

**The Predictability of the Sahelian Climate :  
Seasonal Sahel Rainfall and Onset over Senegal.**

Ousmane Ndiaye

Submitted in partial fulfillment of the  
Requirements for the degree  
of Doctor of Philosophy  
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2010



## **ABSTRACT**

### **The Predictability of the Sahelian Climate : Seasonal Sahel Rainfall and Onset Over Senegal.**

**Ousmane Ndiaye**

The first part of this study addresses the predictability of seasonal rainfall totals in the Sahel region of West Africa. We show that, when driven with observed sea surface temperature (SST), atmospheric General Circulation Models (GCMs) do usually contain information about the observed interannual Sahel rainfall variability, through their predictions of the low-level tropical Atlantic / West Africa wind field. Most current GCMs are unable to directly represent the SST impact on the Sahel's rainfall variability. Therefore, this new result is significant for GCM prediction of Sahel rainfall at seasonal timescales, and potentially other timescales as well. For seasonal predictions, lead-time beyond a month is difficult due to SST developments. The second major result found here is the potential for coupled models to break through the lead-time barrier, with skill superior to that previously reported using empirical or GCM approaches. For the CFS model, skill levels only slightly less than those achieved by the best approaches using observed SST, have been achieved at lead-times of up to six months, utilizing the same low-level regional wind field.

The onset of the rains is critical for agriculture. Two homogeneous onset regions over Senegal are defined (southern region and northern region). Both regions experience frequent false onsets. We find relationships between onset date and a mix of large scale boundary layer and regional atmospheric synoptic features. Variables containing information include low level wind, SST, precipitable water and moist static energy. The southern region onset is influenced by the SST gradient over the tropical Atlantic. The northern region is more related to SST in the global tropics. At the synoptic time scale, we find clear propagation of signals from the eastern Sahel: relatively slow for the southern region onset (about 2.5 m/s), with faster propagation typical for the northern region (about 8.5 m/s). The risk of a false onset is increased if a negative anomaly of PWAT is located in the vicinity of Senegal and nearby tropical Atlantic, with a strong advection of cold (convection inhibiting) air from the north, a further danger sign. The CFS coupled model is found to also have some skill for onset prediction for the southern region when initialized in early May. It is concluded that combining boundary forcing and synoptic propagation can provide useful advance information about onset.

## TABLE OF CONTENTS

Table of contents .....	i
List of figures .....	iv
List of tables .....	vii
Acknowledgments .....	vii
Dedication .....	x
<b>Chapter I : Introduction</b> .....	1
I.1 Socio-economical context.....	1
I.2 Seasonal forecasting of Sahel rainfall.....	4
I.3 Rainfall Onset.....	11
I.4 Objectives of this study.....	13
<b>Chapter II : Data and Methods</b> .....	15
II.1 Data.....	15
II.1.a Rainfall data and Basic Indices.....	15
II.1.b Sea Surface Temperature (SST) .....	17
II.1.c NCEP Reanalysis data.....	17
II.1.d Outgoing Long wave Radiation (OLR) .....	17
II.1.e Atmospheric and Coupled GCM simulations.....	18
II.2 Methods .....	21
II.2.a Empirical Orthogonal Function .....	21
II.2.b Statistical measures of association and statistical significance.....	23
II.2.c Statistical Prediction Methodology .....	24
II.2.d Model validation .....	25

Reference chapters I and II .....	26
<b>Chapter III : Using regional GCM wind to forecast seasonal rainfall over Sahel.....</b>	<b>30</b>
III.1 Model performance with observed SST: direct GCM output versus a MOS system.....	31
III.1.a ECHAM4.5 AGCM simulation skill .....	31
III.1.b Using MOS to correct GCM Sahel rainfall .....	36
III.1.b.1 Robustness in the choice of the domain .....	38
III.1.b.2 Robustness in the choice of the GCMs variable .....	39
III.1.b.3 Robustness in the choice of the GCM .....	40
III.2 Performance of the MOS through the wet and dry Sahel epochs.....	48
III.3 Summary.....	54
Reference chapter III .....	55
<b>Chapter IV : Model forecast skill and sensitivity to SST developments.....</b>	<b>56</b>
IV.1 Performance with persisted SSTA .....	56
IV.2 Key SST developments between May and June.....	61
IV.3 The problem of SST development in empirical prediction schemes.....	66
IV.4 Summary.....	70
Reference chapter IV .....	71
<b>Chapter V : Predictability of seasonal Sahel rainfall using coupled models and lead- time improvements .....</b>	<b>72</b>
V.1 Predictability from April/May initial conditions.....	73
V.2 Further analysis of the CFS predictability and comparisons at up to 6 month lead-time.....	87

V.3 The trend component of Sahel rainfall 1981-2008 in observations and CFS seasonal predictions.....	94
V.4 Summary .....	104
Reference chapter V.....	105
<b>Chapter VI : Onset of the Sahelian Rains: A study over Senegal .....</b>	<b>106</b>
VI.1 Method.....	107
VI.1.a Definition of onset date.....	107
VI.1.b regional indices of onset date.....	111
VI.2 Predictability of monthly rainfall around the mean onset date.....	120
VI.3 Large scale teleconnection with regional onset indices : SST and wind.....	123
VI.4 Atmospheric conditions around the onset date.....	129
VI.4.a Circulation around the onset.....	131
VI.4.b Atmospheric propagating features around the onset.....	135
VI.5 True versus false onset date .....	138
VI.6 Toward predicting the onset date.....	145
VI.6.a Precursor signals in the annual cycle.....	146
VI.6.b Using large scale ocean/atmosphere relationships to forecast the onset date.....	149
VI.6.c Using propagating signals as an early warning.....	153
VI.6.d Forecasting the onset date using CFS .....	155
VI.7 Summary .....	161
Reference chapter VI and conclusion .....	164
<b>Conclusion .....</b>	<b>167</b>

## LIST OF FIGURES

### Chapter I : Introduction

**Fig. I.1** : Climatology of : a) NDVI during October and b) JAS Sahel rainfall indices ....7

### Chapter II : Data and Methods

**Fig. II.1** : Location of the 20 rain gauge stations in Senegal .....16

### Chapter III : Using regional GCM wind to forecast seasonal rainfall over Sahel

**Fig. III.1** : Climatology wind simulated by ECHAM4.5 and from NCEP .....32

**Fig. III.2** : Skill of the ECHAM4.5 GCM wind and rainfall .....33

**Fig. III.3** : JAS WAMI from NCEP and ECHAM4.5 GCM .....34

**Fig. III.4** : Standardized July-September Sahel rainfall indices .....35

**Fig. III.5** : EOF1 of the 925hPa GCM zonal wind during JAS 1968-2002.....37

**Fig. III.6** : Sahel rainfall skill from direct GCMs and using MOS .....41

**Fig. III.7a** : The 925 hPa zonal wind EOF from 8 GCMs.....42

**Fig. III.7b** : The 925 hPa climatological wind simulated by AGCMs.....43

**Fig. III.8** : Time-series of each of the EOFs shown in Fig. III.7a .....45

**Fig. III.9** : Correlation between observed Sahel rainfall index and tropical SST during JAS a)1950-1969 and b)1970-2002.....48

**Fig. III.10** : Regression between Observed Sahel index and each component of the GCM wind at 925hPa during JAS for a) 1950-1969 and b) 1970-2002.....50

**Fig. III.11** : 925hPa zonal wind a) EOF1 1950-02 b) EOF3 1950-02 .....51

**Fig. III.12** : JAS Sahel rainfall observed and predicted using the GCM MOS .....51

**Fig. III.13** : EOF1 during a) wet period 1950-69 and b) dry period 1970-02 .....52

### Chapter IV : Model forecast skill and sensitivity to SST developments

**Fig. IV.1** : EOF1 of the GCM's zonal wind at 925hPa during JAS 1968-2002. GCM is driven with: persisted April, May, June and JAS SSTA .....57

**Fig. IV.2** : Skill of the GCM's zonal wind at 925hPa during JAS 1968-2002. GCM is driven with: persisted April, May, June and JAS SSTA .....59

**Fig. IV.3** : Sahel JAS rainfall skill (1968-02) from direct GCM output and MOS .....60

<b>Fig. IV.4 :</b> Sahel JAS rainfall indices: observed (solid) compared to predictions using the MOS system applied to the GCM driven by , June and JAS SST .....	61
<b>Fig. IV.5 :</b> Composite evolution of anomalies from May to June for (a) Observed SST (b) surface wind response in the GCM and (c) 200hPa wind response in the GCM.....	63
<b>Fig. IV.6 :</b> Time-series of Sahel rainfall (open circles) and the four SST predictors used in empirical forecast models in West Africa .....	67
<b>Fig. IV.7 :</b> Predictions using the empirical predictors in Fig. IV.6 during March, June and JAS SST .....	68
<b>Fig. IV.8 :</b> Cross-validated skill of forecasting the JAS Sahel rainfall index over different sub-periods and for different lead-times.....	69
 <b>Chapter V : Predictability of seasonal Sahel rainfall using coupled models and lead-time improvements</b>	
<b>Fig. V.1 :</b> Same as Fig. III.6 but for the CFS model (with 850hPa wind).....	74
<b>Fig. V.2 :</b> Same as Fig. III.7 but for the CFS model (with 850hPa wind).....	75
<b>Fig. V.3a.b :</b> Time-series of predicted Sahel rainfall from coupled GCMs.....	77
<b>Fig. V.3c:</b> The 850 hPa climatological wind simulated by CFS and DEMETER .....	78
<b>Fig. V.4 :</b> Observed JAS Niño3 correlation with observed JAS SST and wind.....	80
<b>Fig. V.5 :</b> Observed JAS Niño3 correlation with observed June, May and April SST....	81
<b>Fig. V.6 :</b> Obs. JAS Niño3 index correlation with CFS JAS SST .....	84
<b>Fig. V.7 :</b> Same as Fig. V.5 but for CFS near-surface wind predictions.....	85
<b>Fig. V.8 :</b> Same as Fig. V.6, but for CFS initialized from December to March.....	88
<b>Fig. V.9 :</b> Same as Fig. V.7 but for CFS initialized from December to March .....	89
<b>Fig. V.10 :</b> Observed and CFS predicted Sahel rainfall indices. (a) Predictions are from the MOS correction applied to Jan-Feb-Mar initialized forecasts .....	92
<b>Fig. V.11 :</b> Composite difference 1994-2008 MINUS 1981-1993 observed JAS fields...97	
<b>Fig. V.12 :</b> Same as Fig. V.11 but for CFS predictions fields initialized in JFM.....	98
<b>Fig. V.13 :</b> Same as Fig. V.11 but for CFS predictions fields initialized in AMJ.....	99
<b>Fig. V.14 :</b> Time-series of observed JAS Niño3 index (thick solid line), and predicted JAS Niño3 index by CFS (standardized units). .....	101
<b>Fig. V.15 :</b> Time-series of JAS North Atlantic MINUS South Atlantic (AtlN-S) SST : Observed (thick line), predicted by the CFS from JFM and AMJ initializations .....	102

## **Chapter VI : Onset of the Sahelian Rains: A study over Senegal.**

<b>Fig. VI.1 :</b> Mean evolution from South to North of observed mean daily rainfall.....	112
<b>Fig. VI.2 :</b> Defining two homogeneous sub-regions for rainfall onset in Senegal.....	114
<b>Fig. VI.3 :</b> Onset time series from 1950 to 2008 .....	117
<b>Fig. VI.4 :</b> Standardized onset index over northern region of Senegal correlated with tropical SST (left panels) and 925 hPa wind (right panels). .....	126
<b>Fig. VI.5 :</b> same Fig. VI.4 but for the onset index capturing the first significant rainy event recorded over the Southern region of Senegal.....	127
<b>Fig. VI.6 :</b> Sequences of composite fields of anomalies of PWAT, MSE and 925 hPa wind around the first onset event ( $t_0$ ) Northern and Southern Senegal .....	133
<b>Fig. VI.7 :</b> same as fig. VI.6 but $t_0$ represents the date of the first large scale event .....	134
<b>Fig. VI.8 :</b> same as fig. VI.6 but $t_0$ represents the date of the true onset index .....	135
<b>Fig. VI. 9 :</b> Time-longitude variation of OLR, MSE and PWAT anomalies averaged over 10N-20N, from +/- 30 days around the first event onset date .....	137
<b>Fig. VI.10 :</b> Illustrative example of false onset (1972) and a good start (1989). .....	139
<b>Fig. VI.11 :</b> Average number of dry days observed in the 10-days after the first large scale onset event (Left panels) and their frequency distribution (right panels).....	140
<b>Fig. VI.12a :</b> Composites showing systematic difference in sequence between false and true onsets for Northern and southern Senegal.....	142
<b>Fig. VI.12b :</b> as Fig VI.12a but for potential temperature anomalies at 925hPa.....	143
<b>Fig. VI.13 :</b> Correlation between the large scale onset index and the wind shear calculated in each month from April to July. ....	148
<b>Fig. VI.14 :</b> Prediction of the first rainfall event index for the southern region .....	150
<b>Fig. VI.15 :</b> Prediction of the true onset index for the northern region in Senegal.....	152
<b>Fig. VI.16 :</b> Relationship of PWAT with the first event onset index in the south.....	154
<b>Fig. VI.17 :</b> CFS grid boxes in the area over and near Senegal. ....	158
<b>Fig. VI.18 :</b> Mean Annual cycle during MJJASO for observed (black line), CFS raw output (red line) and CFS corrected (blue line). ....	159
<b>Fig. VI.19 :</b> Forecasting onset date with the CFS model. ....	160

## LIST OF TABLES

### Chapter II : Data and Methods

<b>Table II.1</b> : General Circulation Models (GCMs) characteristics .....	18
<b>Table II.2</b> : Schematic of CFS integrations.....	20

### Chapter III : Using regional GCM wind to forecast seasonal rainfall over Sahel

<b>Table III.1</b> : GCM wind MOS skill of predicting Sahel rainfall 1968-2002.....	37
<b>Table III.2</b> : GCM MOS skill of predicting Sahel rainfall 1968-2002 with different variables .....	39
<b>Table III.3</b> : Correlation between observed Sahel Index and GCM EOF U .....	45
<b>Table III.4</b> : Correlation skill of the raw GCM rainfall and the MOS system.....	49

### Chapter V : Predictability of seasonal Sahel rainfall using coupled models and lead-time improvements

<b>Table V.1</b> : CFS skill for JAS Sahel rainfall from AMJ and JFM initializations.....	93
<b>Table V.2</b> : Composite difference of JAS, 1994-2008 MINUS 1981-1993.....	95

### Chapter VI : Onset of the sahelian rains: a study over Senegal

<b>Table VI.1</b> : Mean onset date $\pm$ 1 STD .....	118
<b>Table VI.2a</b> : Cross-Correlation between regional onset indices .....	119
<b>Table VI.2b</b> : Correlation between onset indices and the regional index .....	120
<b>Table VI.3</b> : Predictability of monthly rainfall total .....	121
<b>Table VI.4</b> : Skill of predicting the first rainfall event index in the southern region of Senegal .....	151
<b>Table VI.5</b> : Skill of predicting the first rainfall event index in the southern region of Senegal .....	152

## **Acknowledgements**

There is a saying in West-Africa if you have many people to thank just don't cite names because you might miss some important figures. This list of acknowledgements just touches a few people amongst many whom I want to thank.

I owe a multitude of thanks to my adviser Dr Neil Ward, who first believed in me and then for his inspiration, encouragement, patience and guidance throughout this research.

I also thank my advisory committee members and professors at Columbia University Dr Doug Martinson and Dr Adam Sobel for their availabilities and useful discussions. I extend those thanks to the other members of my thesis defense Dr Peter deMenocal and Dr Paul Roundy. I want also to thank Mrs Mia Leo and her team for being there to answer all questions as well as the ISSO during hard times while getting my papers in order.

My special thanks go to Dr Steve Zebiak and Dr Antonio Moura for supporting my visit in the USA and PhD at Columbia University.

There are many IRI folks whom I also want to thank for their friendships : Rashed, Lareef, Michael Bell, Alessandra, Óli Sveinsson, Pauline Dibi, Madeleine Thompson, Amor, Misrah, Matayo, Yasir, Jim Hansen, Liqiang, Arame, Althea and Maria Salgado.

I profited from the intellectual support of Dr Lisa Goddard who revised my application and other scientific works, Dr Benno Blumenthal through constructive discussion on his wonderful data library interface which I will miss the most. I thank also Dr Simon Mason for hooking me up to CPT as well as some statistical supports, Dr Mike Tippet for matlab, my cousin Dave Dewitt for Grads and some modeling issues and Dr Andy Roberston for useful discussion. Life at IRI was a wealth of experiences for me in terms of paradigm of thinking on how to convey climate information to the climate-user

community. I feel very fortunate to have the opportunity to work at IRI. I would also like to thank IRI administrative, technical and financial support team, including Carolyn, Leo and Ann Binder.

I owe also more than I can express adequately to some people who supported me morally and inspired me, I want to name Mohamed Senouci, Alioune Ndiaye, Mohammed Boulahya, Prof. Peter Lamb, Prof. Mamadou Diouf, Prof Fongang, Prof Ben Mohammed, Moussa Yoro Thiam, Moustapha Ciss, Dr Mansour Thiaw, Dr Muhammad Ahmad Lo, Shakiel Humayun, grand Saliou Lo, Dr Wassila Thiaw, Lucien, Mour Gueye, Issa lele, Mourtada, Jibril Ba, Malanding, Yussuf abakar and Ebrahim ndure.

I received some strong support form Senegalese National Weather Service (ANAMS) in terms of data and guidance and want to thank his DG Mactar Ndiaye, Aïda, Ibou ndiaye, shoua, and Sana.

Many people make my stay at Lamont a very enjoyable experience and I'm very grateful to them especially all those soccer folks, classmates Shala, Louisa, Abby, David Wang, Kandaga, Zahid, David Grass, James Laatsch and our ivy-league driver Fernando,

I deeply thank my wife khadijatou for her forbearances, love, her patience, support and encouragement and my three kids (Muhammad al fatiha, princess Maryam and the little big-trouble maker Abdurrahman) who kept me busy.

Finally my profound gratitude and affection to my best friends in the Harlem community who really keep me from being home sick while being busy on social work : Lamine Agba, Khoule, laye diouf, Lamine Toure, amir Dame Tall, Ahmad Fall, Aboubabcar Cisse, Muhammad Niang, Yacine Diop, Nooh Seydou and all XamSaDine staff member.

## Dedication

*This thesis is dedicated to my elder sister Ngoné Ndiaye who fulfilled mom and dad for me since I was a kid. She has very early on a very clear vision for her little oussou banèe. I owe her a lot. May Allah guide and reward her.*

*Cette thèse est dédiée à ma grande sœur Ngoné Ndiaye qui a joué le rôle de père et de mère pour moi et qui a su tôt avoir un plan très clair pour l'avenir de son oussou banèe, je lui dois beaucoup. Qu'Allah le guide et le paye.*

## **Chapter I : Introduction**

Predicting the climate over the Sahel region in West Africa is crucial due to its impacts on the economy in an environment of scarcity of natural resources. The shortness of the rainy season from about July to September makes these months critical for society, and given the distinct single peak (unimodal) climatology, the onset is also crucial. The main concern of the population, dominated by farmers and shepherds, is the character of the rainy season and its starting date. In many meetings between climate-users and climate forecasters in the Sahel, two questions usually emerge: How much will it be raining this coming season, and when will the rainfall start? Those two questions are at the center of the climate-related concerns of the population and decision makers in the region. Our work is a contribution to addressing these two questions, developing knowledge and improving methods for robust seasonal and onset forecasts that can be made with high scientific confidence and a skill that is as high as possible.

### **I.1 Socio-economical context**

The severe Sahelian droughts in the 1970s stimulated in the global and local community a wake up call to assist and help with drought effects. As a result the AGRHYMET center was created by 9 Sahel countries in 1974 to study and mitigate the impact of drought in the region. The recurring drought has affected natural resources and the population's way of life. It created a series of migrations, from the now poor countryside to the suburb areas of big cities, in search of better jobs. The following droughts in the 80s set in

motion another series of migrations from the Sahel to Europe and America for better living, the process is amplified by the devaluation of the local currency Franc CFA in 1994 (Batterbury and Warren, 2001) with respect to the French currency (FF) which is the CFA guarantee. Two "old" FF became one "new" FF, which over-rated the FF currency thus contributing to increased migration.

Weather and climate forecasts (5-10 days to months ahead) can have high value for Sahel society in varying sectors such as: health, agriculture and hydrology. Incorporating the weather-climate information can increase their efficiency to respond to many calamities (flooding, disease outbreaks, and locust invasions) whose impacts can be disastrous for the African economy and society as a whole.

Agriculture is one of the most important sectors in West Africa due to its underpinning role in the region's food security and the mass of people involved in agricultural activities (farmers and shepherds). West Africa has many dams (especially on the Niger, Senegal and Gambia rivers) which impact many countries, so a meticulous management is needed, at the least, to avoid conflicts. This is related to the demand, amongst neighboring countries, for electricity production, irrigation, and river navigation, each requiring a specific strategy. For example, the management of water resource allocation at the border between Senegal and Mauritania triggered a near-war conflict situation in 1989, with disastrous consequences such as expulsions, border tensions, embargos and refugees on each side of the border (Parker, 1991). A critical decision to make is when the reservoir is getting full (low), do we release (keep) water to avoid going beyond the

maximum (minimum) holding (required) capacity of the dam, in the case that more (less) runoff occurs. These day to day and long term decision making processes can be better managed by using weather and climate information.

Many diseases in the Sahel occur during the rainy season. To fight against those diseases a good approach is to be prepared ahead of time : for example, to move specialists, order and distribute drugs, and to call for international help. Vector born diseases are usually carried out by vectors that need very specific environmental conditions. Those conditions can be identified in real-time by either monitoring the weather conditions or using the weather forecast (rainfall, wind and temperature). A forecast of a long rainy period could make the period prone to an outbreak of many vectors born diseases : for example, malaria, dengue and rift valley fever. Weather forecasts can help to anticipate any probable disease outbreak.

Rapidly growing cities reflect the rural exodus into the metropolis, combined with population growth. There is great demand for more housing and infrastructure, and these are often not well developed. These aspects are contributing to make flooding now a major issue in many big cities in the Sahel, in turn making information about rainfall extremes increasingly important.

All these issues in the context of scarce resources explain the motivation for studying rainfall variability over the Sahel. We will first review the previous studies related to the

predictability of (i) Sahel seasonal rainfall variability and (ii) rainfall onset, before presenting the specific objectives of our study.

## **I.2 Seasonal forecasting of Sahel rainfall**

The rationale behind seasonal forecasting lies in the fact that the atmosphere could be constrained by external forcing. Major sources of forcing on the atmosphere are :

- ✓ Variability of incoming solar radiation (seasonal and decadal)
- ✓ Variability of surface vegetation (year to year and longer term)
- ✓ Variability of sea surface temperature (year to year and longer term)
- ✓ Variability of soil moisture content (seasonal)
- ✓ Variability of chemical compounds in the atmosphere (mostly very long term ie climate change)

Among these forcing agents, the variability of the Sea Surface Temperature (SST) is most used to successfully predict the inter-annual climate variability. Spatial differences in the SST field can induce pressure gradients in the lower boundary of the atmosphere which drive anomalous wind circulation in the near-surface layer. Regions where wind converges (rising air and cloud formation) or diverges (descending air or high pressure region) induce climate anomalies. To illustrate this phenomenon we can refer to the thermal wind relationship :

$$\left. \frac{\partial U}{\partial \ln(P)} = \frac{R}{f} \times \frac{\partial T}{\partial y} \right)_p \quad (Eq. 1.1) \quad [\text{Holton, 2004, 3.30}].$$

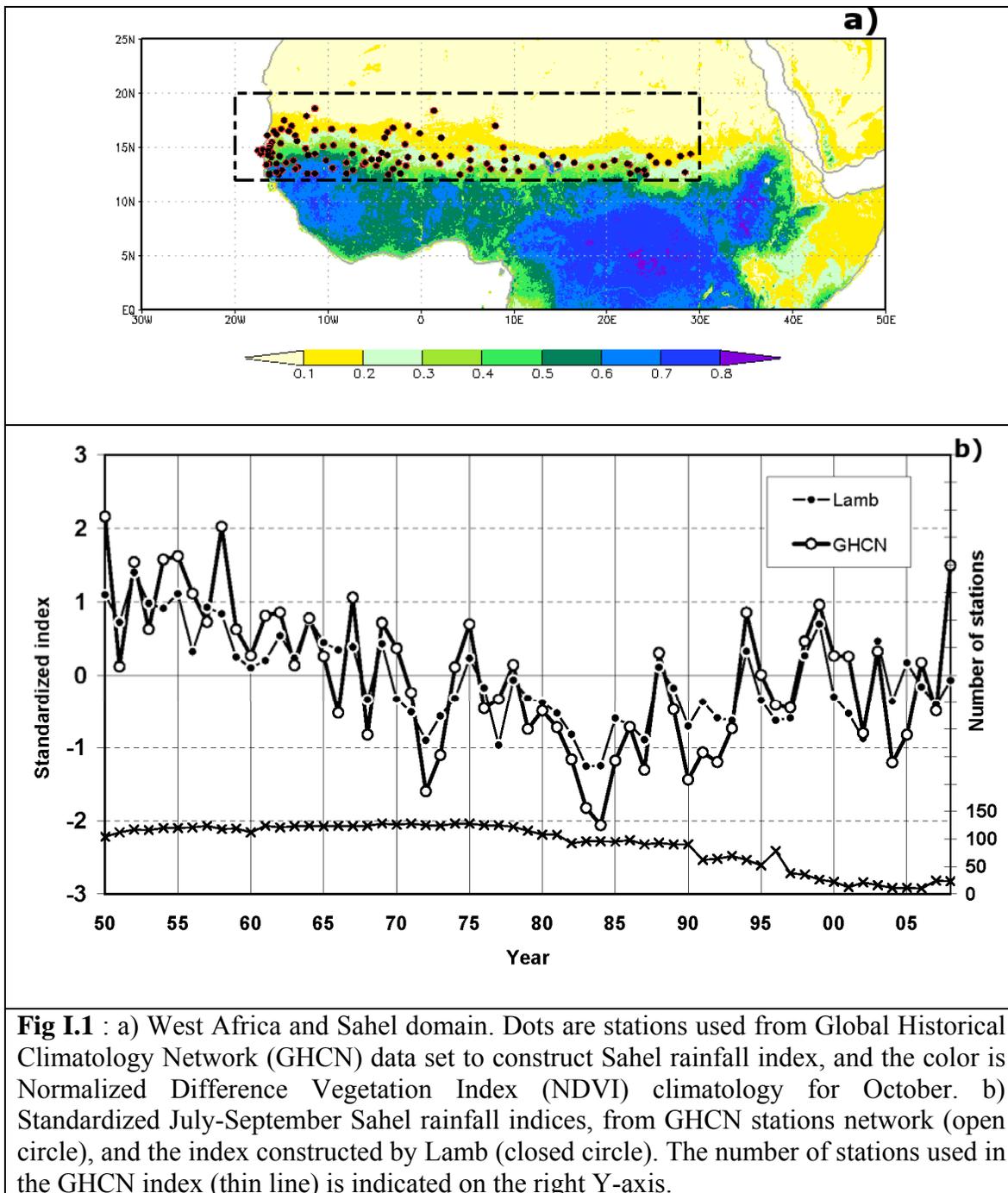
Latitudinal variation of temperature (as can be created by SST variation) at a given pressure level must induce specified wind changes as a function of height. The

dependence on the Coriolis parameter  $f$  means sensitivity to temperature variations is greatest in low latitudes. This is one reason why anomalous SST in the tropical Pacific during ENSO has an impact on the atmospheric general circulation throughout the tropics. Knowing the occurrence of such phenomena well in advance gives us a tool to anticipate fluctuations in the climate. It is the case that large-scale SST anomaly patterns can persist for several months or longer, offering lead time to predict the climate. Further lead-time is gained by explicitly predicting the SST, such as through ENSO prediction (Cane and Zebiak, 1985).

Many approaches are currently used to make seasonal forecasts, reflecting different ways in which the SST information can be related to the predictand (such as rainfall). The basic SST field can be used directly in a statistical model to predict the atmospheric response (statistical models). The observed SST anomaly can be used as the boundary conditions for an Atmospheric General Circulation Model (AGCM), which predicts the dynamical response of the atmosphere (two-tiered forecast system). Finally, coupled ocean-atmosphere GCMs, initialized with prevailing oceanic and atmospheric conditions, project the climate forward in time with full coupling between the ocean and the atmosphere. Both the ocean and atmosphere interact fully in time, so the indirect atmospheric forcing on ocean processes is also taken into consideration. Most statistical models that are used (such as linear regression models) have the advantage that they can be easily used, because they require minimal computing power and are simple to understand and diagnose. However, there is the important problem in statistical modeling of selecting the appropriate predictors for the models, so that they accurately translate the

physical forcing factors, and thus avoid statistical “fishing”. There is the danger of “over-fitting” the statistical model if there are too many predictors. Another danger that can inflate the model skill is if the predictors are related (see Mason and Baddour, 2008). To help cope with these pitfalls in statistical modeling, General Circulation Models (GCMs) can be used, dynamically relating the SST variability to the atmospheric response through the physical linkages in the model. GCMs can be used as a tool to understand the various processes and interactions (e.g. Palmer, 1986), and to evaluate the fraction of variability attributable to teleconnections with SST, therefore estimating the potential for seasonal prediction (such as in Rowell et al., 1995). We will further review and compare the performance of each of these methods over our region of study.

Our study focus is on the Sahel region in tropical West Africa (Fig. I.1a ). The rainy season is marked by strong interannual variability embedded in low frequency multidecadal variability (Nicholson, 1979; Nicholson and Grist, 2001; and see Fig. I.1b).



**Fig I.1** : a) West Africa and Sahel domain. Dots are stations used from Global Historical Climatology Network (GHCN) data set to construct Sahel rainfall index, and the color is Normalized Difference Vegetation Index (NDVI) climatology for October. b) Standardized July-September Sahel rainfall indices, from GHCN stations network (open circle), and the index constructed by Lamb (closed circle). The number of stations used in the GHCN index (thin line) is indicated on the right Y-axis.

Many studies have related Sahel rainfall variability to different components of the near-surface ocean and atmospheric boundary layer. The most investigated part is the SST through its coupling with the atmosphere. The key SST regions identified to influence Sahel rainfall include the tropical Atlantic (Lamb, 1978a; Hastenrath, 1990; Ward, 1998),

the Pacific Ocean El Niño/Southern Oscillation (ENSO) region (e.g. Janicot et al., 1996; Ward, 1998) and, mostly for the decadal and longer timescale, an interhemispheric gradient of SST in the Atlantic (captured by EOF3 of the near global SST in Folland et al., 1991; and used in Sahel linear regression seasonal forecasting), with components of the gradient also present in the Indian and Pacific Oceans leading to its interpretation as a near-global interhemispheric contrast (Folland et al., 1986; Rowell et al., 1995; Ward et al., 1993). Some mechanisms relating the dynamics of Sahel rainfall and SST variability in different basins have been proposed. Anomalous high SST over the tropical South Atlantic basin corresponds to an Intertropical Convergence Zone (ITCZ) anomalously south, with excess rainfall south of 10N in West Africa at the expense of rainfall in the Sahel (Lamb, 1978a, 1978b). With regard to the tropical Pacific, the teleconnection between Sahel rainfall and the ENSO is now well established (Semazzi et al., 1988; Ward, 1998; Janicot et al., 2001). Tropical Pacific warm events (El Niño) are related to reduced rainfall in the Sahel, and cold events (La Niña) are associated with above normal rainfall in Sahel. One proposed mechanism is that warming over the tropical Pacific becomes associated with deep convection and upper level divergence over the Indian Ocean with a descending branch over the Sahel (Shinoda and Kawamura, 1994; Trzaska and Janicot, 2003; Giannini et al., 2003). It is likely that other mechanisms also contribute to the teleconnection between the African sector and the tropical Pacific as well.

These various teleconnection processes in individual years combine in different ways and with different intensities to shape the climate over the Sahel. In addition, on the

interannual timescale, there also appears to be substantial variability internal to the regional climate system and not related to remote teleconnection processes. When driven with observed SST, we would not expect GCMs to accurately capture this internal part of the observed Sahel rainfall variation, but GCMs could be expected to capture the SST-forced component.

However, most GCMs, when forced with the observed SST, continue to have difficulty in simulating the observed variations of seasonal rainfall over the Sahel. This was noted in many early simulations (Sud and Lau, 1996) and is generally noted to still be the case today. This is despite the early realization that SST in all ocean basins could potentially impact Sahel rainfall in a GCM (Palmer, 1986), and good skill demonstrated in one or two models (such as Giannini et al., 2003). The model's analyzed in later chapters further contribute to the discussion on the performance of models over the Sahel.

It is thought that one of the reasons for model failure is that the climate in the Sahel results from complex interaction between SST, regional atmospheric circulations and land-surface interactions which are difficult to model and resolve (Zeng, 2003). One important process that is difficult to represent is recycling of moisture from the soil and vegetation back into the atmosphere. Indeed, vegetation can have a strong influence on rainfall variability through transpiration of moisture, canopy retention and albedo modification. One possible mechanism is that a decrease in rainfall leads to less soil moisture leading to higher surface albedo and reduced evapo-transpiration which has a positive feedback by reducing energy and water flux into the atmospheric column (Zeng

et al., 1999; Xue and Shukla; 1993; Zeng and Neelin; 1999). Because the Sahel lies in a unique geographical zone of sharp gradients of soil moisture and vegetation, its climate is especially sensitive to land-surface variation. Vegetation acts as a low-pass filter on soil moisture by delaying and prolonging its recycling into the atmosphere. Vegetation feedback is thought to be dominantly positive in both wet and dry years, and such amplification of wet and dry periods has been simulated in models (Zeng et al. 1999). Also, recovery from dry years has been argued to be slower than the re-setting of the system after wet years, which might explain the particularly strong persistence of drought over the Sahel (Nicholson, 2000). Another impact of dryness is enhancement of dust production. Dust has a direct effect on radiation (scattering and absorption) and indirect effect through its interaction with clouds. Prospero and Nees (1986) have found atmospheric dust content at the beginning of one rainy season is strongly correlated with rainfall of the previous year which could convey the interannual memory.

Many modeling exercises have been conducted to show the impact of land surface process on Sahelian rainfall. Xue (1997) induced a reduction of precipitation, soil moisture and runoff during the July-September season by replacing savanna and shrub vegetation over the Sahel with bare soil. Zeng et al. (1999) performed a series of model simulations to identify the relative importance of SST forcing, land surface and vegetation by switching on and off different aspects of land and vegetation interaction. They found the decadal variation was reproduced better with inclusion of interactive soil moisture ( $r=0.44$ ) and especially interactive vegetation ( $r=0.67$ ). Field studies, such as during HAPEX-Sahel (Taylor and Lebel, 1998) and recently the AMMA campaign, have shown strong evidence of a link between surface processes and rainfall in the Sahel. This

is leading to better understanding of the processes, which can lead to more realistic representation in GCMs. The importance of the land surface feedback is important to have in mind in any study of the region's climate, even when (like in this thesis) the focus is on the primary forcing from the ocean and the associated predictability.

### **I.3 Rainfall Onset**

The rainfall distribution in the Sahelian region of West Africa is characterized by a single peak in the annual cycle. The rains typically start in early May (at least for the southern parts of the Sahel), reach a maximum in August and fall off in October following the migration of the ITCZ from the Equator to the North and returning back southward toward the Equator. The bulk of the rain is recorded in July-August-September. The shortness of the rainy season makes the understanding of its start relevant and crucial for any activities related to seasonal rainfall, especially crop production and pastoralism. The Sahel is one of the poorest regions in the world and its economy is dependent on subsistence agriculture for its survival (Sivakumar, 1988). In Senegal, agriculture activities mobilize almost 70% of the active population, and have a strong contribution to the GDP. The population growth rate is one of the highest, so managing resources in agriculture is expected to be a key component for the future success of the country.

In Senegal, the start of the rainy season guides when to plant crops, in a rain fed agriculture subjected to strong year to year rainfall total variations, as well as variations in the length of the rainy season. Especially because of the highly variable rainfall distribution within the season, agriculture planning should be carefully framed according to the start of the rainy season. Planting too early may lead to crop failure, whereas planting too late may reduce the growing season and hence crop yield (Dood and Jolliffe, 2001). Sivakumar (1988) showed that an early onset of rains offers a longer growing season while delayed onset results in a considerable shorter growing season. Many studies have related early establishment of crops with higher yields (Kassam and Andrews, 1975; Sivakumar, 1990). For example, studies have shown for sites near Niamey, that in years with an early onset of rainfall, dry matter production of millet reached 7 t/ha while late onset of rains resulted in only 3 t/ha (Sivakumar, 1990). The longer growing season enabled by early onset is related to the fact that the length of the rainy season is highly correlated with the onset of the rainy season. Studies of agricultural users have revealed that the information of most interest to the user is not the rainfall total but the start of the rainy season (Ingram et al., 2003; Ziervogel and Calder, 2003).

All these results show a dire need to provide better information about the rainfall onset, and especially to provide a forecast of the start of the rainy season. This would be of great value for better preparedness, including selecting the appropriate seed to plant, farmland preparation and manpower return from cities to the country side where agriculture is mostly practiced. There is limited knowledge on the nature and predictability of onset. Coupling of atmospheric conditions with SST during build up to the rainy season has

largely not been documented. A few studies have begun to document the atmospheric conditions around the onset over the Sahel. For example, Sultan and Janicot (2003) define a ‘pre-onset’ by the arrival first of the inter-tropical discontinuity (ITD) at 15°N. An initial approach to forecasting onset attempted to capture anomalies in the evolution of the atmosphere’s annual cycle in the weeks leading up to onset (Omotosho, 1992). These studies have not distinguished characteristics of onset for northern versus southern parts of the Sahel. Much therefore remains to be learned about rainfall onset in the Sahel, and given its importance for agriculture, it is a key topic to investigate.

#### **I.4 Objectives of this study**

The overall objective of this study is to develop a reliable and useable seasonal forecasting system and an early warning system for its onset which can lead to effective response strategies. This study will be focused into four objective areas, addressed in Chapters III, IV, V and VI respectively.

In chapter III we develop the basis of using tropical Atlantic wind to correct poor GCM forecasts of JAS Sahel rainfall. Such a Model Output Statistics (MOS) approach is developed first for the ECHAM4.5 GCM, and then the analysis is extended to seven other models, allowing insights into the performance of models generally over the Sahel and tropical Atlantic. In chapter IV we focus on the lead-time limitations for Sahel rainfall prediction due to SST developments in boreal spring. The limitation is shown for a GCM using persisted SST, and for empirical models using established SST predictors.

Chapter V addresses the potential for overcoming these limitations and gaining lead time improvements by using coupled models. Performance of coupled models over the Sahel is studied, and the ability of the same MOS approach as developed in Chapter III is assessed. Experiments from one model (the NCEP CFS) are available up to 2008, which leads the analysis to include the most recent fluctuations in Sahel rainfall, which contain a modest recovery from the strongest drought years of the 1980s.

The onset of Sahel rainfall is addressed in chapter VI. We seek a better understanding of the onset of the rainfall, and explore its predictability, using high resolution station data over Senegal, and various observational and model datasets.

Regional definitions of true and false start are developed, regional atmospheric sequences heralding onset are identified, and distinctive features of true versus false starts are identified in the atmospheric sequences. In addition, the extent to which the onset timing is influenced by ocean-atmosphere coupling is diagnosed. Implications for predictability are drawn from the various possible avenues (slowly-propagating atmospheric features, annual cycle anomalies, SST, coupled model predictions), and some experimental prediction models are presented. In all cases, distinctive results are produced for the northern and southern part of Senegal. Onset in the southern parts is typically about one month earlier, and it is therefore reasonable to expect differing characteristics given that the background state has evolved substantially over that month. Before beginning the research results, Chapter II presents the data used in the study and some of the key methods deployed.

## Chapter II : Data and Methods

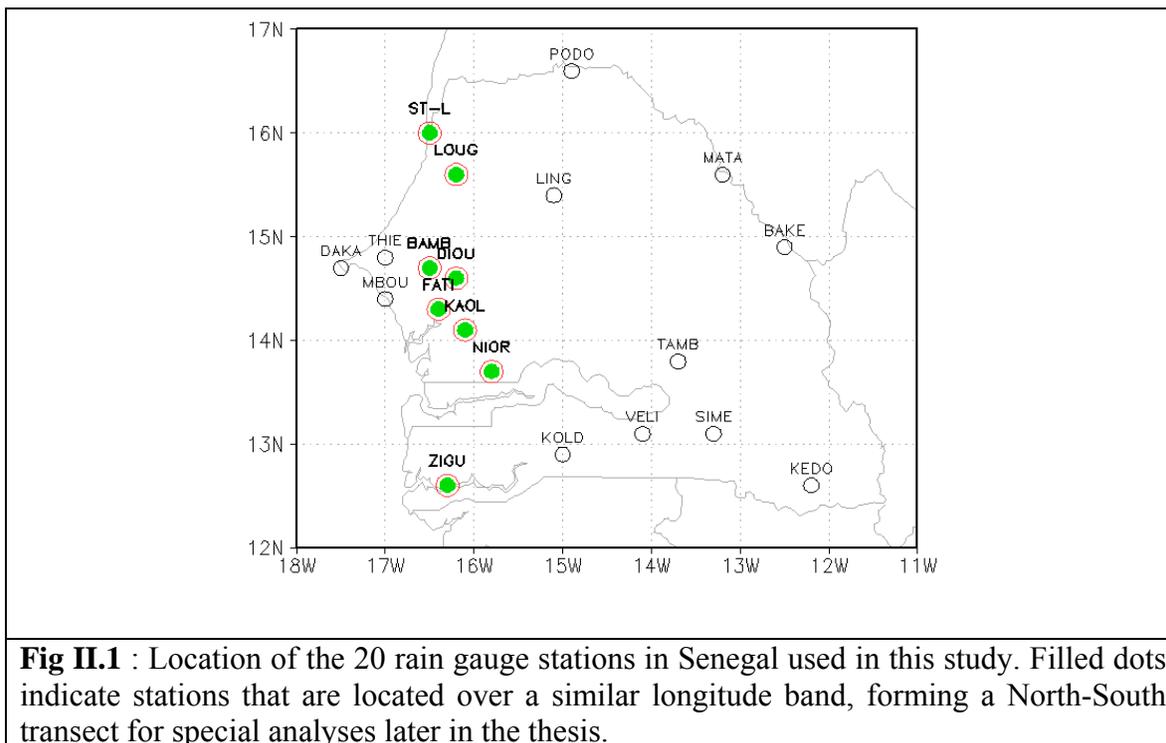
### II.1 Data

#### II.1.a Rainfall data and basic indices

Rainfall data are taken from the Global Historical Climatology Network (GHCN). A set of 131 stations within the domain bounded by 18°W-30°E and 12°N-20°N are used to construct a Sahel rainfall index. The index for July-September is calculated by averaging the standardized seasonal rainfall anomaly at each station (hereafter referred to as the GHCN series). Particularly since the number of stations available to construct this index decreases in recent years (Fig. I.1b), results are compared with a widely used July-September Sahel index (hereafter referred to as the Lamb index, as described in Bell and Lamb, 2006; with updated series to 2008 provided through Lamb, personal communication). The series are shown in Fig. I.1b for the years 1950-2008, which is the maximum period over which GCM runs are analyzed in this paper. The two time-series are very well correlated over 1968-2001 ( $r=0.96$ ), and discrepancies are mostly confined to 2005 to 2008, with only 2006 having a discrepancy of substance. The correlation over the whole period is  $r=0.88$ . The GHCN time-series is used as the primary target predictand, but results are cross-checked with the Lamb index and tests undertaken to check for sensitivity to inclusion of the less certain last few years. Overall, the series show good consistency, both in terms of interannual variability, and in terms of the modest recovery of rains since the peak of the extended drought in the early 1980s (Nicholson, 2005; Giannini et al., 2008). Most analyses in chapters III to V are confined to the use of these observed Sahel rainfall indices. However, for the global precipitation composite in chapter V, data are taken from the Climate Prediction Center (CPC) Merged

Analysis of Precipitation (CMAP) (Xie and Arkin, 1997) without any model input is used as a proxy for rainfall observations. The CMAP rainfall is on a  $2.5^\circ \times 2.5^\circ$  grid and is available starting from January 1979.

In chapter VI we used daily data to calculate the onset date. The station rain gauge data come from the Agence Nationale de la Météorologie Sénégalaise (ANAMS, National Meteorological Service). It covers 20 stations of daily data from 1950 to 2008 from May to October. Figure II.1 shows the spatial distribution of the stations. Most of the stations are located in the Western fringe of the country where the major cities are found. The daily data were subjected to quality control by cross-checking with the ANAMS and station officers for any dubious values. Stations having more than 15% of missing data over the period 1950-2008 are disregarded.



### **II.1.b Sea Surface Temperature (SST)**

SST data are taken from the NOAA/CDC dataset (Smith and Reynolds, 2003). The SST data are a blended (satellite and in situ observations) analysis, with a  $2^\circ$  lat  $\times$   $2^\circ$  long resolution. Recent improvements are made in the bias correction of the satellite data starting in 1985 (ERSST.v3).

### **II.1.c NCEP Reanalysis data**

Atmospheric data were extracted from archives of the National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR) reanalysis project (Kalnay et al., 1996). This data set consists of a dynamical model interpolation of global meteorological observed variables at  $2.5 \times 2.5$  grid resolution from 1950 to 2008. Daily data are used to characterize weather patterns around the onset of the rainy season. Monthly mean data are averaged over the JAS season. The fields of interest are precipitable water (PWAT) and wind and specific humidity at 925, 850 and 200 hPa.

### **II.1.d Outgoing Long wave Radiation (OLR)**

The OLR data (Liebman and Smith, 1996), which will be used as a proxy for tropical convection, are from the Advanced Very High Resolution Radiometer (AVHRR) of NOAA operational polar orbiting satellites. The OLR data of 1978 are missing in the record due to technical problem with the satellite.

### II.1.e Atmospheric and Coupled GCM simulations

Multi-year (mostly over 20 years) simulations from eight different atmospheric GCMs and eight different coupled ocean-atmosphere GCMs are analyzed in this study. The characteristics of the eight different atmospheric GCMs are summarized in Table II.1. All these atmospheric GCMs were forced with observed July-August-September SST.

**Table II.1** : General Circulation Models (GCMs) characteristics used in the wind MOS approach.

	<b>*EC4.5</b>	<b>GFDL</b>	<b>COLA</b>	<b>ECPC</b>	<b>*EC5</b>	<b>*EC5</b>	<b>NSIPP</b>	<b>CCM3</b>
<b>Runs Provider</b>	IRI	NASA	COLA	SCRIPPS	IRI	IRI	NASA	NCAR
<b>Members</b>	24	10	10	12	16	24	9	24
<b>Spatial resolution</b>	T42 ~2.81°	2.5°x2 °	T63 ~1.87°	2.5°x2.5°	T85 ~1.41°	T42 2.81°	2.5°x2°	T42 ~2.8°

\*EC stands for ECHAM

The atmospheric General Circulation Model used extensively in this study is the ECHAM4.5, for which ensemble simulations of 24 members and retrospective forecasts of 12 members have been produced at the International Research Institute for climate and society (IRI). The ensemble mean is used in this study. Simulation ensemble mean results with 12 members were also analyzed, and no difference of substance was found compared to the results with the 24 members that are shown in the analyses in this thesis. ECHAM4.5 is a spectral model at T42 resolution, approximately 2.8 degrees horizontal resolution, with 18 levels in the vertical. More details of the model can be found in Roeckner et al. (1996). Two sets of ensemble runs are used in this study as in Goddard

and Mason (2002). “Simulation” runs were generated by forcing the ECHAM4.5 with observed simultaneous monthly-mean SSTs. “Persisted” runs refer here to forecasts by the GCM using SST anomaly persistence to generate the SST fields to drive the GCM. The SST persistence is designed as follows: for each month the SST anomaly (SSTA) is added to each of the following 5 months SST climatology, which are then used to force the atmospheric GCM. The simulation runs are available from 1950 to present. The persisted runs available to this study cover the period 1968-2002. Seven other AGCMs forced with observed SST are also studied (the models are described in Table II.1).

The DEMETER project (Development of a European Multimodel Ensemble system for seasonal to inTERannual prediction) provides seasonal hindcasts using 7 global coupled ocean-atmosphere models with different atmospheric and oceanic components (Palmer et al., 2004). The set of integrations that allow evaluation of July-September Sahel rainfall forecasts are initialized around May 1<sup>st</sup> each year, forecasting the period from May to October during each year. Thus only one lead-time on the JAS Sahel rainfall season can be investigated. The years for which forecasts are available are of varying length, with all ending in 2001. Details of the models are documented in Palmer et al. (2004).

A set of coupled model seasonal hindcasts are also analyzed from the NCEP Climate Forecast System (CFS) (Saha et al., 2006). The atmospheric component is the NCEP-GFS model (Moorthi et al., 2001) of T62 spatial resolution, and the ocean component is the GFDL Modular Ocean Model version 3 (MOM3) (Pacanowski and Griffies, 1998) with a zonal resolution of 1 degree and a meridional resolution that varies from 1/3

degree near the Equator to 1 degree near the poles. The runs cover 28 years from 1981 to 2008 and each integration corresponds to a 9 months lead-time forecast. For each month, the CFS run is initialized at 00Z and 12Z in 3 different periods of 5 consecutive days each separated by 10 days (Table II.2). Monthly data are created by averaging the 15 daily initializations during each month. Results on the model's ability to represent ENSO have been documented (Wu et al., 2009). Analysis here represents a test of the ability to represent information that is related to Sahel rainfall in the July-September season. The model's climatology in the Sahel has been studied (Thiaw and Mo, 2005), noting a dry bias in the model's rainfall that is corrected for here using standardized units throughout, and more substantially, by employing the MOS approach with the tropical Atlantic winds. A MOS approach offered some promise for CFS forecasts of Sahel rainfall at very short (1-month) lead-time (Mo and Thiaw, 2002), so results here extend those findings to the longer lead-times.

**Table II.2 :** Schematic of CFS initial atmospheric (IACs) and ocean (OBCs) conditions. Each month contains 15 members divided into 3 groups, each of 5 consecutive days (pentad). Each pentad is initialized with the same OBCs taken at the middle day of the pentad (11<sup>th</sup>, 21<sup>st</sup> and 1<sup>st</sup> of the month respectively) using two IACs each day at 00Z and 12Z. L denotes last day of the month.

OBCs	Pentad of the 11 <sup>th</sup>					Pentad of the 21 <sup>th</sup>					1 <sup>st</sup> of the Pentad				
IACs	9	10	11	12	13	19	20	21	22	23	L-1	L	1	2	3
<b>Member</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>
Dec	9	10	11	12	13	19	20	21	22	23	29	30	1	2	3
Jan	9	10	11	12	13	19	20	21	22	23	30	31	1	2	3
Feb	9	10	11	12	13	19	20	21	22	23	30	31	1	2	3
Mar	9	10	11	12	13	19	20	21	22	23	28	29	1	2	3
Apr	9	10	11	12	13	19	20	21	22	23	30	31	1	2	3
May	9	10	11	12	13	19	20	21	22	23	29	30	1	2	3
Jun	9	10	11	12	13	19	20	21	22	23	30	31	1	2	3

## II.2 Methods

In this section we describe only the more general methods applied through the thesis. More specific methods will be present within respective chapters.

### II.2.a Empirical Orthogonal Function

We extensively use near-global gridded fields (GCMs, SST, Re-analysis) as predictors. Contiguous grid boxes are highly correlated to each other. Therefore, to reduce the number of predictors, we pre-filter each field by applying an Empirical Orthogonal Function (EOF) analysis, which retrieves regularly coherent spatial patterns through time. The EOF patterns are constructed to be orthogonal to each other so each one captures a unique part of the variance of the initial fields. The EOFs are an orthogonal basis for the raw fields. EOFs are classified by the amount of variance explained from the raw field. The first EOF (EOF1) captures more variance of the fields than does the second, and so on so forth. Each EOF pattern or shape is contained in each raw field with different weights translated in the expansion coefficients' time series. There are slightly differing formulations of EOF analysis, and a widely used one is Principal Component Analysis (PCA). It is extensively used in climate data analysis (Pearson, 1902; Lorenz, 1956; Bretherton, 1992), and has been widely applied in empirical and MOS approaches to seasonal prediction (e.g. Ward et al., 1993; Feddersen et al., 1999).

The EOF or PCA analysis proceeds by finding eigen values and eigen vectors. Let  $X(m,n)$  be a data field varying in time ( $m$ ) and space ( $n$ ). Time refers here to the season of interest (JAS) during each year and space to the number of grid boxes. To obtain the eigen values and eigen vectors we start here with anomalies by removing the temporal

mean. From the anomaly vector we calculate the variance-covariance matrix  $M=(1/n)*X^T.X$ , then obtain the eigen vectors( $V^T$ ) and eigen values ( $\Sigma$ ) by diagonalizing  $M$  as  $X^T.X= V\Sigma V^T$  and then get  $U = XV \Sigma^{-1}$ .

In this thesis we used a Singular Value Decomposition (SVD) approach to find the EOFs. The advantage of using SVD is that it is computationally more efficient than the covariance approach and avoids dealing with matrix inversion. We just apply a SVD on  $X$ , the SVD transformation is as follow  $X^T=USV^T$ . The singular values ( $S$ ) are the normalized square root of the eigen values ( $\Sigma$ ) :  $SS^T=\Sigma$ . The eigen values are proportional to the variance explained by each EOF. This transformation can be achieved by using widely available numerical recipes libraries (LAPACK routine) through a standard compiler like matlab (SVDA function) or using the Climate Predictability Tool<sup>1</sup> (CPT) package.

From the above,  $U(m,n)$  is an orthogonal matrix of loading patterns for the EOFs, with each column giving a spatial representation of an EOF.  $S(n,n)$  is the diagonal matrix of singular values which are nonnegative real numbers, usually listed in decreasing order.  $V(n,n)$  is an orthogonal matrix of EOF coefficients. Each column gives the temporal weight of the EOFs for each given year. Both  $U$  and  $V$  are orthogonal matrices in other words  $U^T U=I$  and  $V^T V=I$ .

---

<sup>1</sup> CPT : <http://iri.columbia.edu/outreach/software/index.html>

## **II.2.b Statistical measures of association and statistical significance**

### **Pearson correlation coefficient**

To quantify the strength of the relationship between time-series we use the Pearson correlation coefficient. Correlation is a good indicator of the relationship between two variables. When a base index is correlated with each time-series in a gridded field, the map can be an effective description of teleconnection patterns.

Statistical significance of the correlation coefficient is assessed using the standard t-value approach. Caution is needed in interpreting statistical significance in the presence of auto-correlation due to the loss of degrees of freedom. In this study, exploratory correlation analyses do not span the period of Sahel rainfall that contains the strong downward trend of rainfall (that gives rise to strong autocorrelation). Rather, analyses focus on 1968-2008, and when including the earlier period, a separate analysis is undertaken for 1950-1970, so that the periods of analysis don't present any substantial autocorrelation .

### **Composite Analysis**

We used in this study composite fields to see systematic differences between two sub-sets of years. Sometimes a specific sub-set of years is compared to all years (in which case, all years form the second sub-set). On other occasions, two specific sub-sets of years are compared. To test for any significant difference between the two subsets, we apply a Mann-Whitney U-test (Mann and Whitney, 1947; Castello and Shelton, 2004; and Kahya and Dracup, 1994) which is a nonparametric test alternative to the t-test. Mann-Whitney

U-test is sometimes also called the Mann-Whitney-Wilcoxon test. The null hypothesis can be set that the two-sub sets are not significantly different (ie equal) or drawn from the same population. The test is performed by combining the two subsets and then ranking them from the smallest to the highest, and calculating the U statistic :

$$U_1 = R_1 - \frac{m(m+1)}{2} \quad \dots \text{where} \quad R_1 = \sum_{i=1}^n r_i \quad (\text{Eq 2.1});$$

$r_i$  is the rank of each year

which has a mean  $\mu_U$  and a standard deviation  $\sigma_U$

$$\mu_U = \frac{mn}{2}; \quad \sigma_U = \left( \frac{mn(m+n+1)}{12} \right)^{1/2}$$

Where m and n are respectively the size of the two sub-sets, and  $R_1$  is the sum of the ranks assigned to the composite. Very large or small values of U is indicative of the significance of (the years in) the composite maps.

The Mann-Whitney U-test uses the rank instead of the actual values, so is less affected by outliers, and since it's a nonparametric test, it doesn't require assumptions on the distribution of the data set. This is a valuable property, as the number of years in the composite is usually very small and difficult to be fitted with any distribution.

### **II.2.c Statistical Prediction Methodology**

For both model output statistics (MOS) and prediction methods using empirical SST predictors, the routines from the Climate Predictability Tool (CPT) are applied. For MOS, a predictor variable and domain is first selected (such as the tropical Atlantic zonal wind field at 925hPa). The predictor variable is decomposed into EOFs. These EOFs are

incrementally added into a multiple linear regression, first using EOF1 alone, then EOF1 plus EOF2, and so on. For each model, a cross-validated set of predictions are generated, withholding the target year plus one year before and after (i.e. “leave-three-out” cross validation). The skill of the cross-validated prediction is consulted to assess the best model. For the empirical SST predictors used in Chapter III, a comparable approach is taken with multiple linear regression models evaluated applying leave-three-out cross validation. When a model has no skill, it is the case that cross-validation can lead to the forecast and observed series having an artificially induced large negative correlation. Therefore, all cross-validated skills of less than zero are displayed as zero.

#### **II.2.d Model validation**

The main goal of this study in terms of prediction, is to establish the existence of predictability in the various approaches and problems studied, and to provide indication of the magnitude of the predictability that is found. A well-known and generally robust indicator of the degree of association between a set of predictions and a set of observations is the correlation coefficient. This score is therefore used as the basic indicator of skill, both for direct GCM output, for MOS predictions, and for empirical forecast models. Evaluations are also possible that are closer to user requirements for operational utility, including assessments in terms of the reliability of probabilistic forecasts (Mason and Chidzambwa, 2009). Such approaches, many of which are available in CPT, are not presented here, since our goals are mostly the research of basic prediction signals.

## Reference Chapters I and II

- Batterbury S., and A. Warren, 2001: The African Sahel 25 years after the great drought: assessing progress and moving towards new agendas and approaches. *Global Environmental Change*, **11**(1), 1-8.
- Bell M. A., and P. J. Lamb, 2006: Integration of weather system variability to multidecadal regional Climate Change: The West African Sudan-Sahel Zone 1951-98. *J. Climate*, **19**, 5343-5365.
- Bretherton C. S., A. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate*, **5**, 541-560.
- Cane M. A., and S. E. Zebiak, 1985: A theory for El Niño and the Southern Oscillation. *Science*, **228**, 1085-1087.
- Castello A. F., and M. L. Shelton, 2004: Winter precipitation on the US Pacific coast and El Niño-Southern Oscillation events. *Int. J. Climatol.*, **24**, 481-497.
- Dodd D. E. S., and I. T. Jolliffe, 2001: Early detection of the start of the wet season in semiarid tropical climates of Western Africa. *Int. J. Climatol.*, **21**(10), 1251-1262.
- Feddersen H., A. Navarra, and M. N. Ward, 1999: Reduction of model systematic error by statistical correction for dynamical seasonal prediction. *J. Climate*, **12**, 1974-1989.
- Folland C. K., J. A. Owen, M. N. Ward, and A. W. Colman, 1991: Prediction of seasonal rainfall in the Sahel region of Africa using empirical and dynamical methods. *J. Fore.*, **10**, 21-56.
- Folland C. K., T. N. Palmer, and D. E. Parker, 1986: Sahel rainfall and world-wide sea temperatures, 1901-85. *Nature*, **320**, 602-607.
- Giannini A., R. Saravanan, and P. Chang, 2003: Oceanic Forcing of Sahel rainfall on interannual to interdecadal time scales. *Science*, **302**, 1027-1030.
- Giannini A., M. Biasutti and M. M. Verstraete, 2008: A climate model-based review of drought in the Sahel: Desertification, the re-greening and climate change. *Global Planetary Change*, **64**, 119-128.
- Goddard L., and S. J. Mason, 2002: Sensitivity of seasonal climate forecasts to persisted SST anomalies. *Climate Dyn.*, **19**, 619-631.
- Hastenrath S., 1990 : Decadal-scale changes of circulation in the tropical Atlantic sector associated with Sahel drought. *Int. J. Climatol.*, **10**, 459-472.
- Holton J. R., 2004: An Introduction to Dynamic Meteorology, 4th Edition, ISBN 0123340151.
- Ingram K. T., M. C. Roncoli, and P. H. Kirshen, 2003: Opportunities and constraints for farmers of west Africa to use seasonal precipitation forecasts with Burkina Faso as a case study. *Agricultural Systems*, **74**(3), 331-349.
- Janicot S., V. Moron, and B. Fontaine, 1996: Sahel droughts and ENSO dynamics. *Geophys. Res. Lett.*, **23**, 515-518.
- Janicot S., S. Trzaska and I. Pocard, 2001: Summer Sahel-ENSO teleconnection and decadal time scale SST variations. *Climate Dyn.*, **18**, 303-320.
- Kahya E., and J. A. Dracup, 1994: The influences of Type 1 El Niño and La Niña events on streamflows in the Pacific southwest of the United States. *J. Climate*, **7**, 965-976.
- Kalnay E., et al., 1996: The NCEP-NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.

- Kassam A. H., and D. J. Andrews, 1975: Effects of sowing date on growth, development and yield of photosensitive sorghum at Samaru, northern Nigeria. *Exp. Agric.*, **11**, 227-240.
- Lamb P. J., 1978a: Case study of Tropical Atlantic surface circulation patterns during recent sub-saharan weather anomalies : 1967 and 1968. *Mon. Wea. Rev.*, **106**, 482-491.
- Lamb P. J., 1978b: Large-scale Tropical Atlantic circulation patterns associated with Subsaharan weather anomalies. *Tellus*, **A30**, 240-251.
- Liebman B., and C. A. Smith, 1996: Description of a complete (interpolated) outgoing longwave radiation dataset, *Bull Am Meteor Soc*, **77**, 1275-1277.
- Liu, C., and E. J.
- Lorenz E. N., 1956: Empirical orthogonal functions and statistical weather prediction. Technical report, Statistical Forecast Project Report 1. Department of Meteorology, MIT: 49.
- Mann H. B., and D. R. Whitney, 1947: On a test on whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.*, **18**, 50–60.
- Mason S. J., and S. Chidzambwa, 2009: Verification of RCOF Forecasts. *WMO RCOF Review 2008 Position Paper*, 26 pp.
- Mason S. J., and O. Baddour, 2008: Statistical modeling. In Troccoli A., M. S. J. Harrison, D. L. T. Anderson, and S. J. Mason (Eds), *Seasonal Climate Variability: Forecasting and Managing Risk*, NATO Science Series, Springer Academic Publishers, 467 pp.
- Mo K. C., and W. M. Thiaw, 2002: Ensemble canonical correlation prediction of precipitation over the Sahel. *Geophys. Res. Lett.*, **29**, DOI: 10.1029/2002GL015075
- Moorthi S., H.-L. Pan, and P. Caplan, 2001: Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast system. *NWS Tech. Procedures Bulletin*, **484**, 14 pp.
- Nicholson S. E., 2000: Land Surface Processes and Sahel Climate. *Rev. Geophys.*, **38**(1), 117-139.
- Nicholson S. E., 2005: On the question of the “recovery” of the rains in the West African Sahel. *Journal of Arid Environments*, **63**, 615-641.
- Nicholson S. E., and J. P. Grist, 2001: A conceptual model for understanding rainfall variability in the West African Sahel on interannual and interdecadal timescales. *Int. J. Climatol.*, **21**, 1733-1757.
- Nicholson S. E., 1979: Revised rainfall series for the West African subtropics. *Mon. Wea. Rev.*, **107**, 620–623.
- Omotosho J. B., 1992: Long-range prediction of the onset and end of the rainy season in the West African Sahel. *Int. J. Climatol.*, **12**, 369-382.
- Pacanowski R. C., and S. M. Griffies, cited 1998: MOM 3.0 manual. NOAA/GFDL. [Available online at [http://www.gfdl.noaa.gov/\\_smg/MOM/web/guide\\_parent/guide\\_parent.html](http://www.gfdl.noaa.gov/_smg/MOM/web/guide_parent/guide_parent.html).]
- Palmer T. N., et al., 2004: Development of a European multimodel ensemble system for seasonal-to-interannual prediction DEMETER. *Bull. Am. Meteor. Soc.*, **85**, 853–872.
- Palmer T. N., 1986: Influence of the Atlantic, Pacific and Indian Oceans on Sahel rainfall. *Nature*, **322**, 251-253.

- Parker R., 1991: The Senegal Mauritania conflict of 1989 - A Fragile Equilibrium. *Journal Of Modern African Studies*, **29**(1), 155-171.
- Pearson K., 1902: On lines and plans of closest fit to system of points in space. *Philos. Mag.*, **6**, 559-572.
- Prospero J. M., and R. T. Nees, 1986: Impact of the North African drought and El Niño on mineral dust in the Barbados trade wind. *Nature*, **320**, 735-738.
- Roeckner E., et al., 1996: The atmospheric general circulation model ECHAM4: Model description and simulation of present day climate. Rep. 218, Max-Planck-Institut für Meteorologie, 90 pp.
- Rowell D. P., C. K. Folland, K. Maskell, and M. N. Ward, 1995: Variability of summer rainfall over tropical North Africa (1906-92): Observations and modeling. *Quart. J. Roy. Meteor. Soc.*, **121**, 669-704.
- Saha S., et al., 2006: The NCEP Climate Forecast. *J. Climate*, **19**, 3483-3517.
- Semazzi F. H. M., V. Mehta, and Y. C. Sud, 1988: An investigation of the relationship between sub-Saharan rainfall and global sea surface temperatures. *Atmos. Ocean*, **26**, 118-138.
- Shinoda M., and R. Kawamura, 1994: Tropical rainbelt, circulation and sea surface temperatures associated with the Sahelian rainfall trend. *J. Meteor. Soc. Japan*, **72**, 341-357.
- Sivakumar M. V. K., 1988: Predicting rainy season potential from the onset of rains in Southern Sahelian and Sudanian climate zones of West Africa. *Agric. For. Meteorol.* **42**, 295-305.
- Sivakumar M. V. K., 1990: Exploiting rainy season potential from the onset of rains in the Sahelian zone of West Africa. *Agric. For. Meteorol.*, **51**, 321-332.
- Sivakumar M. V. K., (ed.). 1998 : Climate Variability Prediction, Water Resources and Agricultural Productivity: Food Security Issues in Tropical Sub-Saharan Africa. Report of a Joint START/WCRP/SCOWAR Workshop, 22-25 July 1997, Cotonou, Benin. Washington DC, USA: International START Secretariat.
- Smith T. M., and R. W. Reynolds, 2003: Extended Reconstruction of Global Sea Surface Temperatures Based on COADS Data (1854-1997). *J. Climate*, **16**, 1495–1510.
- Sud Y. C., and W. K-M. Lau, 1996: Comments on paper “Variability of summer rainfall over tropical North Africa (1906-1992): observations and modelling.” *Quart. J. Roy. Meteor. Soc.*, **122**, 1001-1006.
- Sultan B., and S. Janicot, 2003: The West African Monsoon Dynamics. Part II: The “Preonset” and “Onset” of the Summer Monsoon. *J. Climate*, **16**, 3407-3427.
- Taylor C. M. and T. Lebel, 1998: Observational evidence of persistent convective-scale rainfall patterns. *Mon. Wea. Rev.*, **126**, 1597-1607.
- Thiaw W. M. and K. C. Mo, 2005: Impacts of sea surface temperature and soil moisture on the Sahel seasonal rainfall predictions. *J. Climate*, **18**, 5330-5343.
- Trzaska S., and S. Janicot, 2003: The role of the Indian Ocean in modulation of ENSO impacts on the West African Monsoon. *CLIVAR exchanges*, **8**, 17-18.
- Ward M. N., C. K. Folland, K. Maskell, A. W. Colman, D. P. Rowell, and K. B. Lane, 1993: Experimental seasonal forecasting of tropical rainfall at the U.K. Meteorological Office. *Prediction of Interannual Climate Variations*, J. Shukla, Ed., Springer–Verlag, 197-216.

- Ward M. N., 1998: Diagnosis and short lead time prediction of summer rainfall in tropical North Africa at interannual and multidecadal timescale. *J. Climate*, **11**, 3167-3191.
- Wu R., B. P. Kirtman, and H. Van Den Dool, 2009: An Analysis of ENSO Prediction Skill in the CFS Retrospective Forecasts. *J. Climate*, **22**, 1801-1818.
- Xie P., and P. Arkin, 1997: Global Precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539-2558.
- Xue Y. K., 1997: Biosphere feedback on regional climate in tropical North Africa. *Quart. J. Roy. Meteor. Soc.*, **123**, 1483-1515.
- Xue Y., and J. Shukla, 1993: The influence of land surface properties on Sahel climate. Part I: desertification. *J. Climate*, **6**, 2232-2244.
- Zeng N., 2003: Drought in the Sahel. *Science*, **302**, 999-1000.
- Zeng N., J. D. Neelin, K. M. Lau, and C. J. Tucker, 1999: Enhancement of interdecadal climate variability in the Sahel by vegetation interaction. *Science*, **286**, 1537-1540.
- Zeng N., and J. D. Neelin, 1999: A land-atmosphere interaction theory for the tropical deforestation problem. *J. Climate*, **12**, 857-872.
- Ziervogel G., and R. Calder, 2003: Climate variability and rural livelihoods: assessing the impact of seasonal climate forecasts in Lesotho. *Area*, **35**(4), 403-417.

### **Chapter III : Using regional GCM wind to forecast seasonal rainfall over Sahel**

The first objective of this chapter is to develop and apply a methodology to approach seasonal prediction in situations when there is strong evidence that in the real climate system, the SSTs do influence rainfall variability. For this, a MOS methodology is applied (e.g. Feddersen et al., 1999; Landman and Goddard, 2005; Friederichs and Paeth, 2006), using large-scale fields from the GCM that, based on regional climate knowledge, should contain the relevant teleconnection information from the SST. The choice of wind field for the MOS predictors is first guided by the knowledge that in West Africa, the low level wind is one of the major features of the rainfall dynamics. It is through the low level wind that moisture is carried from the ocean in the monsoon flow to constitute precipitable water over land. Camberlin et al. (2001) relate SST variations to changes in East-West wind circulation patterns which impact rainfall through the monsoon flow. For example, Camberlin et al. (2001) hypothesize that El Niño events tend to enhance low-level northeasterlies and to reduce the monsoon flow, which coupled to a weakened upper easterly structure, favors dry conditions over West Africa. Thus, after documenting the generally poor skill of direct GCM output for Sahel rainfall (section III.1.a), a MOS prediction system is constructed (section III.1.b) that uses a dynamically plausible field from GCMs as the predictor for the target Sahel seasonal rainfall. The initial focus is on the ECHAM4.5-GCM, a state of the science model for which most complete archives are available to this study. Results with ECHAM4.5 draw from and build on the MOS analysis in Ndiaye et al. (2009). For key findings, we also analyze 7 other atmospheric GCMs, allowing more general conclusions to be drawn.

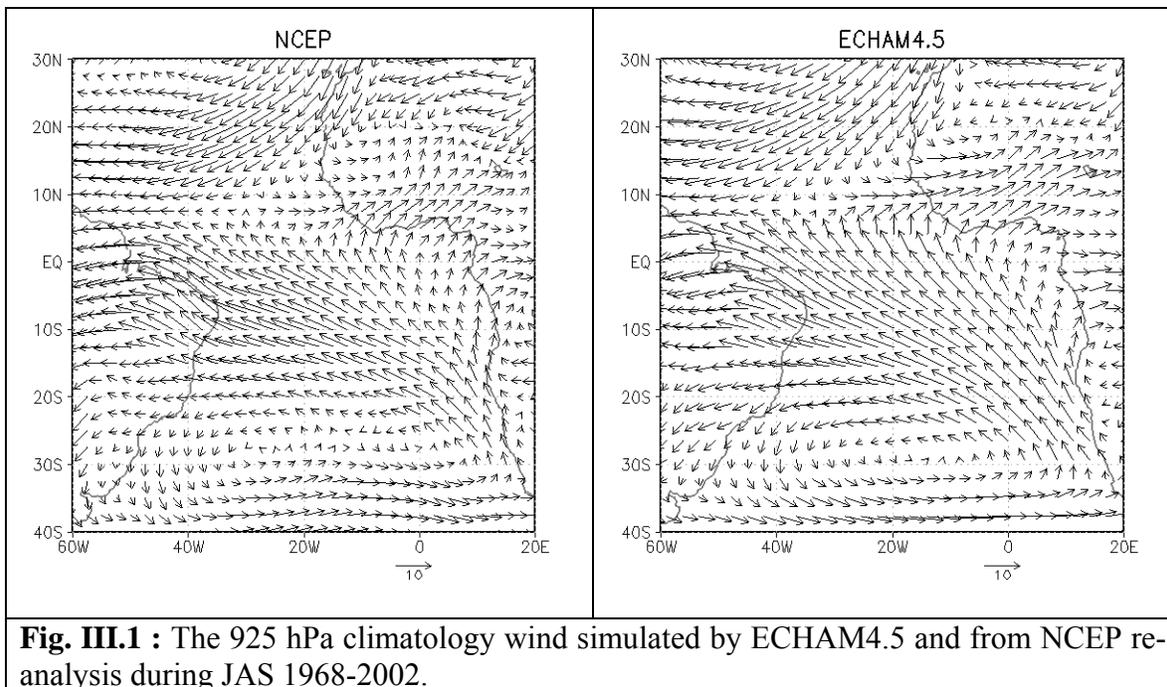
In addition, this chapter also investigates (in section III.2) the sensitivity of the MOS prediction system to sub-period for the Sahel rainfall forecast (using simulations from ECHAM4.5). One of the key issues in Sahel prediction has been the representation and character of predictability through the contrasting wet and dry epochs. Using the GCM simulation experiments driven with observed SST over the longer period 1950-2002, the stability of the MOS through the wet and dry epochs is investigated.

### **III.1 Model performance with observed SST: direct GCM output versus a MOS system**

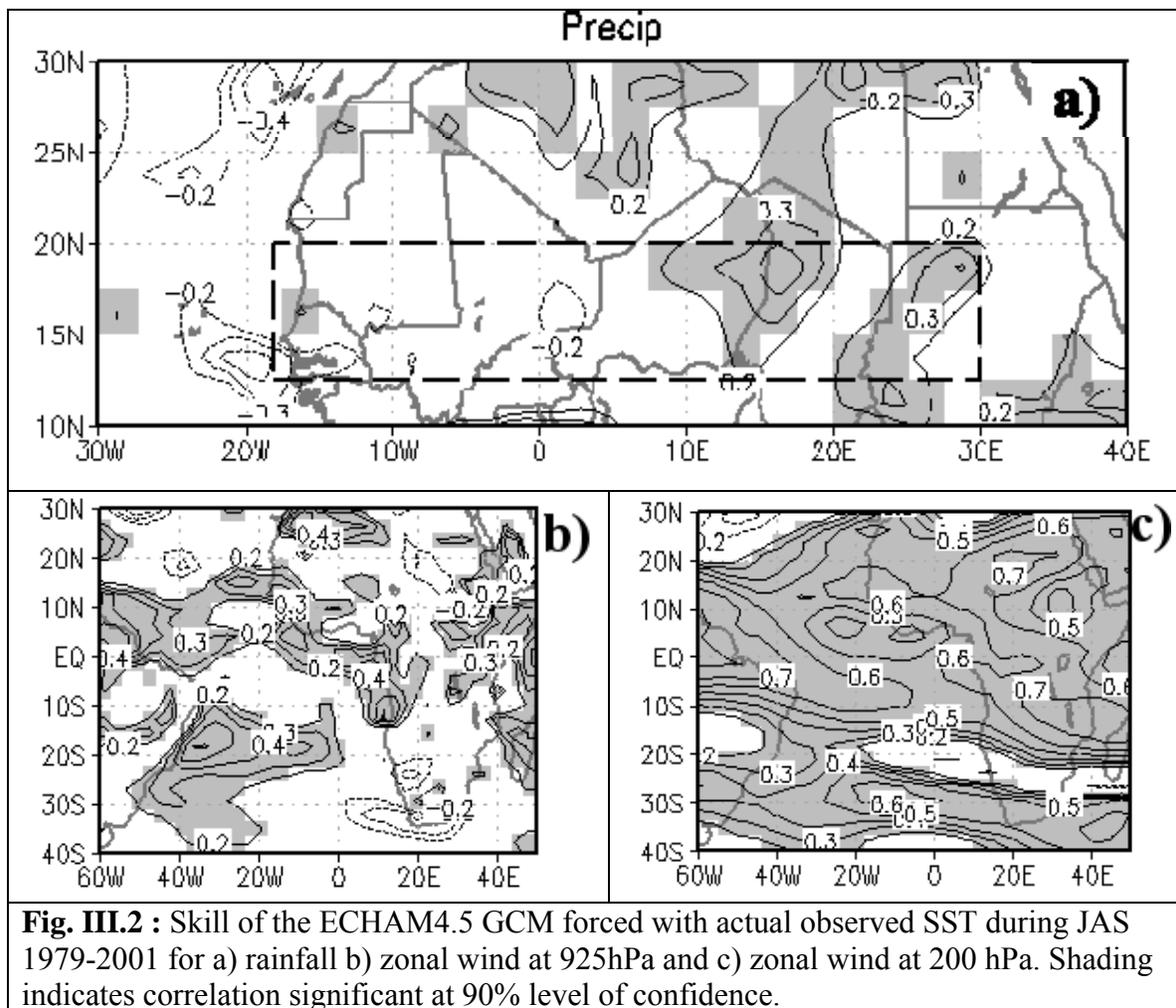
#### **III.1.a ECHAM4.5 AGCM simulation skill**

In this sub-section we will assess the skill level of seasonal rainfall achieved by direct ECHAM4.5 simulated rainfall output. It is known that most AGCMs fail to skillfully simulate the interannual variability of Sahel rainfall when driven with the observed SST (further documented and discussed in Section III.1, see Fig. III.6). In this section we also evaluate the skill of the direct ECHAM4.5 GCM output of some wind fields, to set the stage for the MOS application in the next section.

First, it was checked that the ECHAM4.5 simulates a reasonable climatological pattern of low-level winds over the Sahel and surrounding region (Fig. III.1). Indeed, the main structures of the wind circulation are all well simulated by the model. The large-scale monsoon flow is well reproduced by the model. The strength of the south-easterly flow over the tropical South Atlantic and near the Brazilian coast is correctly simulated.



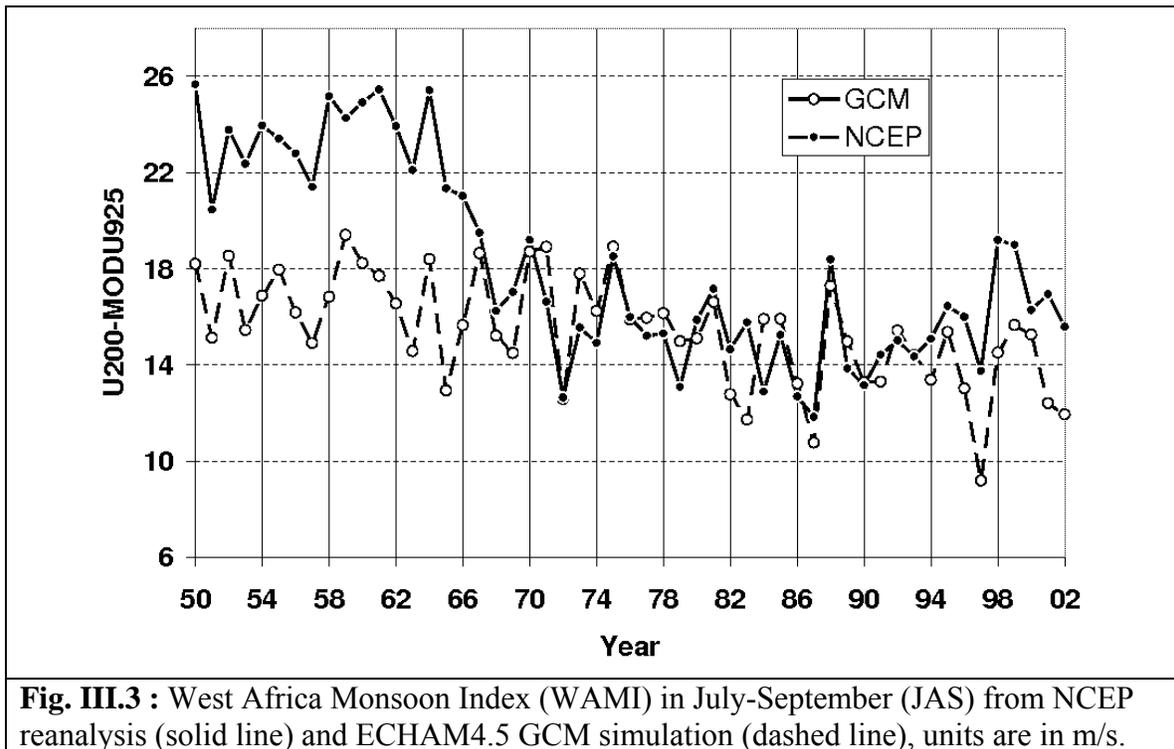
The skill of the ECHAM4.5 in simulating year-to-year seasonal rainfall variations is evaluated for each grid-box by calculating the correlation between ECHAM4.5 simulated rainfall and the observations. The GCM does not simulate the interannual rainfall variations accurately over the Sahel during 1968-2002. On a grid-box by grid-box basis, the average correlation skill over the Sahel is 0.04 (Fig III.2.a). On the other hand, the year-to-year variation of the wind is much better captured over much of the Atlantic basin and, with somewhat weaker skill, extending into West Africa. At 925hPa the model captures the wind variability over the tropical North Atlantic and over the South Atlantic basin especially near the east coast of Brazil (Fig. III.2.b). At 200hPa the skill is also high (Fig. III.2.c), especially for the zonal (U) component, particularly around 10N near the axis of the Tropical Easterly Jet (TEJ), whose variability is an important component related to Sahel rainfall (Fontaine and Janicot, 1992).

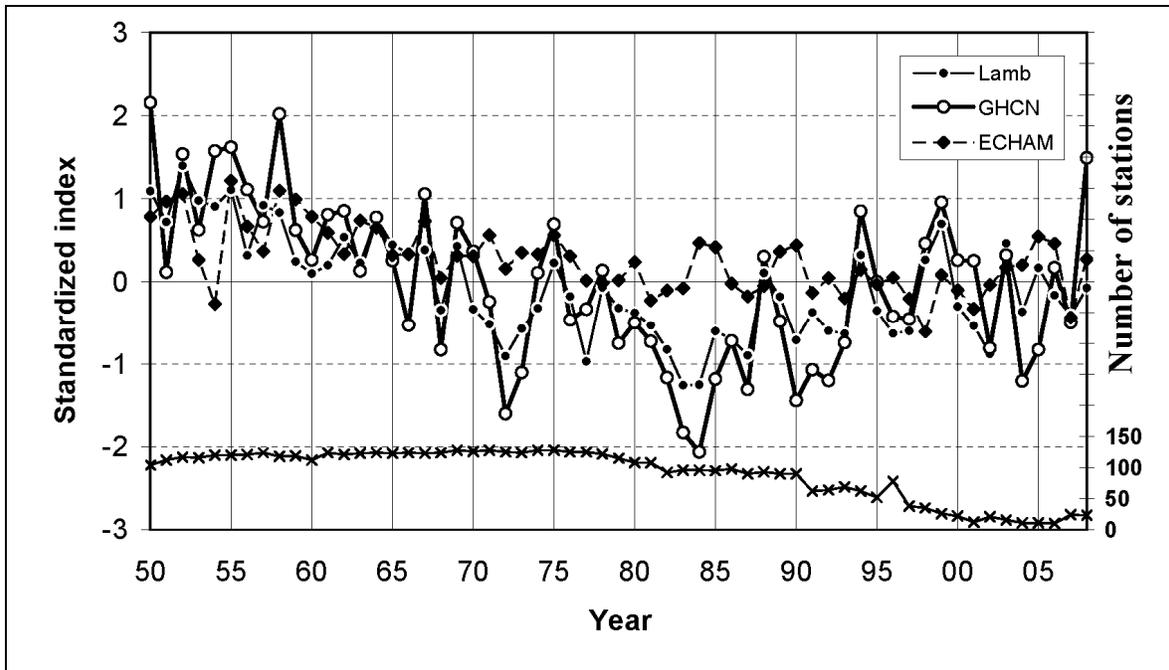


The use of indices, such as the West African Monsoon Index (WAMI), facilitates the examination of regional variability, removing the stringent requirement of accurately reproducing the local variability. The dynamical monsoon index, WAMI, is defined here as the shear between the low level monsoon wind modulus at 925hPa and the zonal wind component at 200hPa, each averaged from 20W to 20E and from 3N to 13N as defined by Moron et al. (2004) and Garric et al. (2002). Figure III.3 shows the WAMI during JAS as simulated by the GCM and observed in the NCEP reanalysis. The correlation between the two indices is 0.53 over the 1968-2002. This indicates a reasonable agreement

between the reanalysis and the GCM prediction of the WAMI based on observed SST.

The step change in the reanalysis data prior to 1968 in Fig. III.3 may at least in part be related to a change in the mix of observations entering the reanalysis.





**Fig. III.4 :** Standardized July-September Sahel rainfall indices, from GHCN station network (open circle), the index constructed by Lamb (closed circle) and from ECHAM4.5 GCM (diamond). The number of stations used in the GHCN index (thin line) is indicated on the right Y-axis.

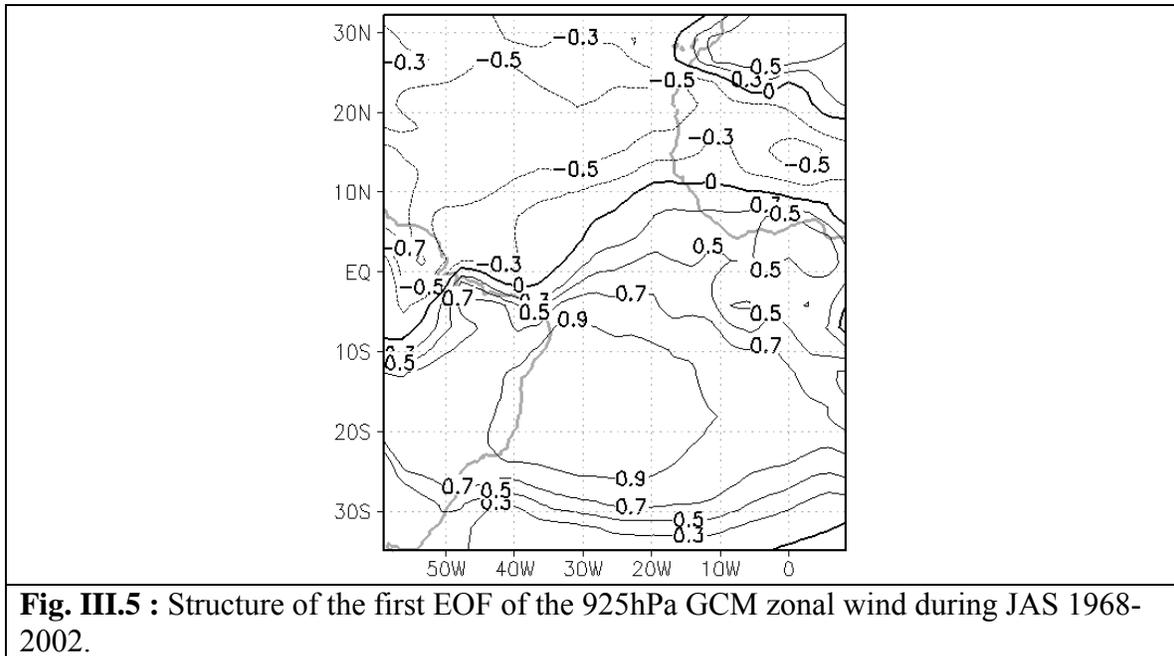
Figure III.4 shows the time series of the observed index and the index calculated from GCM simulations. On the multi-decadal timescale, though the GCM has the correct sign of variation, it is unable to capture the magnitude of the difference between the wet period (1950-70) and the more recent dry period. Interannual variability in the GCM is also too weak, and the year-to-year fluctuations show no agreement with observations. For the recent period dominated by interannual variability (1968-2002) the correlation is 0.07. The GCM is therefore able to better simulate variability in wind across both the tropical Atlantic and West Africa, than it is the Sahel rainfall ( $r = 0.07$ ). The next thing to investigate is if the variability in the wind can be used to correct GCM rainfall over Sahel.

### **III.1.b Using MOS to correct GCM Sahel rainfall**

Given the importance of regional wind variability (tropical Atlantic and West Africa) for rainfall variability over the Sahel (Lamb, 1978; Hastenrath, 1990; Camberlin et al., 2001; Fontaine et al., 1995; Fontaine and Janicot, 1992), and especially the low-level monsoon circulation, we hypothesize that a statistical transformation of the region's low-level wind can be used to forecast the Sahel rainfall. For the MOS, the approach taken is to make a regression or multiple regression between the MOS predictors and the target predictand, the Sahel rainfall index. The candidate MOS predictors are the leading EOFs of the region's low-level wind field. Selection of which EOFs to include in the model needs careful consideration. The MOS systems in this and subsequent sections use cross-validation in which all information on the year being forecast (plus the year before and after) is withheld from model estimation. In other words, the window length for the cross-validation is three years. While a step-wise procedure has been applied to assist in predictor selection, the choice in this study is also made based on physical interpretation of the predictors in relation to known influences on Sahel rainfall. However, results turn out to be very clear-cut and not sensitive across reasonable criteria for inclusion.

Table III.1 provides a summary of EOF predictors and model performance, as the number of EOFs in the model is incrementally increased from one to ten. The EOF1 of 925hPa zonal GCM wind across the tropical Atlantic and surrounding region (Fig. III.5) is found to contain essentially all the predictive information for the Sahel rainfall. The time-series of the first EOF of the zonal wind is very similar to that of the first EOF of the meridional

wind (correlation of 0.94 over 1968-2002), and no additional skill is achieved through use of meridional in place of zonal wind.



**Table III.1:** Skill of predicting Sahel rainfall 1968-2002, using a GCM MOS with zonal wind at 925mb over the window 40S-30N latitude, 65W-10E longitude. EOFs are incrementally added to the prediction system. The row headings are defined as follows: %var indicates the % of wind-field variance explained by each EOF, CV skill is the cross-validated skill of the multiple regression model in predicting Sahel rainfall, R(EOFvSahel) is the individual correlation between each EOF and Sahel rainfall.

	1968-2002									
EOF number	1	2	3	4	5	6	7	8	9	10
%var	34	19	12	7	5	4	3	2	2	2
CV Skill	0.57	0.52	0.57	0.59	0.58	0.55	0.55	0.56	0.55	0.48
R(EOFvSahel)	<u>-0.63</u>	-0.02	0.13	-0.28	-0.03	-0.11	-0.01	0.15	-0.13	-0.09

A small peak (cross-validation skill  $r = 0.59$ ) is found with the addition of zonal wind EOF4 (Table III.1). No physical explanation could be associated to EOF4, which explains only 7% of the total variance in the wind field (Table III.1). Furthermore, a pattern like EOF4 was not found to add skill in longer-period analyses (1950-02, see

section III.2). Even with cross-validation, the risk of choosing a model that contains spurious fitting in the historical period is still present, and in this case, it is proposed that there is not a sufficiently strong case to include EOF4. Note that when a model is constructed using EOF4 alone, and cross-validated, its skill is just  $r = 0.06$ , clearly not statistically significant, and so inclusion of EOF4 is not considered justified based on these statistical grounds either.

Before settling on the model in Table III.1 with EOF1 of zonal wind for ECHAM4.5, other checks were made to confirm robustness of the result and that substantial skill was not being lost through choice of MOS domain and variable. Finally, the generality of the result will be tested across seven other GCMs, to establish the extent to which this MOS performance with low-level wind appears to be a systematic general property of the state of the science GCMs.

#### **III.1.b.1 Robustness in the choice of the domain**

First the regional domain for the MOS predictor was perturbed in each direction by up to 20 degrees latitude and 20 degrees longitude. The first EOF continued to be the dominant skill source, with cross-validated skill ranging from  $r = 0.53$  to  $r = 0.58$ . It is proposed to maintain the original choice of the domain for the MOS, considering the slight peaks of skill with other domains as reflecting statistical noise in the results, and not likely to lead to increases in skill when the system is applied in real-time.

A further aspect to check is whether a more global domain for the MOS predictors would improve skill. This was not expected to be the case. While near-global SST patterns are

considered to influence Sahel rainfall, the transmission of the global SST influence is considered to primarily occur through the regional circulation system, both in the real climate system and in the GCM. Thus, with an appropriately selected domain to capture the regional system, any global SST influences should be captured by the regional MOS. This indeed appears to be the case in the problem studied here. For example, taking a global tropics (40N-40S) domain for the U-wind, cross validation skill with EOF1 was found to be  $r = 0.46$ , with a small peak ( $r = 0.47$ ) when three EOFs are used.

### III.1.b.2 Robustness in the choice of the GCMs variable

The possibility that other model variables may perform better as MOS predictors has been explored. A commonly used variable in MOS systems is model precipitation itself, under the hypothesis that the model can better represent the large-scale regional precipitation patterns which can then form useful predictors for particular target precipitation predictands. Two other variables closely related to tropical precipitation are vertical motion at 850mb and low-level humidity. The skill achievable with each of these three variables has therefore been tested (Table III.2).

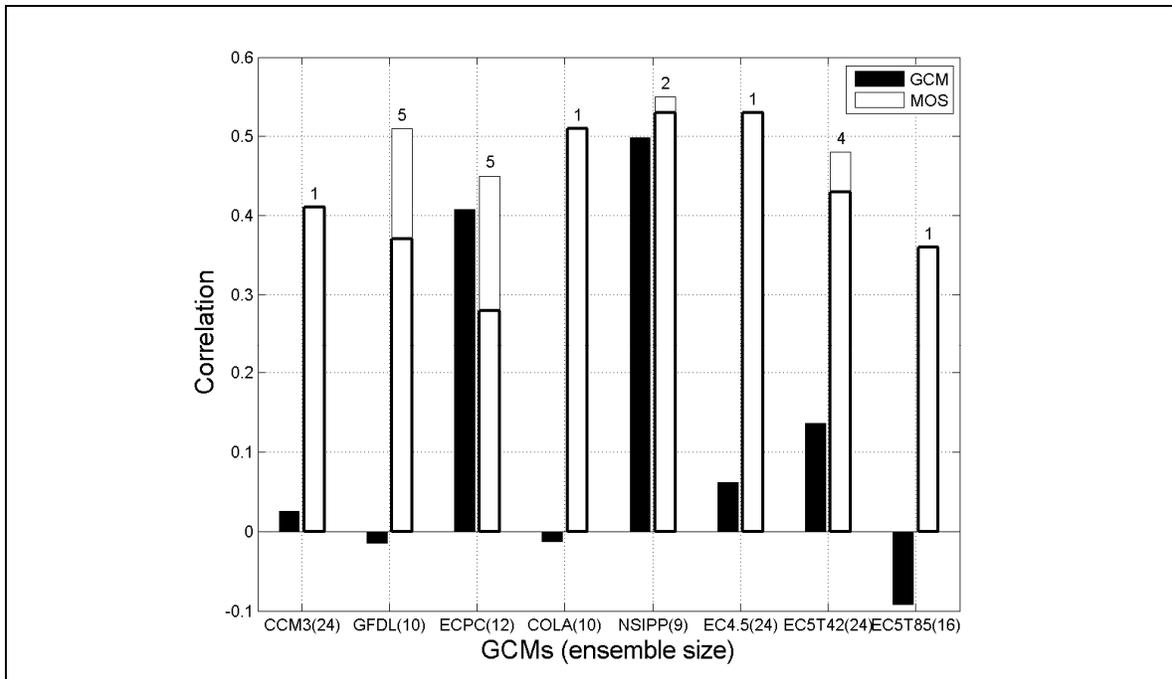
**Table III.2 :** Skill of predicting Sahel rainfall 1968-2002 using a GCM MOS with different variables over the window 40S-30N latitude, 65W-10E longitude. The column headings are defined as follows: U is zonal wind at 925mb, precip is the regional precipitation field, omega is the vertical motion at 850mb, and humidity is at 850mb level. The top row uses the first MOS EOF. The second row shows the best cross-validated result after incrementally adding additional EOFs.

	<b>U</b>	<b>Precip</b>	<b>Omega</b>	<b>Humidity</b>
EOF1	0.57	0.32	0.37	0.10
Best model	0.59 (4 EOFs)	0.37 (4 EOFs)	0.52 (3 EOFs)	0.35 (4 EOFs)

All models' skill are inferior to that using the low-level zonal wind. Regional humidity and precipitation are quite substantially less skillful (best model correlations are less than  $r = 0.40$ ). This suggests the model has most difficulty with these variables in this region at this time of year. A model with 3 EOFs of vertical motion achieves the best result of the new models ( $r = 0.52$ ) but still inferior to the model with EOF1 of zonal wind. A further consideration was whether an index such as WAMI could be used as a MOS predictor to improve upon skill. The GCM's simulated WAMI has a cross-validated skill of  $r = 0.36$  with the predictand observed Sahel rainfall index, substantially less than that achieved by the EOF1 of zonal wind. Furthermore, when the model's WAMI is added to the zonal wind EOF1 in a regression model, the cross-validated skill was found to be less than that with the model of wind EOF1 alone.

### **III.1.b.3 Robustness in the choice of the GCM**

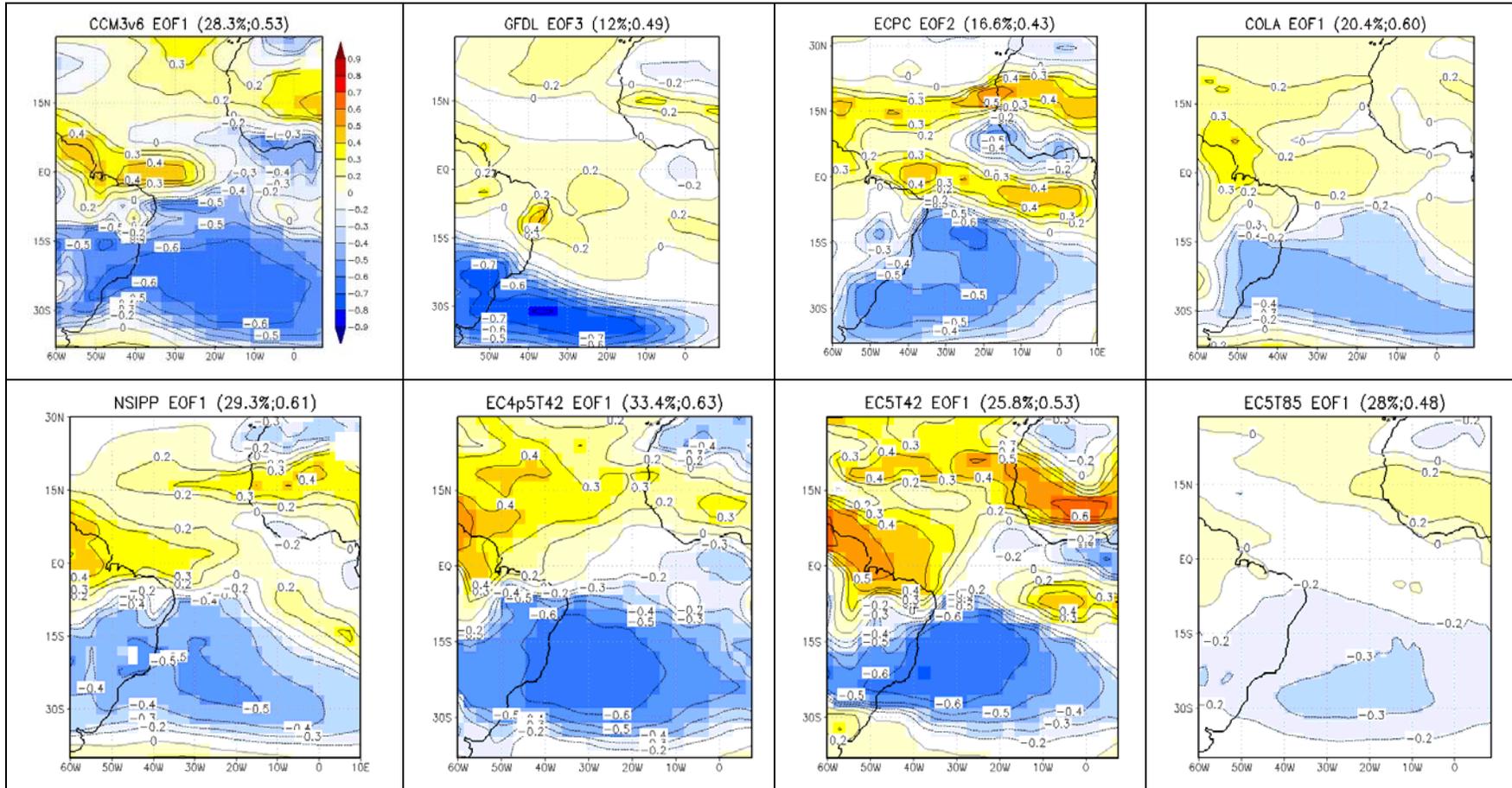
The consistency of our MOS system is now explored across different models, using simulations from seven other available AGCMs (Table II.1). We keep the same domain for calculation of the wind EOFs, and perform the MOS over the same time period 1968-2001. Figure III.6 summarizes the results of the MOS applied to the wind field from each model. The MOS correlation skill is compared to that achieved by direct GCM rainfall output.



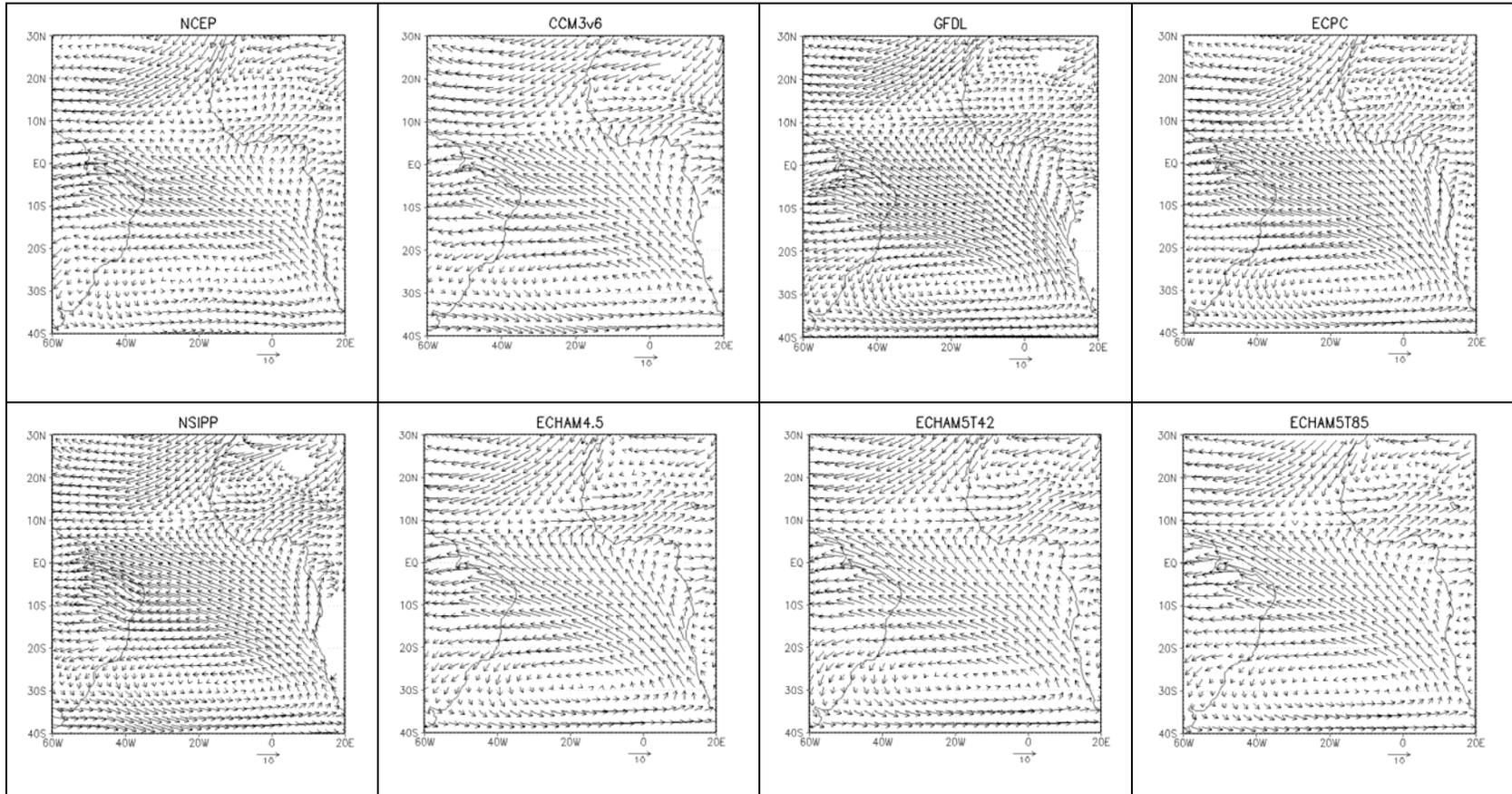
**Fig. III.6** : Sahel rainfall skill during 1968-2001 from various GCMs forced with observed SST. Skill is shown for GCM rainfall (shaded bars), and after applying a Model Output Statistics (MOS) to the 925hPa tropical Atlantic wind field (open bars). The first open bar indicates the skill when the MOS predictor is the best single EOF, usually the leading EOF (see Fig. III.7a). If a combination of up to 5 EOFs improves the skill further, the additional skill is indicated by the additional (thin line) open bar, and the number of EOFs yielding the best skill is indicated by the digit on the top of the open bar.

The GCM precipitation has very low skill in 6 of the 8 GCMs. In contrast, the MOS skill is quite consistent throughout all the GCMs, with correlation skill scores generally in the approximate range of  $r=0.4$  to  $r=0.55$ .

In Fig. III.6, the skill shown in the lower open bar is that using EOF1 alone, with the exception of the results for the GFDL and ECPC models. For these models, a lower order EOF contains the leading signal for predicting Sahel rainfall (EOF2 for ECPC and EOF3 for GFDL).

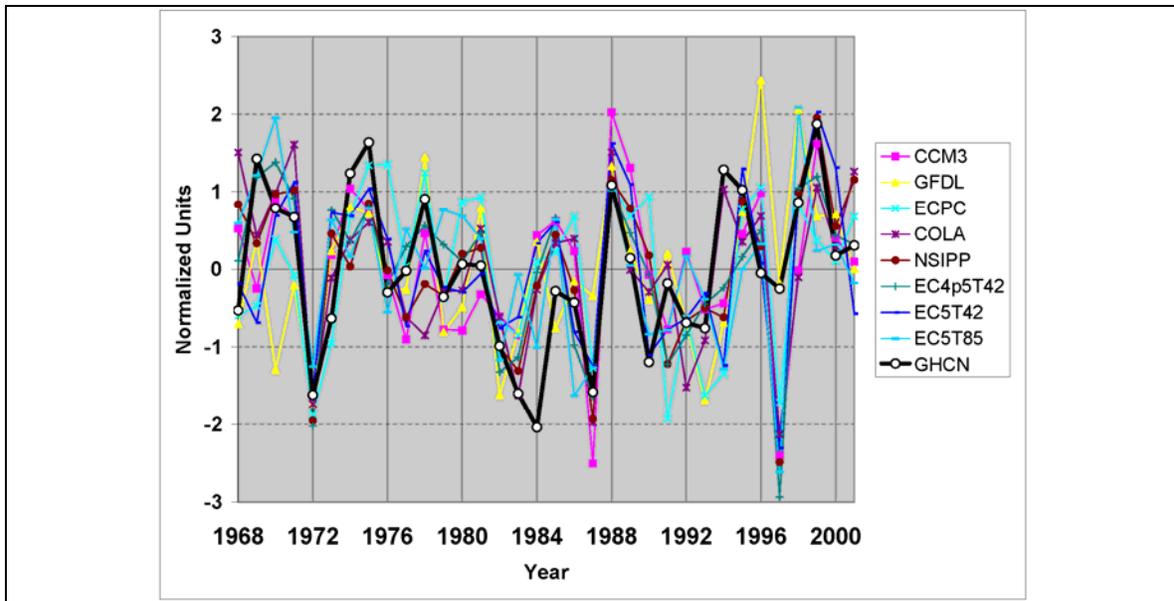


**Fig III.7a :** The 925 hPa zonal wind EOF that is most powerful as a MOS predictor of Sahel rainfall, in each of 8 GCMs. The percentage of variance and correlation with observed Sahel rainfall index are in brackets.



**Fig. III.7b :** The 925 hPa climatological wind simulated by AGCMs depicting the low level monsoon flow during JAS 1968-2001.

Figure III.7a shows the spatial pattern of the leading explanatory EOF in each of the models (EOF2 for ECPC, EOF3 for GFDL and EOF1 for all other models), and Fig. III.7b shows the basic climatological wind field for each model. The similarity of the large-scale spatial pattern in the EOFs (Fig. III.7a) is clear, with zonal wind anomalies of one sign in the tropical and sub-tropical South Atlantic, and a (somewhat more varying) pattern of opposite sign in the tropical North Atlantic and into West Africa. A change of sign is common around the tip of NE Brazil on the west side of the Atlantic. The EOF is capturing large-scale and coherent fluctuations in the complex of the inter-tropical convergence zone, the West African monsoon, and their interactions with the Northern and Southern Hemisphere Hadley Circulations in the Atlantic sector. It is interesting that ECPC and NSIPP, the only two models that have any success of substance with their direct model-output rainfall, have particularly similar patterns over West Africa and especially Sahel latitudes. Most other models vary considerably in their low-level wind weights over West Africa, likely reflecting their difficulty in translating the large-scale circulation over the tropical Atlantic into model rainfall over the Sahel. The time-series of each of the EOFs in Fig. III.7a is shown in Fig. III.8, strongly reinforcing the idea that the ability to represent the interannual variation of the tropical Atlantic wind field associated with Sahel rainfall variations is a robust feature across models, even if the spatial expression varies somewhat over and close to West Africa.



**Fig III.8 :** Time-series of each of the EOFs shown in Fig. III.7a (colored lines), along with GHCN Sahel rainfall index (black line).

Table III.3 documents the diagnostic correlation of the time-series in Fig. III.8 with Sahel rainfall, with the skill shown in Fig. III.6 for the leading EOF being slightly lower than the correlation values in Table III.3 due to the cross-validation evaluation procedure used in the MOS prediction system.

**Table III.3** Correlation between observed Sahel Index and GCM EOF of U wind at 925hPa as shown in Fig. III.8 The last column is the correlation between the average of the eight model time-series and Sahel rainfall

	CCM3	GFDL	ECPC	COLA	NSIPP	EC4.5	EC5T42	EC5T85	Average
corr	0.53	0.49	0.43	0.61	0.61	0.63	0.53	0.48	0.65

In Section III.1.b with ECHAM4.5, the use of one EOF could not be substantially improved upon by adding further EOFs to the MOS when applied over the period of 1968-2001. Figure III.6 shows that this result is largely reproduced here for the additional

seven GCMs. For the models where EOF1 was not successful (GFDL, ECPC) a leading explanatory EOF does emerge at lower order (EOF2 for ECPC, EOF3 for GFDL). However, for GFDL and ECPC, it is also clear that the prediction signal has become split across more than one EOF, such that more substantial improvements in skill are obtained for multiple EOF models (0.37 to 0.51 for GFDL, 0.28 to 0.45 for ECPC). Nonetheless, with the multiple EOFs applied in this way for the ECPC and GFDL models, the MOS approach using tropical Atlantic winds is able to recover skill comparable to the other models.

So the result is very consistent across the GCMs studied here, and the low level zonal (tropical Atlantic and western part of West Africa) wind is a very strong candidate to correct poor GCM Sahel rainfall.

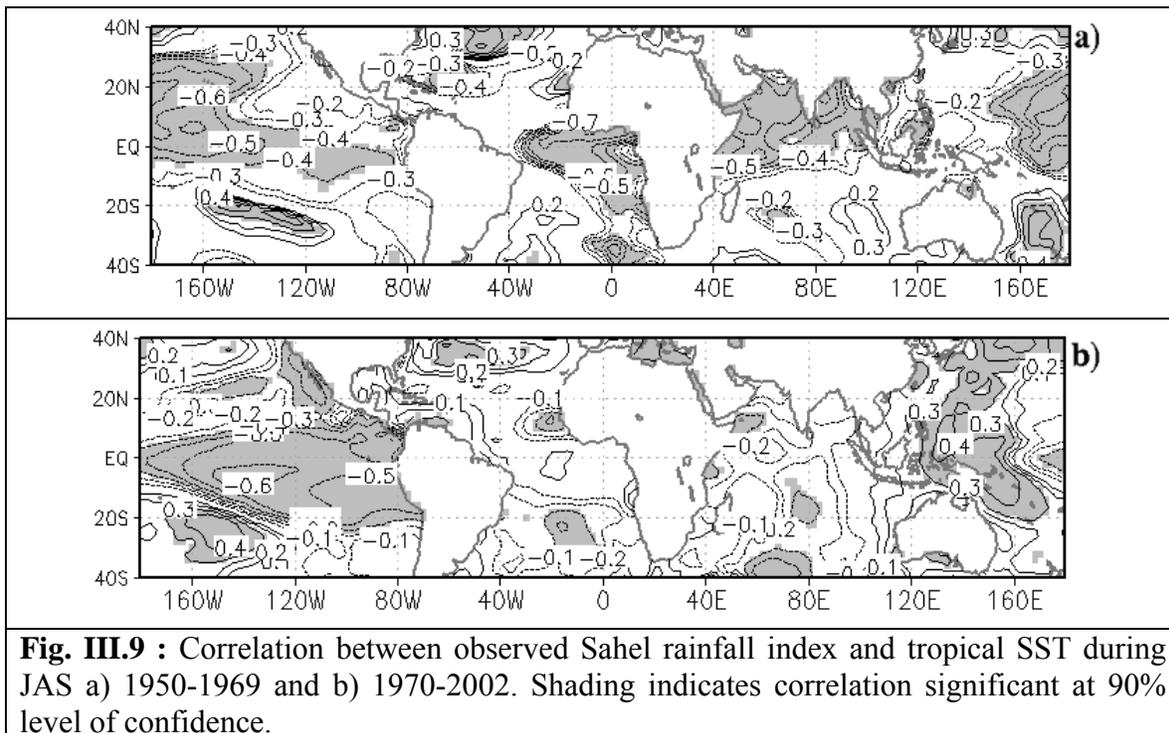
We have also made an initial study of the potential improvement from applying a multi-model approach, by combining the zonal wind from all available GCMs into a single set of candidate predictors. The rationale behind the multi-model approach is that some GCMs may represent better certain physical processes (i.e. different GCMs have differing strengths and weaknesses), and by combining different GCMs we can best capture the different physical features associated with Sahel rainfall variability. For each of the 8 GCMs (Table II.1) we selected the first 10 EOFs of the zonal wind. We pre-filter the 80 new variables by applying another EOF analysis, to avoid co-linearity between variables, concentrating the main information into a smaller number of variables. Then we apply a multi-linear regression on the new EOFs. By using this procedure we find it

very difficult to outperform the top four MOS results for individual GCMs (Fig. III.6). Furthermore, it is necessary to select 13 EOFs in the multi-linear regression in order to have a skill ( $r = 0.51$ ) that is comparable to the better individual GCM-MOS results. This means that selecting the best MOS system for a couple of individual GCMs may be a better practical approach for this rainfall prediction problem, than using a multi-Model GCM MOS. When we apply the multi-model approach on GCM regional rainfall, a good skill is even harder to find, even with 13 EOFs the skill is only  $r = 0.24$ . The result suggests that when a GCM MOS is successful, it tends to capture the same variability as other successful GCM MOS models.

In summary, over the period 1968-2002, the MOS system using the regional low-level wind EOF1, is capable of substantially improving the rainfall skill of the ECHAM4.5 GCM simulation, increasing the anomaly correlation coefficient of the Sahel index from 0.07 to 0.57. A similar result is reproduced using 7 other GCMs. This indicates that GCMs can simulate correctly the low level winds associated with the Sahel rainfall, even though they usually fail to directly simulate the interannual variability of the rainfall. It indicates also that low level wind (925hPa or 850 hPa) can be used to correct poor GCM rainfall skill over the Sahel. This is based on the role of SST in influencing the regional climate system and most robustly, the winds, and draws on the linkage of wind fields to Sahel rainfall. Tropical Atlantic zonal wind extended over 30N-40S and 60W-10E turns out to be very robust for MOS forecasts of Sahel rainfall.

### III.2 Performance of the MOS through the wet and dry Sahel epochs

It is known that during the wetter period (approximately 1950-1969), Sahel rainfall shows a much stronger linkage with the tropical Atlantic SST (Fig. III.9a), whereas during the drier post 1970 period, the influence of ENSO has been stronger (Fig. III.9b) (and see e.g. Janicot et al., 1996; Ward 1998). While forecast ECHAM4.5 GCM experiments were available only for 1968-2002, the simulations with observed SST were available 1950-2002, and allowed investigation of the performance of the MOS over the whole period and in the sub-periods 1950-1969 and 1970-2002.



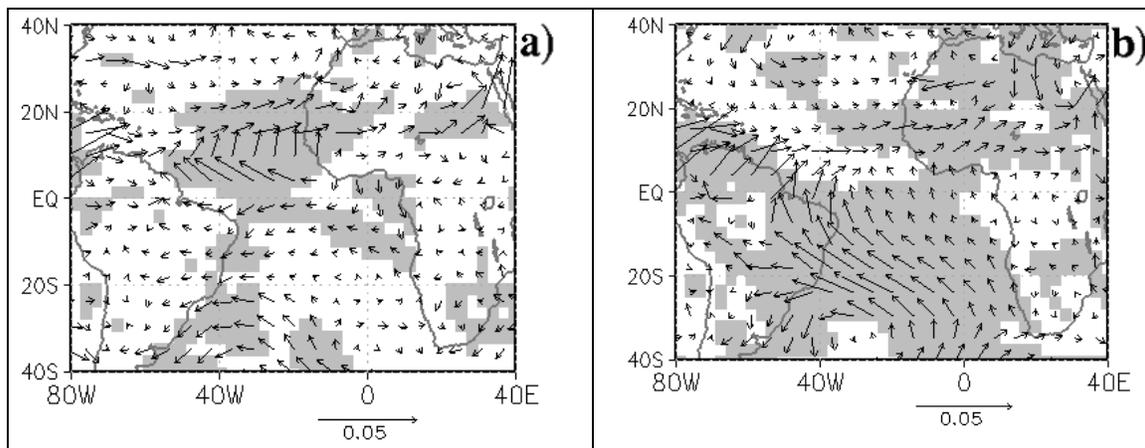
**Table III.4 :** Correlation skill of the raw GCM rainfall and the MOS system, based on GCM experiments driven with observed SST.

	<b>1970-2002</b>	<b>1950-1969</b>	<b>1950-2002</b>
<b>Raw GCM</b>	0.03	0.36	0.55
<b>MOS 1950-02</b>	0.57	0.74	0.78
<b>MOS 1950-69</b>		0.57	
<b>MOS 1970-02</b>	0.52		

First, the performance of the raw model output is reviewed (Table III.4, row 1). For the period 1970-2002, as previously discussed, the raw output has near zero skill. In the period 1950-1969, the raw output turns out to have some modest skill ( $r = 0.36$ ). One interpretation is that the model is better able to directly simulate the effects of the tropical Atlantic SST on the rainfall, as compared to the effects of the tropical Pacific which dominate in the 1970-2002 period. Over the whole period 1950-2002, the correlation skill is  $r = 0.55$ . At first sight this appears a reasonable performance, though as previously noted, the trend that is substantially contributing to the correlation is greatly underestimated (Fig. III.4). It will be shown below that applying the MOS methodology based on the regional low-level wind field is able to substantially improve the skill in all sub-periods.

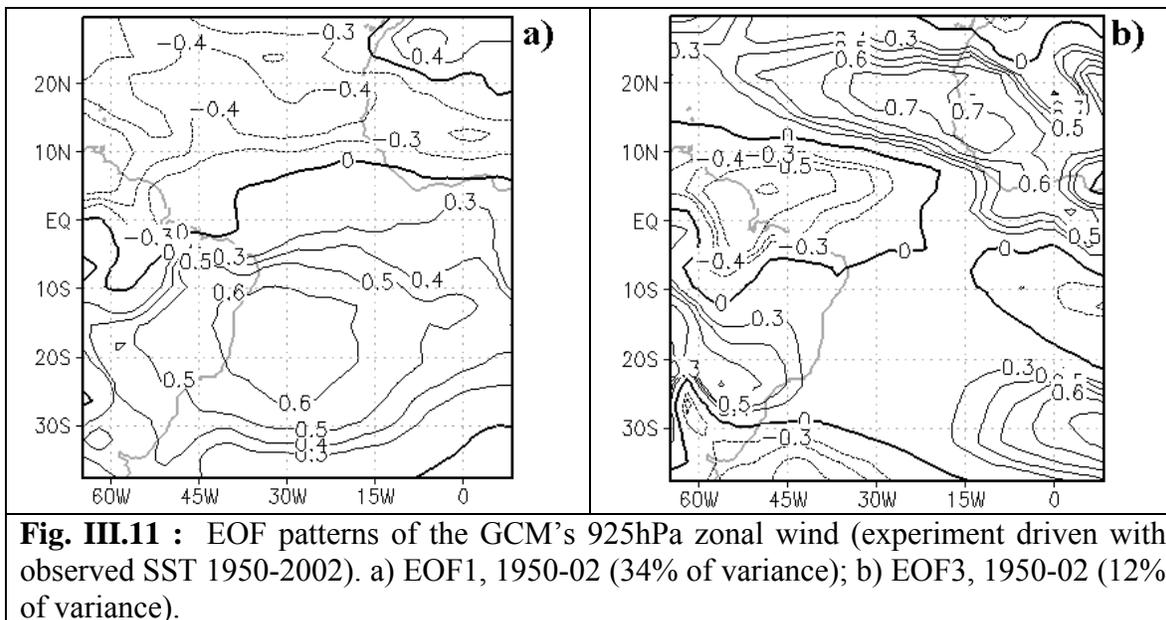
As a preparation for interpretation of the MOS results, Fig. III.10 shows that the teleconnection between observed Sahel rainfall and the simulated GCM low-level wind have some differences between the two epochs. In the latter period 1970-2002 (Fig. III.10b) the pattern resembles a basin-wide modulation of the trade wind - monsoon circulation, resembling closely the EOF1 pattern of model wind that was used in the

MOS in section III.1.b (Fig. III.5). In the early period (Fig. III.10a), some of the basin-wide modulation is again seen, but in addition, there is more strongly superimposed a smaller scale structure with divergence from the Equatorial and South tropical Atlantic, and anomalous westerlies across the northern tropical Atlantic sector into the Sahel of Africa.

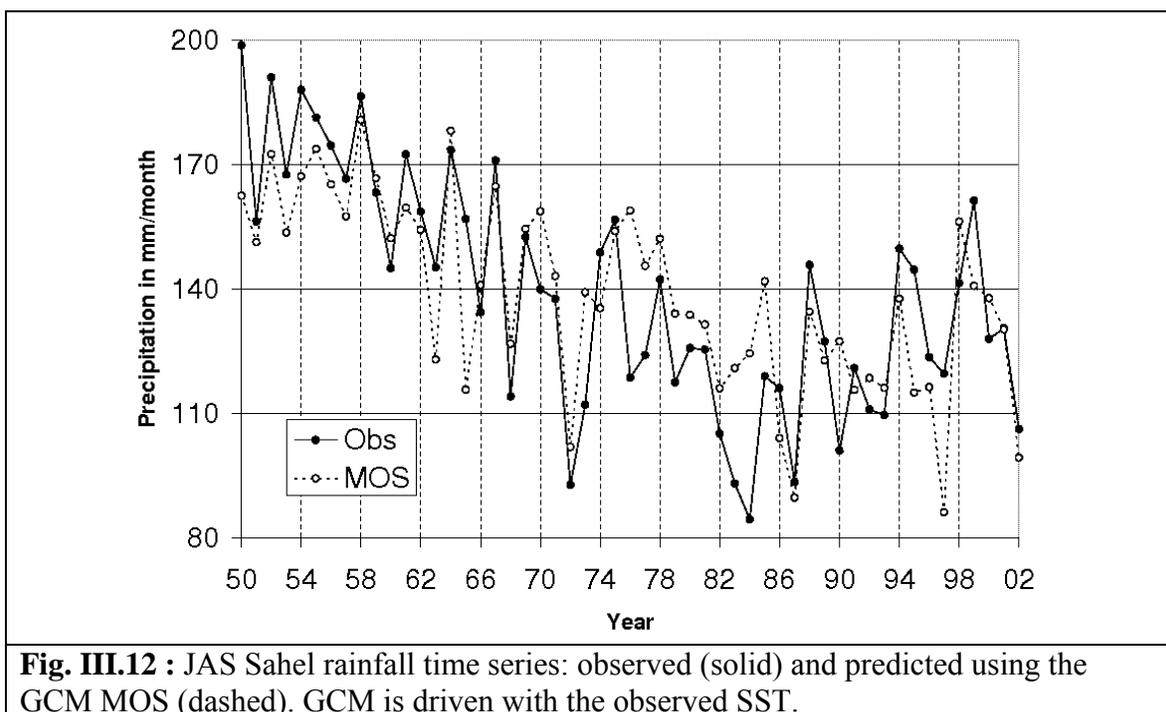


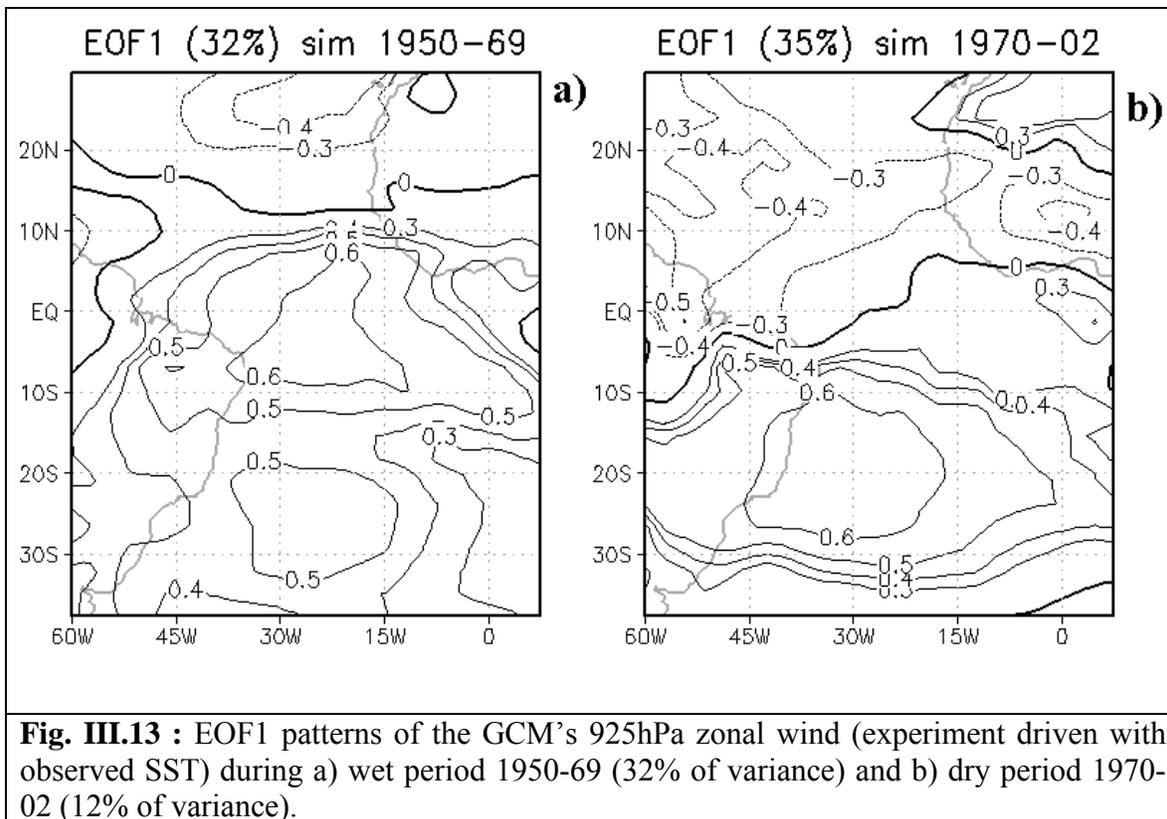
**Fig. III.10** : Regression between Observed Sahel index and each component of the GCM wind at 925hPa during JAS for a) 1950-1969 and b) 1970-2002. Highlighted are areas where the correlation with one of the wind component is significant at 5% error level.

When the MOS is applied to the whole period 1950-2002, two modes provide skill in specification of Sahel rainfall. The first mode (Fig. III.11a) is similar to the one found previously (Fig. III.5). The other mode that adds further skill (Fig. III.11b) is centered through the Equatorial and northern tropical Atlantic and resembles a more localized modulation of the ITCZ and associated monsoon flow. It contains aspects of the model teleconnection map for the early period (Fig. III.10a).



The first mode contains most of the skill ( $r = 0.68$ ), but addition of the other mode does notably raise skill further ( $r = 0.78$ , Fig. III.12).





These two modes, with slightly modified form, are recovered in the two sub-periods (Fig III.13). The first mode in 1950-1969 (Fig. III.13a) has weight shifted somewhat further north than EOF1 of 1950-02 (Fig. III.11a), but over the period 1950-1969, the two time series correlate extremely high ( $r = 0.92$ ), indicating they are essentially capturing the same variability. The period 1950-1969 does however show itself to be rather short for defining the patterns for a stable MOS. The correlation skill of the MOS fitted over these years is 0.57 (Table III.4, row 3). The skill over the 1950-1969 period rises to  $r = 0.74$  using the output from the cross-validated MOS analysis on the whole period 1950-2002 (Table III.4, row 2).

For the period 1970-2002, while the additional mode is indeed recovered in the analysis it turns out to not add to the skill of the MOS. This is interpreted as an expression of the weaker Equatorial and South tropical Atlantic SST influence on Sahel rainfall in this period.

Overall, these results show a strong robustness of the MOS over a period when the background Sahel rainfall was slowly evolving, along with a changing balance of relationships with SST across the different ocean basins. The absolute level of skill for the 1950-02 period needs to be interpreted with caution, since over this period, a strong low frequency component is present in the time series, compromising somewhat the effectiveness of the cross-validation testing of the prediction system. Nonetheless, the stability of the modes and the performance in the two sub-periods are very encouraging. The results build confidence in the application of the MOS system for real-time prediction.

### III.3 Summary

Taken together, the results of this chapter could represent an advance of considerable practical significance for diagnosing Sahel rainfall in state of the science GCMs, and for Sahel rainfall forecasting generally. Firstly, it would indicate that the key information from SSTs was being captured by most models, at least in the regional circulation system. In other words, the poor performance of models could be expected to be improved through addressing the better translation of the tropical Atlantic wind field anomalies into the deeper monsoon anomalies and associated precipitation over West Africa. Secondly, even without further model improvement, it would imply most models could be used now in real-time to extract relevant information on Sahel rainfall, at least at short lead-time, by applying the same MOS approach.

Overall, the use of tropical Atlantic wind fields permits recovery of skill levels at or approaching the best ones found for specifying Sahel rainfall from SST over this approximate post-1970 period, suggesting that GCMs are generally able to capture a substantial part of the interannual variability of the tropical Atlantic wind fields that are tied to Sahel rainfall variability. In most cases, this can be achieved through use of a single EOF, and in all cases, a single EOF captures a significant fraction of the predictable Sahel variance. This encourages use of this approach in operational seasonal forecast settings. However, the results in this chapter have used observed SSTs. The next chapter studies the challenge of achieving lead-time for forecasts, in the presence of SST evolution through boreal spring.

### Reference chapter III

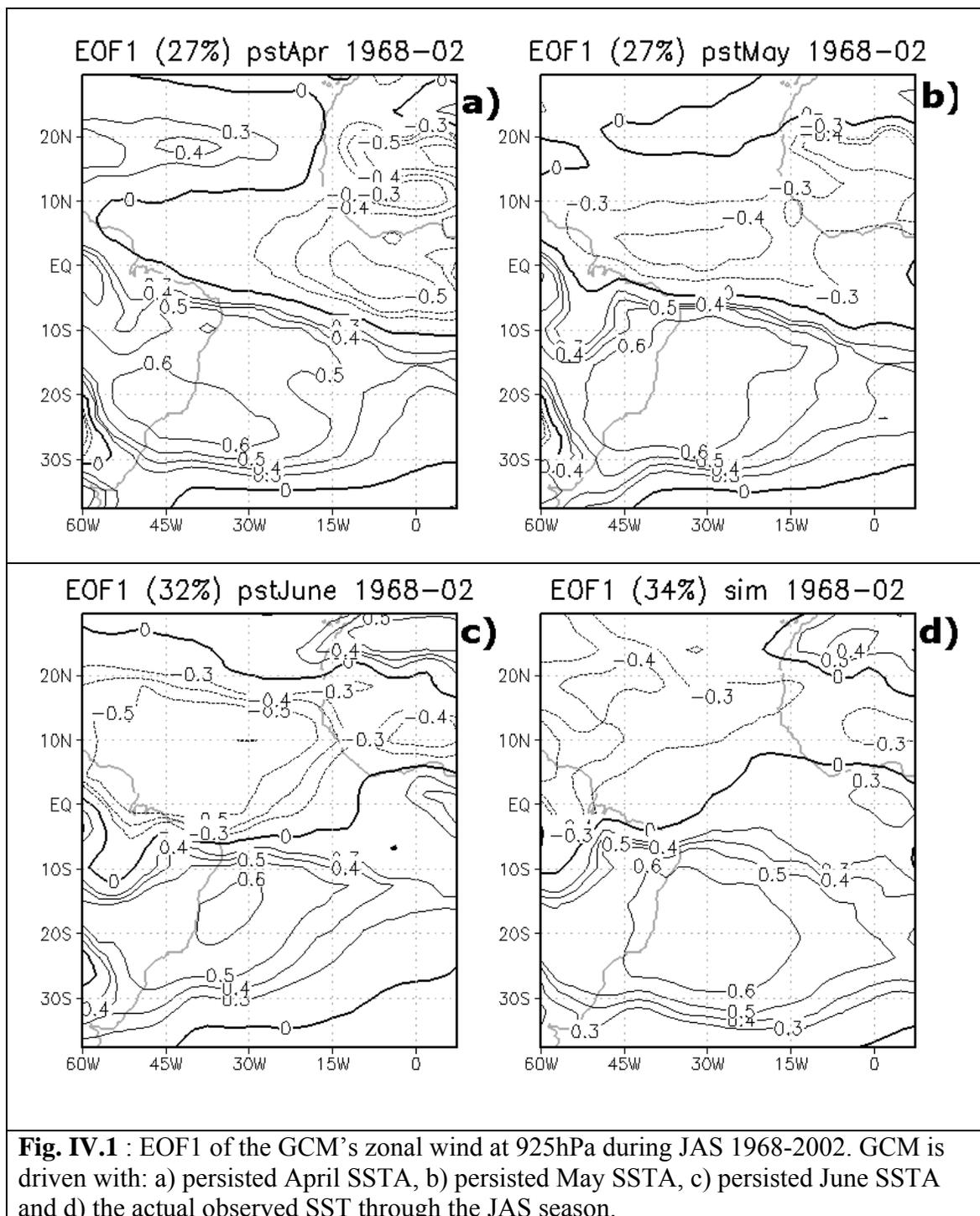
- Camberlin P., S. Janicot, and I. Pocard, 2001: Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea-surface temperature : Atlantic vs ENSO. *Int. J. Climatol.*, **21**, 973-1005.
- Fedderson H., A. Navarra, and M. N. Ward, 1999: Reduction of Model Systematic Error by Statistical Correction for Dynamical Seasonal Predictions. *J. Climate*, **12**, 1974-1989.
- Fontaine B., and S. Janicot, 1992: Wind-Field coherence and its variation over West Africa. *J. Climate*, **5**, 512-524.
- Fontaine B., S. Janicot, and V. Moron, 1995: Rainfall anomaly patterns and wind field Signals over West Africa in August (1958-1989). *J. Climate*, **8**, 1503–1510.
- Friederichs P., and H. Paeth, 2006: Seasonal prediction of African precipitation with ECHAM4-T42 ensemble simulations using a multivariate MOS re-calibration scheme. *Climate Dyn.*, **27**, 761-786.
- Garric G., H. Douville, and M. Deque, 2002: Prospects for improved seasonal prediction of monsoon precipitation over West-Africa. *Int. J. Climatol.*, **22**, 331-341.
- Hastenrath S., 1990: Decadal-scale changes of circulation in the tropical Atlantic sector associated with Sahel drought. *Int. J. Climatol.*, **10**, 459-472.
- Janicot S., V. Moron, and B. Fontaine, 1996: Sahel droughts and ENSO dynamics. *Geophys. Res. Lett.*, **23**, 515–518.
- Lamb P. J., 1978: Case study of Tropical Atlantic surface circulation patterns during recent sub-saharan weather anomalies : 1967 and 1968. *Mon. Wea. Rev.*, **106**, 482-491.
- Landman W. A., and L. Goddard, 2005: Predicting southern African summer rainfall using a combination of MOS and perfect prognosis. *Geophys. Res. Lett.*, **32**, L15809, doi:10.1029/2005GL022910.
- Moron V., N. Philippon, and B. Fontaine, 2004: Simulation of West African Monsoon circulation in four atmospheric circulation models forced by prescribed sea surface temperature. *J. Geophys. Res.*, **109**, DOI:10.1029/2004JD004760.
- Ndiaye O., L. Goddard, and M. N. Ward, 2009: Using regional wind fields to improve general circulation model forecasts of July-September Sahel rainfall. *Int. J. Climatol.*, **29**, 1262-1275.
- Ward M. N., 1998: Diagnosis and short lead time prediction of summer rainfall in tropical North Africa at interannual and multidecadal timescale. *J. Climate*, **11**, 3167-3191.

## **Chapter IV : Model forecast skill and sensitivity to SST developments**

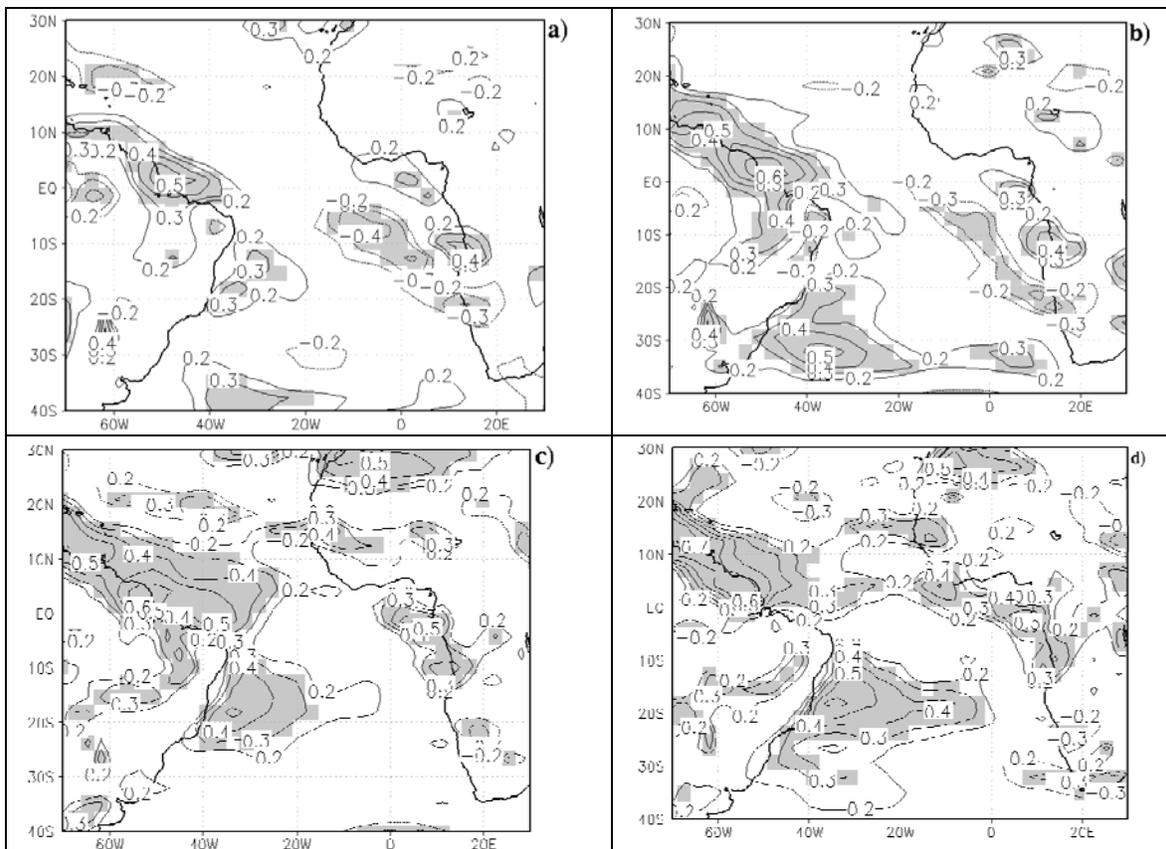
In chapter III we demonstrated the potential of correcting poor GCM Sahel rainfall by using the low level wind over the tropical Atlantic; however, the GCM experiments in chapter III used observed SST boundary conditions. In this chapter we will begin by applying the MOS system to GCM experiments that assume persisted SST from April, May and June, to create true forecasts. We will look at the skill as lead-time increases. Section IV.1 therefore assesses the skill of the MOS using ECHAM4.5 forced with the persisted SST of April, May and June. Next (section IV.2) we will investigate the source of the loss of skill during the spring. Finally (section IV.3), comparisons will be made with the sensitivity of empirical predictions systems to the SST developments.

### **IV.1 Performance with persisted SSTA**

Here we will study the skill of the rainfall forecast (over 1968-2002) using SST persistence at lead times of zero (June) to two months (April) before the main Sahel rainy season in JAS. The first EOF mode of low-level tropical Atlantic wind is also used for the runs with persisted SST. The total wind variance explained by EOF1 is 32%, 27% and 27%, respectively, for the forecast experiments with June, May and April persisted SSTA, and the spatial pattern of the leading mode is similar to the result with simulated SST (Fig. III.5) in all cases (Fig. IV.1).



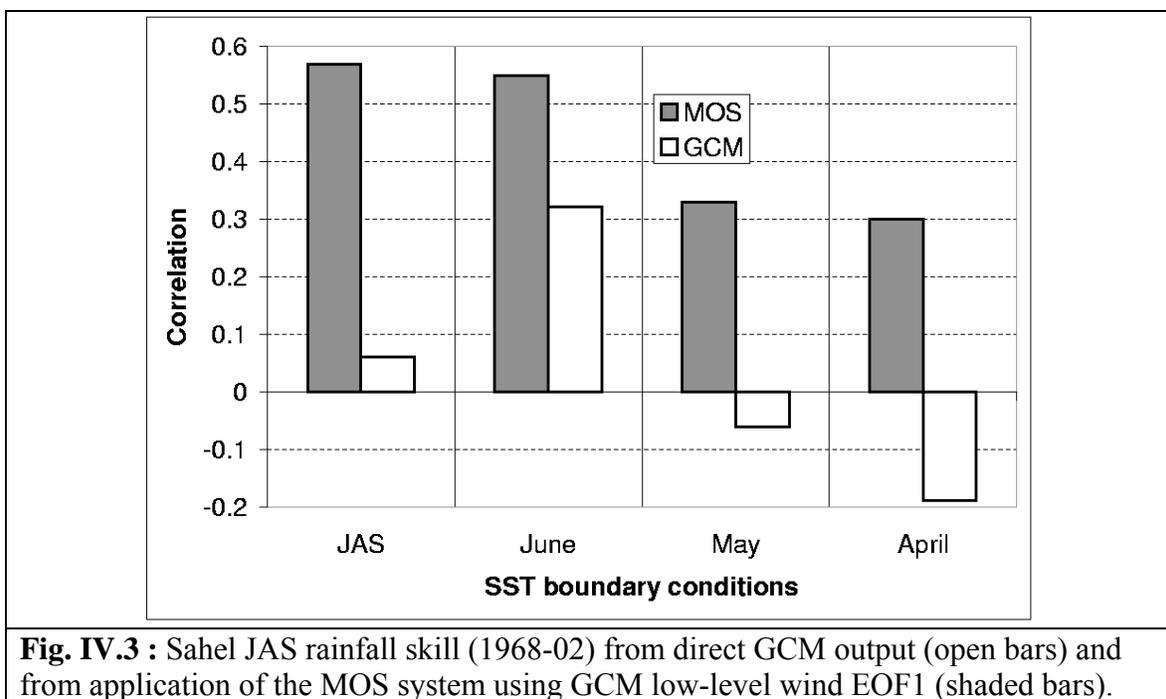
Before reviewing the skill of the MOS rainfall forecast system, we will consider the skill of the model wind fields, since they form the basis of the MOS system. The skill is calculated between observation and model at each grid point. Figure IV.2 shows skill maps for the zonal wind component at 925hPa. The skill maps for the different lead times are very consistent in space, generally just weakening with longer lead: centers of skill are always located over the east coast of tropical Northern and Southern America, and over the Gulf of Guinea near the equator. The skill is very similar for both U and V components. Results for the simulation (Fig. IV.2d) and with persisted June SSTA (Fig. IV.2c) are very similar. However, the persisted May map (Fig. IV.2b) shows a notable drop in skill overall, and especially in areas in the eastern tropical North Atlantic and West Africa. The map for the persisted April SSTA experiments (Fig. IV.2a) shows a further decline in skill.



**Fig. IV.2 :** Skill of the GCM's zonal wind at 925hPa during JAS 1968-2002. GCM is driven with: a) persisted April SSTA, b) persisted May SSTA, c) persisted June SSTA and d) the actual observed SST through the JAS season. Shading indicates correlation significant at 90% level of confidence.

Figure IV.3 compares MOS Sahel rainfall skill for the GCM forced with different SST boundary conditions (JAS, June, May and April). The MOS corrected rainfall always performs better than the raw GCM rainfall. Persisted June SSTA experiments have almost the same skill as those with the observed SST. In other words, the SST boundary condition in June is as good as that of JAS for building a prediction model for the Sahel region targeting the JAS season. However, this provides very short lead time for the targeted rainy season, as observations of June SSTs are not available until early July. The skill deteriorates from 0.55 with June persisted SST to 0.33 and 0.30 with May and April

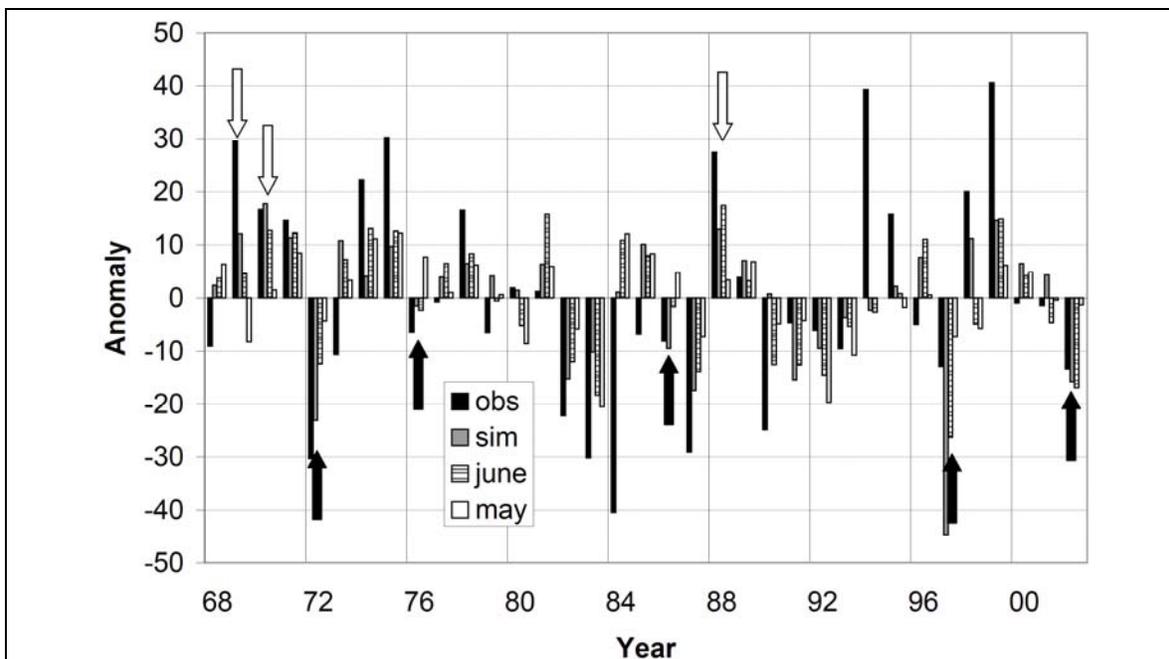
persisted SST, respectively. From June to May, in terms of percent variance explained, the drop is from 32% to 12%, compared to no drop between JAS and June SST. These results are qualitatively similar to those found in Ward et al. (1993) where there was also a very substantial drop in skill when using May as compared to June persisted SSTs. Though those results were based on a much smaller sample of years, it is noteworthy that the same pattern emerges here.



This loss of skill from June to May appears to be real and substantial, and is crucial for the potential uses of the climate forecasts for those who need more lead time to plan for their activities. The next section investigates the source of the loss of skill by looking at the SST evolution from May to June.

## IV.2 Key SST developments between May and June

The set of years is identified where the MOS system is accurate using June (and JAS) SSTA, but fails using May SSTA. We hypothesize that these forecast failures are not due to random internal atmospheric dynamics, but rather, are due to systematic and rapid changes in the ocean boundary forcing from May to June. The analyses below identify the structure of the SST change and the sensitivity of the GCM's global tropics circulation to those SST changes.



**Fig. IV.4 :** Sahel JAS rainfall indices: observed (solid) compared to predictions using the MOS system applied to the GCM. Sets of GCM experiments are shown using observed SST (sim), persisted June SSTAs (jun) and persisted May SSTA (may). Units are in mm/month anomaly.

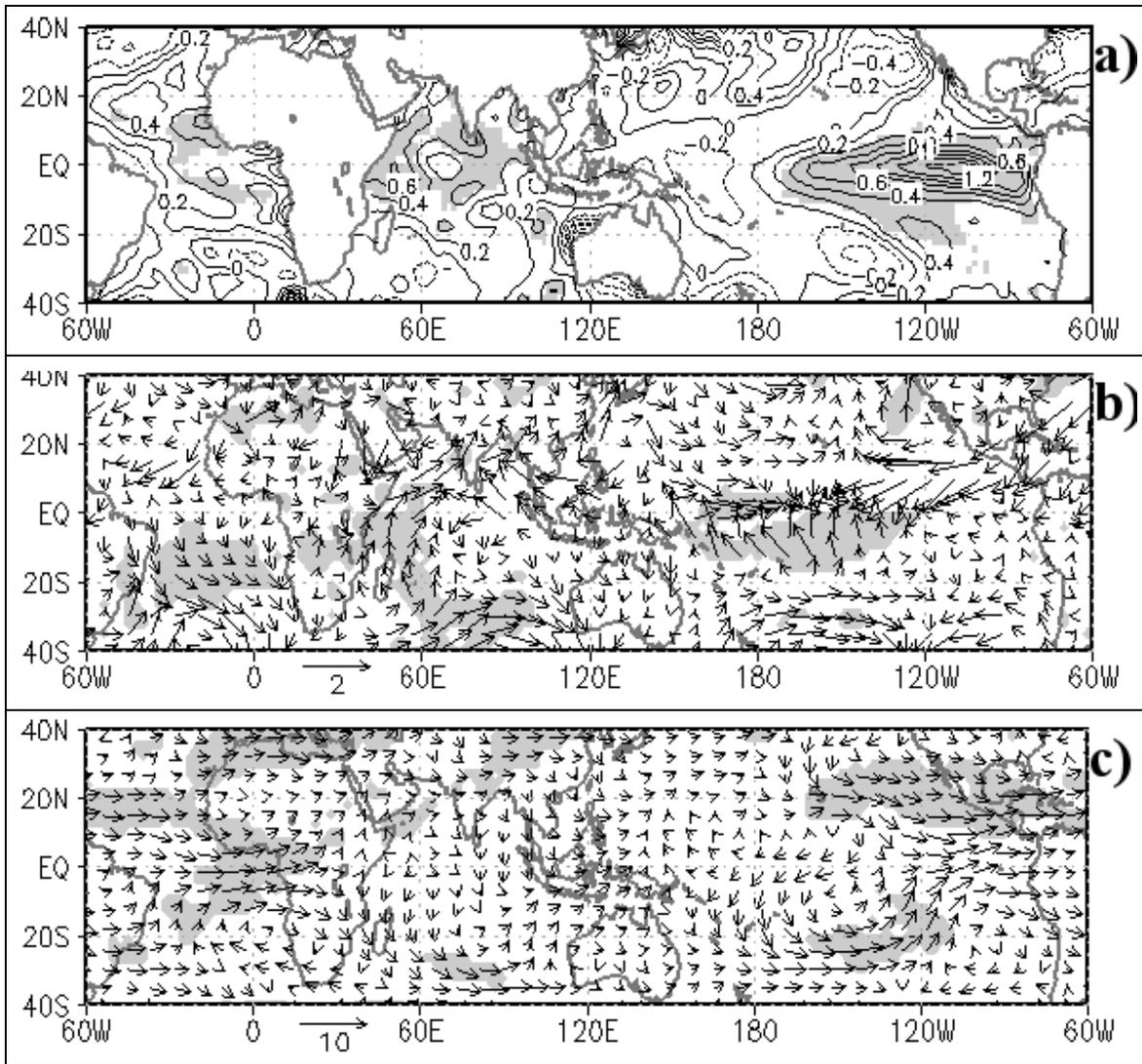
The years identified for the analysis are indicated in Fig. IV.4. In the wet years 1969, 1970 and 1988, the June persisted SSTA MOS gets the right rainfall response, as does the simulation (i.e. the MOS for the GCM run with JAS SSTA). The May persisted SSTA MOS, however, produces either a very weak or an opposite response (open arrows, Fig.

IV.4). Likewise, a set of dry years (1972, 1976, 1986, 1997 and 2002) is identified when the simulation and the June persisted SSTA MOS capture the dryness in the Sahel, but the May persisted SSTA MOS produces a much weaker or opposite response (Fig. IV.4, black arrows). To see the systematic changes in SST during these years, the following composite is constructed:

$$\begin{aligned} \text{SST}' = & [(\text{JUN}_{72} - \text{MAY}_{72}) + (\text{JUN}_{76} - \text{MAY}_{76}) + (\text{JUN}_{86} - \text{MAY}_{86}) + (\text{JUN}_{97} - \text{MAY}_{97}) \\ & + (\text{JUN}_{02} - \text{MAY}_{02})]/5 - [(\text{JUN}_{69} - \text{MAY}_{69}) + (\text{JUN}_{70} - \text{MAY}_{70}) + (\text{JUN}_{88} - \text{MAY}_{88})]/3 \end{aligned}$$

(Eq. 4.1)

where  $\text{JUN}_{69}$  is the SSTA field for June 1969, etc. The field of  $\text{SST}'$  is shown in Fig. IV.5a. It can be interpreted as the systematic pattern of SST change (tendency) from May to June, in years when the May forecast fails but the June forecast successfully forecasts dry. The field multiplied by -1 can be considered the systematic pattern of SSTA change (tendency) from May to June in years when the May forecast fails, but the June forecast successfully forecasts wet. The wet and dry years are combined in this way to increase sample size and allow more robust estimate of statistical significance (shading on Fig. IV.5a).



**Fig. IV.5 :** Composite evolution of anomalies from May to June for years when the MOS system fails in May but succeeds in June (see years marked on Fig. IV.4 (we combine the result for wet years (69-70-88) and dry years (72-76-86-97-02) and show the sign of the plot for the development in dry years. (a) Observed SST (b) surface wind response in the GCM and (c) 200hPa wind response in the GCM. Shading indicates significance at the 90% level of confidence (for either u or v wind) using the Man-U-Whitney test

A substantial SSTA tendency from May to June in these years is found in the central and eastern Equatorial Pacific. The tendency is substantial both in terms of magnitude and statistical significance. The suggestion is that the forecast failures are associated with rapid developments from May to June in the ENSO phenomenon, with La Niña developing in wet years and El Niño developing in dry years. In addition, statistically

significant SSTA tendencies are found in the tropical Atlantic and northwestern Indian Ocean, with SSTA cooling (warming) in these locations in the Sahel wet (dry) years.

To assess the implications of the SSTA tendencies for global tropics atmospheric circulation, a similar composite approach to Equation 4.1 has been taken with the GCM experiments:

$$\text{JAS Circulation}' = [(\text{JUN}_{72} - \text{MAY}_{72}) + (\text{JUN}_{76} - \text{MAY}_{76}) + (\text{JUN}_{86} - \text{MAY}_{86}) + (\text{JUN}_{97} - \text{MAY}_{97}) + (\text{JUN}_{02} - \text{MAY}_{02})]/5 - [(\text{JUN}_{69} - \text{MAY}_{69}) + (\text{JUN}_{70} - \text{MAY}_{70}) + (\text{JUN}_{88} - \text{MAY}_{88})]/3 \quad (\text{Eq. 4.2})$$

where  $\text{JUN}_{69}$  is now the JAS GCM circulation in the experiment driven with persisted June SST anomalies for 1969. These fields are shown for near-surface wind (Fig. IV.5b) and 200hPa wind (Fig. IV.5c). They can be viewed as the response of the model to the change in boundary forcing from May to June in years when the forecast from May SST fails, but the forecast from June SST successfully forecasts dry conditions in the Sahel. Again, as with the SST, when the circulation fields are multiplied by -1, they show the counterpart for the wet years. It can be seen that the SST tendencies are sufficient to drive a substantial change in the GCM low-level wind response over the central and eastern Equatorial Pacific (magnitude 2m/s), with statistically significant changes in the 200mb upper level easterlies extending from the eastern Pacific into the tropical Atlantic. At the surface in the tropical Atlantic, there are also significant changes in circulation, with anomalous northwesterlies south of the Equator, and with some weakening of the monsoon circulation north of the equator. This pattern resembles the one used in the MOS, and is consistent with the better performance of the MOS using June SST compared to May SST. The suggestion then, is the teleconnection from the tropical

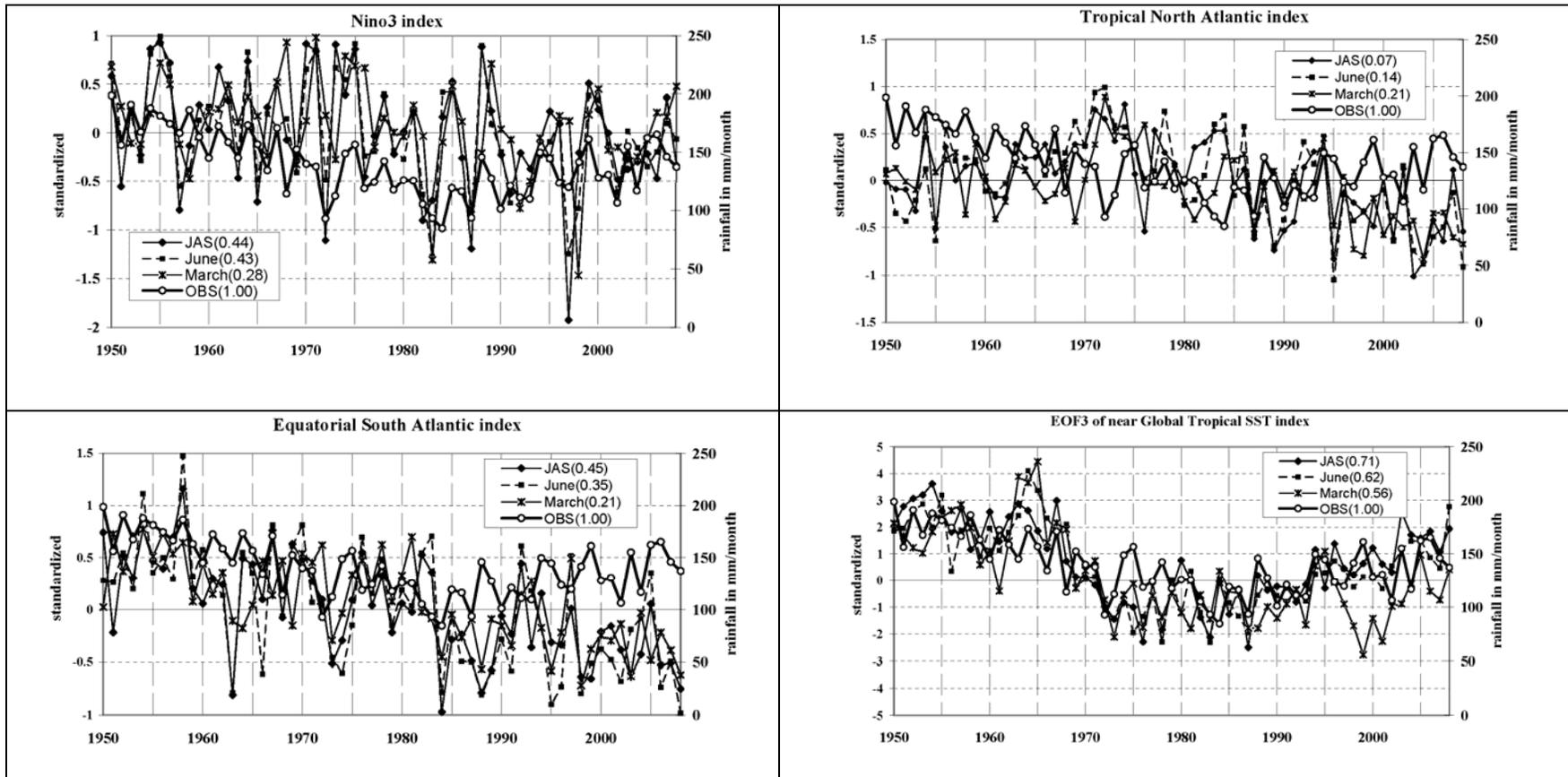
Pacific to the tropical Atlantic basin is better established in the forecasts from June SSTs. The wind circulation depicted in fig. IV.5c matches very well with the dynamical response expected for putting a localized tropical heat source (as in Fig. IV.5a), according to the Gill model (Gill, 1980).

It is also possible that the tendencies in SST in the tropical Atlantic and tropical Indian Ocean (Fig. IV.5a) are also playing a role in the better establishment of the circulation anomalies in the tropical Atlantic. Further investigation of this will assist in evaluation of whether SSTs outside the tropical Pacific need to be forecast from May to June in order to improve Sahel forecasts, and if so, which aspects of the SSTs. From the results here, the situation in the Indian Ocean appears particularly problematic. The tendency for increasing of SST in the northwestern Indian Ocean (as seen in Fig. IV.5a) induces in the GCM a significant increase of near-surface monsoon flow over the western Indian Ocean (Fig. IV.5b). Rather than a clue to the SST field that needs to be forecast, this is more likely a clue to a more fundamental problem with driving the GCM with observed SST, since warming of the SST in this location is often a response to a weakened Indian monsoon, not a primary driving agent for atmospheric anomalies. This problem of monsoon circulations driving SST changes, that if used to drive a GCM, give the opposite response in monsoon circulation, has now been quite widely investigated (e.g. Wang et al., 2004).

### **IV.3 The problem of SST development in empirical prediction schemes**

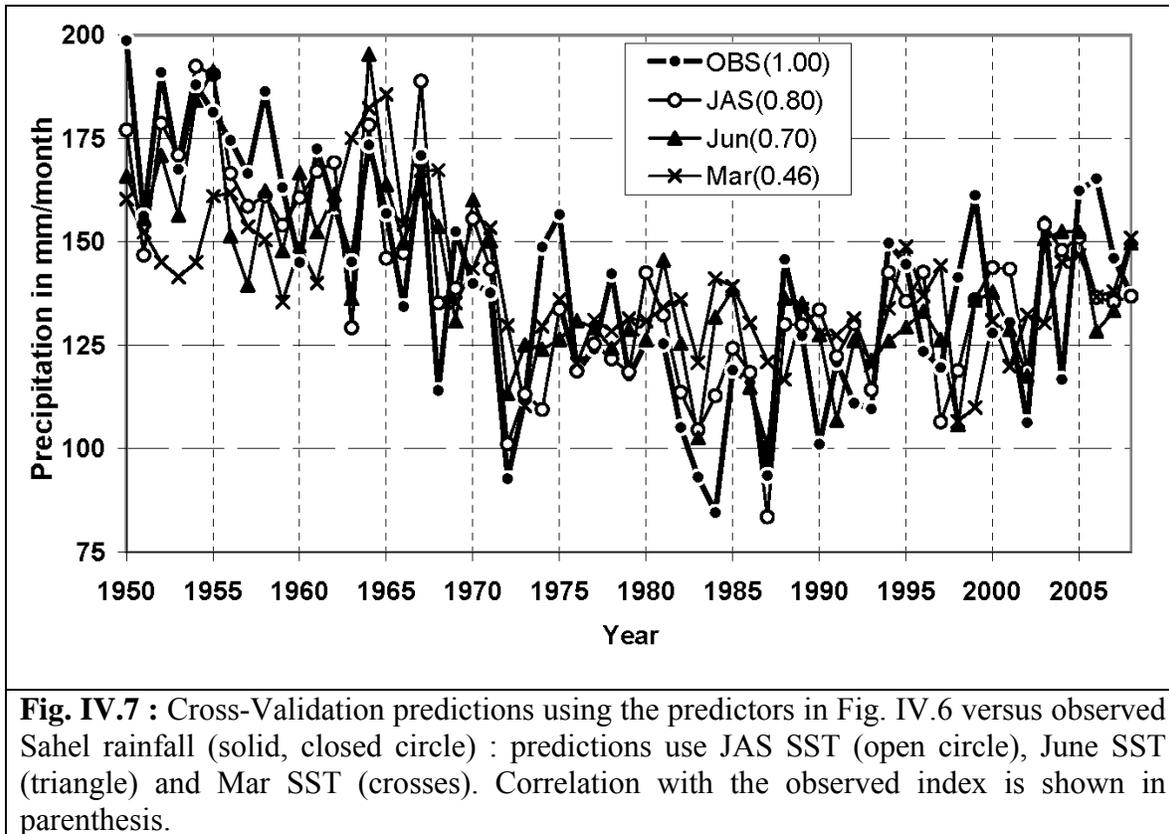
Previous predictability studies of Sahel rainfall have used pre-rainfall season SSTs as predictors in empirical prediction models (e.g. Folland et al., 1991; Ward, 1998; Thiaw et al., 1999). Especially when focusing on the interannual variability of rainfall, a very substantial loss of skill has generally also been found when using SST information in April and May, as compared to using SST information in June. This result was isolated using empirical approaches in Ward (1998). Such an empirical result is consistent with that found in section IV.1 when using the MOS system applied to the ECHAM4.5 GCM forced with persisted SST.

Since 1998, under the supervision of the African Centre of Meteorological Application for Development (ACMAD), there has been produced an operational seasonal forecast for the region of West Africa as a whole (PRESAO). The forecast is for the July-August-September seasonal rainfall. In PRESAO, linear statistical models between SST and rainfall are developed and applied by National Meteorological and Hydrological Services (NMHS) to make seasonal forecasts. Generally, four predictors are applied : Central/Eastern tropical Pacific (Fig. IV.6 upper left), Equatorial South Atlantic (Fig. IV.6 lower left), tropical Northwestern Atlantic (Fig. IV.6 upper right) and an EOF time series capturing the (mostly) multi-decadal North-South SST contrast (Fig. IV.6 lower right) (definitions and details are in WMO/ACMAD, 1998). These selected indices were based on early work applying both the statistical and dynamical method (Ward et al., 1993; Folland et al., 1991; and Ward, 1998; in turn drawing on the many diagnostic and modeling studies, as listed in the introduction Chapter).



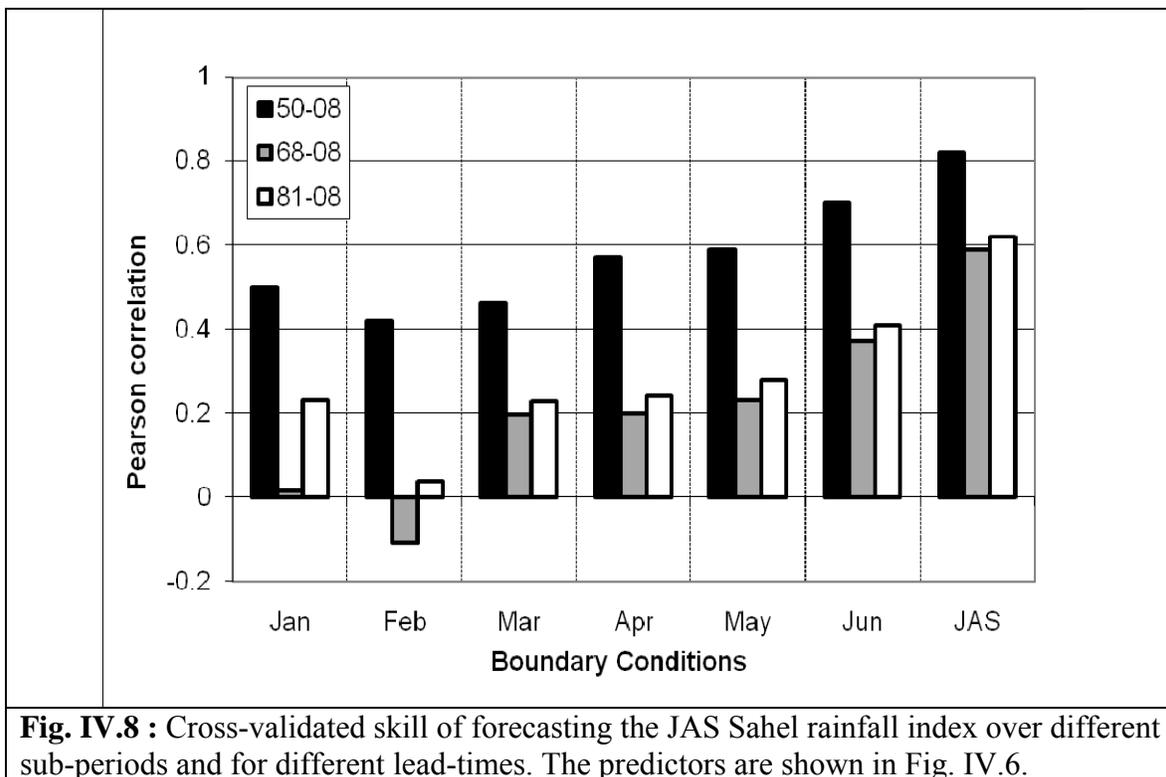
**Fig. IV.6 :** Time-series of Sahel rainfall (open circles) and the four SST predictors used in empirical forecast models. The predictors are as given in WMO/ACMAD (1998), namely, central/eastern tropical Pacific (10S-10N, 150W-90W), tropical NW Atlantic (20N-40N, 30W-10W), Equatorial South Atlantic (20W-10E, 0-10S) and the third EOF of global SST representing near-global north-south interhemispheric variation of SST (positive values in the figure correspond to a relatively warm Northern Hemisphere and relatively cool Southern Hemisphere (EOF3 time-series coefficient provided by Andrew Colman, personal communication). The correlation of each predictor with Sahel rainfall is given in brackets on each panel.

Fig IV.7 and IV.8 summarize the overall skill using these four predictors in a cross validated multi-linear regression.



The period 1950-2008 is strongly influenced by the strong downward trend from about 1950-60 to 1970-90 in the observations. The ability of the seasonal forecasts to track the trend (Fig. IV.7) enables the forecasts to gain a generally higher correlation skill score, compared to the recent period (see Fig. IV.8), that is more dominated by inter-annual variability. In the recent periods more dominated by interannual variability, the role of SST development and the spring barrier becomes very clear in these empirical models (Fig. IV.8). The skill is very weak from January to around April/May and starts increasing faster in June and shows the upper limit of predictability using JAS SST of

around  $r=0.6$  for both periods 1968-2008 and 1981-2008. The pattern of these results largely mirror the lead-time barrier found in our MOS approach using persisted SST (Fig. IV.2). The skill scores in Fig. IV.8 are conservative estimates of that achievable in the PRESAO, since predictions of SST are considered when the predictors are operationally applied. However, the similarity between the empirical and GCM MOS approach as quantified here, points to a common need for SST prediction. The similarity also reinforces the idea that SST information is translated into Sahel rainfall through the tropical Atlantic wind and that the first EOF of the zonal wind in GCMs captures very well the global SST signal.



#### **IV.4 Summary**

This chapter has identified that systematic SST changes are found in the tropical Pacific, Indian and Atlantic Oceans during boreal spring, which have significant impact of the predictability of Sahel rainfall in the absence of explicit predictions of SST. This is true in GCM forecasts based on persistence, and empirical systems using SST indices. The changes are especially focused from May to June. This has been studied by focusing on years when the June Sahel forecast is successful and the May forecast fails. The response of the ECAHM4.5 GCM to those SST changes shows the establishment of significant circulation anomalies extending from the tropical Pacific into the tropical Atlantic, consistent with the MOS system being able to better forecast Sahel rainfall anomalies. The importance of rapid SST changes during boreal spring is established to be very real for Sahel rainfall prediction, and frames an important research question for understanding and predicting SST tendencies at this time of the year.

The drop in skill is likely to have substantial implications for the usability of seasonal forecast information. This motivates further study to understand better the source of the loss of skill and the possibility of achieving a better lead time for skillful forecasts. In the next chapter, we will address these issues using coupled ocean-atmosphere models.

#### Reference chapter IV :

- Folland C. K., J. A. Owen, M. N. Ward, and A. W. Colman, 1991: Prediction of seasonal rainfall in the Sahel region of Africa using empirical and dynamical methods. *J. Fore.*, **10**, 21-56.
- Gill A. E., 1980: Some simple solutions for heat-induced tropical circulation. *Quart. J. Roy. Meteor. Soc.*, **106**, 447-462.
- Thiaw W. M., A. G. Barnston, and V. Kumar, 1999: Predictions of African rainfall on the seasonal timescale. *J. Geophys. Res.*, **104**(D24), 31589-31597.
- Wang B., I. S. Jang, and J. I. Lee, 2004: Ensemble simulations of Asian-Australian Monsoon Variability by 11 AGCMs. *J. Climate*, **17**, 802-818.
- Ward M. N., C. K. Folland, K. Maskell, A. W. Colman, D. P. Rowell, and K. B. Lane, 1993: Experimental seasonal forecasting of tropical rainfall at the U.K. Meteorological Office. *Prediction of Interannual Climate Variations*, J. Shukla, Ed., Springer-Verlag, 197-216.
- Ward M. N., 1998: Diagnosis and short lead time prediction of summer rainfall in tropical North Africa at interannual and multidecadal timescale. *J. Climate*, **11**, 3167-3191.
- WMO/ACMAD, 1998: Prevision climatique en Afrique. African Centre of Meteorological Applications for Development Rapport Tech. 927, 210 pp. [Available from ACMAD, B.P. 13184, Niamey, Niger.].

## **Chapter V : Predictability of seasonal Sahel rainfall using coupled models and lead-time improvements**

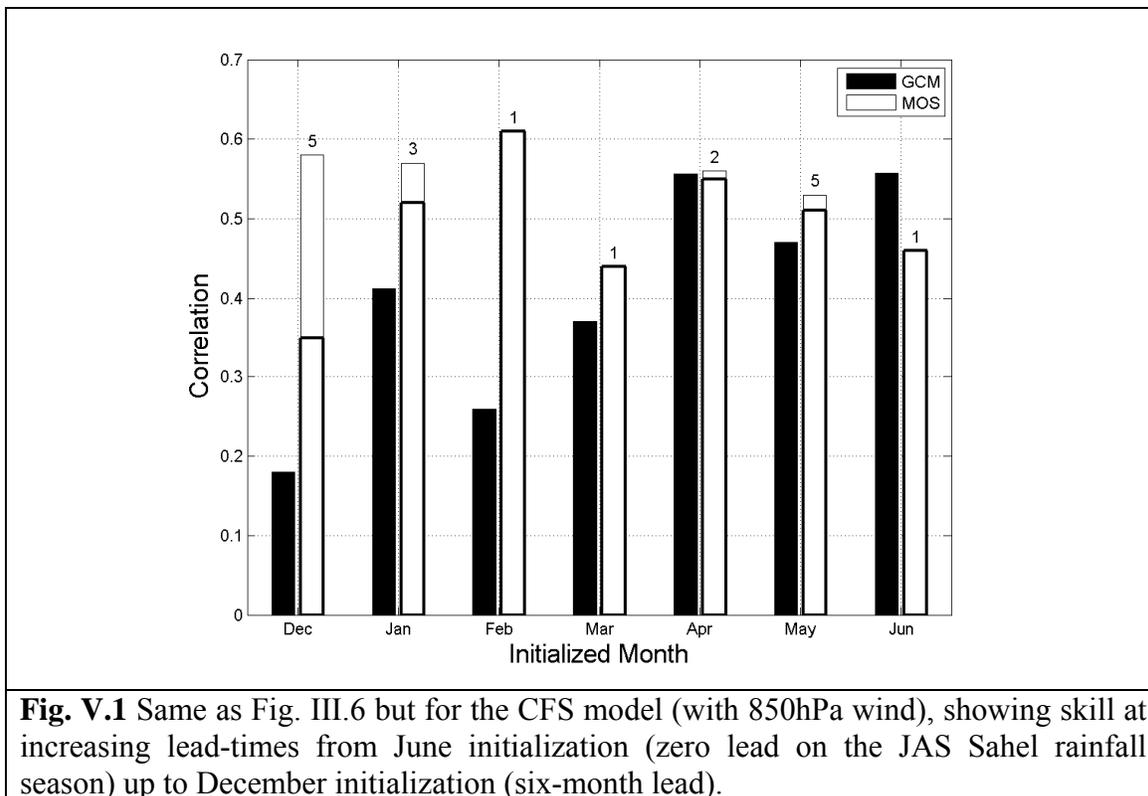
In the preceding chapter we have demonstrated the inability of a GCM to predict JAS Sahel rainfall when forced with SST boundary conditions of April and May, while it successfully forecasts with SST boundary conditions of June. This failure is consistent with the behavior of empirical SST predictors. We have also identified the SST changes, especially over the entrant and eastern tropical Pacific, that lead to the forecast failures in the GCM MOS system. It is therefore of interest to evaluate the ability of coupled ocean-atmosphere models to predict Sahel rainfall, assessing both model rainfall and using the MOS system discussed in the previous chapters, to see if skill can be gained at a longer lead-time.

It was shown in chapter III.1.b that the tropical Atlantic low-level wind was the best approach using the AGCM simulations, compared to using a range of different domains and variables. Given that the low-level tropical Atlantic wind is known to be strongly tied to monsoonal circulation and rainfall, the choice of this approach was also validated based on physical climate insight. Also, it was shown that over the period from 1970, use of just the first EOF over this domain provided a model that captured most of the prediction signal. As a first step, a similar approach is taken here with the coupled model. The first EOF of the coupled model zonal wind is used to capture the climate variability over the tropical Atlantic (including the western part of West Africa) bounded by 30°N-40°S and 60°W-10°E (the same domain as in section III.1.b.1). Partly driven by

the available archived fields, analyses used 850hPa instead of 925hPa as used with the AGCMs. For runs where both fields were available, skill was found to be insensitive to the choice of these two levels. In addition to considering the skill using just one EOF, the optimum number of EOFs (up to a maximum of five) is also explored.

### **V.1 Predictability from April/May initial conditions**

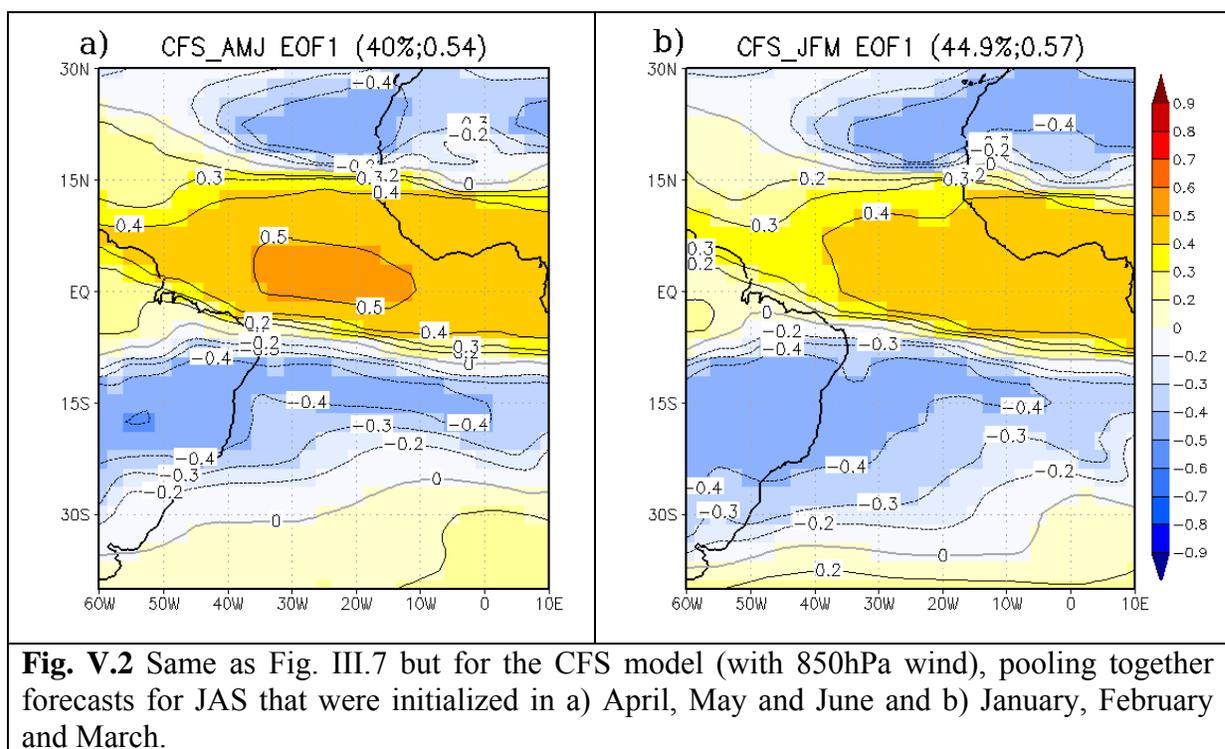
First, we focus on the results from the CFS, for which the most comprehensive set of hindcasts are available to this study, initialized at monthly resolved lead-times and available for the years 1981-2008. The skill of the model precipitation predictions and the MOS system predictions are shown in Fig. V.1. For now we focus on the right side of the panel, showing the skill of simulations that are termed as initialized in April, May and June. Table II.2 gives the exact timing of initializations, showing that strictly, the runs termed as April initialized also contain some runs from the first few days of May. Nonetheless, it is considered reasonable to compare the skill of runs that we term April initialized, with skill previously achieved by systems that used April SST. Similarly, we will compare the May initialized runs with skill achieved from May SST, and June initialized with skill achieved from June SST.



Firstly, for the GCM precipitation (shaded bars, Fig. V.1), it is noted that runs initialized in June, May and April all display substantial skill, at a level that is comparable with the best simulations from observed SST (Fig. III.6). Furthermore, the skill does not decay, but rather is maintained as the lead-time increases. The most likely interpretation of the slight dip in skill in May, compared to June and April initializations, is considered due to sampling error rather than a repeatable phenomenon.

Applying the MOS system (open bars in Fig. V.1) does not substantially alter the skill levels, but rather, skill is generally at approximately the same level as achieved with the model rainfall. However, these MOS results do demonstrate that use of the low-level wind (mostly EOF1) is again effective for Sahel prediction, here in a coupled forecast system. Furthermore, as diagnosed subsequently in sections V.2 and V.3, the skill of the

MOS system and the model rainfall are found to contain somewhat different aspects of the rainfall variability. The spatial pattern of the EOF1 of the low-level wind from the CFS runs initialized in June, May and April is shown in Fig. V.2a, confirming a similar large-scale spatial pattern compared to the ones in Fig. III.7. Again, consistency is greatest south of the Equator, with more model-specific features emerging north of the Equator, especially over West Africa itself.

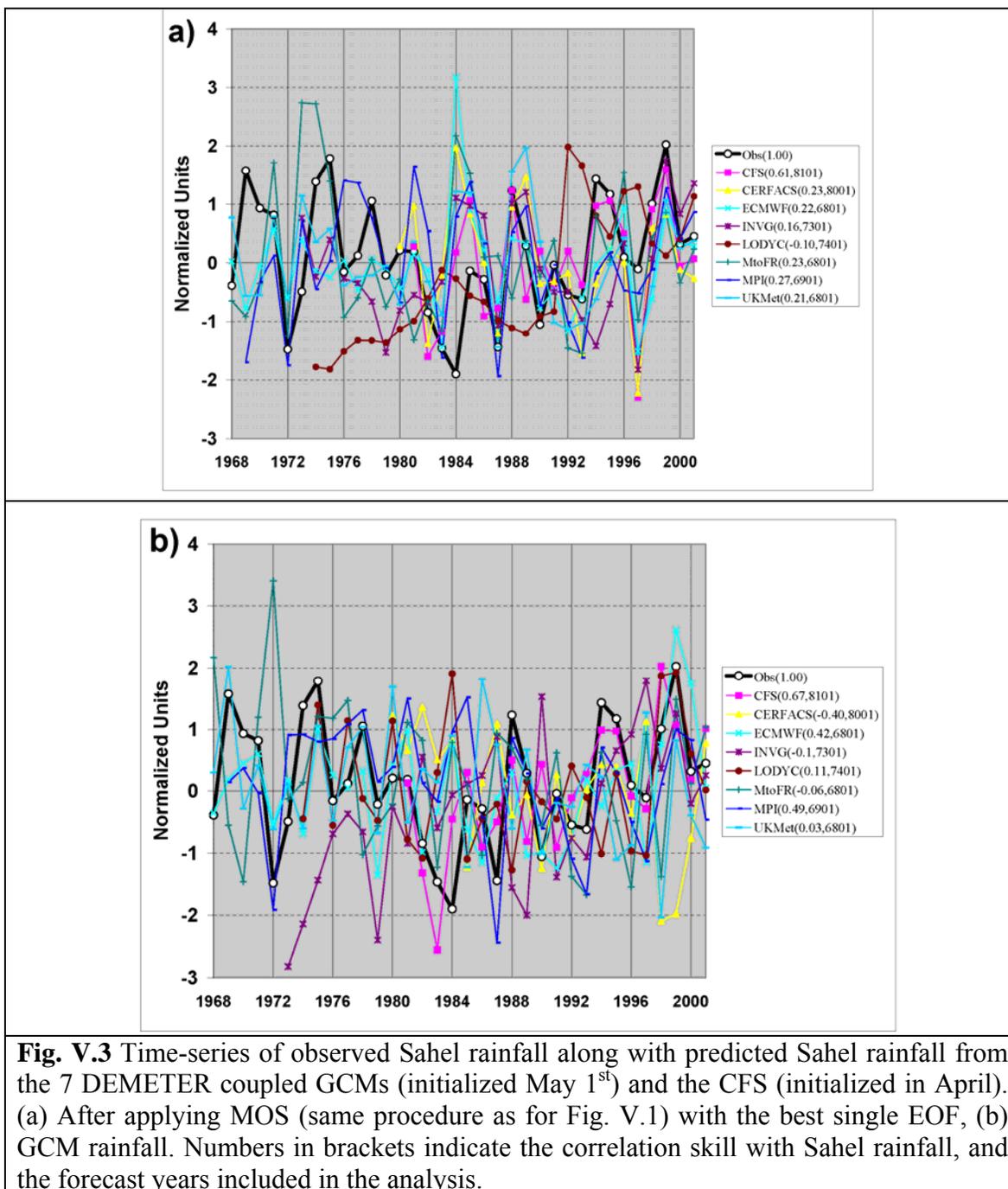


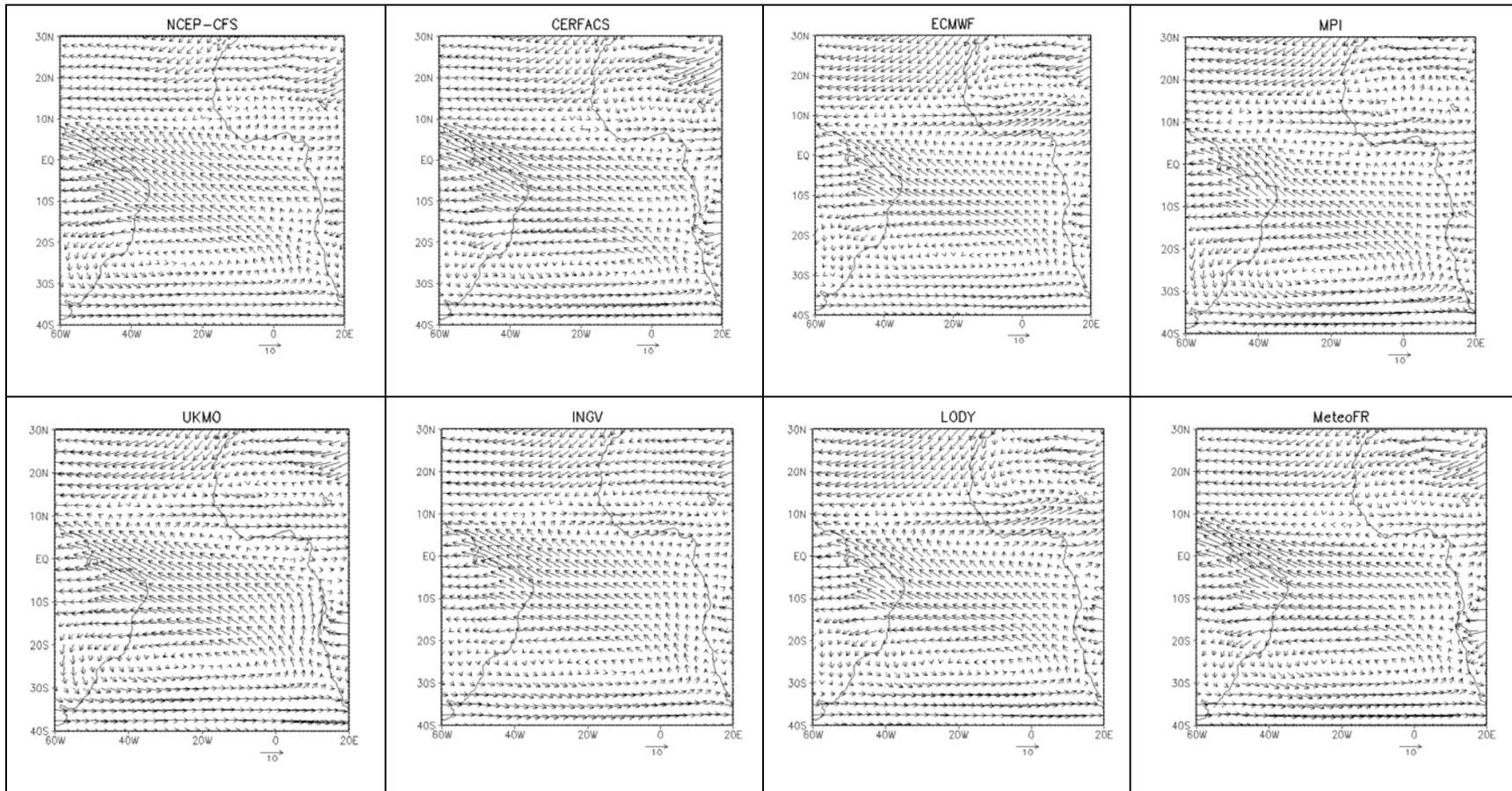
The forecast skill shown in Fig. V.2 is reproduced when the Lamb Sahel rainfall index is used as the target predictand. For example, for the April, May and June initializations, the average MOS EOF1 skill using GHCN is 0.51, whereas using the Lamb index, the average correlation skill is 0.56. The conclusion that is drawn is that the CFS contains skill comparable to that achieved with observed SST, and appears to have broken through

the problem of losing skill substantially when forecasts are made using April information as compared to June information.

The DEMETER coupled GCMs were all initialized around May 1<sup>st</sup>, so their Sahel rainfall skill can be compared with that achieved by systems using April information, and specifically here, the April initialized runs of the CFS (Fig. V.3c compares the low-level climatological tropical Atlantic winds of all models, showing good basic consistency). The model rainfall and the MOS predicted rainfall from low-level tropical Atlantic EOF1 were analyzed in the same way as reported above for the CFS. For all DEMETER experiments, 2001 is the last forecast year. MOS predicted rainfall (Fig. V.3a) and GCM rainfall (Fig. V.3b) are plotted for all available years 1968-2001 for the DEMETER models, and for the April initialized CFS. Fig. V.3 also indicated the years available for each model, and the correlation skill for each model. Most models (five of seven) show very low levels of skill in their Sahel rainfall predictions ( $r < 0.20$ ), with five of seven showing small increases in skill when the MOS is applied, but with levels of correlation skill still generally below  $r = 0.3$ . There are two models (ECMWF,  $r = 0.42$  and MPI,  $r = 0.49$ ) that emerge from the set as most promising in their Sahel rainfall predictions. Though these skill levels are only comparable or slightly higher than that generally achieved using persistence approaches, the results do support the assertion that coupled models are becoming important contributors to the Sahel rainfall prediction problem. Two aspects motivate a more detailed analysis of the CFS experiments. Firstly, the skill levels make a jump that separates them clearly from the other models. Secondly, due to the start-times available for hindcasts, only the skill of models from initialization around

the May 1<sup>st</sup> can be considered for the DEMETER runs, whereas for the CFS, a much larger set of hindcasts are available, permitting skill levels to be established and diagnosed more robustly across varying lead-times.

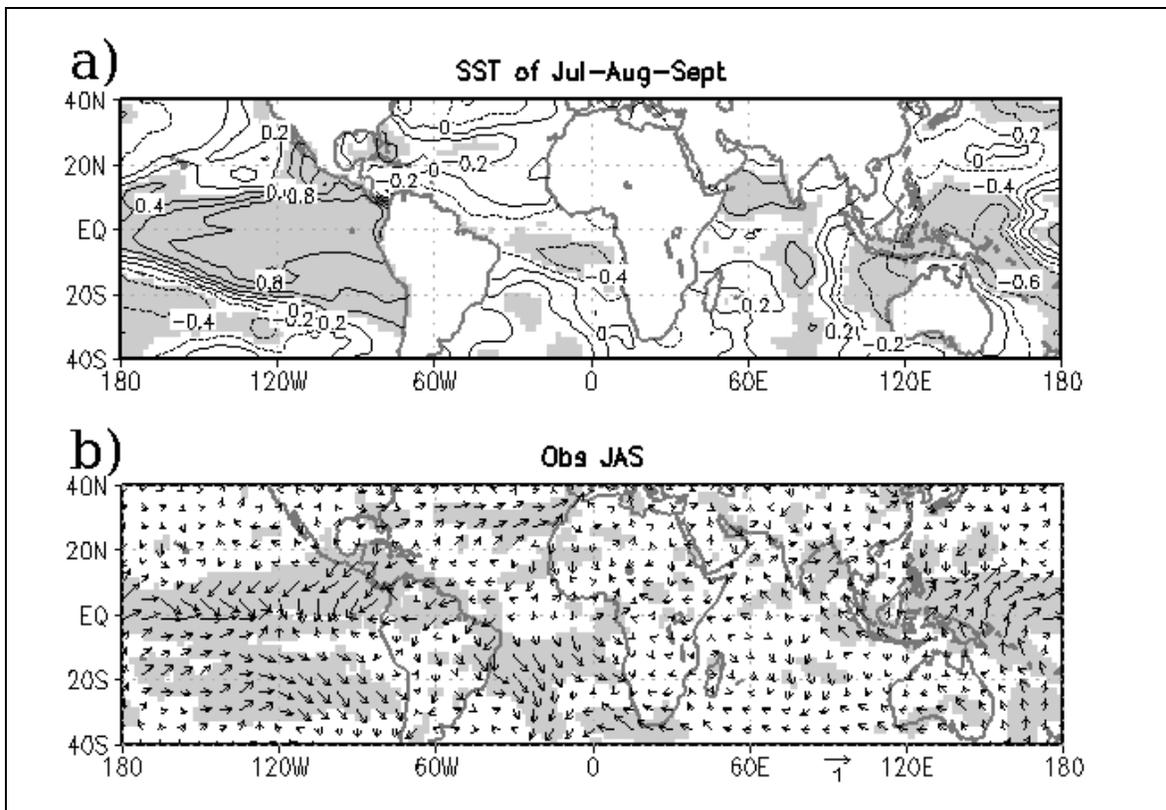




**Fig. V.3c :** The 850 hPa climatological wind simulated by CFS and DEMETER coupled-GCMs initialized in April / early May, depicting the low level monsoon flow during JAS 1981-2001.

The level of skill achieved with the CFS suggests that the model has been able to capture the evolution of the coupled ocean-atmosphere system through the period April to June and arrive at a correct representation of the atmospheric teleconnection response in the tropical Atlantic wind field and indeed the model output Sahel precipitation itself. Verification of this hypothesis is now sought. The following diagnostic analysis is designed to lend support to the skill findings, and confirm that the results are rooted in recognizable coupled ocean-atmosphere developments.

In section IV.2, when the persistence approach led to Sahel forecast failures in the period 1968-2001, the key SST evolution from April to June was shown to be associated with developments in the central and eastern Equatorial Pacific. To diagnose the ability of the CFS to capture such evolution of SST and develop the key teleconnection structures in the tropical Atlantic, the following approach is taken. We first calculate the observed teleconnection between JAS Niño 3 and the pattern of JAS SSTs (Fig. V.4a) and JAS near-surface (the 10-meter level) winds (Fig. V.4b). These represent the teleconnection structures that need to be in place in the coupled model forecasts, especially the linkages from the tropical Pacific to the tropical Atlantic. First we address the SST aspects, and then address the wind aspects.



**Fig. V.4** : Observed JAS Niño3 index correlation 1981-2008 with (a) observed JAS SST and (b) reanalysis JAS near-surface u and v wind. A vector is formed using the u-correlation (zonal component of the vector) and the v-correlation (meridional component of the vector). Correlations significant at the 95% level are shaded (for the wind map, shading is applied if either the u or the v is significant).

First, to highlight the failure of the persistence approach for this climate forecast problem, we correlate the JAS Niño 3 index to each SST grid box from April to June (Fig. V.5). While the June map (Fig. V.5a) contains much of the JAS teleconnection structure achieved with JAS Niño 3 (Fig. V.4a), the maps for May and April are dramatically weaker in the central and eastern tropical Pacific. The SST correlation over the Niño 3 region is strong in June ( $r=0.8$ , Fig. V.5a) weakens in May (Fig. V.5b) and drops drastically to about  $r=0.4$  and is only significant over a much smaller area during April (Fig. V.5c).

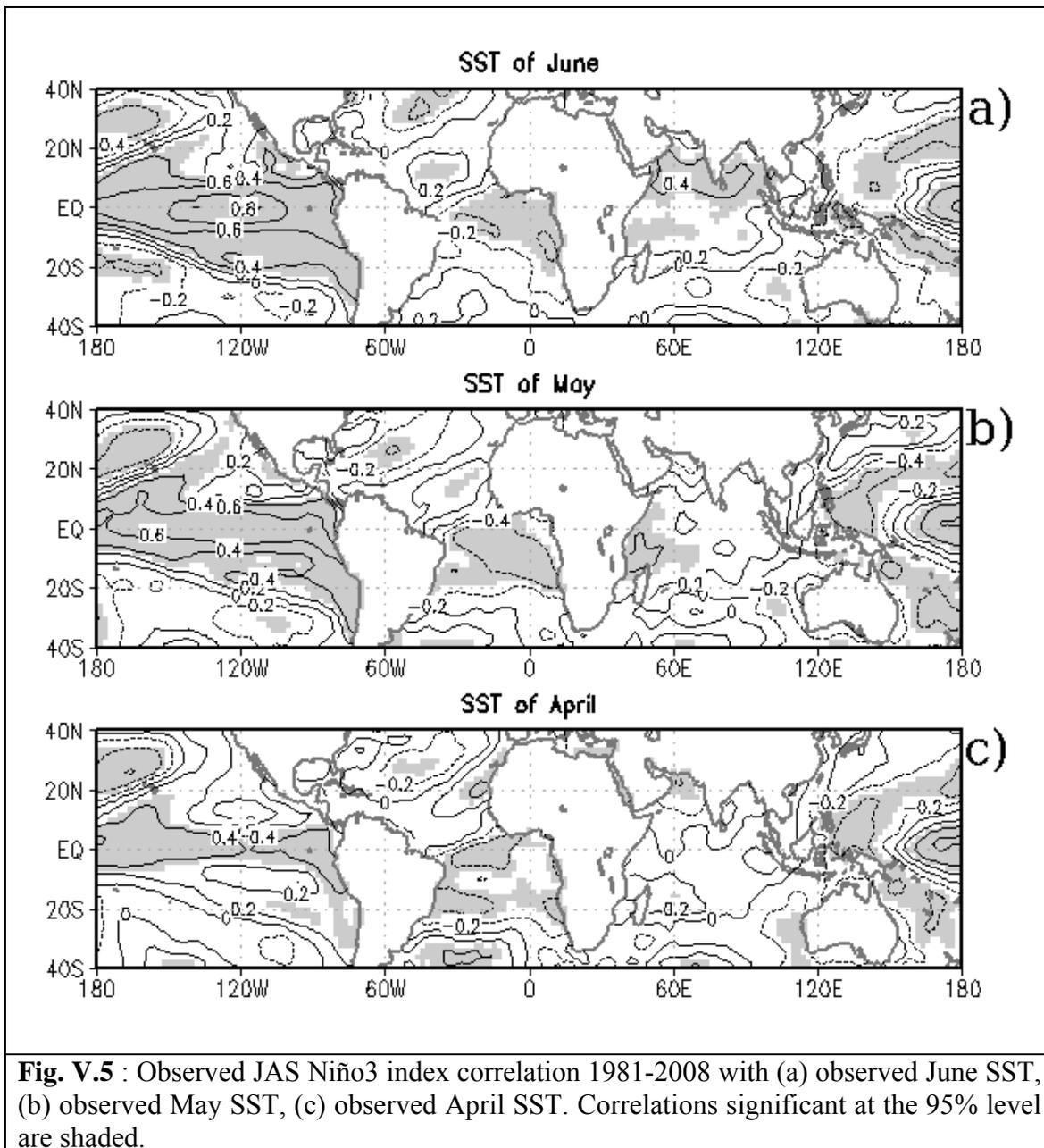


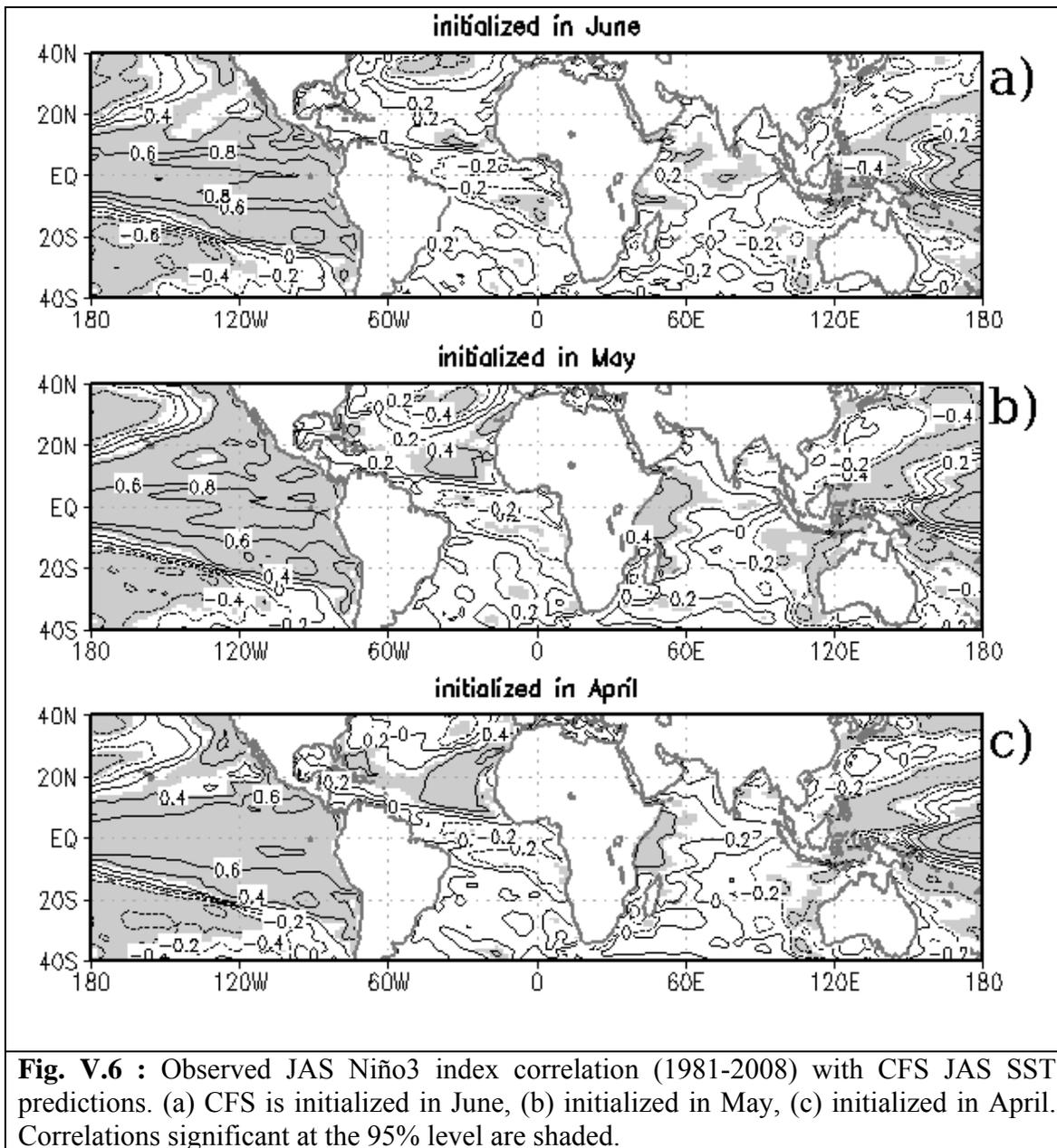
Figure V.4a shows the known teleconnection between El Niño (La Niña) and warmth (cool conditions) in the tropical North Atlantic (Enfield and Mayer, 1997; Lau and Nath, 2001). The panel with April SST (Fig. V.5c) contains a much weaker signal for this tropical North Atlantic teleconnection, which indicates that the warmth is usually not already in place in the North Atlantic in April, in those years that have strong El Niño

presence in JAS (and vice-versa, cool conditions are not in place in those years when La Niña is present in JAS). Figure V.4a also shows the teleconnection between El Niño (La Niña) and warmth (cool conditions) in the northwestern Indian Ocean. Again, Fig. V.5c shows that in April, the warmth is not present in the northwestern Indian Ocean, in those years when July-September sees the presence of El Niño (and vice-versa, cool conditions are not in place in those years when La Niña is present in JAS). So these Indian and North Atlantic teleconnection features can be considered as ones that develop during boreal spring along with Niño 3 development. The tropical North Atlantic development implies that during an El Niño, the North Atlantic actually warms. In terms of local SST forcing on Sahel rainfall, this is consistent with a wetter Sahel. The warming can therefore be considered a negative feedback on the large-scale forcing which generates atmospheric teleconnection structures from the Pacific to the Atlantic that favor drier conditions in the Sahel (Janicot et al., 1996). This was noted in the coupled mode connecting ENSO to Sahel rainfall in Ward (1998), and may be one of the reasons why GCMs have difficulty with representing the impacts of ENSO directly on the rainfall, since they need to correctly balance these processes in the tropical Pacific and Atlantic domains.

A further aspect of Fig. V.4a is the negative correlation with tropical South Atlantic SST. This is not a widely reported phenomenon in the literature, and when noted, has been not as easily interpreted in terms of physical mechanism (Enfield and Mayer, 1997). Its presence here may be amplified through chance sampling in this period. Further uncertainty is cast by the fact that the teleconnection shows as strongest in Fig. V.5, which suggests tropical South Atlantic SSTs in May are a precursor of ENSO

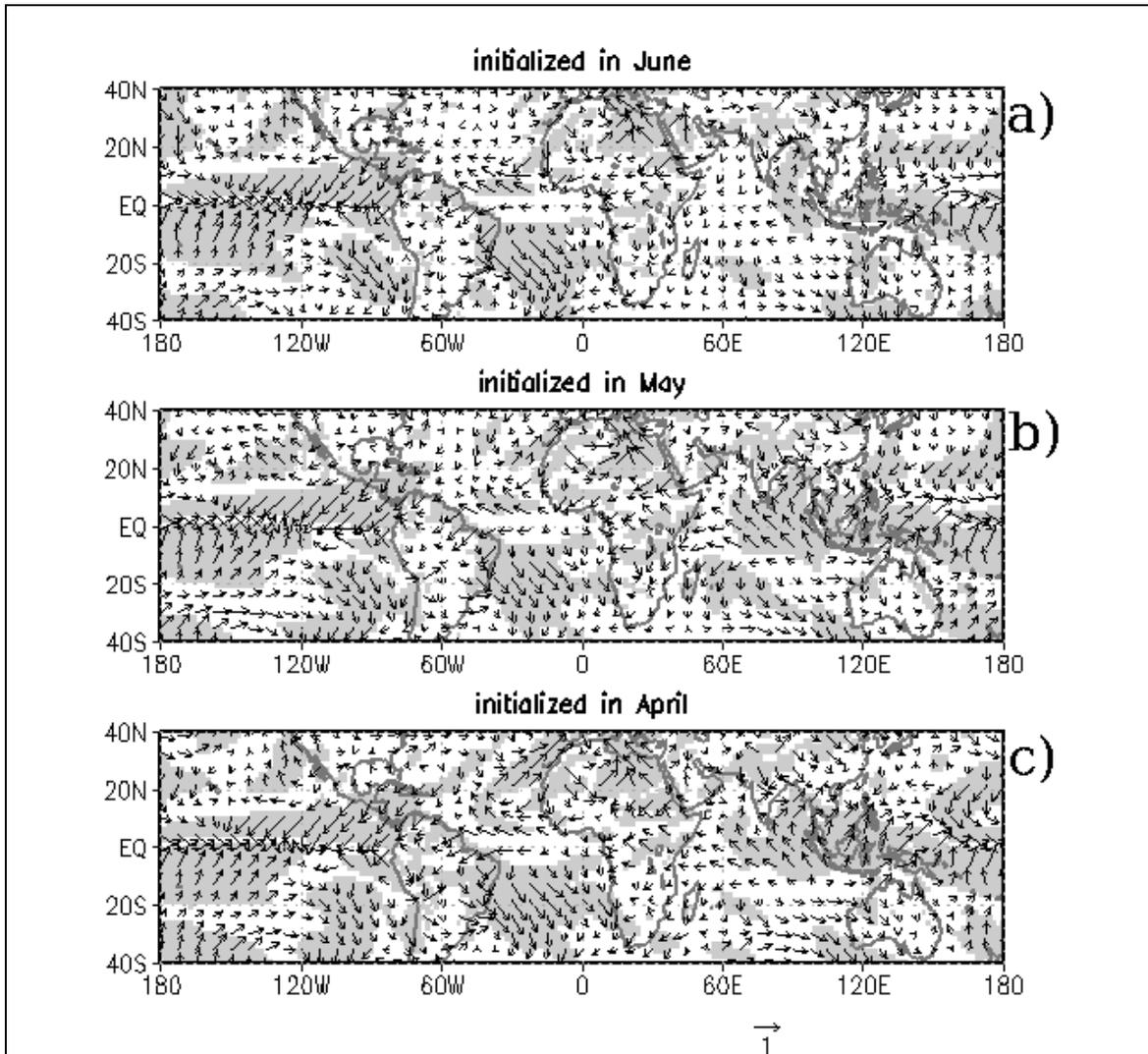
development into July-September. Given the difficulties in interpretation, less emphasis is placed on this SST teleconnection structure, though it could be worthy of future investigation.

Next, we consider the extent to which CFS forecasts contain teleconnection structures with the observed JAS Niño 3. The observed JAS Niño 3 is correlated with the predicted JAS SST field that is generated by each of the initialization times. First, for the tropical Pacific, it is clear that runs initialized in June (Fig. V.6a), May (Fig. V.6b) and April (Fig. V.6c) all successfully represent the evolution of tropical Pacific SST to match the observed Niño 3 in July-September and its teleconnection across the Pacific basin. The performance represents a substantial improvement upon persistence, as evidenced by comparing Fig. V.5c (April persistence used to predict JAS SST) with Fig. V.6c (April initialized model prediction of JAS SST).



The CFS also reproduces aspects of the SST development in the western Indian Ocean and tropical Atlantic Oceans. The aspects in the North Atlantic and western Indian Ocean are actually least clear in the June initialized results (Fig. V.6a), suggesting the longer lead may allow the GCM to better develop the teleconnections between a developing El Niño (La Niña) in the tropical Pacific and warming (cooling) in the North Atlantic and

the western Indian Ocean (Figs. V.6b,c). There is also a tendency for the negative correlation to be present in the South Atlantic, such that Figs. V.6b,c have all the basic elements in the tropical Atlantic and western Indian Ocean that are found in the observed teleconnection structure (Fig. V.4a).



**Fig. V.7** : Same as Fig. V.5 but for CFS near-surface wind predictions. A vector is formed using the u-correlation (zonal component of the vector) and the v-correlation (meridional component of the vector). Shading is applied if either the u or the v correlation is significant at the 95% level.

These results on representing the evolution of SST, encourage the expectation that the atmospheric teleconnection evolutions may also be well captured. This is largely

confirmed in Fig. V.7. For example, Fig. V.7c shows that CFS runs initialized in April produce JAS Pacific wind fields that correlate strongly and consistently with the observed Niño3 JAS index. The levels of teleconnection representation are as strong for the April initializations (Fig. V.7c) as they are for the May (Fig. V.7b) and June (Fig. V.7a) initializations. The pattern is comparable to that achieved between the observed Niño 3 and the observed JAS reanalysis winds (Fig. V.4b).

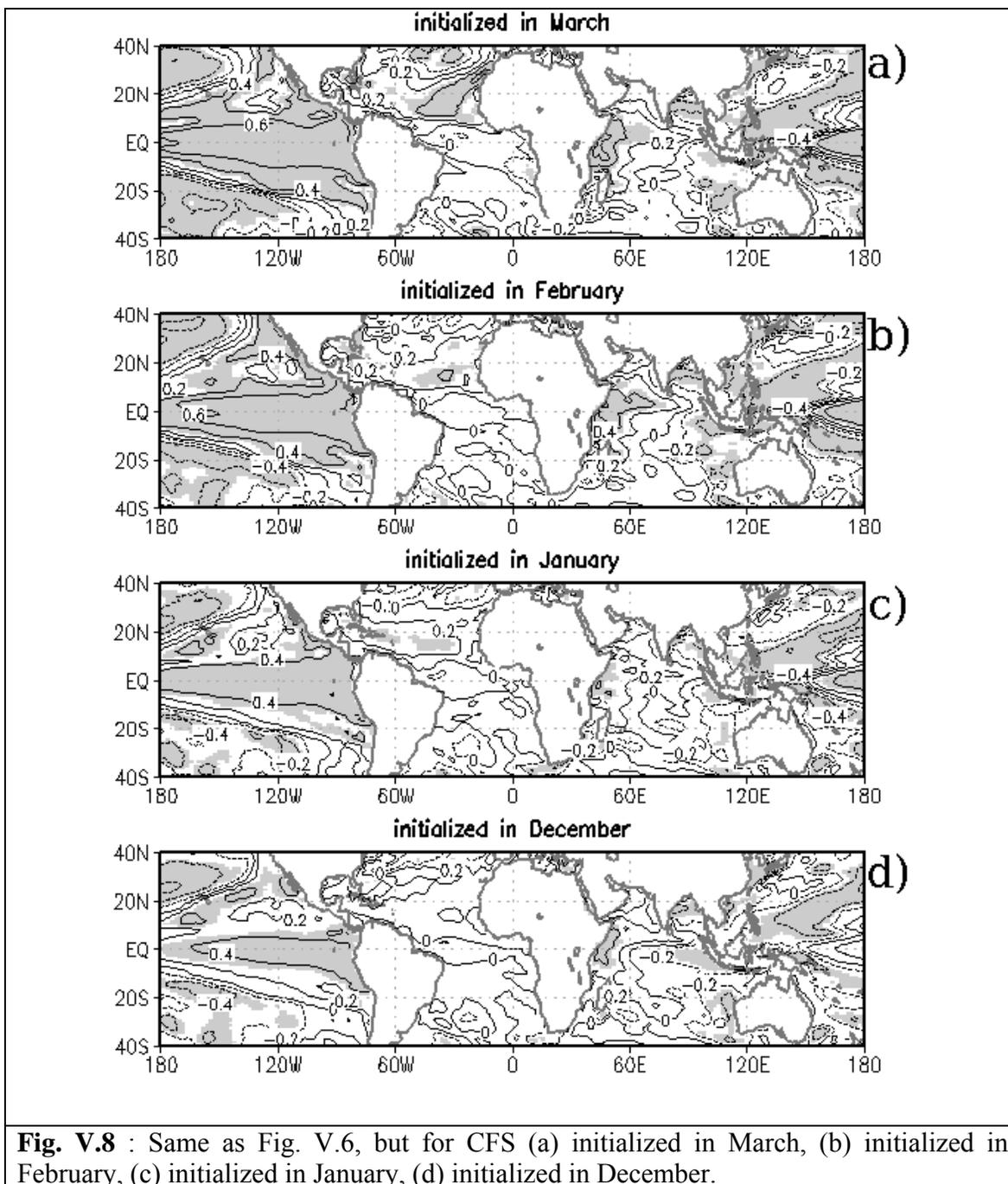
In support of the good skill in predicting Sahel rainfall (Fig. V.1), JAS teleconnection structures in the tropical Atlantic are also reproduced in the model runs initialized in April, May and June. In other words, the runs initialized during boreal spring produce forecast JAS tropical Atlantic wind fields that correlate with the observed JAS Niño 3 in ways that are largely consistent with the observed teleconnection structures. For the near-surface zonal wind, opposite sign correlations are found in the tropical North and tropical South Atlantic, with the tip of NE Brazil being the latitude of the change of sign in most maps (Fig. V.7). The results in Figs. V.6 and V.7 are interpreted as strong evidence that the results of the good skill in Sahel prediction are indeed connected to very large-scale Pacific and Atlantic developments that are being successfully simulated by the CFS. It lends considerable support to the hypothesis that the CFS model is able to forecast the key developments from April for the predictable part of the Sahel rainfall variance that is related to ENSO.

The CFS forecasts initialized in April project tropical Atlantic wind conditions for July-September that correlate strongly with the observed JAS Sahel rainfall index. A major part of this is achieved through developments in the tropical Pacific and tropical Atlantic

that are consistent with the observed climate system teleconnections across the Pacific-Atlantic sector. This suggests the coupled model is able to skillfully project forward the coupled system over this challenging period in the Pacific-Atlantic sector. Indeed, there is no decline in Sahel rainfall prediction skill when lead-time increases from zero months (June initialization) to two months (April initialization). This motivates investigation of the additional CFS runs that are available out to a lead-time of six months (December initialization).

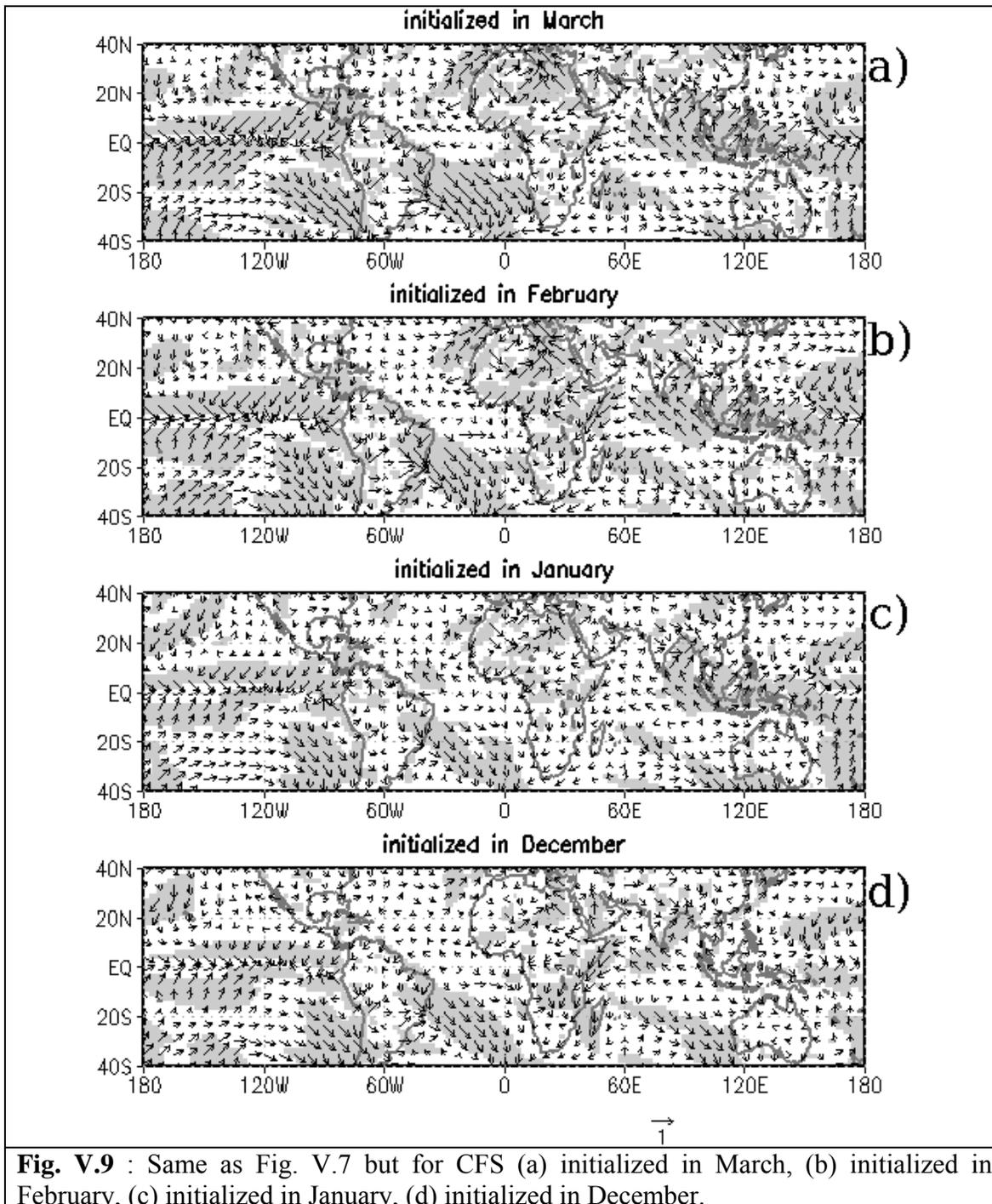
## **V.2 Further analysis of the CFS predictability and comparisons at up to 6 month lead-time**

The skill of the model precipitation and MOS system are plotted in Fig. V.1 for all runs for which a forecast for JAS is available from the CFS. The longest lead-time available corresponds to approximately 6 months before the JAS rainfall season, being initialized in December and available from January the 5<sup>th</sup>. Figure V.1 shows that generally, for initialization before April, the model precipitation skill falls quite substantially compared to the shorter lead-time forecasts. However, for the MOS on the regional circulation, skill levels are maintained remarkably stable just using EOF1, even up to the 5-month lead-time (January initialization). The leading EOF spatial pattern is very stable for these initializations and similar to that found for the shorter lead-time initializations (compare Fig. V.2a and Fig. V.2b). For December initialization, the skill drops to  $r=0.35$  when using just EOF1 in the MOS system. However, comparable skill levels are recovered using the leading 5 EOFs ( $r=0.58$ ).



To confirm that these long-lead forecast skill improvements are reflected in large-scale model fields, the analyses with observed JAS Niño 3 are repeated using the predicted JAS SST (Fig. V.8) and circulation (Fig. V.9) for each of the initialization months. For example, Fig. V.8b shows that even for forecasts initialized in February, predicted JAS

SST in the central and eastern tropical Pacific correlates over a wide area at over 0.5 with the observed JAS Niño 3, while predicted SST in the tropical North Atlantic and western Indian Ocean continues to correlate positively with the observed JAS Niño 3, indicating representation of these teleconnection structures at long lead-time.

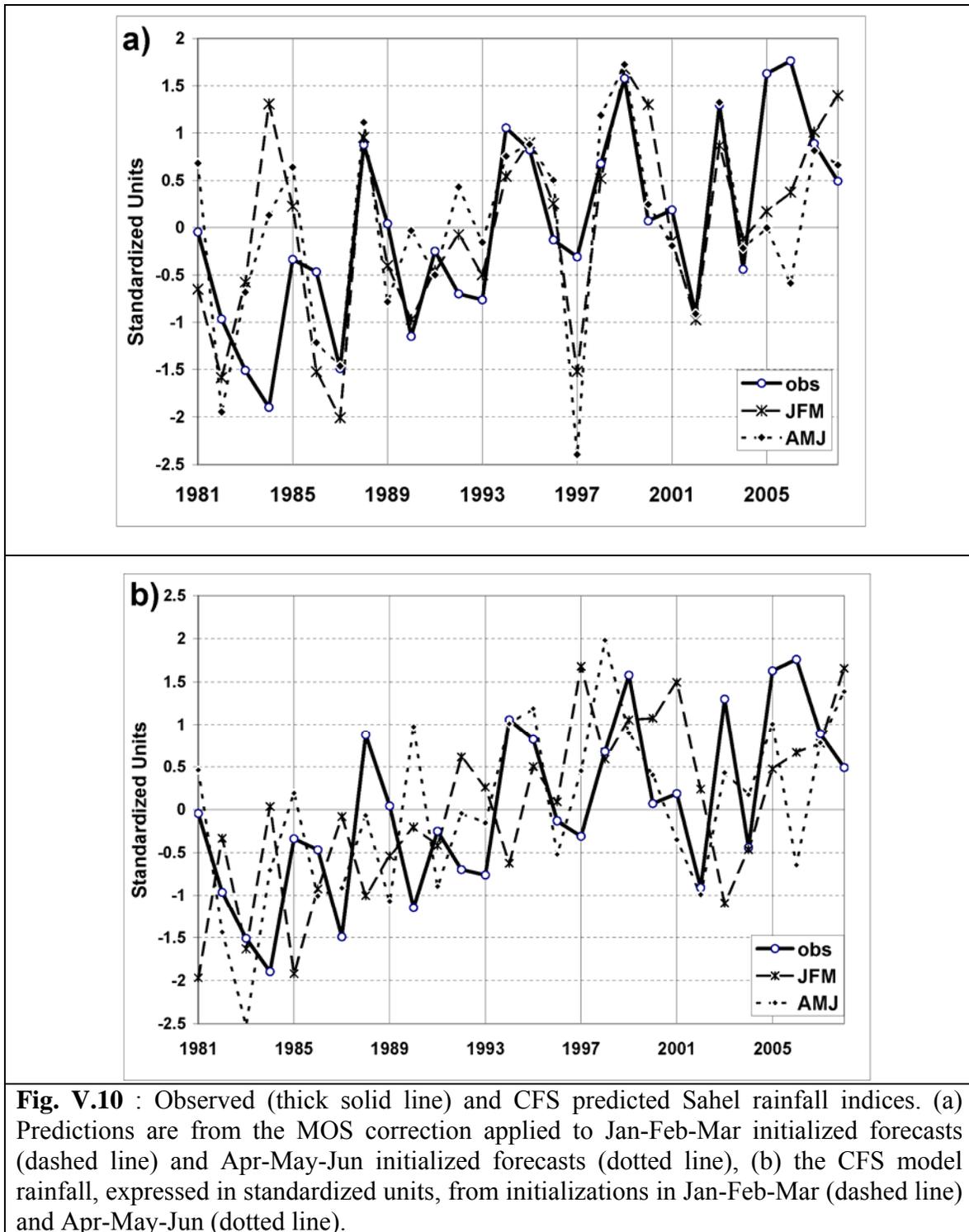


The predicted JAS near-surface winds from December initialization (Fig. V.9d) still show a strong and widespread teleconnection structure with the observed JAS Niño 3 SST. However, the pattern is weakening, and especially so to the north of the Equator (compare Figs. V.9d and V.9a). This is true in both the Pacific and in the tropical Atlantic, and may reflect the greater difficulty in recovering the MOS skill for Sahel rainfall using December initializations. However, for initializations from at least January onward (Figs. V.9a,b,c), the teleconnection structure response of the CFS supports the assertion that the model runs contain useful information at long-lead for Sahel rainfall. Even from December initialization, information is still present that allows skilful MOS predictions, even if the information is in a somewhat different form, with less of the relevant information found in fields north of the Equator.

Another aspect of Figs. V.6 and V.8 is to probe the question of whether the predicted SST fields may serve as more effective MOS predictors than the tropical Atlantic wind fields. Such a possibility was not an option with the persisted SST runs in chapter IV, where different variables in an atmosphere-only model were explored for making MOS predictions of Sahel rainfall. A detailed investigation is beyond the scope of this paper, but a check was made on whether there were any obvious gains to be achieved from substituting SST for wind in the MOS system. One of the challenges in trying to use the SST as a MOS predictor is that SST information is likely entering the tropical Atlantic wind field and West African monsoon from multiple geographic sources, even though a primary source in the 1981-2008 period appears to be from the tropical Pacific. The advantage of using the wind field in the tropical Atlantic is that it effectively integrates the relevant SST information from around the globe, and provides a distinct geographic

domain over which to construct the MOS predictors. Various SST domains with physical reasoning were tried. None of the domains improved upon the skill of the low-level tropical Atlantic wind field. One of the best SST options was to take the global tropics SST (40°N-40°S). For this experiment, the average correlation skill score for SST for Jan-Feb-Mar initializations was 0.49, compared to 0.54 for the tropical Atlantic wind field. For Apr-May-Jun initializations results were very similar (0.49 versus 0.52). The main conclusion is that the low-level wind field remains a robust integrator of relevant forecast information in this coupled model setting. With a coupled model, the variable SST is a candidate for use in a MOS, but it does not here usurp the tropical Atlantic low-level wind field's position as the favored variable on which to perform MOS.

The skill levels in Fig. V.1 from January to June initializations do move around modestly from month to month. A substantial part of that is considered to be due to sample size. To partly confirm that hypothesis, and also to provide a year-by-year summary of the forecast system, forecasts have been averaged together across initializations in Jan-Feb-Mar and across initializations in Apr-May-Jun. The model precipitation and MOS results (using EOF1) are shown as time-series plots in Fig. V.10. The general consistency of the MOS forecasts for the shorter and longer lead-time can now be seen (Fig. V.10a). The skill is comparable for the pooled Jan-Feb-Mar ( $r=0.57$ ) initialized forecasts and the pooled Apr-May-Jun ( $r=0.54$ ) initialized forecasts.



**Fig. V.10** : Observed (thick solid line) and CFS predicted Sahel rainfall indices. (a) Predictions are from the MOS correction applied to Jan-Feb-Mar initialized forecasts (dashed line) and Apr-May-Jun initialized forecasts (dotted line), (b) the CFS model rainfall, expressed in standardized units, from initializations in Jan-Feb-Mar (dashed line) and Apr-May-Jun (dotted line).

Inspection of Fig. V.10 gives the impression that the MOS contains substantial skill on the interannual timescales (Fig. V.10a), whereas much of the correlation skill achieved by

the model precipitation from the CFS is derived from the upward trend in both the observation and the CFS rainfall predictions (Fig. V.10b). This is indeed confirmed in Table V.1 which compares the correlation skill for the series in Fig. V.10 before and after detrending. After detrending, the MOS skill is maintained at substantial levels ( $r=0.39$ ,  $0.33$  for initializations in AMJ and JFM respectively). For the model precipitation predictions, after detrending there is no skill for JFM initialized forecasts ( $r=-0.22$ ), while skill drops substantially from  $r=0.57$  to  $r=0.31$  when predictions from the AMJ initializations are detrended.

**TABLE V.1** : Comparing the CFS forecast skill (correlation) 1981-2008 for July-September Sahel rainfall, from Apr-May-Jun initializations (AMJ) and Jan-Feb-Mar initializations (JFM). Results are shown for GCM rainfall and for the MOS system using low-level tropical Atlantic wind EOF1. Results are also shown when all time-series are de-trended over the 1981-2008 period.

	<b>JFM</b>	<b>AMJ</b>
<b>Model Precipitation Predictions</b>	0.24	0.57
<b>Model Precipitation Predictions after</b>	-0.22	0.31
<b>MOS EOF1 Predictions</b>	0.57	0.54
<b>MOS EOF1 Predictions after Detrending</b>	0.33	0.39

The levels of skill achieved by the MOS on the detrended series are slightly higher than the magnitude of the correlation between the observed JAS Niño3 and Sahel rainfall for the period ( $r=-0.36$ , after series are detrended). For comparison, using Niño 3 as a predictor for Sahel rainfall in this period loses skill very rapidly as lead-time increases, in a manner consistent with that described in Ward (1998). The Niño3 monthly correlations with JAS Sahel rainfall for the 1981-2008 period are (all series detrended),  $r=-0.34$  (Jun),  $r=-0.28$  (May),  $r=-0.26$  (Apr),  $r=-0.17$  (Mar),  $r=-0.08$  (Feb),  $r=-0.01$  (Jan). These results reinforce the idea that SST persistence provides some interannual skill in

predicting Sahel rainfall based on ENSO at short lead-times. The results also reinforce the idea that persistence performs very poorly at longer lead-times. This highlights the progress of achieving, even after detrending, levels of skill of 0.33 (Jan-Mar initialization) and 0.39 (April-June initialization) with the CFS MOS system.

### **V.3 The trend component of Sahel rainfall 1981-2008 in observations and CFS seasonal predictions**

It is well known that Sahel rainfall has a strong low frequency component of variance (Nicholson, 1980) which has been expressed in recent decades in an extended relatively dry period through the 1970s and 1980s, with a widely document modest recovery after the 1980s (Nicholson, 2005; Herrmann et al., 2005). In terms of the shift of rainfall from the 1950s to the 1980s, the recent upward trend is often considered relatively modest. However, it is clearly present visually for the 1981-2008 period for which the CFS forecasts are available (e.g Fig. I.1b, Fig. V.10), and it is confirmed to be a highly significant statistical trend, measured in terms of linear regression or composite difference. It is therefore considered a real physical aspect of the climate system, likely to have expression in large-scale ocean-atmosphere fields, and it is of interest to consider how the CFS forecast system has performed in this context. Following analysis and inspection of the Sahel rainfall series, the period has been divided into the relatively dry early set of years (1981-1993) and the more recent wetter years (1994-2008), for the purpose of constructing composite differences. Table V.2 summarizes such differences for the observed Sahel series, the MOS predictions using EOF1, and the CFS rainfall predictions, providing quantification and statistical significance for the trends in these quantities. In the observation, the difference is 1.244 standardized units, which is

significant at the 99% level using a t-test. This trend is well captured in the CFS model rainfall in both JFM and AMJ initializations. Both show a strong trend, slightly larger for the JFM initializations, with a difference of 1.166 and 1.034 standardized units respectively, with both significant at the 99% level (Table V.2).

**Table V.2.** Composite difference of July-September values, 1994-2008 MINUS 1981-1993.

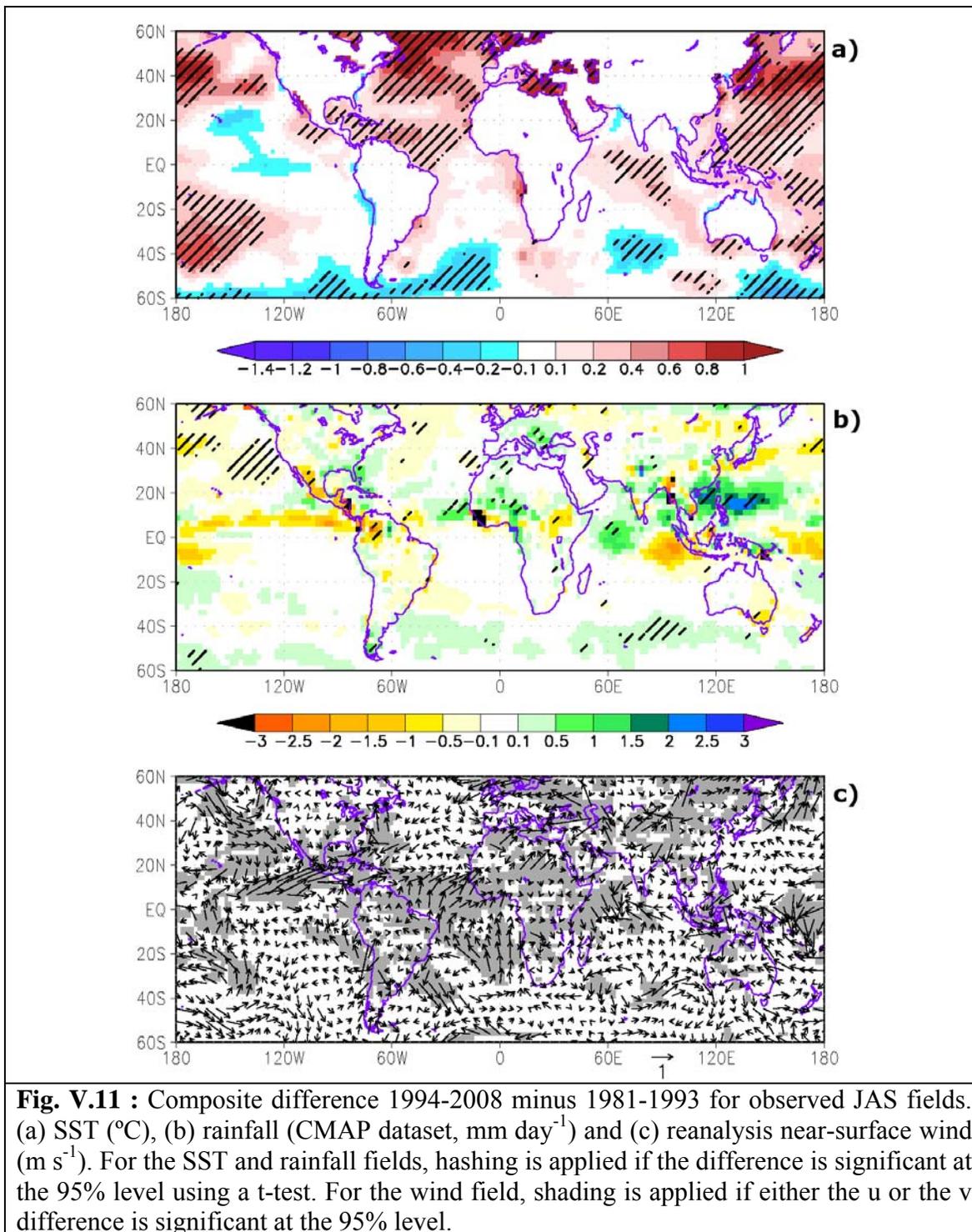
Results are for observed and CFS-predicted rainfall and SST indices. Predictions are initialized in Jan-Feb-Mar (JFM) and Apr-May-Jun (AMJ). Rainfall results are in standardized units, for observed, CFS model predictions, and predictions using the MOS on the low-level tropical Atlantic wind field. The SST index is the difference between the North and South Atlantic (AtlN-S, defined in text), and units are degrees Celsius. Statistical significance is estimated using a t-test.

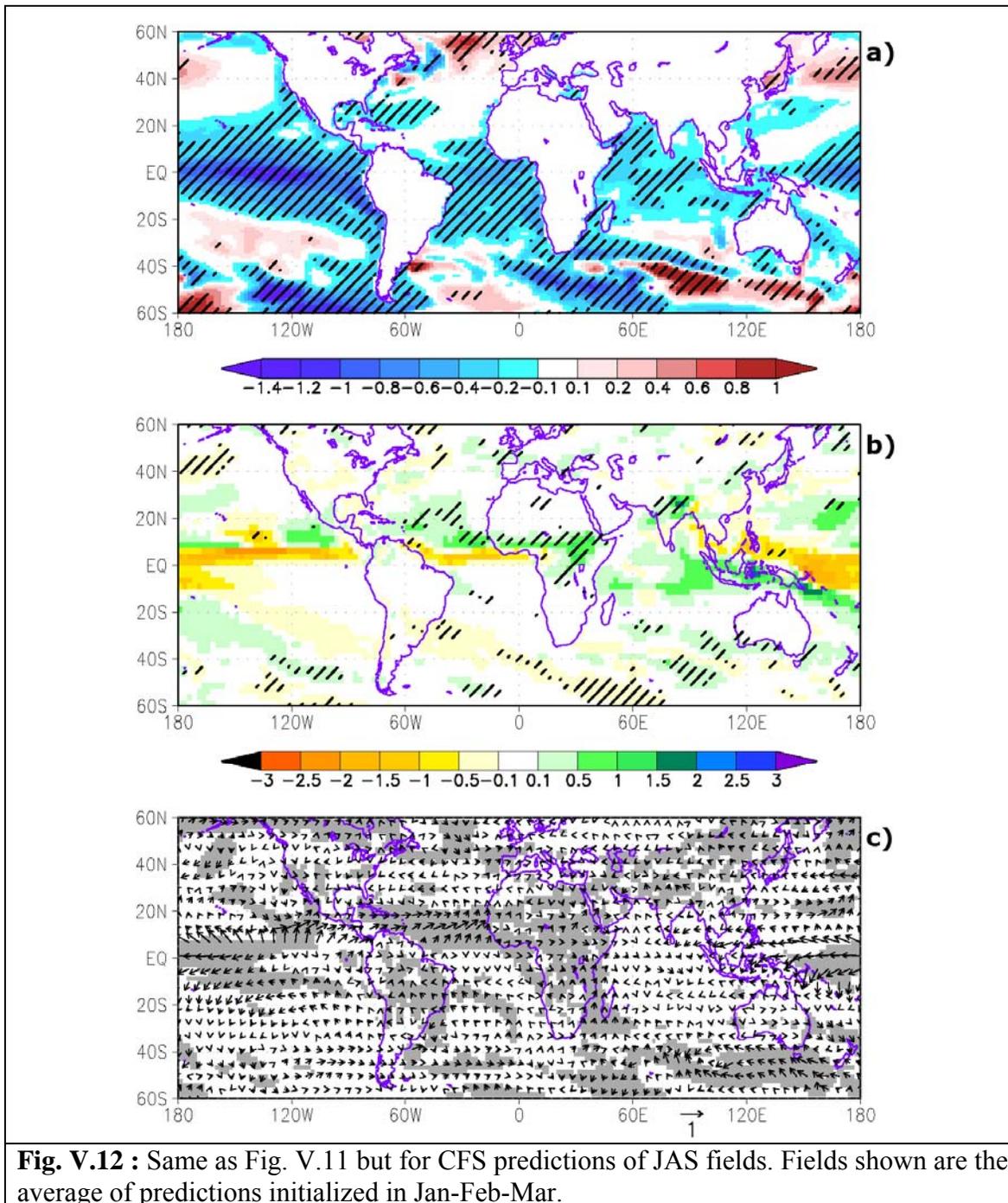
	PRECIPITATION				SST			
	Obs	GCM		MOS Prediction		North – South Atlantic		
		CFS	CFS	CFS	CFS	Obs	CFS	CFS
	JFM	AMJ	JFM	AMJ		JFM	AMJ	
Diff.	1.244	1.1663	1.0338	0.8985	0.5444	0.316	0.327	0.288
p-value	0.0005	0.0005	0.0001	0.0090	0.0084	0.0003	0.0001	0.0007

The composite analysis also confirms a positive trend in the MOS predictions using EOF1, though the trend is weaker, with a difference between the two periods of 0.898 and 0.544 standardized units, for respectively JFM and AMJ initializations. The implication is that CFS wind EOF1 does contain an upward trend, but of a magnitude that leads to a somewhat underestimated upward trend in Sahel rainfall when the MOS prediction system is applied.

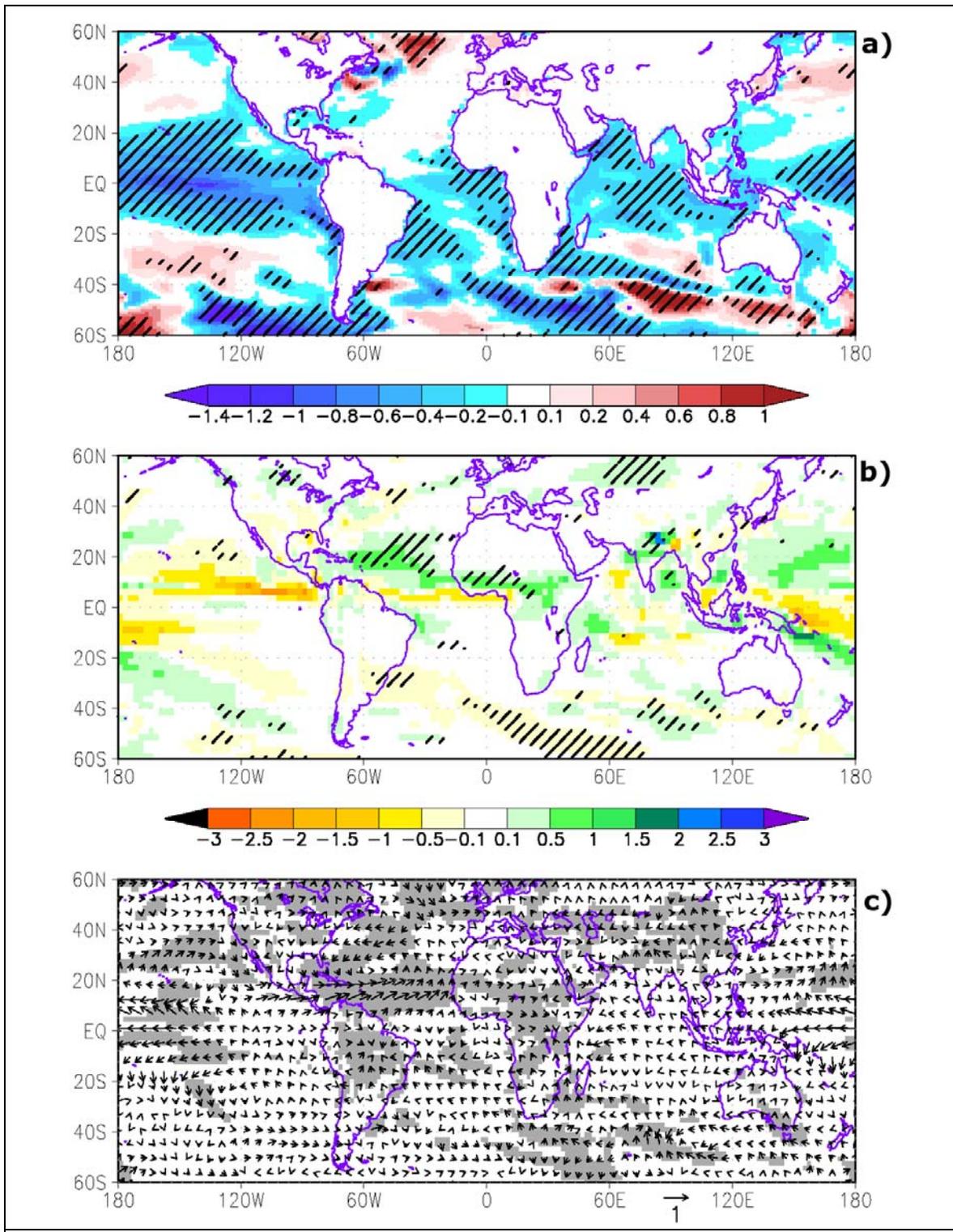
Further insight is gained from the composite difference JAS fields for the low-level wind, precipitation and SST for observations (Fig. V.11), the JFM initialized forecasts (Fig.

V.12) and the AMJ initialized forecasts (Fig. V.13). In the observed SST composite difference (Fig. V.11a), the first impression is of a general warming signal, as would be expected given the known global warming through the period (Solomon et al., 2007). However, the warming is far from spatially uniform, and in general, the warming in the North Atlantic is substantially greater than that in the South Atlantic. The role of such an arrangement of relative temperature gradient (including the Indian Ocean with the South Atlantic) has been previously proposed in explaining low-frequency Sahel rainfall fluctuations through the 20<sup>th</sup> Century (Folland et al., 1991; Rowell et al., 1995; Hoerling et al., 2006), and an index is used in many empirical Sahel seasonal prediction systems to represent this effect (WMO/ACMAD, 1998). The North Atlantic part of the interhemispheric SST contrast has been recognized to project strongly on the Atlantic Multidecadal Oscillation (AMO, Enfield et al., 2001) which itself has been directly associated with low frequency Sahel rainfall variations (e.g. Zhang and Delworth, 2006; Knight et al., 2006). The implication from Fig. V.11a is that a distinct North Atlantic versus South Atlantic temperature difference has been evolving through the 1981-2008 period, with a strong contribution from the AMO (e.g. Ting et al., 2009), and with a sign that would support the increase in Sahel rainfall precipitation. Consistent changes are found in the tropical Atlantic wind field (Fig. V.11c), changes that resemble the leading wind EOF discussed earlier in models, with large-scale fluctuation in the North and South Atlantic trade-wind systems.





**Fig. V.12 :** Same as Fig. V.11 but for CFS predictions of JAS fields. Fields shown are the average of predictions initialized in Jan-Feb-Mar.

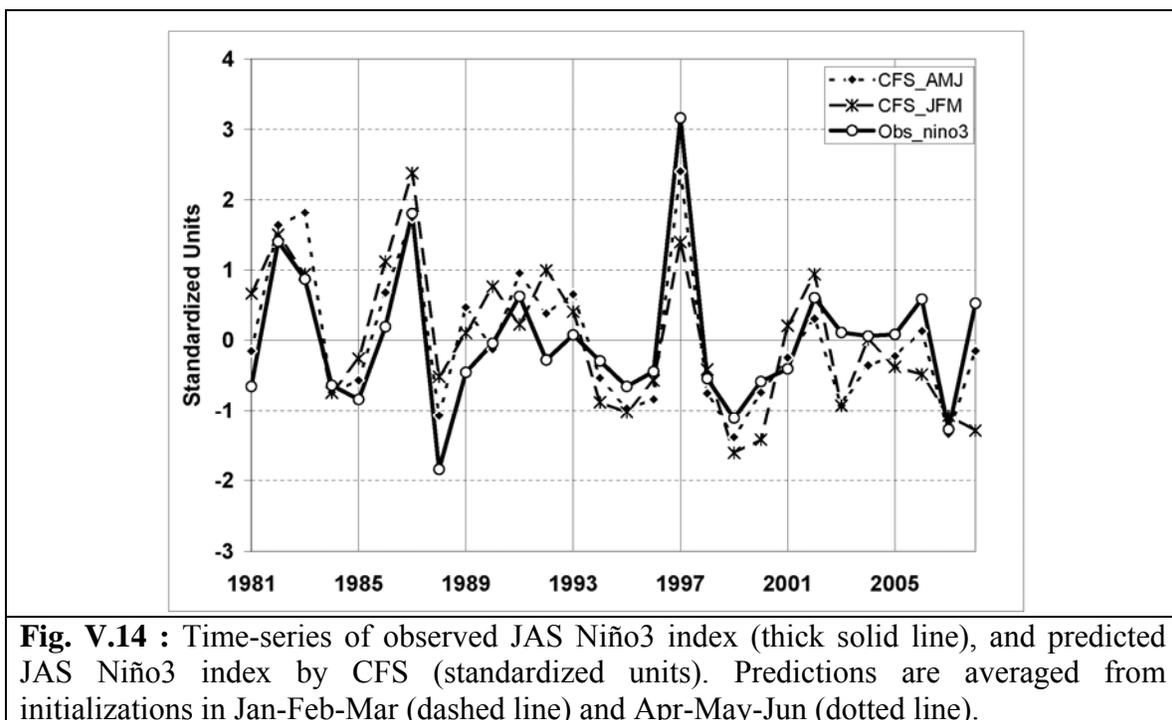


**Fig. V.13** : Same as Fig. V.11 but for CFS predictions of JAS fields. Fields shown are the average of predictions initialized in Apr-May-Jun.

Figures V.12 and V.13 reveal the extent to which the CFS forecasts contain the above low frequency climate fluctuations over 1981 to 2008. Firstly, for the SST composite, the CFS (Fig. V.12a, Fig. V.13a) is generally cooler than observed (Fig. V.11a) in the recent period. However, since the North-South gradient of SST is hypothesized to be a key factor for Sahel rainfall, it is important to assess the extent to which the CFS forecasts are capturing this aspect. The composite maps show that the North Atlantic forecasts for JAS clearly tend to be warmer than the South Atlantic for both JFM (Fig. V.12a) and AMJ (Fig. V.13a) initializations. The precipitation (Fig. V.12b, Fig V.13b) over West Africa and across the tropical Atlantic is also modified in a manner that is broadly consistent with observations (Fig. V.11b), with enhanced precipitation across the Sahel and northern side of the ITCZ across the tropical Atlantic, with a partial compensation in areas to the south (the partial compensation on the south side is more pronounced in the model than in observations). Of significance for interpreting the MOS rainfall results, the tropical Atlantic low-level wind field (Fig. V.12c, Fig. V.13c) is found to have a pattern that has resemblance to the EOF1 modes used in the MOS system, with anomalous easterly South of the Equator and, more noticeably, anomalous westerly North of the Equator, although the wind expression is somewhat weak, especially South of the Equator. Thus, these results are consistent with the EOF1 of the CFS wind containing a positive trend over the period, but more weakly than observed, leading to the weaker trends in the MOS rainfall prediction than observed.

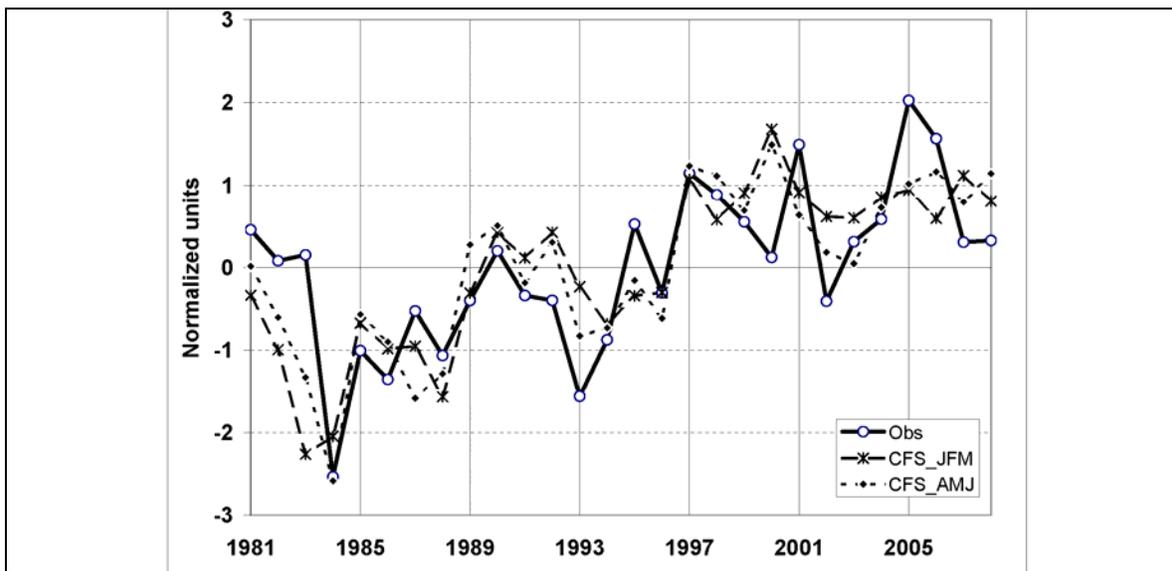
A further aspect of the CFS composites (Figs. V.12a, V.13a) is that they reveal an apparent tendency to predict cool conditions in the central and eastern Equatorial Pacific during the more recent period. Though it appears the cooling is more pronounced in the CFS forecasts than it is in the observations, a tendency to forecast La Niña is not the

dominant feature of the composite trend in the forecasts. Inspection of the predicted and observed Niño 3 time-series (Fig. V.14) shows that there is not a dramatic bias toward forecasting La Niña development and associated cold conditions in the latter period. Furthermore, if a dominant aspect of the composite difference for the CFS were a tendency for La Niña in the recent period, then from the teleconnection structures of the CFS ENSO in Figs. V.6 and V.7, warmer conditions in the South Atlantic and cooler conditions in the North Atlantic would be expected in Figs V.12 and V.13. Therefore, in the Atlantic sector, the trends appear to represent features that are not a direct response to a trend in the CFS toward cooler conditions in the eastern tropical Pacific.



To further investigate the observed and model Atlantic SST variations over 1981-2008, an index of North minus South SST over the Atlantic basin (AtlN-S) has been calculated (Fig. V.15). The index is calculated for the Atlantic domain 30°S to 60°N, and is calculated as (60°N-10°N) MINUS (10°N-30°S). The dividing line for the difference is

taken at 10°N given the general results in the literature that show this to be the approximate latitude at which the zero line of the North-South contrast usually emerges in July-September teleconnection analyses with Sahel rainfall (Folland et al., 1986; Rowell et al., 1995), and represents a measure that contrasts SST approximately to the North and South of the latitude of the ITCZ at this time of year. As expected from the composite maps (Figs. V.11-V.13), the observed and model-predicted time-series of AtlN-S show significant upward trends (Table V.2). The observed JAS index has a positive correlation with observed Sahel rainfall over this period ( $r=0.54$ ), consistent with the relationship between such indices and Sahel rainfall throughout the 20<sup>th</sup> Century (Folland et al., 1986; 1991). The CFS JAS Sahel rainfall is strongly related to the model's AtlN-S ( $r=0.55$  for AMJ initializations,  $r=0.65$  for JFM initializations), supporting the relationship to be an integral feature of the model's regional climate system, as also suggested in the composite maps (Figs. V.12,V.13).



**Fig. V.15** : Time-series of JAS North Atlantic MINUS South Atlantic (AtlN-S) SST. Observed (thick line), predicted by the CFS from initializations in Jan-Feb-Mar (dashed line,  $r=0.64$  with observed) and initializations in Apr-May-Jun (dotted line,  $r=0.78$  with observed).

Furthermore, it is clear from Fig. V.15 that the CFS is quite effective at containing the AtlN-S information in its JAS SST predictions, from both JFM initialization times ( $r=0.64$ ) and AMJ initialization times ( $r=0.78$ ). Persistence in the observed AtlN-S index is quite high ( $r=0.53$  for AMJ to JAS,  $r=0.72$  for JFM to JAS) so the above skill for the CFS represents a modest improvement on persistence. However, it is noteworthy to find that the CFS not only maintains the initialized AtlN-S, but in fact enhances information to slightly beat persistence in its predictions of the JAS values of AtlN-S at several month lead-times.

The interpretation is that CFS seasonal forecasts successfully track many of the observed trends in the climate of the study region and Atlantic Ocean. This is achieved because the CFS contains and skillfully projects forward in its seasonal forecasts, variations of AtlN-S, low-level tropical Atlantic winds, and Sahel rainfall. The low-level tropical Atlantic winds yield a modest underestimation of the MOS-predicted Sahel rainfall. The reason for this underestimation is beyond the scope of the current paper and requires further investigation. While adding AtlN-S as a predictor in the MOS system might be suggested from the above analysis to yield a better trend in the MOS system, results to this point do not indicate that such an approach leads to any major immediate improvement.

#### **V.4 Summary**

In summary, the MOS system using EOF1 of tropical Atlantic winds has been able to display skill in predicting Sahel rainfall using the CFS model in the period 1981-2008 at lead-times up to six months. The skill derives from the interannual timescale (primary relationship with ENSO), and also from the low frequency timescale, with predictions tracking the upward trend in Sahel rainfall over this period, though with a trend magnitude somewhat less than observed. This low frequency component has a close relationship with AtIN-S over this period of analysis. The direct model-output Sahel rainfall from the CFS seems to contain a response to the model's skilfull representation of AtIN-S, giving the model's Sahel rainfall good skill on the low frequency timescale. However, the skill on the interannual timescale related to ENSO, at lead-times of up to six months, is for the most part only achieved through applying the MOS on the tropical Atlantic wind field.

**Reference chapter V :**

- Enfield, D. B. and D. A. Mayer, 1997 : Tropical Atlantic sea surface temperature variability and its relation to El Nino-Southern Oscillation. *J. Geophys. Res.*, **102**, 929–945.
- Enfield D. B., A. M. Mestas-Nunez, and P. J. Trimble, 2001: The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S. *Geophys. Res. Lett.*, **28**, 2077-2080
- Folland C. K., T. N. Palmer, and D. E. Parker, 1986: Sahel rainfall and world wide sea temperatures, 1901-85. *Nature*, **320**, 602-607.
- Folland C. K., J. A. Owen, M. N. Ward, and A. W. Colman, 1991: Prediction of seasonal rainfall in the Sahel region of Africa using empirical and dynamical methods. *J. Fore.*, **10**, 21-56.
- Herrmann S. M. et al., 2005: Recent trends in vegetation dynamics in the African Sahel and their relationship to climate. *Global Environmental Change*, **15**, 394-404.
- Hoerling M., J. Hurrell, J. Eischeid, and A. Phillips, 2006: Detection and Attribution of Twentieth- Northern and Southern African Rainfall Change. *J. Climate*, **19**, 3989-4008.
- Janicot S., V. Moron, and B. Fontaine, 1996: Sahel droughts and ENSO dynamics. *Geophys. Res. Lett.*, **23**, 515–518.
- Knight J., C. K. Folland, and A. A. Scaife, 2006: Climate impacts of the Atlantic Multidecadal Oscillation. *Geophys. Res. Lett.*, **33**, L17706, doi:10.1029/2006GL026242.
- Lau N. C., and M. J. Nath, 2001: Impact of ENSO on SST variability in the North Pacific and North Atlantic: Seasonal dependence and role of extratropical air–sea interaction. *J. Climate*, **14**, 2846–2866.
- Nicholson S. E., 1980: The nature of rainfall fluctuations in sub-tropical West Africa. *Mon. Wea. Rev.*, **108**, 473-487.
- Nicholson S. E., 2005 : On the question of the “recovery” of the rains in the West African Sahel. *Journal of Arid Environments*, **63**, 615-641.
- Rowell D. P., C. K. Folland, K. Maskell, and M. N. Ward, 1995: Variability of summer rainfall over tropical North Africa (1906-92): Observations and modeling. *Quart. J. Roy. Meteor. Soc.*, **121**, 669-704.
- Solomon S., D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. L. Miller Jr., and Z. Chen, 2007: Climate Change 2007: The Physical Science Basis. Cambridge University Press, 1009 pp.
- Ting M., Y. Kushnir, R. Seager, and L. Cuihua, 2009: Forced and Internal Twentieth-Century SST Trends in the North Atlantic. *J. Climate*, **22**, 1469-1481.
- Ward M. N., 1998: Diagnosis and short lead time prediction of summer rainfall in tropical North Africa at interannual and multidecadal timescale. *J. Climate*, **11**, 3167-3191.
- WMO/ACMAD, 1998: Prevision climatique en Afrique. African Centre of Meteorological Applications for Development Rapport Tech. 927, 210 pp. [Available from ACMAD, B.P. 13184, Niamey, Niger.].
- Zhang R., and T. L. Delworth, 2006: Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophys. Res. Lett.*, **33**, L17712, doi:10.1029/2006GL026267.

## **Chapter VI : Onset of the Sahelian Rains: A study over Senegal**

In this chapter we try to understand and predict the onset of the rainy season in the western most part of the Sahel, Senegal. The study draws on daily rain gauge data. The aim is to contribute to knowledge that can underpin alternatives that could be offered to farmers, based on information about the onset of the rains. For example by tailoring management tactics to weather conditions in years with an early onset of rains, it is possible to establish a second crop along with the traditional one. The main objectives are: (i) understand the mechanism and dynamics around the time of an onset, (ii) establish the relationship between onset and atmospheric circulation prior to the onset with separation between true and false onset and (iii) propose a model to forecast the regional onset date at a synoptic time scale, and at a longer time scale spanning a month or two, using large scale information, drawing on any SST influence.

We will first define the onset of the rainy season incorporating a practical user perspective. From onset dates we will calculate large scale onset indices over two well defined homogeneous regions of Senegal. The next sections contain the results. First, we will look at the predictability of monthly rainfall around the mean onset date using large scale information, to form a bridge from the results of previous chapters. The possibility of large-scale forcing (SST) influencing the onset date itself is then explored by calculating teleconnection maps with the onset indices. Then, a more synoptic perspective is taken, documenting atmospheric circulation around the onset including propagating features which are the most dominant features during this time of the season.

The next section develops the synoptic perspective further, and identifies atmospheric sequences that are distinctive of false starts, as compared to true starts. The last results section considers prediction methods.

## **VI.1 Method**

We use data sets with daily resolution to calculate the onset of the rainfall. The period of analysis is from April to October from 1950 to 2008. A set of 20 stations with a complete record during that period was selected. We will especially focus on the latter part of the record when atmospheric datasets (reanalysis, satellite) are more reliable and available for diagnostic analyses, and when the coupled CFS-GCM runs are available.

### **VI.1.a Definition of onset date**

Many definitions of onset are found in the literature for West Africa, reflecting the considerable interest in the topic. In meteorology, the onset is the period when deep convection is enabled and the broad scale atmospheric system is in place. The hydrologist is more concerned by the period when the first rainy events are generating a runoff even if it is very confined in space. The agronomist needs an event that (i) sufficiently moistens the soil, and (ii) is not followed by a dry spell. These considerations show the difficulty of defining a consensual onset date. In this study we consider onset definitions which try to capture different aspects of an onset: local effect (first rainfall event at any station), large scale signal (simultaneous rain event reported in at least 3 stations over a large domain), and an actual onset guaranteeing continuity in the rainfall sequences (onset not followed by a long dry spell). Each has its own relevance for different users especially for farmers, hydrologists and pastoralists. We consider a rainfall event to be significant when it amounts to 20 mm over a three-day period. 20mm seems to be a

number which allows early-season runoff and moistening of the soil, taking account of the daily evapo-transpiration in the shiny and windy climate of the Sahel (Taylor, 2000; Gash et al., 1991). For example Davy et al. (1976) have documented that the planting date for millet was observed to coincide with the first occurrence of 20mm of rain over a two-day period. Such a threshold is similar to that proposed by Stern et al. (1981) for an onset definition over the Sahel, and to that proposed by Sivakumar (1998) based on observations of the establishment of the millet crop at the ICRISAT Sahelian Center.

Some previous studies, especially focused on meteorology and dynamics of onset, used onset indices calculated directly from regional scale indicators e.g. OLR. In our study, we seek to combine the user-oriented indices above with regional indices that lead to insight into the meteorology behind the aspects of most interest to users. We look more closely into daily station data, resolving distinctive features for rainfall onset in the northern and southern sub-regions, as well using indices that are based on rain events and dry spells, that are important for users.

**First rainfall event in the sub-region** : is defined as the first time after May the 1<sup>st</sup> when, at any station, there is a 3-day rainfall accumulation of at least 20mm. With such a definition, we emphasize more on the end of the dry season and the first outbreak of rain, with no constraint applied to the scale of the rain event, nor to the character of the days following the event. This definition is very important for a local population because it sets the seasonal clock, their mind set and their psychology, which switch from the long dry season up to now. They start thinking and reacting in terms of wet season. Usually farmers start making practical decisions after the first significant rainy event: moving

from cities (where they may have been doing a little business to survive during the long dry season) back into the country side to prepare their fields, to start borrowing funds or seeds from the government or private operators, to start working the fields, or even some planting to profit from a probable long season according to their own judgment. Studying such an event is relevant for these reasons.

**First large scale onset event in the sub-region** : primarily to enable a better study of the large-scale atmospheric circulation and related conditions associated with onset, we now require the rainfall event to be observed at a larger spatial scale. From the definition of “First rainfall event” we add a further condition which requires the onset rainfall event to be observed simultaneously across at least three stations in the given homogeneous zone. This second condition guarantees that the event was not isolated and must be associated with large-scale atmospheric conditions. This large scale rainfall event onset can be partitioned further into two categories: a true onset and a false onset.

**True onset for Type 1 and Type 2:** for true onset type 1, an additional restriction was placed on the large scale rainfall event. It requires that no dry period of 7 or more consecutive days will occur in the following 10 days, a dry day being when less than 0.1 mm is recorded. In other words, for a rainfall event to be qualified as a true onset, the following needs to be satisfied simultaneously across at least 3 stations : i) 3-day period of rain greater than 20 mm and ii) no dry spell of 7 consecutive days or more at any time in the next 10 days. The true onset definition is intended to guarantee optimal condition for crop germination, establishment and growth during its early stage, without encountering any water stress. It’s difficult during such conditions to imagine farmers

having difficulty in planting and establishing their crop, so it's a good proxy for a preferred planting date. A similar definition of a true onset has been suggested by many authors Benoit et al. (1977), Stern et al. (1981,) Jolliffe and Sarria-Dodd (1994), Hulme (1987), Nicholls (1984) and Sivakumar (1988). However, in all cases, the focus was on a single station. In this study we extend the definition from a single location to a larger homogeneous domain. For the true onset type 1, as discussed above, we have made the extension by focusing on at least a three-station rain event, and checking for dry spells at each of the three stations. A more flexible variation on this is also introduced (true onset type 2), which considers the extent of dryness across the whole homogeneous region after the rainfall event. The true onset type 2 takes into consideration the spatial mean of the number of dry days in the 10 day-period following the large scale onset event. True onset type 2 requires this mean number of dry days to be less than a specified threshold.

**False onset, and the definitions by Type 1 and Type 2 :** It's very important to be able to clearly distinguish false start from true start. With a false start, it is likely that farmers have already prepared their farms, applied fertilizer and planted their crops, only to then experience a dry spell which their crop cannot survive. With very limited resources (manpower and money to buy seeds) the agricultural season is then jeopardized. Detecting false start is therefore of huge practical significance, and any reliable information or forecast of such events would be of great value for the farmers and for the general economy as a whole.

False onset type 1 is when the three-station rainfall event criterion is satisfied, but the event fails the dry spell criterion (as discussed in defining true onset type 1). False onset type 2 provides a more flexible approach. As with true onset type 2, we consider the 10

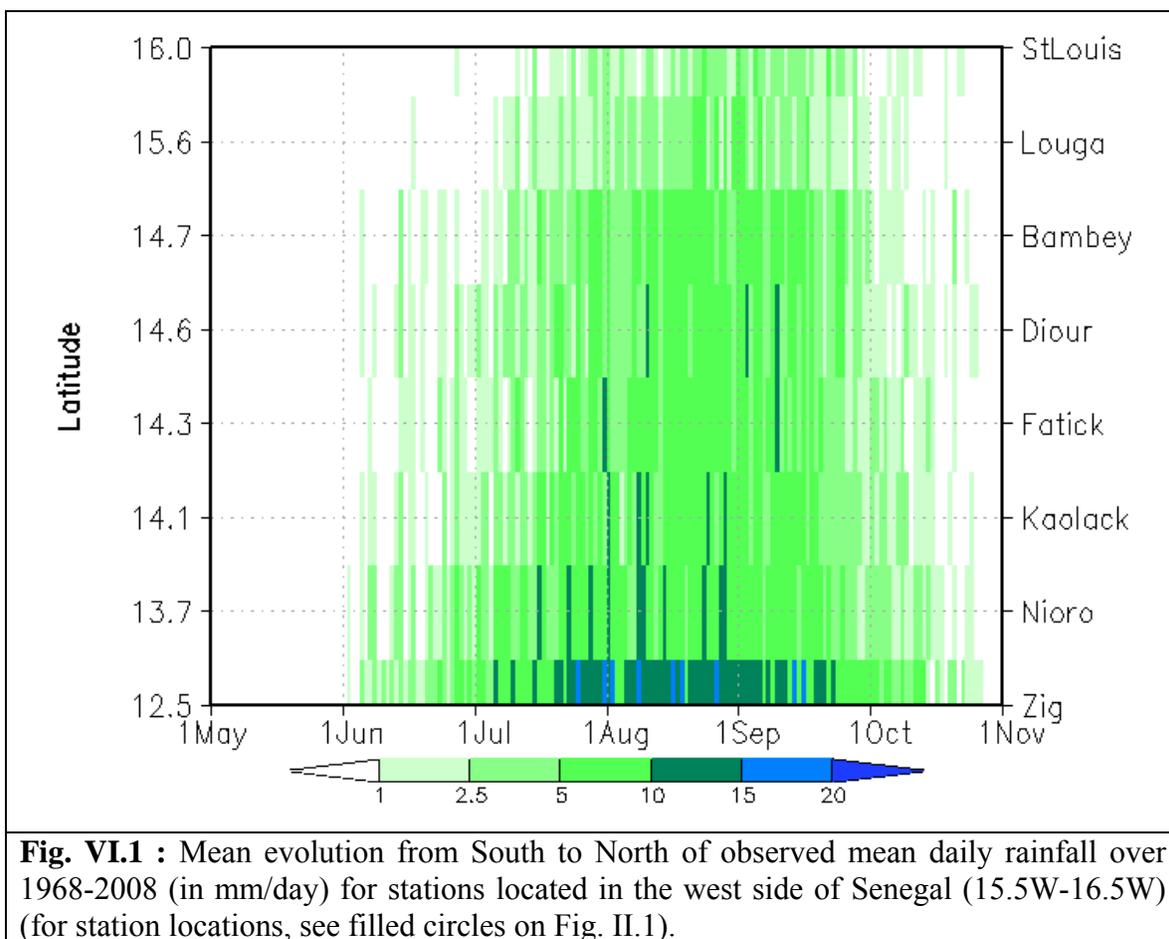
days after the rain event, and take the average number of dry days across the whole homogeneous region. False onset type 2 is when the average number of dry days is greater than a certain threshold. If this threshold is set very high, it will focus on severe false starts. Therefore, this approach can be implemented to guarantee especially clear demarcation between false and true onset to help highlight differences in atmospheric sequences between true and false starts. This type 2 definition will therefore be used in section VI.6 where we will try to separate true versus false onsets and find precursor signals to differentiate between true and false starts. Prior to that section, reference to true and false start will be according to Type 1 above.

In the next section we will apply the definitions (first event, first large scale event, and true onset type I) to station rain gauges and propose to classify them into homogeneous zones with respect to onset date.

### **VI.1.b regional indices of onset date**

It's well known that West Africa rainfall is controlled by the North-South movement of the Inter-Tropical Convergence Zone (ITCZ), which is concomitant with the progression of the West African monsoon. Senegal is located in the western most part of the Sahel, and has its seasonal rainfall march from May to October controlled by these factors. Fig VI.1 shows the mean annual cycle for stations located over a narrow band of longitude, 15.5W and 16.5W, and aligned from North to South 12.5N to 16.0N. There is a decrease of the length of the rainy season from South to North. In the south it starts in June and lasts until October. The climatological rainfall can reach more than 15mm per day and the bulk of the rain falls in July and September. For northern stations, the rainy season

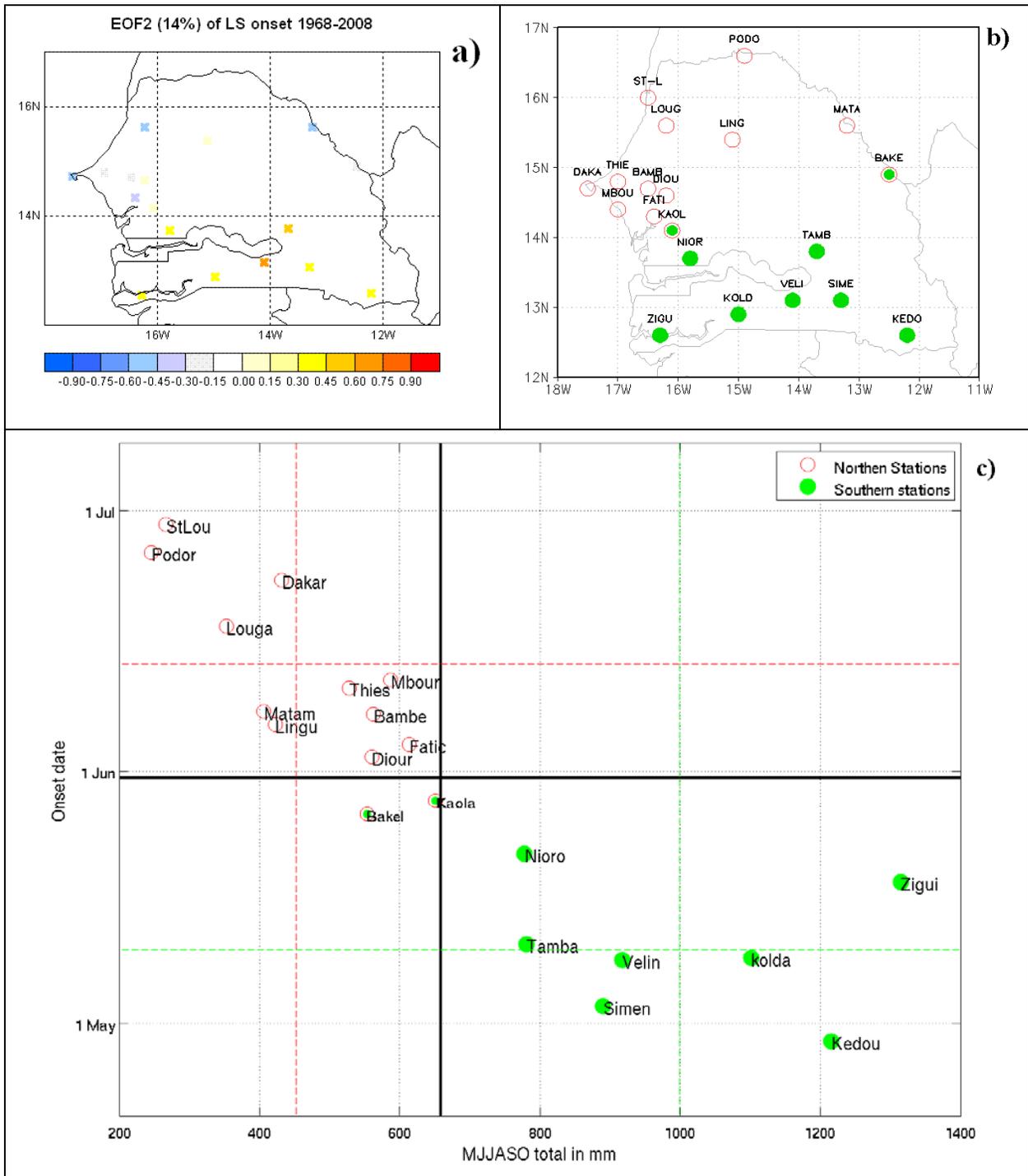
starts in July and lasts until the end of September. Rainfall climatology rarely reaches 15mm per day. A salient feature is the gradual establishment of the rainy season from South to North. The progression is very clear. On the other hand, the end of the season is very abrupt and occurs more simultaneously in all stations from North to South (Fig. VI.1). Lele and Lamb (2010) have documented in detail this asymmetry in the climatological speed of progression of the annual cycle for the central Sahel (Niger).



First, some preparatory analysis is undertaken, to define homogeneous rainfall onset regions across Senegal. We calculate the first rainfall event at each of the 20 selected stations over Senegal. To diagnose the spatial homogeneity of the first rainfall event

amongst stations, we apply an EOF analysis. Fig. VI.2a shows the second EOF pattern of the first rainfall event capturing 14% of the total variance. It shows clear separation between northern and southern stations. The first EOF contained weights with the same sign across the whole domain. This second EOF may partly reflect a tendency to be a bipolar pattern. However, as discussed below, in this case the pattern has a physical interpretation and provides guidance on the separation of the domain into two parts based on the correlation structure in the data.

Fig. VI.2c shows a scatter plot of mean rainfall total during MJJASO versus mean first rainfall event date by station. We see a near-perfect line up of the stations from North (St-Louis, Podor) to South (Kedougou). This is consistent with the location of the stations in a trans-section South-North (Fig. VI.2b) and reflects how the ITCZ and monsoon progress from South to North. Southern stations receive a higher seasonal rainfall total (around 1000 mm on average), the first rainfall event occurs earlier in the season (around mid May) and the rainy season lasts longer. Northern stations receive less rainfall total during the season, on average around 450mm, and the first rainfall event occurs later in the season (around the end of June).



**Fig. VI.2 :** Defining two homogeneous sub-regions for rainfall onset in Senegal : a) the second EOF of the first onset date time-series for each station, b) the northern and southern regions that are used for calculating regional onset indices and c) mean first event onset date plotted against seasonal rainfall total.

Using both the climatology, and the year to year variability of the first rainfall event (EOF2, Fig. VI.2a), Senegal can be divided into two main regions which mirror the migration of the ITCZ. A northern region having 11 stations (open circle in fig. VII.2b) and the southern region having 7 stations (close circle in fig. VII.2b). The dashed lines in Fig. VII.2c indicate the center of each region in terms of mean rainfall total during MJJASO and mean first rainfall onset date.

We will calculate regional onset-date indices for each of these two regions. A time-series will be produced applying each of the definitions introduced above: first event, first large-scale event and true onset type 1. In addition, we will also calculate for each region, a spatial mean standardized index for the date of the first event:

$$I_i = \frac{1}{(n-4)} \sum_{j=1}^{n-2} (X_i(j) - \overline{X_j}) / \hat{\sigma}_{X_j} \quad (6.1)$$

$I_i$  = standardized onset value during year  $i$ .

$n$  = number of stations in the region

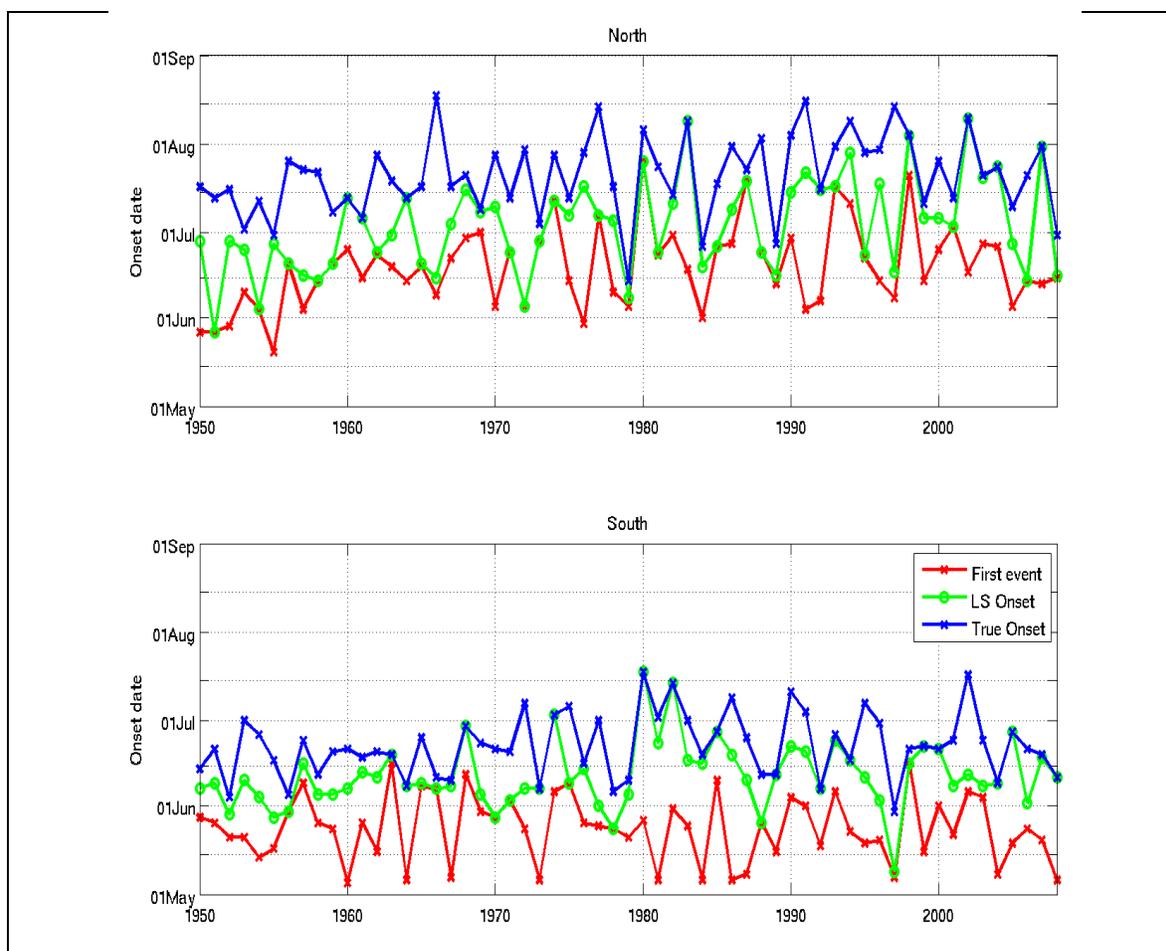
$X_i(j)$  = onset date at station  $j$  during year  $i$ .

$\overline{X_j}$  = mean onset for station  $j$

As the range of dates of the first rainfall event tends to contain outliers within the same region during each year, before calculating the spatial average, we remove during each year the two extreme dates, earliest and latest date, ( hence  $n-4$  in the above equation counting for estimation of the mean and the standard deviation). This standardized index is considered to be well-suited to identifying large-scale signals in the

climate system associated with early/late onset. The standardized index will be used in section VI.3 to calculate the large scale teleconnection with SST and circulation.

Figure VI.3 shows time series of the onset date using the three definitions for the northern and southern regions from 1950 to 2008. For both regions, there is not a single year when all the three onset definitions coincide. In other words, usually the season starts with first a local significant rainfall event, followed later by a large scale event and then afterwards full establishment of the actual rainy season. The first rainfall event may correspond with the large scale onset or the large scale onset may be the actual start of the rainy season but there is never coincidence of the three. This increases further the interest of studying, understanding and predicting the rainfall onset. It's also interesting to note the very weak parabola trend in the onset dates. This contrasts with the known rainfall total trends: the strong downward trend from the 50s toward the 70s, and the moderate recent (about 20 years) upward trend in Senegal and the Sahel region as a whole. Kniveton et al. (2009) , using a global data set, found similarly weak trends in the onset date over Senegal during 1978-2002 using different rainfall thresholds in defining the onset.



**Fig. VI.3** : Onset time series from 1950 to 2008 for the northern Senegal region (top panel) and southern Senegal region (bottom panel). Indices are as defined in the text, for first rainfall event in the region (red), first large scale event in the region occurring simultaneously over at least 3 stations (green), and the true start on the rainy season, not followed by dry spells (blue).

Table VI.1 gives some descriptive statistics of the onset dates. There is no systematic shift or trend in the onset date regardless of the definition. In some senses, the standard deviation can be considered not so large (being in the range 11-17 days for these indices), in a region with erratic rainy events and strong year to year variation. However, the variation is very significant for agricultural practice. The large scale onset over the North has the highest standard deviation of about 17 days.

**Table VI.1** : Mean onset date  $\pm$  1 standard deviation in days for, different regions, periods and onset definitions.

		<b>1950-2008</b>	<b>1968-2008</b>	<b>1981-2008</b>
<b>NORTH</b>	<b>First event</b>	20 June $\pm$ 14	23 Jun $\pm$ 14	23 Jun $\pm$ 13
	<b>Large scale onset</b>	4 Jul $\pm$ 17	9 Jul $\pm$ 16	11 Jul $\pm$ 17
	<b>True onset</b>	22 Jul $\pm$ 14	24 Jul $\pm$ 14	25 Jul $\pm$ 14
<b>SOUTH</b>	<b>First event</b>	25 May $\pm$ 12	25 May $\pm$ 11	23 may $\pm$ 12
	<b>Large scale onset</b>	13 Jun $\pm$ 12	14 Jun $\pm$ 13	15 Jun $\pm$ 12
	<b>True onset</b>	23 Jun $\pm$ 11	25 Jun $\pm$ 12	25 Jun $\pm$ 12

False onsets are very frequent in both regions and it's quite seldom to find the first large scale event coinciding with true onset (Fig. VI.3). In the northern region only 10 out of 59 years experience such a coincidence and in the south, the ratio is 15 out 59. This weak probability of having a true onset with the first large scale rainy event could lead to low crop production. Crop failure is not necessarily related to lack of enough total rainfall through the season, but instead the occurrence of short dry periods causing water stress during early stage of plant growth.

There is, over all the indices, a generally positive correlation between onset in the south and onset in the north. The magnitude of the correlation varies across definition of the onset, but for 1981-2008, is in the range 0.36 to 0.22 (Table VI.2a). This small, but positive, association could be used to contribute to prediction for the north, given the event in the south is about 1 month before the event in the north.

Within the regions, there is some common variation among the regional indices, but there is also some substantial independent variation. For example, first event and first large-scale event correlate at 0.31 for the south and 0.36 for the north. There is a closer association between first large-scale event and true onset ( $r=0.60$  for the south and  $r=0.56$  for the north).

The standardized indices based on the first event at each station (Eq. 6.1) have the most common variance across the regional indices (Table VI.2b). Only the first event regional index for the northern region has a relatively weak relation with the northern standardized index ( $r=0.37$ ). All other correlations are in the range 0.54 to 0.78. The good correlation of the standardized indices with the other indices is supportive of the idea that the standardized indices best capture large-scale signals associated with onset. However, the other indices have the advantage that they can be related to specific onset dates, allowing analysis of the evolution of synoptic circulation features around the time of onset. The respective advantages of the indices are used in the following sections.

**Table VI.2a** Cross-Correlation between regional onset indices 1981-2008.

		North			South		
		First event	Large scale	True onset	First event	Large scale	True onset
North	First event	1	<b>0.36</b>	<b>0.24</b>	<b>0.36</b>	<b>0.20</b>	<b>0.17</b>
	Large scale		1	<b>0.56</b>	<b>0.43</b>	<b>0.22</b>	<b>0.31</b>
	True onset			1	<b>0.35</b>	<b>-0.22</b>	<b>0.23</b>
South	First event				1	<b>0.31</b>	<b>0.37</b>
	Large scale					1	<b>0.60</b>
	True onset						1

**Table VI.2b** Correlation between the area-averaged standardized onset indices (Equation 6.1) and the regional index defined in the three different ways (first event, large scale event, true onset). Correlations are for 1981-2008.

		North	South
Regional indices	First event	0.37	0.61
	Large scale	0.70	0.54
	True onset	0.78	0.56

In this section we have divided Senegal into two homogeneous regions with respect to the first rainfall event as well as using the climatological onset and the seasonal rainfall total. Before using onset indices in diagnostic and prediction analyses, we will first check the predictability of monthly rainfall around the climatological time of the onset using large scale signals from SST, applying the same approaches as used in the previous chapters for the bulk of the rainy season.

### **VI.2 Predictability of monthly rainfall around the mean onset date**

In this section we propose to look for any relationship between monthly rainfall around the time of the onset (May, June and July) and large scale signals in the SST. This was done directly through correlations with SST, or indirectly, by analyzing a GCM forced with observed SST (ECHAM4.5) or a fully coupled GCM (CFS). The objective is to check for any predictability of the monthly rainfall total during the time of onset. This is done having in mind that monthly rainfall might be related to the onset, or both onset and monthly rainfall may be related to the same ocean-atmosphere dynamics. Table VI.3 shows the skill of predicting monthly rainfall from May to July for the Sahel and each

region of Senegal. We choose SST because predictable large scale information at this time scale is primarily driven by the SST. We also apply the MOS approach on the GCM wind field, in exactly the same way as in previous chapters. The MOS approach could subsequently be investigated further, since it is possible that the domain window might be improved upon for the pre-rainfall season months.

**Table VI.3 :** Predictability of monthly rainfall total in May, June and July. Predictands are rainfall indices for the Sahel, Southern Senegal and Northern Senegal. SST indicates an EOF regression with tropical SST (using SST observed at the same time as the rainfall). ECHAM4.5 indicates GCM simulation with observed SST, either direct model output (Prcp) or applying the regional wind MOS approach. CFS indicates 0-month lead forecast using the fully coupled CFS. Results shown are the correlation skill. When more than 1 EOF is used, the value in brackets indicates the number of EOFs to achieve the maximum skill (with a maximum of 5 EOFs allowed).

		MAY						
		SST			ECHAM4.5 sim			CFS
		50-08	68-08	81-08	50-08	68-08	81-08	81-08
Sahel	Prcp				0.11	0.18	0.14	-0.07
	MOS	0.18(5)	0.00	0.00	0.41(3)	0.37(4)	0.19(3)	0.00
Northern Senegal	Prcp							
	MOS	No rain during this period						
Southern Senegal	Prcp				0.35	0.45	0.44	-0.19
	MOS	0.22(4)	0.15(5)	0.46(5)	0.31(3)	0.06(5)	0.2(5)	0.00

		JUNE						
		SST			ECHAM4.5 sim			CFS
		50-08	68-08	81-08	50-08	68-08	81-08	81-08
Sahel	Prcp				0.2	0	0	0.05
	MOS	0.18(4)	0.06	0.09(5)	0.26(2)	0.17(2)	0.25(5)	0.00
Northern Senegal	Prcp				0.2	0.0	0.16	-0.12
	MOS	0.08	0.00	0.00	0.00	0.00	0.00	0.00
Southern Senegal	Prcp				0.37	0.22	0.28	-0.01
	MOS	0.17	0.00	0.09(5)	0.03(5)	0.00	0.09(5)	0.16(4)

		JULY						
		SST			ECHAM4.5 sim			CFS
		50-08	68-08	81-08	50-08	68-08	81-08	81-08
Sahel	Prcp				0.3	0.28	0.14	0.69
	MOS	0.25 0.48(4)	0.35(2)	0.35	0.51 0.64(2)	0.39 0.43(2)	0.08 0.21(2)	0.65
Northern Senegal	Prcp				0.49	0.46	0.28	0.24
	MOS	0.31 0.44(4)	0.06 0.30(2)	0.14 0.2(3)	0.40 0.46(4)	0.20 0.44(5)	0.15(3)	0.23
Southern Senegal	Prcp				0.34	0.26	0.12	0.37
	MOS	0.16 0.23(4)	0.12(2)	0.00	0.28	0.14	0.10(3)	0.18 0.27(4)

The skills are quite modest or nil for May and June rainfall indices. There is possibly some suggestion of modest skill in May rainfall for the Sahel (with ECHAM4.5) and Southern Senegal (with SST and direct ECHAM4.5 rainfall). This skill, although modest, shows origins in the tropical Atlantic SST (not shown), and is consistent with prediction signals for rainfall onset in Southern Senegal that emerge in the next section. In contrast to the modest skills found in May, the skills for June are all very weak.

In July we start seeing more general positive skills, with some high correlation skills especially over 1950-2008 (when skill is amplified by the trend). July is the period when the known teleconnections with Sahel rainfall begin to deliver high skill. Results are noisier for the Senegal indices than for the Sahel, but there is some general skill, especially for the northern Senegal index. We can notice the high correlation with CFS of  $r=0.65$  which we documented already in chapter V.

This picture provides a backdrop against which to interpret large-scale signals with the rainfall onset indices in the next section.

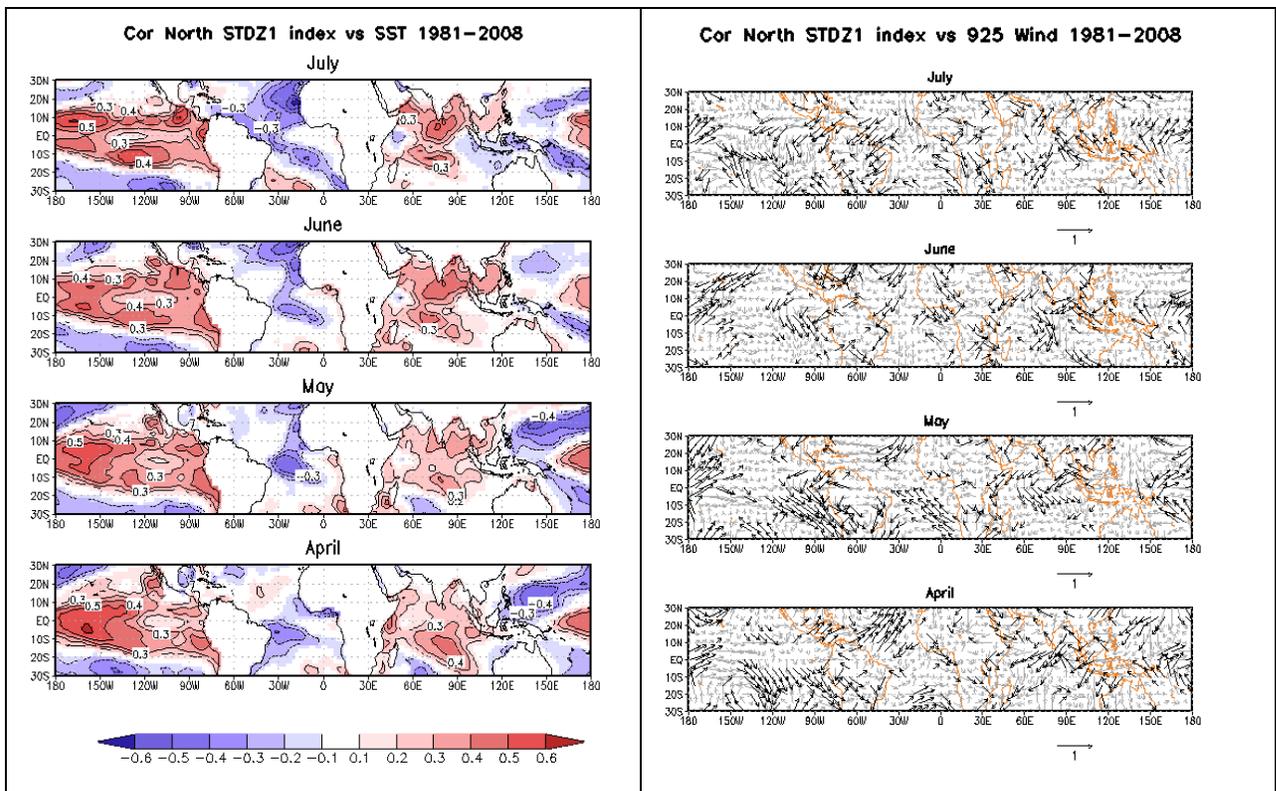
### **VI.3 Large scale teleconnection with regional onset indices : SST and wind**

This section addresses the relationship between rainfall onset and any large scale signal that might be contained in monthly observed SST and low level wind patterns. The area-average standardized onset date indices for each region (Eq. 6.1) are correlated with observed SST and wind from April to July during 1981 to 2008.

Figure VI.4 shows the teleconnection between the onset index over northern Senegal and the tropical SST (Fig. VI.4a) and low level wind (Fig. VI.4b). The teleconnection with the SST depicts positive correlation with the Pacific and Indian oceans and negative correlation with the northeastern Atlantic Ocean near the Senegalese coast. These correlations generally increase as we get closer to the climatological onset month which is July in this region.

Positive teleconnection with the Pacific and Indian oceans means early onset (negative onset anomaly) is associated with low SST (negative anomaly) and late onset with high SST. Late onset thus coincides with El Niño events which is consistent with relatively dry July conditions during El-Niño in the Sahel (Janicot et al., 1996; Ward, 1998), and suggests the known teleconnection relationship is having some impact on the onset date in this northern region. However, the result is neither very strong, nor are the patterns perfectly in line with classical ENSO expectations. While winds do converge from the western equatorial Pacific into the central equatorial Pacific, the wind pattern in the eastern Pacific is somewhat mixed. Also, correlations in the eastern Equatorial Pacific are actually somewhat weaker than those in the central Equatorial Pacific, and off the equator. Whether these departures from classical ENSO signature are simply due to the sampling of a weak relationship, or whether they are real, requires further investigation. However, some influence of the tropical Pacific, with overall warming/cooling consistent with previous teleconnections with July Sahel rainfall on the monthly timescale, seems reasonable to conclude from fig. VI.4.

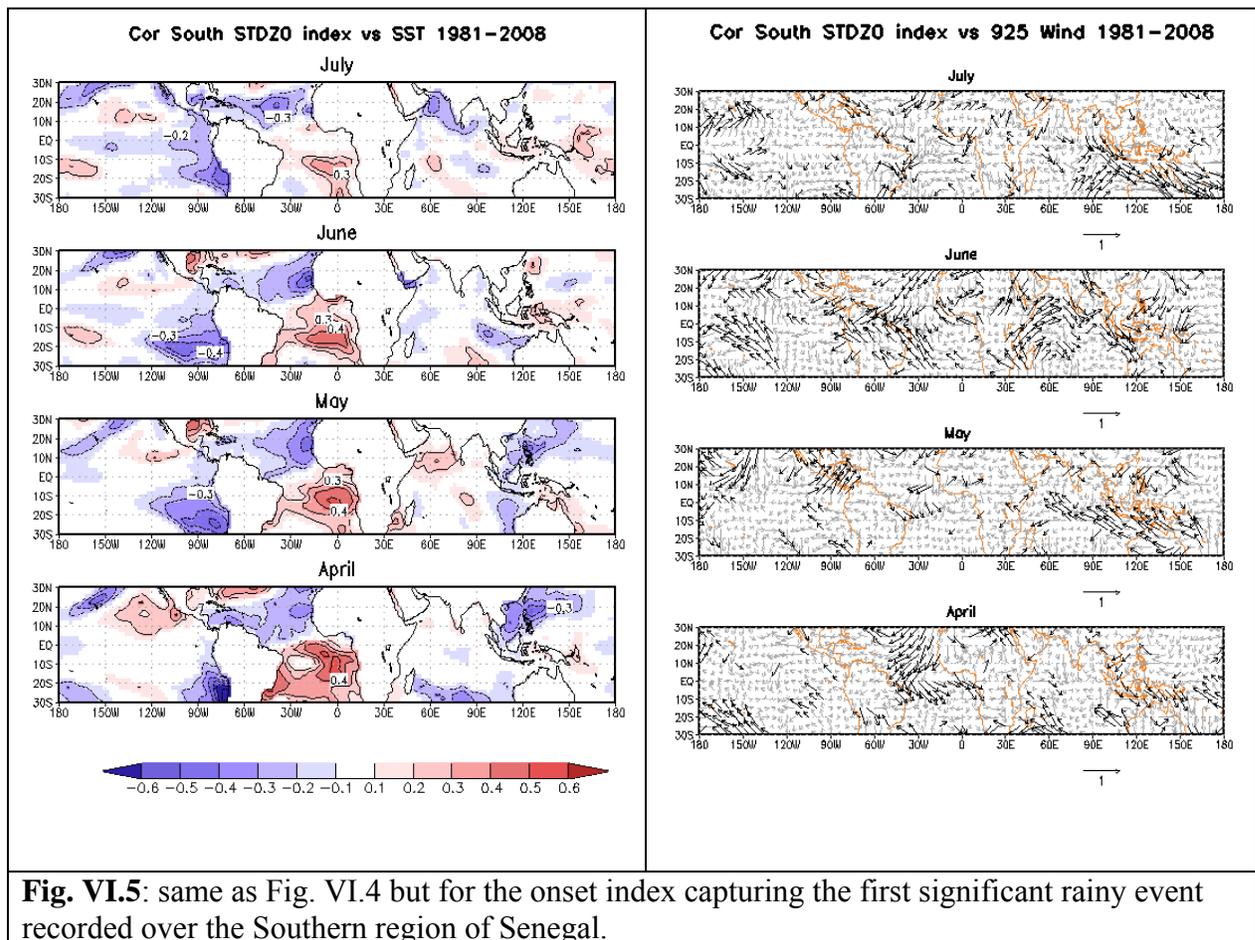
The negative correlation with the northeastern tropical Atlantic is particularly strong along the coast of Senegal. Relatively low SST in the tropical North Atlantic is known to be associated with relatively dry conditions in the western Sahel and Senegal in July (Lamb, 1978; Ndiaye et al., 1999), and so is consistent with a delayed onset as implied by the negative correlation in Fig. VI.4 for July. However, the strength of the relationship is particularly marked, especially close to the coast of Senegal. It could be amplified due to a local effect such as a “sea-breeze” type interaction favorable for convective activities. Sensible heat input during summer is extremely large especially over the continent compared to the neighboring ocean. High SST would be a source of water vapor. During the day the temperature gradient creates a lower level sea breeze flow from the warm ocean (relatively to its climatology) to the even warmer continent due to solar heating. The result is that significantly increased moisture flux is created in the lower troposphere resulting in either local enhancement of convective clouds or strengthening of upcoming eastward propagating meso-scale convective systems. This relationship appears to develop during the time of onset, that is, it is especially strong in the July SST field, when teleconnections also emerge with northerly near-surface wind flow close to the Senegalese coast. Therefore, this teleconnection may be more of a diagnostic tool, than something that can be used for prediction.



**Fig. VI.4** : Standardized onset index over northern region of Senegal correlated with tropical SST (left panels) and 925 hPa wind (right panels). Bold arrows and contoured areas are significant at the 95% level of the t-test. Bottom panels are with April SST and April wind, therefore with the longest lead on the onset index.

Figure VI.5 depicts teleconnection between the standardized onset for the southern region with SST and near-surface wind. There is a striking difference from the northern region teleconnection pattern (Fig. VI.4). Now, the signal is mostly confined in the Atlantic basin, and in the months of April and May. In the SST, there is a clear north-south gradient : positive correlation around the Gulf of Guinea and negative correlation in the northern Atlantic. In other words, warming in the Gulf of Guinea and cooling in the northern Atlantic favors late or delayed onset. High SST in the Gulf of Guinea makes the adjacent surface air lighter creating a low pressure situation while the cooling in the northern Atlantic makes the surface air denser which favors high pressure, and such a pressure gradient will set up a northerly wind flow anomaly (Fig. VI.5) which will work

against the monsoon surge, and which can explain a delay in the onset due to lack of moisture supply (Goëtlogon et al., 2010; Okumura and Xie, 2004). It is promising that the tropical Atlantic teleconnection signal (Fig. VI.5) is present in April, about a month ahead of the typical onset date in the southern region, offering a clear predictive prospect from this teleconnection relationship.



This striking difference between the teleconnections associated with the southern and northern Senegal onsets, can be interpreted according to the physical mechanisms that are operating during the time of southern and northern onset respectively. The onset in the south, which occurs earlier around the end of May, is directly related to the monsoon

development, which can be more controlled by the SST gradient over the Atlantic Ocean. Any blocking or delay in the monsoon system can lead to late onset. Whereas over the northern part of Senegal, the onset occurs around the end of June, one month later, when most of the atmospheric features controlling the dynamics of the rainfall in Sahel are already in place, such as the jets (AEJ and TEJ), creating wind shear which can favor rainfall dynamics and large scale events. The relationship between the Jets, the large scale boundary layer (Fontaine and Janicot, 1992) and the tropical Pacific SST may explain why the onset around July in the northern region has an expression in the tropical Pacific and Indian Ocean SST.

The evolution of the strength of the teleconnection between large scale SST from April to June is opposite in the two regions. Teleconnection between SST and southern large scale onset is stronger during April, May and June, and then weakens in July, after the date of the rainfall onset. In the north the teleconnection is stronger during June and July when the onset usually occurs, and becomes weaker in April. It's worth noting the strong coupling between observed wind and SST in April during the first rainfall event in the southern region. The SST gradient in April is consistent with the wind circulation over the Tropical Atlantic (Fig. VI.5). We will use this relationship to establish a prediction of the onset using tropical Atlantic SST and wind during April.

Another striking feature is the opposite sign of the teleconnection between the Pacific Ocean and the onset of each region. The northern region shows a relatively strong and positive correlation with the Pacific whereas for the southern region, the correlation is weak and negative (Figs VI.4 and VI.5). This is consistent with the change of sign in the teleconnection between ENSO and Sahelian rainfall from July (the well-established negative correlation, that prevails in August and September as well) to June (some

evidence of a reverse in sign to positive correlation, Ward, 1998). Therefore in June, warm condition associated with El Niño (+), favors wet conditions which is consistent with early onset (-) in the Southern Senegal, as described in the slight negative correlation in the eastern tropical Pacific in fig. VI.5. The ENSO-Sahel teleconnection reverses sign during July-August-September which includes the time when the onset over the northern region takes place. This explains the positive correlation : warm conditions of El Niño (+) favors late onset (+).

We have shown some large scale teleconnections between SST/wind and rainfall onset. It is possible to identify consistencies and extensions from previous knowledge, as well as some new features, such as the signal in April/May with southern Senegal. We will return to these findings in the final section in this chapter that explicitly focuses on prediction. In the next section, we will focus on atmospheric circulation at a finer, more synoptic, time scale around the onset. We will look for atmospheric configurations conducive for an onset.

#### **VI.4 Atmospheric conditions around the onset date**

The basis of this section is to composite the time evolution of various atmospheric fields relative to the onset, for each region in Senegal. The objective is to follow atmospheric evolution around the onset date. Some previous studies have begun to document the atmospheric conditions around the onset over the Sahel. Sultan and Janicot (2003) define a 'pre-onset' by the arrival first of the inter-tropical discontinuity (ITD) at 15°N. Then, the onset stage corresponds to an abrupt latitudinal shift of the Inter-Tropical Convergence Zone (ITCZ) from 5°N to 15°N with a climatological date around 24 June for the continental Sahel (between 10W and 10E). Hagos and Cook (2007) using a

regional climate model, document the “monsoon jump” theory, arguing on the important role of the boundary layer circulation and supply of moisture preconditioning the atmosphere. Couvreux et al. (2009) documents the role of the Saharan Heat Low and the simultaneous acceleration of the low level meridional wind during the onset.

All these studies focus on a large time scale and coarse spatial grid and most of the time have filtered the time scale by focusing on the frequency scale of the phenomenon of interest. A filtering approach is difficult to implement in an operational context since the filtering needs past and future data. Here we focus on the variability, at the synoptic time scale, of the lower levels of the atmosphere, without any frequency filtering. Furthermore, we generate results separately for the onset in northern and southern parts.

The 850 hPa is considered as the middle of the monsoon layer and close enough to the surface to depict variability in the boundary layer. We first form daily anomalies by subtracting the multi-year daily mean climatological values. The daily climatology is first smoothed by applying a running average of 7 days to take into consideration the smoothly varying nature of the seasonal cycle. All the composites of wind, moist static energy (MSE) and precipitable water (PWAT) are formed using anomalies calculated in this way.

The moist static energy, in  $\text{KJ Kg}^{-1}$ , is calculated as follows :

$$MSE = C_p T + gZ + L_c q \quad (6.2)$$

where  $C_p T$  is the entropy ( $C_p = 1000 \text{ m}^2 \text{ s}^{-2} \text{ deg}^{-1}$ , is the specific heat capacity of dry air at constant temperature;  $T$  the temperature in Kelvin at 850hPa),  $gZ$  is the potential energy ( $g = 9.8 \text{ m s}^{-2}$  is the gravitational acceleration and  $Z$  the geopotential height of the 850 hPa)

and  $L_c q$  the latent heat energy associated with evaporation and condensation of water ( $L_c=2501 \text{ J g}^{-1}$  is the latent heat of evaporation and  $q$  the specific humidity at 850hPa). The MSE is a useful measure of instability in the monsoon layer. Eltahir and Gong (1996) show that the dynamics of the monsoon are strongly regulated by the meridional gradient of the boundary-layer entropy and the MSE accounts for entropy as well as geopotential and latent heat. The MSE is widely used to characterize monsoon and boundary layer variation (Dalu et al., 2008; Chakraborty et al., 2006; Fontaine et al. 2008, Peyrillé et al., 2007).

#### **VI.4.a Circulation around the onset**

In this section we propose to characterize the onset by looking at the atmospheric configuration during and around the onset. We create composite fields of wind, MSE and PWAT averaged over periods of three days centered over the onset date ( $t=t_0$ ). That is,  $t=t_0$  corresponds to a 3-day average centered on the onset date, and all lead and lag fields are also three-day averages centered on the indicated day. We follow the evolution of the spatial structure of the composite at a time step of two days starting 6 days prior to the onset ( $t=t_0-6$ ) to 2 days after the onset ( $t=t_0+2$ ). Figures VI.6, VI.7 and VI.8 show such sequences for each region and each of the three onset definitions : the first rainy event, the large scale event and the true onset type 1. At  $t_0$  the onset is characterized by strong southerly wind convergence into Senegal and a general area of positive MSE and PWAT anomalies (with two maxima embedded). All these conditions contribute to provide a warmer air temperature and greater low level moisture supply, which can be expected to enhance convection at the local scale. Déme et al. (2003) have documented the relationship at the daily time step between PWAT and rainfall recorded in Dakar

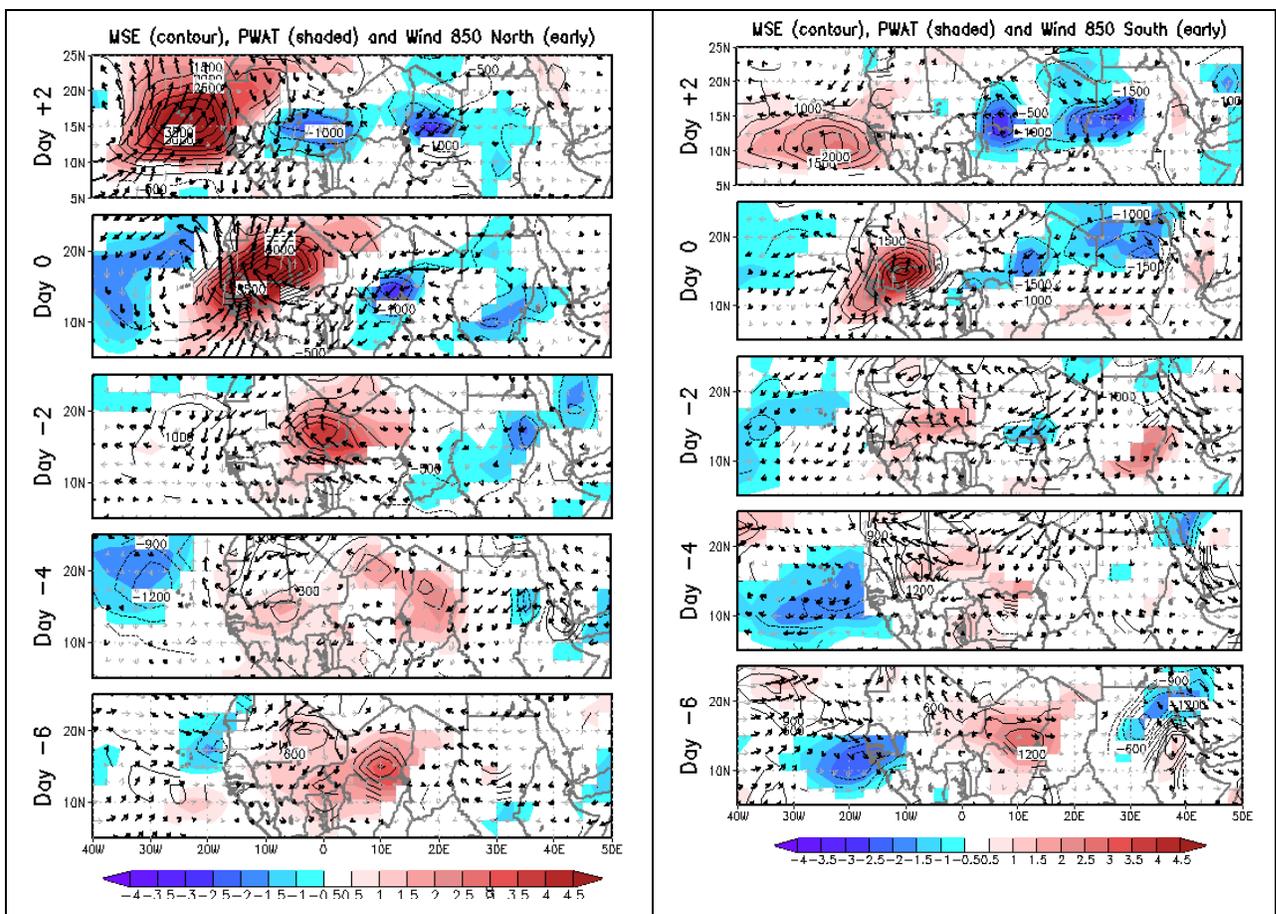
(Senegal) with a maximum of  $r=0.43$  for a time lag of one day, with PWAT leading rainfall. Notice that, to the East and to the West of the positive anomaly of PWAT and MSE, there is a negative anomaly, which is indicative of a large scale wave pattern. At  $t_0$  the low-level wind is characterized by a cyclonic circulation (vortex) centered just off the coast of Senegal, and an anticyclonic circulation inland to the East. These two circulations combine to enhance the magnitude of southerly flow into Senegal.

Another notable feature of Figs. VI.6-8 is that the northern onset composite map shows a clearer signal in this time window, than does the onset map of the southern region. This could indicate a different balance of physical processes behind each region. The northern region may therefore be more related to dynamic wave propagation whereas in the south, where the onset occurs earlier, the onset may be more related to the slowly propagating monsoon surge. This interpretation would be consistent with the large scale SST teleconnection (Figs. VI.4-5).

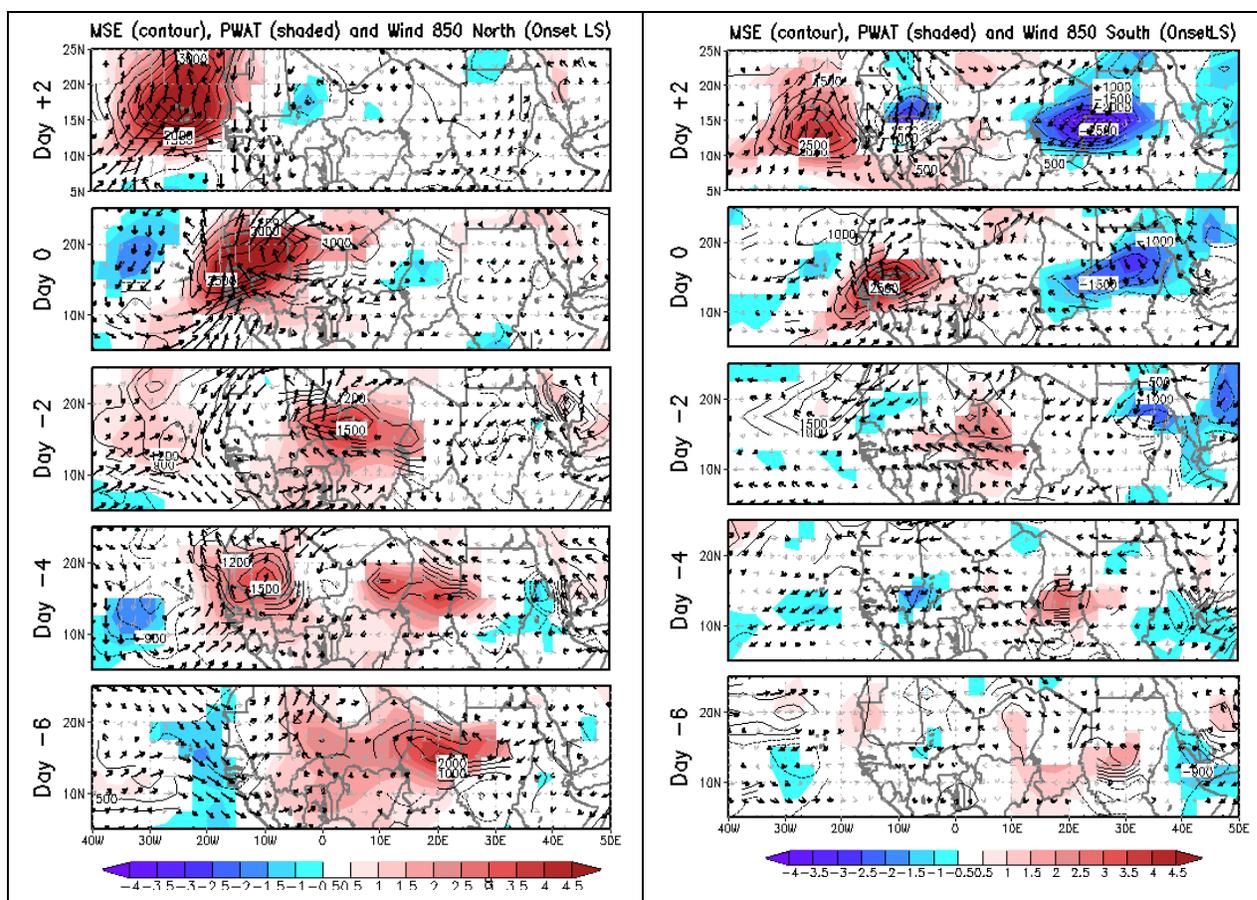
Figures VI.6, VI.7 and VI.8 show, especially for the northern region, a distinct system moving from East to West from  $t_0-6$  to  $t_0+2$  associated with the onset. The progression is confined within a latitudinal band between about 10N and 20N. The signal is strengthening as the system moves westward close to the onset date. All these results show that a robust and well-organized synoptic signal takes place during and before the onset. The nature of the circulation around the onset date is, to first approximation, very consistent regardless of the definition we choose as an onset date. However, for the northern region, which has a stronger propagation signal in this time window, it is possible to detect some variations according to the onset definition type. The first large-scale event has a clearer expression at  $t_0-6$  (even extending to 30E), compared to the first isolated event, which becomes mostly clearly organized at  $t_0-2$  days. The sequence

associated with the true onset is a little less defined in the days building up to  $t_0$ , but shows continuing strengthening at  $t_0+2$  for the MSE and PWAT positive anomalies.

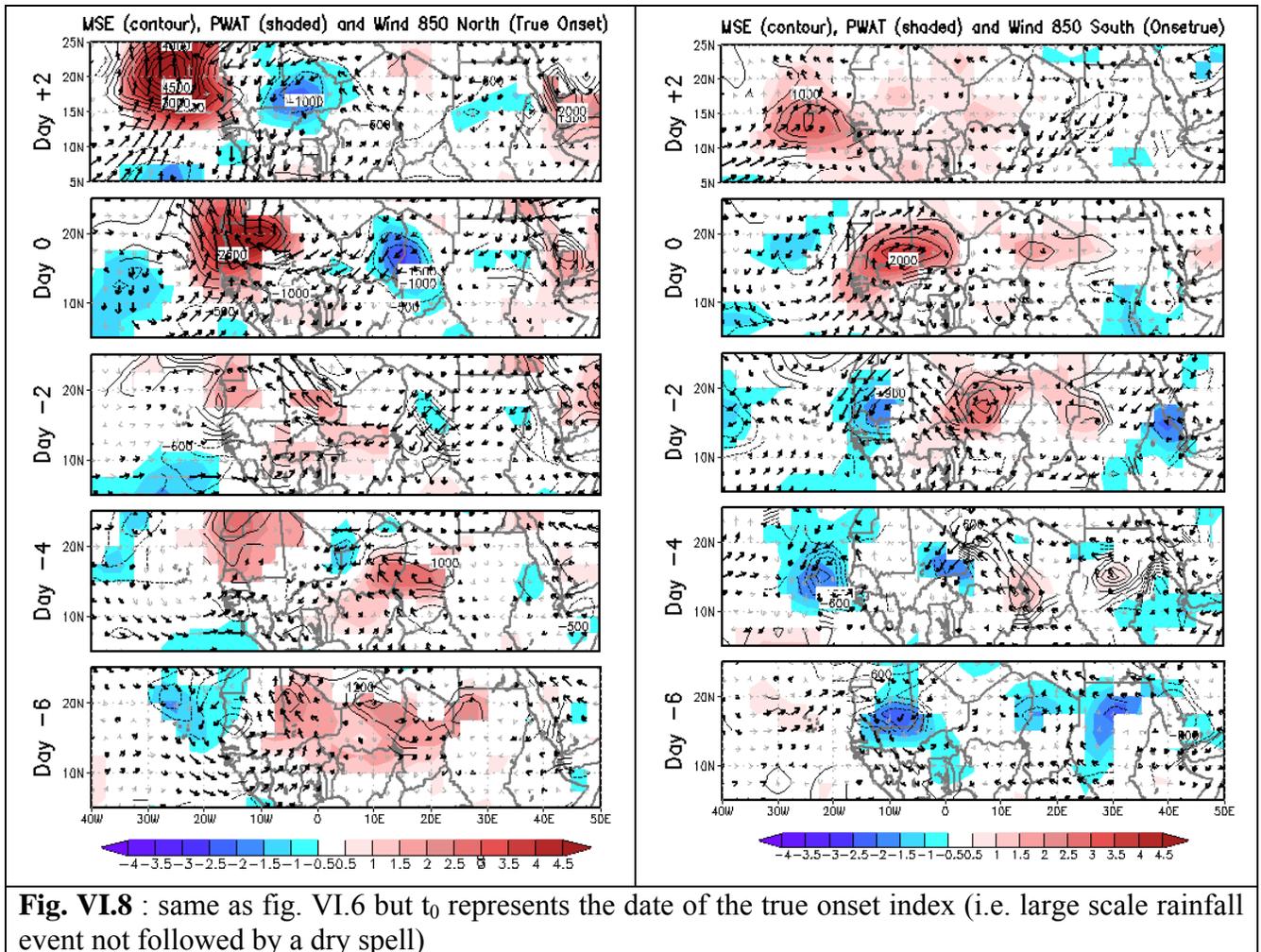
Overall, the results suggest these data are able to detect a strong relationship between local scale (rainfall events over Senegal) and large scale circulation, with strong propagation seen in the northern region. A synoptic and statistical prediction of onset date based on large the scale signal may therefore be feasible. The next section will investigate more on the propagation of the signal at a lead time of up to 30 days.



**Fig. VI.6** : Sequences of composite fields of anomalies of PWAT (shaded), MSE (contour) and 925 hPa wind winds for  $t_0-6$  days, through  $t_0+2$  days, where  $t_0$  represents the date of the first rainfall event index. Left panels are for the Southern Senegal onset index, right panels for the Northern Senegal onset index. Shaded, contoured and bold vector values are when the mean composite value is in the upper 10% or lower 10% of values from all years (sampled from the appropriate sequence location relative to onset).



**Fig. VI.7** : same as fig. VI.6 but  $t_0$  represents the date of the first large scale onset event index (i.e. rainfall event simultaneously across at least three stations).



#### VI.4.b Atmospheric propagating features around the onset

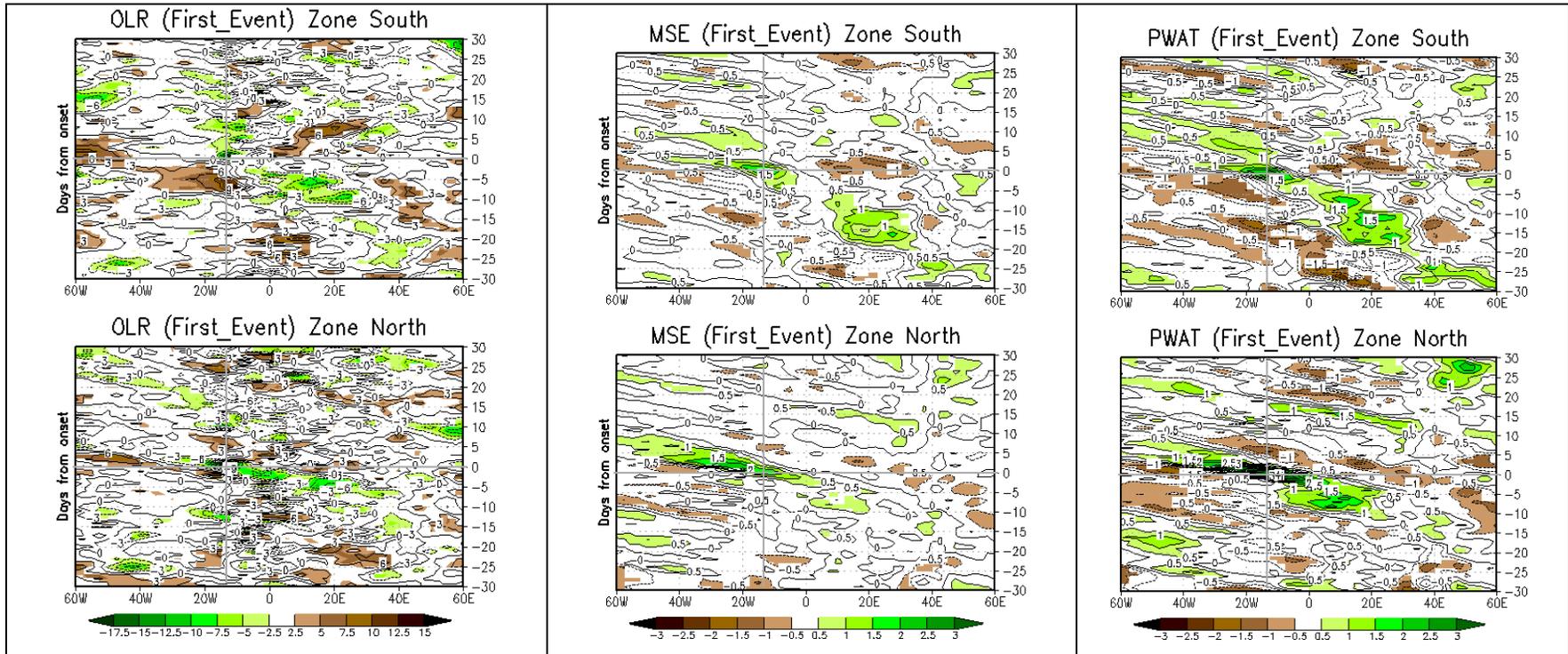
As the propagating signals related to the onset appear confined within a narrow band (10N to 20N), we compute the longitude-time evolution of the averaged signal between these two latitudes to investigate further the propagation around the onset. Figure VI.9 shows the hovmöller diagram depicting such propagation for PWAT, MSE and OLR, at  $\pm 30$  days around the onset date ( $t_0$ ).

For the northern region, the result found in the previous section is clearly visible, and the propagating feature around the time of onset is confirmed to be relatively fast moving,

first emerging at 5-10 days before onset. The extension of the time-window to +/- 30 days adds little to the figures for the northern region.

However, for the southern region, the signal progression is much more organized and clear over a longer time-period, extending beyond the 6 days studied in the previous section. It is present in the large-scale onset chart as well (not shown). The MSE and PWAT positive anomalies can be traced from the time of onset over Senegal, back to 20-30E, 15 to 20 days prior to the onset. The propagation path is very clear and it takes 10 days approximately to travel from 20E to 15W which corresponds of a phase speed of 37 degrees in 15 days ( $\approx 2.8 \text{ m s}^{-1}$ ). This is much less than the squall line speed  $6 \text{ m s}^{-1}$  to  $33 \text{ m s}^{-1}$ , and is notably less than the speed typically associated with easterly waves. By contrast, the wave speed for the north is 35 degrees in 5 days which is  $8.5 \text{ m/s}$ , which approximately matches the speed of the easterly wave of 3-5 days centered at 15N, as defined in Janicot et al. (2008).

The rainfall onset over the southern region of Senegal is clearly more related to a large scale slow moving westward system, with anomalous high supply of PWAT and MSE. The supply of moist and warm air can be traced back clearly to 20 days in the composite, and possibly even up to 30 days extending across 40-60E. More clearly defined, there is a build up and strengthening of the system 10 to 20 days before the onset around the latitude 20E. Around 20E the energy covers a large area and increases in magnitude to reach a certain level then progresses eastward. We can see in the similar plot for the OLR that there is high convective activity taking place with perhaps a few days lag after the energy maximum at 20E. It may be possible to use this slowly moving feature as an early warning signal of the onset in the southern region.



**Fig. VI.9** : Time-longitude variation of OLR, MSE and PWAT anomalies averaged over 10N-20N, from +/- 30 days around the first event onset date for the southern region of Senegal (top panels) and northern region of Senegal (bottom panels). Vertical grey line locates the central position of Senegal, and horizontal grey line the onset date.

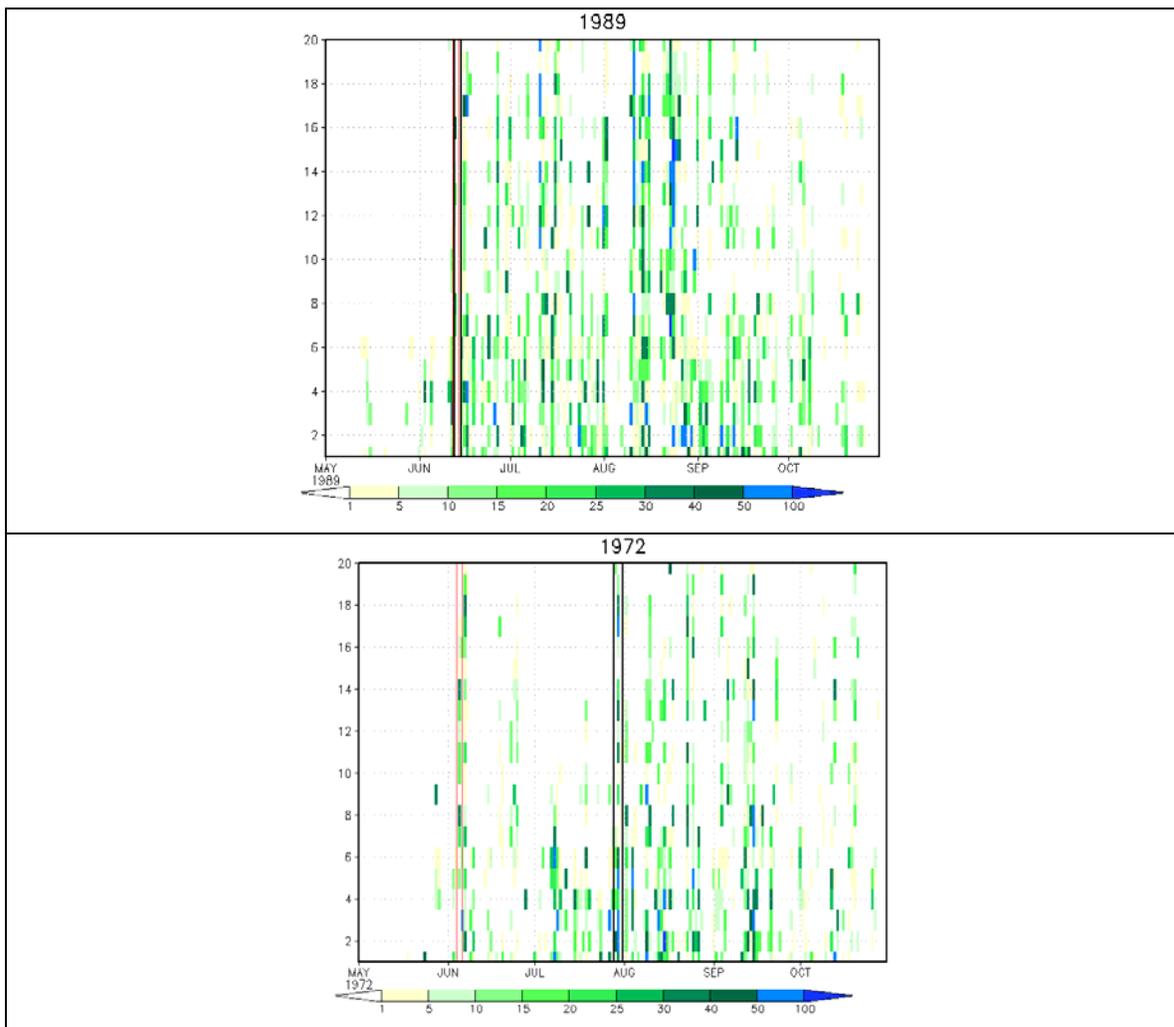
### **VI.5 True versus false onset date**

We have documented in section VI.2b the high frequency of false onset in both the northern and southern regions of Senegal. Any first rainfall event, even at large scale, is likely to be a false start of the rainy season, at least to some extent. Knowing the impact of false onset on agriculture in general and crop growth in particular we will now study separately the false onset and true onset. The goal is to differentiate them based on any precursor atmospheric signals. To set the stage, Fig. VI.10 shows an example of a very clear false start (1972) and a very strong true start (1989). We wish to be able to give warning when there is a risk of a start like in 1972, and to give early information on an expected good start, such as in a year like 1989.

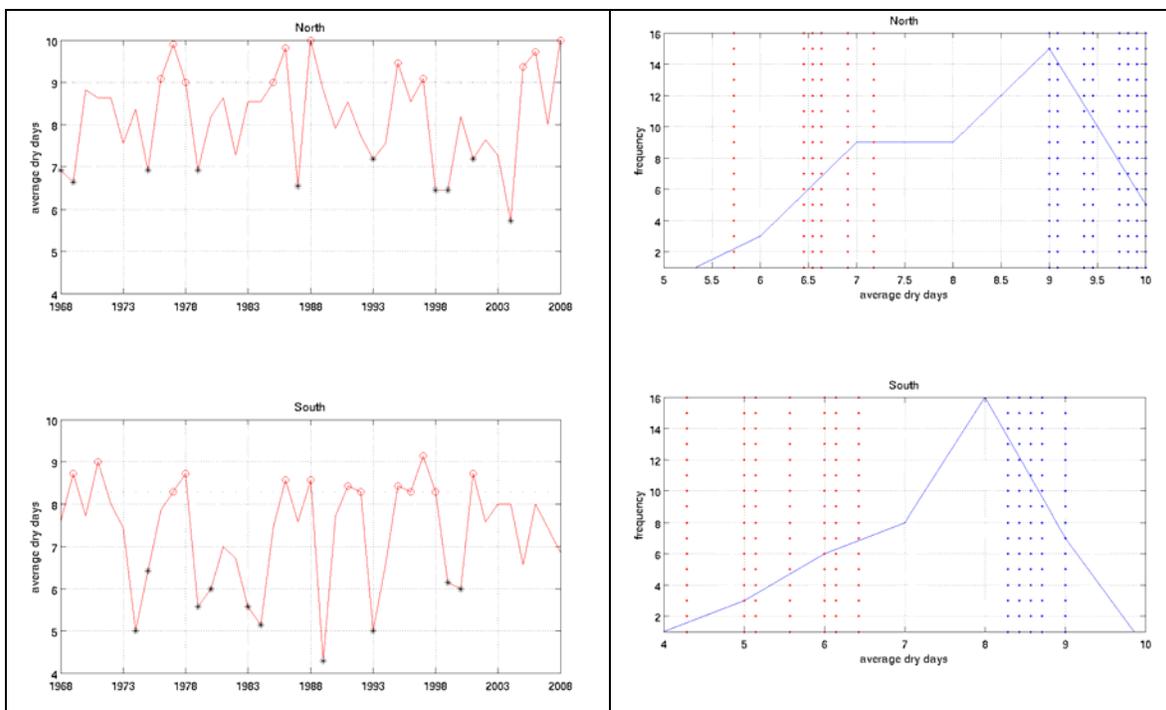
We create composite maps around the rainfall onset event, for (i) events that turn out to be false starts, and (ii) events that turn out to be true starts. For this separation, we have found using true and false type 2 definitions to be the most effective. The separation between false and true onset is therefore based on the number of dry days averaged across the region, following the first large scale rain event. Figure VI.11 shows the time series and PDF of the mean number of dry days in the 10 days following the large scale event. The number of dry days just after the large scale onset is a good indication of the severity of the dry spells and therefore, the severity of any false start.

The graphs show that there is a clear distinction between years of high number of dry days following the large scale onset, and years with low number of dry days. Obviously, there are generally less dry days in the south. However, the south region still clearly shows years when there are many dry days after the first rain event, even averaging

between 8-9 dry days out of 10. The years with the most dry days (upper quartile) on Fig. 11 are taken to define the most severe false starts, while the years with the least dry days are taken to define the best true starts. Fig VI.12 shows composite maps of the difference between the severest false starts (upper quartile) minus the best true starts (lower quartile).



**Fig. VI.10 :** Illustrative example of daily rainfall during a season experiencing a severe false onset (1972) and a season with a good start (1989). Stations used on the y-axis are the filled circles shown in Fig II.1. Vertical red and grey bars mark the first rainy event and the true onset event respectively.



**Fig. VI.11** : Left panels: Average number of dry days observed in the 10-days after the first large scale onset event. Top left is for the Northern Senegal region, bottom left is for the Southern Senegal region. The years with the highest (open circles) and lowest (stars) number of dry days are classified as severe false onset years and best onset years respectively, Right panels: frequency distribution of the series shown in the left panels. The severe false start years are indicated by blue dots, and the best onset years are indicated by red dots.

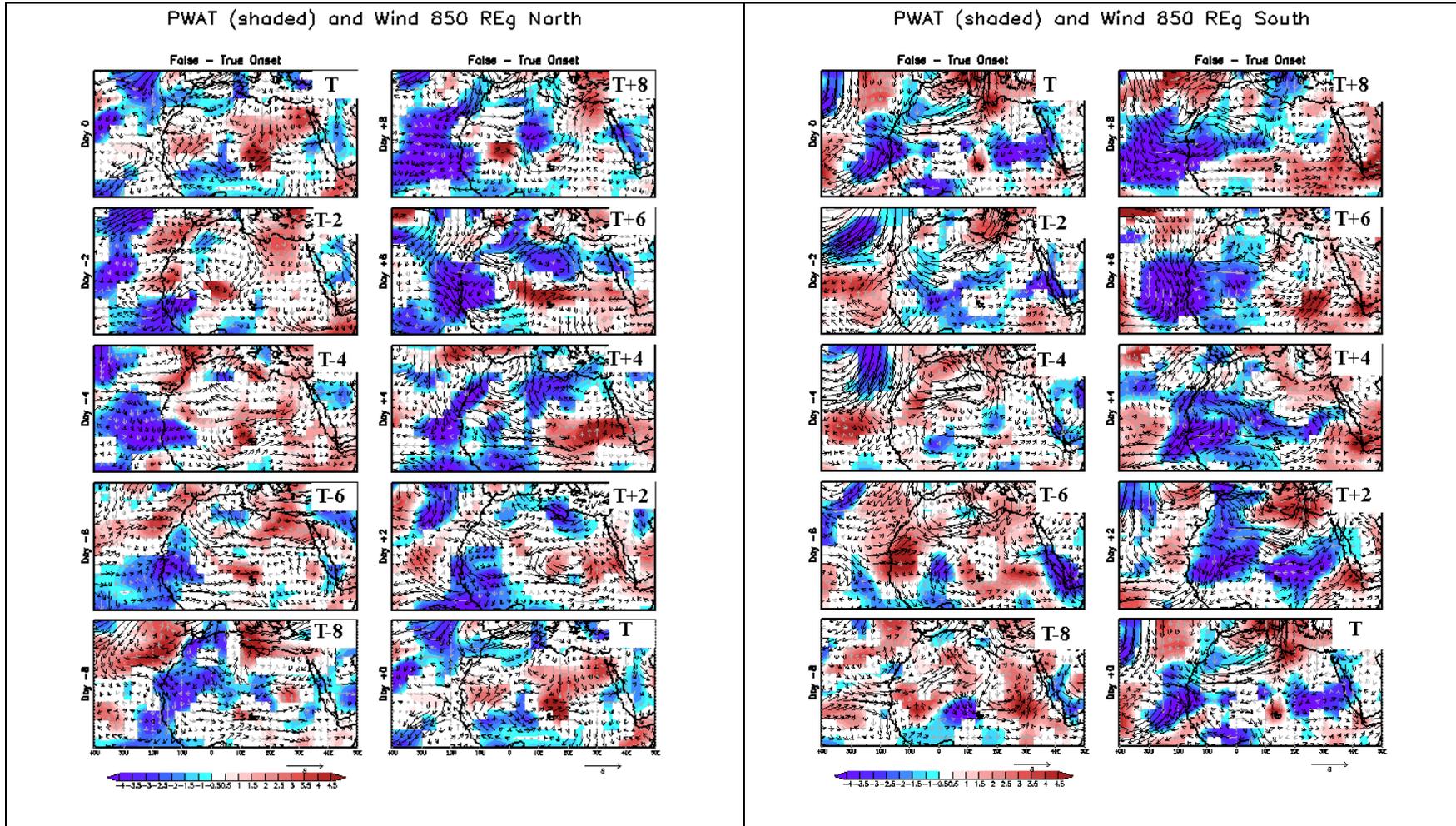
The aim is to see if there is any systematic tendency which discriminates false versus true onset. This way of compositing will allow detection of any ‘danger signals’ to watch out which might increase the risk of a false start.

The striking feature in fig. VI.12a is the high negative anomaly of PWAT located at various times in the sequence, over and to the west of the region. The sequence ahead of the rainfall onset is clearer for the northern region, and this will be focused upon first.

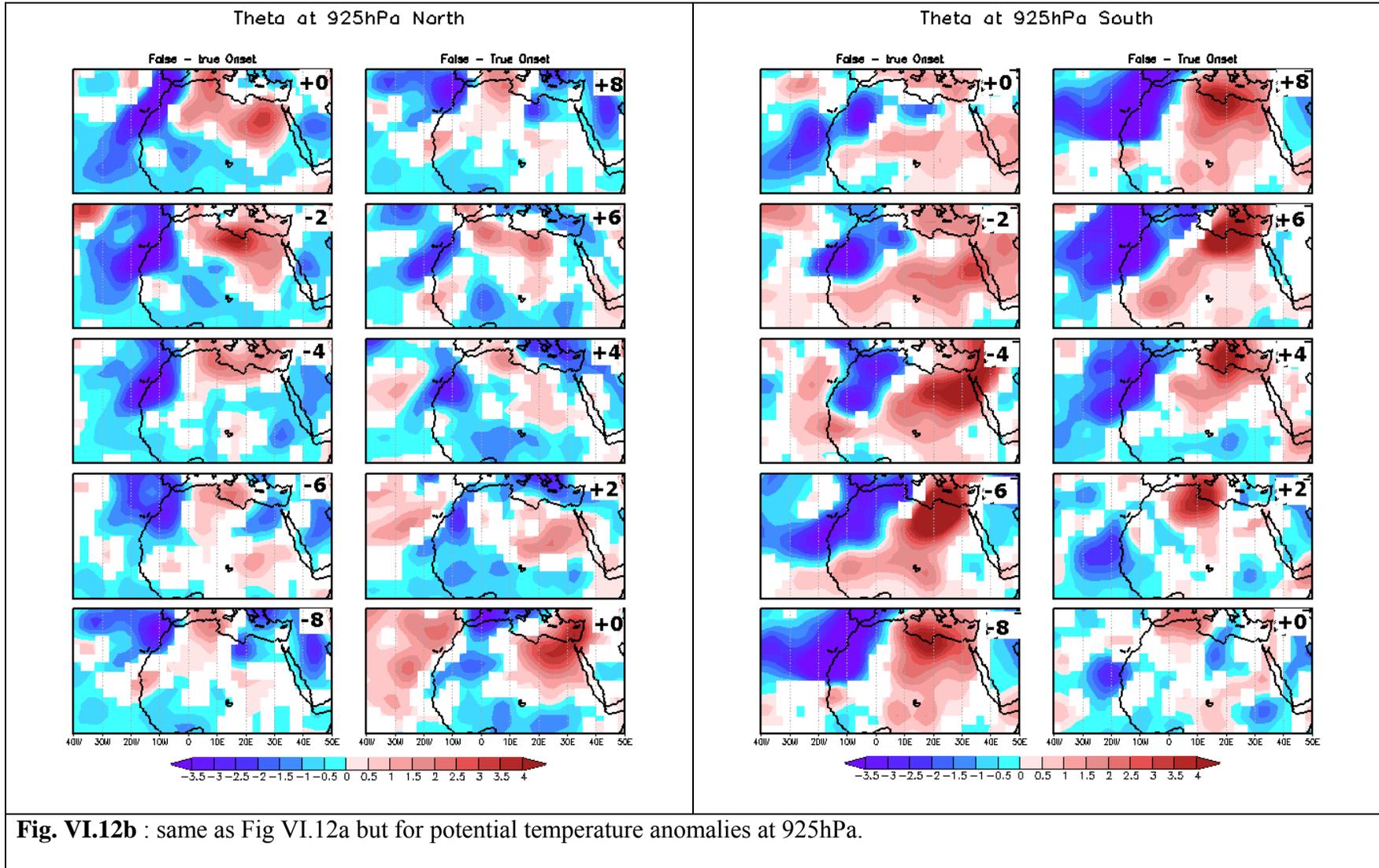
Well before the onset event date there is a systematic anomaly of low PWAT (blue color) located in the vicinity of Senegal and surrounding areas. During the northern region

onset, the negative anomaly of PWAT is seen 8 days prior to the onset, when it extends further North and East. It persists, shifting its axis southward and westward, until two days before the onset date. This negative anomaly of PWAT is concomitant with a northeasterly wind South of Senegal indicative of reduction in the monsoon flow which is southwesterly. Those two signals are very internally consistent, since monsoon flow supplies water vapor. However, the reduction in PWAT does not look to be solely related to the wind flow. The existence of low PWAT values over Senegal and extending over the tropical North Atlantic appears to be a warning sign that an impending rainfall event may correspond to a false start.

During the time of the rainfall onset event, areas of reduced PWAT values remain in the region, but sit to the North and South of Senegal. The low values to the North are over the eastern North Atlantic and to the South are over Equatorial West Africa. Two days after the onset, negative PWAT anomalies become reestablished over Senegal, and a strong northerly / northeasterly wind emerges in the composite. This N and NE wind is advecting cold air from the North, and this is seen in Fig VI.12b which shows the potential temperature anomalies at low level, 925 hPa. This cold air surge at low levels will contribute to stabilization of the air and having an inhibiting effect on the monsoon surge while killing any convection. The relationship between mid-latitude and tropical air interaction has been well documented in the literature (Vizy and Cook, 2009; Mounier et Janicot, 2004; Knippertz 2007; Roca et al., 2005).



**Fig. VI.12a** : Composites showing systematic difference in sequence between false onset (rainfall event followed by significant number of dry days) and true onset (continuous rainfall after rainfall onset event). Sequence is shown for  $t_0-8$  days, through  $t_0+8$  days, where  $t_0$  represents the date of the first large scale rainfall event index. Left set of panels is for Northern Senegal region, right set of panels is for southern Senegal region. Anomalous PWAT is shaded and vectors show wind at 925 hPa.



For the southern region, there are weaker precursors showing up in the composite (Fig. VI.12a) at the longer leads of day-8 and day-6. However, at day-4, there emerges strong northerly flow, with the same implications for convection as discussed above. There is a clear cold air surge (Fig. VI.12b) with strong negative PWAT anomalies. These become established over Senegal and intensify following the onset date. Therefore, for the southern region, a slightly different picture emerges, with less general PWAT anomalies in the days leading up to a false start, but a very clear cold surge event taking place to the North, appearing to signal an increased risk that a rainfall event may correspond to a false start.

In West Africa, the water vapor variability is indicative of interaction between phenomena of various time and space scales, such as moist convection, wave activity, dry intrusion and monsoon flow. Any deficit in water vapor supply can affect these phenomena. Fig VI.12a shows a clear signal of PWAT prior to the onset related to false onset over the northern and southern region. A close monitoring of the PWAT during the time of onset could help identify incoming rain systems as likely representing true or false starts. The methodology presented here can be implemented easily in an operational framework, all that is needed are the climatological fields and close monitoring of atmospheric conditions around the onset. In particular three aspects to watch for are : 1) the PWAT deficit over and nearby, especially West, of Senegal (nearby tropical Atlantic) 2) the  $\theta$  at 925hPa North of Senegal and 3) the northerly wind operating on the temperature anomaly, with cold advection into Senegal.

We have diagnosed the atmospheric conditions occurring around the onset, as well as the onset's teleconnection with the large scale boundary layer including SST and wind. We will now try to look for any predictability of the onset using these diagnostic studies.

### **VI.6 Toward predicting the onset date**

In this section we will try and develop models to predict the onset of the rainy season in each region of Senegal. First we will explore the method proposed by Omotosho (1992) which is based on the vertical wind shear before the onset. Although this approach was developed for the central Sahelian zone, it is natural to explore whether this variable is applicable to the western most part of the Sahel, as represented by indices here. Next, we will build upon the relationships between rainfall onset and ocean-atmosphere coupling, as identified in section VI.4. The propagation of PWAT and MSE around the onset, especially the signal building up 10 to 20 days before onset around 20E, seems to be a precursor signal for onset and offers another tool set for an early warning system based more on synoptic evolution, and with a lead-time of weeks rather than a month or so, as is being sought through the ocean-atmosphere coupling. Finally the CFS coupled model skill is explored for predicting the onset using an ensemble of simulations (Saha et al., 2006). A coupled model, used in this way, could potentially capture skill from SST-forcing / atmosphere-ocean coupling, and from synoptic/dynamic atmospheric evolution on the timescale of weeks. All these approaches here are explored and reviewed, for an indication of whether such forecast systems offer the prospect of being feasible and reliable.

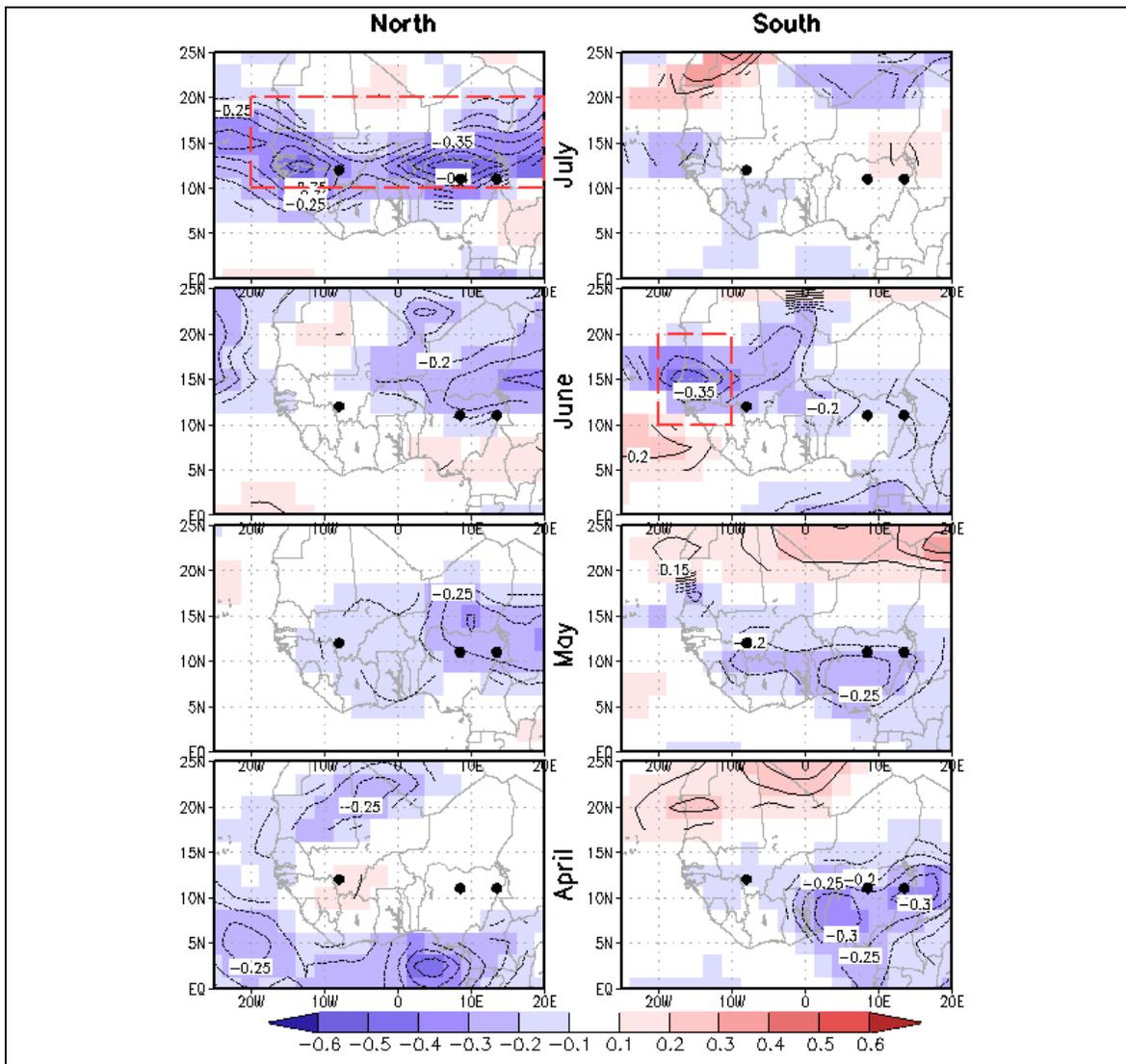
### VI.6a Precursor signals in the annual cycle

Omotosho et al. (1992) have established an onset prediction scheme for some stations in West Africa (Bamako, Kano, Maidiguri) using upper-wind daily data, with more than two months lead-time on the onset. Fig. VI.13 also shows the locations of the stations studied by Omotosho. The method builds on the relationship between Meso-scale Convective Systems (MCS) and rainfall over West Africa as MCS are themselves controlled by the vertical wind shear associated with the African Easterly Jet (Omotosho, 1987). The vertical wind shear empirical conditions needed to indicate pending onset are: that the wind shear below the AEJ (surface and 700hPa,  $\Delta U_L$ ), and above the AEJ (700 and 400 hPa,  $\Delta U_M$ ) simultaneously pass a certain threshold which is,  $|\Delta U_L| - 2\Delta U_M \geq 0$  (Omotosho, 1992). The rationale behind these wind shear requirements is the establishment of both the AEJ and the TEJ, and the poleward retreat of the subtropical jet, which creates a rather sudden wind change from westerlies to easterlies above 400hPa. It is argued that onset occurs 10 weeks after these wind conditions become established at a given location. This approach was presented first at ACMAD during 1994 (Omotosho, 1996). This technique in practice suffers from lack of observed upper-wind data over West Africa.

In this study we test this method by calculating the zonal wind shear threshold over West Africa using NCEP reanalysis, limiting ourselves to the latest period of reanalysis data which is considered to be the most reliable. For each grid box, we calculate  $|\Delta U_L| - 2\Delta U_M$  as described above. We correlate Senegalese onset indices (the area-average standardized indices) with the wind shear indices, calculated at all reanalysis grid boxes (Fig. VI.13).

The correlations are generally negative which agrees with the way the wind shear index was developed by Omotosho (1992) to anticipate onset. It is also encouraging that the strongest relationships are found in July for the northern region (around the time of the onset in the northern region) and in June for the southern region (around the time of the onset in the southern region), and the correlation values are larger over Senegal. As presented here, these are primarily diagnostic relationships. They encourage further analysis of the reanalysis data at finer temporal resolution. At the longer-lead monthly resolution presented here, it is not clear that there are any predictive relationships. For the southern region, a belt of negative correlations are found to the South and East of Senegal in April and May that may capture evolution of the wind shear in the annual cycle and warrant further investigation.

The diagnostic relationships in Fig. VI.13 have been used to assess potential predictability of onset from the wind shear. The regions identified by red-dash lines in Fig. VI.13 were selected and an EOF analysis is applied over each region to capture the signal in the wind shear indices. For the north region, the selected area is 10N-20N and 20W-20E in July. For the south region, the selected area is 20W-10W and 10N-20N during June. Applying a cross-validated EOF regression, the skill for southern region onset using June wind shear is  $r=0.26$ . The skill for the northern index based on wind shear during July is  $r=0.35$ . Therefore, there appears to be some potential predictability from this wind shear approach using reanalysis data. Limitations found here may be related to the reanalysis data used, or that the approach is less applicable to the westernmost part of the Sahel.



**Fig. VI.13** : Correlation between the large scale event onset index and the wind shear index constructed from NCEP reanalysis wind. Correlation maps are shown for the wind shear index calculated in each month from April to July. Correlations are calculated for 1981-2008. The red box indicates the area used for creation of wind shear indices to use as predictors for onset.

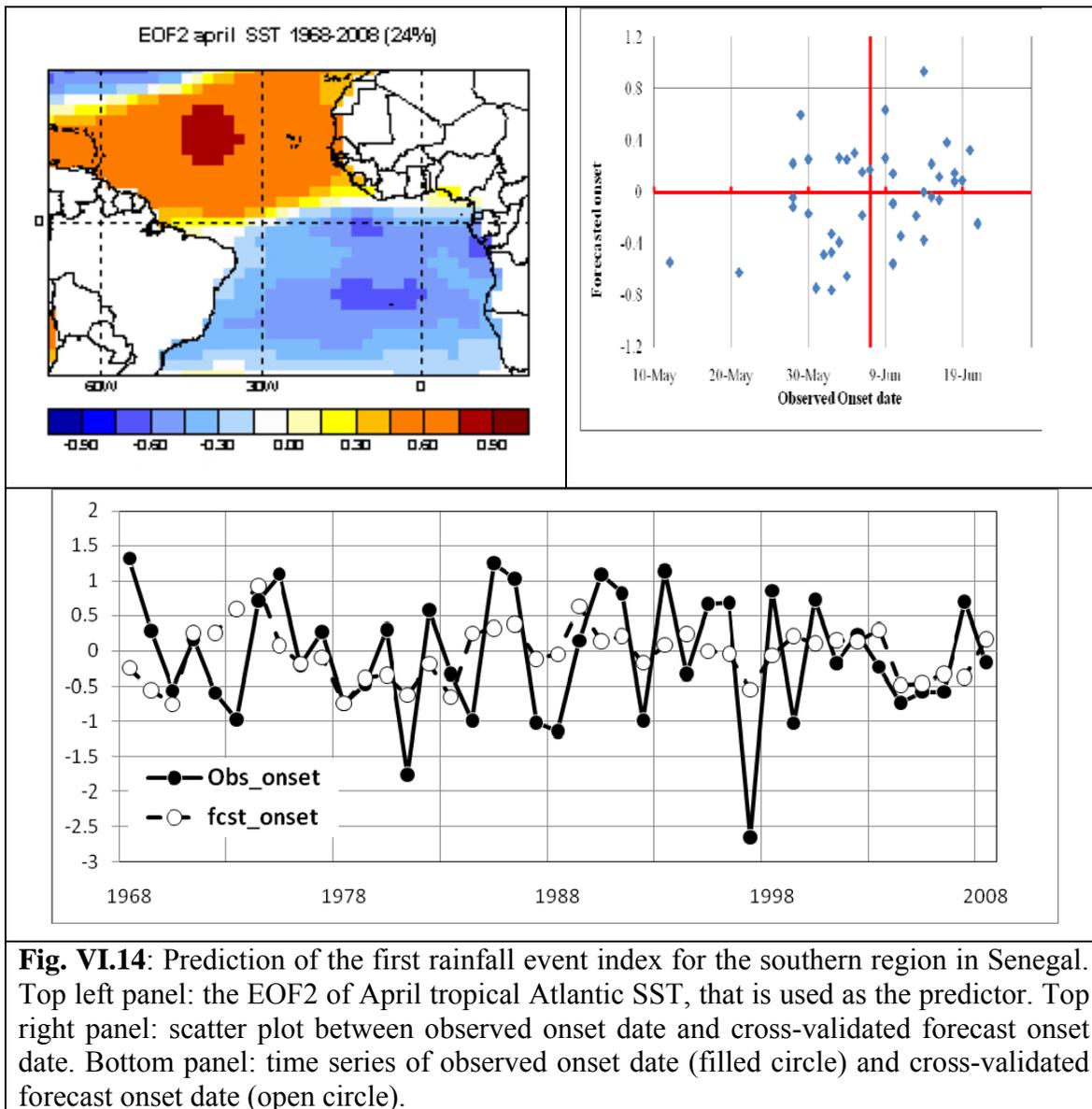
### **VI.6.b Using large scale ocean/atmosphere relationships to forecast the onset date**

The results found in section VI.4 suggest a relationship between rainfall onset in the southern part of Senegal and both SST and Wind at 925hPa in April over the Atlantic, whereas for the northern region, large scale onset is related to near global tropical SST variation in July. We will establish a regression between each onset index and its relation to the large scale signal.

For the southern region the second EOF (EOF2) of the SST over the tropical Atlantic basin bounded by 30N and 30S (fig. VI.14) captures very well the signal explaining the variability of the rainfall onset. The EOF2 pattern has strong and opposite weight on each side of the equator which captures the North-South gradient that modulates the monsoon flow. EOF2 is used to predict the area-averaged standardized onset index for the southern part of Senegal. Figure VI.14 shows the observed and cross-validated forecast time series. The cross validated Pearson correlation skill is  $r=0.34$  over 1968-2008 and 0.43 over 1981-2008.

There is a reasonable match between the predicted and observed time series (Fig. VI.14). The same signal is captured by the first EOF of the wind in the same tropical Atlantic sector with a skill of  $r=0.46$  and  $r=0.34$  for respectively 1968-2008 and 1981-2008. The dipole of SST and near-surface wind over the tropical Atlantic in April is therefore a good indicator of the first rainfall event in the southern part of Senegal. This is consistent with the relationship between monsoon flux and the SST Atlantic dipole (Coëtlogon et al., 2010; Okumura et Xie, 2004; Reason et Rouault, 2006). The skill obtained using the

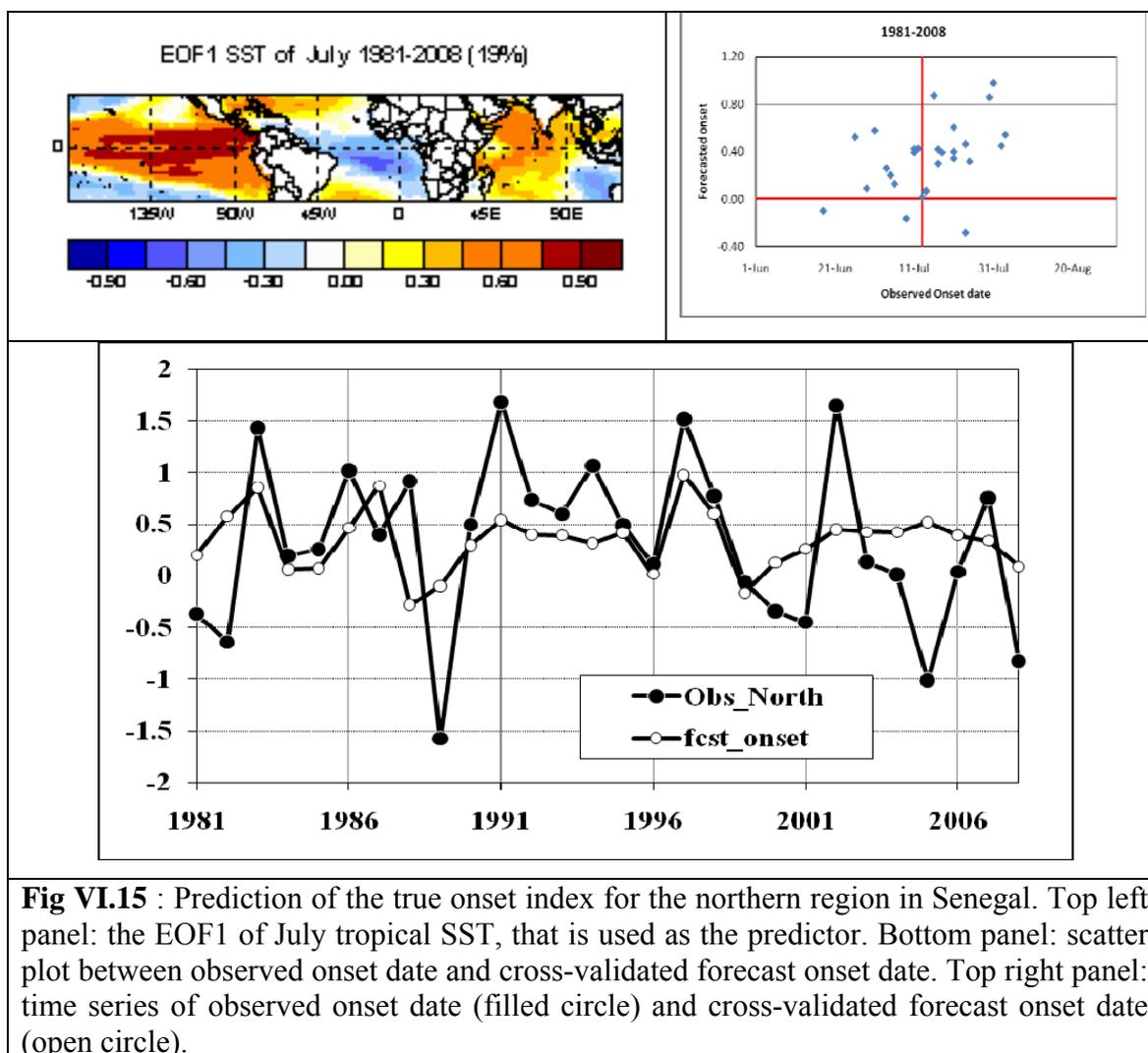
SST from May to July declines to near nil (Table VI.4) consistent with the climatological occurrence of the onset date over the southern part of Senegal, i.e. the SST after the onset is not related to the onset date.



**Table VI.4** : Skill of predicting the first rainfall event index in the southern region of Senegal using the EOF2 of tropical Atlantic SST. Cross-validated correlation skill is shown for 1981-08 (top row) and 1968-08 (bottom row). The SST used to make the forecast is indicated in the columns.

<b>First event South</b>	<b>Apr</b>	<b>May</b>	<b>Jun</b>	<b>Jul</b>
<b>R (81-08)</b>	0.34	0.24	0.15	-0.035
<b>R (68-08)</b>	0.46	0.23	0.15	0.022

Over the northern part of Senegal, the true onset date is the index that is most strongly related to near global variation of the SST, which is encouraging from a practical perspective (and all indices display the same basic relationship). Table VI.5 shows the cross-validated correlation skill score for true onset date, using large scale SST predictors for April to July. Opposite to the timing of the skill for the southern region, the skill is higher later in the season in July, close to the climatological time of the true onset (see table VI.1). Fig VI.15 summarizes the prediction model and cross-validated skill between large scale July SST and observed true onset date. Using July SST is a diagnostic result (no lead-time), but results in Table VI.5 suggest some skill in June SST (cross-validated  $r=0.36$  with 5 EOFs, 1981-2008) and even May SST (cross-validated  $r=0.30$  for just 1 EOF, 1981-2008).



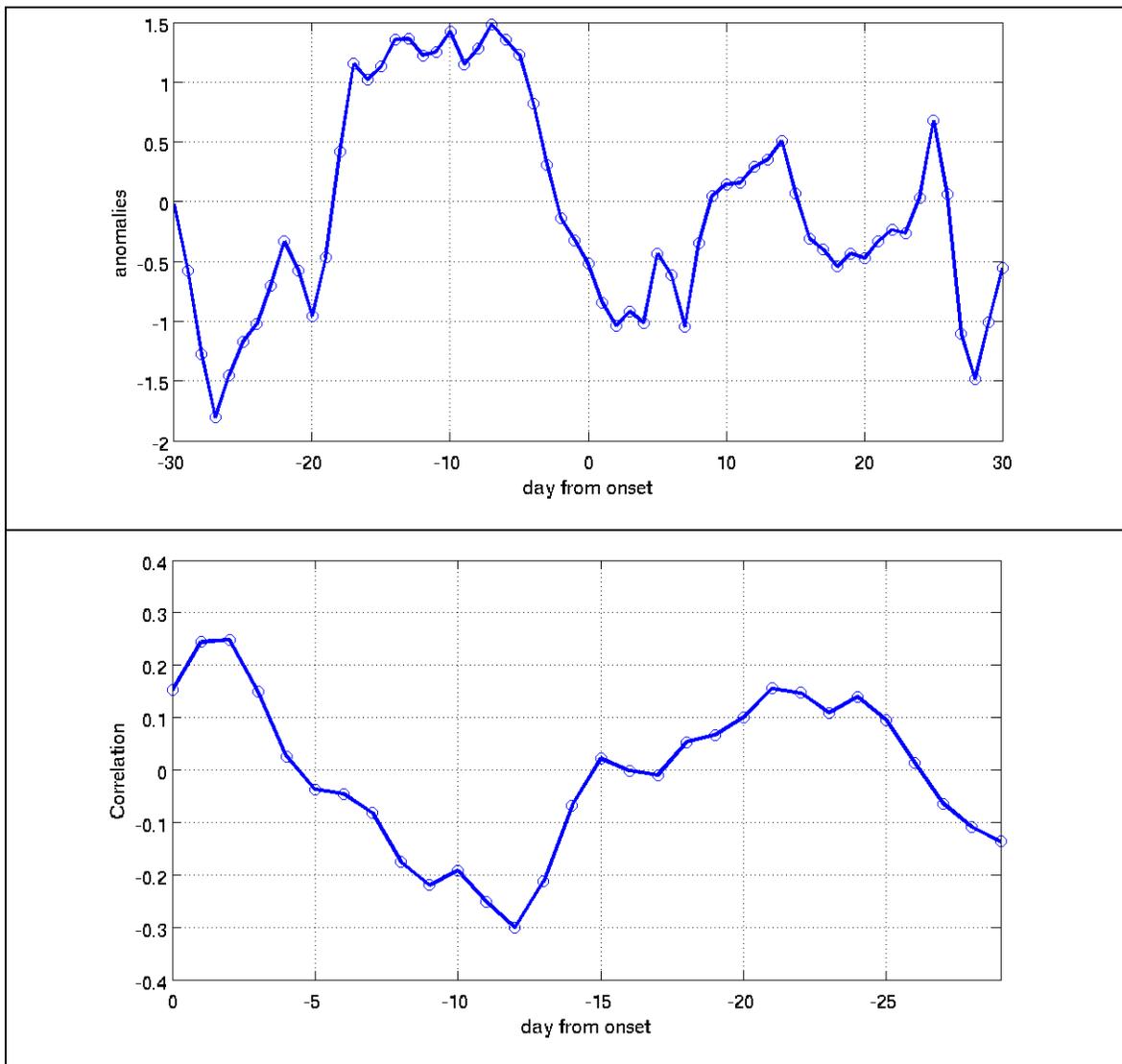
**Table VI.5** : Skill of predicting the true onset event index in the northern region of Senegal using EOFs of tropical SST. Cross-validated correlation skill is shown for 1981-08 (top row) and 1968-08 (bottom row). First value shown is for skill with the first EOF. If more than one EOF provided the best cross-validated skill, that is shown as the second value (with the number of EOFs used in brackets). The SST used to make the forecast is indicated in the columns.

True Onset North	Apr	May	Jun	Jul
<b>R(68-08)</b>	0.06/0.22(2)	0.19/0.33(5)	0.30	0.27/0.36(2)
<b>R(81-08)</b>	0.12/0.23(2)	0.30	0.28/0.36(5)	0.43

### **VI.6.c Using propagating signals as an early warning**

In section VI.5.b we presented the PWAT building up around 20E and its progression eastward in association with the rainfall onset in southern Senegal. In this section we try to build upon these results toward a synoptic early warning system for the rainfall onset. Fig VI.16a shows the mean PWAT anomaly (averaged over 20N-10N and 20E-10E) evolving during +/- 30 days around the onset date (first rainfall event) over the southern part of Senegal. We can see clearly a maximum of the water vapor 7 to 15 days prior to the first rainfall event over southern Senegal. This further confirms the result in section VI.5.b seen on the hovmöller diagram. The water vapor in this region seems to be critical to the onset. One factor may be the land-surface source of water vapor through plant transpiration. When we compute the daily lag correlation the maximum correlation magnitude ( $r=-0.3$ ) is found around 12 days before the onset event (Fig. VI.16b). This shows a relationship, albeit weak, between early onset and PWAT in this region. Future work can explore variations on the result shown in Fig. VI.16 for different locations or combinations of locations, given the promising result in Fig. VI.9.

The objective is not to build a model per se with just of order 10% of explained variance, but to provide preliminary results of a conceptual model. Nonetheless it can offer an early signal for the first significant rainfall event in the southern region. It's good to note in all these calculations we use the full period of 41 years which might contain poor data in early years, years where the onset signal is weak, or years when the first event was due to other phenomenon like the Atlantic dipole as seen in section VI.4.



**Fig. VI.16** : Relationship of PWAT averaged over 20N-10N and 20E-10E with the first event onset index in the southern Senegal region. Results are for 1968-2008. Top panel: composite average variation around the onset date. Bottom panel: daily lag correlation with the onset date index.

#### **VI.6.d Forecasting the onset date using CFS**

As presented in chapter V the fully coupled model CFS captures very well the interaction between ocean and atmosphere related to variations in the tropical Atlantic / West Africa in July-September, and permits longer-lead skill than other GCMs studied here, in predicting the seasonal rainfall over the Sahel. In this section, we test the skill of the CFS to predict the onset of the rainy season over Senegal. Another interesting feature of the CFS is it assimilates two kinds of information both relevant for the forecast of the onset date. One is ocean boundary conditions which take into account the slowly delayed forcing of the ocean on the evolution of the atmosphere and the other is the given initial observed atmospheric conditions at a particular date important for a good weather forecast, and perhaps even capturing the slowly evolving synoptic/annual cycle anomaly features (e.g. Fig. VI.9) that are beyond the traditionally viewed timescale limits of weather prediction. Both might have an impact on the onset date depending on if the onset is strongly conditioned by ocean conditions, or due to weather/synoptic propagating systems, or a mixture of both. In this section we choose CFS runs initialized between 30Apr and 04May (before the climatological onset date in both region) forecasting MJJASO at a daily time step (there are two runs initialized for each day and we averaged those two runs to give forecast series starting in each of the five initialization days of 30Apr to 04May). Each of the five forecasts from 30 April to 04 May (five days) is considered here as an initialization run (i.e. there is an ensemble of 5). All these runs are initialized with the same ocean boundary condition taken from the 1<sup>st</sup> of May during each year. Forecasts are made for each year from 1981 to 2008.

First the ability of the CFS to simulate the character of daily rainfall variability over Senegal is checked. Fig VI.17 shows the CFS grid boxes. Four land grid boxes cover the Senegal, two in the northern region and two in the southern region as defined previously. The blue grid boxes in Fig. 17 are represented in the model as the ocean. Fig VI.18 shows time series of daily evolution of the mean annual cycle averaged over the Northern Region (the two northern land grid boxes) and the Southern Region (the two southern land grid boxes). The CFS is capable of differentiating the Northern and Southern regions, rainfall increasing faster in the early part of the season in the southern region, such that the mean annual cycle is very distinct over the two regions. The peak of the rainy seasons are also well captured which is in July-August-September in the South and August over the North. However, we notice that the CFS over estimates the daily rainfall over both regions, and very much more so over the South. There is clearly potential for bias correction to improve the CFS daily rainfall. We will evaluate CFS onset forecasts using both raw CFS data and bias corrected CFS data.

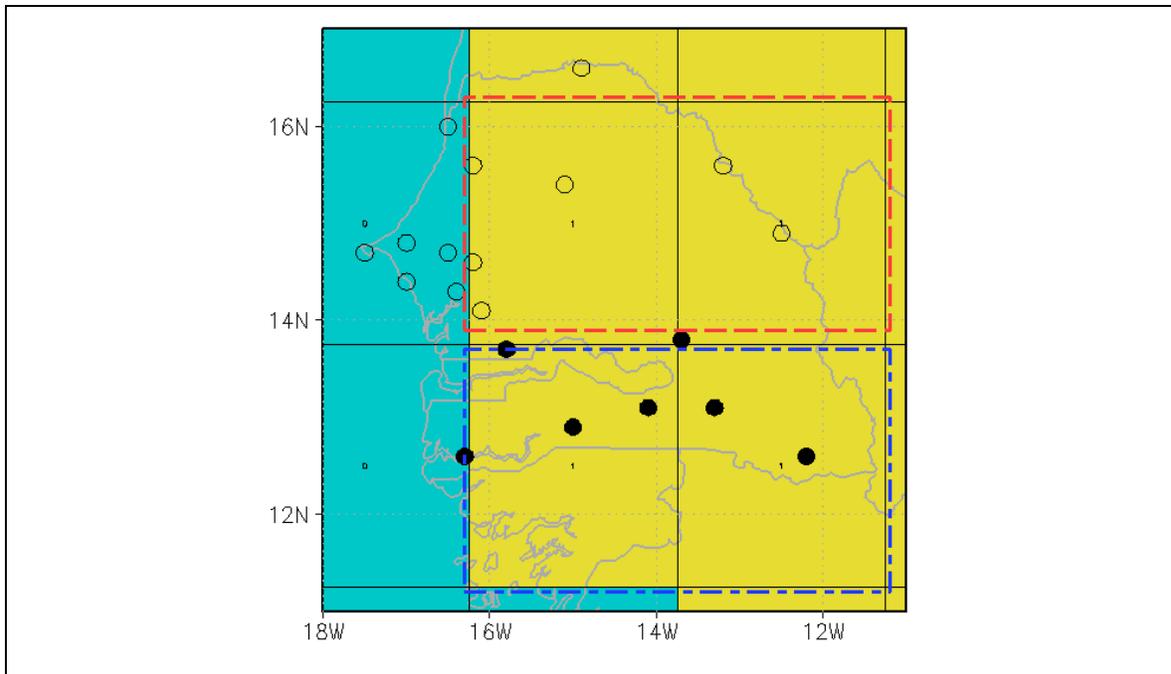
We use a technique developed by Ines and Hansen (2006) to correct the biases in both the frequency and the intensity distribution of daily CFS rainfall. The daily distribution of each GCM grid cell over each region is mapped into each single observed station within the region. The bias of the rainfall frequency is corrected by truncating the empirical distribution of the GCM daily rainfall above a certain threshold calculated from the empirical observed and GCM cumulative rainfall distribution. Thus the mean frequency of rainfall above the threshold matches the observed mean rainfall frequency of each station in the same region. The CFS rainfall intensity is corrected by mapping its intensity

distribution function into the observed intensity distribution function. The mapping is done here by using an empirical-gamma transformation.

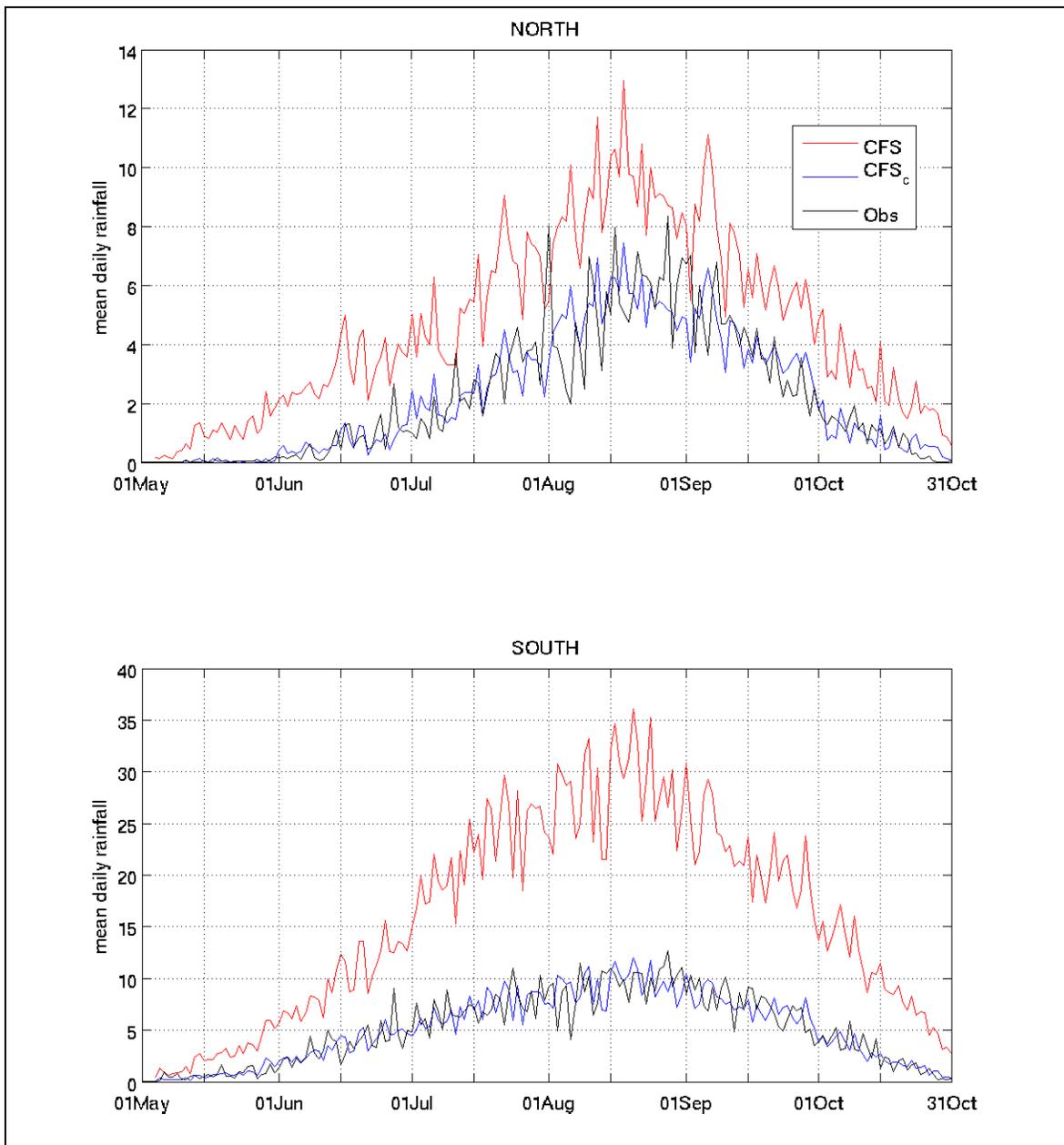
The correction is applied to each of the 5 runs at each grid box in order to match frequency and intensity of each station in the same region. So for each run at each grid box we will generate  $n$  corrected time series,  $n$  being the number of stations in the region. With the 13 stations in the northern region we will get 130 corrected (13 stations x 5 runs x 2 grid boxes) time series and 70 time series for the 7 stations located in the south region. This approach has the advantage to create a sample large enough to capture all the specificity at each station. Fig VI.18. shows that the climatological CFS rainfall is much improved by the correction

Next each corrected time series is used as a station, and the onset definitions are applied. The regional onset dates calculated from these CFS-downscaled time series are compared with regional onset dates derived from the observed stations. Figure VI.19 shows observed onset dates in black, onset dates from direct CFS output (blue) and after applying the correction approach (red). Dots represent the onset for the 5 members x 2 grid boxes over each homogeneous region. The peculiar feature is the spread of the onset date (blue dot) for northern and southern regions. The spread is larger for northern region. It turns out also there is no skill over the northern region. Low skill in the north may be explained by the initialization in end of April beginning of May which is very far to the onset in the northern region. Even if we apply the correction on the time series there is still no skill (Fig. VI.19). The CFS does have some skill in predicting the onset of the

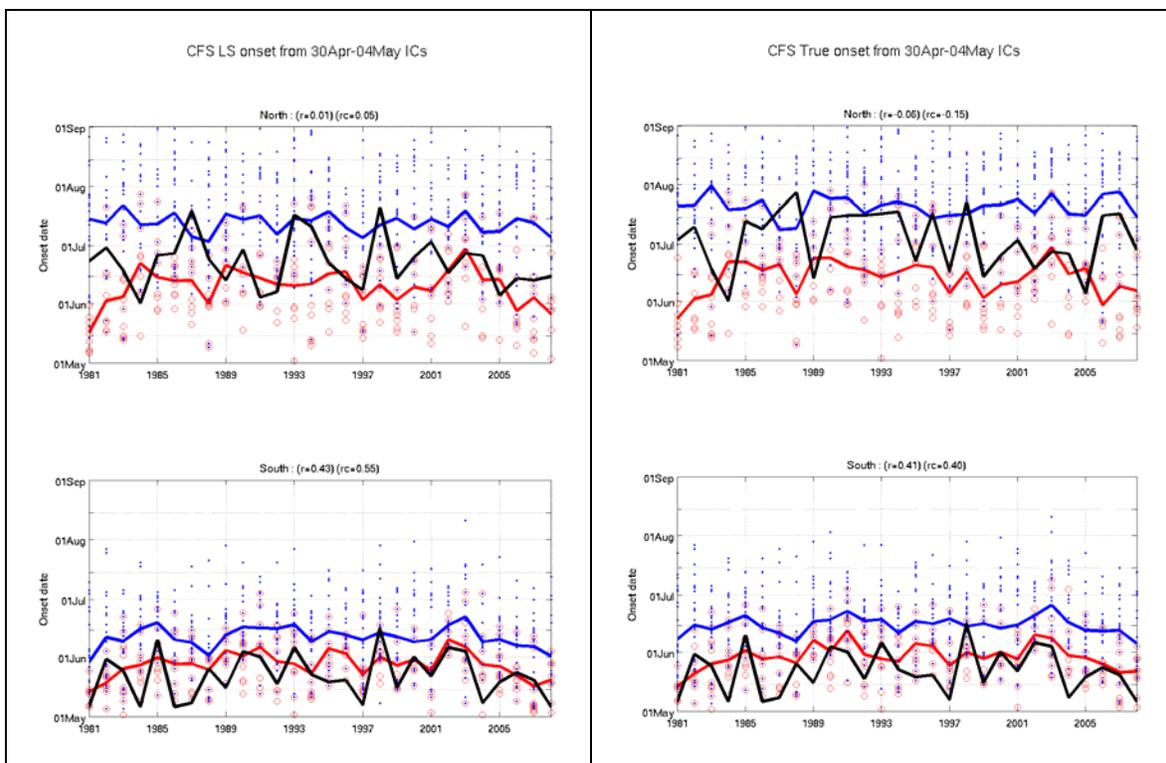
rainy season over the southern region. The skill is  $r=0.43$  and  $r=0.41$  for respectively the large-scale onset and the true onset. After applying the bias correction the skill become respectively 0.55 and 0.40. These represent very encouraging first results on the ability of a coupled prediction system to provide information about rainfall onset in West Africa.



**Fig. VI.17** : CFS grid boxes in the area over and near Senegal. Boxes shaded blue are represented as ocean in the CFS and yellow as land. Red dashed line defines the two CFS grid-boxes used as northern Senegal in the CFS, and blue dashed line defines the two grid-boxes used as southern Senegal. Open circle are stations in the northern Senegal region, and filled circle are in the southern Senegal region.



**Fig VI.18** : Mean Annual cycle during MJJASO for observed (black line), CFS raw output (red line) and CFS corrected (blue line). Results are shown for the Northern Senegal region (top panel) and Southern Senegal region (bottom panel). The CFS is for the average of forecasts initialized during 30 Apr - 04May for each year, 1981-2008.



**Fig. VI.19** : Forecasting onset date with the CFS model. Solid red line is the ensemble average of direct model output, solid blue line is ensemble average of bias corrected model output, and solid black is observed regional onset date. Red dots are for each uncorrected CFS series (5 runs x 2 grid-boxes in each sub-region = 10 dots), blue dots are for each corrected series (after the downscaled estimate for each station is averaged back to the two model grid boxes, therefore also 2 grid-boxes x 5 runs = 10 dots). Top 2 panels are for the Northern region, bottom two panels for the Southern region. Left panels are for large-scale event onset and bottom two panels for true onset date. Values in brackets above each panel indicate the correlation of the observed index with the CFS uncorrected onset (red line) and corrected onset (blue line).

## VI.7 Summary

In this section we identify two homogeneous onset regions over Senegal. In each region, three basic onset indices were constructed : first rainfall event, a large scale onset event, and a true onset which considers whether any long dry spells occurred after the first large scale rainfall onset event. One striking feature that we found is the large number of false onset events. It's quite difficult to find years where the rainfall sequence is completely continuous. The first event is always followed by dry spells.

Using these various regional indices we established connection with the large scale boundary layer including SST and low level wind. The onset over the northern region of Senegal is related to the large scale signal developing in June–July. It includes correlation with tropical Pacific and Indian ocean warming favorable for late onset.

For the southern region the teleconnection is mostly confined over the Atlantic basin with a strong dipole, occurring earlier in the season (around April/May). This teleconnection operates through a strengthening/weakening of the monsoon flow, and we argue this is through the thermal (SST) wind gradient mechanism.

We also looked at the dynamics which takes place at the finer synoptic time scale around the onset. The onset over Senegal is strongly controlled by the large scale atmospheric system with propagating features prior to, and after, the onset. For the northern region, features propagate relatively faster (signals up to 6 days prior to onset), consistent with easterly wave propagation from the eastern Sahel. For the southern region, features

evolve more slowly, with build up of instability in the eastern Sahel 10-20 days prior to onset, and then relatively slow propagation across the Sahel.

Building upon all these findings we explicitly explored predictability of onset. For the southern region of Senegal a good statistical model is established between tropical Atlantic SST and the first rainy event, with a skill of  $r=0.46$  over 1968-2008. For the northern region the true onset is the most predictable using the tropical SST. The skill is  $r=0.43$  over 1981-2008. The SST is a good candidate for forecasting the onset over both North and South Senegal and the skill offered are encouraging, and quite close to the levels of seasonal rainfall total forecast skill in empirical models from April/May SST predictors.

One important aspect we also investigated is the extent to which it is possible to separate and dissociate true versus false onset. The composite analysis of the difference in anomaly shows a clear ‘danger’ signal to watch for. False onset (as compared to true onset) can be characterized by pre-event anomalies of PWAT located at the western fringe of Senegal and the tropical North Atlantic, with strong north easterly wind located North of the region (Mauritania to western Europe) advecting low potential temperature favorable for a cold air surge. This low level cold air surge can act to kill any convection. For the northern region, the PWAT anomalies may be established over the tropical Atlantic and Senegal more generally several days ahead of the “false” onset event. For the southern region, the PWAT anomalies appear to establish a few days ahead associated with a cold surge in the eastern Atlantic from mid-latitudes.

In the last part of this chapter, we took a first look at coupled model representation and forecasts of rainfall onset. We assessed the performance of the CFS model. We found that that CFS is able to separate the basic climatological onset character of these two regions, albeit with positive rainfall bias, especially in the southern region. This initial study focused only on forecasts initialized in early May. There is no skill over the northern part of Senegal, which is challenging given the long-lead nature of such forecasts. For the southern region, the skill of these forecasts is very encouraging, at  $r=0.43$  for the large scale onset and  $r=0.46$  for the true onset. These results give hope for using coupled GCMs to forecast onset of the rainfall. Applying a bias correction on the daily sequence does improve dramatically the representation of the daily PDF. In terms of correlation skill for onset dates, the skill improvement here is small but positive. Further investigation of coupled model forecasts of onset, with and without bias correction, is motivated by the findings here.

## Reference chapter VI and conclusion

- Baron C., B. Sultan, M. Balme, B. Sarr, S. Traore, T. Lebel, S. Janicot, and M. Dingkuhn, 2005 : From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. *Philos. T. Roy. Soc. B.*, **360**, 2095-2108.
- Benoit P., 1977: The start of the growing season in Northern Nigeria. *Agric. Meteorol.* **18**, 91-99.
- Biasutti M., and A. Giannini, 2006: Robust Sahel drying in response to late 20th century forcings. *Geophys. Res. Lett.*, **33**, L11706, doi:10.1029/2006GL026067
- Chakraborty A., R. S. Nanjundiah, and J. Srinivasan, 2006: Theoretical aspects of the onset of Indian summer monsoon from perturbed orography simulations in a GCM. *Ann. Geophys.*, **24**, 2075–2089.
- Coëtlogon G., S. Janicot, and A. Lazar, 2010: Intraseasonal variability of the ocean atmosphere coupling in the Gulf of Guinea during boreal spring and summer. *Quart. J. Roy. Meteor. Soc.*, **136(s1)**, 426-441.
- Couvreux F., F. Guichard, O. Bock, B. Campistron, J. P. Lafore, and J. L. Redelsperger, 2009: Intraseasonal variability of the monsoon flux over West Africa prior to the monsoon onset. *Quart. J. Roy. Meteor. Soc.*, **136(s1)**, 160-174.
- Dalu G. A., M. Gaetani, and M. Baldi, 2008: A hydrological onset and withdrawal index for the West African monsoon. *Theor. Appl. Climatol.*, **96**, 179-189.
- Davy E., G. F. Mattei, and S. L. Solomon, 1976: An evaluation of the climate of water resources for development of agriculture in the Sudano-Sahelian Zone of West Africa. *Special Environmental Report 9. World Meteorological Organization: Geneva, Switzerland, WMO-No459.*
- Déme A., A. Viltard, and P. de Félice, 2003: Daily Precipitation Forecasting in Dakar Using the NCEP–NCAR Reanalyses. *Wea. Forecasting*, **18**, 93–105.
- Eltahir E. A. B., and C. Gong, 1996: Dynamics of wet and dry years in West-Africa. *J. Climate*, **9**, 1030-1042.
- Folland C. K., T. N. Palmer, and D. E. Parker, 1986: Sahel rainfall and world wide sea temperatures, 1901-85. *Nature*, **320**, 602-607.
- Fontaine B., and S. Janicot, 1992: Wind-Field coherence and its variation over West Africa. *J. Climate*, **5**, 512-524.
- Fontaine B, S. Louvet, and P. Roucou, 2008: Definition and predictability of an OLR-based West African monsoon onset. *Int J Climatol.*, **28**, 1787-1798.
- Gash J. H. C., J. S. Wallace, C. R. Lloyd, A. J. Dolman, M. V. K. Sivakumar, and C. Renard, 1991: Measurements of evaporation from fallow Sahelian Savannah at the start of the dry season. *Quart. J. Roy. Meteor. Soc.*, **117**, 749–760.
- Giannini A., R. Saravanan, P. Chang, 2003: Oceanic Forcing of Sahel rainfall on interannual to interdecadal time scales. *Science*, **302**, 1027-1030.
- Hagos S. M., and K. H. Cook, 2007: Dynamics of the West African Monsoon Jump. *J. Climate*, **20**, 5264-5284.
- Hansen J. W., A. Challinor, A. W. M. Ines, T. Wheeler, and V. Moron, 2006: Translating climate forecasts into agricultural terms: Advances and challenges. *Climate Res.*, **33(1)**, 27-41.
- Hulme M., 1987: Secular changes in wet season structure in central Sudan. *J. Arid Environ.*, **13**, 31–46.

- Ines A. V. M., and J. W. Hansen, 2006: Bias correction of daily GCM rainfall for crop simulation studies. *Agric. For. Meteorol.*, **138**, 44-53
- Janicot S., F. Mounier, and A. Diedhiou, 2008: Les ondes atmosphériques d'échelle synoptique dans la mousson d'Afrique de l'Ouest et centrale : ondes d'est et ondes de Kelvin. *Secheresse*, **19(1)**, 13-22.
- Janicot S., S. Trzaska, and I. Pocard, 2001: Summer Sahel-ENSO teleconnection and decadal time scale SST variations. *Climate Dyn.*, **18**, 303-320.
- Janicot S., F. Mounier, N. M. J. Hall, S. Leroux, B. Sultan, and G. N. Kiladis, 2009: Dynamics of the West African Monsoon. Part IV: Analysis of 25-90-Day Variability of Convection and the Role of the Indian Monsoon. *J. Climate*, **22**, 1541-1565.
- Janicot S., V. Moron, and B. Fontaine, 1996: Sahel droughts and ENSO dynamics. *Geophys. Res. Lett.*, **23(5)**, 515-518.
- Jolliffe I. T., and D. E. Sarria-Dodd, 1994: Early detection of the start of the wet season in tropical climates. *Int. J. Climatol.*, **14**, 71-76.
- Kniveton D., et al., 2009: Trends in the start of the wet season over Africa. *Int. J. Climatol.*, **29(9)**, 1216-1225.
- Knippertz P., 2007: Tropical-extratropical interactions related to upper-level troughs at low latitudes. *Dyn. Atmos. Ocean*, **43**, 36-62.
- Lamb P. J., 1978: Large-scale Tropical Atlantic circulation patterns associated with Subsaharan weather anomalies. *Tellus*, **A30**, 240-251.
- Lélé M. I., and P. J. Lamb, 2010: Variability of InterTropical Front (ITF) and Rainfall over West African Soudano-Sahel. *J. Climate*, in press.
- Lu J., and T. L. Delworth, 2005: Oceanic Forcing of the Late 20th Century Sahel Drought. *Geophys. Res. Lett.*, **32**, L22706, doi:10.1029/2005GL023316.
- Matthews A. J., 2004: Intraseasonal variability over tropical Africa during northern summer. *J. Climate*, **17**, 2427-2440.
- Mounier F., S. Janicot, 2004: Evidence of two independent modes of convection at intraseasonal time scale in the West African summer monsoon. *Geophys. Res. Lett.*, **31**, L16116. doi:10.1029/2004GL020665.
- Ndiaye O., L. Goddard, M. N. Ward, 2009: Using regional wind fields to improve general circulation model forecasts of July-September Sahel rainfall. *Int. J. Climatol.*, **29**, 1262-1275.
- Ndiaye O., M. N. Ward, and W. M. Thiaw, 1999: Interannual rainfall variability in Senegal and teleconnexion with the Pacific Ocean in a monthly time scale. *Pub. Ass. Int. Clim.*, **12**, 298-305.
- Nicholls N., 1984: A system for predicting the onset of the North Australian wet-season. *J. Climatol.*, **4**, 425-435.
- Okumura Y., and S. P. Xie, 2004: Interaction of the Atlantic Equatorial Cold Tongue and the African Monsoon. *J. Climate*, **17**, 3589-3602.
- Omotosho J. B., 1987: Richardson number, vertical wind shear and storm occurrences at Kano, Nigeria. *Atmos. Res.*, **21**, 123-137.
- Omotosho J. B., 1992: Long-range prediction of the onset and end of the rainy season in the West African Sahel. *Int. J. Climatol.*, **12**, 369-382.
- Omotosho J. B., 1996: The validation of prediction models for onset of rainfall and prospects for West African Sahel. Report. A contribution to the ACMAD Demonstration period (**96**), ACMAD, Niamey, Niger.

- Peyrillé P., J. P. Lafore, and J. L. Redelsperger, 2007: An idealized two-dimensional framework to study the West African monsoon. Part I: Validation and key controlling factors. *J. Atmos. Sci.*, **64**(8), 2765–2782.
- Pohl B., S. Janicot, B. Fontaine, and R. Marteau, 2009: Implication of the Madden–Julian Oscillation in the 40-Day Variability of the West African Monsoon. *J. Climate*, **22**, 3769–3785.
- Reason C. J. C., and M. Rouault, 2006: Sea surface temperature variability in the tropical southeast Atlantic Ocean and West African rainfall. *Geophys. Res. Lett.*, **33**, L21705. 10.1029/2006GL027145.
- Roca R., Lafore J-P., Piriou C., and J-L. Redelsperger, 2005: Extratropical dry-air intrusions into the West African monsoon midtroposphere: An important factor for the convective activity over the Sahel. *J. Atmos. Sci.*, **62**, 390-407.
- Saha S. et al., 2006: The NCEP Climate Forecast. *J. Climate*, **19**, 3483–3517.
- Sivakumar M. V. K., 1988: Predicting rainy season potential from the onset of rains in Southern Sahelian and Sudanian climate zones of West Africa. *Agric. For. Meteorol.* **42**, 295-305.
- Sivakumar M.V.K. (ed.), 1998: Climate Variability Prediction, Water Resources and Agricultural Productivity: Food Security Issues in Tropical Sub-Saharan Africa. Report of a Joint START/WCRP/SCOWAR Workshop, 22-25 July 1997, Cotonou, Benin. Washington DC, USA: International START Secretariat.
- Stern R. D., M. D. Dennett, and D. J. Garbutt, 1981: The start of the rains in West Africa. *J. Climatol.*, **1**, 59-68.
- Sultan B., and S. Janicot, 2003: The West African Monsoon Dynamics. Part II: The “Preonset” and “Onset” of the Summer Monsoon. *J. Climate*, **16**, 3407-3427.
- Taylor C. M., 2000: The influence of antecedent rainfall on Sahelian surface evaporation. *Hydrol. Processes*, **14**(7), 1245-1259.
- Vizy E. K., and K. H. Cook, 2001: Mechanisms by which the Gulf of Guinea and eastern North Atlantic sea surface temperature anomalies can influence African rainfall. *J. Climate*, **14**, 795-821.
- Vizy E. K., and K. H. Cook, 2009: A mechanism for African monsoon breaks: Mediterranean cold air surges. *J. Geophys. Res.*, **114**, D01104, doi:10.1029/2008JD010654.
- Ward M. N., 1998: Diagnosis and short lead time prediction of summer rainfall in tropical North Africa at interannual and multidecadal timescale. *J. Climate*, **11**, 3167-3191.
- Ziervogel G, and R. Calder, 2003: Climate variability and rural livelihoods: assessing the impact of seasonal climate forecasts in Lesotho. *Area*, **35**(4), 403-417.

## Conclusion

The Sahel region is very challenging for modeling rainfall, especially using coarse resolution GCMs. The model must represent the impact on the atmosphere of the complex and spatially varying land and ocean surface, with the Sahara desert transitioning southward into Equatorial rain forest, and continuing southward, the presence of an Equatorial ocean, these changes all typically occurring over a small number of model grid boxes. They are expressed in tight North-South gradients of such atmospheric variables as temperature, wind, humidity and rainfall. Six of the eight AGCMs studied here show no ability of substance to represent the interannual variability of Sahel rainfall when driven with observed SST. However, the broader tropical Atlantic wind field is known to be closely tied to Sahel rainfall variations. It is hypothesized that the large-scale tropical Atlantic wind field is easier for a GCM to represent. Results presented here show consistency amongst eight GCMs on reproducing Sahel rainfall skill by using the low-level zonal wind over the tropical Atlantic / West Africa domain as a MOS predictor. This result builds confidence in the use of a GCM's regional circulation and encourages GCM Sahel rainfall correction for seasonal predictions in real-time. In most cases, the first EOF (EOF1) captures all the important wind variability over the tropical Atlantic / West Africa domain (30°N-40°S and 60°W-10°E) and using more EOFs adds no further skill, or only marginal skill, which makes the MOS correction very simple and less prone to statistical artifacts in estimated skill levels. A leading EOF (the EOF1 of most GCMs) is particularly consistent across the tropical Atlantic. Weights across the included area of West Africa are weaker and more variable, reflecting the

varying way in which models translate the large-scale signal into their own West African monsoon.

This study has focused on post-1967, the period for which most model forecasts are available, allowing coherent pictures to be drawn by studying models over comparable years. On the interannual timescale, this period permits a close assessment of skill associated with ENSO, but does not allow a good assessment of the ability to capture the tropical Atlantic mode of SST variability that was particularly important for interannual variability in Sahel rainfall during the wetter regimes of 1950-1970 and earlier in the century (Ward, 1998; Janicot et al., 2001). However, the MOS approach of using the tropical Atlantic winds does in principle appear capable of capturing the tropical Atlantic SST influence, as illustrated through analyses with the ECHAM4.5 GCM simulations that were available for the whole period 1950 to present. In this analysis, when the MOS correction are focused on GCM predictions for 1950-1970, or including the full period 1950-2002, a second model wind mode emerged, that specified rainfall variability specifically associated with tropical Atlantic SST. Further consideration of ensuring MOS approaches capture this aspect is encouraged. It is particularly important to address, given the results of Vizzy and Cook (2001), showing that coarse resolution GCMs can be expected to have particular difficulty with resolving the impact of this tropical Atlantic SST mode that gives rise to enhanced (reduced) rainfall in the Sahel with compensating reduced (enhanced) rainfall over the Guinea Coast region.

While good skill with the MOS approach was generally found for GCMs forced with observed SST, the problem of rapid SST development in boreal spring is known to be a

factor limiting operational seasonal forecasts. This was reproduced here using the ECHAM4.5 GCM driven with SST that is persisted from values in April ( $r=0.30$ ), May ( $r=0.33$ ) and June ( $r=0.55$ ). We have diagnosed the SST developments that occur in forecast failures and showed the largest changes (at least for forecast failures in the 1968-2002 period) are associated with tropical Pacific SST developments, with additional significant changes in the tropical Atlantic and Indian Oceans as well. For a failure to forecast observed dry (wet) conditions, warming (cooling) occurs between May and June in the central and eastern tropical Pacific, along with warming (cooling) in parts of the tropical Atlantic and Indian Oceans. This has a strong signature of ENSO development. We have diagnosed that the SST development from May to June has a strong and dynamically understandable response in the ECHAM4.5 GCM. Models that use empirical SST predictors are similarly limited in their lead-time by the SST developments during boreal spring.

Therefore, this study has explored coupled models, for their ability to increase the lead-time of Sahel rainfall predictions. Several coupled models were examined for skill. Most confirm the continuing challenge in predicting Sahel rainfall from April/May initial conditions, and indeed only a small number showed promise to match the level of skill achieved by persistence approaches to the forecast problem from April/May initial conditions. The CFS model revealed most promising results, and also has available a large set of hindcasts at varying lead-times, allowing detailed diagnosis of the prediction skill. The model, especially with the application of the MOS approach, is found to contain long-lead forecast information for JAS Sahel rainfall on the interannual

timescale, primarily related to ENSO, and also a trend component, that is found to accompany the lower-frequency fluctuation found in observed and model time-series of the North Atlantic versus South Atlantic SST difference.

The skill on the interannual timescale is dependent on applying the MOS approach, and the key information is captured effectively in the first EOF of the low-level tropical Atlantic winds. The predictive skill on the interannual timescale is largely not contained in the model's rainfall, which after detrending, has only very modest skill at short lead-times (0-2 months), and no skill at the longer lead-times (3-6 months). The information in the tropical Atlantic wind field is confirmed to be part of large-scale developments in the coupled model involving recognizable teleconnection structures from the tropical Pacific to the tropical Atlantic in the SST and low-level wind fields. Documenting these large-scale teleconnection structures in the model, reinforces confidence that the results illustrate the potential of coupled models to represent the evolution of the climate system through this period of the annual cycle, and specifically, to represent key aspects that are relevant for anticipating Sahel rainfall anomalies.

The period of study with the CFS forecasts (1981-2008) also contains a significant upward trend in Sahel rainfall, continuing the low-frequency component of the region's climate that has been well documented through the 20<sup>th</sup> Century. That low frequency component has been associated with an interhemispheric contrast in SST, largely expressed, during wet periods, with enhanced warmth in the North Atlantic, and contrasting relatively cool temperatures in the South Atlantic and Indian Ocean. For the

period of 1981-2008, aspects of North-South contrast appear to have operated again, with relative warmth in the North Atlantic accompanying the increasing rainfall through the period. The CFS forecasts do contain this tendency for a relatively warm North Atlantic compared to South Atlantic. There is also a tendency for changes in the forecast tropical Atlantic low-level wind field, and this does become translated into an upward trend in the MOS seasonal forecasts of Sahel rainfall, albeit with a trend magnitude that is between 50-75% weaker than observed. Overall, the results presented here are interpreted as evidence that the CFS forecasts contain information about interannual and multidecadal aspects of Sahel rainfall at lead-times up to 6 months.

An aspect that is not strongly present in the 1981-2008 period, but which has been shown to be a strong forcing agent on low-frequency Sahel rainfall fluctuations, is low-frequency change in the Indian Ocean SST (Folland et al., 1986; Giannini et al., 2003). For the purpose of diagnosis in this thesis, an index of Atlantic North-South contrast has been used. It was found that adding Indian Ocean SST into the index did not contribute to explanatory power in the period 1981-2008, although it is known that including the Indian Ocean is important for indices that explain the low-frequency rainfall fluctuations through the whole 20<sup>th</sup> Century. Indeed, in the presence of continued global warming, it has been noted that there is potential for trends in all ocean basins to play a significant role (e.g. Lu and Delworth, 2005; Biasutti and Giannini, 2006) in driving anomalies in regional circulation such as in the tropical Atlantic / West African monsoon. The robustness of GCMs in general, and a MOS system such as proposed here, requires careful monitoring in the presence of such trends.

The findings here are however considered to be of practical significance for the generation of predictive information about Sahel rainfall. At short lead-time, it is encouraging that most models can be expected to contain information about JAS Sahel rainfall. The consistency in Fig. III.6 (and average result in Table III.3) suggests most models are capturing a common signal, and any gains from combining a large-number of models can be expected to be relatively modest in this case. Preliminary analyses have confirmed this to be the case. Further analysis is needed to propose a solution to achieving a practical way of dealing with information from multiple models for this prediction problem, and to assess the best ways to combine information from GCMs and information from statistical approaches with SST predictors. It is generally held that the GCM and statistical methods are complementary and analysis of performance in combination would provide valuable information for operations.

For the longer-lead information, results here suggest at least one model (CFS) already contains information that could be usefully consulted at lead-times up to six months. This should motivate further experiments and diagnosis of other coupled models for this prediction problem. Information for Sahel rainfall on the interannual timescale at lead-times between 3 and 6 months has to this point not been generally demonstrated on the scale reported here, and so the findings here could represent new information available to the operational West Africa forecasting community.

This result opens new horizons for diagnosis of the Sahel rainfall in GCMs, using the wind as a surrogate for the rainfall. The MOS approach could be well-suited to assessing decadal timescale predictions and climate change predictions of Sahel rainfall by using the low-level regional wind simulated by models, and this is considered to be a promising avenue for further investigation.

One important area for further study remains the development of systems to operationally downscale this skill into information at smaller scales within the Sahel, and also to develop other predictands like water satisfaction indices or frequency of rainy events greater than a certain threshold. Such scales of information, and such variables, can be more suited to many user needs and can be expected in general, to be more effectively translated into action.

For arid regions that rely on rain-fed agriculture, the start date of the wet season is a fundamental piece of information that aids decision making at a number of levels including the individual farmers, local NGOs, farmer associations and cooperatives, government agencies, insurance companies, and inter-governmental organizations responsible for regional food security and supply. In the last chapter of this study, we focused on rainfall onset in Senegal. We started by defining the rainfall onset with a particular approach of defining it over homogeneous sub-regions. We are unaware of any previous study of onset taking such an approach. Previously onset was either defined at an individual station level (and taking account of local moisture conditions) or was defined as a single large-scale West Africa phenomenon (e.g. using regional OLR). We have also tried to take into consideration various aspects of the onset by using three

definitions: (i) the first rainfall event sufficient to trigger thinking in terms of rainy season farming practices, after the long dry season, (ii) a large scale onset which extends the first definition to cover a rainfall event observed simultaneously over at least three stations within the sub-region, and (iii) an onset index that registers onset only when there is no long dry spell after the rainfall event. One striking feature that we found is the large number of false onset events. It's quite difficult to find years where the rainfall sequences are continuous. Almost all first rainfall events are followed by dry spells. This makes this study especially relevant.

Using these various regional indices we established connection with the large scale boundary layer including SST and low level wind. We were able to identify two mechanisms which control the timing of the onset over northern and southern Senegal. The onset over the northern region of Senegal is related to large scale signals developing in June–July. It includes moderate correlation with tropical Pacific and Indian Ocean warming events. Such a pattern is favorable for late onset. It is consistent with the known relationship between July SST and July rainfall in the Sahel (El Niño related to dry conditions in the Sahel). In addition, there is a negative correlation with the adjacent tropical North Atlantic Ocean, which might involve water vapor advection through a sea-breeze type mechanism. This relationship appears primarily a simultaneous one, emerging around the time of onset, rather than providing long-lead information.

For the southern region the teleconnection is mostly confined over the tropical Atlantic basin with a strong North-South dipole. This teleconnection operates through

strengthening/weakening the monsoon flow by a thermal (SST) wind gradient mechanism. This mechanism has expression in early April before the climatological onset in the south (May-June). It is also consistent with skill that emerged for the rainfall total of May in Southern Senegal (related to tropical Atlantic SST), and provides a promising avenue for further investigation. The Southern Senegal onset relationship with the tropical Pacific is much weaker, and actually of opposite sign, as compared to the relationship for the Northern region. Such a weak opposite sign would be consistent with a previously proposed opposite relationship for June Sahel rainfall (weak positive correlation with ENSO at the time of onset of the southern region), as compared to July rainfall (stronger negative correlation with ENSO at the time of onset of the northern region).

We also look at the atmospheric mechanisms which take place around the onset, including analysis of atmospheric circulation, PWAT and MSE. The onset over Senegal is strongly controlled by the large scale atmospheric system, with propagating features prior to and after the onset. The PWAT, MSE and wind circulation all appear to exert a strong control on the onset. Circulation, MSE and PWAT anomalies are detected ahead of the onset, clearly present as far away as 20-30E (and possibly further East still), and propagating westward within the corridor of 10N-20N. The PWAT and MSE build up to present a clear high maximum at 20-30E, 8 to 15/20 days prior the onset date. They then travel westward into Senegal. This feature offers a good way of providing advance information, with onset likely when such clear build-up and propagation is seen in the system.

One important aspect to take into consideration is to separate and dissociate true with false onset. To study the contrast, we developed a further way of defining true / false onset. For this, we analyze, over all stations in the sub-region, the 10 days after the onset rainfall event. We calculate the average number of dry days in the 10-day period. Years with a high number of dry days following the onset are selected as severe false onset. We compare such years, with ones with a low number of dry days after the onset event (strong true start years). The composite analysis of the difference in anomaly shows clear 'danger' signals to watch for. The danger signals are in the wind, PWAT and potential temperature at 925hPa. False onset with respect to true onset can be characterized by anomalies of PWAT located at the western fringe of Senegal and over the adjacent tropical Atlantic. The PWAT appears to be the most salient feature contributing to a false onset, and can inhibit the progression of any propagating system after the onset. There is also strong North easterly wind located near and North of the region (Mauritania to western Europe) concomitant with low potential temperature favorable for an advection of cold air. This low level cold air surge contributes further to kill any convection. This second mechanism operates more after the false onset event. With such clear signals, we can provide warning of possible false onset by following these key identified parameters.

Part of chapter VI considered explicit prediction systems for the onset indices. Building upon the findings with teleconnections, we propose statistical models. For the southern region of Senegal the best model is established between tropical Atlantic SST and the first rainy events with a skill of  $r=0.46$  using April SST over 1968-2008. For the northern region the true onset is the most predictable using the tropical SST. The skill is  $r=0.43$

using July SST over 1981-2008, though lead-time is harder to achieve for the northern region, with skill declining to about  $r=0.28-0.36$  using June SST. However, overall these results are encouraging and the SST is a good candidate for forecasting the onset over both North and South Senegal, and the skills offered are within the range of those achieved for the longer-lead seasonal forecasting.

A further aspect investigated was anomalies in the pre-onset annual cycle of the atmosphere. Previously proposed vertical shear does not emerge as a strong predictor for Senegal rainfall onset (at least, not using the reanalysis data analyzed here). However, onset in the southern region did have some association with onset in the northern region, which suggests some predictive skill may be possible to establish from the annual cycle developments, at least about a month in advance for the northern region. At the shorter lead-times of several days, the propagating features that have been found offer the prospect of early information, including the separation of true and false starts.

In the last part of chapter VI we assess the level of skill of CFS model on predicting the onset. The CFS is able to distinguish the climatology of the Northern and Southern regions of Senegal. Forecasts analyzed were initialized in early May, representing about a two months lead on the onset date for the Northern region. At this lead-time, there was no skill found for the northern part of Senegal. However, for the southern region of Senegal, the skill is  $r=0.43$  for the large scale onset and 0.46 for the true onset event. These results are very encouraging and give hope for using coupled models to forecast onset of the rainfall. We also investigated the application of a statistical correction on the forecast

rainfall sequence, by mapping the intensity and distribution to observed station characteristics within the same homogeneous sub-region (Hansen et al., 2006; Baron et al., 2005). This improved the basic daily characteristics of the CFS rainfall series. In terms of skill of predicting onset, the correction did not change the skill of the true onset date forecast, but did raise the skill of the first large scale event from  $r=0.43$  to  $r=0.55$ . The correction approach is a promising avenue for further investigation.

For an operational setting, it is proposed to use all these approaches in a multi-model information system providing guidance about the onset. This reflects the fact that we are trying to forecast a single event (onset), and it may be controlled by signals from diverse sources. The PWAT, which seems to be a strong atmospheric indicator, is for example controlled by interacting phenomena such as moist convection, wave activity, dry intrusion, and monsoon flux. So one aspect captured by a given approach might work in a year when that mechanism is dominant, while in other years, other aspects might control the onset.

One aspect that did not emerge in our results is the Madden Julian Oscillation (MJO). Studies have proposed its role in intraseasonal Sahel rainfall variations, mostly focusing within the main rainy season (Mathews, 2004; Pohl et al., 2009; and Janicot et al., 2009). It is possible that filtering data may lead to better identification of MJO signals in onset. One feature found here that may connect to the findings in Janicot et al. (2009) is the interaction of the slow propagating feature for southern Senegal onset with the Indian Ocean sector, given their suggestion of a role for active/break phases in the Indian

monsoon. Another avenue might be to focus on special onset cases and use models which contain basic mechanism of the monsoon to study and better understand the interaction between onset and larger scale processes such as MJO and Indian monsoon. Another large scale feature that deserves further attention for its role in the onset is the Saharan heat low (Couvreur et al., 2009).

This study is a contribution to the understanding and the forecasting of the onset and seasonal rainfall total, over the Sahel. For the seasonal rainfall total, we have found an approach of correcting poor GCM performance over the Sahel. It has been tested thoroughly that this approach is reproducible across many GCMs. With the easy access of GCM output online, and the increase in local computing power and internet connection, this approach can be readily adopted in an operational framework in Africa. We have also contributed by exploring a promising way of increasing the lead time of seasonal forecasts over the Sahel (up to six month lead-time) by using coupled models, with the CFS runs at six-month lead time showing skill levels close to that previously only achieved at zero or one-month lead-time.

The chapter on the onset of the seasonal rainfall has demonstrated the relationship between the synoptic scale and the large scale boundary layer, which combine to influence the timing of the onset. The early indicators and explicit prediction skill found for rainfall onset over Senegal is very encouraging. This study over Senegal can be implemented in other neighboring countries.

These two major results are a non negligible response to the two climate-related questions that most concern the population every year in the Sahel region: how much will it be raining and when will it start? Skillful seasonal forecasting at lead times up to six months, and forecasting the timing of the onset, is demanded by the population. The information is expected to be of great value for planning over the Sahel, in a region where the resources are very rare and need to be used with care.