# Statistical and dynamical assessment of land-ocean-atmosphere interactions across North Africa

By

Yan Yu

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Atmospheric and Oceanic Sciences)

at the

UNIVERSITY of WISCONSIN-MADISON

2017

Date of final oral examination: 4/24/2017

The dissertation is approved by the following members of the Final Oral Committee:

Zhengyu Liu, Professor, Atmospheric and Oceanic Sciences

Michael Notaro, Senior Scientist, Nelson Institute Center for Climatic Research

Ankur Desai, Professor, Atmospheric and Oceanic Sciences

Daniel Vimont, Professor, Atmospheric and Oceanic Sciences

Tristan L'Ecuyer, Associate Professor, Atmospheric and Oceanic Sciences

# Table of Contents

**Chapter 1 Introduction ..... 1**

**1.1. Study region ..... 1**

**1.2. North African climate variability and its drivers ..... 3**

1.2.1. Oceanic drivers ..... 6

1.2.2. Terrestrial feedbacks ..... 8

**1.3. Limitations in past observational studies and GEFA development ..... 11**

**1.4. Overview of the current study ..... 14**

**Chapter 2 Data and method ..... 18**

**2.1. GEFA methodology ..... 19**

2.1.1. Traditional GEFA (or full GEFA, FGEFA) ..... 19

2.1.2. Stepwise GEFA (SGEFA) ..... 20

2.1.3. GEFA-based estimate of percent variance explained by ocean and land forcings ..... 23

**2.2. GEFA validation within CESM ..... 23**

2.2.1. Fully coupled CESM control run (CTRL) ..... 24

2.2.2. Validation of GEFA in capturing oceanic impacts on regional climate ..... 25

2.2.2. Validation of GEFA in capturing terrestrial impacts on regional climate ..... 31

**2.3. Observational GEFA framework ..... 41**

**2.4. Observational datasets and multi-data observational benchmark ..... 44**

2.4.1. Multi-dataset bootstrapping method ..... 44

2.4.2. Station dust observations and MERRA-2 dust reanalysis ..... 45

**Chapter 3 GEFA validation ..... 51**

**3.1. GEFA's capability in capturing oceanic impacts on regional climate ..... 51**

3.1.1. Stepwise GEFA versus full GEFA ..... 51

3.1.2. GEFA's capability at capturing key oceanic impacts on North African climate ..... 53

**3.2. GEFA's capability in capturing terrestrial impacts on regional climate ..... 68**

3.2.1. Stepwise GEFA versus full GEFA in capturing the seasonal cycle of local responses to LAI anomalies ..... 69

3.2.2. Terrestrial impacts over Sahel and WAM in CESM ..... 75

**3.3. Discussion: Guidance for future application of GEFA ..... 81**

**Chapter 4 Observed oceanic drivers of North African climate ..... 84**

**4.1. Important observed oceanic drivers of North African climate ..... 84**

**4.2. Observed impacts of tropical Pacific SSTs on North African climate .....86**
**4.3. Observed impacts of tropical Atlantic SSTs on North African climate .....88**
**4.4. Observed impacts of tropical Indian SSTs on North African climate.....90**
**Chapter 5 Observed terrestrial feedbacks across North Africa..... 94**
**5.1. Relative contribution of oceanic drivers and terrestrial drivers to North African climate.94**
**5.2. Observed terrestrial impacts on North African climate .....97**
5.2.1. Observed moisture feedbacks .....98
5.2.2. Observed albedo feedbacks.....102
5.2.3. Observed momentum feedbacks .....105
5.2.4. Observed remote responses to HOA vegetation anomalies .....106
5.2.5. Observed dust responses to the Sahel vegetation anomalies .....107
**5.3. Evaluating terrestrial impacts on North African climate simulated by CESM .....109**
**Chapter 6 Summary and future work ..... 117**
**Reference ..... 122**

## Abstract

North Africa is highly vulnerable to hydrologic variability and extremes, including impacts of climate change. The current understanding of oceanic versus terrestrial drivers of North African droughts and pluvials is largely model-based, with vast disagreement among models in terms of the simulated oceanic impacts and vegetation feedbacks. Regarding oceanic impacts, the relative importance of the tropical Pacific, tropical Indian, and tropical Atlantic Oceans in regulating the North African rainfall variability, as well as the underlying mechanism, remains debated among different modeling studies. Classic theory of land-atmosphere interactions across the Sahel ecotone, largely based on climate modeling experiments, has promoted positive vegetation-rainfall feedbacks associated with a dominant surface albedo mechanism. However, neither the proposed positive vegetation-rainfall feedback with its underlying albedo mechanism,

nor its relative importance compared with oceanic drivers, has been convincingly demonstrated up to now using observational data. Here, the multivariate Generalized Equilibrium Feedback Assessment (GEFA) is applied in order to identify the observed oceanic and terrestrial drivers of North African climate and quantify their impacts.

The reliability of the statistical GEFA method is first evaluated against dynamical experiments within the Community Earth System Model (CESM). In order to reduce the sampling error caused by short data records, the traditional GEFA approach is refined through stepwise GEFA, in which unimportant forcings are dropped through stepwise selection. In order to evaluate GEFA's reliability in capturing oceanic impacts, the atmospheric response to a sea-surface temperature (SST) forcing across the tropical Pacific, tropical Indian, and tropical Atlantic Ocean is estimated independently through ensembles of dynamical experiments and compared with GEFA-based assessments. Furthermore, GEFA's performance in capturing terrestrial impacts is evaluated through ensembles of fully coupled CESM dynamical experiments, with modified leaf area index (LAI) and soil moisture across the Sahel or West African Monsoon (WAM) region. The atmospheric responses to oceanic and terrestrial forcings are generally consistent between the dynamical experiments and statistical GEFA, confirming GEFA's capability of isolating the individual impacts of oceanic and terrestrial forcings on North African climate. Furthermore, with the incorporation of stepwise selection, GEFA can now provide reliable estimates of the oceanic and terrestrial impacts on the North African climate with the typical length of observational datasets, thereby enhancing the method's applicability.

After the successful validation of GEFA, the key observed oceanic and terrestrial drivers of North African climate are identified through the application of GEFA to gridded observations, remote sensing products, and reanalyses. According to GEFA, oceanic drivers dominate over

terrestrial drivers in terms of their observed impacts on North African climate in most seasons. Terrestrial impacts are comparable to, or more important than, oceanic impacts on rainfall during the post-monsoon across the Sahel and WAM region, and after the short rain across the Horn of Africa (HOA). The key ocean basins that regulate North African rainfall are typically located in the tropics. While the observed impacts of SST variability across the tropical Pacific and tropical Atlantic Oceans on the Sahel rainfall are largely consistent with previous model-based findings, minimal impacts from tropical Indian Ocean variability on Sahel rainfall are identified in observations, in contrast to previous modeling studies. The current observational analysis verifies model-hypothesized positive vegetation-rainfall feedback across the Sahel and HOA, which is confined to the post-monsoon and post-short rains season, respectively. However, the observed positive vegetation feedback to rainfall in the semi-arid Sahel and HOA is largely due to moisture recycling, rather than the classic albedo mechanism.

Future projections of Sahel rainfall remain highly uncertain in terms of both sign and magnitude within phases three and five of the Coupled Model Intercomparison Project (CMIP3 and CMIP5). The GEFA-based observational analyses will provide a benchmark for evaluating climate models, which will facilitate effective process-based model weighting for more reliable projections of regional climate, as well as model development.

# Chapter 1 Introduction

## 1.1. Study region

North Africa is characterized by pronounced ecological and moisture gradients. The land surface varies remarkably from the tropical Congo rainforest, to steppe vegetation across the Horn of Africa (HOA), to savanna and croplands in the Sahel, to the Saharan Desert, with large spatial variability in the remotely sensed Normalized Difference Vegetation Index (NDVI) (Figure 1). The Sahel represents the southern margin of the Saharan Desert and northern extent of the region affected by the boreal summer African monsoon (Giannini et al. 2008a). Typically, the West African monsoon (WAM) onset occurs in late June in conjunction with an abrupt northward push of the Intertropical Convergence Zone (ITCZ) (Sultan and Janicot 2000; Le Barbé et al. 2002), bringing about 70% and 50% of the annual rainfall to the Sahel and WAM region, respectively, during July to September (Figure 2). The seasonal cycle of HOA rainfall is characterized by two wet seasons following the latitudinal movement of the ITCZ, namely the long rains during March-May and short rains during October-December. The semi-arid Sahel and HOA exhibit large rainfall variability on the intra-seasonal, interannual [e.g. linked to El Niño-Southern Oscillation (ENSO) and Indian Ocean sea-surface temperatures (SSTs)], and decadal time scales (Nicholson 1978, 1980; Brooks 2004; Janicot et al. 2011).

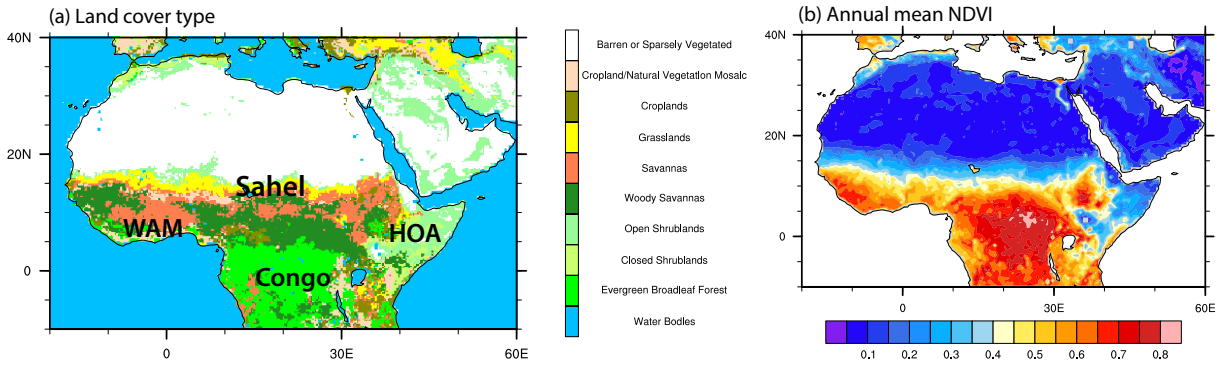


**Figure 1** (a) Remotely-sensed land cover type from the International Satellite Land Surface Climatology Project (ISLSCP) initiative II International Geosphere-Biosphere Project (IGBP) DISCover and SiB Land Cover dataset (1992-1993). (b) Annual mean Normalized Difference Vegetation Index (NDVI) from Advanced Very High Resolution Radiometer (AVHRR) Global Inventory Modeling and Mapping Studies (GIMMS) NDVI3g dataset (1981-2011).

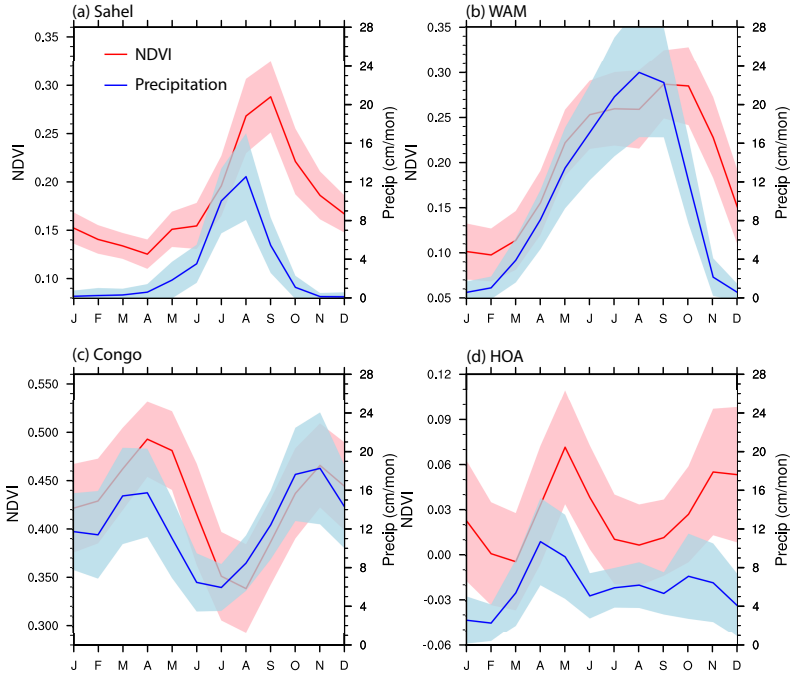


**Figure 2** Seasonal cycle in NDVI (red) and precipitation (cm month$^{-1}$, blue) across (a) Sahel, (b) WAM region, (c) Congo, and (d) HOA. Precipitation data is from Global Precipitation Climatology Centre

(GPCC), and NDVI data is from the AVHRR GIMMS NDVI3g dataset. Shading represents ± one interannual standard deviation.

### 1.2. North African climate variability and its drivers

Rainfall variability in the Sahel is characterized by both interannual and decadal variability, with an increasing contribution of low-frequency variability since the mid-20$^{th}$ century (Farmer and Wigley 1985; Nicholson and Entekhabi 1986; Hulme 2001; Brooks 2004). The Sahel experienced one of the most pronounced climatic shifts worldwide in the observational record (Figure 3), as abundant rains during the 1950s-mid-1960s transitioned to extreme drought during the late 1960s-1990s (Nicholson 1979; Lamb 1982; Katz and Glantz 1986; Lamb and Peppler 1992; Hulme 1996; Giannini et al. 2008 a,b; Flato et al. 2013). Lack of planned irrigation in agriculture (You et al. 2011) and limited medical care resources (Nayar 2012) make the Sahel one of the most vulnerable regions to hydrologic extremes. The prolonged drought and its resulting famine from the late 1960s to early 1980s led to 100,000 fatalities, left 750,000 people dependent on food aid, and affected most of the Sahel's 50 million people (United Nations Environmental Programme 2002). The HOA is currently experiencing a severe drought, with 12 million people across Ethiopia, Kenya, and Somalia in need of food assistance, together with rising debt, low cereal and seed stocks, and low milk and meat production (FAO 2016).

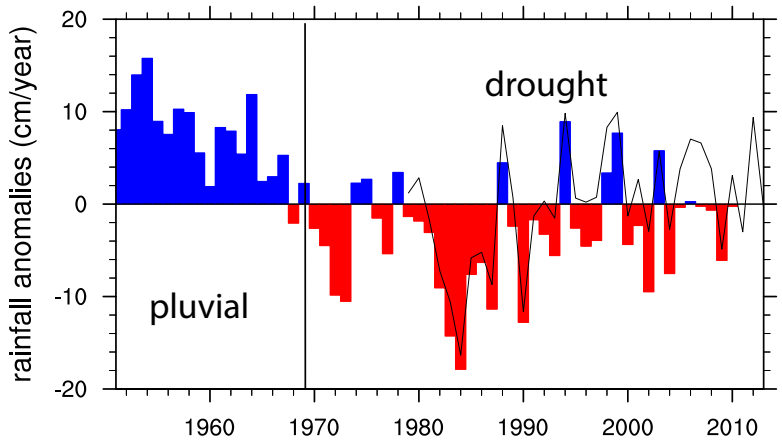


**Figure 3** Time series of annual Sahel rainfall anomalies (cm year$^{-1}$) during 1951-2013 [bars from GPCC and line from Global Precipitation Climatology Project (GPCP)]. In addition to experiencing large interannual variability in rainfall, the Sahel rapidly transitioned from pluvial conditions during the 1950s-mid-1960s to drought during the late 1960s-1990s.

Since the 1970s, a hierarchy of modeling studies have attributed the persistent drought across the Sahel to either oceanic drivers, namely regional or global SST anomalies (Folland et al. 1991; Rowell et al. 1995; Giannini et al. 2003), or terrestrial drivers, especially land cover and land use changes and associated vegetation feedbacks (Charney 1975; Charney et al. 1977; Shukla and Mintz 1982; Xue 1997; Zheng and Eltahir 1997; Clark et al. 2001; Taylor et al. 2002; Wang et al. 2004). Pioneer observational studies have linked Sahel rainfall anomalies to SST anomalies in the tropical Atlantic (Lamb 1978; Lamb and Peppler 1992), tropical Pacific (Palmer 1986; Ward 1992; Janicot et al. 1996), Indian Ocean (Palmer 1986; Janicot et al. 1996), and the Mediterranean Sea (Rowell 2003; Rodriguez-Fonseca et al. 2011), as well as inter-hemispheric SST contrast (Folland et al. 1986; Folland et al. 1991), based on composite, correlation, and regression analyses. Meanwhile, modeling studies have approximately reproduced the multi-decadal variability in Sahel rainfall, although producing weaker and shorter droughts than

observed, by forcing the global climate models (GCMs) with observed SST anomalies alone (Folland et al. 1986; Palmer 1986; Rowell et al. 1995), suggesting that oceanic forcings are the primary driver of Sahel rainfall anomalies. On the other hand, among the earliest modeling efforts, Charney (1975) introduced the concept of a positive vegetation feedback on rainfall regarding North African desertification, suggesting that a reduction in greenness leads to an increase in surface albedo, resulting in surface cooling, increased atmospheric stability, sinking motion, and drying. By including soil moisture and vegetation feedbacks, recent modeling studies were better able to capture the magnitude and duration of the observed Sahel drought, leading to the conclusion that Sahel rainfall variability is amplified by land-atmosphere interactions (Zeng et al. 1999; Giannini et al. 2003; Wang et al. 2004; Held et al. 2005; Scaife et al. 2009; Kucharski et al. 2012). Brooks (2004) and Lu and Delworth (2005) outlined the arguments for the cause of the late 20$^{th}$ century drought in the Sahel, including exogenous factors, which consist of external forcing from SST anomalies (Folland et al. 1986; Palmer 1986; Rowell et al. 1995), and endogenous factors, which consist of land degradation (e.g. overgrazing) and desertification and associated local feedbacks (Charney 1975). However, the relative contribution of oceanic and terrestrial impacts to the interannual and decadal variability in Sahel rainfall has never been studied in observations. Furthermore, the apparent recovery from the multi-decadal drought during the early 21$^{st}$ century and the underlying recovery mechanism remain highly debated, partly due to an observed increase in the interannual variability in Sahel rainfall (Nicholson 2013). In particular, the recent interannual variability in Sahel rainfall was not successfully predicted by the SST forcings that explained the late 20$^{th}$ century decadal drought in most state-of-the-art climate models, implying either recent changes in oceanic drivers of the interannual Sahel rainfall variability or elevated importance of other regulators, including land

surface feedbacks (Biasutti et al. 2008). Despite this model-based hypothesis regarding the change in oceanic regulations, there has never been observational quantification of the relative contribution from oceanic versus terrestrial drivers of the recent interannual variability in Sahel rainfall. Given uncertainty and inherent biases in GCMs, it is necessary to advance methodologies that can isolate the relative contributions of SST anomalies and regional LAI anomalies toward driving Sahel rainfall in observations.

### 1.2.1. Oceanic drivers

While generally agreeing that tropical SST anomalies are critical drivers of Sahel rainfall, modeling studies have continually debated the relative contribution of the nearby tropical Atlantic and Indian Oceans and distant Pacific Ocean (Giannini et al. 2003; Wang et al. 2004; Lu and Delworth 2005) and associated dynamic mechanisms. GCMs generally indicate that positive SST anomalies in the eastern equatorial Atlantic Ocean, including the Gulf of Guinea, favor an anomalously dry Sahel by reducing the land-ocean temperature contrast and pushing the ITCZ and associated deep convection southward (Janicot et al. 1998; Giannini et al. 2003). The tropical Atlantic SST anomaly dipole has been shown to affect Sahel rainfall (Giannini et al. 2008 a,b), with a warmer tropical South Atlantic compared to tropical North Atlantic leading to Sahel drought by three proposed mechanisms: (1) diminished cyclonic flow and moisture flux into the Sahel (Hagos and Cook 2008), (2) southward shift of the ITCZ towards warmer waters (Folland et al. 1986; Hoerling et al. 2006), and (3) inability to sustain deep convection over the land as the ocean temperature rises, which warms up the nearby tropical troposphere, thereby leading to a more stable atmosphere (Chou et al. 2001; Chiang and Sobel 2002; Neelin et al. 2003; Giannini et al. 2005; 2008a,b). Through the generation of divergence and an anomalous anticyclonic circulation, positive SST anomalies in the tropical Indian Ocean are believed to

favor dry conditions in the Sahel (Bader and Latif 2003, 2005; Giannini et al. 2003; Hagos and Cook 2008). GCMs generally indicate that an anomalously warm tropical eastern Pacific Ocean supports Sahel drought (Giannini et al. 2003, 2005) by generating an amplified east-west divergent circulation across the tropical Atlantic Ocean with anomalous subsidence over West Africa (Janicot et al. 1998).

A number of critical uncertainties remains among modeling studies of oceanic drivers of North African rainfall. One of the major concerns is regarding the inconsistent results produced by different models. For instance, Lu and Delworth (2005) identified the Indian and Pacific Oceans as the main regulators of Sahel rainfall anomalies, with minimal contribution from the Atlantic Ocean, while Hoerling et al. (2006) simulated a strong influence from the Atlantic Ocean, with minimal contribution from the Indian Ocean. The attributed mechanisms behind these relationships likewise remain inconsistent and uncertain among modeling studies. The relative contribution of different oceanic forcings also appears to change over time (Janicot et al. 1996; Janicot et al. 1998; Rodriguez-Fonseca et al. 2011). For example, the observed impacts of SST anomalies across the Gulf of Guinea on the north-south displacement of the WAM rain belt appear to have largely diminished since the late 1970s (Rodriguez-Fonseca et al. 2011). Furthermore, these oceanic drivers can interact nonlinearly; specifically, Janicot et al. (1998) simulated an influence on the WAM rainfall from eastern equatorial Pacific and Atlantic Ocean SST anomalies, with the influence from SST anomalies in one basin modulated by those from the other basin. Due to the clear divergence in model-based findings, Lu and Delworth (2005) encouraged further investigation of the relative contribution of different tropical ocean basins in driving Sahel rainfall variability. Past observational studies did not specifically deal address covariability among SST forcings, including the tropical Pacific and tropical Indian Oceans

(Yang et al. 2007); therefore, the conclusions about their relative contributions are somewhat uncertain. A deeper understanding of oceanic drivers of North African rainfall is needed through more sophisticated observational analyses, so that regional climate predictability can be enhanced and climate models can be properly evaluated and improved.

Beyond the Sahel, rainfall anomalies across the HOA, especially in October-December (short rains), are believed to be regulated by ENSO and the Indian Ocean Dipole (IOD). El Niño conditions are believed to support positive rainfall anomalies (Indeje et al. 2000; Mason and Goddard 2001; Hastenrath 2007). The IOD, with anomalously warm SST in the western and anomalously cold SSTs in the eastern tropical Indian Ocean, appears to induce anomalous southeasterly moist trade winds, thereby favoring above-average rainfall (Behera et al. 2005; Saji et al. 1999). However, the relative importance of tropical Pacific and tropical Indian SST anomalies to HOA rainfall anomalies is still under debate (Liebmann et al. 2014).

### 1.2.2. Terrestrial feedbacks

Vegetation and climate interact through a serious of complex feedbacks. Vegetation affects the climate directly through biophysical feedbacks, consisting of moisture, energy, and momentum exchanges with the atmosphere, and indirectly through biogeochemical processes that alter atmospheric $CO_2$ levels (Pielke et al. 1998; Bonan 2002) (Figure 4). Through the moisture feedback, an increase in evapotranspiration potentially leads to greater atmospheric precipitable water and precipitation, further enhancing plant growth. According to the albedo (energy) feedback, changes in vegetation cover alter the surface albedo and radiative fluxes, leading to a local temperature change and eventually a vegetation growth response. The albedo (energy) feedback is particularly important when forests mask snow cover or grass spreads into

desert (Robinson and Kukla 1985; Bonan et al. 1992; Betts and Ball 1997; Bonan 2002). According to Charney (1975), a reduction in greenness leads to an increase in surface albedo, resulting in low-level cooling, increased atmospheric stability, sinking motion, and drying. According to previous modeling studies, a reduction in Sahel vegetation cover, as a result of an ocean-induced prolonged drought, may amplify and extend the duration of droughts by increasing the surface albedo and decreasing evapotranspiration (Charney 1975; Charney et al. 1977; Zeng et al. 1999; Wang et al. 2004). According to the momentum feedback, variations in the surface roughness of vegetation alter the wind speed, moisture convergence, turbulence, and depth of the atmospheric boundary layer, which then affect vegetation growth (Sud and Smith 1985; Sud et al. 1998; Buermann 2002).

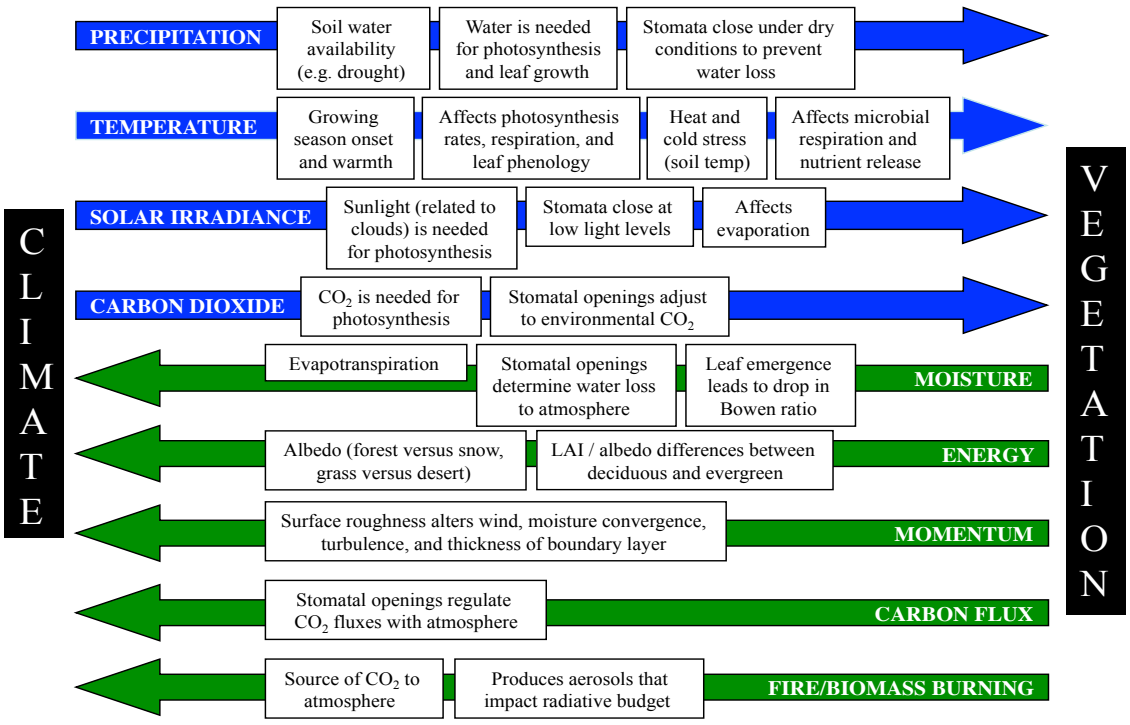


**Figure 4** Schematic of vegetation-climate interactions (adapted from Notaro et al. 2006).

The current understanding of biophysical vegetation feedbacks has largely come from running and analyzing coupled vegetation-climate model simulations, which have several key limitations. Simulated feedbacks are model dependent, given that climate models largely differ in terms of their dynamical cores, numerical schemes, parameterizations, spatial resolution, and resulting simulation lengths. For example, using the fully coupled National Center for Atmospheric Research (NCAR) Community Climate System Model Version 3.5 (CCSM3.5) with dynamic vegetation, Notaro et al (2011) concluded that reduced vegetation cover leads to an earlier subtropical Chinese monsoon and delayed, weaker tropical Australian monsoon. In contrast, in the Regional Climate Model Version Four (RegCM4), reduced LAI leads to diminished rainfall during Australia's pre-mid monsoon season but not for China (Notaro et al. 2017). The inconsistent findings regarding China's monsoonal response were attributed to CCSM's excessive forest cover and LAI, exaggerated roughness mechanism, and deficient ET response (Notaro et al. 2017). In addition to the model biases, nearly all modeling studies have applied extreme sensitivity experiments, such as a complete replacement of a specific vegetation type with bare ground or another vegetation type, either locally or globally. Such extreme experiments are unrealistic, since vegetation changes are typically heterogeneous and occur over time. Furthermore, the vast majority of modeling studies has focused on the long-term equilibrium response of climate to an imposed vegetation change, rather than the climatic response to intra-seasonal to interannual variations in vegetation abundance, despite its importance to short-term climate prediction (Wang et al. 2014). Owing to these limitations in modeling studies, observational studies of land surface feedbacks are critically needed for confirming the model-based findings (O'Brien 1996), which requires the development and validation of a powerful statistical tool for extracting observed vegetation feedbacks.

### 1.3. Limitations in past observational studies and GEFA development

Spurred by the aforementioned limitations in previous modeling studies of oceanic and terrestrial feedbacks, an observational study is needed to identify the oceanic and land surface drivers of North African climate. Such an observational study would have to address several key challenges (Wang et al. 2013, 2014). First, most observational records are short in duration and contain measurement errors. Second, it is challenging to extract the observed signal of vegetation forcing on atmosphere, given the existence of large atmospheric internal noise and the fact that the atmospheric forcing on vegetation outweighs vegetation's feedback to the atmosphere (Notaro et al. 2006; Liu et al. 2006). Third, given that the regional climate is affected by variability in both the ocean and land, potential covariability between the ocean and land makes it difficult to clearly separate their individual impacts on the atmosphere. Liu et al. (2006) used a statistical method, equilibrium feedback assessment (EFA), to quantify vegetation influence on the atmosphere over the globe. They confirmed the positive vegetation-rainfall feedbacks in a few isolated regions, including part of Sahel and Horn of Africa. However, the identified feedbacks are trivial and insignificant, largely due to EFA's inability to exclude the impacts associated with oceanic forcings. Other recent observational studies of North African vegetation feedbacks used a statistical vegetation index simulation, in which the NDVI is expressed as a linear function of antecedent and current local precipitation and temperature (Los et al. 2006), and Granger causality analysis, in which NDVI and precipitation are expressed as a linear function of each other at different time lags (Lee et al. 2015; He and Lee 2016). These observational studies found evidence of the proposed positive vegetation-rainfall feedback across the Sahel. However, they do not tease out potential oceanic impacts on the variability in both vegetation and precipitation. The ignored oceanic impacts are either due to the absence of

oceanic predictors in their analysis or limitations of the applied methodology itself, namely that multiple linear regression-based methods are not able to deal with highly correlated predictors. Therefore, the assessed terrestrial impacts based on the multiple regression-based methods are biased. Furthermore, their analyses are generally based on a single observational or reanalysis product for each variable, e.g. precipitation from the Climatic Research Unit (CRU), which has limited gauge coverage and thus limited reliability in the Sahel when compared with recent remotely sensed products (Table 5). Moreover, none of the previous observational studies have quantified the influence of observational uncertainty on the identified vegetation-rainfall feedbacks. Therefore, both the observed positive vegetation-rainfall feedback and the underlying mechanisms require more sophisticated and comprehensive investigations.

A multivariate statistical method, the Generalized Equilibrium Feedback Assessment (GEFA), was developed by Liu et al. (2008) to address the aforementioned challenges in the observational analysis of the oceanic and terrestrial forcings on the atmosphere (Wen et al. 2010). The primary purpose of this statistical method is to extract the forcing of a slowly-evolving environmental variable (e.g. SST, LAI) on the rapidly-evolving atmosphere, either in climate model output or observational data. GEFA was developed to assess non-local climate feedbacks as a multivariate generalization of the univariate EFA (Frankignoul et al. 1998), based on the stochastic climate theory of Hasselmann (1976) and Frankignoul and Hasselmann (1977). The GEFA methodology addresses both local and non-local feedbacks simultaneously, which is critical given that vegetation and SST anomalies can remotely affect atmospheric conditions (Chen et al. 2012; Wen et al. 2013; Wang et al. 2013, 2014). GEFA is capable of separating the individual impacts of different ocean basins and vegetated regions on climate in select regions. Atmospheric responses to surface forcings estimated by GEFA have been validated with two

independent statistical methods, linear inverse modeling (LIM) and fluctuation-dissipation theorem (Liu et al. 2012 a,b). Within the fully coupled CCSM3.5, Wang et al. (2013, 2014) performed dynamical ensemble experiments, with imposed anomalies of tropical Pacific or tropical Atlantic SSTs, as well as regional LAI across North America, and compared the atmospheric response to that predicted by the statistical GEFA approach, as applied to the CCSM3.5 control run. Generally consistent feedback estimates between the independent statistical and dynamical approaches, in the same climate model, demonstrated GEFA's ability to extract the impacts of oceanic and land surface forcings on North America's regional climate. After validating GEFA in the model, the method was confidently applied to observational and remote sensing data to assess controls on North America's observed climate and develop an observational oceanic and vegetation feedbacks benchmark against which climate models may be assessed (Wang et al. 2013, 2014). Beyond Wang et al. (2013, 2014), GEFA has been applied to examine the impacts of North Pacific SST variability in NCAR's Community Climate System Model Version 3 (CCSM3) (Zhong and Liu 2008), and of global SST variability on observed patterns of geopotential height (Wen et al. 2010) and United States' precipitation (Zhong et al. 2011).

GEFA is characterized by several key strengths (Wang et al. 2013, 2014) but needs further exploration in terms of its applicability and potential improvement. First, it can estimate the impact of individual forcings within a unified framework, particularly suitable for separating oceanic and terrestrial impacts on the regional climate. Second, the method is easily applied to model output or observations, without need for computationally expensive simulations. However, although GEFA's capability at isolating the contribution of terrestrial forcings from oceanic forcings on North America's climate has been successfully validated (Wang et al. 2013, 2014),

GEFA's applicability to other ecoregions, e.g. semi-arid regions, tropical and subtropical regions, has never been demonstrated by previous studies. In particular, the Sahel and WAM region exhibit pronounced decadal variability in their regional climate. It remains uncertain if the decadal variability affects the accuracy of GEFA. In addition, semi-arid regions like the Sahel and HOA exhibit strong vegetation-soil moisture-atmosphere coupling (Koster et al. 2004). GEFA is likely to capture the combined impact of coupled fluctuations in vegetation and soil moisture in those regions, which needs to be demonstrated with dynamical experiments. Furthermore, the sampling error grows quickly with increasing number of forcings to be assessed by the traditional GEFA method, thereby requiring relatively long data records to achieve reliable feedback estimates (Wang et al. 2013, 2014). For example, with only roughly 30 years of remotely sensed vegetation indices, the traditional GEFA exhibits limited credibility in assessing the observed vegetation feedbacks (Wang et al. 2014). Moreover, with more oceanic forcings to be assessed, since North African climate is believed to be affected by oceanic forcings in both hemispheres and the Mediterranean Sea, the sampling error likely grows even faster for North Africa than North America. The potential for improving the GEFA method, especially reducing the required data length, needs to be explored if observed land-atmosphere interactions are to be assessed. Particularly for the data-sparse North Africa, the application of GEFA also need to incorporate an effectively way to consider multiple observational datasets and their differing uncertainties.

### 1.4. Overview of the current study

The current study will address the following primary questions:

1) *How accurate is the multivariate statistical tool, GEFA, in separating feedbacks induced by variability across individual oceanic basins and terrestrial ecoregions of North Africa?*

   Hypotheses: Based on studies by Liu et al. (2012 a,b) and Wang et al. (2013, 2014), which demonstrated GEFA's reliability using conceptual and fully coupled climate models, it appears that GEFA is a potentially useful tool for isolating specific oceanic and terrestrial drivers of North African regional climate. Despite the initial demonstration of GEFA's reliability, further evaluation of the method is needed, particularly for the tropical/subtropical semi-arid regions in North Africa, which exhibits pronounced decadal rainfall variability and strong vegetation-soil moisture-atmosphere coupling. In addition, GEFA's accuracy in the assessed observed vegetation feedbacks will be largely constrained by the length of available remote sensing data, which is only about three decades long. Stepwise-selection is expected to improve GEFA's reliability especially when applied to short data, but the improvement needs to be assessed and quantified. Furthermore, in semi-arid regions like Sahel where vegetation and soil moisture are tightly coupled, GEFA feedback estimates will likely capture the combined impact of coupled variability in vegetation and soil moisture, which needs to be demonstrated by comparing the statistically-assessed response to both dynamical experiments in which regional LAI and coupled LAI-soil moisture are modified (Notaro et al. 2008).

2) *What are the primary oceanic drivers of the observed variability in North African climate and under what mechanisms?*

   Hypotheses: Based on previous modeling studies, the observed North African rainfall is expected to be primarily regulated by tropical oceans. The importance of tropical oceanic drivers for the Sahel and WAM region has been suggested by several modeling studies

(Janicot et al. 1998; Giannini et al. 2003, 2005), but the relative importance of each oceanic basin have never been examined in observations. Significant impacts of tropical Pacific and tropical Indian Ocean SSTs on HOA rainfall have been suggested by previous modeling studies, but the relative contribution of the two basins is still under debate (Liebmann et al. 2014).

3) *How do vegetation feedbacks affect the observed North African climate, and how important are they compared to oceanic drivers?*

   Hypotheses: Vegetation feedback is expected to act as a secondary regulator with regional hotspots over the semi-arid Sahel and HOA, where intra-annual to interannual variability in LAI is substantial (Koster et al. 2006; Liu et al. 2006). The dominance of oceanic drivers over terrestrial drivers for the Sahel and WAM regions has been suggested in previous modeling studies (Zeng et al. 1999; Giannini et al. 2003; Wang et al. 2004; Scaife et al. 2009), although never verified in observations. While the classic theory about vegetation-rainfall feedback promotes a dominant albedo feedback mechanism (Charney 1975), recent modeling studies highlight both the albedo and moisture feedback mechanism (Zeng et al. 1999; Wang et al. 2004). Therefore, the vegetation feedback mechanism need to be convincingly demonstrated with observational data.

In the current study, a combined observational and modeling assessment of land-ocean-atmosphere interactions across North Africa is conducted. The reliability of the multivariate statistical method GEFA over North Africa is first assessed using the NCAR CESM by comparing the GEFA-based statistical feedback assessment against an ensemble-based dynamical feedback assessment. In particular, the traditional GEFA approach is refined through stepwise GEFA, in which the size of the GEFA forcing matrix is reduced by dropping

unimportant drivers, so that the accuracy of the feedback estimates is expected to be improved. The potential improvements of stepwise GEFA over traditional GEFA are quantified by comparing to the dynamical assessment. GEFA is then applied to observational data to identify the primary oceanic and terrestrial drivers of North African regional climate. Multiple observational datasets are considered in the observational GEFA analysis to quantify feedback uncertainty in the data-sparse North Africa. The GEFA-based observational oceanic and terrestrial influence will serve as a benchmark for assessing the reliability of the climate models, e.g. CMIP5 models, across North Africa towards the generation of process-based weights of future projections from different models and reduce the overall uncertainty in the climate projections.

## Chapter 2 Data and method

The overall flowchart of the current work is summarized in Figure 5.

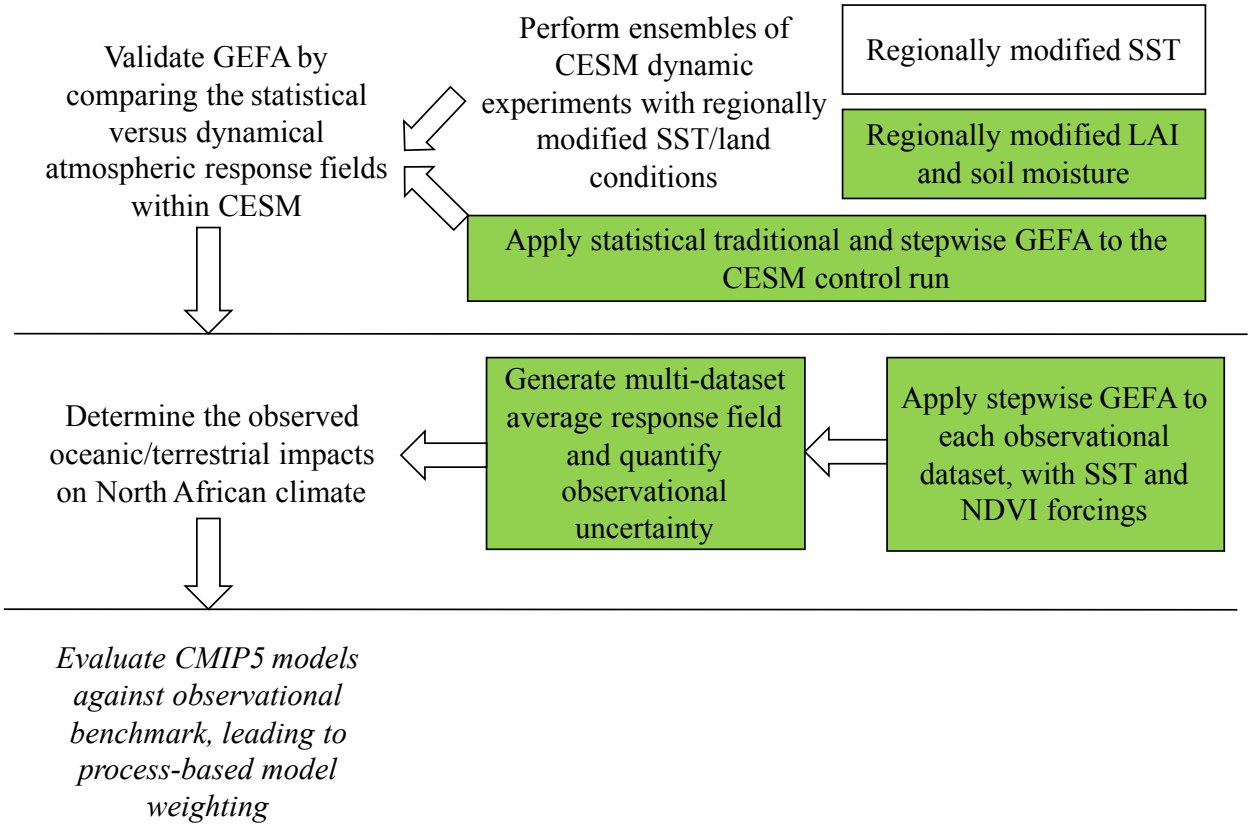


**Figure 5** General workflow. Text without a box states the objectives, with the italicized text representing future work. Each box represents a specific task, with green boxes indicating tasks largely completed by myself.

## 2.1. GEFA methodology

### 2.1.1. Traditional GEFA (or full GEFA, FGEFA)

The statistical GEFA approach extracts the forcing of a slowly-evolving environmental variable, such as SST or LAI, on the rapidly-evolving atmosphere, either in climate model output or observational data. GEFA was developed to assess both local and non-local climate feedbacks. More details on the GEFA methodology are provided by Liu et al. (2008).

At time scales longer than the atmospheric memory (about one week), the atmospheric variable (e.g. precipitation) at time t, A(t), as a response to an array of slowly-evolving variables (e.g. SST, LAI), O(t), can be approximated as (Liu et al. 2008):

$$A(t) = B \cdot O(t) + N(t) \qquad (1)$$

where B is the feedback matrix, and N(t) is the atmospheric internal noise. Right multiplying $O^T(t-\tau)$ on both sides of equation (1) and applying the covariance yield:

$$C_{AO}(\tau) = B \cdot C_{OO}(\tau) + C_{NO}(\tau) \qquad (2)$$

where $\tau$ is the time scale, exceeding the atmospheric adjustment time, and C is a covariance matrix. Given the time series' length L of the atmospheric and oceanic variables, the lagged covariance matrices are estimated as:

$$C_{AO}(\tau) = \frac{1}{L}A(t)O^T(t-\tau), C_{OO}(\tau) = \frac{1}{L}O(t)O^T(t-\tau), C_{NO}(\tau) = \frac{1}{L}N(t)O^T(t-\tau) \qquad (3)$$

The superscript "T" indicates a transpose. Since oceanic variability or land surface variability cannot be forced by atmospheric internal variability at a later time and the

atmospheric internal noise is not driven by oceanic or land surface forcings by definition in equation (1), $C_{NO}(\tau) = 0$. As a result, the feedback matrix can be estimated as:

$$B = C_{AO}(\tau) \cdot C_{OO}^{-1}(\tau) \qquad (4)$$

Before applying GEFA, the seasonal cycle and linear trend are removed from all forcing and response fields. The statistical significance of GEFA feedback matrices is assessed using the Monte Carlo bootstrap method with 1000 random iterations in which the time series of the response variable is scrambled (Wang et al. 2013, 2014).

The estimated feedback matrix represents the instantaneous influence of slowly-evolving variables (e.g. SST, LAI) on an atmospheric variable. Theoretically, B does not change with $\tau$, but due to the sampling error (insufficient L), the magnitude of B and the sampling error always increase with greater $\tau$. Here, $\tau$ is assigned to be one month. Indeed, larger $\tau$ leads to deteriorating estimates of the magnitude of oceanic and terrestrial feedbacks compared to the dynamical experiments, especially when considering short data records (not shown).

### 2.1.2. Stepwise GEFA (SGEFA)

In order to minimize the sampling error associated with relatively short datasets, it is necessary to reduce the list of forcings before estimating the feedback matrix. Here, negligible forcings are eliminated using a backward-selection stepwise method (Hocking, 1976), which builds a statistical prediction model by selecting the most important forcings as predictors of the atmospheric variable through an automated procedure. The stepwise selection has been widely applied to predictor selection in developing linear models as prediction tools for the climate or ecosystems (Yin et al. 2014; Segele et al. 2015; Yu et al. 2015). Akaike information criterion (AIC) (Akaike, 1974), which measures the relative quality of a statistical model by estimating

the goodness of fit and penalizing the complexity of the model (number of predictors), is used here as the selection criterion in the stepwise process.

$$AIC = 2\times N_f - 2\times \ln(\hat{L}) \quad (5)$$

$$\hat{L} = -\frac{L}{2}\ln(\sum_{t=1}^{L}\left(\hat{A}(t) - A(t)\right)^2/L) + C_1 \quad (6)$$

$$\hat{A}(t) = B \cdot O(t) \quad (7)$$

In equation (5), $N_f$ represents the number of forcings in the forcing matrix, and $\hat{L}$ stands for the maximized likelihood function of the statistical model in (1), which represents the likelihood of the statistical model in equation (1) with parameters (B matrix) estimated from real data by equation (4). $\hat{L}$ in (5) is calculated by equations (6-7), based on linear theory (Wonnacott and Wonnacott et al. 1972). In equation (6), L stands for the length of data, and $C_1$ is a constant value independent of either the actual time series of forcing matrix and response variable or B estimates. In equation (7), $\hat{A}(t)$ refers to the predicted atmospheric state at time t, based on equation (1). If AIC decreases after removing a select forcing, then this forcing has no significant contribution to explaining the variability of the atmospheric variable and can be eliminated from the forcing matrix. The procedure of stepwise selection continues until the AIC does not decrease after removing any of the remaining forcings. After the stepwise selection, the feedback matrix is estimated by equation (4) using the reduced forcing matrix. In this way, the number of forcings to be assessed by GEFA is reduced, thereby allowing more reliable estimates of feedbacks from the remaining, significant forcings.

In SGEFA, there are several ways to deal with forcings that are of interest but eliminated by the stepwise selection. If one is interested in a particular forcing, it can be “forced” to remain in

the forcing matrix. For example, when assessing the terrestrial feedbacks over the Sahel, if Sahel LAI is not selected by SGEFA, the local response to variability in Sahel LAI can be estimated by manually adding Sahel LAI into the forcing matrix. This approach is used in the GEFA evaluation when comparing to the dynamical assessments. Although this approach does not yield significant responses to local LAI in the Sahel or WAM region in CESM (discussed later), the assessed insignificant responses are still meaningful when compared with the dynamical experiments in evaluating GEFA's performance in distinguishing the significant from insignificant forcings. An alternative way to deal with the forcing that is of interest but not selected by SGEFA is to assign the response as zero, which is adopted in the observational study.

In addition to the sampling error from the short data record, GEFA's reliability is also limited by additional factors. GEFA is a linear statistical method, so its feedback estimates are a first-order approximation, given that both vegetation and SSTs can potentially induce non-linear feedbacks (Hoerling et al. 1997, 2001; Zhou et al. 2003). Furthermore, in semi-arid regions, including the Sahel, where vegetation and soil moisture are tightly coupled, GEFA feedback estimates will likely capture the combined impact of coupled fluctuations in vegetation and soil moisture, while a dynamical experiment with an imposed vegetation reduction, without a complimentary imposed soil moisture reduction, will result in a weaker feedback response than GEFA under the assumption that soil moisture and vegetation feedbacks are of the same sign (Notaro et al. 2008). This apparent inconsistency between the two methodologies has been addressed by performing additional dynamical experiments, in which both vegetation amount and soil moisture are reduced in recognition of their tight coupling (Notaro et al. 2008).

### 2.1.3. GEFA-based estimate of percent variance explained by ocean and land forcings

According to GEFA, the percent variance in a select response variable, as explained by either an oceanic or terrestrial forcing, is calculated similarly to the Analysis of Variance (ANOVA) approach in multiple linear regression (Wonnacott and Wonnacott 1972). For example, the percentage of explained variance by oceanic forcings, $V_O$, is calculated by

$$V_O = C_{A_O A}/V_A \quad (8)$$

where $C_{A_O A}$ is the covariance between the observed atmospheric time series (A) and predicted atmospheric time series by oceanic forcings ($A_O$), and $V_A$ is the variance in the atmospheric time series. The predicted atmospheric time series is reconstructed by

$$A_O = B_O \cdot O \quad (9)$$

where $B_O$ is the GEFA feedback matrix when only oceanic forcings are included in the forcing matrix, and O is the forcing matrix containing the oceanic forcings. The percentage of explained variance by terrestrial forcings is calculated similarly.

## 2.2. GEFA validation within CESM

In the current study, the reliability of GEFA and its refinement, stepwise GEFA, in capturing the oceanic and terrestrial impacts on North Africa's regional climate is evaluated against three ensemble sets of dynamical experiments within CESM. In the statistical assessment, both traditional GEFA and stepwise GEFA are applied to a 300 year fully coupled control run (CTRL), yielding statistically assessed atmospheric responses to SST anomalies in key oceanic basins and to LAI anomalies in key North African ecoregions. In the dynamical assessment, one ensemble set of dynamical experiments is developed with modified SSTs in the key oceanic

basins that affect North African climate, namely the tropical Pacific, tropical Atlantic, and tropical Indian Oceans ($EXP_{SST}$), which are determined by applying GEFA to the CESM CTRL (shown in Chapter 3). Two additional ensemble sets of dynamical experiments are developed in the Sahel and WAM region, in which either regional LAI ($EXP_{LAI}$) or coupled soil moisture-LAI is modified ($EXP_{SOIL}$). The dynamically-assessed atmospheric responses to either the oceanic or terrestrial anomalies are treated as the "truth" for evaluating the statistical GEFA method. By comparing the statistically- and dynamically-assessed atmospheric responses to SST and LAI anomalies, GEFA's applicability in assessing the oceanic and terrestrial feedbacks across North Africa is evaluated and expected improvements of stepwise GEFA over traditional GEFA are quantified.

### 2.2.1. Fully coupled CESM control run (CTRL)

The statistical assessment of land surface feedback is performed by applying GEFA to the 300-year fully coupled control run (CTRL) generated with CESM version 1.2 (Hurrell et al. 2013). The active components applied in CTRL include the Community Atmosphere Model version 5.3 (CAM5.3) (Neale et al. 2010), Community Land Model version 4 with Carbon-Nitrogen Dynamic Global Vegetation Model (CLM4-CNDV) (Oleson et al. 2010; Lawrence et al. 2011), Parallel Ocean Program version 2 (POP2) (Smith et al. 2010), and Community Ice CodE version 4 (CICE4) (Hunke et al. 2008). The model resolution is 0.9 latitude x 1.25 longitude (gx1v6). CESM generally captures the mean climatology and interannual variability over the majority of global regions (Hurrell et al. 2013). Potential model biases in mean climatology, land cover type, and vegetation feedbacks do not affect the approach applied here for GEFA evaluation, since both the statistical and dynamical assessments in the current work are conducted in the same model; consistency between atmospheric response fields from the two

assessments, even in the presence of large model biases, is sufficient to demonstrate GEFA's reliability. Therefore, the current study does not include a rigorous evaluation of CESM's simulated North African climate and ecosystems against observations.

A 2050-year offline CLM4-CNDV spin-up simulation is first generated to allow terrestrial carbon and nitrogen pools to reach equilibrium, providing the restart files for the initial fully coupled run, which is spun up for another 100 years. At the end of the 2150 year spin-up, the absolute value of globally-averaged monthly mean net ecosystem exchange oscillates around $\pm 0.05$ Pg C year$^{-1}$, which is regarded as a sign for equilibrium (Hoffman et al. 2008), with no significant trend in any of the key Earth system variables (e.g. SST, LAI, soil moisture, and fractional cover of plant functional types). The run was extended for another 300 years to serve as the control simulation for GEFA evaluation.

### 2.2.2. Validation of GEFA in capturing oceanic impacts on regional climate

GEFA's reliability at capturing the oceanic impacts on the North African climate is evaluated within NCAR CESM (Wang et al. 2017). The length of data record needed to obtain stable atmospheric responses by traditional GEFA (FGEFA) and stepwise GEFA (SGEFA) is estimated using the CESM Large Ensemble Community Project (LENS) data (Kay et al. 2015), which provides a sufficiently long data record for estimating the data record length required to obtain stable GEFA responses to oceanic forcings. GEFA's reliability is further validated by comparing the statistically- and dynamically-assessed atmospheric responses to SST anomalies within the tropical Pacific, tropical Atlantic, and tropical Indian Oceans, which are identified as important oceanic basins for North African climate by GEFA in the CESM.

#### 2.2.1.1 Estimating the minimum required data length for stable GEFA responses

As mentioned earlier, the accuracy of the GEFA approach is affected by the length of available data record. The CESMLENS (Kay et al. 2015), which includes 40 ensemble members under historical radiative forcing from 1920-2005, is regarded as a sufficiently long dataset containing 3440 years (40 ensemble members x 86 years), thereby providing a test bed to estimate the data record length required for obtaining stable GEFA responses to oceanic forcings. Each model realization differs from another by only small round-off level variations in their atmospheric initial conditions. All simulations are performed with CESM version 1 (CESM1, Hurrell et al. 2013) with Community Atmosphere Model version 5 (CAM5, Neale et al. 2010) as its atmospheric component. The model resolution is 0.9 latitude x 1.25 longitude (gx1v6). GEFA is applied to the full 3440-year time series, which is constructed by connecting the monthly anomaly outputs of all 40 members. In order to avoid potential problems from discontinuity between ensemble members, the seasonal cycle and linear trend are removed from raw model output for each ensemble member to obtain the monthly anomalies. Note that the discontinuity in the SST anomalies between ensemble members does not affect the GEFA results, as long as the SST memory is similar in all ensemble members and significantly longer than that of the simulated atmosphere. The removal of trend ensures that the GEFA-based feedbacks are on the seasonal-to-interannual time scales. Indeed, if we remove higher order trends (e.g. third order polynomial trend), the GEFA-based atmospheric responses are almost identical to those represented in the current study.

#### 2.2.1.2 Statistical assessments of oceanic impacts

In the statistical assessment of oceanic impacts, the GEFA forcing matrix is populated by the leading two empirical orthogonal function (EOF) modes of SST anomalies from eight non-overlapping ocean basins and area-average Mediterranean SSTs from CTRL. The purpose of

performing GEFA in truncated SST EOF space is to reduce the sampling error from highly correlated forcing fields (Wen et al. 2013). Past modeling studies have suggested the potential impacts of SST variability across the tropical Pacific (TP) (Folland et al. 1986), North Pacific (NP) (Folland et al. 1986), tropical Atlantic (TA) (Giannini et al. 2003; Giannini et al. 2005; Hoerling et al. 2006), tropical Indian (Lu and Delworth 2005), North Atlantic (Rodriguez-Fonseca et al. 2015), South Pacific (Folland et al. 1986), South Indian (Folland et al. 1986), South Atlantic (Folland et al. 1986), and Mediterranean Sea (Rowell 2003) on the North African climate. The leading two EOF modes explain a large portion of the total variance in SST anomalies, typically varying from 30 to 60% by basin and month. Moreover, for most basins, the leading two EOF models have clear physical meanings, such as El Niño-Southern Oscillation (ENSO), Indian Ocean Basin Mode, and Atlantic Niño mode. Therefore, the leading two EOFs from these oceanic basins are considered, such that the full forcing matrix consists of:

O = [TP1 TP2 NP1 NP2 TI1 TI2 TA1 TA2 NA1 NA2 SP1 SP2 SA1 SA2 SI1 SI2 MED]

For example, TP1 and SI2 represent the first principal component (PC) of tropical Pacific SST and the second PC of South Indian SST, respectively. The spatial patterns for all 16 oceanic modes for January, April, July, and October are shown in Figures 6-7.

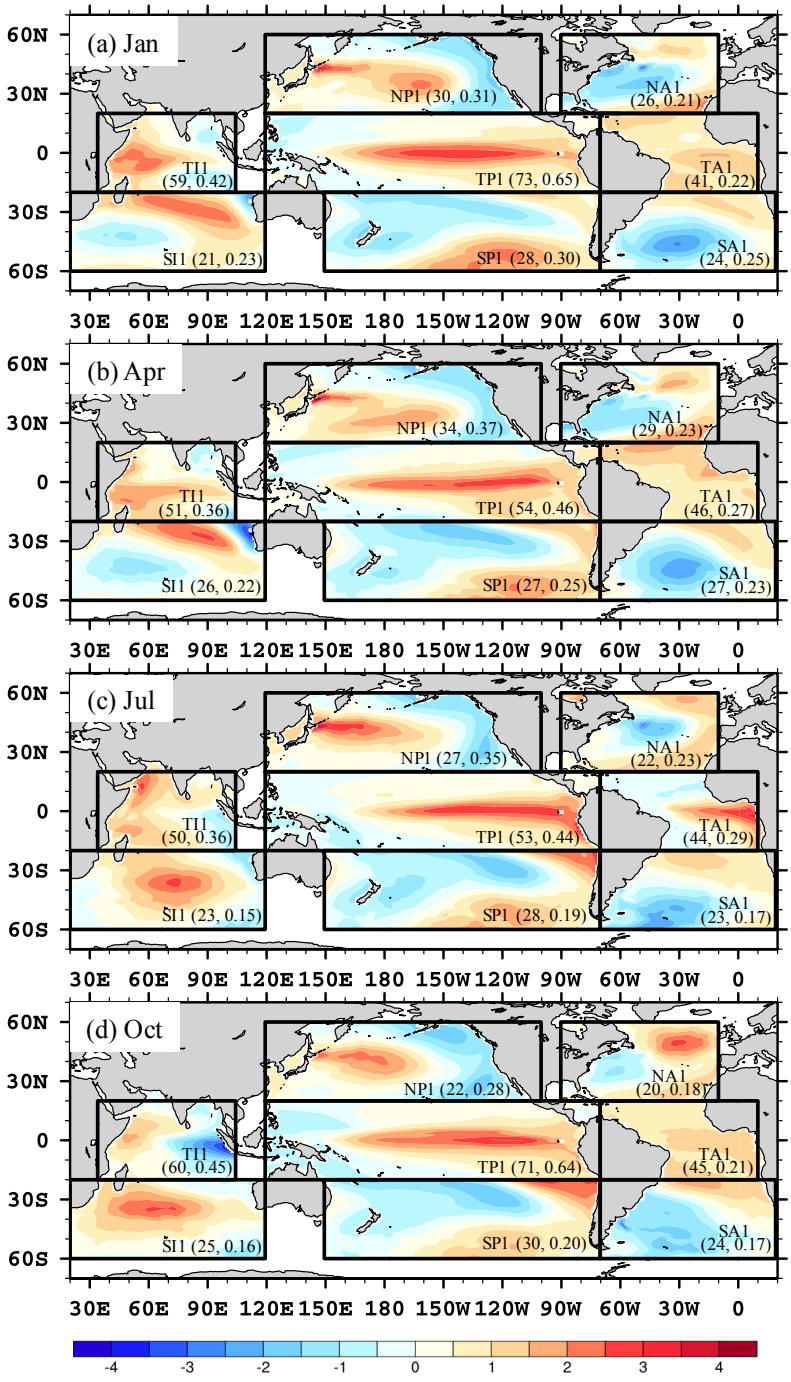


**Figure 6** Simulated spatial pattern of the first EOF mode (unitless) in the CESM CTRL in (a) January, (b) April, (c) July and (d) October of SST anomalies in eight ocean basins, namely the tropical Pacific (TP), tropical Indian (TI), tropical Atlantic (TA), North Pacific (NP), North Atlantic (NA), South Pacific (SP), South Atlantic (SA) and South Indian (SI). The explained variance (%) and standard deviation (℃) of

corresponding PCs of each ocean basin are indicated in parentheses. EOF analysis is performed separately over each boxed ocean basin, and the results are plotted together as a collage for convenience.

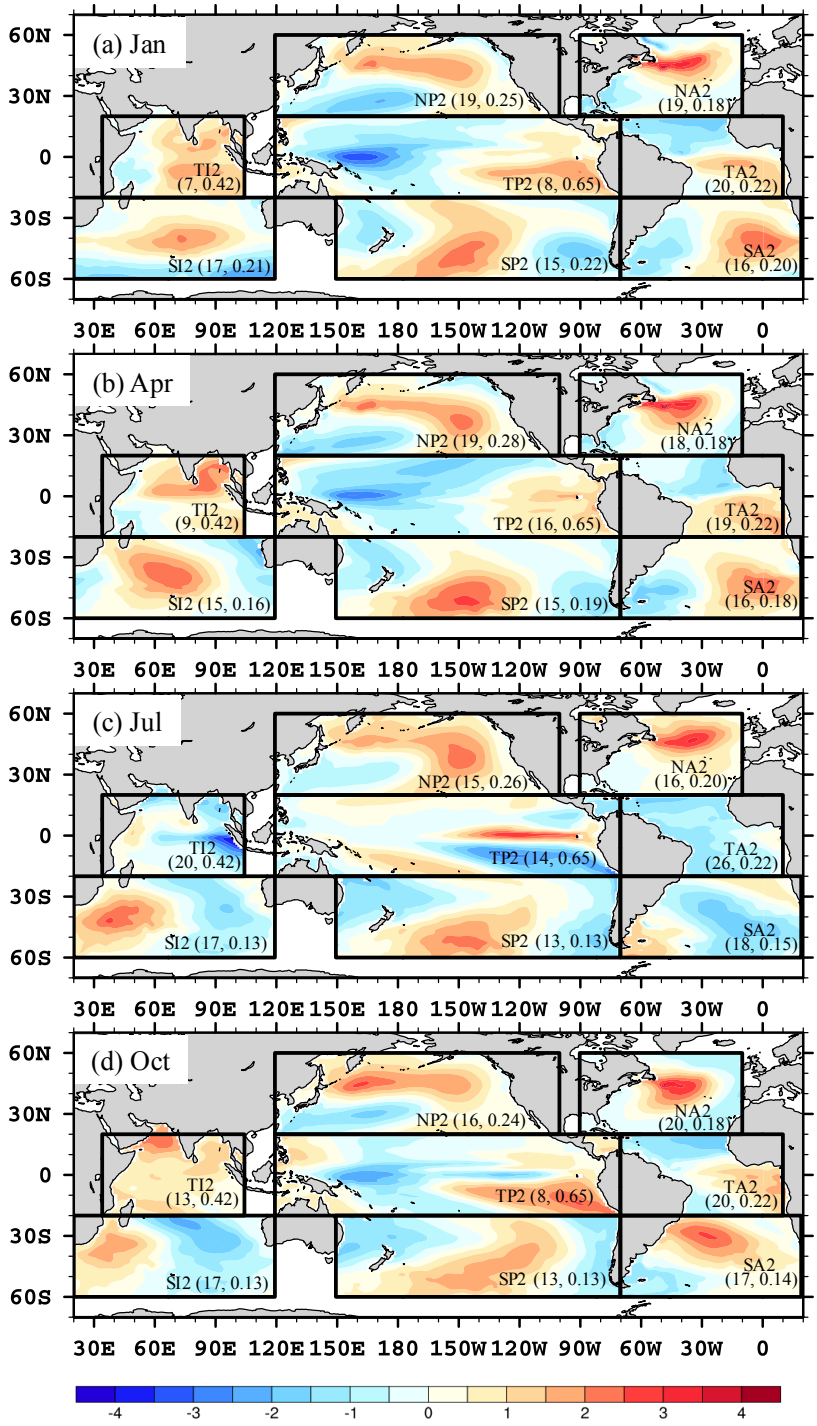


**Figure 7** Same as Figure 6, except for the second EOF modes.

#### 2.2.1.3 Dynamical assessment of oceanic impacts on regional climate

The "true" response of the atmosphere to a specified oceanic forcing is obtained dynamically through ensemble sensitivity experiments, which provide a benchmark for evaluating the statistically-assessed atmospheric responses by GEFA. Since TP1, TI1, and TA1 are identified as the most important oceanic forcings of North African climate in CESM (Chapter 3), the dynamical experiments focus on these three modes. Two-month-long, data-ocean dynamical experiments are conducted for each calendar month, with each member starting from its previous month's restart file from the CTRL. Each ensemble member is initialized from a different year in the CTRL. The prescribed sea ice field is the climatological sea ice fraction from the CTRL. For each oceanic forcing and calendar month, an ensemble set of 20 members with positive SST EOF anomaly pattern (P) and 20 members with negative SST EOF anomaly pattern (N) over the focal ocean basin is created. The SST anomalies are imposed onto the global climatological SSTs from CTRL. For TP1 and TI1, one standard deviation anomalies from the 300-year CTRL are applied, while for TA1, due to its weak interannual variability simulated by CESM, two standard deviation anomalies are applied. Both P and N are compared with the climatology from the CTRL, with (P-climatology) and (climatology-N) considered as the response of each ensemble member. The Student's t-test is used to determine the significance of the responses at the 90% level. After the first month of the simulation, the global atmospheric anomalies in 850-hPa and 200-hPa geopotential heights remain stable for all three oceanic forcings and in all calendar months. Therefore, the atmospheric responses are analyzed in the second month of the simulation. 40 ensemble members are demonstrated to be enough by examining the stability of the ensemble mean with increasing number of ensemble members. The

dynamical experiments are completely independent of the statistical method, thereby providing an independent check of the statistical method.

### 2.2.2. Validation of GEFA in capturing terrestrial impacts on regional climate

GEFA's reliability at isolating the terrestrial feedbacks from oceanic impacts on the North African climate is evaluated in CESM (Yu et al. 2017a). GEFA's reliability, including expected improvements of SGEFA over FGEFA, is evaluated by comparing the statistically- and dynamically-assessed response to LAI anomalies in the Sahel and WAM region, which represent contrasting landscapes and climates. Two ensembles of dynamical experiments, against which the GEFA-based vegetation feedbacks are evaluated, are developed for the Sahel or WAM region. In $EXP_{LAI}$, regional LAI is modified in one-month-long simulations, while in $EXP_{SOIL}$, regional LAI and soil moisture are modified together during winter-spring, motivated by the strong soil moisture-LAI coupling.

#### 2.2.2.1. Statistical assessment of terrestrial impacts on regional climate

In the statistical assessment of land surface feedbacks, the GEFA forcing matrix is comprised of the leading SST EOF modes from non-overlapping basins, area-average Mediterranean SSTs, and time series of area-average LAI across the Sahel, WAM region, HOA, and Congo. The purpose of including oceanic forcings when assessing terrestrial impacts is two-folded: first, to tease out potential contamination from ocean-land covariability on the estimated terrestrial feedbacks (Sun and Wang 2012); second, it allows for a comparison of the relative importance of oceanic versus terrestrial drivers. In terms of the oceanic forcings, the leading two EOF modes from the eight ocean basins are included in the forcing matrix. In addition, higher-order ($3^{rd}$ – $10^{th}$) SST EOFs that are correlated (with a temporal correlation of monthly anomalies

exceeding 0.2 for n = 300 years) with any of the four LAI anomaly time series in the fully coupled control run are also included in the GEFA forcing matrix. In term of the terrestrial forcings (Figure 8), the four ecoregions represent different North African landscapes in CESM, varying from mainly forest in the Congo and WAM region, to a combination of forests, shrublands, and grasslands in the Sahel and HOA. The geographic extent of the four ecoregions is determined through rotated EOF (REOF) analysis of monthly LAI anomalies in the CTRL. Since rotated EOF usually identifies regionally-coherent monopole patterns of variability, it is particularly suitable for determining ecoregions whose area-average vegetation growth are included in the GEFA forcing matrix.

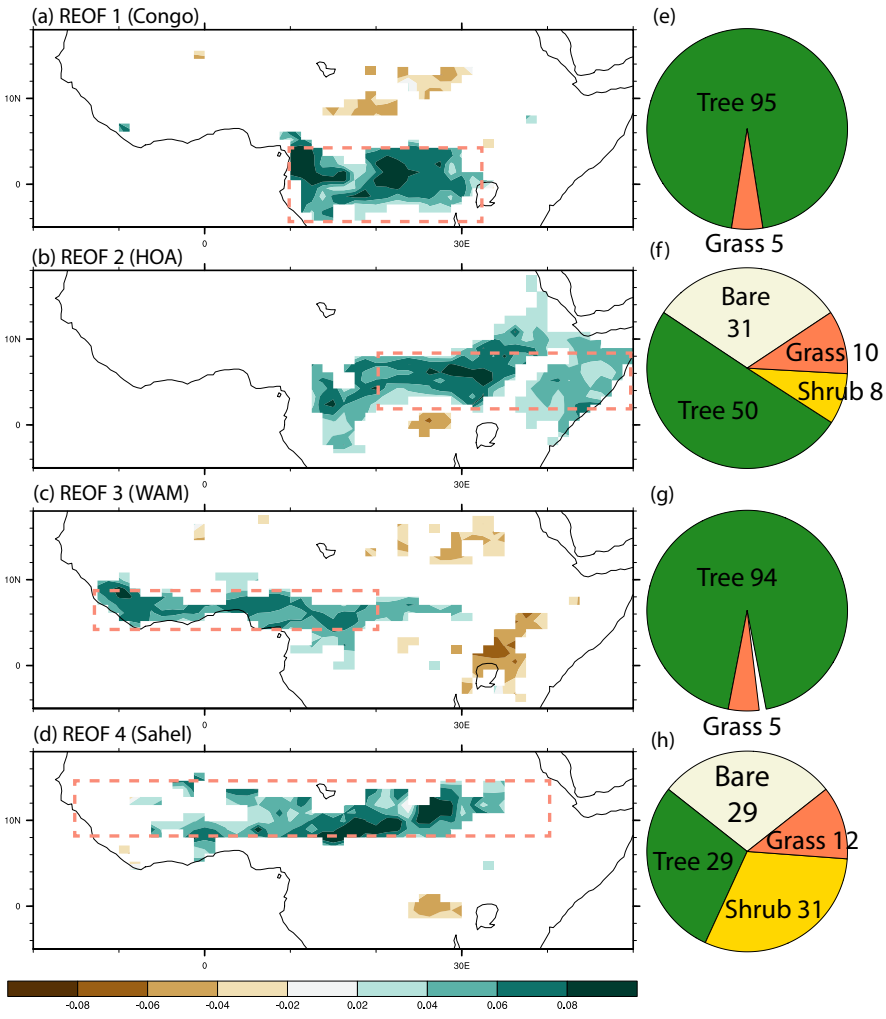


**Figure 8** Spatial pattern of the (a) first, (b) second, (c) third, and (d) fourth leading modes of variability in standardized monthly LAI anomalies from CTRL, according to REOF analysis. REOF 1-4 mainly represent LAI variability across the Congo (4˚S-4˚N, 10˚E-33˚E), broader HOA (2˚N-8˚N, 20˚E-50˚E), WAM region (4˚N-8˚N, 20˚W-20˚E), and Sahel (8˚N-15˚N, 15˚W-40˚E), respectively, and combined explain 22.8% of the total LAI variance. (e-h) The pie charts denote the mean vegetation distribution per region, as percent cover of tree, shrub, grass, and bare ground in CESM. The REOF analysis helps define regions (red boxes) for area-average LAIs to be included in the GEFA forcing matrix.

#### 2.2.2.2. Dynamical assessments of terrestrial impacts on regional climate

In order to evaluate GEFA's performance in capturing terrestrial impacts on the atmosphere, two ensemble sets of fully coupled CESM dynamical experiments are created, by either modifying LAI ($EXP_{LAI}$) or coupled soil moisture-LAI ($EXP_{SOIL}$) across the Sahel or the WAM region. The reasons for choosing both the Sahel and WAM ecoregions are two-folded. First, these two regions represent different climate and vegetation types and likely have unique land surface feedbacks (Liu et al. 2010; Wang et al. 2013). The Sahel is a semi-arid region covered by shrublands, forests, grasslands, and bare ground, while the WAM region receives greater annual rainfall and is mainly covered by forests in CESM (Figure 8). Second, these two ecoregions are geographically close to each other with highly correlated area-average LAIs during most months. The high correlation between the two LAI forcings potentially brings challenge for GEFA to accurately separate their individual impacts. The temporal correlation of monthly area-averaged LAI anomalies in the Sahel and WAM region in CTRL varies from 0.28 in December to 0.66 in March and has an annually averaged correlation of 0.52. Although Wang et al. (2013) demonstrated GEFA's capability of separating significantly correlated oceanic forcings, e.g. ENSO and Indian Ocean Basin Mode in winter-spring, GEFA's reliability in separating individual impacts from geographically nearby and statistically correlated vegetation fields has never been evaluated before.

##### 2.2.2.2.1 $EXP_{LAI}$

An ensemble set of one-month fully coupled dynamical experiments with modified regional LAI ($EXP_{LAI}$) is created for GEFA validation. In $EXP_{LAI}$, in each month, LAI at each grid cell within one of the focal sub-regions is either increased and fixed to the long-term monthly 95$^{th}$

percentile from CTRL in the positive (LAIINC) experiments or decreased and fixed to the long-term monthly 5$^{th}$ percentile from CTRL in the negative (LAIDEC) experiments (Figure 9). Ensembles with 30 positive and 30 negative members are created for the Sahel and the WAM region in $EXP_{LAI}$ in each month. The ensemble-mean area-averaged LAI anomalies vary from 0.47 $m^2$ $m^{-2}$ in August to 1.37 $m^2$ $m^{-2}$ in December across the Sahel and from 3.02 $m^2$ $m^{-2}$ in May to 4.07 $m^2$ $m^{-2}$ in January across the WAM region in LAIINC. The magnitudes of LAI anomalies in LAIDEC are generally slightly smaller than in LAIINC for both regions, since LAI cannot be lower than zero.

The response to the LAI anomalies across a given sub-region is represented by the regression coefficient of atmospheric anomalies upon LAI anomalies across the 60 ensemble members (30 LAIINC and 30 LAIDEC), and later compared with the response statistically estimated by GEFA. The statistical significance of the responses is evaluated by the Student's t-test. The 90% confidence interval (5$^{th}$ – 95$^{th}$ percentiles), which provides the uncertainty in the response among ensemble members, is determined through statistical bootstrapping, in which a probability distribution function of the response is generated by randomly sampling 60 out of the 60 ensemble members with replacement for 1000 times (Efron and Efron 1982).

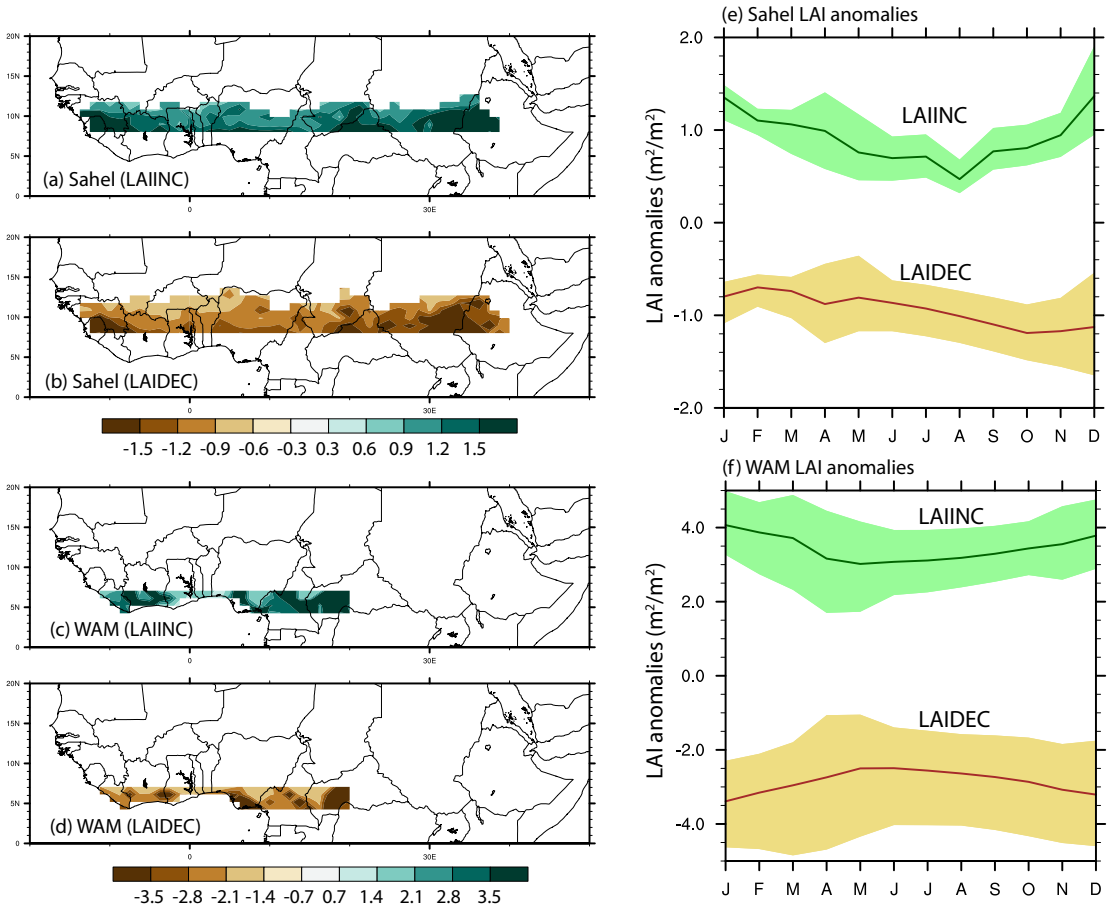


**Figure 9** LAI anomalies ($m^2$ $m^{-2}$) in the (a,c) LAIINC and (b,d) LAIDEC dynamical experiments with modified LAI ($EXP_{LAI}$) across the (a,b) Sahel and (c,d) WAM region with the (e, f) seasonal cycle of area-average LAI anomalies. Green and brown lines in (e-f) represent area-average LAI anomalies averaged across LAIINC and LAIDEC members, respectively, and the shading denotes the minimum and maximum anomalies among all members.

#### 2.2.2.2.2 $EXP_{SOIL}$

In semi-arid regions (e.g. Sahel) where vegetation and soil moisture are tightly coupled, GEFA feedback estimates capture the combined impact of coupled fluctuations in vegetation abundance and soil moisture in CTRL, while $EXP_{LAI}$ is expected to result in a weaker feedback response than GEFA under the positive soil moisture-rainfall feedback assumption (Koster et al. 2004), since the LAI increase (decrease) in $EXP_{LAI}$ is not accompanied by an increase (decrease) in soil moisture (Notaro and Liu 2008; Liu et al. 2010). In order to address the apparent inconsistency between the GEFA-based and $EXP_{LAI}$-based feedback assessments, additional dynamical experiments, in which both vegetation amount and soil moisture are modified together, are performed in the current study.

Motivated by the strong soil moisture-LAI coupling across North Africa in CESM during the winter-spring dry season, another ensemble set of dynamical experiments, in which the coupled soil moisture and LAI are modified together ($EXP_{SOIL}$), is developed for select months for each focal region. In December, January, March, and April for the Sahel, and February and March for the WAM region, the temporal correlation in the CTRL between simulated monthly LAI and antecedent monthly soil moisture within the top 1 m largely exceeds 0.7 (Figure 10). During the dry season, water supply is likely the main factor that limits vegetation growth across the study regions, thereby leading to a strong control of the antecedent soil moisture anomalies on LAI anomalies. In other months, there is no statistically significant correlation in CESM between LAI and either the antecedent or current top 1 m column average soil moisture across the focal regions. In $EXP_{SOIL}$, the initial soil moisture in the focal region is modified, thereby allowing the soil moisture and LAI anomalies to naturally evolve together in two-month fully coupled runs. In contrast to $EXP_{LAI}$ in which the LAI anomalies are fixed during the entire month of simulation, in $EXP_{SOIL}$ the initial soil moisture is modified in $EXP_{SOIL}$ and the

anomalies in soil moisture and LAI naturally evolve, which takes about one month to reach equilibrium. Therefore, one-month simulations are performed in $EXP_{LAI}$, but two-month simulations are performed in $EXP_{SOIL}$, in which the responses in the second month are analyzed.

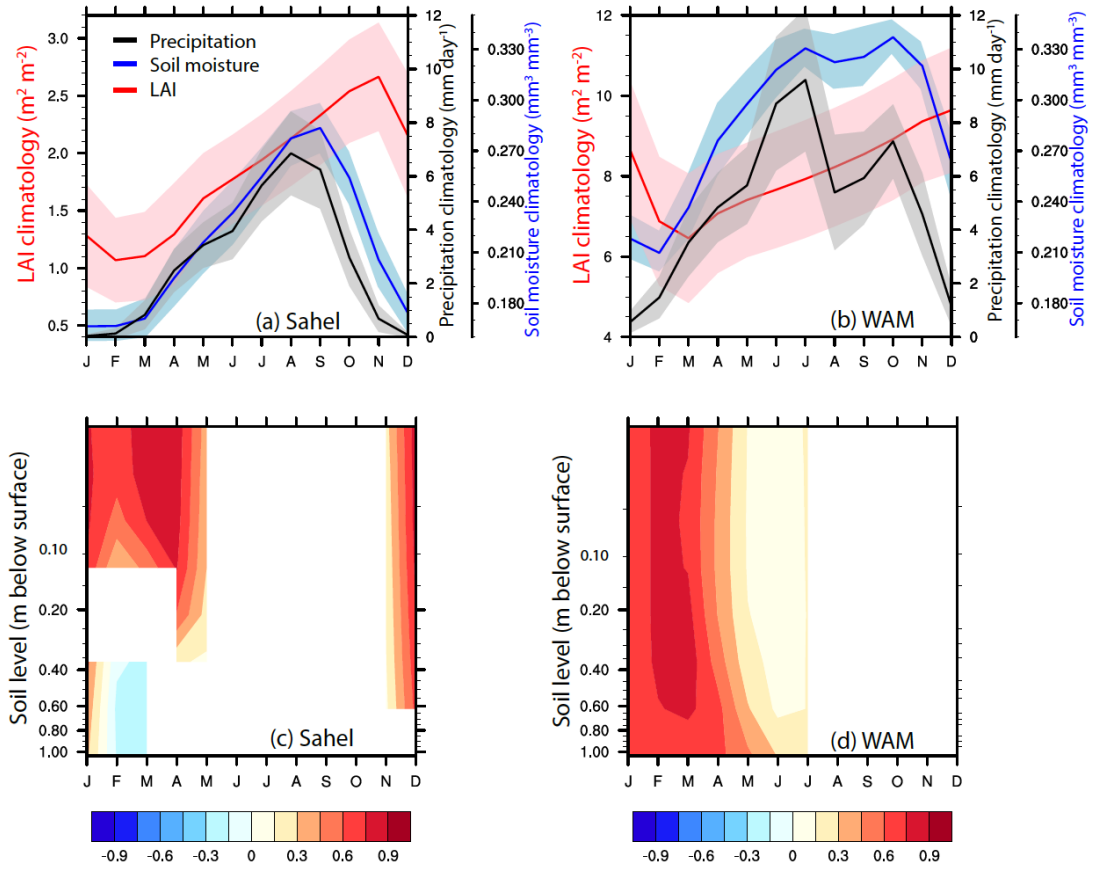


**Figure 10** LAI-soil moisture coupling in CESM. (a,b) Mean seasonal cycle of area-average LAI ($m^2$ $m^{-2}$, red), near-surface soil moisture within top 1 m ($mm^3$ $mm^{-3}$, blue) weighted by layer depth, and precipitation (mm $day^{-1}$, black), with their interannual ±1 standard deviation (shading) across the (a) Sahel and (b) WAM region. (c,d) Temporal correlation between area-average LAI and previous month's soil moisture by depth across the (c) Sahel and (d) WAM. Dynamical experiments with modified initial soil moisture are conducted for January, March, April, and December across Sahel, and February and March across WAM. In (c-d), only significant correlations are shown based on the Student's t-test ($p<0.1$).

On the first day of each two-month run and at each grid cell, soil moisture within the top 1 m, which contains over 70% of the simulated total root mass in both the Sahel and WAM region, is increased to the long-term 95$^{th}$ percentile from CTRL in the SOILINC experiments or decreased to the long-term 5$^{th}$ percentile from CTRL in the SOILDEC experiments. The anomalously wet (dry) soil initially enhances (inhibits) vegetation growth (Figure 11a,d), which in turn depletes (augments) the soil moisture within the rooting depth and causes a drying (wetting) trend in soil moisture in the later days (Figure 11b,e). By the second month of the simulations, the initial soil moisture anomalies are notably reduced compared to the control run, but still significantly more pronounced than in $EXP_{LAI}$, i.e. wetter in SOILINC than in LAIINC and drier in SOILDEC than in LAIDEC (Figure 11c,f). Therefore, $EXP_{SOIL}$ successfully represents the effects of coupled soil moisture-LAI anomalies on the atmosphere by allowing vegetation anomalies to naturally evolve. Atmospheric responses to the coupled soil moisture-LAI anomalies are analyzed in the second month of the experiment, when the soil moisture and LAI anomalies reach equilibrium and LAI anomalies are maximized (Figure 11a,d).

In order to directly compare with the statistically-assessed feedbacks, the responses in the $EXP_{SOIL}$ are represented by the regression coefficient of atmospheric anomalies upon LAI anomalies across the 60 ensemble members (30 SOILINC and 30 SOILDEC). The statistical significance and 90% confidence interval (5$^{th}$ -95$^{th}$ percentile) are obtained similarly as in $EXP_{LAI}$.

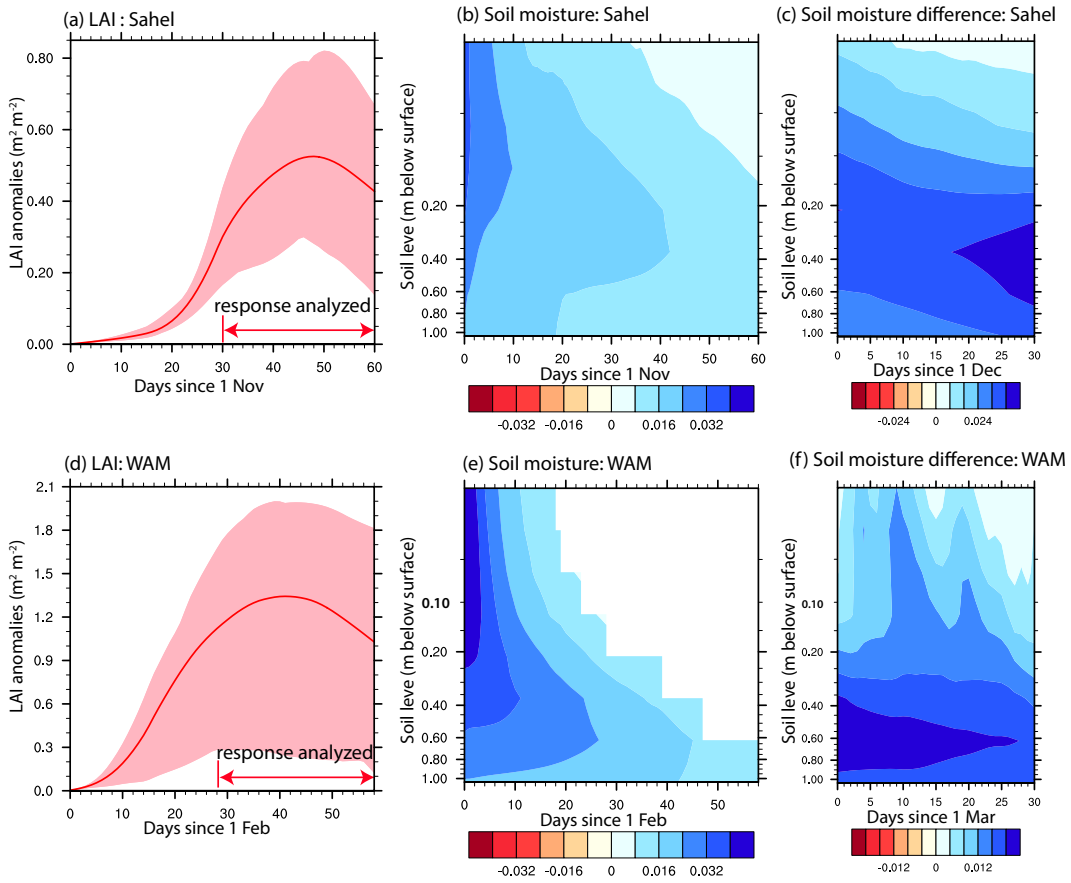


**Figure 11** Evolution of anomalies [(SOILINC-SOILDEC)/2-CTR] in (a,d) LAI ($m^2$ $m^{-2}$), (b,e) soil moisture ($mm^3$ $mm^{-3}$) across the (a-b) Sahel since 1 November and (c-d) WAM region since 1 February in $EXP_{SOIL}$. (c,f) Evolution of soil moisture responses ($mm^3$ $mm^{-3}$ $LAI^{-1}$) in the top 1 m between SOIL experiments and LAI experiments [(SOILINC-SOILDEC)/2-(LAIINC-LAIDEC)/2] across the (c) Sahel since 1 December and (f) WAM region since 1 March. The lines in (a-b) represent ensemble average, and shading represents the minimum and maximum LAI anomalies among all ensemble members. In (b,c,e,f), only significant differences are shown in color based on the Student's t test ($p<0.1$). Although soil moisture anomalies in SOIL are not sustained into the second month in the WAM region, the SOILINC (SOILDEC) experiments produce significantly wetter (drier) soil than the LAIINC (LAIDEC) experiments, which reflects the strong positive coupling between soil moisture and LAI.

### 2.3. Observational GEFA framework

In order to identify the key oceanic and terrestrial drivers of North African climate, the multivariate statistical method, GEFA, is applied to gridded observations, remote sensing products, and reanalyses (Table 1, Yu et al. 2017b). The GEFA forcing matrix is comprised of the observed leading two SST EOF modes from eight non-overlapping basins, area-average Mediterranean SSTs, and time series of area-average NDVI across the Sahel, WAM region, and HOA. (Figures 12-13):

$$O = [TP1\ TP2\ NP1\ NP2\ TI1\ TI2\ TA1\ TA2\ NA1\ NA2\ SP1\ SP2\ SA1\ SA2\ SI1\ SI2\ MED\ NDVI_{Sahel}\ NDVI_{WAM}\ NDVI_{HOA}]$$

In terms of terrestrial forcings, the Sahel, WAM, and HOA represent unique North African landscapes, i.e. mainly savanna and woody savanna across the WAM, savanna and grasslands across the Sahel, and shrubs and bare ground across the HOA (Figure 13). The geographic extent of the three ecoregions is determined through rotated EOF analysis (Figure 13) of monthly remote sensing standardized NDVI anomalies from Boston University's 30-year NDVI3g dataset (Zhu et al. 2013). Note that the geographic extent of these three ecoregions in the observational analysis are different from that in the CESM CTRL, given the slightly biased vegetation distribution and mean vegetation growth simulated by CESM. Grid cells with seasonal-average vegetation greenness in the Congo are not included in the GEFA analysis due to uncertainties in remotely sensed NDVI within tropical rainforests caused by the low sensitivity of remotely sensed NDVI to variations in high-density vegetation and cloud contamination (Samanta et al. 2012). NDVI, instead of LAI in the GEFA application to CESM CTRL, is used as an index for vegetation growth in the observational study, because NDVI is directed sensed by satellite

instruments while LAI is calculated from NDVI and thus contains another source of uncertainty introduced by the conversion from NDVI to LAI (Zhu et al. 2013).

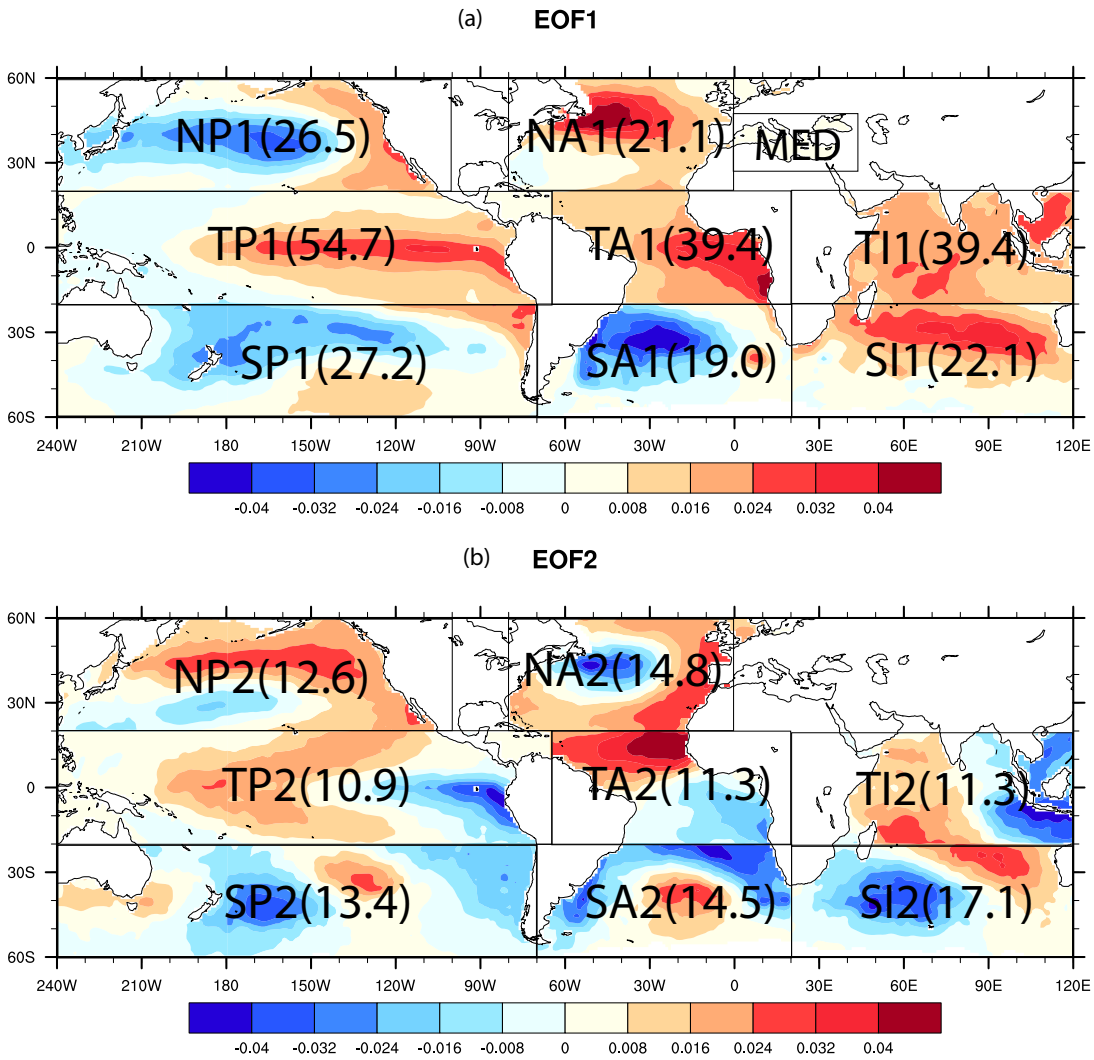


**Figure 12** Collage of the spatial patterns of the observed (a) first and (b) second EOF modes (unitless) of SST anomalies in eight ocean basins, namely the tropical Pacific (TP), North Pacific (NP), tropical Indian (TI), tropical Atlantic (TA), North Atlantic (NA), South Pacific (SP), South Indian (SI), and South Atlantic (SA). Analysis is based on the Met Office – Hadley Centre Global Sea Ice Coverage and SST dataset for 1900-2011. The percent explained variance in SST anomalies across a specific oceanic basin by each mode is identified in parentheses. EOF analysis is performed within each oceanic basin separately.

In each season and at each grid cell, the relatively unimportant forcings are dropped from the forcing matrix according to stepwise selection using AIC. By reducing the number of forcings considered simultaneously by GEFA, the reliability of estimated feedbacks associated with the remaining forcings is enhanced. Atmospheric response fields, including circulation, moisture, vertical motion, temperature, precipitation, and clouds, are obtained from the observational, remote sensing, and reanalysis datasets. The seasonal cycle and linear trend are removed from all fields in order to focus on seasonal and interannual variability (Wang et al. 2013). The statistical significance of GEFA feedback matrices is assessed using the Monte Carlo bootstrap method with 1000 random iterations in which the atmospheric time series are scrambled (Czaja and Frankignoul 2002). In order to achieve sufficient length of data and obtain reliable estimates of the feedback matrices, seasonal feedbacks are estimated by aggregating data from the consecutive three months, so that the effective sample size is three times the number of years. For example, in order to assess the feedbacks in January-March (JFM), the atmospheric data in January, February, and March of each year is pasted together to form a new time series; instantaneous forcing matrix contains instantaneous oceanic and vegetation fields in January, February, and March of each year.

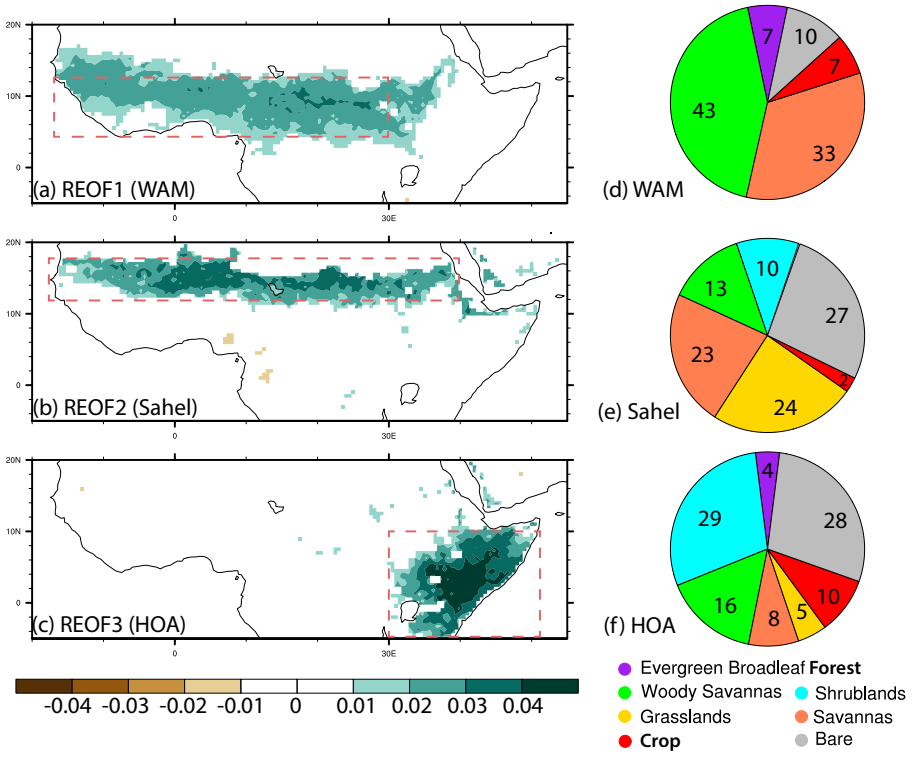


**Figure 13** Spatial pattern of the leading three REOFs in standardized remotely sensed NDVI, representing variability across the (a) WAM region (5˚N-12˚N, 20˚W-30˚E), (b) Sahel (12˚N-17˚N, 20˚W-40˚E), and (c) HOA (5˚S-10˚N, 30˚E-52˚E). Percent area of land cover types across the (d) WAM region, (e) Sahel, and (f) HOA. Analysis is based on the AVHRR GIMMS NDVI3g dataset for 1982-2011. The REOF analysis helps define regions (red boxes) for area-average NDVI to be included in the GEFA forcing matrix.

## 2.4. Observational datasets and multi-data observational benchmark

### 2.4.1. Multi-dataset bootstrapping method

Multiple observational, remote sensing, and reanalysis datasets (Table 1) are analyzed for the Sahel in the GEFA framework. By applying the Monte Carlo bootstrapping approach (Efron

and Efron 1982), the potential impacts of observational measurement errors across the data-sparse North Africa on estimated GEFA response fields are reduced. Furthermore, this approach facilitates a reliable estimation of the multi-dataset mean and quantification of observational uncertainty in the GEFA-based atmospheric responses to oceanic and terrestrial forcings. The GEFA-based response is first obtained from each dataset, and then a probability distribution function (PDF) of the weighted-average response of all datasets is generated by the Monte Carlo bootstrap approach with 1000 random iterations. With each iteration, weights are randomly generated from a uniform distribution with values from 0 to 1 and standardized so that they sum up to one. The standardized weights are assigned to the datasets in the order of their regional reliability, with the highest weight assigned to the most reliable dataset. Based on the multi-dataset PDF of the 1000 weighted-averages, the multi-dataset average and uncertainty range of the responses to oceanic and terrestrial forcings are obtained. The regional reliability of each observational dataset across the Sahel is evaluated against an independent data in terms of the temporal correlation and root-mean-square-difference, as outlined in Tables 2-5, leading to a practical ranking of all datasets to be applied in the Monte Carlo bootstrap approach.

### 2.4.2. Station dust observations and MERRA-2 dust reanalysis

In order to investigate potential vegetation-dust-precipitation feedbacks, the observed terrestrial impacts on dust emission and concentration are examined. Beyond the Modern-Era Retrospective Analysis for Research and Application-2 (MERRA-2) reanalysis of dust-related fields, dust observations are retrieved from the National Climatic Data Center (NCDC) hourly global and U.S. Integrated Surface hourly data set for 1982-2015 at 502 North African stations. At each station, a “dust day” is defined as a day in which either dust/sand storm or severe dust/sand storm is reported at least once, or dust suspension is reported for at least a quarter of

the total number of observations during the daytime (Yu et al. 2013, 2015). Therefore, the dust day metric is a combined measure of the frequency and intensity of dust activity.

Station observations are first interpolated to a 0.25˚ x 0.25˚ grid. In each grid cell, a "regional dust day" is defined if at least one station within that grid cell indicates a "dust day". Monthly dust frequency in each grid cell is calculated when dust observations are available on more than half of the days during that month, or otherwise left as a missing value.

In addition to station dust observations, dust aerosol reanalysis from MERRA-2 is also analyzed. The MERRA-2 reanalysis of aerosols includes assimilation of bias-corrected Aerosol Optical Depth (AOD) from AVHRR over the oceans, Moderate Resolution Imaging Spectroradiometer (MODIS) sensors on both Terra and Aqua satellites, Multi-angle Imaging SpectroRadiometer (MISR) over bright surfaces, and Aerosol Robotic Network (AERONET) data. The vertical structure of MERRA-2 aerosol reanalysis has been successfully validated using Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) data, including over North Africa (Buchard and Da Silva 2016). Furthermore, the MERRA-2 surface dust concentration reanalysis exhibits similar seasonal cycle and interannual variability with the station dust frequency across the Sahel (not shown).

**Table 1** List of analyzed observational, remote sensing, and reanalysis datasets for each GEFA response variable. Asterisks denote datasets that incorporate remotely-sensed information.

| **Variables** | **Dataset** | **Analyzed Years** | **Spatial Resolution** | **Reference** |
|---|---|---|---|---|
| Vertical motion, precipitable water, 2-m specific humidity, 10-m | National Aeronautics and Space Administration (NASA) Modern-Era Retrospective Analysis for Research and Applications (MERRA)* | 1982-2011 | 0.5° x 0.66° | Rienecker et al. 2011 |

| | | | | |
|---|---|---|---|---|
| wind speed, sea-level pressure (SLP) | Japanese 55-year Reanalysis (JRA55)* | 1982-2011 | 0.63° x 0.63° | Kobayashi et al. 2015 |
| | European Centre for Medium-Range Weather Forecast (ECMWF) Interim Reanalysis (ERA-In)* | 1982-2011 | 0.75° x 0.75° | Dee et al. 2011 |
| | National Centers for Environmental Prediction (NCEP)-Climate Forecast System Reanalysis (CFSR)* | 1982-2010 | 0.5° x 0.5° | Saha et al. 2010 |
| 2-m air temperature | University of Delaware (UDEL) Terrestrial Air Temperature | 1982-2011 | 0.5° x 0.5° | Matsuura and Willmott 2012 |
| | University of East Anglia Climatic Research Unit (CRU) Time Series (TS3.22) | 1982-2011 | 0.5° x 0.5° | Harris et al. 2014 |
| Precipitation | UDEL Terrestrial Precipitation | 1982-2011 | 0.5° x 0.5° | Matsuura and Willmott 2012 |
| | Global Precipitation Climatology Centre (GPCC) | 1982-2011 | 0.5° x 0.5° | Schneider et al. 2008 |
| | CRU TS3.22 | 1982-2011 | 0.5° x 0.5° | Harris et al. 2014 |
| | Global Precipitation Climatology Project (GPCP)* | 1982-2011 | 2.5° x 2.5° | Huffman 1997 |
| | Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP)* | 1982-2011 | 2.5° x 2.5° | Xie and Arkin 1997 |
| | Global Historical Climatology Network stations - daily | 1982-2011 | | |
| Outgoing longwave radiation (OLR) | Advanced Very High Resolution (AVHRR) Pathfinder Atmospheres-Extended (PATMOS-x)* | 1982-2011 | 0.1° x 0.1° | Heidinger et al. 2014 |
| | National Oceanic and Atmospheric Administration (NOAA) Climate Data Records (CDR)* | 1982-2011 | 1° x 1° | National Research Council 2004 |
| Evapotranspiration | Mao's merged diagnostic ET product* | 1982-2010 | 0.5° x 0.5° | Mao et al. 2015 |

| | | | | |
|---|---|---|---|---|
| (ET) | Global Land Evaporation Amsterdam Model (GLEAM) Global Evapotranspiration* | 1982-2010 | 0.5° x 0.5° | Mu et al. 2007 |
| Dust emission, concentration, and transport | MERRA-2* | 1982-2011 | 0.5° x 0.63° | Bosilvich et al. 2015 |
| Dust frequency | Global Historical Climatology Network stations - hourly | 1982-2011 | | |

**Table 2** Summary of multi-dataset evaluation and ranking. In order to reduce the observational error, weights are assigned to different data products according to their regional reliability when generating the multi-dataset PDF of the response. The listed weights are specific to the entire North Africa region.

| **Variable** | **Ranking criteria** | **Weights for North Africa** |
|---|---|---|
| SLP, 10-m wind, vertical motion | Evaluate surface u- and v-wind against data from 502 North African stations in terms of the temporal correlation and root-mean-square error | JRA55 = ERA-In > CFSR > MERRA |
| Precipitation | Evaluate against GPCP satellite-gauge merged product in terms of the temporal correlation and root-mean-square error | GPCP > GPCC > CMAP > UDEL> CRU |
| 2-m air temperature | Number of North African stations included in dataset | UDEL > CRU |
| Precipitable water, 2-m specific humidity | Evaluate against NVAP precipitable water in terms of the temporal correlation | ERA-In = CFSR > JRA55 > MERRA |
| ET | Number of data sources included in dataset | Mao's ET > GLEAM |
| OLR | Number of included satellite instruments | PATMOS-x = CDR |

**Table 3** Evaluation of daily surface u- and v-wind from reanalyses against 502 stations across the four sub-regions in North Africa (1979-2013): regional average temporal correlation and root-mean-square-error (RMSE, m $s^{-1}$) in daily u- and v-wind between station wind observations and the nearest grid cell from reanalysis. The determined weights are applied to SLP, 10-m wind speed, and vertical motion, since these variables are closely related to surface u- and v-wind.

| **Variable** | **u-wind (Temp. Correlation/RMSE in m/s)** | | | | **v-wind (Temp. Correlation/RMSE in m/s)** | | | |
|---|---|---|---|---|---|---|---|---|
| Region | Sahel | HOA | WAM | Congo | Sahel | HOA | WAM | Congo |
| MERRA | 0.53/2.39 | 0.45/2.21 | 0.34/2.12 | 0.34/1.99 | 0.55/2.56 | 0.50/2.12 | 0.32/2.38 | 0.28/2.17 |
| CFSR | 0.51/2.27 | 0.45/2.09 | 0.37/2.05 | 0.34/1.96 | 0.57/2.54 | 0.45/2.17 | 0.37/2.31 | 0.39/1.99 |
| ERA-In | 0.55/2.32 | 0.48/2.23 | 0.40/2.02 | 0.35/1.89 | 0.59/2.09 | 0.50/2.22 | 0.38/2.48 | 0.41/1.97 |
| JRA-55 | 0.59/1.93 | 0.49/2.19 | 0.39/1.94 | 0.36/1.77 | 0.58/2.11 | 0.51/2.19 | 0.38/2.03 | 0.39/2.01 |

**Table 4** Evaluation of daily precipitable water from reanalyses against NVAP (1988-2001) across the four sub-regions in North Africa: regional average temporal correlation between daily NVAP and each reanalysis. NVAP is chosen as a benchmark because of its spatial and temporal coverage, especially over the North Africa region where radiosonde data is sparse. The determined weights are also applied to 2-m specific humidity, which is closely related to precipitable water.

| **Temp. Correlation*** | **Sahel** | **HOA** | **WAM** | **Congo** |
|---|---|---|---|---|
| MERRA | 0.35 | 0.59 | 0.54 | -0.08 |
| JRA-55 | 0.41 | 0.66 | 0.51 | 0.29 |
| ERAIn | 0.49 | 0.67 | 0.55 | 0.30 |
| CFSR | 0.46 | 0.68 | 0.63 | 0.43 |

*RMSE is not computed due to the wet bias in NVAP (Amenu et al. 2005).

**Table 5** Evaluation of monthly precipitation from gridded gauge observations against the satellite-gauge merged GPCP (1979-2014) across the four sub-regions in North Africa: regional average temporal

correlation and RMSE (cm $mon^{-1}$) between monthly precipitation anomalies from each precipitation dataset and GPCP. GPCP is considered as a benchmark because it incorporates both gauge observations and multiple remote-sensing products (Huffman, 1997).

| **Correlation/RMSE (cm $mon^{-1}$)** | **Sahel** | **HOA** | **WAM** | **Congo** |
|---|---|---|---|---|
| CRU | 0.68/1/40 | 0.62/2.54 | 0.61/3.08 | 0.48/3.34 |
| UDEL | 0.72/1.32 | 0.73/2.38 | 0.68/2.65 | 0.57/3.25 |
| CMAP | 0.75/1.09 | 0.74/2.10 | 0.72/2.43 | 0.69/3.33 |
| GPCC | 0.79/1.18 | 0.80/2.32 | 0.73/2.75 | 0.70/3.01 |

# Chapter 3 GEFA validation

## 3.1. GEFA's capability in capturing oceanic impacts on regional climate

### 3.1.1. Stepwise GEFA versus full GEFA

In the comparison between FGEFA and SGEFA, a diverse set of variables is considered, including local latent (LHFLX) and sensible (SHFLX) heat fluxes over the corresponding forcing oceanic basin, 850-hPa geopotential height (Z850), 200-hPa geopotential heights (Z200), 500-hPa vertical velocity (ω500) over the tropics-subtropics (35˚S-35˚N, 0-360˚E), and 2-m air temperature (Temp) and precipitation (Precip) over North Africa (10˚S-17˚N, 20˚W-55˚E).

The advantage of SGEFA over FGEFA is more substantial when analyzing short data records. With long data records, such as the full 3440-year CESMLENS time series, FGEFA and SGEFA yield nearly identical response fields, with an average spatial correlation between the two GEFA approaches of 0.96 across the seven variables (SHFLX, LHFLX, Z850, Z200, ω500, Temp, Precip) and 12 months. With short records, SGEFA generally outperforms FGEFA, with higher spatial correlations between the response fields from the short record and the full 3440-year CESMLENS records. Therefore, SGEFA can achieve the same spatial correlation as FGEFA with shorter minimum data lengths (Figure 14). For instance, using a 100-year data record, SGEFA can improve the spatial correlation between the response fields from the short record and the full record from 0.8 (TP1, TA1) and 0.75 (TI1) to 0.85. The difference in response fields between FGEFA and SGEFA diminishes with increasing length of data, and finally their response fields are nearly identical when the complete 3440 years of data is used. It is noteworthy that the monotonic relationship between the threshold spatial correlation and required data length is not linear, as the spatial correlation between the response fields from the

long and short records increases more rapidly with increasing sample size when the data is short. For example, for TP1 (also true with TI1 and TA1), it only takes an extra 54 years to increase the spatial correlation from 0.7 to 0.85, while it takes another 213 years to increase the spatial correlation from 0.85 to 0.95.

In summary, SGEFA does not require as long of a time series as FGEFA to achieve stable GEFA response estimates, and the benefit is more obvious when working with short data records. If a 100-year data record is available, about 72% ($0.85^2$) of the monthly GEFA response field using the full time period can be reproduced using SGEFA, while only about 64% ($0.8^2$) can be reproduced using FGEFA. These results suggest that SGEFA is preferred over FGEFA when performing observational analyses with the approximate 30 years of satellite data. Another implication is that with roughly 30 years of observational data, monthly response fields estimated by either the FGEFA or SGEFA are unreliable, while the seasonal response fields are largely trustworthy from SGEFA.

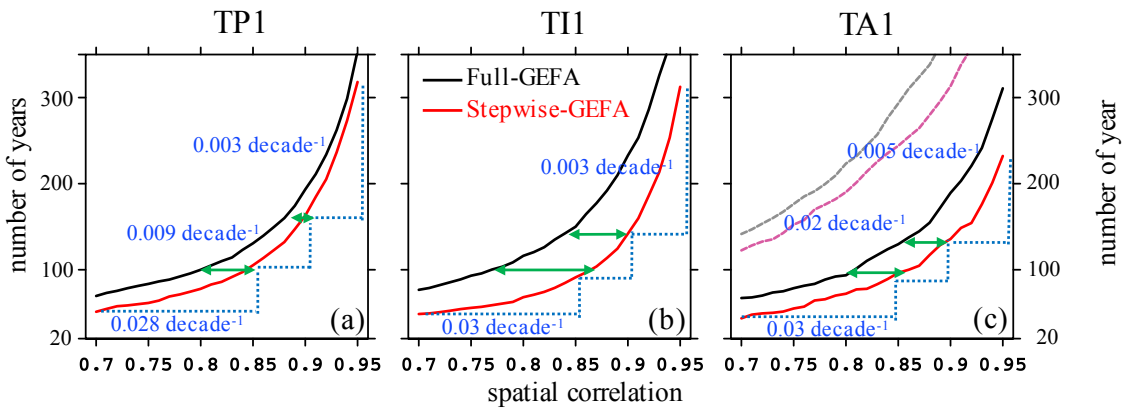


**Figure 14** Length of data record required for stable estimates of GEFA response fields averaged across seven variables (SHFLX, LHFLX, Z850, Z200, ω500, Temp, Precip) and 12 months to tropical oceanic forcings from (a) the tropical Pacific (TP1), (b) tropical Indian (TI1), and (c) tropical Atlantic (TA1)

Oceans by full-GEFA (black lines) and stepwise-GEFA (red lines) corresponding to different thresholds in the spatial correlation between the full 3440-year CESM-LENS data. The blue numbers indicate the growth rate of the spatial correlation with increasing data record length. For TA1, results are shown both for the full annual cycle of 12 calendar months (gray: full-GEFA, pink: stepwise-GEFA) and only JJA (black: full-GEFA, red: stepwise-GEFA), due to weak forcing and no clear physical meanings of TA1 during September to May.

### 3.1.2. GEFA's capability at capturing key oceanic impacts on North African climate

#### 3.1.2.1 Identifying key oceanic forcings in CESM

According to SGEFA, tropical oceanic forcings are the dominant drivers of North African climate in CESM (Figure 15), consistent with previous studies (Folland et al. 1986; Giannini et al. 2003; Lu and Delworth 2005; Hoerling et al. 2006). TP1, which represents ENSO, substantially affects air temperature and precipitation for all three North African sub-regions in nearly all months. TI1, which consists of the Indian Ocean Basin (IOB) mode during February to July and the Indian Ocean Dipole (IOD) mode during August to January in CESM, affects temperature and precipitation across the HOA almost all-year-round, especially the short rains (OND). TA1, which represents the Atlantic Niño mode during June to August, regulates air temperature and precipitation over the WAM region and Sahel during the pre- and peak-monsoon seasons. TP2, which is characterized by the tropical Pacific meridional mode during October-May in CESM, affects temperature and precipitation over the HOA. TI2, which represents the IOB mode during August-January, influences HOA temperature and precipitation. In summary, in CESM, TP1 and TA1 exert the dominant impacts on the Sahel and WAM region, and the leading two EOF modes across the TP and TI are important oceanic regulators for HOA.

Since the magnitude of the impacts from TI2 and TP2 on North African climate are small, the GEFA validation effort that compares dynamically- and statistically-assessed atmospheric responses to oceanic anomalies mainly focuses on TP1, TI1, and TA1.

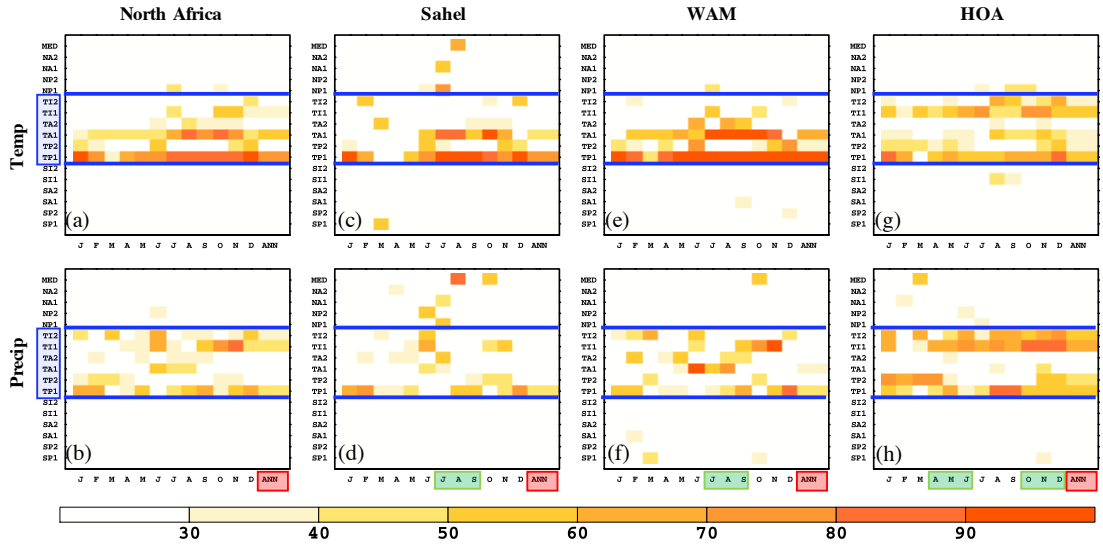


**Figure 15** Percentage of area with significant responses ($p<0.1$, based on Monte Carlo bootstrapping method) in (a-d) 2-m air temperature (Temp) or (e-h) precipitation (Precip) by month across the Sahel (12°N – 17°N, 20°W – 40°E), WAM region (5°N – 12°N, 20°W – 30°E), HOA (10°S – 10°N, 30°E – 50°E), or entire North Africa (10°S – 20°N, 20°W – 50°E) to 17 individual oceanic modes, based on the application of stepwise-GEFA to the CTRL run. The annual mean percentage of area with significant responses is shown as the last column of each panel figure (red boxes). The tropical oceanic modes are indicated by blue lines. The green boxes denote the wet seasons per region.

### 3.1.2.2 General comparison of statistically- and dynamically-assessed responses

Three metrics are considered in the comparison between statistically- and dynamically-assessed responses: sign and significance consistency, spatial correlation, and response magnitude, which reflects the accuracy of GEFA in capturing the sign, spatial distribution, and

magnitude of atmospheric responses. In order to measure the sign and significance consistency, the relationship between statistical and dynamical assessments for each grid cell is classified into five categories: (1) both significant ($p<0.1$) and of the same sign, (2) neither significant, (3) dynamically significant but statistically insignificant, (4) dynamically insignificant but statistically significant, (5) both significant but of opposite sign. The percent consistency for each response variable is calculated as the ratio between the number of grid cells in which statistical and dynamical assessments fall in the same category and total number of grid cells. The spatial correlation between the statistically- and dynamically estimated response fields is calculated for each oceanic forcing using only grid cells that attain statistically significance in both assessments. The response magnitude is represented by the seasonal cycle of area-average responses across Sahel, WAM region, and HOA.

Overall, the statistical and dynamical assessments yield fairly good agreement in terms of the atmospheric responses to tropical oceanic forcings in CESM (Figure 16). The percentage of consistency, as the sum of categories (1) and (2) in Figure 16, is 69% for TP1, 61% for TI1, and 62% for TA1, averaged across seven variables (SHFLX, LHFLX, Z850, Z200, ω500, Temp, Precip) and 12 months. Low consistency between the statistically- and dynamically-assessed responses, such as for TP1 in March, TI1 in February and March, or TA1 during September-May, is largely caused by category (3), in which the dynamical response is significant, but the statistical response is not. The percentage of absolutely inconsistency, or category (5), is extremely low across all three forcings and each calendar month (< 5%).

Almost all seven variables exhibit low consistency (<50%) in February and March to TI1 forcing. The disagreement between the statistical and dynamical assessments during February and March is caused by the extremely high temporal correlation between TP1 and TI1 in CESM.

The instantaneous correlation is about 0.88, while the correlation peaks at 0.92 when TP1 leads TI1 by three months. In other words, the SST anomalies over the tropical Indian Ocean are unlikely naturally independent, but rather a product of ENSO. As a result, the statistical stepwise GEFA indicates that only 20% (24%) of the global response area is significant in February (March), while the dynamical assessment suggests that 67% of the response area is significant in both February and March. Therefore, when the temporal correlation between two forcings is extremely high, it is more appropriate to evaluate their combined impacts, as they are not naturally independent of each other. To test this hypothesis, we perform another set of dynamical experiments (TP1TI1), following the same general approach as the other dynamical experiments. In this case, EOF anomalies are applied to the tropical Pacific and tropical Indian Ocean basins together (20˚S-20˚N,35˚E-60˚W), and the imposed SST anomalies in TP1TI1 is based on this expanded EOF region; the corresponding time series is also used as one of the forcings in SGEFA. SGEFA successfully captures the combined impact of TP1 and TI1 produced by the TP1TI1 experiment that positive TP1 and TI1 together leads to positive anomalies in air temperature across the majority of tropical and subtropical Africa (Figure 17).

Sensible and latent heat fluxes exhibit low consistency to TA1 forcing, especially during September-May when the TA1 SST forcing is weak. The average absolute SST anomaly in these months is only about 0.23˚C in the CTRL, leading to generally insignificant local responses in the sensible and latent heat fluxes in GEFA but marginally significant responses in the dynamical assessment. However, for the remote variables, including geopotential height and vertical velocity, both GEFA and dynamical assessments indicate insignificant responses, thereby leading to high consistency between these two approaches.

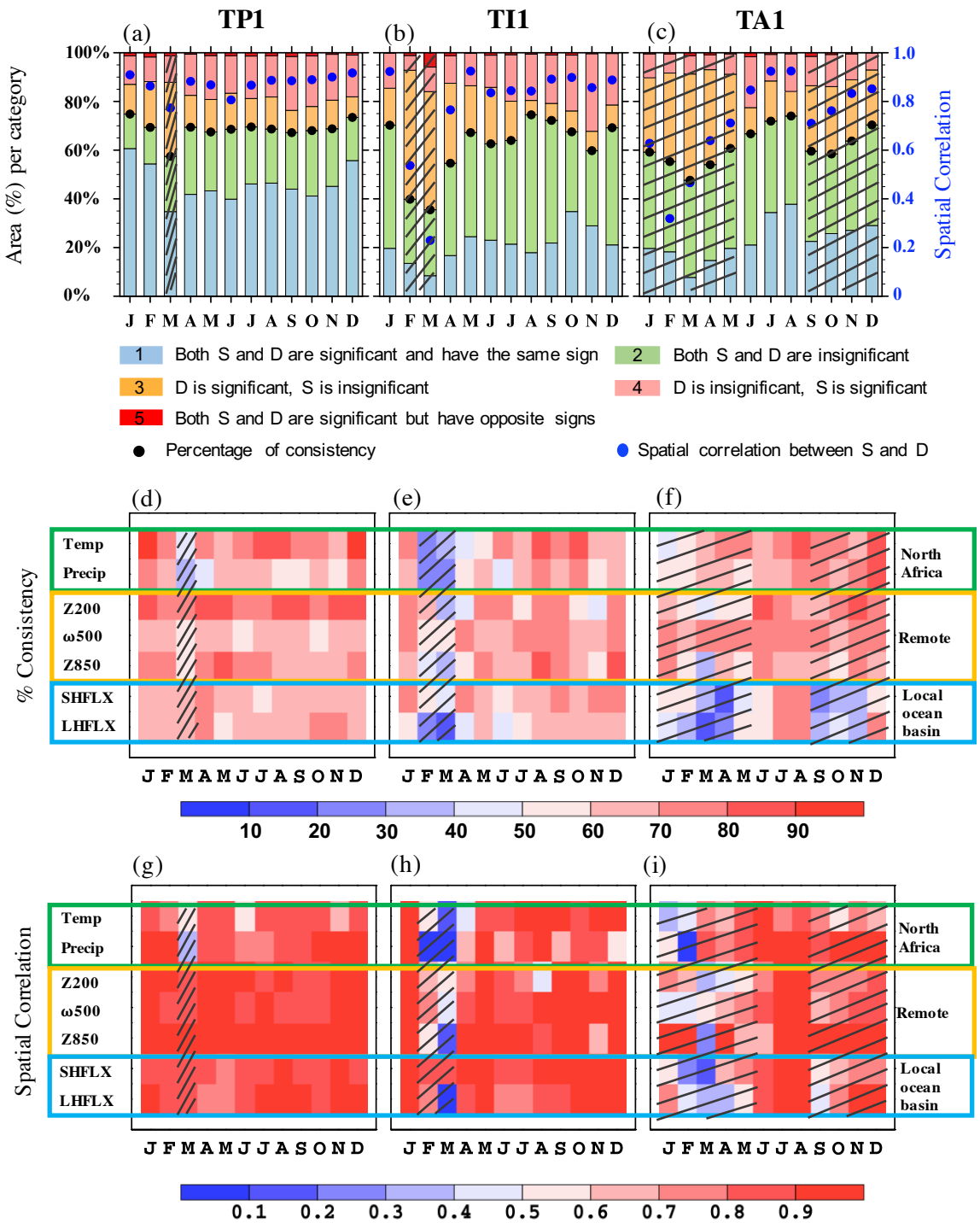


**Figure 16** Statistical (S) and dynamical (D) assessment comparison in CESM. Percentage of area covered by each category for (a) TP1, (b) TI1, and (c) TA1 is averaged across seven variables. Black dots indicate the percentage of consistency by month. Blue dots represent the spatial correlation between the

statistically and dynamically assessed response fields averaged by month. Percentage of area with consistent response between statistical and dynamical assessment in CESM for (d) TP1, (e) TI1, and (f) TA1 forcings. Spatial correlation of response patterns between statistical and dynamical assessment in CESM for (g) TP1, (h) TI1, and (i) TA1 forcings. Hatching indicates that the inconsistency is caused by a high correlation between TP1 and TI1, or the weak forcing of TA1.

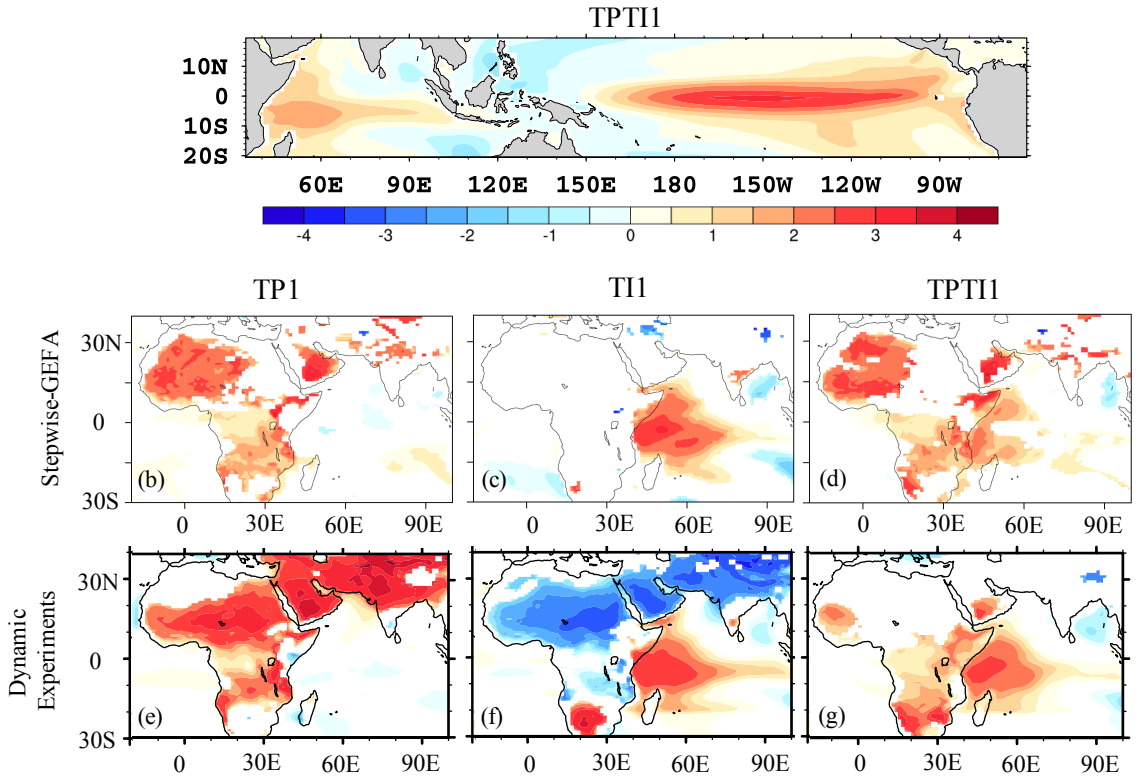


**Figure 17** (a) Spatial pattern of SST anomalies associated with the first EOF mode (unitless) of tropical Pacific and tropical Indian SSTs (TPTI1). 2-m air temperature response (unit: ºC $\sigma_{PC}^{-1}$) to TP1 (b), TI1 (c), and TPTI1 (d) forcing in stepwise-GEFA. 2-m air temperature response (unit: ºC $\sigma_{PC}^{-1}$) to TP1 (e), TI1 (f), and TPTI1 (g) forcing in dynamical experiments. Only statistical significant response fields ($p<0.1$) are shown.

The spatial correlation between the significant responses in the statistically- and dynamically-estimated response fields is generally high among the three oceanic forcings for most variables, with an average spatial correlation of 0.87 for TP1, 0.79 for TI1, and 0.72 for TA1 across all variables and all months. Leaving out the inconsistent cases discussed earlier, the average spatial correlation is 0.87 for TI1 during April-January and 0.90 for TA1 during June-August, which confirms that the apparent inconsistencies between the statistically- and dynamically-assessed responses are mainly due to the reasons previously discussed.

In terms of the response magnitude, the seasonal cycle of responses in 2-m air temperature (Figure 18) and precipitation (Figure 19) over the three North African ecoregions to the leading oceanic forcings, namely TP1, TI1, and TA1, are generally consistent between the statistical and dynamical assessments. The statistically and dynamically assessments are in quantitative agreement, with an root-mean-square-error of 0.15, 0.09 and 0.11 ℃ $\sigma_{PC}^{-1}$ in 2-m air temperature response, and 0.18, 0.19, and 0.14 $mm\ day^{-1}\ \sigma_{PC}^{-1}$ in precipitation response, across the Sahel, WAM region, and HOA, respectively.

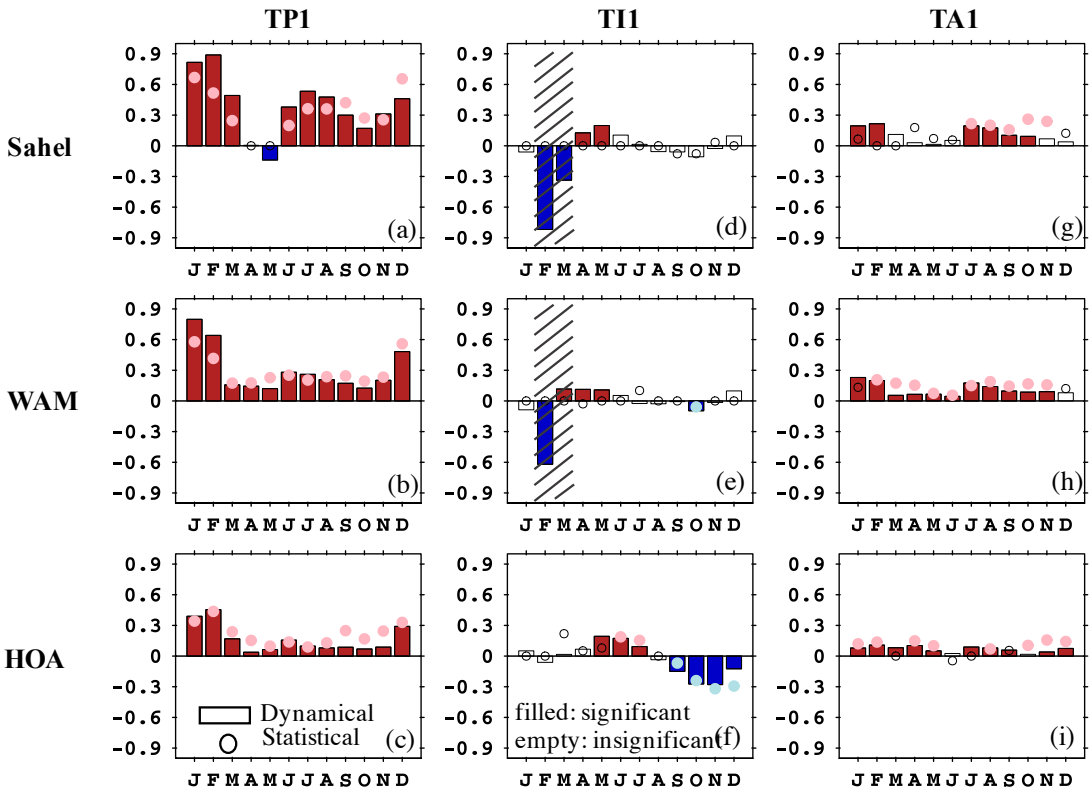


**Figure 18** Seasonal cycle of 2-m air temperature response (℃ $\sigma_{PC}^{-1}$) to (a – c) TP1, (d – f) TI1, and (g – i) TA1 over the (a, d, g) Sahel, (b, e, h) WAM region, and (c, f, i) HOA in CESM. Bars and dots indicate responses according to the dynamical experiments and statistical stepwise-GEFA, respectively. Filled bars and dots indicate that the responses are statistically significant (>90%) based on Student's t-test (dynamical experiments) and Monte Carlo bootstrapping method (stepwise-GEFA), respectively. The inconsistent surface air temperature response to TI1 between dynamical and statistical assessments is caused by the high temporal correlation between TP1 and TI1 (hatching area).

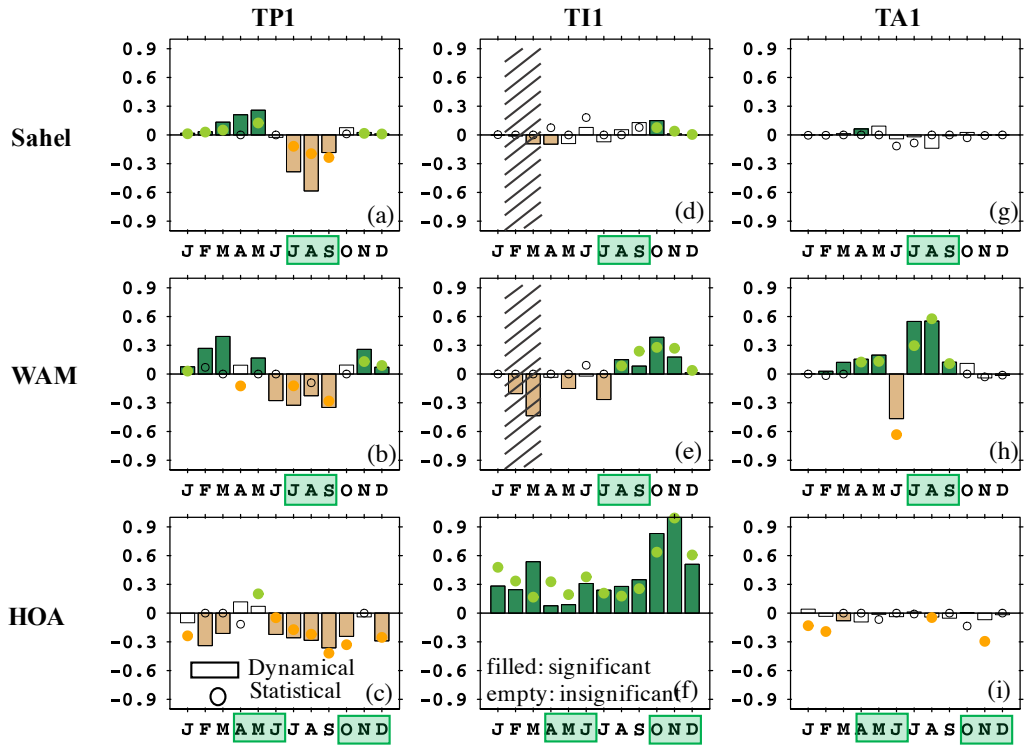


**Figure 19** The same as Figure 18, except for precipitation (unit: $mm\ day^{-1}\ \sigma_{PC}^{-1}$). Green boxes on the x-axis indicate the wet seasons per region in CESM.

### 3.1.2.3 Examples of individual oceanic impacts on North African climate in CESM: Statistical versus dynamical assessment

#### a. TP1

In CESM, under historical radiative forcing, El Niño favors warm anomalies during all seasons and dry anomalies during JAS over most of North Africa (Figure 20). Here, August is chosen as an example, as the response patterns and associated mechanisms are similar with June and July. In response to positive TP1 mode in CESM, both SGEFA and the dynamical experiments indicate a broad-scale reduction in rainfall over the tropical Atlantic Ocean (around 10˚N), extending partly over land to the Sahel, on the west side of the Ethiopian Highlands, and

into the western tropical Indian Ocean. The atmospheric anomalies resemble a baroclinic Rossby wave response to deep tropical heating and barotropical Rossby wave propagation into the extratropics, mainly across the Southern Hemisphere, since it is austral winter. In other words, over the tropical Pacific, negative height anomalies at 850-hPa and positive height anomalies at 200-hPa are generated by El Niño, while over the subtropical Pacific, negative height anomalies are produced in both the lower and upper troposphere.

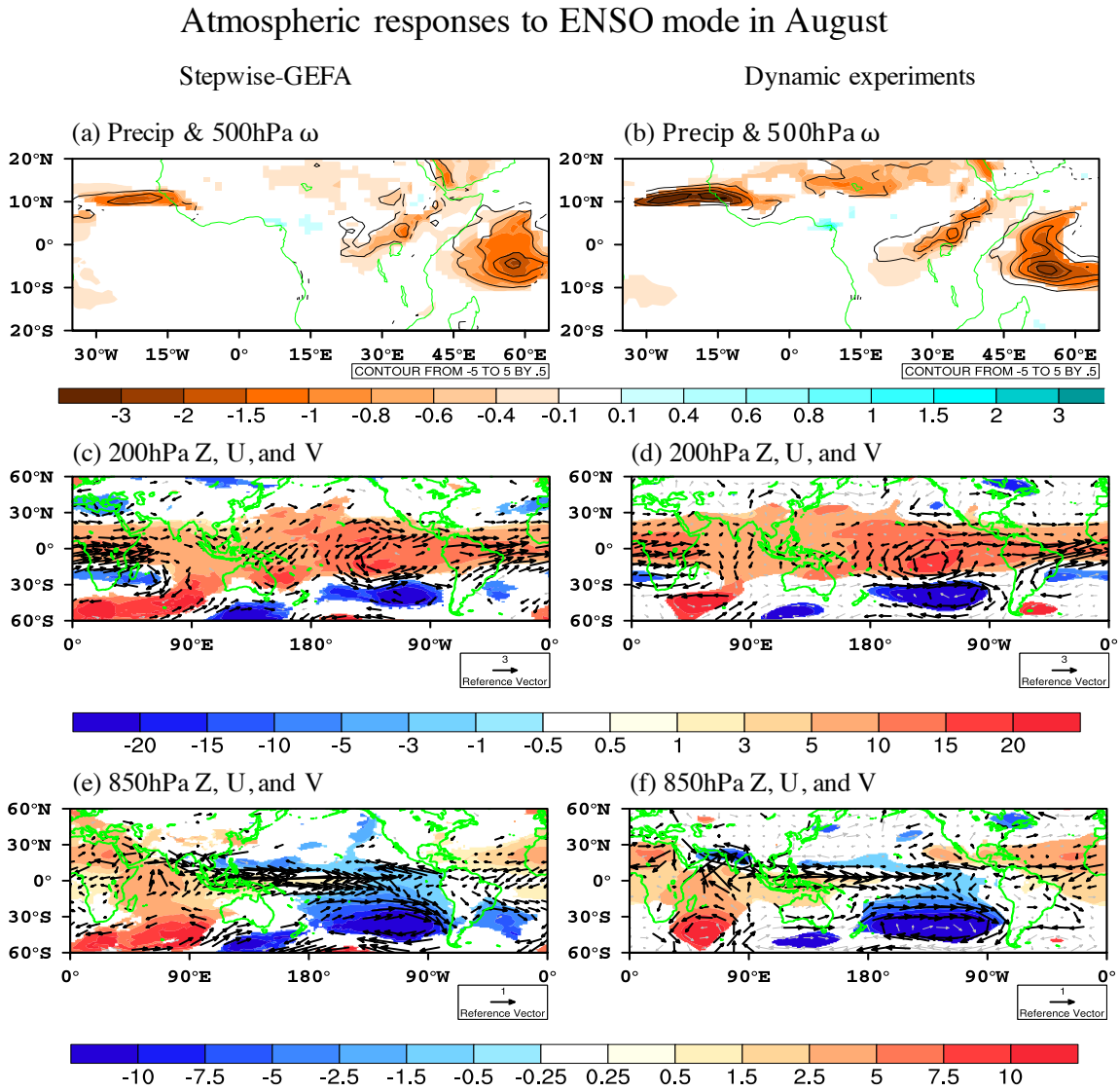

**Figure 20** Atmospheric responses in CESM to TP1 forcing using (a, c, e) SGEFA and (b, d, f) dynamical assessment in August. Atmospheric response fields include precipitation (shading, unit: $mm\ day^{-1}\ \sigma_{PC}^{-1}$) and 500-hPa vertical motion (contour, unit: Pa $s^{-1}\ \sigma_{PC}^{-1}$, solid lines indicate descending motion and dash lines indicate ascending motion), (c – d) 200-hPa and (e – f) 850-hPa geopotential heights (unit: $m\ \sigma_{PC}^{-1}$) and wind (gray vectors, unit: $m\ s^{-1}\ \sigma_{PC}^{-1}$). Only statistical significant ($p<0.1$) response fields are shown, except for the wind field, where black arrows represent significant wind responses. The three sub-regions, namely the Sahel, WAM region, and HOA, are shown in blue boxes in (a).

**b. TI1**

The TI1 mode, which represents the IOB (IOD)I mode during February-July (August-January) in CESM, mainly affects rainfall over the HOA almost year-round, especially during the short rains season (OND) and the relative dry summer season (JJAS) between the long and short rains (Figure 21). In June, both statistically and dynamically assessments agree that precipitation increases over the western slopes of the Ethiopian Highlands in response to positive phase of IOB in CESM. Over the Indian Ocean, the simulated atmospheric response is baroclinic, with negative height anomalies in the lower troposphere (850-hPa) and positive height anomalies in the upper troposphere (200-hPa). This anomalous low-level convergence supports increased rainfall over the tropical Indian Ocean. The enhanced rainfall over land is largely attributed to orographic lift, where enhanced atmospheric moisture content carried by the anomalous westerly wind is condensed and precipited out.

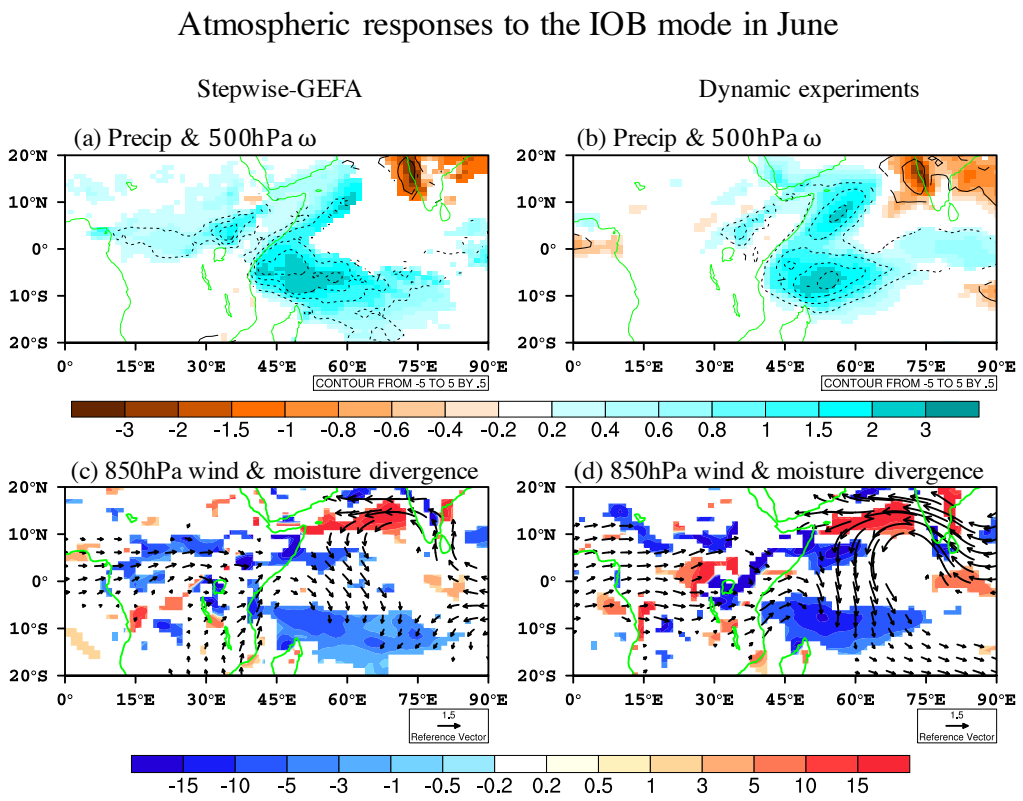


**Figure 21** Precipitation (shading, unit: $mm\ day^{-1}\ \sigma_{PC}^{-1}$) and 500-hPa vertical motion (contour, unit: Pa $s^{-1}\ \sigma_{PC}^{-1}$) response to TI1 forcing (resembles Figure 6c) using (a) stepwise-GEFA and (b) dynamical experiments during June for CESM. 850-hPa moisture divergence (shading, unit: $10^{-6}\ m\ s^{-1}\ \sigma_{PC}^{-1}$) and 850-hPa wind response (vector, $m\ s^{-1}\ \sigma_{PC}^{-1}$) to TI1 forcing using (c) SGEFA and (d) dynamical experiments. Only statistical significant response ($p<0.1$) fields are shown.

During the short rains, including November, both statistical and dynamical assessments indicate enhanced precipitation over the entire HOA in response to the positive IOD mode in CESM (Figure 22). Two mechanism could explain the enhanced precipitation over HOA: first, positive SST anomalies across the western tropical Indian Ocean increase air temperature, atmospheric moisture, and instability over HOA; second, as in June, enhanced convergence over

the anomalously warm western tropical Indian Ocean generate westerly wind anomalies over west and central Africa and easterly wind anomalies over the Indian Ocean. The zonally-orientated SST gradient enhances the zonal surface pressure gradient, leading to stronger easterly winds, which brings more moisture to Somalia than in June. Enhanced atmospheric moisture and wind convergence support an increase in HOA rainfall.

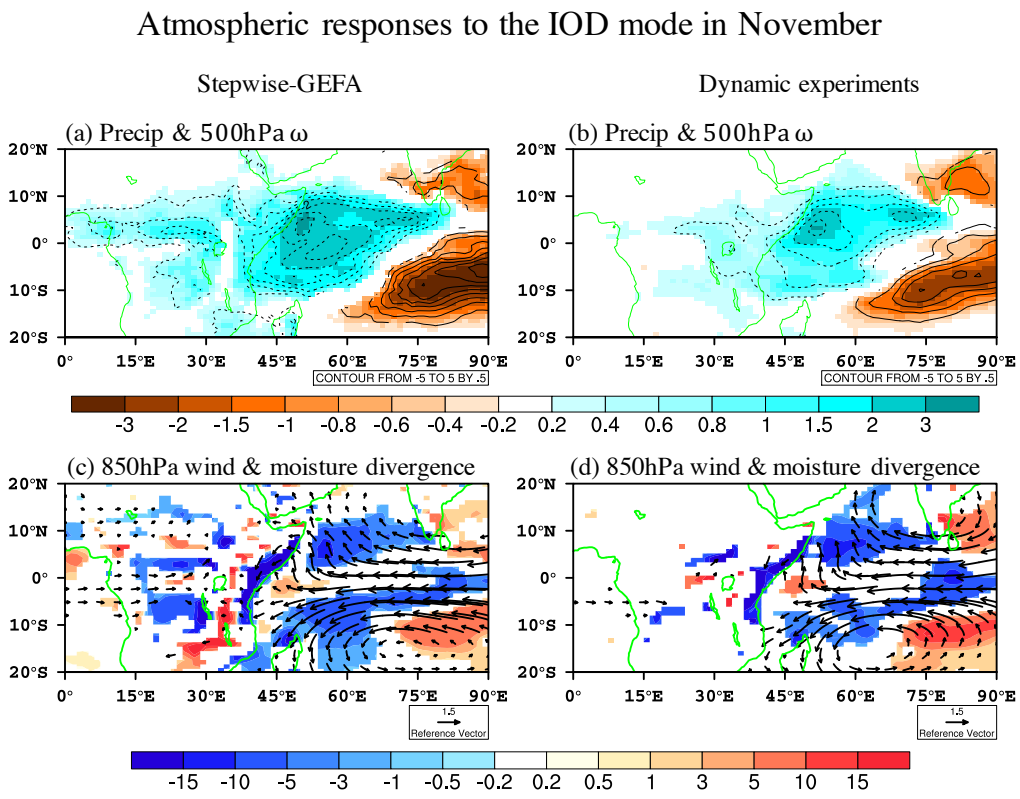


**Figure 22** The same as Figure 21, except for November. The TI1 forcing resembles Figure 6a.

### c. TA1

In CESM, both statistical and dynamical assessments indicate that the Atlantic Niño mode mainly regulates rainfall over the WAM region during the pre- and peak-monsoon seasons,

specifically reducing rainfall in June, while enhancing rainfall in July-August. In June (Figure 23), the climatological location of the Atlantic ITCZ, based on outgoing longwave radiation (OLR), is 4˚N in CESM, and the positive SST anomalies associated with the positive Atlantic Niño mode peak along the southern flank of the climatological ITCZ. According to both SGEFA and the dynamical experiments, the anomalously warm eastern tropical Atlantic Ocean favors a southward shift of the ITCZ, resulting in anomalous low-level moisture convergence near the equator, favoring greater precipitation to the south of the Gulf of Guinea and extending westward into the tropical Atlantic Ocean. Accompanying the southward-shifted ITCZ, anomalous descending motion emerges north of the band of anomalous ascent. A belt of anomalous moisture divergence is generated from the Gulf of Guinea into the tropical Atlantic, which results in reduced precipitation over the WAM region. In August (similar to July) (Figure 24), the climatological location of the Atlantic ITCZ is around 10˚N in CESM, with positive TA1 characterized by anomalously warm waters to the south and cold waters to the north of the climatological ITCZ. The warm eastern tropical Atlantic Ocean favors a southward shift of the ITCZ and its associated ascent. However, distinct from June, the belt of the anomalous moisture convergence and corresponding ascending motion is just over the WAM region, and the anomalous descending motion is confined to the north, over the tropical Atlantic Ocean. The moisture convergence over the WAM region strengthens the West African monsoon flow and enhances precipitation.

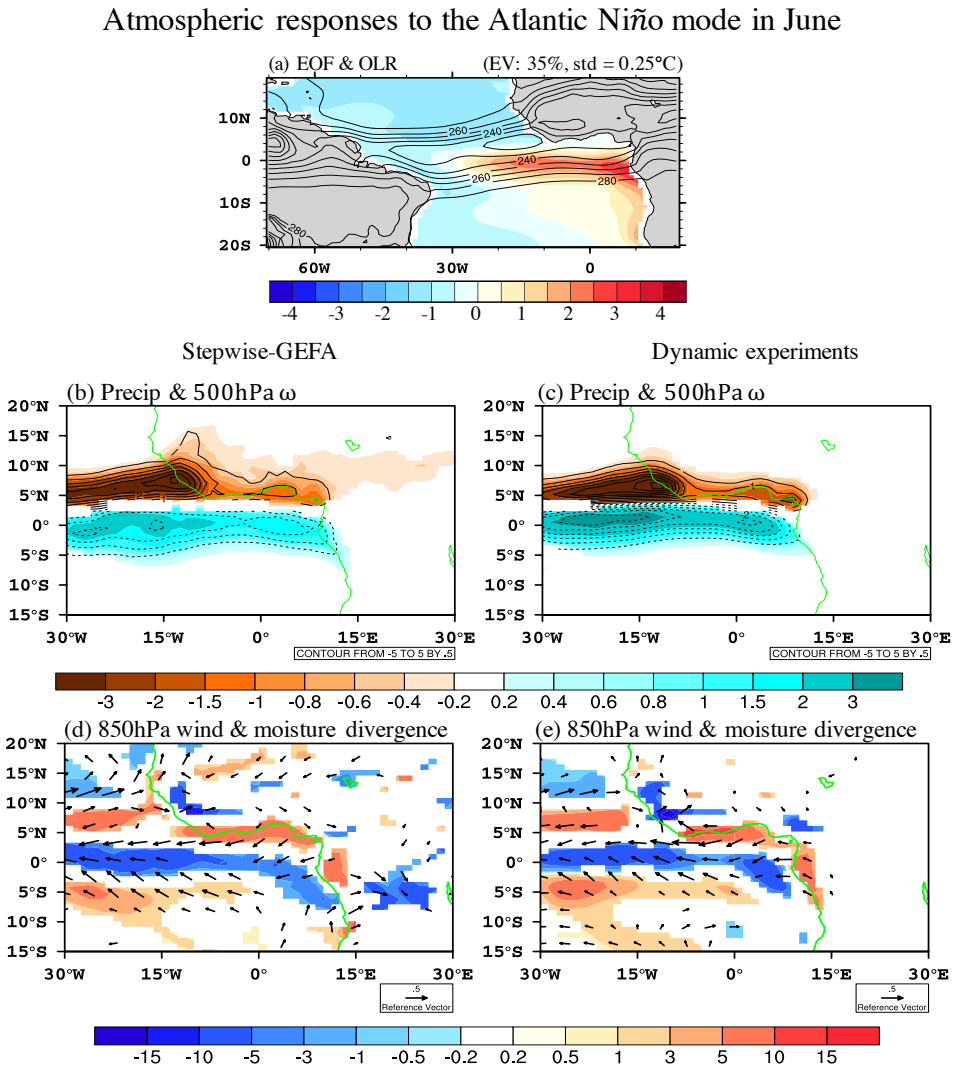


**Figure 23** (a) Climatological OLR (contour, unit: W $m^{-2}$) and leanding EOF of SST over the tropical Atlantic Ocean based on 300-year CTRL for June in CESM. The magnitude of the spatial pattern is normalized with a standard deviation of 1℃. The explainced variance (EV) and standard deviation (std) of corresponding PCs are indicated in parentheses. Precipitation (shading, unit: $mm\ day^{-1}\ \sigma_{PC}^{-1}$) and 500-hPa vertical motion (contour, unit: Pa $s^{-1}\ \sigma_{PC}^{-1}$) response to TA1 forcing using (b) stepwise-GEFA and (c) dynamical experiments during June. 850-hPa moisture divergence (shading, unit: $10^{-6}\ kg\ m\ kg^{-1}\ s^{-1}\ \sigma_{PC}^{-1}$) and 850-hPa wind response (vector, $m\ s^{-1}\ \sigma_{PC}^{-1}$) for TA1 forcing using (d) SGEFA and (e) dynamical experiments. Only statistical significant response ($p<0.1$) fields are shown.

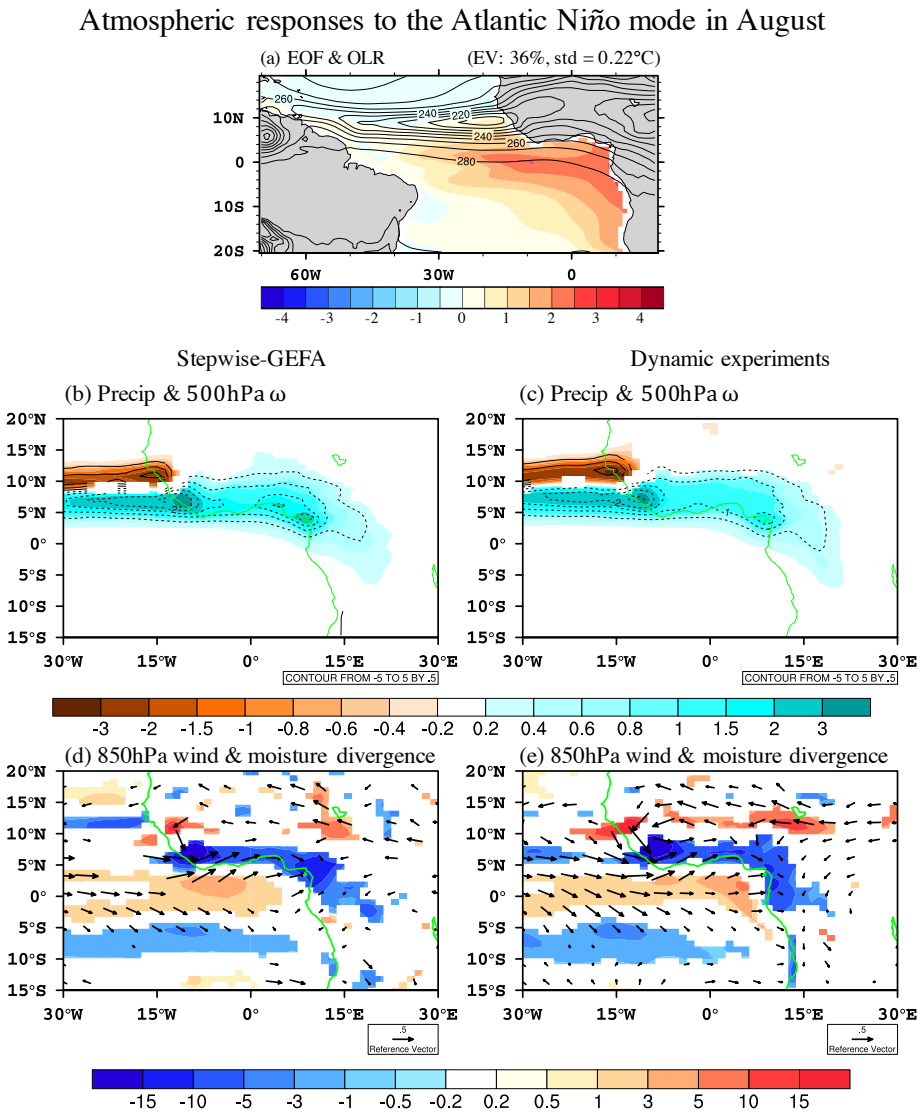


**Figure 24** Same as Figure 23, but for August.

### 3.2. GEFA's capability in capturing terrestrial impacts on regional climate

Unlike oceanic impacts, which usually exerts significant remote impacts on the atmosphere through teleconnections as demonstrated in the previous section, the impacts of terrestrial forcings are mostly local in CESM. For example, in the dynamical experiments with modified

Sahel LAI, an annual average of 97% of significant responses in ET over land, 89% in latent heat flux over land and ocean, 88% in sensible heat flux, 91% in surface air temperature, and 81% in precipitation occur within the Sahel box. For FGEFA (SGEFA), these numbers are 81% (89%), 79% (92%), 74% (90%), 72% (87%), and 69% (84%), respectively. The responses to WAM LAI anomalies are also largely local in both statistical and dynamical assessments, with smaller percentage of significant responses occurring within the WAM box, which is likely due to the less pronounced local impacts of WAM vegetation compared with Sahel (discussed later). Since most of the response to LAI anomalies occurs locally, the evaluation of statistical GEFA mainly focuses on the local atmospheric responses to LAI anomalies across the Sahel or WAM region.

### 3.2.1. Stepwise GEFA versus full GEFA in capturing the seasonal cycle of local responses to LAI anomalies

In this section, SGEFA and FGEFA are compared in terms of their consistency with $EXP_{LAI}$ and $EXP_{SOIL}$ in the seasonal cycle of the assessed local atmospheric responses to LAI anomalies in the Sahel or WAM region. For each focal region, in the months when $EXP_{SOIL}$ is conducted, the dynamical responses are extracted from $EXP_{SOIL}$, or otherwise from $EXP_{LAI}$. The comparison involves three metrics: temporal correlation (N = 12 months), signal-to-noise ratio (SNR), and sign consistency, which reflects the general consistency, magnitude consistency, and sign consistency, respectively. SNR is calculated as the ratio between the mean absolute response in the dynamical experiments and the root-mean-square-difference between responses from GEFA and the dynamical experiments. Sign consistency is calculated as the percentage of 12 calendar months in which the signs of estimated responses by GEFA and dynamical experiments are the same (both significantly positive, significantly negative, or insignificant). With this rigorous comparison between SGEFA and FGEFA, the expected benefit of SGEFA is quantified.

SGEFA generally performs better than FGEFA in capturing the sign and magnitude of local atmospheric responses to LAI anomalies across the Sahel and WAM region (Figures 25-26). SGEFA yields higher temporal correlations, SNRs, and sign consistency than FGEFA, in terms of estimating the seasonal cycle of all atmospheric variables in response to Sahel LAI anomalies, when compared with the dynamical experiments. With 300 years of data, the seasonal cycle of the local response estimated by SGEFA is moderately to highly correlated (N = 12 months) with that from the dynamical experiments, with a temporal correlation of 0.90 for ET, 0.81 for 2-m air temperature, 0.89 for planetary boundary layer height (PBLH), 0.68 for 2-m specific humidity, 0.74 for 10-m wind, and 0.53 for precipitation, compared with weaker correlations of 0.62, 0.71, -0.10, 0.03, 0.12, 0.48, respectively, from FGEFA. In terms of the SNR, the benefits of SGEFA over FGEFA are even more pronounced. With 300 years of data, the SNRs with dynamical experiments in terms of the estimated seasonal cycle of local Sahel responses according to SGEFA are 3.3, 2.7, 3.0, 1.4, 2.0, and 1.4 for ET, 2-m air temperature, PBLH, 2-m specific humidity, 10-m wind, and precipitation, respectively, compared with only 2.0, 2.4, 0.5,0.1, 0.7, and 0.4 according to FGEFA. SGEFA also boasts a higher sign consistency than FGEFA with the dynamical experiments, although the benefit is not as pronounced as with the SNR. This indicates that SGEFA offers a greater advantage over FGEFA in terms of estimating the magnitude of responses rather than the sign of the responses. In estimating responses to LAI anomalies in the WAM region, SGEFA generally outperforms FGEFA as well, but the advantage is less pronounced than with the Sahel, especially in terms of the sign consistency, probably because the atmospheric responses to the WAM LAI anomalies are largely insignificant and small in magnitude in CESM (Figure 28).

Among the different response variables considered here, both FGEFA and SGEFA obtain the highest accuracy in estimating the ET response, which is directly affected by LAI and soil moisture. The precipitation response is one of the most challenging variables to estimate with GEFA, especially in terms of the response magnitude, likely because precipitation is indirectly affected by vegetation and soil moisture anomalies through multiple competing mechanisms (Figure 27, discussed later). Partly due to the competing vegetation-rainfall feedback mechanisms, the precipitation responses are generally small in magnitude according to both GEFA and dynamical experiments, compared with ET, thereby leading to low SNRs and temporal correlations of the response seasonal cycle.

The improvement of stepwise GEFA compared to traditional GEFA is mainly due to the size reduction of the forcing matrix, which leads to an increase in the effective degrees of freedom of the feedback estimation. For example, the total number of forcings in the traditional GEFA forcing matrix varies from 29 to 41 by month, and is dramatically reduced to 5 to 11 after selection by stepwise GEFA for Sahel precipitation. Among the selected forcings for Sahel precipitation in stepwise GEFA, 30%-100% (varying by month) are statistically significant ($p<0.1$) based on the Monte Carlo bootstrap test, compared with 10%-17% among the full list of forcings in traditional GEFA.

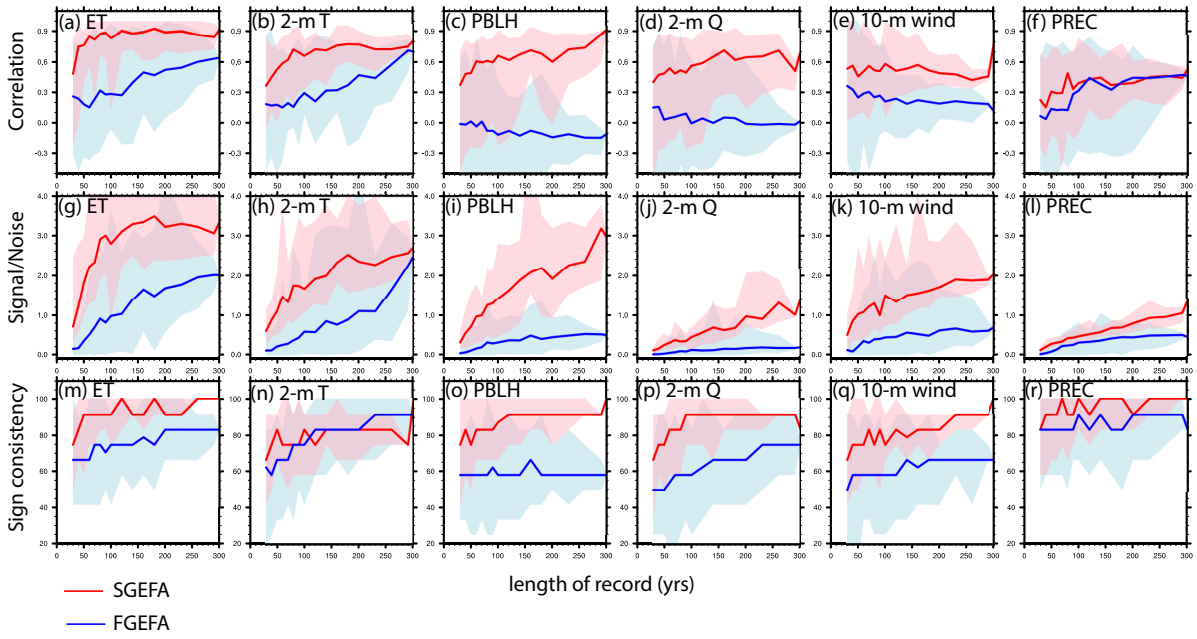


**Figure 25** Consistency between statistical GEFA (blue: FGEFA, red: SGEFA) and dynamical experiments (January, March, April, and December from $EXP_{SOIL}$; other months from $EXP_{LAI}$) in the seasonal cycle (N = 12 months) of local area-average responses to LAI anomalies across the Sahel, in terms of (a-f) temporal correlation, (g-l) SNR, and (m-r) sign consistency, as a function of length of record (number of years) on the x-axis. Comparisons are shown for variables with significant responses to Sahel LAI in the dynamical experiment for 10+ months, including ET, 2-m air temperature (2-m T), PBLH, 2-m specific humidity (2-m Q), 10-m wind speed (10-m wind), and precipitation (PREC), in the order of relative consistency between GEFA and dynamical experiments. Lines (shading) represent the median (minimum and maximum) across different periods with which GEFA is applied.

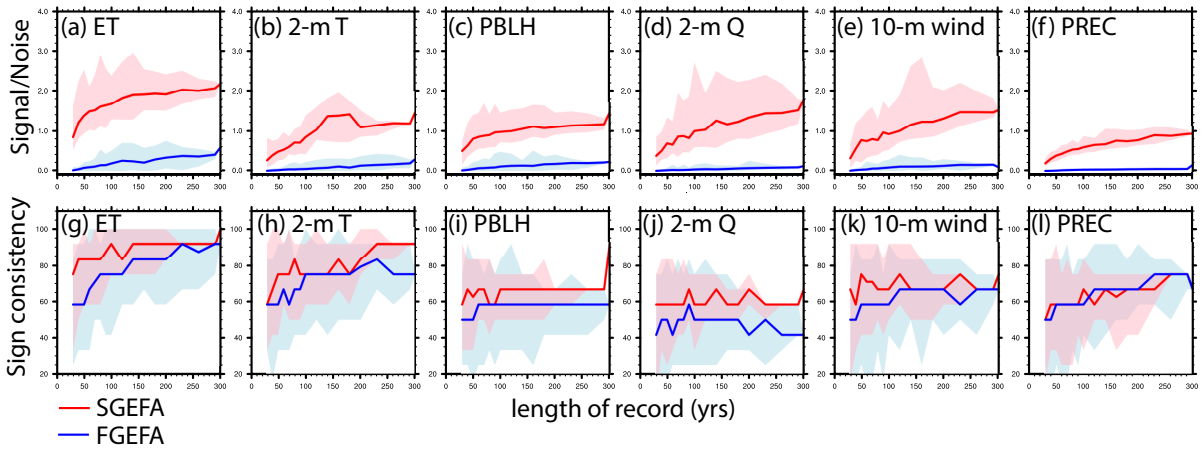


**Figure 26** Similar to Figure 25, but for the WAM region. Temporal correlation of the response seasonal cycle is not included in the comparison, because significant response to LAI and soil moisture anomalies across the WAM region mainly occur in just two months, February and March.

Both SGEFA and FGEFA are sensitive to record length, although with significantly shorter minimum data records needed for SGEFA to obtain reliable feedback estimates. For the responses to Sahel LAI anomalies among most variables, when the data record length is short (<100 years, which is the typical length of observations), FGEFA struggles to capture the terrestrial impacts on the local atmosphere, while SGEFA achieves reasonable sign consistency (>70%) and temporal correlation (typically 0.6-0.8, N = 12) for most response variables, compared with the dynamical assessment. In order to obtain a reliable estimate of the seasonal cycle of local responses with a temporal correlation of 0.7 with the dynamical experiments for Sahel, SGEFA requires approximately 50, 70, 90, 140, 280, and >300 years of data for ET, 2-m air temperature, PBLH, 2-m specific humidity, 10-m wind, and precipitation, respectively, compared with at least 290 years for all variables with FGEFA. Again, precipitation requires relatively long data records to achieve an acceptable temporal correlation between the response

seasonal cycle from the statistical GEFA assessment and dynamical assessment, mainly because of its small response magnitudes. Indeed, if the goal is to achieve a sign consistency of 90% between SGEFA and dynamical experiments, only 50 years of precipitation data are needed, which is similar to that of ET. Given the significant improvement of SGEFA compared with FGEFA, the subsequent detailed evaluation against the dynamical experiments focuses only on SGEFA.

GEFA-based atmospheric responses to LAI anomalies are generally consistent with both the $EXP_{LAI}$ and $EXP_{SOIL}$ ensembles, with higher consistency in both sign and magnitude with the $EXP_{SOIL}$ ensemble than the $EXP_{LAI}$ ensemble during months with strong soil moisture-LAI coupling in the CTRL. For example, in the Sahel, LAI anomalies alone cause a response of +0.40 mm $day^{-1}$ $LAI^{-1}$ in local ET during the dry season (averaged among December, January, March, and April when $EXP_{SOIL}$ experiments are performed), while the coupled soil moisture-LAI anomalies support a greater response of +0.85 mm $day^{-1}$ $LAI^{-1}$ in local ET, compared with +0.74 mm $day^{-1}$ $LAI^{-1}$ from SGEFA (Figure 27 and Table 6). In the WAM region, GEFA is more consistent with $EXP_{SOIL}$ than $EXP_{LAI}$ in both the sign and magnitude of the estimated responses (Figure 28 and Table 6). For example, a positive LAI anomaly alone in $EXP_{LAI}$ causes a slight drying during March in the WAM region (-0.09 mm $day^{-1}$ $LAI^{-1}$), while synchronous positive anomalies in LAI and soil moisture in $EXP_{SOIL}$ cause a weak increase in precipitation (+0.15 mm $day^{-1}$ $LAI^{-1}$), which is successfully captured by SGEFA (+0.08 mm $day^{-1}$ $LAI^{-1}$) in both sign and magnitude. The higher consistency between GEFA and $EXP_{SOIL}$, in terms of the estimated response in most atmospheric variables for both study regions, indicates that GEFA captures the combined impact of coupled LAI and soil moisture anomalies on atmospheric conditions.

**Table 6** Local responses in select variables to LAI anomalies across the Sahel and WAM region assessed by $EXP_{LAI}$, $EXP_{SOIL}$, and SGEFA. The responses are averaged during the winter-spring dry season, namely December, January, March, and April for the Sahel and February and March for the WAM region. Values in italics/bold indicate whether $EXP_{LAI}$ or $EXP_{SOIL}$ is more consistent with SGEFA.

| | Sahel | | | WAM | | |
|---|---|---|---|---|---|---|
| | $EXP_{LAI}$ | $EXP_{SOIL}$ | SGEFA | $EXP_{LAI}$ | $EXP_{SOIL}$ | SGEFA |
| ET (mm $day^{-1}$ $LAI^{-1}$) | 0.40 | ***0.85*** | 0.74 | 0.11 | ***0.27*** | 0.20 |
| Latent heat flux (W $m^{-2}$ $LAI^{-1}$) | 11.57 | ***24.42*** | 20.17 | 2.85 | ***7.17*** | 5.45 |
| Sensible heat flux (W $m^{-2}$ $LAI^{-1}$) | -7.12 | ***-16.02*** | -16.14 | -2.69 | ***-9.11*** | -6.12 |
| 2-m air temperature (°C $LAI^{-1}$) | -0.60 | ***-1.13*** | -0.97 | -0.1 | ***-0.19*** | -0.16 |
| PBLH (m $LAI^{-1}$) | -37.54 | ***-77.18*** | -58.58 | ***-10.59*** | -36.82 | -22.22 |
| 850-hPa vertical motion (Pa $s^{-1}$ $LAI^{-1}$) | 0.009 | ***0.018*** | 0.017 | 0.004 | ***0.005*** | 0.005 |
| OLR (W $m^{-2}$ $LAI^{-1}$) | 0.94 | ***-0.75*** | -0.89 | 0.87 | ***-2.01*** | -2.27 |
| 10-m wind speed (m $s^{-1}$ $LAI^{-1}$) | -0.06 | ***-0.13*** | -0.15 | -0.02 | ***-0.05*** | -0.04 |
| 2-m specific humidity (g $kg^{-1}$ $LAI^{-1}$) | 0.34 | ***0.93*** | 0.75 | 0.08 | ***0.45*** | 0.50 |
| Precipitation (mm $day^{-1}$ $LAI^{-1}$) | -0.11 | ***-0.16*** | -0.17 | -0.09 | ***0.15*** | 0.08 |

### 3.2.2. Terrestrial impacts over Sahel and WAM in CESM

Negative vegetation-precipitation feedbacks across the Sahel in spring-summer are consistently identified in CESM according to both the dynamical experiments ($EXP_{LAI}$ and $EXP_{SOIL}$) and statistical assessments (Figure 27). Positive Sahel LAI anomalies cause enhanced ET, mainly due to increased canopy transpiration, leading to increased latent heat flux to the atmosphere and decreased Bowen ratio, which causes surface and low-level cooling. The low-level cooling and increased atmospheric stability, indicated by decreased PBLH and anomalous descending motion in the low-mid troposphere, inhibit deep convection (as indicated by an increase in OLR), which causes reduced rainfall, particularly in the pre- and early-monsoon season. The low-level cooling also causes a decrease in land-ocean temperature contrast, which delays and weakens the monsoon circulation, contributing to the drying response to positive LAI anomalies. The increase in atmospheric moisture associated with greater ET is limited to the

winter-spring dry season and near the surface and does not support an enhancement in precipitation in CESM. Decreased surface wind speed is associated with increased surface roughness, most notable during the dry season. A modest decline in surface albedo associated with increased greenness is present in the dynamical experiments (not shown), but the response is not enough to trigger surface warming and increased convergence, as proposed in the classic theory by Charney (1975). Indeed, the negative relationship between LAI and surface albedo in CTRL is mostly confined to the southern portion of the Sahel, where shrubs are more abundant than grasses. In summary, enhanced atmospheric stability associated with an increase in Sahel LAI and soil moisture dominates over the moisture, momentum, and albedo feedbacks, leading to a negative vegetation-precipitation feedback across the Sahel in CESM, as revealed consistently by both SGEFA and the dynamical experiments.

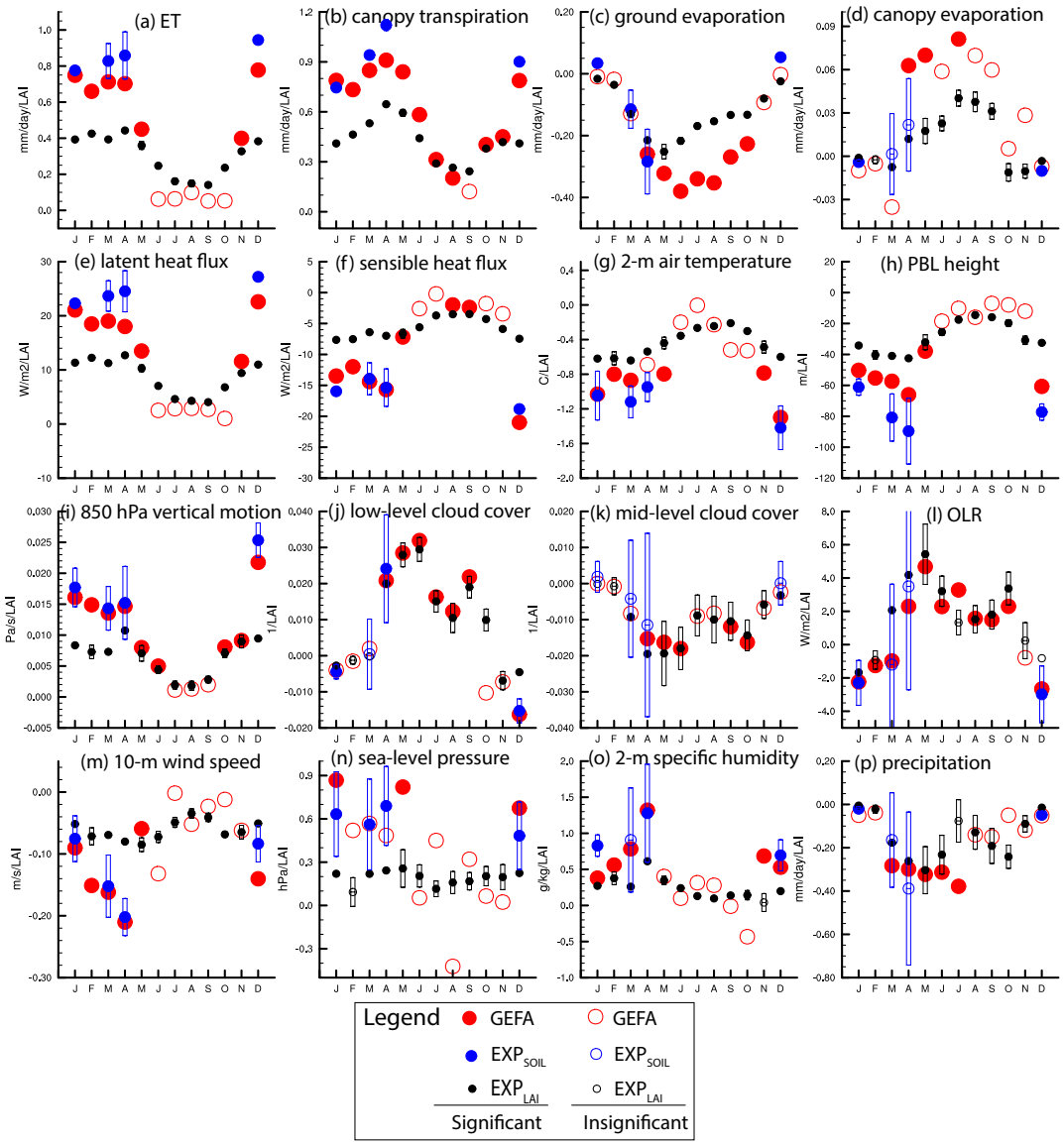


**Figure 27** Seasonal cycle of local area-average responses to LAI anomalies across the Sahel in select variables assessed from $EXP_{LAI}$ (black circles), $EXP_{SOIL}$ (blue circles), and SGEFA (red circles) in CESM. Filled circles indicate statistically significant ($p < 0.1$) responses, according to the Student's t-test for the dynamical experiments and Monte Carlo bootstrap test for GEFA. The boxes represent 90% confidence interval of the responses from $EXP_{soil}$ (blue) and $EXP_{LAI}$ (black).

In contrast to the Sahel, weak positive vegetation-precipitation feedbacks are present across the WAM region during the dry season in CESM, as consistently captured by both the dynamical experiments and statistical assessments (Figure 28). This positive feedback is largely due to the moisture recycling mechanism. Positive LAI anomalies and associated wet soil favor enhanced ET, largely attributed to increased canopy transpiration, which leads to greater atmospheric moisture, convective activity, and precipitation. Increased atmospheric stability is also present in response to the positive anomalies in LAI and soil moisture, as indicated by surface and low-level cooling, decreased PBLH, and anomalous low-level descending motion. However, the impact of enhanced moisture recycling dominates over the effect of increased atmospheric stability in terms of the WAM rainfall response to LAI anomalies. Indeed, the positive vegetation-rainfall feedbacks are mostly attributed to the soil moisture anomalies, since $EXP_{LAI}$, in contrast to $EXP_{SOIL}$ and SGEFA, indicates a weak negative precipitation response to positive LAI anomalies in March. In other seasons, the vegetation feedbacks in the WAM region are trivial on most atmospheric variables.

Positive simulated vegetation-rainfall feedbacks in the WAM region, in contrast to the Sahel, is likely due to the extensive tree cover in the former region, which induces a greater reduction in surface albedo in response to positive anomalies in coupled soil moisture-LAI and associated surface warming, thereby inhibiting the increase in atmospheric stabilization. The differing vegetation feedbacks induced by trees and grasses are also noted by the modeling studies of Liu et al. (2010) and Notaro et al. (2011).

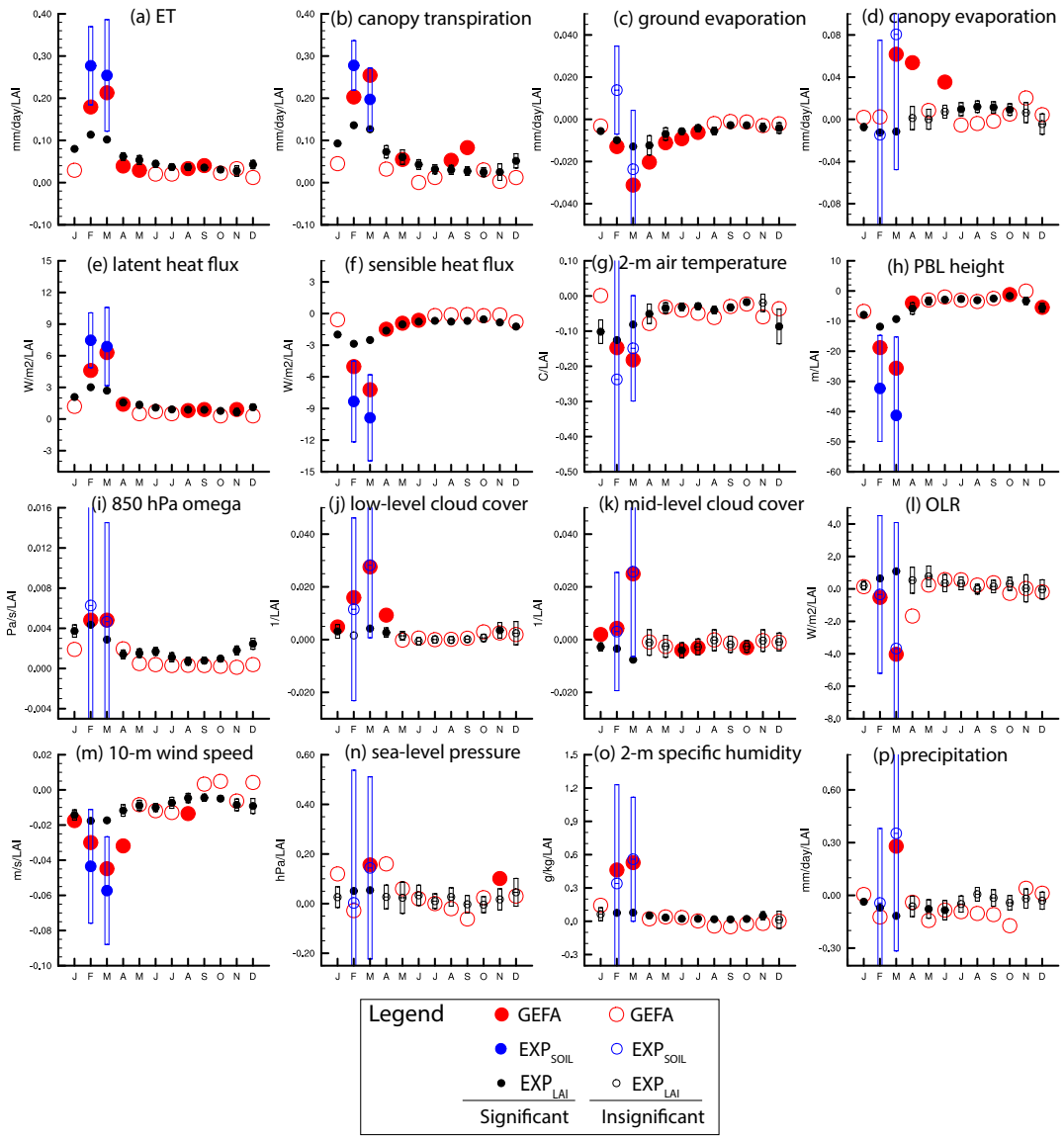


**Figure 28** Similar to Figure 27, but for the WAM region.

Remote impacts of Sahel vegetation variability on tropical Atlantic rainfall are identified in CESM, but limited to December and January, according to both the statistical and dynamical assessments (Figure 29). In December, positive LAI anomalies and associated wet soil in the Sahel cause local low-level cooling and anomalous subsidence. The cooling leads to an enhanced,

and slightly southward-displaced, ITCZ. In addition, the response of subsidence across the Sahel initiates an anomalous northerly outflow and enhances low-level convergence over the tropical eastern Atlantic Ocean. Again, GEFA captures the combined soil moisture-LAI effects, as the GEFA-based response patterns and magnitudes agree better with those from $EXP_{SOIL}$ than $EXP_{LAI}$. The seasonality of these remote impacts is largely determined by both the magnitude of local feedbacks and the mean climatological position of the ITCZ in CESM. In December and January, positive LAI anomalies and associated wet soil in the Sahel generate substantial local stabilization (Figure 27). Meanwhile, the climatological mean ITCZ is at its southernmost position, far away from the Sahel region. Therefore, the remote impacts of the Sahel vegetation on rainfall over the tropical Atlantic Ocean are maximized during the winter months.

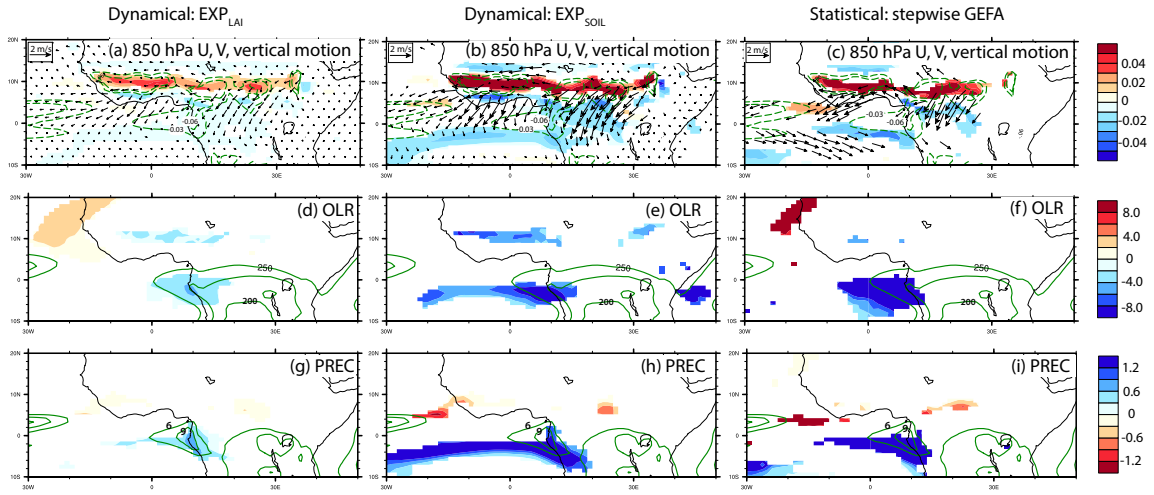


**Figure 29** Dynamically and statistically assessed responses to LAI anomalies across the Sahel in December. Response fields include (a-c) 850-hPa wind (m $s^{-1}$ $LAI^{-1}$, vector) and vertical motion (Pa $s^{-1}$ $LAI^{-1}$, shading), and (d-f) OLR (W $m^{-2}$ $LAI^{-1}$, shading), and (g-i) and precipitation (PREC) (mm $d^{-1}$ $LAI^{-1}$, shading). Climatology of December 850-hPa vertical motion (Pa $s^{-1}$), OLR (Pa $s^{-1}$), and precipitation (mm $d^{-1}$) in CTRL are shown in (a-c), (d-f), and (g-i), respectively, by green contours. Dynamical responses are assessed from the (a,d,g) $EXP_{LAI}$ and (b,e,h) $EXP_{SOIL}$ ensembles. Statistical responses are assessed based

on (c,f,i) stepwise GEFA. Only statistically significant responses are shown based on the Student's t-test and Monte Carlo test ($p<0.1$) for dynamical and statistical assessments, respectively.

### 3.3. Discussion: Guidance for future application of GEFA

Although GEFA's capability in extracting the atmospheric responses to oceanic and terrestrial drivers across both North America (Wang et al. 2013, 2014) and North Africa has been demonstrated, further applications of the method to other study regions or datasets should consider the following recommendations. First, it is necessary to the check the covariability among forcings. As demonstrated in the TP1 and TI1 case, if the temporal correlation between two forcings in the forcing matrix is extremely high (>0.8), indicating that these two forcings are not physically independent, it is hard to separate individual impacts from these two forcings. In this case, one might consider combining the highly-correlated forcings into a single driver and carefully interpreting the results. As an example, since LAI and soil moisture are so tightly coupled during the dry season across either the Sahel and WAM region in CESM, GEFA cannot separate their individual impacts. Therefore, the GEFA-based feedbacks are interpreted as the combined impacts of LAI and soil moisture in the current study.

Second, in order to clearly separate the impacts from correlated forcings, all forcings that are closely related to the forcings of interest must be included in the forcing matrix. For example, in the present study, if WAM LAI, which is correlated with Sahel LAI in most months (temporal correlation of monthly anomalies $\geq 0.28$, $N = 300$), is not included in the forcing matrix, then the responses in ET, air temperature, atmospheric moisture, and precipitation over the WAM region to Sahel LAI anomalies will not be correctly captured by GEFA. For example, significant

positive ET responses to positive Sahel LAI anomalies incorrectly appear across the WAM region when WAM LAI is absent from the forcing matrix, but correctly disappear if WAM LAI is included in the forcing matrix. Therefore, when applying GEFA to assess the oceanic and terrestrial controls on regional climate, all of the forcings that are of interest, as well as other forcings that are closely related to them, need to be included in the forcing matrix. In assessing terrestrial impacts, it is recommended to first identify regions that have moderate-to-high correlations with the focal region in monthly LAI anomalies before determining the GEFA forcing matrix. This can be done by correlating the area-average LAI anomalies across the focal region with LAI anomalies at each grid cell across the globe.

Third, beyond additional LAI forcings to be considered in the forcing matrix, the potential impacts of other land surface forcings, e.g. snow and ice cover, on the assessed vegetation feedbacks over the high-latitude regions need to be explored in future GEFA validation and application studies.

Although GEFA successfully predicts both the sign and magnitude of the terrestrial feedbacks across North Africa in the model, the statistical significance is usually weaker in GEFA than in the dynamical experiments, especially $EXP_{LAI}$. The application of different significance tests, namely the Monte Carlo bootstrap method for GEFA versus the Student's t-test for the dynamical experiments, is not the likely cause for this discrepancy, since the alternative application of the Monte Carlo bootstrap method to the dynamical experiments yields nearly the same significance levels as indicated by the Student's t-test. Likely reasons for the weaker significance in GEFA-based estimates include: 1) larger magnitudes of imposed terrestrial forcings in the dynamical assessment than in the statistical assessment, given that extreme anomalies ($5^{th}$ and $95^{th}$ percentiles) are applied in the dynamical experiments for

computational efficiency, while GEFA considers the full spectrum of forcing magnitudes from CTRL to estimate the feedbacks, and 2) insufficient length of CTRL (300 years) for application of GEFA to obtain significant responses, especially for those variables, e.g. sea-level pressure, that are less directly affected by terrestrial anomalies.

# Chapter 4 Observed oceanic drivers of North African climate

After successful validation of GEFA, especially the demonstration of benefits introduced by stepwise GEFA (Chapter 3), stepwise GEFA is now applied to multiple observational, reanalysis and remote sensing data to assess the observed oceanic (Chapter 4) and terrestrial (Chapter 5) impacts on North African climate. This chapter is motivated by the inconsistent conclusions among previous modeling studies regarding the dominant oceanic drivers on North African rainfall variability. The main objective of the work presented in this chapter is to identify the observed important oceanic drivers and the underlying mechanisms for their impacts on North African rainfall.

## 4.1. Important observed oceanic drivers of North African climate

Tropical oceans exert dominant impacts on the observed variability in North African precipitation and air temperature (Figure 30). The primary oceanic drivers of variability in the observed precipitation across the Sahel and WAM region are the tropical Pacific and tropical Atlantic Oceans (Figure 30), consistent with past modeling studies (Folland et al. 1986; Giannini et al. 2003, 2005; Lu and Delworth 2005; Hoerling et al. 2006). The Mediterranean Sea also affects Sahel rainfall variability in the monsoon season, consistent with previous model-based findings (Rowell et al. 2003; Park et al. 2016). The Indian Ocean, which was identified as another important basin regulating the Sahel and WAM rainfall variability by previous modeling studies (Lu and Delworth 2005), is seemingly not important for either the Sahel and WAM region according to the application of SGEFA to observations. The inconsistency between the model-based and GEFA-based findings regarding the impacts of Indian Ocean on the Sahel's monsoon rainfall is likely due to the high covariability between the tropical Indian and tropical

Pacific SSTs (Yang et al. 2007). As demonstrated in the GEFA validation section, the Indian Ocean Basin Mode is primarily driven by variability in tropical Pacific SSTs, and therefore the apparent impacts from the Indian Ocean are indeed driven by the tropical Pacific Ocean. For the HOA, both the tropical Indian and tropical Pacific Oceans exert strong controls on observed precipitation variability in almost all seasons, consistent with previous modeling studies (Mason and Goddard 2001; Behera et al. 2005). The impacts on HOA rainfall by Indian Ocean SST variability are greatest during the pre- and early-short rains, while the influence from Pacific Ocean SST variability peaks after the HOA short rains.

The primary oceanic drivers of observed variability in North African air temperature are the tropical Atlantic and tropical Pacific Oceans. The extratropical oceans also influence air temperature over the Sahel and HOA. Air temperature over the Sahel and WAM is more strongly controlled by oceanic forcings than over the HOA. In general, oceanic drivers exert larger impacts on North African air temperature than precipitation.

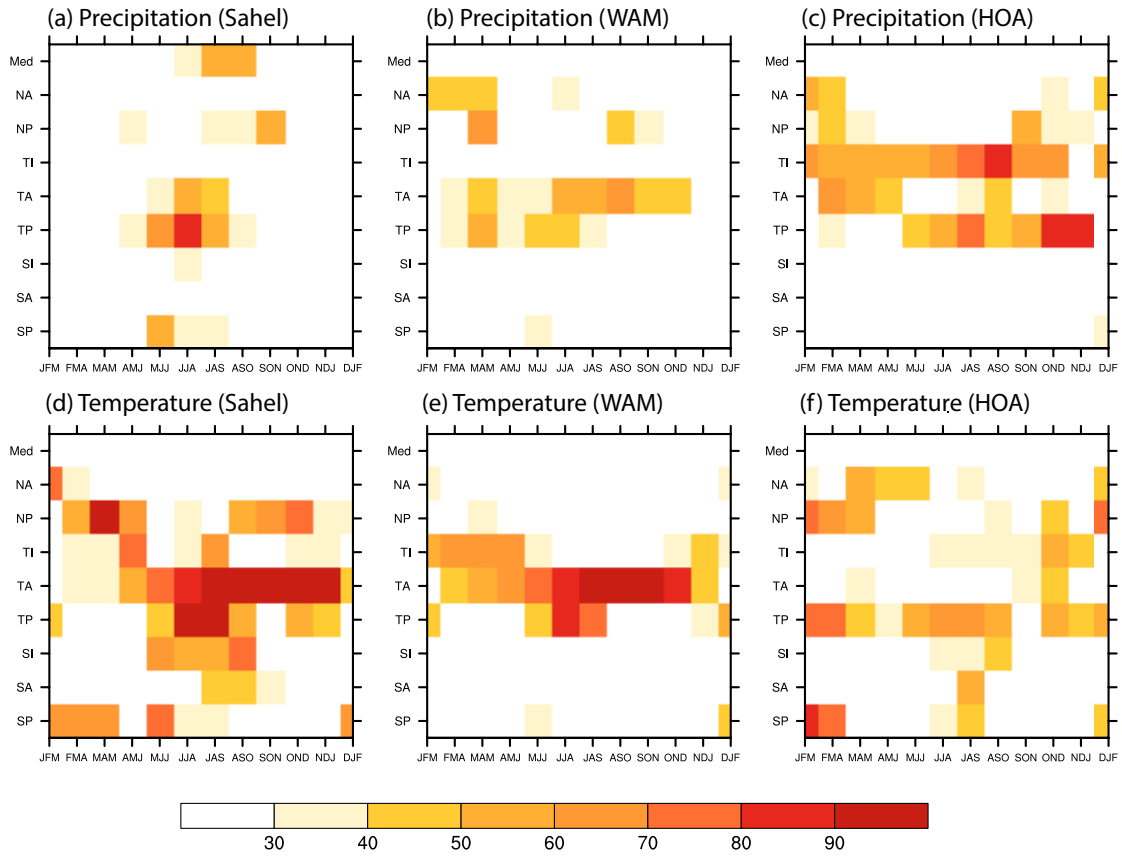


**Figure 30** Multi-dataset weighted-average percentage of area with significant observed responses ($p<0.1$) in (a-c) precipitation and (b-d) surface air temperature across the (a,d) Sahel, (b,e) WAM region, and (c,f) HOA to at least one out of two leading EOFs in each oceanic basin, by season.

## 4.2. Observed impacts of tropical Pacific SSTs on North African climate

A warm tropical eastern Pacific Ocean (El Niño) during the early monsoon season causes anomalous subsidence across the Sahel, as a part of the canonical large-scale see-saw response in pressure to El Niño (Figure 31). The anomalous subsidence leads to diminished oceanic flow from the Atlantic Ocean and associated moisture advection, which reduces observed precipitable water and rainfall over the Sahel. This observed mechanism has been noted by previous modeling studies (Janicot 1998; Giannini et al. 2003, 2005). Outside of the monsoon season, the

impact of tropical Pacific SST variability on Sahel precipitation is minimal, largely due to low climatological rainfall across the Sahel.

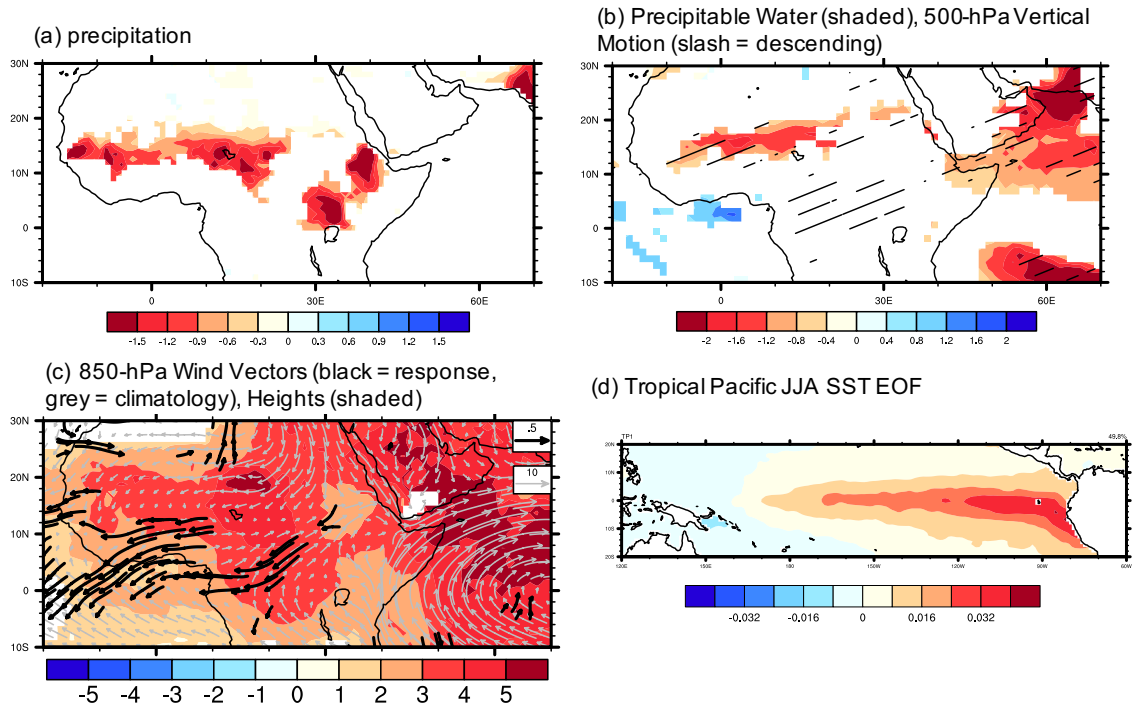


**Figure 31** GEFA-based multi-dataset observational response to tropical Pacific EOF1 in JJA: (a) over-land precipitation (cm mon$^{-1}$ $\sigma_{SST}^{-1}$) (b) precipitable water (shading, g cm$^{-2}$ $\sigma_{SST}^{-1}$) and 500-hPa vertical motion (Pa s$^{-1}$ $\sigma_{SST}^{-1}$, slash indicates descending motion), (c) 850-hPa wind (response in black vectors, climatology in grey vectors, m s$^{-1}$ $\sigma_{SST}^{-1}$) and height (shading, m $\sigma_{SST}^{-1}$). Only significant ($p<0.1$) response fields are shown.

Besides Sahel rainfall, tropical Pacific SSTs also regulate rainfall variability during the short-rains season of the HOA (Figure 32). El Niño favors wet anomalies over the HOA due to anomalous ascending motion and enhanced moisture over the tropical western Indian Ocean. The additional moisture is transported to the HOA by the mean wind, which is enhanced by

anomalous easterlies as induced by the large-scale see-saw response in pressure to El Niño.

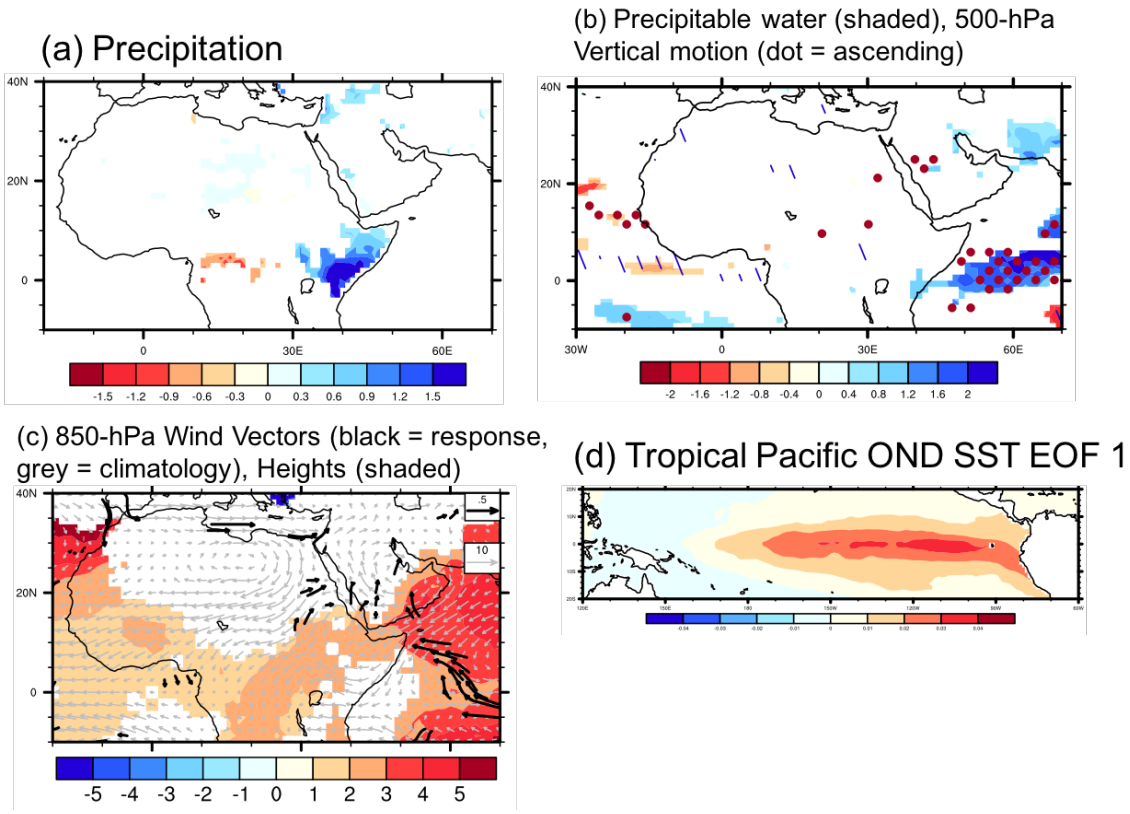


**Figure 32** GEFA-based multi-dataset observational response to tropical Pacific EOF1 in OND (similar to Figure 31).

### 4.3. Observed impacts of tropical Atlantic SSTs on North African climate

Observed tropical Atlantic SSTs mainly affect precipitation across the Sahel and WAM region, through influence on the mean latitudinal position of the ITCZ (Figure 33). An anomalously warm tropical eastern Atlantic Ocean to the south of the mean climatological ITCZ typically induces an observed southward shift of the ITCZ in JJA, characterized by anomalous mid-tropospheric subsidence and increased OLR (not shown) over the Sahel and anomalous ascent and decreased OLR (not shown) over the WAM region. The southward shift of ITCZ in response to positive TA1 leads to enhanced rainfall over the WAM region and reduced Sahel

rainfall, consistent with the model-based findings of Gianini et al. (2003). Anomalous atmospheric moisture over the Gulf of Guinea associated with enhanced evaporation over warm water also contributes to greater precipitation over the WAM region, consistent with the model-based results of Hagos and Cook (2008). The impacts of TA1 are maximize during the monsoon season, when the SST pattern exhibits its greatest anomalies around the Gulf of Guinea.

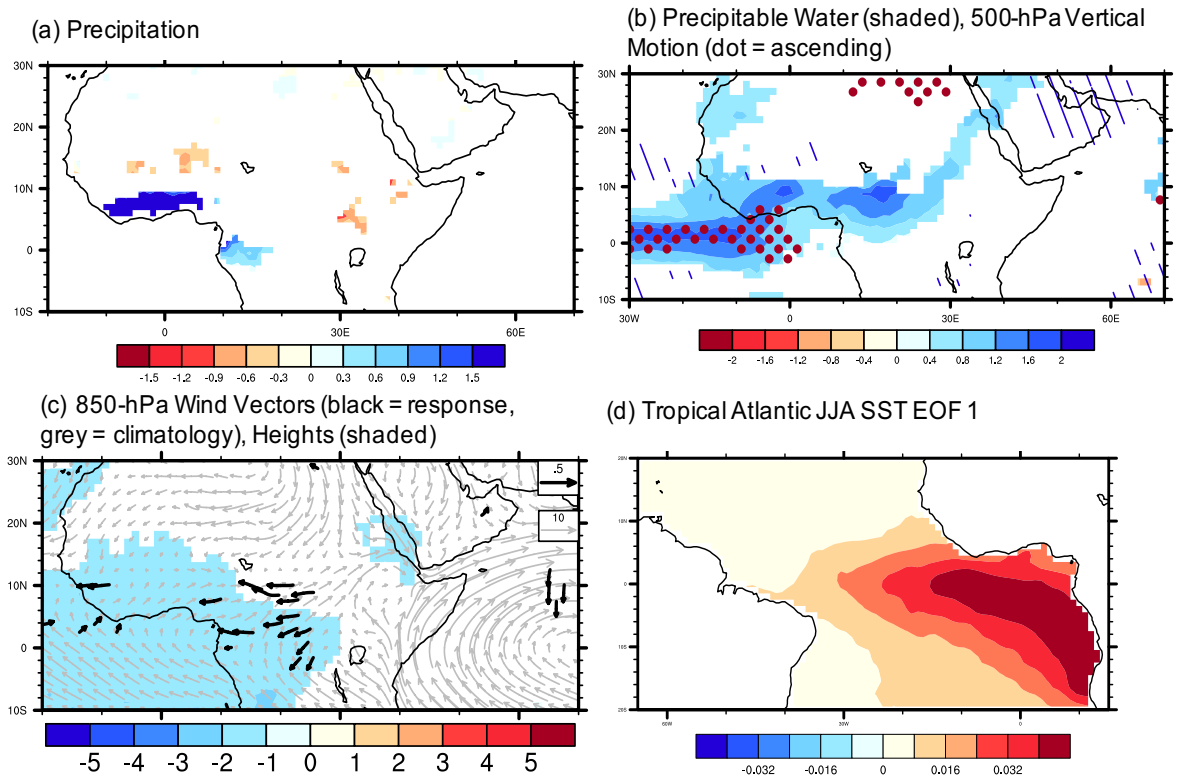


**Figure 33** GEFA-based multi-dataset observational response to tropical Atlantic EOF1 in JJA (similar to Figure 31).

The tropical Atlantic SST meridional mode (TA2), consisting of positive SST anomalies in the south and negative anomalies in the north, usually favors an observed enhancement of rainfall over the WAM region and Congo and reduced rainfall over the Sahel (Figure 34). The north-south dipole anomalies in the SST pattern induces a southward shift of the ITCZ,

consistent with model-based findings (Hoerling et al. 2006; Giannini et al. 2008a,b). The responses in precipitable water and vertical motion peak during the dry season (November-May), when the TA2 pattern exhibits most pronounced meridional SST gradient.

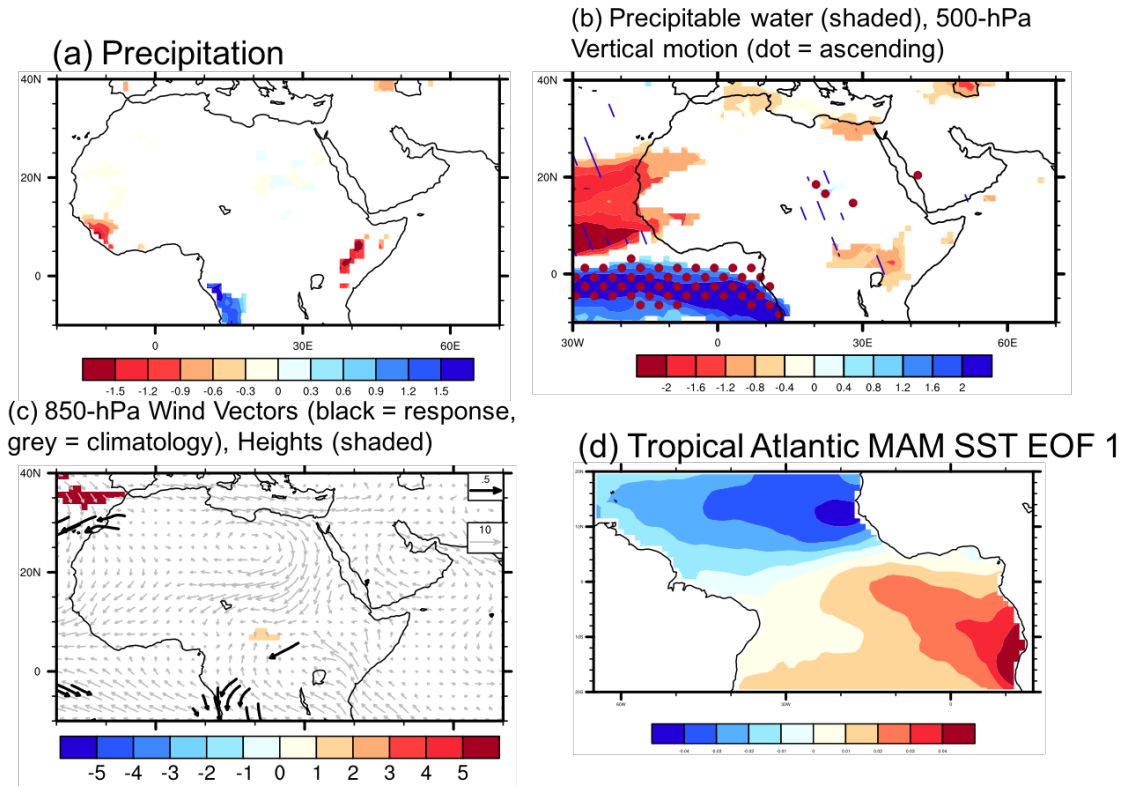


**Figure 34** GEFA-based multi-dataset observational response to tropical Atlantic EOF2 in MAM (similar to Figure 31).

### 4.4. Observed impacts of tropical Indian SSTs on North African climate

In contrast to the modeling study of Lu and Delworth (2005), the tropical Indian Ocean is found to exert minimal observed impact on Sahel and WAM precipitation, due to the lack of significant responses in mid-to-low level divergence and anticyclonic circulation as noted by previous modeling studies (Bader and Latif 2003, 2005; Giannini et al. 2003; Hagos and Cook

2008). Observed precipitation variability across the HOA is primarily controlled by SST anomalies across the tropical Indian and tropical Pacific Oceans, consistent with past modeling studies (Indeje et al. 2000; Mason and Goddard 2001; Hastenrath 2007). During the early short-rains (ASO), a warm tropical western Indian Ocean often triggers enhanced local evaporation and ascending motion over the warm SST region (Figure 35). Anomalous moisture advection by the mean wind leads to greater early short rains across the HOA. Indeed, a warm tropical western Indian, whether existing as part of the monopole mode or dipole mode, always supports greater short rains across the HOA through enhanced evaporation and moisture advection to the HOA (Figure 36). Since the impacts of tropical Indian and tropical Pacific SSTs on the HOA are mainly observed during the boreal autumn and winter, when SST variability between these two basins is sufficiently independent of each other, the individual impacts are physically independent and can be separated by GEFA.

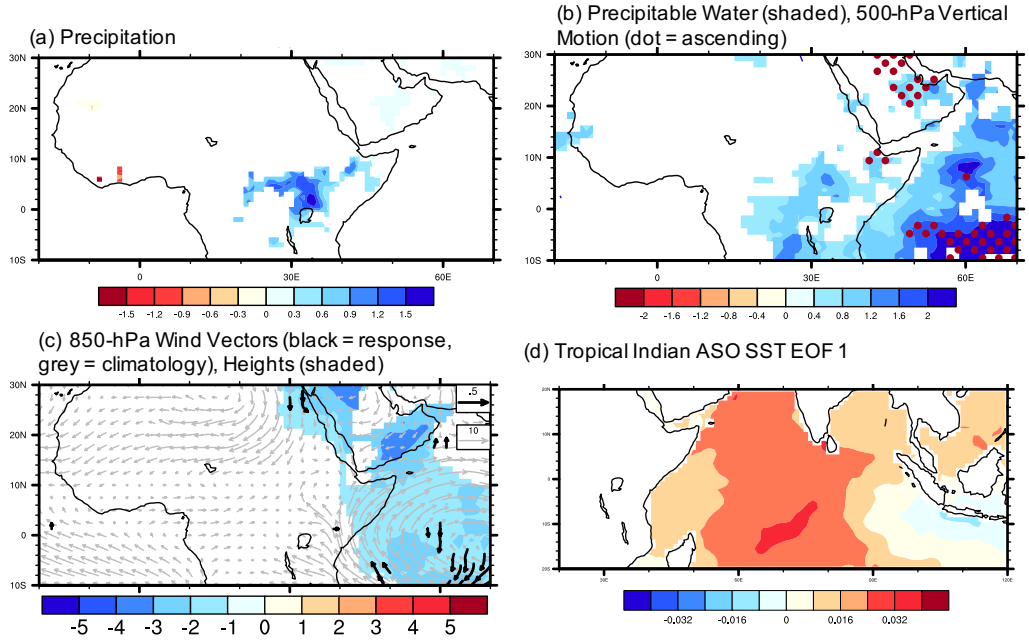


**Figure 35** GEFA-based multi-dataset observational response to tropical Indian EOF1 in ASO (similar to Figure 31).

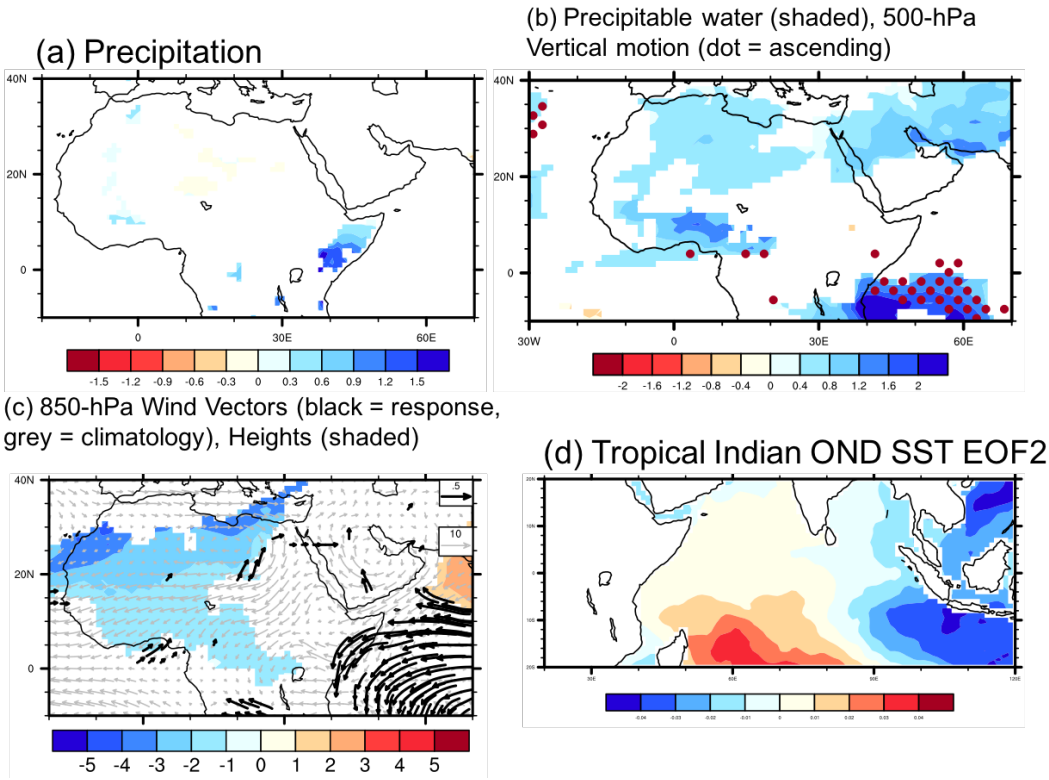


**Figure 36** GEFA-based multi-dataset observational response to tropical Indian EOF2 in OND (similar to Figure 31).

### 4.5 Observed impacts of Mediterranean SST on North African climate

Consistent with previous model-based findings (Rowell et al. 2003; Park et al. 2016), Mediterranean SSTs affect observed rainfall variability across the central and eastern Sahel in the monsoon season (Figure 37). Anomalously high Mediterranean SSTs seemingly support enhanced local evaporation. The additional moisture is transported to the central and eastern Sahel by the climatological northerly winds, which are induced by the climatological Saharan heat low during summer.

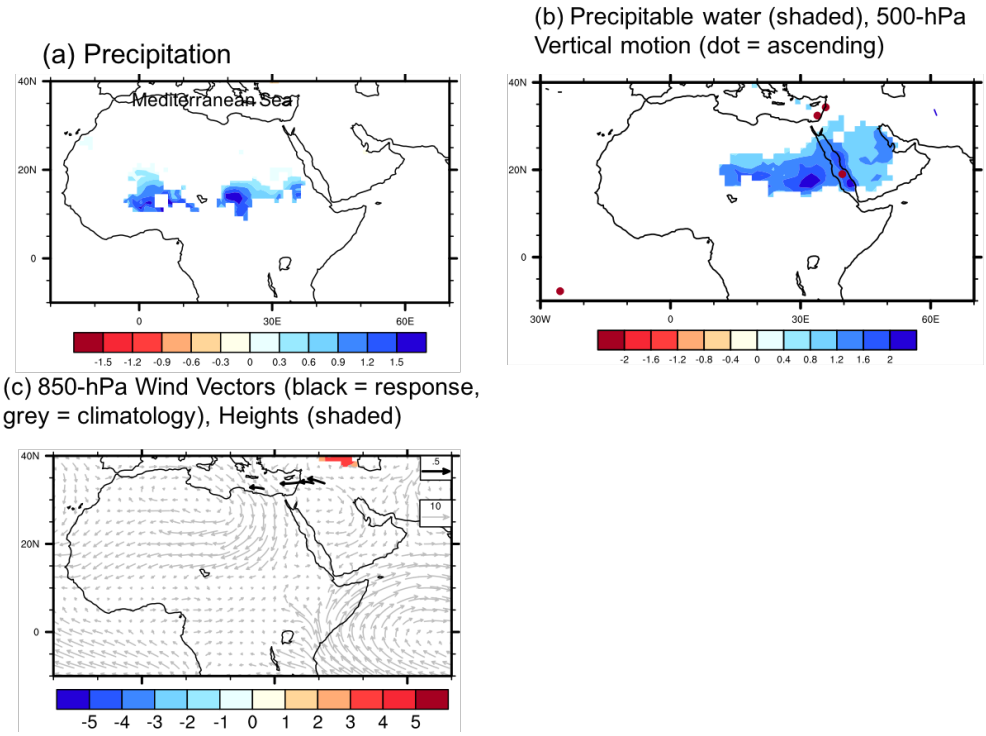


**Figure 37** GEFA-based multi-dataset observational response to anomalously high Mediterranean SSTs in ASO (similar to Figure 31a-c).

# Chapter 5 Observed terrestrial feedbacks across North Africa

In this chapter, the observed terrestrial feedbacks across the Sahel, WAM region, and HOA are assessed using GEFA (Yu et al. 2017b). This part of the analysis is motivated by the previous model-based hypothesis of positive vegetation-rainfall feedbacks across North Africa and the underlying mechanism, especially the debate on the dominance of either the albedo feedback or moisture feedback in the simulated positive vegetation-rainfall feedbacks. The analysis presented in this chapter mainly addresses the following questions: How important are vegetation feedbacks compared with oceanic drivers on observed North African climate variability? What mechanism is responsible for observed North African vegetation feedbacks? NDVI is applied here as an index for vegetation growth in the observational study, rather than LAI as in the model-based GEFA application to CESM CTRL. The radiation used to compute NDVI is directly detected by satellite instruments, while LAI is a derived variable calculated from NDVI and thus contains an additional source of uncertainty introduced by the conversion from NDVI to LAI (Zhu et al. 2013). The observed terrestrial feedbacks further serve as a benchmark for evaluating the terrestrial feedbacks simulated by CESM, as a demonstration of planned future evaluation of the full set of CMIP5 models.

## 5.1.Relative contribution of oceanic drivers and terrestrial drivers to North African climate

According to GEFA, oceanic and terrestrial drivers together explain about 20-50% of the observed total variance in North African precipitation and 30-80% of the total variance in surface air temperature on the seasonal to interannual time scale, varying by season and sub-region (Figure 38). Consistent with findings from previous modeling studies (Zeng et al. 1999; Wang et

al. 2004), oceanic drivers exert the dominant control on North African precipitation and temperature in most seasons, especially during the pre- and early-monsoon season across the Sahel and WAM region, when vegetation is emerging and the interannual variability in vegetation growth is small (Figure 2). In addition, the variability of ENSO, one of the dominant oceanic forcings for the Sahel and WAM region, decays during boreal summer, leading to diminished oceanic controls on Sahel and WAM rainfall during the post-monsoon season (Figure 30). The terrestrial impacts exceed oceanic impacts on observed precipitation during the post-monsoon (SON) period across the Sahel and WAM regions, following the mean seasonal cycle of soil moisture (Liu et al. 2013). Largely consistent explained variance is obtained from multiple datasets, demonstrating the robustness of the observed relative contribution of oceanic and terrestrial influences to variability in North African climate.

Among the three sub-regions, the Sahel is observed to be most sensitive to terrestrial forcings. The explained variance in precipitation by terrestrial forcings is 8% on the annual mean and ranges from 0% in April-June (AMJ) to 18% in August-October (ASO) and SON, which is comparable in magnitude with the terrestrial impacts on air temperature. SON is the only season in which the land surface forcings are observed to dominate over oceanic forcings, with land surface variability explaining 18% (17.5-18.2% among observational datasets) of the total variance in precipitation. The relatively enhanced contribution from land-atmosphere interactions in autumn is likely attributed to two factors: (1) the reduced amplitude and broad-scale atmospheric circulation impacts of key ocean-atmosphere teleconnection patterns (Frederiksen and Branstator 2005), including ENSO, and (2) seasonally wet soils and consequential vegetation growth in response to the antecedent monsoon that support significant ET fluxes (Figure 39).

The residual portion of the total variance in precipitation and air temperature includes the atmospheric internal variability (equation 1), non-linear impacts of oceanic and terrestrial forcings which are not detected by the linear GEFA statistical method, and impacts from other oceanic (e.g. higher order SST EOFs or SST EOFs from other basins) or terrestrial (e.g. NDVI from other ecoregions) forcings absent from the forcing matrix. Measurement errors, which have been partly accounted for in the analysis, potentially contribute to this residual explained variance.

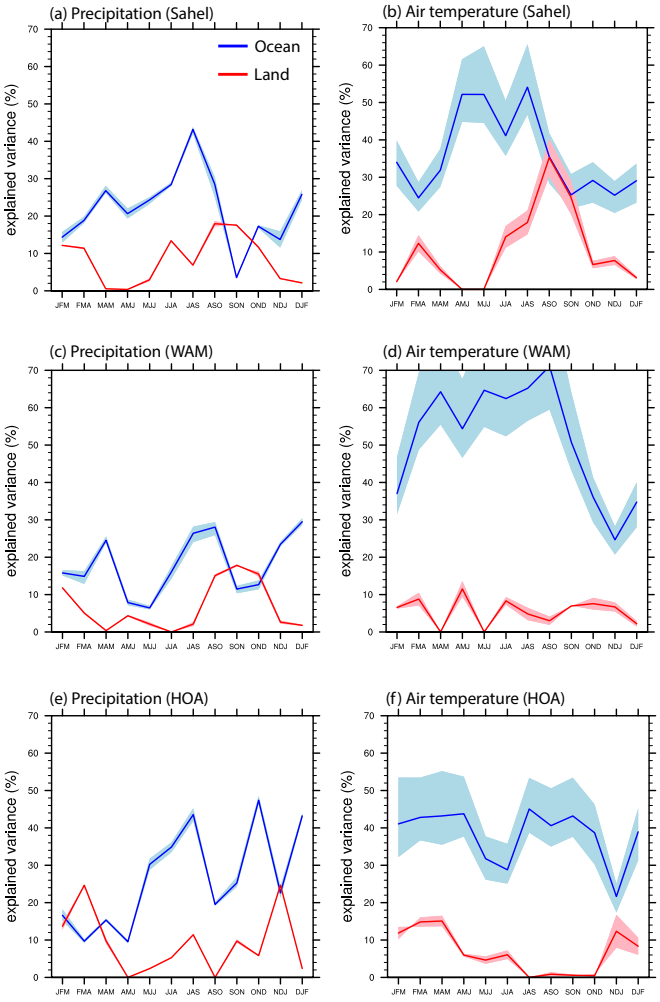

**Figure 38** Total percent variance in observed (a,c,e) precipitation and (b,d,f) 2-m air temperature across the (a,b) Sahel, (c,d) WAM region, and (e,f) HOA explained by oceanic (blue) versus terrestrial (red) forcings in each season. The lines represent the multi-dataset average, and the shading represents the $10^{th}$ and $90^{th}$ percentiles of the multi-dataset uncertainty range.

The smaller multi-dataset uncertainty range in precipitation than temperature is due to the inclusion of a greater number of precipitation datasets in the analysis. By randomly selecting two out of the four precipitation datasets, the typical multi-dataset uncertainty largely increases and becomes comparable with that of air temperature.

### 5.2. Observed terrestrial impacts on North African climate

The proposed positive vegetation feedbacks on precipitation by previous modeling studies (Charney 1975) are confirmed in observations across the semi-arid Sahel during the post-monsoon seasons and across the HOA during and after the short rains. However, these positive vegetation feedbacks are largely due to the moisture recycling mechanism, rather than the classic albedo feedback mechanism (Charney 1975). Positive anomalies in NDVI across Sahel and HOA favor enhanced ET, precipitable water, convective activity, and rainfall, indicative of amplified moisture recycling. In the Sahel, diminished vegetation growth and accompanying dry soils lead to enhanced dust emissions, which also potentially contribute to the positive vegetation-rainfall feedback. Observed vegetation feedbacks are trivial in the wetter WAM region. The subsequent sub-sections examine the moisture, albedo, and momentum feedbacks, that are featured in past modeling studies, and potential other mechanism that contribute to the positive vegetation-rainfall feedbacks in North Africa.

### 5.2.1. Observed moisture feedbacks

Across the Sahel, positive NDVI anomalies favor enhanced ET, precipitable water, convective activity (reduced OLR), and total precipitation, indicative of amplified moisture recycling in the observations (Figure 39). The ET and precipitation responses are of comparable magnitude during SON, namely +0.43 (0.40 - 0.44 among datasets) cm $month^{-1}$ per standard deviation of NDVI anomalies ($\sigma^{-1}_{NDVI}$) and +0.49 (0.48 - 0.50 among datasets) cm $month^{-1}$ $\sigma^{-1}_{NDVI}$, respectively, implying the dominance of the moisture recycling mechanism underlying the positive vegetation-rainfall feedback in the post-monsoon season across the Sahel.

Vegetation imposes a greater influence on the frequency, rather than intensity, of convective activity across the Sahel. While this has not been previously reported, the findings are consistent with previous observational results regarding the enhanced probability of afternoon precipitation across the eastern United States and Mexico due to anomalously high evaporation rates (Findell et al. 2011). Specifically, the increase in precipitation amount of +7.7 (7.5 - 7.9 among datasets) % $\sigma^{-1}_{NDVI}$ during SON is largely due to a +7.8% $\sigma^{-1}_{NDVI}$ increase in precipitation frequency, with minimal response in precipitation event intensity to NDVI anomalies. A likely explanation for the unique response in precipitation frequency versus intensity is that surface turbulent flux partitioning, associated with vegetation and soil moisture anomalies, shifts the local atmosphere between non-convective to convective states, while other controls, such as tropospheric moisture content or large-scale moisture convergence, largely determine rainfall intensity (Romps and Kuang 2010). Furthermore, positive NDVI anomalies favor a higher frequency of moderately low OLR days in the Sahel (not shown), implying that anomalous vegetation growth supports an increased chance of convective events with moderate intensity. Although the GEFA-based analysis agrees with previous regression-based analyses on the existence of a positive vegetation-

rainfall feedback in the Sahel, the statistical vegetation index analysis concluded that the vegetation feedback peaks during the monsoon season (Los et al. 2006), rather than the post-monsoon season as identified by GEFA. A potential reason for the inconsistent conclusion is that the regression-based analysis did not account for the oceanic impacts on Sahel rainfall, which peak during the monsoon season (Figure 38) and likely bias the estimated influence of vegetation variability.

Enhanced moisture recycling associated with positive NDVI anomalies is also observed across the HOA, leading to a wetter and extended short-rain season (Figure 40). Unlike the Sahel, which has more distinct dry and wet seasons, significant ET responses to positive NDVI anomalies occur all-year-round across the HOA. However, significant enhancement in 2-m specific humidity, precipitation, convective activity, and rainfall in response to positive NDVI anomalies is confined to the mid- to post- short-rain season, when both ET and precipitation responses exhibit similar magnitudes at about 0.5 cm month$^{-1}$ $\sigma^{-1}_{NDVI}$.

Across the climatologically wetter WAM region, a positive ET response to increased NDVI is observed during the pre- and post-monsoon seasons, but with a smaller magnitude than observed across the Sahel and HOA. However, vegetation-rainfall feedbacks across the WAM region are largely insignificant during most seasons (Figure 41), consistent with findings regarding soil moisture feedbacks from previous modeling studies that the semi-arid Sahel and HOA exhibit stronger soil moisture-rainfall coupling than the climatologically wetter WAM region (Koster et al. 2004), and previous model-based findings that grasses and shrubs induce

stronger moisture feedbacks than trees (Liu et al. 2010; Notaro et al. 2011).

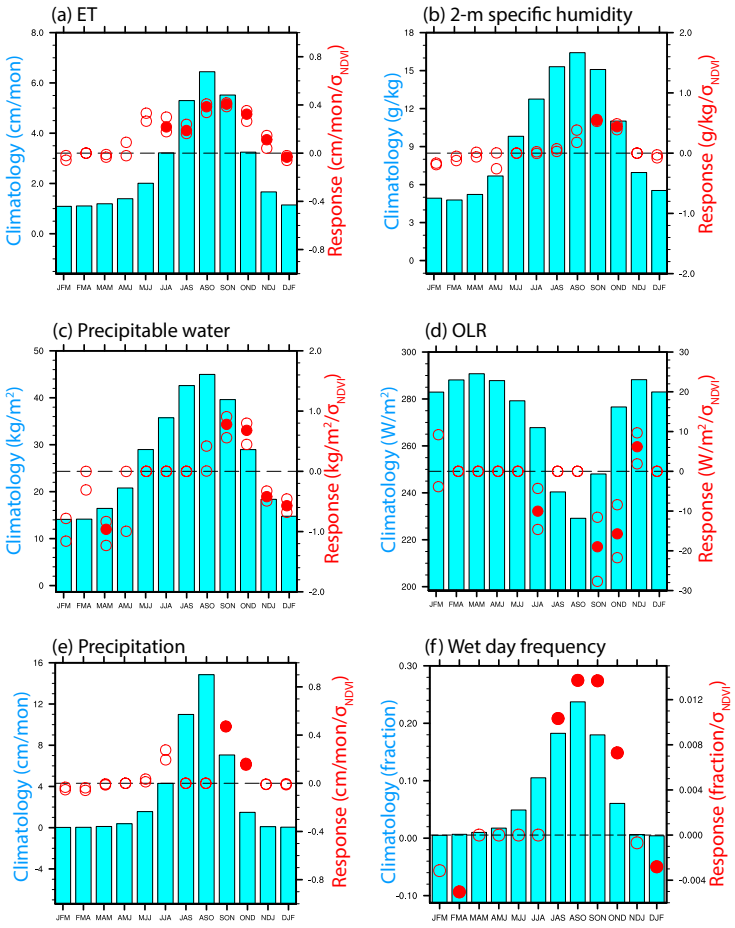


**Figure 39** Multi-dataset observational moisture response to positive local NDVI anomalies across the Sahel in (a) ET, (b) 2-m specific humidity, (c) precipitable water, (d) OLR, (e) precipitation, and (f) frequency of wet days. In (a-e), dots indicate statistically significant ($p<0.1$) multi-dataset average responses, referring to the right y-axis; open circles represent the 10$^{th}$ and 90$^{th}$ percentiles of the multi-dataset responses, regardless of their statistical significance. In (f), dots and open circles indicate significant and insignificant responses, respectively, using daily station precipitation data. Bars indicate the multi-dataset climatology of the response variable, referring to the left y-axis. "$\sigma_{NDVI}$" in the units of the response variables refers to one standard deviation of the Sahel NDVI anomaly. The statistical significance is determined based on Monte Carlo bootstrap testing.

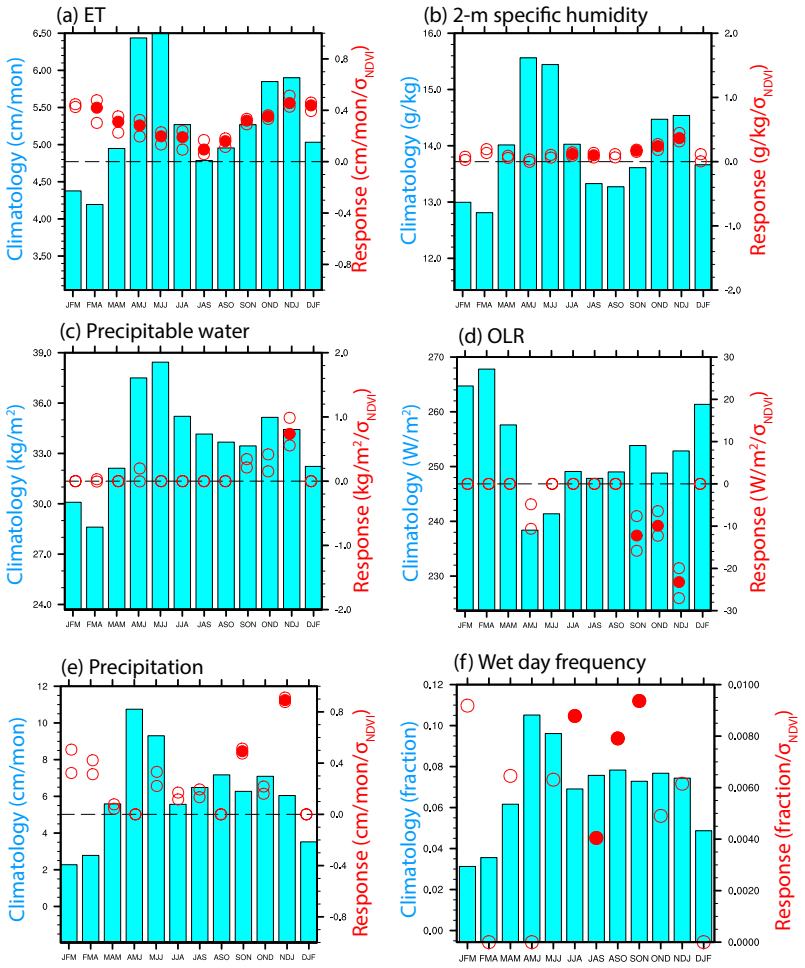


**Figure 40** Same as Figure 39, but for HOA.

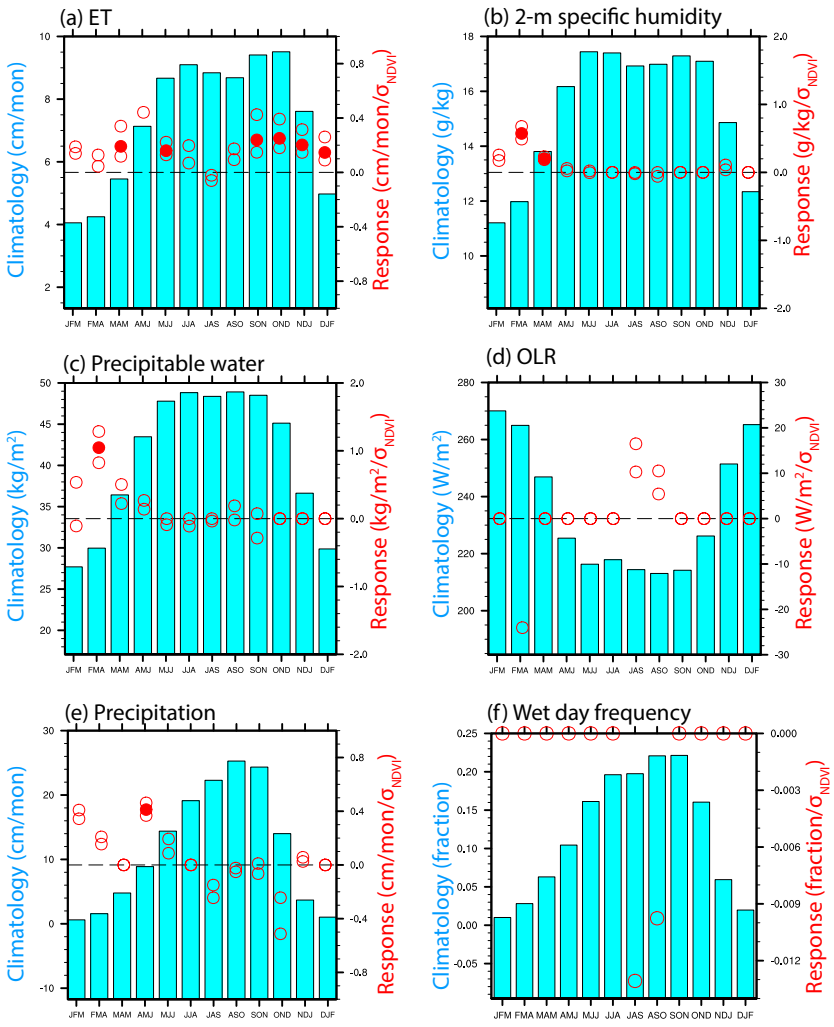


**Figure 41** Same as Figure 39, but for the WAM region.

### 5.2.2. Observed albedo feedbacks

In contrast to Charney's model-derived hypothesis (1975), the observed NDVI and surface albedo are not strongly coupled over the shrublands or grasslands of the Sahel and HOA on the seasonal to interannual time scales, which is unsupportive of the traditional albedo feedback (Figure 42). During autumn, positive NDVI anomalies trigger modest declines in surface albedo,

confined to the savanna, woody savanna, and evergreen broadleaf forest portions of the Sahel and WAM region, consistent with previous model-based findings (Notaro et al. 2008). However, their magnitudes, on the order of 0.01-0.04 $\sigma^{-1}_{NDVI}$, are too small to generate significant surface warming or ascending motion, compared with the imposed change of 0.21 in surface albedo applied in the pioneering experiments of Charney (1975). Indeed, increased NDVI leads to surface cooling across the Sahel, WAM region, and HOA when positive vegetation-rainfall feedbacks are present, in opposition to Charney's hypothesis (Charney 1975), likely due to increased ET and latent heat flux and thus decreased Bowen ratio in response to positive NDVI anomalies. The albedo responses are trivial over the more widespread grasslands and shrublands of the Sahel and HOA, probably because the seasonal-interannual time scale is too short for the grass-desert or shrub-desert conversions proposed by Charney (1975). However, at long time scales associated with land cover / land use change, the albedo impact might be significantly more important, especially if pronounced grass-desert conversion or soil degradation occurs.

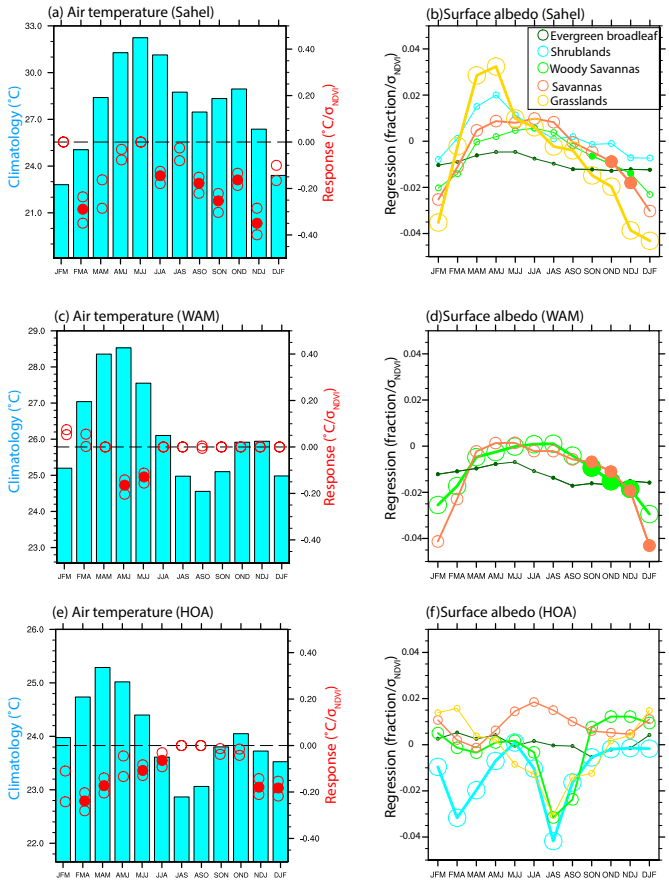


**Figure 42** Observed local energy (albedo) responses to positive NDVI anomalies across the (a,b) Sahel, (c,d) WAM region, and (e,f) HOA. (a,c,e) Surface air temperature, with similar figure elements as Figure 39, and (b,d,f) temporal regression coefficient of monthly surface albedo upon standardized local NDVI anomalies, averaged by biome type. In (b,d,f), filled dots indicate statistically significant ($p<0.1$) correlations, according to the Student's t-test; the size of the circles denotes the abundance of the corresponding biome across the focal region, which is based on the remotely-sensed land cover type from the International Satellite Land Surface Climatology Project (ISLSCP) initiative II International Geosphere-Biosphere Project (IGBP) DISCover and Simple Biosphere (SiB) Land Cover dataset.

### 5.2.3. Observed momentum feedbacks

The momentum (roughness) feedback mechanism is not responsible for the positive vegetation-rainfall feedbacks across the three studies regions. Although decreased near-surface wind speeds are observed in response to increased NDVI during the growing seasons across the Sahel and HOA, anomalous low-level convergence as indicated by negative anomalies in sea-level pressure and anomalous low-level ascent, which is typical in the classical positive momentum feedback, does not occur during any season for any region (Figure 43). Indeed, the observed anomalous low-level subsidence in response to increased vegetation abundance is likely caused by increased atmospheric stability, which is induced by surface and low-level cooling.

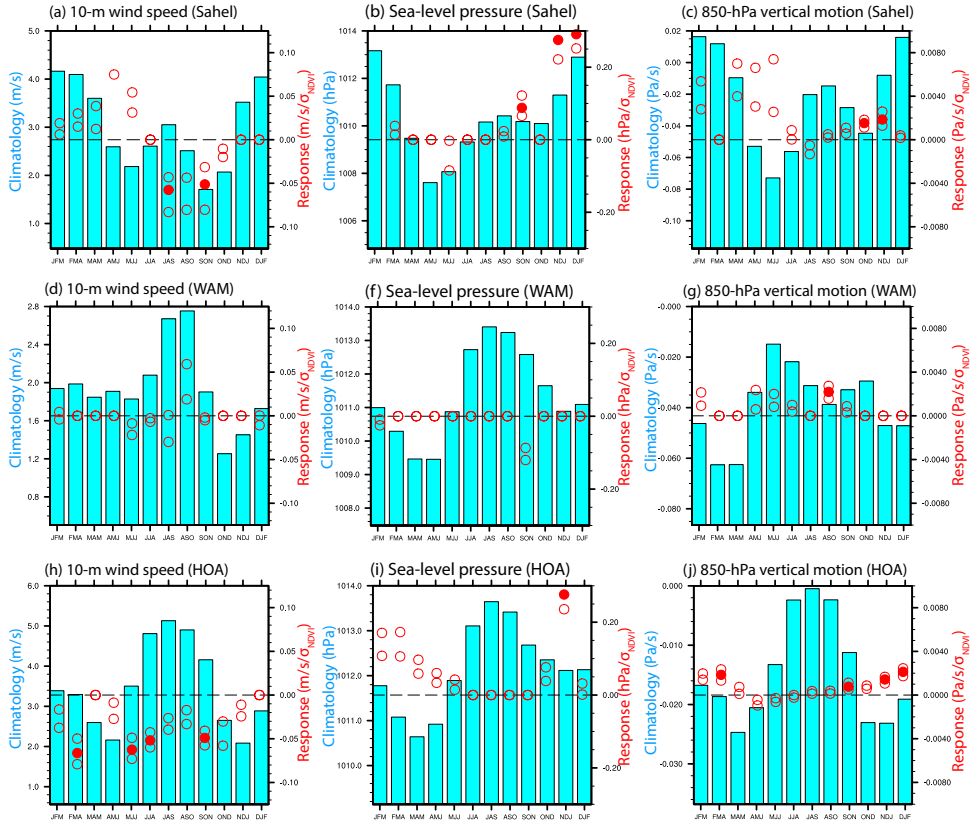

**Figure 43** Multi-data observed momentum response to positive local NDVI anomalies across the (a-c) Sahel, (d-g) WAM region, and (h-j) HOA in (a,d,h) 10-m wind speed, (b,f,i) sea-level pressure, and (c,g,j) 850-hPa vertical motion (positive for anomalous subsidence). Figure elements are similar to Figure 39 (a-e).

### 5.2.4. Observed remote responses to HOA vegetation anomalies

Vegetation anomalies in the Sahel, WAM region, and HOA generally exert limited remote impacts on the atmosphere according to observations. A significant remote atmospheric response is only identified over the tropical western Indian Ocean in response to HOA NDVI anomalies after the short rains (Figure 44). Enhanced ET in response to positive NDVI anomalies across the HOA leads to local low-level cooling and descending motion, which induces anomalies westerlies along the terrain gradient from the Eastern Rift mountains in Ethiopia and Kenya to the western Indian Ocean coast. The anomalous low-level westerlies lead to low-level convergence and mid- to low-level ascending motion over the tropical western Indian Ocean. As a consequence of the increased mid- to low-level atmospheric instability, deep convection is enhanced and precipitable water is increased over the tropical western Indian Ocean.

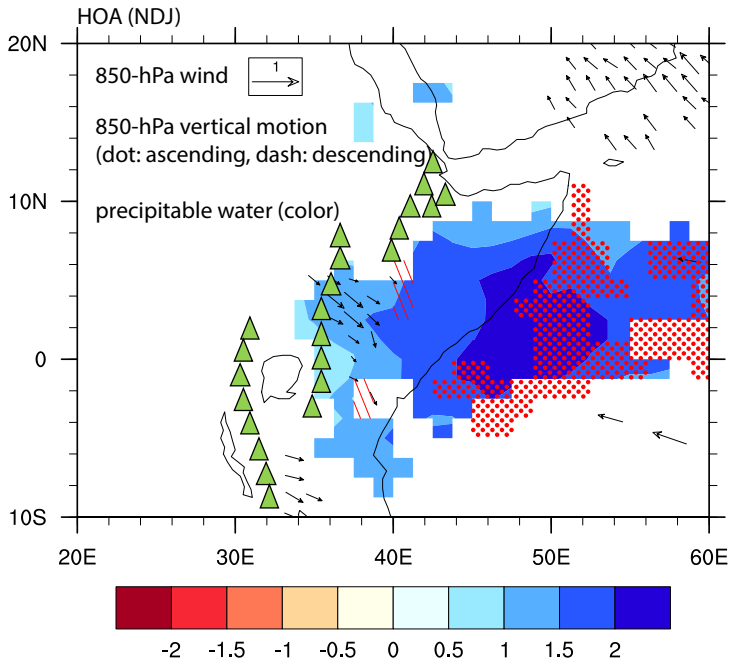


**Figure 44** GEFA-based weighted multi-dataset average observed response to positive HOA NDVI anomalies during NDJ: 850-hPa wind (m $s^{-1}$ $\sigma_{NDVI}^{-1}$, vector), 850-hPa vertical motion (Pa $s^{-1}$ $\sigma_{NDVI}^{-1}$, dot: ascending, dash: descending), and precipitable water (g $cm^{-2}$ $\sigma_{NDVI}^{-1}$, color). The green triangles indicates mountain areas. Only statistical significant responses ($p<0.1$), based on the Monte Carlo bootstrap test, are shown.

### 5.2.5. Observed dust responses to the Sahel vegetation anomalies

The observational GEFA analysis further verifies that diminished vegetation growth and accompanying dry soils across the Sahel lead to enhanced dust emissions and dust storm activity during the mid- to post-monsoon season (Figure 45), as suggested by previous correlation-based observational studies (Kim et al. 2017). In addition, the current observational analysis reveals the remote impacts of Sahel vegetation and soil moisture on dust concentration over the tropical Atlantic Ocean. The enhancement in dust emissions is most pronounced across the southern

boundary of North Africa's major dust source regions, including the Bodélé Depression (Evan et al. 2015), where dust emissions are substantially increased by more than 60% during SON corresponding to a negative anomaly in NDVI on the order of one standard deviation. Increased surface wind speed and soil bareness, as caused by inhibited vegetation growth, as well as the accompanying decreased soil moisture, all likely contribute to the enhanced dust emissions. The enhanced dust emissions across the Sahel support increased southward dust transport and thereby elevated surface and column dust concentrations, as well as frequency of dust days according to station observations over land, across tropical and subtropical North Africa and the eastern tropical Atlantic Ocean (Figure 45). The increase in atmospheric dust concentration and its associated radiative cooling could inhibit tropical cyclone development (Emanuel 2005) by lowering tropical Atlantic SSTs. This observed vegetation-dust feedback acts as a potential secondary mechanism for the positive vegetation-rainfall feedback in the Sahel, given the direct effects of dust aerosols that cause low-level cooling and atmospheric stabilization of the atmosphere (Rosenfeld et al. 2001), and the indirect radiative effects of dust aerosols that increase the number of cloud condensation nuclei and inhibit precipitation efficiency (Gu et al. 2015).

There are minimal dust responses to NDVI anomalies across the WAM region, which is mostly covered by forests and exhibits minimal dust emission, or HOA, where dust emission

occurs mostly from bare ground rather than vegetated regions.

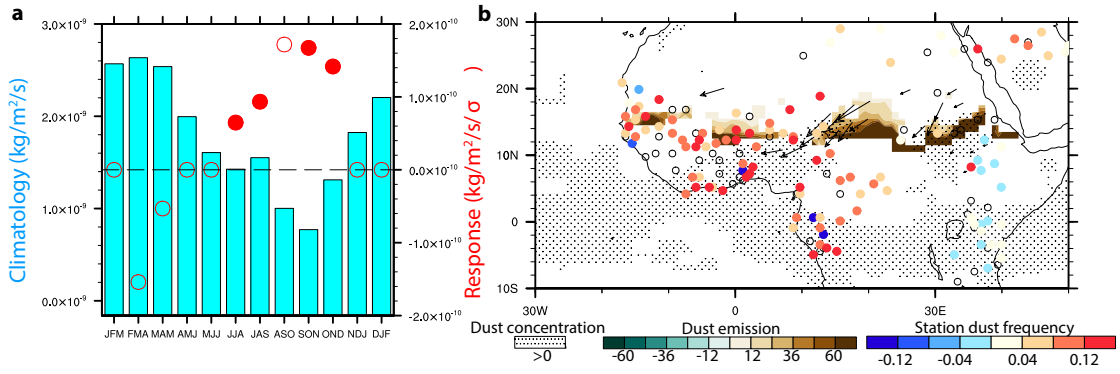


**Figure 45** Observed dust responses to negative Sahel NDVI anomalies and corresponding dry soil anomalies. (a) Total dust emission, with similar figure elements as Figure 39, and (b) spatial pattern in dust emission (% climatology $\sigma_{NDVI}^{-1}$, green-brown color), column dust concentration (kg $m^{-2}$ $\sigma_{NDVI}^{-1}$, stitching and hatching), column dust transport (kg $m^{-1}$ $s^{-1}$ $\sigma_{NDVI}^{-1}$, vector), and regional dust frequency (fraction $\sigma_{NDVI}^{-1}$, blue-red color) in September-November. Only statistical significant ($p<0.1$) responses are shown in b, according to the Monte Carlo bootstrap test. Open circles in (b) represent stations with insignificant responses in dust frequency.

### 5.3.Evaluating terrestrial impacts on North African climate simulated by CESM

In order to evaluate the simulated terrestrial impacts on North African climate within CESM, the simulated annual cycle of vegetation growth and mean vegetation distribution are first evaluated against observations. The model-simulated LAI and remotely sensed NDVI/LAI exhibit different seasonal cycles, which reflect the biased timing of vegetation growth in the North African sub-regions simulated by CESM (Figure 46). Across the Sahel and WAM region, the simulated LAI peaks in October and November, later than the September peak in remotely

sensed NDVI and LAI, potentially due to abnormally longer memory in soil moisture in these regions in CESM. Across the HOA, vegetation growth peaks after the short rains in CESM, probably due to the excessive rainfall received in the short rains in CESM (Figure 48d). CESM also simulates a greater interannual variability in boreal springtime vegetation growth in all three study regions than in observation. Furthermore, CESM generally simulates less vegetation greenness in the Sahel and HOA, and excessive vegetation greenness in the WAM region, as reflected by the difference in mean LAI between satellite data and CESM. In addition to the biases in the mean and interannual variability of vegetation growth, CESM simulates a very different vegetation distribution across North Africa than observed. Across the Sahel and HOA, CESM with dynamic vegetation produces excessive bare ground and less shrublands than observed, which explains the negative bias in mean LAI compared with remote sensing. In the WAM region, excessive shrublands and less tree cover are present in CESM. The biased vegetation distribution supports the rotated EOF patterns in CESM-simulated LAI (Figure 8) characterized by a southward-displaced Sahel box and westward-extended HOA box compared to observations, given that the northern Sahel and eastern HOA are mostly bare ground in the model.

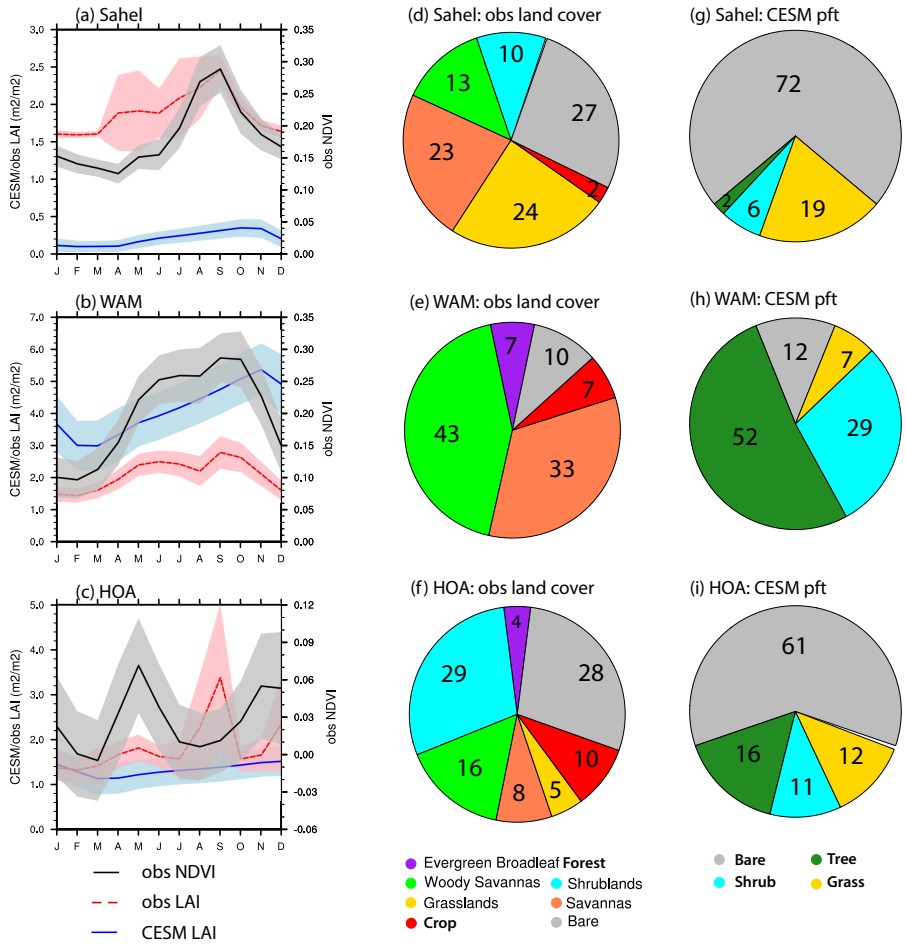


**Figure 46** (a-c) Seasonal cycle in vegetation growth [remotely sensed NDVI: black lines (average) with grey shading (interannual standard deviation) during 1982-2011; remotely sensed LAI: red dashed lines (average) with pink shading (interannual standard deviation) during 1982-2011; LAI in CESM: blue lines (average) with light blue shading (interannual standard deviation) in the 300-year CTRL] averaged across the (a) Sahel, (b) WAM region, and (c) HOA. (d-f) Observed percent area of land cover types across the (d) Sahel, (e) WAM region, and (f) HOA. (g-i) Percent cover of main vegetation types in CESM across the (g) Sahel, (h) WAM region, and (f) HOA. The geographic regions of Sahel, WAM region, and HOA are defined the same as the current observational analysis.

Corresponding to the biases in the simulated mean vegetation growth and distribution, the simulated vegetation feedbacks in CESM are substantially different from those detected in observations through SGEFA. Across the Sahel, positive vegetation-rainfall feedbacks are only present during spring in CESM, at the time of peak ET response to LAI anomalies, in contrast to the observed strongest vegetation-rainfall feedback during autumn (Figure 47). A likely cause for the spring peak in vegetation-rainfall feedback in CESM is the excessive spring precipitation and accompanying wet soil. During autumn, although a significant ET response is present in CESM, the magnitude is much smaller than in observations, leading to trivial responses in precipitable water and precipitation. The smaller magnitude in ET responses during autumn is likely due to the biased vegetation distribution in CESM, in which the Sahel is most covered by bare ground and exhibits limited response in canopy transpiration and canopy evaporation. The underestimated vegetation growth and interannual variability in autumn is likely caused by a drier simulated monsoon in the Sahel, which leads to subsequent drier soil in autumn than observed. The biased vegetation distribution also explains underestimated moisture responses in the HOA (Figure 48). Across the WAM region, smaller biases in vegetation distribution likely attribute to smaller model-observational inconsistencies in moisture feedbacks compared to the Sahel (Figure 49). In addition, CESM-DGVM only represents natural vegetation types, which also potentially contributes to the bias in simulated feedbacks. The potential influence of vegetation distribution biases on the simulated vegetation feedbacks can be tested in the future by forcing the model with observed vegetation seasonal cycle and vegetation type.

Another potential reason for the simulated biases in vegetation-rainfall feedback is the overwhelming atmospheric stabilization response to vegetation anomalies in CESM. For example, across the HOA, substantial positive ET responses to local vegetation anomalies during

the boreal winter (after the short rains and before the long rains) do not cause significant positive responses in precipitable water or precipitation (Figure 48). This is likely due to the mid- to low-level atmospheric stabilization, as indicated by anomalous descending motion, in response to positive LAI anomalies. Indeed, across the Sahel (Chapter 3), the atmospheric stabilization effect dominates over other mechanisms and produces a negative vegetation-rainfall feedback in CESM. This is not the first time that negative vegetation-rainfall feedbacks across North Africa have been identified in a modeling study. For example, Notaro et al. (2011) simulated an increase in post-monsoon rainfall in response to negative LAI anomalies over the North African monsoon region with CESM's predecessor, CCSM3.5. Negative vegetation-rainfall feedbacks were also identified during the mid-Holocene in the Fast Ocean Atmosphere Model-Lund Potsdam Jena dynamic global vegetation model (FOAM-LPJ) (Notaro et al. 2008).

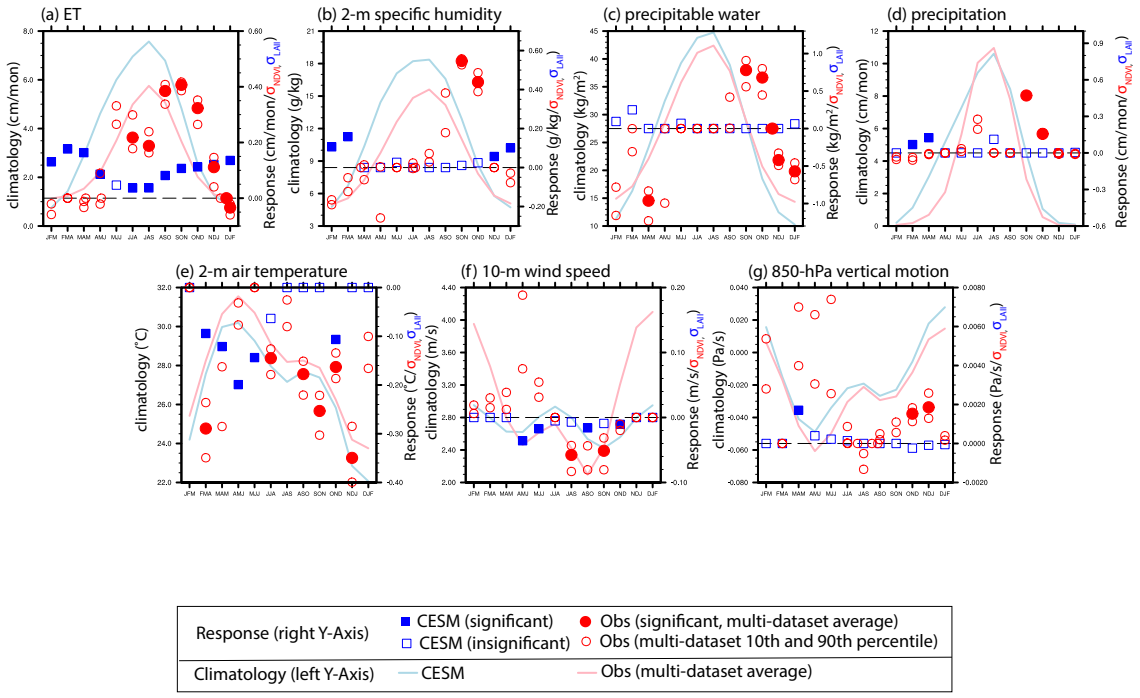


**Figure 47** Observed and CESM-simulated response to positive local vegetation anomalies (NDVI in observations and LAI in CESM) across the Sahel in (a) ET, (b) 2-m specific humidity, (c) precipitable

water, (d) precipitation, (e) 2-m air temperature, (f) 10-m wind speed, and (g) 850-hPa vertical motion. Lines represent the climatology in observations (1982-2011, multi-dataset average, pink) and CESM (300-year CTRL, light blue), referring to the left y-axis. Circles (observation) and squares (CESM) represent the response, referring to the right y-axis. Filled (open) squares indicate significant (insignificant) responses in CESM, based on the Monte Carlo bootstrap test ($p<0.1$). Filled (open) circles indicate the multi-dataset average ($10^{th}$ and $90^{th}$ percentiles) responses in observations.

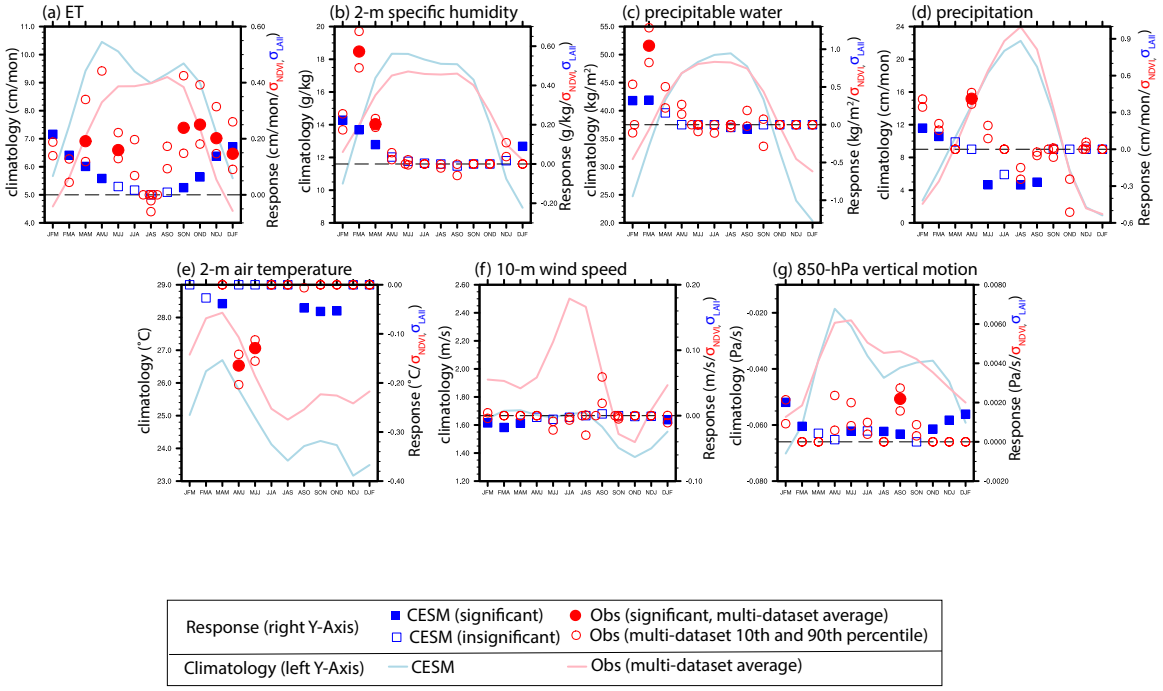


**Figure 48** Same as Figure 47, but for local responses to vegetation anomalies across HOA.

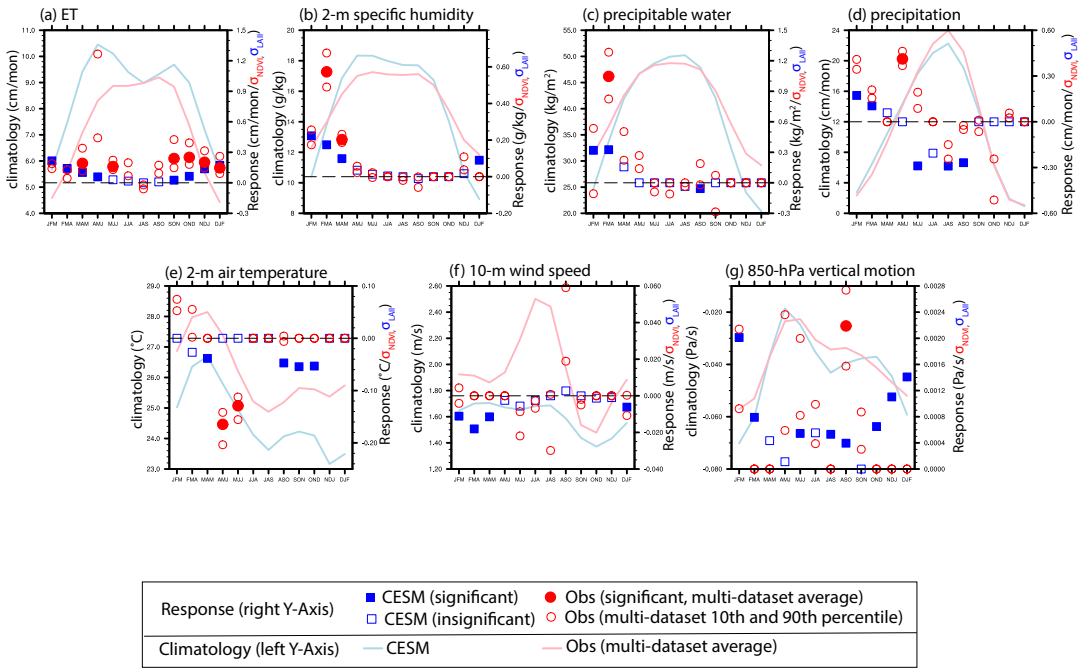


**Figure 49** Same as Figure 47, but for local responses to vegetation anomalies across WAM region.

Model deficiencies in the representation of terrestrial feedbacks, as indicated by the application of CESM in the current study and CCSM3.5 in the previous study by Notaro et al. (2011), motivate the evaluation of coupled Earth System Models against the observational benchmark established in this study. Land-atmosphere interactions remain a key source of uncertainty in climate modeling and climate change projections (Flato et al. 2013). Given the substantial uncertainties in the sign and magnitude of projected changes in Sahel rainfall among the CMIP5 models (Giannini et al. 2008; Roehrig et al. 2013), most of which contain interactive vegetation phenology and some of which also contain dynamic vegetation cover, it is necessary to first evaluate the representation of vegetation-climate interactions by these coupled Earth System Models against the observational benchmark. With rigorous evaluation of the simulated

vegetation feedbacks on the North African regional climate in the CMIP5 models, we can develop process-based weights for the models and narrow the uncertainty in future climate projections.

# Chapter 6 Summary and future work

The current study examines the observed oceanic and terrestrial drivers of North African regional climate by applying the multivariate statistical method, GEFA, to a spectrum of reanalyses, remote-sensing products, gridded observational datasets, and station observations. Before applying GEFA to observations, the reliability of the statistical method at extracting oceanic and terrestrial forcings on North African regional climate is first demonstrated using CESM. The traditional GEFA approach is refined through stepwise GEFA, in which oceanic and terrestrial forcings that fail to significantly influence North African regional climate are dropped from the forcing matrix through stepwise selection, thereby reducing the size of the matrix and increasing the accuracy of estimated response fields.

Consistent atmospheric responses to SST anomalies across the tropical Pacific, tropical Atlantic, and tropical Indian Oceans, as well as North African LAI anomalies, between the statistical GEFA approach, as applied to the CESM control run, and the dynamical approach, based on CESM ensemble experiments, demonstrate GEFA's capacity to successfully extract oceanic impacts and vegetation feedbacks within the complex Earth system. Stepwise GEFA boasts higher spatial correlations, SNRs, and sign consistency between the seasonal cycles of statistically- and dynamically-assessed atmospheric responses to both oceanic and land surface anomalies across North Africa than those achieved through traditional GEFA. This permits the reliable estimation of atmospheric response fields through stepwise GEFA, even with relatively short data records (e.g. as found with many observational products). In assessing the terrestrial impacts, GEFA-based atmospheric responses to LAI anomalies are much more consistent in magnitude with the $EXP_{SOIL}$ ensemble, in which the coupled soil moisture-LAI in the focal

region is modified, than with the $EXP_{LAI}$ ensemble, in which only the regional LAI is modified, during months with strong soil moisture-LAI coupling. This indicates that GEFA captures the combined impact of both LAI and soil moisture anomalies on atmospheric conditions.

The observational analysis identifies tropical ocean SST variability as the key driver of North African climate, with dominant impacts of SST variability across the tropical Pacific and tropical Atlantic Oceans on Sahel and WAM rainfall and tropical Pacific and tropical Indian Oceans on HOA rainfall. The tropical Indian Ocean, in contrast to previous modeling studies (Lu and Delworth 2005), does not appear to be an important driver of Sahel monsoonal rainfall in the current observational analysis, likely due to the exaggerated forcings applied in such modeling studies (Lu and Delworth 2005). In addition, the simulated impacts of tropical Indian Ocean SSTs in these modeling studies potentially originate from the tropical Pacific Ocean, since the variability in Indian Ocean SSTs are mainly driven by tropical Pacific SSTs.

According to GEFA, oceanic drivers dominate over terrestrial drivers in terms of their observed impacts on North African climate in most seasons. Terrestrial impacts are comparable to, or more important than, oceanic impacts on rainfall during the post-monsoon across the Sahel and WAM region, and after the short rain across the HOA.

The observational analysis further verifies the model-based hypothesis of positive vegetation-rainfall feedbacks in the semi-arid Sahel (Charney 1975; Zeng et al. 1999; Wang et al. 2004) and HOA, which is confined to the post-monsoon and post-short rain season, respectively. However, this observed positive feedback is associated with a moisture recycling mechanism, rather than the classic albedo-based mechanism (Charney 1975), on the seasonal to interannual time scale. Across the Sahel, positive NDVI anomalies during the late and post-monsoon periods

favor enhanced ET, precipitable water, convective activity, and rainfall, indicative of amplified moisture recycling. The identified modest low-level cooling and anomalous atmospheric subsidence in response to positive NDVI anomalies are counter to the responses expected through the classic vegetation-albedo feedback mechanism. The observational analysis further indicates that diminished vegetation growth and accompanying dry soils lead to enhanced dust emissions across the Sahel and transport to tropical/subtropical Africa and the eastern tropical Atlantic Ocean during the mid- to post-monsoon season, which potentially contribute to the positive vegetation-rainfall feedback through the direct and indirect radiative effects of dust aerosols. Across the HOA, positive NDVI anomalies support an extended short rain season through the moisture feedback mechanism.

The current study presents the first comprehensive observational exploration of the key oceanic and terrestrial drivers of North African climate. The relative importance of the land versus oceanic impacts on North African regional climate on the seasonal-to-interannual time scales is quantified for the first time using observational data. Furthermore, the current study presents the first convincing observational evidence for the model-hypothesized positive vegetation-rainfall feedbacks in the Sahel, by successfully isolating terrestrial feedbacks from oceanic drivers and systematically examining multiple observational datasets in order to quantify observational uncertainty in feedback response estimates. The identification of key oceanic and terrestrial drivers will aid in successful seasonal predictions of regional climate in North Africa, a highly vulnerable region to hydrological extremes.

The current study advances the GEFA methodology, although further explorations of GEFA's applicability to other ecoregions need to be accomplished by future studies. The development of stepwise GEFA in the current study leads to substantial improvement over

traditional GEFA when assessing feedbacks from short data records, which is particularly valuable for analyses of remote sensing products. In addition, given the demonstrated reliability of GEFA in extracting the individual oceanic and terrestrial feedbacks on the regional climate of North Africa (in the current study) and North America (Wang et al. 2013, 2014), one can apply GEFA to both tropical/subtropical and mid-latitude regions with confidence. However, both the current and previous GEFA validation studies (Wang et al. 2013, 2017) focused on tropical oceanic forcings. The accuracy of GEFA in capturing the extratropical oceanic impacts on regional climate requires further exploration, given the potentially different memory in extra-tropical SSTs than tropical SSTs and potential interactions between extra-tropical oceans and frontal systems. Moreover, GEFA's applicability to high-latitude regions, where potential impacts from snow and ice cover are present, needs to be explored by future studies.

The framework of the current observational analysis can be applied to address related scientific questions. For example, one remaining question from the current study involves the transient nature of the sign and intensity of oceanic and terrestrial feedbacks. While the model-based findings (Biasutti et al. 2008; Nicholson 2013) suggest weakened oceanic controls on the variability of Sahel rainfall in the late-20$^{th}$ century compared with early- and mid-20$^{th}$ century, it has never been demonstrated through observations, partly due to the limited reliability of observational data in the earlier decades of the 20$^{th}$ century. If reliable observational records are available for earlier time period, one can apply GEFA to different time windows and explore the temporal variability in the oceanic impacts. Similar, modeling studies have suggested that the sign and intensity of land surface feedbacks may not remain static with ongoing changes in climate and terrestrial ecosystems (Notaro et al. 2008; Dirmeyer et al. 2013; Willeit et al. 2013).

This can be explored by applying GEFA to paleoclimate records/simulations, modern climate records/simulations, and future climate projections (e.g. from CMIP5 models).

The current observational study considers a spectrum of observational, reanalysis, and remote sensing products and quantifies the observational uncertainty in an innovative way that takes into account the regional reliability of each dataset. In light of the growing number of observational products being made available, such a multi-dataset analysis approach is recommended for all observational studies. The multi-dataset bootstrapping approach is potentially useful for all quantitative observational studies, especially those with focus on data-sparse regions where the reliability of each available dataset differs substantially.

Future projections of Sahel rainfall, in response to the anthropogenically enhanced greenhouse effect, remain highly uncertain in terms of both sign and magnitude within CMIP5 (Giannini et al. 2008 a,b; Rodriguez-Fonseca et al. 2015; Roehrig et al. 2013). In further studies, the GEFA-based assessment of the key observed oceanic and terrestrial drivers of North African regional climate will serve as an observational benchmark for evaluating the representation of ocean-land-atmosphere interactions within state-of-the-art climate models as applied by the Intergovernmental Panel on Climate Change. This innovative approach will foster model evaluation and development, along with the formulation of regionally-specific, process-based model performance metrics for weighting future climate projections for the Sahel and reducing associated uncertainty.

## Reference

Africa Environment Outlook. Past, present and future perspectives. United Nations Environmental Programme (2002). Retrieved 2015-04-02.

Akaike, H., 1974: A new look at the statistical model identification, IEEE Transactions on Automatic Control, 19 (6), 716-723.

Amenu, G. G., and P. Kumar, 2005: NVAP and Reanalysis-2 global precipitable water products: Intercomparison and variability studies, Bull. Amer. Meteor. Soc., 86 (2), 245.

Bader, J. and Latif, M., 2003: The impact of decadal-scale Indian Ocean sea surface temperature anomalies on Sahelian rainfall and the North Atlantic Oscillation. Geophysical Research Letters, 30, 2165-2169.

Behera, S. K., J. J. Luo, S. Masson, P. Delecluse, S. Gualdi, A. Navarra, and T. Yagamata, 2005: Paramount impact of the Indian Ocean dipole on the East African short rains: A CGCM study, J. Climate, 18, 4514-4530.

Betts, A.K., and J.H. Ball, 1997: Albedo over the boreal forest. J. Geophys. Res., 102, 28901–28909.

Biasutti, M., and A. Giannini, 2006: Robust Sahel drying in response to late 20th century forcings. Geophys. Res. Lett., 11, L11706. doi:10.1029/2006GL026067.

Biasutti, M., I.M. Held, A.H. Sobel, and A. Giannini, 2008: SST forcings and Sahel rainfall variability in simulations of the twentieth and twenty-first centuries. J. Clim., 21, 3471-3486.

Bonan, G.B., 2002: Ecological Climatology: Concepts and Applications, Cambridge University Press, 678 pp.

Bonan, G.B., D. Pollard, and S.L. Thompson, 1992: Effects of boreal forest vegetation on global climate, Nature, 359, 716–718.

Bosilvich, M. G., R. Lucchesi, M. Suarez, 2015: MERRA-2: File Specification, GMAO Office Note No. 9, http://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf.

Brooks, N., 2004: Drought in the African Sahel: Long term perspectives and future prospects. Working Paper 61, Tyndall Centre for Climate Change Research, October 2004.

Buchard, V., & A. Da Silva, 2016, Ensemble-based assimilation of aerosol observations in GEOS-5. https://ntrs.nasa.gov/search.jsp?R=20160005006.

Buermann, W., 2002: The impact and response of vegetation to climate at interannual timescales. Ph.D. thesis, Boston University, 140 pp.

Charney, J.G., 1975: Dynamics of desert and drought in the Sahel. Quarterly Journal of the Royal Meteorological Society, 101, 193-202.

Charney, J.G., W.J. Quirk, S.H. Chow, and J. Kornfield, 1977: A comparative study of the effects of albedo change on drought in semi-arid regions. Journal of Atmospheric Sciences, 34, 1366-1385.

Chen, G., M. Notaro, Z. Liu, and Y. Liu, 2012: Simulated local remote biophysical effects of afforestation over the Southeast United States in boreal summer. J. Climate, 15, 2616-2631.

Chiang, J. C. H., A.H. Sobel, 2002: Tropical tropospheric temperature variations caused by ENSO and their influence on the remote tropical climate, J. Clim., 15(18), 2616-2631.

Chou, C., J.D. Neelin, and H. Su, 2001: Ocean–atmosphere–land feedbacks in an idealized monsoon. Q. J. R. Meteorol. Soc., 127, 1869–1891.

Clark, D.B., Y.K. Xue, R.J. Harding, and P.J. Valdes, 2001: Modeling the impact of land surfacedegradation on the climate of tropical North Africa. J. Clim., 14, 1809–1822.

Dee, D. P. et al., 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quart. J. R. Meteorol. Soc, 137, 553-597.

Dirmeyer, P.A., Y. Jin, B. Sign, and X. Yan, 2013: Trends in land-atmosphere interactions from CMIP5 simulations. Journal of Hydrometeorology, 14, 278-288.

Doherty, R., J. E. Kutzbach, J. Foley, D. Pollard, 2000: Fully coupled climate/dynamical vegetation model simulations over Northern Africa during the mid-Holocene. Climate Dynamics, 16, 561-573.

Efron, B, and B. Efron, 1982: The jackknife, the bootstrap and other resampling plans, Vol. 38 Philadelphia: Society for industrial and applied mathematics.

Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years, Nature, 436, 686-688.

Evan, A. T. et al. 2015: Derivation of an observation-based map of North African dust emission, Aeolian Res., 16, 153-162.

Food and Agriculture Organization of the United Nations, 2016: With continued drought, Horn of Africa braces for another hunger season. http://www.fao.org/news/story/en/item/460996/icode/

Farmer, G., and T.M.L. Wigley, 1985: Climatic trends for tropical Africa. Research Report, University of East Anglia, Norwich, UK, pp. 136.

Findell, K., P. Gentine, B. R. Lintner, and C. Kerr, 2011: Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation, Nat. Geosci., 4, 434-439.

Flato, G. et al., 2013: Evaluation of climate models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Stocker et al. (eds.)). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Foley, J. A., M. T. Coe, M. Scheffer, G. L. Wang, 2003: Regime shifts in the Sahara and Sahel: interactions between ecological and climatic systems in northern Africa. Ecosystems, 6, 524-539.

Folland, C., J. Owen, M.N. Ward, and A. Colman, 1991: Prediction of seasonal rainfall in the Sahel region using empirical and dynamical models. J. Forecast., 10, 21–56.

Folland, C.K., T.N. Palmer, and D.E. Parker, 1986: Sahel rainfall and worldwide sea temperatures, 1901-85. Nature, 320, 602–607. doi:10.1038/320602a0.

Frankignoul, C., and K. Hasselmann, 1977: Stochastic climate models, Part II: Application to sea-surface temperature anomalies and thermocline variability. Tellus, 29, 289-305.

Giannini, A., R. Saravanan, and P. Chang, P., 2003: Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. Science, 302, 1027–1030.

Giannini, A., M. Biasutti, and M.M. Verstraete, 2008a: A climate model-based review of drought in the Sahel: Desertification, the re-greening and climate change. Global and Planetary Change, 64, 119-128.

Giannini, A., M. Biasutti, I.M. Held, and A.H. Sobel, 2008b: A global perspective on African climate. Climatic Change, 90, 359–383. doi:10.1007/s10584-008-9396-y.

Gu, Y., K. N. Liou, J. H. Jiang, J. Su, and X. Liu (2015): Dust aerosol impact on North African climate: a GCM investigation of aerosol-cloud-radiation interactions using A-Train satellite data, Atmos. Chem. Phys., 12, 1667-1679.

Hagos, S., and K.H. Cook, 2008: Ocean warming and late 20th century Sahel drought and recovery. J. Climate, doi:10.1175/2008JCLI2055.1.

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister, 2014, Updated high-resolution grids of monthly climatic observation – the CRU TS3.10, Int. J. Climatol., 34 (3), 623-642.

Hasselmann, K., 1976: Stochastic climate models. Part I: Theory. Tellus, 28, 473-485.

Hastenrath, S., D. Polzin, and C. Mutai, 2007: Diagnosing the 2005 drought in equatorial East Africa, J. Climate, 20 (18), 4628-4637.

He, Y., and E. Lee, 2016: Empirical relationships of Sea Surface Temperature and Vegetation activity with summer rainfall variability over the Sahel, Earth Interactions, 20-006.

Heidinger, A. K., M. J. Foster, A. Walther, and X. Zhao, 2014: The Pathfinder Atmospheres Extended (PATMOS-x) AVHRR climate dataset, Bull. Amer. Meteor. Soc., 95 (6), 909-922.

Held, I.M., T.L. Delworth, J. Lu, K. Findell, and T.R. Knutson, 2005: Simulation of Sahel drought in the 20th and 21st centuries. Proc. Nat. Acad. Sci., 102, 17891–17896.

Henderson-Sellers, A., K. McGuffie, and C. Gross, 1995: Sensitivity of global climate model simulations to increased stomatal resistance and CO2 increases. J. Climate, 8, 1738–1756.

Hocking, R. R., 1976: The analysis and selection of variables in linear regression, Biometrics, 32(1), 49.

Hoerling, M.P., J.W. Hurrell, J. Eischeid, and A. Phillips, 2006: Detection and attribution of twentieth-century northern and southern African rainfall change. J. Climate, 19, 3989-4008.

Hoerling, M.P., A. Kumar, and T. Xu, 2001: Robustness of the nonlinear climate response to ENSO's extreme phases. J. Climate, 14, 1277-1293.

Hoerling, M.P., A. Kumar, and M. Zhong, 1997: El Niño, La Niña, and the non-linearity of their teleconnections. J. Climate, 10, 1769-1786.

Hoffman, F. M. et al. 2008: The carbon-land model project (C-LAMP): a protocol and evaluation metrics for global terrestrial biogeochemistry models. In International Congress on Environmental Modeling and Software 4th biennial Meeting on Integrating Science and Information Technology for Environmental Assessment and Decision Making (eds Sanchez-Marre, M., J. Begar, J. Comas, A. Rizzoli, and G. Guariso), International Environmental Modeling and Software Society, Barcelona, Spain.

Huffman, G. J., 1997: Estimates of root-mean-square random error for finite samples of estimated precipitation, J. Appl. Meteor., 36, 1191-1201.

Hulme, M., 1996: Recent climatic change in the world's drylands. Geophys. Res. Lett., 23, 61–64.

Hulme, M., 1996: Recent climatic change in the world's drylands. Geophys. Res. Lett., 23, 61–64.

Hunke, E. C., W. H. Lipscomb, A. K. Turner, N. Jeffery, and S. Elliott, 2008: CICE: the Los Alamos sea ice model, documentation and software, version 4.0. Los Alamos National Laboratory Tech. Rep., LA-CC-06-012.

Hurrell, J. W., and coauthors, 2013: The Community Earth System Model: a framework for collaborative research, Bulletin of the American Meteorological Society, 94 (9), 1339-1360.

Indeje, M., F. H. M. Semazzi, L.J. Ogallo, 2000: ENSO signals in East African rainfall seasons, International Journal of Climatology, 20 (1), 19-46.

Janicot, S., V. Moron, and B. Fontaine, 1996: Sahel droughts and ENSO dynamics. Geophys. Res. Lett., 23, 515-518.

Janicot, S., et al., 2011: Intraseasonal variability of the West African monsoon. Atmos. Sci. Lett., 12, 58-66.

Katz, R.W., and M.H. Glantz, 1986: Anatomy of a rainfall index. Mon.Wea. Rev., 114, 764–771.

Kay, J. E., et al., 2015: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. Bull. Amer. Meteoro. Soc., 96(8). 1333-1349.

Kaufmann, R. K., D. I. Stern, R. B. Myneni, C. J. Tucker, D. Slayback, N. V. Shabanov, and J. Pinzon, 2003: The effect of vegetation on surface temperature: A statistical of NDVI and climate data. Geophys. Res. Lett., 30, 2147.

Kim, D., M. Chin, L. A. Remer, T. Diehl, H. Bian, H. Yu, M. E. Brown, and W. R. Stockwell, 2017: Role of surface wind and vegetation cover in multi-decadal variations of dust emission in the Sahara and Sahel. Atmospheric Environment 148, 282-296.

Kobayashi, S. et al., 2015: The JRA-55 Reanalysis: General specifications and basic characteristics. J. Meteor. Soc. Japan, 93 (1), 5-48.

Koster, R. D., et al., 2004: Regions of strong coupling between soil moisture and precipitation. Science, 305, 1138.

Koster, R. D., et al., 2006: GLACE: The global land-atmosphere coupling experiment. Part I: Overview, Journal of Hydrometeorology, 7, 590-610.

Koster, R. D., et al., 2010: The contribution of land surface initialization to subseasonal forecast skill: First results from the GLACE-2 project. Geophys. Res. Lett., 37, L02402.

Koster, R. D., et al., 2011: The second phase of the Global Land-Atmosphere Coupling Experiment: Soil moisture contribution to subseasonal forecast skill. J. Hydrometeor., 12, 805-822.

Kucharski, F., N. Zeng, and E. Kalnay, 2012: A further assessment of vegetation feedback on decadal Sahel rainfall variability. Clim. Dyn., doi:10.1007/s00382-012-1397-x.

Kutzbach, J. E., G. Bonan, J. Foley, and S. P. Harrison, 1996: Vegetation and soil feedbacks on the response of the African monsoon to orbital forcing in the early to middle Holocene. Nature, 384, 623-626.

Lamb, P.J., 1978: Case studies of tropical Atlantic surface circulation patterns during recent sub-Saharan weather anomalies: 1967 and 1968. Mon. Wea. Rev., 106, 482-491.

Lamb, P.J., 1982: Persistence of Subsaharan drought. Nature, 299, 46–48.

Lamb, P.J., and R.A. Peppler, 1992: Further case studies of tropical Atlantic surface atmospheric and oceanic patterns associated with sub-Saharan drought. J. Climate, 5, 476–488.

Lawrence, D. M., and coauthors, 2011: Parameterization improvements and functional and structural advances in version 4 of the Community Land Model, Journal of Advances in Modeling Earth Systems, 3, M03001.

Le Barbé, L., T. Lebel, and D. Tapsoba, 2002: Rainfall variability in West Africa during the years 1950–90. *J. Climate*, 15**,** 187–202.

Lee, E., Y. He, M. Zhou, and J. Liang, 2015: Potential feedback of recent vegetation changes on summer rainfall in the Sahel, Physical Geography, 36 (6), 449-470.

Liebmann, B., et al., 2014: Understanding recent eastern Horn of Africa rainfall variability and change, J. Climate, 27, 8630-8645.

Liu, Z., L. Fan, S.-I. Shin, and Q. Liu, 2012a: Assessing atmospheric response to surface forcing in the observations. Part II: Cross validation of seasonal response using GEFA and LIM. J. Climate, 25, 6817-6834.

Liu, Z., N. Wen, and L. Fan, 2012b: Assessing atmospheric response to surface forcing in the observations. Part I: Cross validation of annual response using GEFA, LIM, and FDT. J. Climate, 25, 6796-6816.

Liu, Q., L. Wang, Y. Qu, N. Liu, S. Liu, H. Tang, and S. Liang, 2013, Preliminary evaluation of the long-term GLASS albedo product, International Journal of Digital Earth, 6 (sup 1), 69-95.

Liu, Y. et al., 2013: Trend-preserving blending of passive and active microwave soil moisture retrieval, Remote Sensing of Environment, 123, 280-297.

Liu, Z., M. Notaro, J. Kutzbach, and N. Liu, 2006: Assessing global vegetation-climate feedbacks from observations. J. Climate, 19, 787-814.

Liu, Z., M. Notaro, and R. Gallimore, 2010: Indirect vegetation-soil moisture feedback with application to Holocene North Africa climate, Global Change Biology, 16, 1733-1743.

Liu, Z., N. Wen, and Y. Liu, 2008: On the assessment of nonlocal climate feedback. Part I: The Generalized Equilibrium Feedback Assessment. Journal of Climate, 21, 134-148.

Liu, Z. et al., 2007: Simulating the transient evolution and abrupt change of Northern African atmosphere-ocean-terrestrial ecosystem in the Holocene. Quaternary Science Reviews, 26, 1818-1837.

Los, L. O., G. P. Weedon, P. R. J. North, J. D. Kaduk, C. M. Taylor, and P. M. Cox, 2006: An observation-based estimate of the strength of rainfall-vegetation interactions in the Sahel. Geophy. Res. Lett., 33, L16402.

Lu, J., and T.L. Delworth, 2005: Oceanic forcing of the late 20th century Sahel drought. Geophysical Research Letters, 32, L22706, doi:10.1029/2005GL023316.

Matsuura, K. And C. J. William, 2012: 1900-2008 gridded monthly time series (version 2.01), [Available online at http://climate.geog.udel.edu/~climate/html_pages/Global2_Ts_2009/README.global_p_ts_2009.html.]

Mao, J., W. Fu, X. Shi, D. M. Ricciuto, J. B. Fisher, R. E. Dickinson, and C. R. Schwalm, 2015: Disentangling climatic and anthropogenic controls on global terrestrial evapotranspiration trend, Environmental Research Letters, 10 (9), 0094008.

Mason, S. J. and L. Goddard, 2001: Probabilistic precipitation anomalies associated with ENSO, Bull. Am. Meteorol. Soc., 82, 619-638.

Mu, Q., F. A. Heinsch, M. Zhao, and S. W. Running, 2007: Development of global evapotranspiration algorithm based on MODIS and global meteorology data, Remote Sensing of Environment, 111, 519-536.

National Research Council, 2004: Climate Data Records from environmental satellites: interim report. The National Academies Press, Washington, D. C., pp 136.

Nayar, A., 2012: African land grabs hinder sustainable development, Nature News, February 2012.

Neelin, J.D., C. Chou, and H. Su, 2003: Tropical drought regions in global warming and El Niño teleconnections. Geophys. Res. Lett., 30 (24), 2275. doi:10.1029/2003GL0018625.

Nicholson, S.E., 1978: Climatic variations in the Sahel and other African regions during the past five centuries. Journal of Arid Environments, 1, 3-24.

Nicholson, S.E., 1979: Revised rainfall series for the West African subtropics. Mon. Wea. Rev., 107, 620–623.

Nicholson, S.E., 1980: The nature of rainfall fluctuations in subtropical West Africa. Mon. Wea. Rev.,108, 473–487.

Nicholson, S.E., and D. Entekhabi, 1986: The quasi-periodic behavior of rainfall variability of Africa and its relationship to the Southern Oscillation. Arch. Meteorol. Geophys. Bioclimatol. Ser. A34, 311-348.

Notaro, M., G. Chen, and Z. Liu, 2011: Vegetation feedbacks to climate in the global monsoon regions. J. Climate, 24, 5740-5756.

Notaro, M., G. Chen, Y. Yu, F. Wang, and A. Tawfik, 2017: Regional climate modeling of vegetation feedbacks on the Asian-Australian monsoon systems. J. Climate, doi: 10.1175/JCLI-D-16-0669.1.

Notaro, M., and Z. Liu, 2008: Statistical and dynamical assessment of simulated vegetation feedbacks on climate over the boreal forests. Climate Dynamics, 31, 691-712.

Notaro, M., Z. Liu, and J. W. Williams, 2006: Observed vegetation-climate feedbacks in the United States. J. Climate, 19 (5), 763-786.

Notaro, M., Y. Wang, Z. Liu, R. Gallimore, and S. Levis, 2008: Combined statistical and dynamical assessment of simulated vegetation-rainfall interactions in North Africa during the mid-Holocene. Global Change Biology, 14, 347-368.

O'Brien, K. L., 1996: Tropical deforestation and climate change. Prog. Phys. Geogr., 20, 311–335.

Palmer, T.N., 1986: Influence of the Atlantic, Pacific and Indian Oceans on Sahel rainfall. Nature, 322, 251–253.

Park, J., J. Bader, and D. Matei, 2016: Anthropogenic Mediterranean warming essential driver for present and future Sahel rainfall, Nature Climate Change, 6, 941-945.

Pielke, R.A. Sr., R. Avissar, M. Raupach, A.J. Dolman, X. Zeng, and A.S. Denning, 1998: Interactions between the atmosphere and terrestrial ecosystems: Influence on weather and climate. Global Change Biology, 4, 461-475.

Pollard, D., and S.L. Thompson, 1995: The effect of doubling stomatal resistance in a global climate model. Global Planet. Change, 10, 1–4.

Rayner, N. A. et al., 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, Journal of Geophysical Research, 108, 4407.

Rienecker, M. M. et al., 2011, MERRA: NASA's Modern-Era Retrospective Analysis for Research and Application, Journal of Climate, 24, 3624-3648.

Robinson, D.A., and G. Kukla, 1985: Maximum surface albedo of seasonally snow covered lands in the Northern Hemisphere. J. Climate Appl. Meteor., 24, 402–411.

Roehrig, R., D. Bounioi, F. Guichard, F. Hourdin, and J.-L. Redelsperger, 2013: The present and future of the West African monsoon: A process-orientated assessment of CMIP5 simulations along the AMMS transect. J. Climate, 26, 6471-6504.

Romps, D. M., and Z. Kuang, 2010: Nature versus Nurture in shallow convection, Journal of Atmospheric Sciences, 67 (5), 1655-1666.

Rosenfeld, D., Y. Rudich, and R. Lahav, 2001: Desert dust suppressing precipitation: A possible desertification feedback loop, Proc. Natl. Acad. Sci., USA, 98, 5975-5980.

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. Q. J. R. Meteorol. Soc., 121, 669–704.

Rowell, D.P., 2003: The impact of Mediterranean SSTs on the Sahelian rainfall season, J. Clim., 16(5), 849-862.

Saha, S. et al., 2010: The NCEP Climate Forecast System Reanalysis, Bull. Amer. Meteor. Soc, 91 (8), 1015-1057.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean, Nature, 401, 360-363.

Santanello, J. A, C. D. Peters-Lidard, and S. V. Kumar, 2011: Diagnosing the sensitivity of local land-atmosphere coupling via the soil moisture-boundary layer interaction. J. Hydrometeor., 12, 766-786.

Scaife, A.A., et al., 2009: The CLIVAR C20C project: selected twentieth century climate events. Clim. Dyn., 33, 603–614.

Schneider, U., T. Fuchs, A. Meyer-Christoffer, B. Rudolf, 2008: Global precipitation analysis products of the GPCC, Global Precipitation Climatology Centre (GPCC), DWD, Internet Publication, 112.

Shukla, J., and Y. Mintz, 1982: Influence of land-surface evapotranspiration on the Earth's climate. Science, 215, 1498-1501.

Smith, R., and coauthors, 2010: The Parallel Ocean Program (POP) reference manual ocean component of the Community Climate System Model (CCSM) and Community Earth System Model (CESM). LAUR-01853.

Sud, Y.C., and W.E. Smith, 1985: The influence of surface roughness of deserts on the July circulation. Boundary-Layer Meteorology, 33, 15-49.

Sud, Y.C., J. Shukla, and Y. Mintz, 1988: Influence of land-surface roughness on atmospheric circulation and precipitation: A sensitivity study with a general circulation model. Journal of Applied Meteorology, 27, 1036-1054.

Sultan, B., and S. Janicot, 2000: Abrupt shift of the ITCZ over West Africa and intra-seasonal variability. *Geophys. Res. Lett.*, 27**,** 3353–3356.

Sun, S., and G. Wang, 2012: The complexity of using a feedback parameter to quantify the soil moisture-precipitation relationship. Journal of Geophysical Research, 117, D11113.

Taschetto, A. S., A. S. Gupta, N. C. Jourdain, A. Santoso, C. C. Ummenhofer, and M. H. England, 2014: Cold tongue and warm pool ENSO events in CMIP5: mean state and future projections. J. Climate, 27 (8), 2861-2885.

Taylor, C.M., E.F. Lambin, N. Stephenne, R.J. Harding, and R.L.H. Essery, 2002: The influence of land use change on climate in the Sahel. J. Clim., 15, 3615–3629.

United Nations Environmental Programme, 2002: Africa Environment Outlook. Past, present and future perspectives.

Wang, G., E.A.B. Eltahir, J.A. Foley, D. Pollard, and S. Levis, 2004: Decadal variability of rainfall in the Sahel: results from the coupled GENESIS-IBIS atmosphere-biosphere model. Clim. Dyn., 22, 625–637.

Wang, F., Z. Liu, and M. Notaro, 2013: Extracting the dominant SST modes impacting North America's observed climate. J. Climate, 26, 5434-5452.

Wang, F., M. Notaro, Z. Liu, and G. Chen, 2014: Observed local and remote influences of vegetation on the atmosphere across North America using a model-validated statistical technique that first excludes oceanic forcings. J. Climate, 27, 362-382.

Wang. F., Y. Yu, M. Notaro, J. Mao, X. Shi, and Y. Wei, 2017: Advancing a model-validated statistical method for decomposing the key oceanic drivers of regional climate: Focus on North African climate variability in CESM. In review.

Wen, N., Z. Liu, and Q. Liu, 2013: Observational assessment of nonlocal heat flux feedback in the North Atlantic by GEFA. Journal of Applied Meteorology and Climatology, 52, 645-653.

Wen, N., Z. Liu, Q.Y. Liu, and C. Frankignoul, 2010: Observed atmospheric responses to global SST variability modes: A unified assessment using GEFA. Journal of Climate, 23, 1739-1759.

Willeit, M., A. Ganopolski, and G. Feulner, 2013: Asymmetry and uncertainties in biogeophysical climate-vegetation feedback over a range of CO2 forcings. Biogeosciences Discuss., 10, 12967-13013.

Xie, S. P., and P. A. 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. Q. J. R. Meteorol. Soc., 123, 1483–1515.

Yang, J., Q. Liu, S. P. Xie, Z. Liu, and L. Wu, 2007: Impact of the Indian Ocean SST basin mode on the Asian summer monsoon, Geophysical Research Letters, 34.

Yin, L., R. Fu, Y. F. Zhang, P. A. Arias, D. N. Fernando, W. Li, K. Fernandes, and A. R. Bowerman, 2014: What controls the interannual variation of the wet season onsets over the Amazon? Journal of Geophysical Research: Atmospheres, 119, 2314-2328.

You, L., C. Ringler, U. Wood-Sichra, R. Robertson, S. Wood, T. Zhu, G. Nelson, Z. Guo, and Y. Sun, 2011: What is the irrigation potential for Africa? A combined bipphysical and socioeconomic approach, Food Policy, 36(6), 770-782.

Yu, Y., M. Notaro, Z. Liu, F. Wang, F. Alkolibi, E. Fadda, F. Bakhrjy, 2015: Climatic controls of the interannual to decadal variability in Saudi Arabian dust activity: Towards the development of a seasonal prediction model. *Journal of Geophysical Research-Atmospheres,*120(5), 1739-1758.

Yu, Y., M. Notaro, F. Wang, J. Mao, X. Shi, and Y. Wei, 2017(a): Validation of a statistical methodology for extracting vegetation feedbacks: focus on North African ecosystems in the Community Earth System Model. In review.

Yu, Y., M. Notaro, F. Wang, J. Mao, X. Shi, and Y. Wei, 2017(b): Observed vegetation-climate feedbacks in the Sahel: Is the classic albedo feedback mechanism truly dominant? In review.

Zeng, F. W., G. J. Collatz, J. E. Pinzon, and A. Ivanoff, 2013: Evaluating and quantifying the climate-driven interannual variability in Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) at global scales, Remote Sensing, 5 (8), 3918-2950.

Zeng, N., J. D. Neelin, K. –M. Lau, and C. J. Tucker, 1999: Enhancement of interdecal climate variability in the Sahel by vegetation interaction. Science, 286, 1537-1540.

Zheng, X.Y., and E.A.B. Eltahir, 1997: The response to deforestation and desertification in a model of West African monsoons. Geophys. Res. Lett., 24, 155–158.

Zhong, Y., and Z. Liu, 2008: A joint statistical and dynamical assessment of atmospheric response to North Pacific oceanic variability in CCSM3. Journal of Climate, 21, 6044-6051.

Zhong, Y.F., Z. Liu, and M. Notaro, 2011: A GEFA assessment of observed global ocean influence on US precipitation variability: Attribution to regional SST variability modes. Journal of Climate, 24, 693-707.

Zhou, L., R.K. Kaufmann, Y. Tian, R.B. Myneni, and C.J. Tucker, 2003: Relation between interannual variations in satellite measures of northern forest greenness and climate between 1982 and 1999. J. Geophys. Res., 108, 4004.

Zhu, Z., J. Bi, Y. Pan, S. Ganguly, A. Samanta, L. Xu, A. Anav, S. Piao, and R. R. Nemani, 2013: Photosynthetically active radiation (FPAR) 3g derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the period 1981-2011. Remote Sensing, 5, 927-948.