (1986). Sahel rainfall and worldwide sea temperature, *Nature*, 320, 602-607.
- Fontaine, B., Bigot, S. (1993). West African rainfall deficits and sea surface temperatures. *International Journal of Climatology* 13(3), 271-285
- Fontaine, B., Janicot, S. (1996). Sea surface temperature fields associated with West African rainfall anomaly types. *J. Clim.*, 9, 2935-2940.
- Fontaine, B., Philippon, N., Camberlin, P. (1999). An improvement of June-September rainfall forecasting in the Sahel based upon region April-May moist static energy content (1968-1997), *Geophysical Research Letters*, 26, 2041-2044.
- Fontaine, B., Roucou, P., Monerie, P.A. (2011). Changes in the African monsoon region at medium-term time horizon using 12 AR4 coupled models under A1b emissions scenario. *Atmosph. Sci. Let.*, DOI:10.1002/asl321
- Frees, E.W, Valdez, E.A. (1998). Understanding relationships using copulas. *N. Am. Actuarial J.* 2(1), 1-25.
- French, M.N., Krajewski, W.F., Cuykendall, R.R. (1992). Rainfall forecasting in space and time using neural network, *J. Hydrol.*, 137, 1-31.
- Fuller, D.O., Ottke, C. (2002). Land cover, rainfall and land-surface albedo in West Africa. *Climatic change*. 54, 1816204
- Gado Djibo, A., Karambiri, H., Seidou, O., Sittichok, K., Paturel, J.E., Saley, H.M. (2015c). Statistical seasonal streamflow forecasting using probabilistic approach over West African Sahel. *Nat Hazards*, DOI 10.1007/s11069-015-1866-8
- Gado Djibo, A., Karambiri, H., Seidou, O., Sittichok, K., Philippon, N., Paturel, J.E., Saley, H.M. (2015a). Linear and non-linear approaches for statistical seasonal rainfall forecast in the Sirba watershed region (SAHEL). *Climate*, 3, 727-752; doi:10.3390/cli3030727
- Gado Djibo, A., Seidou, O., Karambiri, H., Sittichok, K., Paturel, J.E., Saley, H.M. (2015b). Development and assessment of non-linear and non-stationary seasonal

- rainfall forecast models for the Sirba watershed, West Africa. *Journal of Hydrology: Regional Studies* (4) 134–152
- Gado Djibo, A., Seidou O., Moussa Saley, H., Karambiri, H., Philippon, N., Sittichok, K., Paturel, J.E. (2016). A copula-based approach for assessing flood protection overtopping associated with a seasonal flood forecast in Niamey, West Africa. *Journal of Earth Sciences* (peer reviewed).
- Gaetani, M., Fontaine, B. (2013). Interaction between the West African Monsoon and the summer Mediterranean climate: An overview. *Fisica de la Tierra*, 25, 41–55.
- Gaetani, M., Fontaine, B., Roucou, P., Baldi, M. (2010). Influence of the Mediterranean Sea on the West African monsoon: intraseasonal variability in numerical simulations. *J. Geophys. Res.*, 115, doi: 10.1029/2010JD014436.
- Gaetani, M., Mohino, E. (2013). Decadal prediction of the Sahelian precipitation in CMIP5 simulations. *J. Clim.* 26, 7708–7719.
- Garcia-Serrano, J., Doblas-Reyes, F.J., Haarsma, R.J., Polo, I. (2013). Decadal prediction of the dominant West African monsoon rainfall modes. *J. Geophys. Res. Atmos.* 118, 5260–5279.
- Garric, G., Douville, H., Déqué, M. (2002): Prospects for improved seasonal predictions of monsoon precipitation over West-Africa, *International Journal of Climatology*, 22, 331-345.
- Gaye, A. (2002). Caractéristiques dynamiques et pluviosité des lignes de grains en Afrique de l'ouest. PhD thesis, *Université Cheikh Anta Diop*, Dakar, Sénégal.
- Gaye, A. (2009). Consultation sous-régionale pour la préparation aux inondations en Afrique de l'Ouest. [En ligne] http://www.sununews.com/index.php?view=article&catid=45%3Aactualite&id=119%3Aconsultation-regionale-pour-la-preparation-aux-inondations-en-afrique-de-louest-pour-lannee-2009&format=pdf&option=com_content&Itemid=50 (consulted 5 may 2015)
- Gelfand, A.E., Hills, S.E., Racine-Poon, A., Smith, A.F.M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling, *J. Am. Stat. Assoc.*, 85, 972-985.
- George, C., Leon, L. F. (2007). WaterBase: SWAT in an open source GIS. The Open Hydrology Journal, 1, 19-24. *Bentham Science Publisher Ltd.* 1874-3781/07 2007.
- George, M.W., Sutton, R.T. (2006). Artificial neural networks (the multilayer perceptron) – A review of applications in the atmospheric sciences. *Atmospheric Environment*, 32 (14-15), 2627-2636
- Giannini, A., Saravanan, R., Chang, P. (2003). Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales, *Science*, 302, 1027-1030.
- Goswami, P., Srividya D. (1996). A novel neural network design for long range prediction of rainfall pattern. *Current science* 70(6), 447-457

- Gwangseob, K., Ana, P.B. (2001). Quantitative flood forecasting using multisensor data and neural networks, *J. Hydrol.*, 246, 45-62.
- Graham, N.E., Michaelsen, J., Barnett, T.P. (1987). An Investigation of the El Niño-Southern Oscillation Cycle with Statistical Models 1. Predictor Field Characteristics, *Geophysical Research Letter*, 92(C13), 14,251-14,270
- Grayson, R.B., Blöschl, G. (2000). Spatial patterns in catchment hydrology: observations and modeling. *Cambridge University Press*, UK.
- Grist, J., Nicholson, S. (2001). A Study of the Dynamic Factors Influencing the Rainfall variability in the West African Sahel. *Journal of climate* 14, 1337-1359.
- Gu, G., Adler, R. (2004). Seasonal evolution and variability associated with the West African monsoon system. *Journal of climate*, 17(582), 3364-3377.
- Gut, A., Steinebach, J. (2002). Truncated sequential change-point detection based on renewal counting processes, *Scand. J. Stat.*, 29(4), 693-719.
- Guangzhe, F., Brian, G.J. (2005). Regression tree analysis using TARGET, *Journal of Computational and Graphical Statistics*, 14:1, 206-218
- Gwangseob, K., Ana, P.B. (2001). Quantitative flood forecasting using multisensor data and neural networks, *J. Hydrol.*, 246, 45-62.
- Hall, N.M.J., Peyrillé, P. (2006). Dynamics of the West African Monsoon, *J. Phys. IV*, France, 139, 81-99. *EDP Sciences*, Les Ulis. DOI: 10.1051/jp4:2006139007.
- Hamatan, M. (2002). Synthèse et évaluation des prévisions saisonnières en Afrique de l'Ouest. DEA Sciences de l'Eau dans l'Environnement Continental. *Université Montpellier 2*. 115pp.
- Hamatan, M., Mahe, G., Servat, E., Paturel, J.E., Amani, A. (2004): Synthèse et évaluation des prévisions saisonnières en Afrique de l'Ouest. *Sécheresse*, 15 (3), 279-86
- Hansen, J.W. (2002). Realizing the potential benefits of climate prediction to agriculture: issues, approaches, challenges. *Agricultural Systems*, 74(3), 309-330.
- Hansen, J.W. et al., (2011). Review of Seasonal Climate Forecasting for Agriculture in Sub-Saharan Africa. *Experimental Agriculture*, 47(2), 205-240.
- Hansen, M., DeFries, R., Townshend, J., Sohlberg, R. (1998). 1 Km Land Cover Classification Derived from AVHRR. <http://glcf.umiacs.umd.edu/data/landcover>
- Hastenrath, S. (1984). Interannual Variability and Annual Cycle: Mechanisms of Circulation and Climate in the Tropical Atlantic Sector. *Monthly Weather Review*, 112, 1097-1107.
- Hastenrath, S. (1990). Decadal-scale changes of the circulation in the tropical Atlantic sector associated with Sahel drought. *International Journal of Climatology*, 10,459-472.
- Hastenrath, S. (1995). Recent advances in tropical climate prediction. *J. Climate*: 8, 1519-1532

- Hastenrah, S., Nicklis, A., Greischar, L. (1993). Atmospheric-hydrospheric mechanisms of climate anomalies in the West equatorial Indian ocean. *J. Geophys. Res.*, 98, 219-235
- Hastenrath, S., Wolter, K. (1992). Large-scale patterns and long-term trends of circulation variability associated with Sahel rainfall anomalies. *Journal of the Meteorological Society of Japan*, 70, 1045-1056.
- Hastie, T., Tibshirani, R., Friedman, J. (2001). *The Elements of Statistical Learning*. Springer, New York.
- Hayes, M., Svoboda, M., Le Comte, D., Redmond, K., Pasteris, P. (2005). Drought monitoring: New tools for the 21st century, in: *Drought and Water Crises: Science, Technology, and Management Issues*, edited by: Wihite, D.A., Talyor Francis, BocaRaton (LA), 53-69
- Hecht-Nielsen, R. (1995). Replicator neural networks for universal optimal source coding. *Science*, 269, 1860-1863.
- Held, I., Delworth, T., Lu, J., Findell, K., Knutson, T. (2005). Simulation of Sahel drought in the 20th and 21st centuries, *Proceedings of the National Academy of Sciences of the United States of America* 102(50), 17891-17891.
- Hingray, B., Picouet, C., Musy, A. (2009). *Hydrologie, Une science pour l'ingénieur*. Première édition. *Presses polytechniques et universitaires romandes*. 607pp
- Hoerling, M., Hurrell, J., Eischeid, J., Phillips, A. (2006). Detection and Attribution of Twentieth-Century Northern and Southern African Rainfall Change. *Journal of Climate*, 19, 3989-4008.
- Hsieh, W.W. (2004). Nonlinear multivariate and time series analysis by neural network methods. *Reviews of Geophysics*, 42. RG1003, doi:10.1029/2002RG000112.
- Hsu, K., Gupta, H.V., Sorooshian, S. (1995). Artificial neural network modeling of the rainfall-runoff process, *Water Resour. Res.*, 31(10), 2517-2530
- Hung, N.Q., Babel, M.S.; Weesakul, S., Tripathi, N.K. (2008). An artificial neural network for rainfall forecasting in Bangkok, Thailand. *Hydrological and Earth Systems Sciences Discussions*, 5, 183-218.
- Hutjes, R.W.A., Kabat, P., Running, S.W., Shuttleworth, W.J., Field, C., Bass, B., Da Silva, A.M.F., Avissar, R., Becker, A. (1988). Biospheric aspects of the hydrological cycle. *J. Hydrology*. 212-213, 1-21
- Ibrahim, B., Karambiri, H., Polcher, J., Yacouba, H., Ribstein, P. (2014). Changes in rainfall regime over Burkina Faso under the climate change conditions simulated by 5 regional climate models. *Clim Dyn* 42:1363-1381
- Ingram, K., Roncoli, C., Kirshen, P. (2002): Opportunities and Constraints for Farmers of West Africa to Use Seasonal Precipitation Forecasts with Burkina Faso as a Case Study, *Agricultural Systems*, 74, 331-349, 2002.
- International Research Institute for Climate and Society (IRI) Data Library. Available Online: <http://iridl.ldeo.columbia.edu> (accessed on 9 September 2015).

- Janicot, S. (1992). Spatiotemporal variability of West African rainfall. Part II: associated surface and air mass characteristics. *Journal of Climate*, 5, 499-511.
- Janicot, S (1997): ENSO Impact on atmospheric circulation and convection over the tropical Atlantic and West Africa. *Annales Geophysicae*, 15, 471-475
- Janicot, S., Fontaine, B. (1993) : L'évolution des idées sur la variabilité interannuelle récente des précipitations en Afrique de l'Ouest, *La Météorologie*, 8, 28-53.
- Janicot, S., Harzallah, Fontaine, A. B., Moron, V. (1998). West African Monsoon Dynamics and Eastern Equatorial Atlantic and Pacific SST Anomalies (1970-88), *Journal of Climate*, 11, 1874-1882.
- Janicot, S., Moron, V., Fontaine, B. (1996). Sahel droughts and ENSO dynamics. *Geophysical Research Letters*, vol. 23, p. 515-518.
- Janicot, S., Trzaska, S., Pocard, I. (2001). Summer Sahel-ENSO teleconnection and decadal time scale SST variations, *Climate Dynamics*, 18, 303-320.
- Johnson, S.K., Dominique, J.O. (1988). Extracting topographic structure from digital elevation data for Geographic Information Systems analysis. *Photogrammetric Engineering and Remote Sensing*. 54(11), 1593-1600.
- Jolliffe, I.T. (1986). Principal Component Analysis. *Springer-Verlag*, New York. 502 pp.
- Jolliffe, I.T., Stephenson D.B. (2012). Forecast verification: A practitioner's guide in Atmospheric science. 2nd edition, *John Wiley & Sons*. 274 pp
- Jung, T., Ferranti, L., Tompkins A.M. (2006). Response to the summer 2003 Mediterranean SST anomalies over Europe and Africa. *J. Climate*, 19, 5439-5454.
- Kannan, S., Ghosh, S. (2010). Prediction of daily rainfall state in a river basin using statistical downscaling from GCM output. *Stoch Environ Res Risk Assess*. DOI 10.1007/s00477-010-0415-y
- Kharin, V.V., Zwiers, F.W. (2002). Climate prediction with multi-model ensembles, *Journal of Climate*, 15, 793-799
- Kirby, M.J., Miranda, R. (1996). Circular nodes in neural networks. *Neural Computation*, 8, 390-402.
- Kirshen, P., Ingram, K., Hoogenboom, G., Jost, C., Roncoli, C., Ruth, M., Knee, K. (2003). Lessons learned for climate change adaptation; part 1 - implementation of seasonal climate forecasting in West Africa; part 2 - impacts from and adaptation to climate change in metro Boston, USA. Prepared for Insights and Tools for Adaptation: *Learning from Climate Variability*, 18-20, Washington, DC
- Klugman, A. S., Parsa, R. (1999). Fitting bivariate loss distribution with copulas. *Inst. Math. Econ*. 24, 139-148.
- Knebl, M.R., Yang, Z.L., Hutchison, K., Maidment, D.R. (2005). Regional Scale Flood Modeling using (NEXRAD). Rainfall GIS, and HEC-HMS/RAS: A Case study for the San Antonio River Basin Summer 2002 storm Event, *Journal of Environment Management*, 75, 325-336

- Knox, R.A. (1987). The Indian Ocean: interaction with the monsoon, in "Moonsoons", J.S. Fein and P.L. Stephens, Wiley & sons, 365-397
- Konte, O. (2011). Vérification des prévisions climatiques saisonnières sur les précipitations en Afrique de l'Ouest (PRESAO) sur la période Juillet-Aout-Septembre (JAS) de 1998-2010, au Sénégal. (draft).
- Korecha, D., Barnston, A.G. (2006). Predictability of June-September rainfall in Ethiopia. *Monthly Weather Review*, 135, 628-650
- Koutsoyiannis, D. (2003). Rainfall disaggregation methods: theory and applications. *Workshop on statistical and mathematical methods for Hydrological analysis-Rome*, Italia
- Kramer, M.A. (1991). Nonlinear principal component analysis using auto-associative neural networks. *AIChE Journal*, 37, 233-243.
- Krishnamurti, T.N., Kishtawal, C.M., Zhan, Z., Larow, T., Bachiochi, D., Williford, E., Gadgil, S., Surendran, S. (2000). Multimodel ensemble forecasts for weather and seasonal climate, *Journal of Climate*, 13, 4196-4216.
- Kumar, K.K., Soman, M.K., Kumar, K.R. (1995). Seasonal forecasting of Indian summer monsoon rainfall: A review. *Weather* 50(12), 449-467
- Kwon, H.H., Brown, C., Xu, K.; Lall, U. (2009). Seasonal and annual maximum streamflow forecasting using climate information: application to the Three Gorges Dam in the Yangtze River basin, China. *Hydrol. Sci. J.*, 54, 582-595.
- Lachenbruch P.A., Mickey, M.R. (1968). Estimation of error rate in discriminant analysis. *Technometrics*. 10, 1-11
- Lambergeon, D. (1977) Relation entre les pluies et les pressions en Afrique occidentale. *Publication* no. 57, Direction de l'Exploitation Météorologique de l'ASECNA, Dakar, Sénégal.
- Lamb, P.J. (1978). Case studies of tropical Atlantic surface circulations patterns during recent sub-Saharan weather anomalies: 1967 and 1968, *Monthly Weather Review*, 106, 482-491.
- Lamb, P.J., Pepler, R.A. (1991). West Africa. Teleconnections Linking Worldwide Climate Anomalies, M. H. Glantz, R. W. Katz, and N. Nicholls, Eds., *Cambridge University Press*, 121-189.
- Lafore, J., Flamant, C., Giraud, V., Guichard, F., Knippertz, P., Mahfouf, J., Mascart, P., Williams, E. (2010). Introduction to the AMMA special issue on Advances in understanding atmospheric processes over West Africa through the AMMA field campaign, *Quarterly Journal of the Royal Meteorological Society* 136 (S1), 2-7.
- Laganier, R., Scrawell, H.J. (2004). Risque d'inondation et aménagement durable des territoires. *Presses Universitaires du Septentrion*, 226 pp.
- LeBarbé, L., Lebel, T. (1997). Rainfall climatology of the HAPEX-Sahel region during the years 1950-1990. *Journal of Hydrology*, 188, 43-73.

- Legates, D.R., McCabe, J.G.J. (1999). Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation, *Water Resour. Res.*, 35, 1, 233-241.
- Li, X., Sailor, D. (2000). Application of tree-structured regression for regional precipitation prediction using general circulation model output. *Climate research* vol. 16: 17-30
- Liu, Z., Alexander, M. (2007). Atmospheric bridge, oceanic tunnel, and global climatic teleconnections. *Rev. Geograph.*, 45, 1–34.
- Livezy, R.E. (1995). The evaluation of forecast, in "Analysis of climate variability", Springer, Berlin, 177-196.
- Lopez-Bustins, J.A., Martin-Vide, J., Sanchez-Lorenzo, A., (2008). Iberia winter rainfall trends based upon changes in teleconnection and circulation patterns. *Global and Planetary Change*, 63, 171-176.
- Lopez-Parages, J., Rodriguez-Fonseca, B. (2012). Multidecadal modulation of El Niño influence on the Euro-Mediterranean rainfall. *Geophys. Res. Lett.*, 39, doi: 10.1029/2011GL050049.
- Lund, R., Reeves, J. (2002). Detection of undocumented changepoints: a revision of the two-phase regression model. *Journal of Climate*, 15, 2547-2554
- Luo, P., Takara, K., He, B., Cao, W., Yamashiki, Y., Nover, D. (2011). Calibration and uncertainty analysis of SWAT model in a Japanese river catchment. *Annual Journal of Hydraulic Engineering*, 55, 61-66.
- Mahé, G., Citeau, J. (1993). Relations océan-atmosphère-continent dans l'espace africain de la mousson atlantique. Schéma général et cas particulier de 1984, *Veille Climatique Satellitaire* Ed. ORSTOM-METEO FRANCE 44, 34-54.
- Mahé, G., Girard, S., New, M., Paturel, J., Cres, A., Dezetter, A., Dieulin, C., Boyer, J., Rouche, N., Servat, E., (2008). Comparing available rainfall gridded datasets for West Africa and the impact on rainfall-runoff modelling results, the case of Burkina-Faso, *Water Sa.* 34(5), 529-536.
- Maheepala, S., Perera, B.J.C. (1996). Monthly hydrologic data generation by disaggregation. *Journal of hydrology* 178, 277-291
- Malardel, S. (2005): Fondamentaux de météorologie, à l'école du temps, CEPADUES.
- Malthouse, E.C. (1998). Limitations of nonlinear PCA as performed with generic neural networks. *IEEE Transactions on Neural Networks*, 9, 165-173.
- Manusthiparom, C., Oki, T., Kanae, S. (2003). Quantitative Rainfall Prediction in Thailand, First *International Conference on Hydrology and Water Resources on Asia Pacific Region* (APHW), Kyoto, Japan, 13-15 March.
- Mara, F. (2010). Développement et analyse des critères de vulnérabilité des populations sahéliennes face à la variabilité du climat: le cas de la ressource en eau dans la vallée de la Sirba au Burkina Faso. Thèse de doctorat. *Université du Québec à Montréal*. 273pp

- Mason, S., Mimmack, G. (2001). Comparison of some statistical methods of probabilistic forecasting of ENSO. *Journal of climate*, 15, 8-29
- Michaelson, J. (1987). Cross-Validation in statistical climate forecast models. *Journal of climate and applied meteorology*. 26: 1589-1600
- Millennium Ecosystem Assessment (MEA), (2005). Current state and trends assessment, Washington D.C., Island Press.
- Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., (2008). Stationarity is dead: Whither water management, *Science* 319, 573-574.
- Mohino, E., Rodriguez-Fonseca, B., Mechoso, C.R., GERVOIS, S. Ruti, P., Chauvin, F. (2011). Impacts of the Tropical Pacific/Indian Oceans on the Seasonal Cycle of the West African Monsoon. *Journal of Climate*, 24, 3878-3891
- Monahan, A.H., Fyfe, J.C., Pandolfo, L. (2003). The vertical structure of winter time climate regimes of the northern hemisphere extratropical atmosphere. *J. Climate*, 16, 2005-2021.
- Mondal, M.S., Wasimi, S.A. (2005). Disaggregation model for synthetic stream-flow generation. *Journal of civil engineering*, 33 (1) 43-54
- Moreno, E., Casella, G., Garcia-Ferrer, A. (2005). An objective Bayesian analysis of the change point problem, *Stochastic Environ. Res. Risk Assess.* (SERRA), 19(3), 191-204.
- Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *American Society of Agricultural and Biological Engineers*, 50, 885-900.
- Moron, V. (1994). Guinean and Sahelian rainfall anomaly indices at annual and monthly scales (1933-1990). *International Journal of Climatology*, 14, 325-341.
- Moron, V., N. Philippon. et B. Fontaine (2004): Simulation of West African monsoon index in four Atmospheric General Circulation Models forced by prescribed sea surface temperature, *Journal of Geophysical Research*, 109, doi:10.1029/2004JD004760.
- Murphy, A.H. (1988). Skill scores based on the mean square error and their relationship to the correlation coefficient. *Mon. Wea. Rev.* 116, 2417-2424
- Napradean, I., Chira, R. (2006). The hydrological modeling of the Usturoi Valley - Using two modeling programs - WetSpa and HecRas. *Carpathian Journal of Earth And Environmental Sciences*. v1.issn (1842-4090.53-62).
- Nash, J.E., Sutcliffe, J.V. (1970). River flow forecasting through conceptual models, Part I - A discussion of principles, *J. Hydrol.*, 10, 282-290
- Ndiaye, O., Goddard, L., Ward, M.N. (2009). Using regional wind fields to improve general circulation model forecasts of July-September Sahel rainfall. *International Journal of Climatology*, 29, 1262-1275.

- Ndiaye, O., Ward, M.N., Thiaw, W.M. (2011). Predictability of seasonal Sahel rainfall using GCMs and lead-time improvements through the use of a coupled model. *Journal of Climate*, 24, 1931-1949.
- New, M., Lister, D., Hulme, M., Makin, I. (2002). A high-resolution data set of surface climate over global land areas, *Climate research* 21(1), 1-25.
- Ngoran, S.D., Dogah, K.E., Xue, X.Z. (2015). Assessing the Impacts of Climate Change on Water Resources: The Sub-Saharan Africa Perspective. *Journal of Economics and Sustainable Development*. (6). 185-193
- Nicholson, S.E. (2013). The West African Sahel: A review of recent studies on the rainfall regime and its interannual variability. *ISRN Meteor.*, 453–521.
- Nicholson, S.E., Kim, J. (1997). The relationship of the El Nino-Southern Oscillation to African rainfall. *International Journal of Climatology*, 17, 117-135.
- O'Connor, K.M. (2006). River flow forecasting. In: River Basin Modelling for Flood Risk Mitigation (D.W. Knight and A.Y. Shamseldin, eds). London, *Taylor and Francis Group*
- Ogallo, L.A., Boulahya, M.S., Keane, T. (2000). Applications of seasonal to interannual climate prediction in agricultural planning and operations. *Agricultural and Forest Meteorology*, 103, 159-166.
- Olivera, F., Valenzuela, M., Srinivasan, R., Choi, J., Cho, H. D., Koka, S., Agrawal, A. (2006). ArcGIS-SWAT: A geodata model and GIS interface for SWAT, *Journal of the American Water Resources Association* 42, 295-309.
- Olivry, J. C. (1993). Evolution récente des régimes hydrologiques en Afrique Intertropicale. L'Eau la terre et les hommes, *Presses Universitaires de Nancy*.
- Ouédraogo, A. (2008). Facteurs de vulnérabilité et stratégies d'adaptation aux risques des maraîchers urbains et périurbains dans les villes de Ouahigouya et Koudougou. Mémoire de fin de cycle, Université polytechnique de Bobo-Dioulasso, Institut du développement rural, Burkina Faso, 78 p.
- Palmen, E. (1951). The role of atmospheric disturbances in the general circulation. *Quart. J. Roy. Meteor. Soc.*, 77; 337-3354
- Palmer T.N., (1986). Influence of the Atlantic, Pacific and Indian Oceans on Sahel rainfall. *Nature*, 322, 251-253.
- Palmer, T., Brankovic, C., Viterbo, P., Miller, M. (1992). Modeling interannual variations of summer monsoons. *Journal of Climate*, 5, 399-417.
- Palmer, T.N. et D.L.T. Anderson (1994): The prospects for seasonal forecasting: A review paper, *Quarterly Journal of the Royal Meteorological Society*, 120, 755–793.
- Palmer, T., N.A. Alessandri, U., Andersen, P., Cantelaube, M., Davey, P., Délécluse, Déqué, M., Diez., E., Doblas-Reyes, J.F., Feddersen, H., Graham, R., Gualdi, S., J.F. Guérémy, J.F., Hagedorn, R., Moshen, M., Keenlyside, N., Latif, M., Lazar, A., Maisonnave, E., Marletto, V., Morse, A.P., Orfila, B., Rogel, P., Terres, J.M.,

- Thomson, M.C. (2004). Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER), *Bulletin of the American Meteorological Society*, 85, 853-872.
- Palmer, T.N., Brankovic, C., Richardson, D.S. (2000). A probability and decision-model Analysis of PROVOST seasonal multi-model ensemble integrations, *Quarterly Journal of the Royal Meteorological Society*, 126, 2013-2034.
- Parker, D. Burton, R., Diongue-Niang, A., Ellis, R., Felton, M., Taylor, C., Thorncroft, C., Bessemoulin, P., Tompkins, A. (2005). The diurnal cycle of the West African monsoon circulation, *Quarterly Journal of the Royal Meteorological Society* 131(611), 2839-2860.
- Pavan, V., Doblas-Reyes, F.J. (2000). Multi-model seasonal hindcasts over the Euro-Atlantic: skills cores and dynamic features, *Climate Dynamics*, 16, 611-625.
- Peng, P., A. Kumar, A., Van den Dool, H. (2002). Analysis of multimodel ensemble predictions for seasonal climate anomalies, *Journal of Geophysical Research*, 107, 4710, doi: 10.1029/2002JD002712.
- Perreault, L., Bernier, J., Bobée, B., Parent, E. (2000). Bayesian change-point analysis in hydrometeorological time series 2. Part 2. Comparison of change-point models and forecasting, *J. Hydrol.*, 235, 242-263
- Perrin, C., Michel, C., Andréassian, V. (2001). Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *J. Hydrol.*, 242:275-301.
- Peterson, T.C. et al. (1998). Homogeneity adjustments in situ atmospheric climate data: A review. *Intl. J. Climatol.*, 18, 1493-1517.
- Peyrillé, P. (2006). Etude idéalisée de la mousson de l'Afrique de l'Ouest à partir d'un modèle numérique bidimensionnel, PhD thesis, *Université Paul Sabatier, Toulouse, France*.
- Philippon, N. (2002). Une nouvelle approche pour la prévision statistique des précipitations saisonnières en Afrique de l'Ouest et de l'Est : méthodes, diagnostics (1968-1998) et applications (2000-2001). *Thèse de doctorat*, CRC - Université de Bourgogne, Dijon, France, 241 pp
- Philippon, N., Fontaine, B. (1999): A new statistical predictability scheme for July-September Sahel rainfall (1968-1994), *Comptes Rendus de Acadmie des Sciences*, 329, p 1-6.
- Philippon, N., Fontaine, B. (2002). The relationship between the Sahelian and previous 2nd Guinean rainy seasons: a monsoon regulation by soil wetness. *Annales Geophysicae*, 20, 575-582.
- Piechota, T.C., Chiew, F.H.S., Dracup J.A., McMahon, T.A., (1998). Seasonal streamflow forecasting in eastern Australia and the El Nino-Southern Oscillation. *Water Resources Research*, 34, 3035-3044.
- Polo, I., Ullmann, A., Roucou, P., Fontaine, B. (2011). Weather regimes in the Euro-Atlantic and Mediterranean sector and relationship with West African rainfall over

the period 1989–2008 from a self-organizing maps approach. *J. Clim.*, 24, 3423–3432.

- Porter, J.W., Pink, B.J. (1991). A method of synthetic fragments for disaggregation in stochastic data generation. *Int. Hydrol. And Water RES. Symp.*, Perth, W.A. The institution of Eng. Australia, Canberra, pp. 187-191
- Quan, X., Hoerling, M., Whitaker, J., Bates, G., Xu, T. (2006). Diagnosis sources of US seasonal forecast skill. *Journal of climate*, 19 (13), 3279-3293
- Raichich, F., Pinardi, N., Navarra, A. (2003). Teleconnections between Indian Monsoon and Sahel rainfall and the Mediterranean. *International Journal of Climatology*, 23, 173-186.
- Raje, D., Mujumdar, P.P. (2011). A comparison of three methods for downscaling daily precipitation in the Punjab region. *Hydrological Process*, 25, 3575-3589.
- Ramel, R. (2005). Impact des processus de surface sur le climat en Afrique de l'Ouest. Thèse de doctorat, *Université Joseph Fourier de Grenoble*, France.
- Reeves, J., Chen, J., Wang, X.L., Lund, R., Lu, Q. (2006). A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology*, 46, 900-914.
- Rodionov, S.N. (2004). A sequential algorithm for testing climate regime shifts. *Geophys.Res. Lett.*, 31, L09204, doi: 10.1029/2004GL019448.
- Roncoli, C., et al. (2009): From accessing to assessing forecasts: An end-to-end study of participatory forecast dissemination in Burkina-Faso West Africa. *Clim. Change*, 92, 433-460.
- Roncoli, C., Ingram, K., Jost, C., Kirshen, P. (2004): Meteorological Meanings: Farmers' Interpretations of Seasonal Rainfall Forecasts in Burkina Faso. In Sarah Strauss and Benjamin Orlove eds. *Weather, Climate, Culture*, Berg.
- Roncoli, C., Ingram, K., Kirshen, P. (2001): The costs and risks of coping with drought: Livelihood impacts and farmers' responses in Burkina Faso. *Clim. Res.*, 19, 119-132.
- Roncoli, C., Ingram, K., Kirshen, P. (2002): Reading the rains: Local knowledge and rainfall forecasting in Burkina Faso. *Soc. Nat. Resour.*, 15-5, 409-427.
- Ropelewski, C.F., Halpert, M.S. (1987). Global and regional patterns associated with the ENSO. *Monthly Weather Review*, 115, 1606-1626.
- Ropelewski, C.F., Halpert, M.S. (1989). Precipitation patterns associated with the high index phase of the Southern Oscillation. *Journal of Climate*, 2, 268-284.
- Rowell, D.P. (2001). Teleconnections between the tropical Pacific and the Sahel. *Q. J. R. Meteorol. Soc.*, 127, 1683–1706.
- Rowell, D.P. (2003). The impact of Mediterranean SSTs on the Sahelian rainfall season, *Journal of Climate*, 16, 849-862.

- Rowell, D.P., Folland, C.K., Maskell, K., Ward, M.N. (1995). Variability of summer rainfall over Tropical North Africa (1906-1992): Observations and modelling, *Quarterly Journal of Royal Meteorological Society*, 121, 669-704.
- Ruiz, J. E., Cordery, I., Sharma, A. (2005). Integrating Ocean Subsurface Temperatures in Statistical ENSO Forecasts. *Journal of climate*; 18, 3571-3586
- Salinger, M. (2005). Climate variability and change: Past, present and future - An overview, *Clim. Change*, 70, 9-29.
- Sahai, A.K., Soman, M.K., Satyan, V. (2000). All India summer monsoon rainfall prediction using an artificial neural network. *Climate Dynamics*, 16, 291-302
- Samimi, C., Fink, A.H., Paeth, H. (2012). The 2007 flood in the Sahel: causes, characteristics and its presentation in the media and FEWS NET. *Natural Hazards and Earth System Science*, 12, 313-325.
- Sagarin, R., Micheli, F. (2001). Climate change in nontraditional data sets, *Science*, 294, 811.
- Saporta, G. (2006) : Probabilités, analyse des données et statistique, 2^{ème} Ed, Editions Technip, Paris.
- Sarr, M.A., Zoromé, M., Seidou, O., Bryant, C.R., Gachon, P. (2013). Recent trends in selected extreme precipitation indices in Senegal-A change point approach, *Journal of Hydrology*, doi: <http://dx.doi.org/10.1016/j.jhydrol.2013.09.032>
- Savadogo, R.C. (2004). Rapport provisoire : adaptation au changement climatique pour le système hydrologique des fleuves sahéliens et bassins versants de leurs affluents: cas de la Sirba au Burkina Faso, Niger. Niamey *AGRHYMET*: 54 p.
- Sawadogo, L., Nyagard, R., Pallo, F. (2002). Effects of livestock and prescribed fire on coppice growth after selective cutting of Sudanian savannah in Burkina Faso. *Annals of Forest Science*. vol. 59, no 2, p. 185-196.
- Schepen, A., Wang, Q.J., Robertson, D.E., (2012). Combining the strengths of statistical and dynamical modeling approaches for forecasting Australian seasonal rainfall. *Journal of Geophysical Research*, 117, 148-227.
- Scholz, M., Martin, F., Joachim, S. (2007). Nonlinear principal component analysis: neural network models and applications. In principal manifolds for Data visualization and dimension reduction, edited by Alexander N. Gorban, Balazs Kegl, Donald C. Wunsch, and Andrei Zinovyev. Vol 58 of LNCSE, pp 44-67. Springer Berlin Heidelberg.
- Seidou, O., Asselin, J. J., Ouarda, T. B. M. J. (2007). Bayesian multivariate linear regression with application to changepoint models in hydrometeorological variables, *Water Resour. Res.*, doi: 10.1029/2005WR004835
- Seidou, O., Ouarda, T.B.M.J. (2007). Recursion-based multiple changepoint detection in multiple linear regression and application to river streamflows. *Water Resources Research*, Vol. 43, W07404, doi: 10.1029/2006WR005021

- Servat, E., Paturel, J.E., Lubès, H., Kouamé, B., Quedraogo, M., Mason, J.M. (1997). Climatic variability in humid Africa along the Gulf of Guinea. Part I: detailed analysis of the phenomenon in Côte d'Ivoire. *Journal of Hydrology* 191, 1-15.
- Shamseldin, A.Y. (1997). Application of a neural network technique to rainfall-runoff modeling, *J. Hydrol.*, 199, 272-294
- Sighomnou D., Descroix L., Genthon P., Mahe G., Bouzou Moussa I., Gautier E., Mamadou I., Vandervaere J.P., Bachir T., Coulibaly B., Rajot J.L., Malam Issa O., Malam Abdou M., Dessay N., Delaitre E., Maiga O.F., Diedhiou A., Panthou G., Vischel T., Yacouba H., Karambiri H., Paturel J.E., Diello P., Mougin E., Kergoat L., Hiernaux P., (2013). La crue de 2012 a Niamey : un paroxysme du paradoxe du Sahel ? *Secheresse* 24 : 3-13. doi: 10.1684/sec.2013.0370 T
- Singh V.P. (2005). Watershed modelling. In: Singh V.P., ed. Computer models of watershed hydrology. Colorado, USA: *Water Ressources Publications*, Highlands Ranch, 1-22.
- Sittichok, K.; Gado Djibo, A.; Seidou, O.; Saley, H.M.; Karambiri, H.; Paturel, J. (2014). Statistical seasonal rainfall and streamflow forecasting for the Sirba watershed, using sea surface temperature. *Hydrol. Sci. J.*, doi:10.1080/02626667.2014.944526.
- Sivakumar, M. (1988). Predicting rainy season potential from the onset of rains in Southern Sahelian and Sudanian climatic zones of West Africa. *Agricultural and Forest Meteorology* 42(4), 295-305.
- Shaman, J., Tziperman, E. (2011). An atmospheric teleconnection linking ENSO and Southwestern European precipitation. *J. Clim.*, 24, 124–139.
- Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* 8, 229–231.
- Smith, J.A., Baeck, M.L., Villarini, G., Krajewski, W.F. (2010). The hydrology and hydrometeorology of flooding in the Delaware River basin. *Journal of hydrometeorology* 11, 841-859
- Soleymani M., Hossein S., Amir K., Fereydoon K., Hossein B. (2014). Determining Flood Zones Using HEC-RAS Model (Case study: Gale Hassan River situated in Atrak watershed). *Bull. Env. Pharmacol. Life Sci.*, 3 (7), 102-107
- Solomon, A., Goddard, L., Kumar, A., Carton, J., Deser, C., Fukumori, I., Greene, A.M., Hegerl, G., Kirtman, B., Kushnir, Y. (2011). Distinguishing the roles of natural and anthropogenically forced decadal climate variability. *Bull. Am. Meteorol. Soc.* 92, 141–156.
- Solow, A. R. (1987). Testing for climate change: An application of the two-phase regression model, *J. Appl. Meteorol.*, 26, 1401-1405.
- SRTM (2004). DEM data from International Centre for Tropical Agriculture (CIAT), available from the CGIAR CSI SRTM 90m Database: <http://srtm.csi.cgiar.org>
- Sultan, B., Janicot, S. (2003). The West African monsoon dynamics. Part II: The "preonset" and "onset" of the summer monsoon, *Journal of climate* 16(21), 3407-3427.

- Sultan, B., Janicot, S. (2004). La variabilité climatique en Afrique de l'Ouest aux échelles intra-saisonnières. 1ère partie : Analyse diagnostique de la mise en place de la mousson et de la variabilité intra-saisonnière de la convection. *Sécheresse*, 15(4), 1-10.
- Svanidze, G. G. (1977). Mathematical modeling of hydrologic series, T. Guerchon, translator, *Water Resources Publications*, Littleton, Colorado.
- Tangang, F.T., Hsieh, W.W., Tang, B. (1998). Forecasting regional sea surface temperatures in the tropical Pacific by neural network models, with wind stress and sea level pressure as predictors. *Journal of Geographical Research* 103(C4), 7511-7522
- Tarhule, A. (2005). Damaging rainfall and flooding: the other Sahel Hazards. *Climatic Change*, 72(3), 355-377.
- Taweye, A. (1995). Contribution à l'étude hydrologique du bassin versant de la Sirba à Garbé-Kourou. Dissertation, Centre Régional AGRHYMET, 96pp.
- Thiaw, W.M., Mo, K.C. (2005). Impact of sea surface temperature and soil moisture on seasonal rainfall prediction over the Sahel. *Journal of Climate*, 18, 5330-5343.
- Tomozeiu, R., Stefan, S., Busuioc, A. (2005). Winter precipitation variability and large-scale circulation patterns in Romania. *Theoretical and Applied Climatology* 81, 193-201
- Toth, E., Brath, A., Montanari, A. (2000). Comparison of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology*, 239, 132-147.
- Trzaska, S., Moron, V., Fontaine, B. (1996). Global atmospheric response to specific Linear-combinations of the main SST modes. Part I: Numerical experiments and preliminary-results, *Annales Geophysicae-Atmospheres Hydrospheres and Space Sciences V*, 14, P, 1066-1077.
- Tuteja, N.K, Shin D., Laugesen, R., Khan, U., Shao, Q., Wang, E., Li, M., Zheng, H., Kuczera, G., Kavetski, D., Evin, G., Thyer, M., MacDonald, A., Chia, T., Le, B. (2011). *Experiment evaluation of the dynamic seasonal streamflow forecasting approach*. Technical Report, Bureau of Meteorology, Melbourne.
- Udall, B., Hoerling, M. (2005). Seasonal forecasting: skill in the intermountain west? Feature Article from Intermountain West Climate Summary, May 2005. A joint project of University of Colorado and NOAA Climate Diagnostics Center.
- UK Met Office Hadley Centre, (2010). Climat sahélien: rétrospective et projections. *SICCS Project Report*. 32pp
- US Army Corps of Engineers-USACE, (2009). *HEC-GeoRAS, GIS Tools for Support of HECRAS using ArcGIS. User's Manual, Version 4.2, U.S. Army Corps of Engineers, Davis, California*, 246 p.
- Vautard, R., Pires, C., Plaut, G. (1996). Long-Range Atmospheric Predictability Using Space-Time Principal Components, *Monthly Weather Review*, 124, 288-307.

- Venkatesan, C., Raskar, S.D., Tambe, S.S., Kulkarni, B.D., Keshavamurty, R.N. (1997). Prediction of all India summer monsoon rainfall using error-back-propagation neural networks. *Meteorology and Atmospheric Physics* 62 (3-4), 225-240.
- Venter, G. (2002). *Tails of copulas*. *Proc. Casualty Actuarial Society*. 89, 68-113.
- Venter, G. (2003). *Quantifying correlated reinsurance exposures with copulas*. *Casualty Actuarial Society Forum*, Spring, 215-229.
- Villarini, G., Serinaldi, F., Smith, J.A., Krajewski, W.F. (2009). On the stationarity of annual peaks in the continental United States during the 20th century. *Water Resources Research*, 45, 1-17.
- Villarini, G., Smith, J.A. (2010). Flood peak distributions for the eastern United States. *Water Resources Research* 46, doi: 10.1029/2009WR007645
- Villarini, G., Smith, J.A., Serinaldi, F., Ntelekos, A.A. (2011). Analyses of seasonal annual and maximum daily discharge records for central Europe. *Journal of hydrology* 399, 299-312
- Vincent, L.A. (1998). A technique for the identification of inhomogeneities in Canadian temperature series, *J. Clim.*, 11, 1094-1105.
- Vizy, E.K., Cook, K.H. (2001). Mechanisms by which Gulf of Guinea and Eastern North Atlantic sea surface temperatures anomalies can influence African rainfall. *Journal of Climate*, 14, 795-821.
- Vizy, E.K., Cook, K.H. (2002). Development and application of a mesoscale climate model for the tropic: Influence of sea surface temperature anomalies on the West African Monsoon. *Journal of Geophysical Research*, 107, doi: 10.1029/2001JD000686.
- Von Storch, H., Zwiers F.W. (1999). Statistical analysis in climate research, *Cambridge University press*, Cambridge, 484 pp.
- Wang, G., Eltahir, E.A.B. (2000). Role of vegetation dynamics in enhancing the low frequency variability of the Sahel rainfall. *Water Resour.*, 36, 1013-1021.
- Wang, E., Zhang, Y., Luo, J., Chie, F.H.S., Wang, Q.J. (2011). Monthly and seasonal streamflow forecasts using rainfall-runoff modeling and historical weather data, *Water Resources Research*, 47, 1-13.
- Wang, X. L. (2003). Comments on "Detection of undocumented changepoints: A revision of the two-phase regression model". *J. Climate*, 16, 3383-3385.
- Wang, Y., Witten, I.H. (1997). Induction of model trees for predicting continuous classes. *Proceedings European Conference on Machine Learning*, Prague, 1997, pp. 128-137
- Ward, M.N. (1992). Provisionally corrected surface wind data, worldwide ocean-atmosphere surface fields and Sahelian rainfall variability. *Journal of Climate*, 5, 454-475.

- Webster, P.J., Magana, V.O., Palmer, T.N., Shukla, J., Tomas, R.A., Yanai, M., Yasunari, T. (1998). The monsoon: Processes, predictability and prediction. *Journal of Geophysical Research*, 103, 14451-14510.
- Weldeab, S., Lea, D., Schneider, R., Andersen, N. (2007). 155,000 years of West African monsoon and ocean thermal evolution, *Science* 316 (5829), 1303.
- Wilks, D. (2006). Statistical Methods in the Atmospheric Sciences. *Academic Press*; 3rd edition. New York, U.S.A. 704pp
- Wilks, D.S. (2011). Statistical Methods in the Atmospheric Sciences, third edition. *Oxford: Elsevier Inc.*
- Willmott, C.J., Matsuura, K. (2001). Terrestrial air temperature and precipitation: monthly and annual time series (1950-1999). Version 1.02. Available online at: <http://climate.geog.udel.edu/~climate/>.
- Woo, M., Thorne, R. (2003). Comment on 'Detection of hydrologic trends and variability' by D. H. Burn and M. A. Hag Elnur, 2002, *Journal of Hydrology* 255, 107-122, *J. Hydrol.*, 277, 150-160.
- World Meteorological Organization - WMO (2011). Manual on Flood Forecasting and Warning. *Chairperson, Publications Board*. 142 pp
- Wu, C.L., Chau, K.W., Fan, C. (2010). Prediction of rainfall time series using singular artificial neural networks coupled with data processing techniques. *Journal of hydrology*, 389, 146-167
- Xiong, L., Guo, S. (2004). Trend test and change-point detection for the annual discharge series of the Yangtze River at the Yichang hydrological station. *Hydrological Sciences Journal* 49 (1), 99-112
- Yao, Y. (1984). Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches, *Ann. Stat.*, 12, 1434-1447.
- Yossef, N. C., Winsemius, H., Weerts, A., Beek, R.V., Bierkens, F.P., (2013). Skill of global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. *Water Resources Research*, 49, 4687-4699.
- Yu, J.R., Tzeng, G.H., Li, H.L. (2001). General fuzzy piecewise regression analysis with automatic change-point detection. *Fuzzy sets and systems* 119, 247-257
- Yun, W.T., Stefanova, L., Krishnamurti, T.N. (2003). Improvement of the super-ensemble technique for seasonal forecasts, *Journal of Climate*, 16, 3834-3840.
- Yuval, W., Hsieh, W. (2002): The impact of time-averaging on the detectability of nonlinear empirical relations. *Quarterly Journal of the Royal Meteorological Society*, 128 (583), 1609-1622
- Zealand, C.M., Burn, D.H., Simonovic, S.P. (1999). Short term streamflow forecasting using artificial neural networks, *J. Hydrol.*, 214, 32-48
- Zeng, N., Neelin, J.D., Lau, K.M., Tucker, C.J. (1999). Enhancement of interdecadal climate variability in the Sahel by vegetation interaction. *Science*, 286, 1537-1540.

- Zheng, X., Eltahir, E.A.B. (1998). The Role of Vegetation in the Dynamics of West African Monsoons. *Journal of Climate*, 11, 2078-2096.
- Zhang, L., Singh, V.P. (2006). Bivariate flood frequency analysis using the copula method. *J. Hydrol. Engng* 11(2), 150-164.
- Zhang, X., Srinivasan, R., Bosch, D. (2009). Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging. *Journal of Hydrology*, 374(3-4): 307-317.
- Zhao, H., Moore, G.W.K. (2007). Trends in the boreal summer regional Hadley and Walker circulations as expressed in precipitation records from Asia and Africa during the latter half of the 20th century, *International Journal of Climatology*, DOI: 10.1002/joc.1580.
- Zwiers, F. W., Von Storch, H. (2004). On the role of statistics in climate research. *International Journal of Climatology*, 24 (6), 665-680.