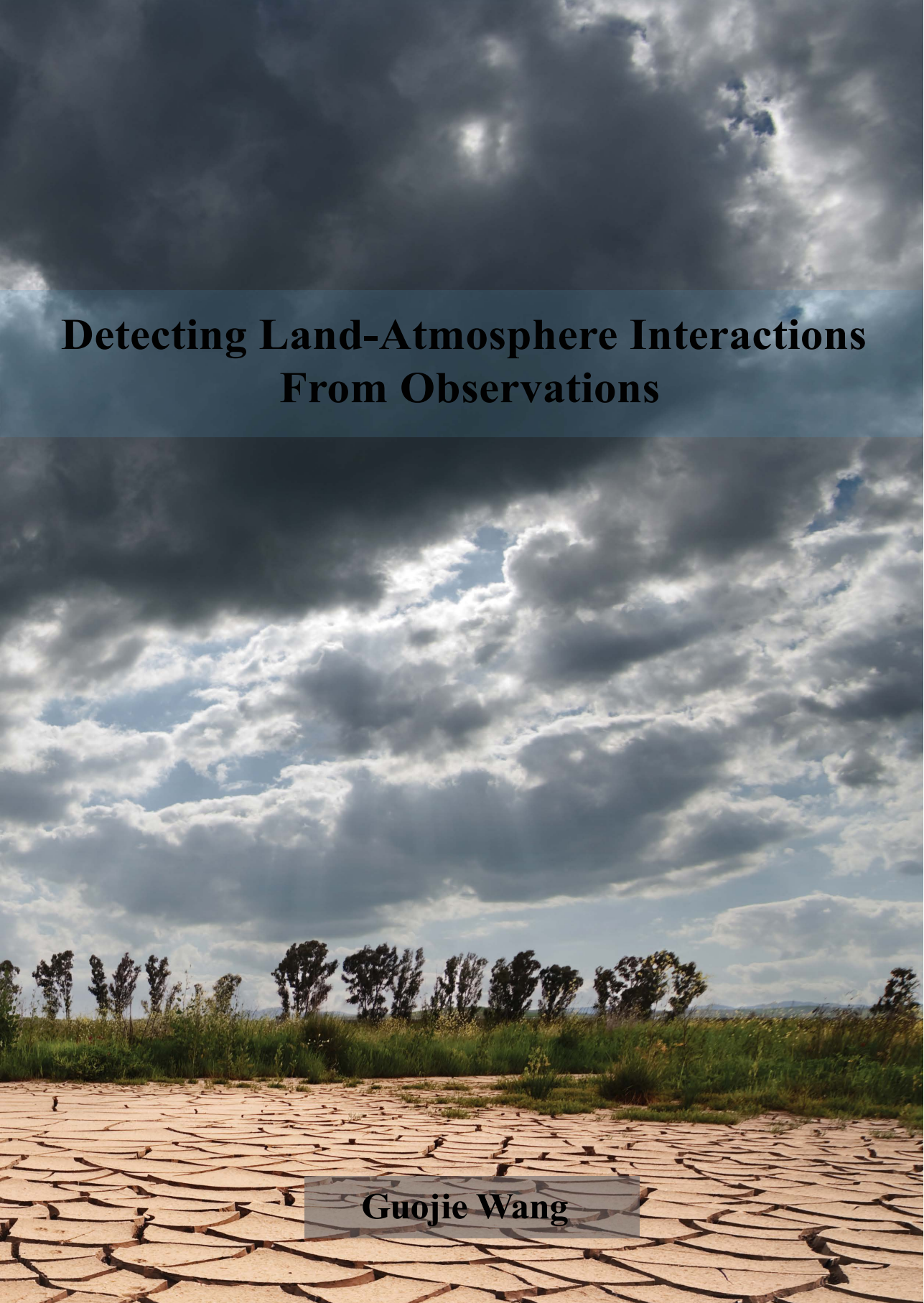

Detecting Land-Atmosphere Interactions From Observations

Guojie Wang

Detecting Land-Atmosphere Interactions from Observations

Guojie Wang

Detecting Land-Atmosphere Interactions from Observations
(PhD Thesis, VU University Amsterdam)
In dutch: Het bepalen van land-atmosfeer interacties uit waarnemingen
(Academisch Proefschrift, Vrije Universiteit Amsterdam)

This research was funded by the Netherlands Organisation for Scientific Research (NWO), entitled "the land-atmosphere coupling and large-scale drought over Europe" under grant nr. 854.00.026.

ISBN 9789086595839
Subject headings:
Soil moisture / hydrology / atmosphere / temperature / precipitation / interactions / climate / Europe

reading committee: prof. dr. Sonia Isabelle Seneviratne
prof. dr. Tong Jiang
dr. Christopher M. Taylor
dr. A.G.C.A. (Antoon) Meesters
dr. Geert Jan van Oldenborgh

VRIJE UNIVERSITEIT

Detecting Land-Atmosphere Interactions from Observations

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Aard- en Levenswetenschappen
op woensdag 30 november 2011 om 9.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Guojie Wang

geboren te Shandong, China

promotor: prof. dr. A.J. Dolman
prof. dr. van den Hurk
copromotor: dr. R.A.M. de Jeu

Contents

Chapter 1

Introduction

1.1 Research Background

Mediterranean Europe has suffered from an increasing frequency of summer droughts and heat waves in recent decades that have had substantial societical and ecological impacts. Hot and dry summers in Europe are generally associated with a specific large-scale anticyclonic regime (*Cassou et al*., 2005). However, model simulations have suggested that the extreme conditions of droughts and heats may be enhanced when soil moisture is depleted (*Seneviratne et al*., 2010; *Seneviratne et al*., 2006; *Teuling et al*., 2010). Soil moisture partitions available net radiation into latent heat for evaporation and sensible heat for temperature. This partitioning thus directly affects the moisture and heat conditions of the lower atmosphere. The depletion of soil moisture results in reduced evaporation and eventually precipitation. The reduced latent cooling amplifies the temperature and the resulting increased temperature may increase the atmospheric demand of evaporation, leading to accelerated soil moisture depletion. Meanwhile, there is also possibility that sufficient soil moisture and thus increased evaporation stabilizes the boundary layer of atmosphere, thus forming a negative feedback on precipitation at particularly local scales (*Seneviratne et al*., 2010). These interactions between land and atmosphere are sketched in Figure 1.1. Among these interactions, the soil moisture influences on precipitation and temperature are of particular interest. From improved understanding of these feedbacks, better prediction of summer droughts and heat waves is expected.

Up to now, our understanding of the soil moisture-atmosphere interaction relies heavily on numerical experiments. However, large unexplained discrepancies exist among model responses, regarding both the location and magnitude of the land-atmosphere interactions (*Dirmeyer et al*., 2006; *Koster et al*., 2004). There is thus an increasing need to investigate soil moisture-atmosphere interactions with observational datasets to benchmark the models. In recent years, the increasing volume and diversity of climate datasets, especially the soil moisture datasets from satellite images, has created such an opportunity (*Dolman and de Jeu*, 2010; *Owe et al*., 2008; *de Jeu et al*., 2008). However, so far, evidence of soil moisture influence on precipitation and temperature from observations remains rare.

The influences of land surface processes on atmosphere are generally subtle, compared to other influences such as sea surface temperatures. The expected signals are often very hard to detect behind the strong background noise. Under such circumstances, the ordinary methods applied in climatology such as correlation or regression analysis may become inefficient. The failure of such ordinary methods can be easily seen from an example. Suppose we have two identical time series, which are unit correlation owing to a certain physical process. We perturb each time series by adding different noise to them until this perturbation leads to a reduced correlation, e.g., say a coefficient of 0.2. This very low correlation tells a story of "no relation", which is false as they are already known to have a certain linkage. Thus better solutions are needed to find the signal.

The case of land-atmosphere interactions is even more complicated. As illustrated in Figure 1.1, soil moisture and atmosphere exhibit often two-way interactions. In climate models, the direction of the influences, in a strict sense of causality, can be seen from the responses by perturbing the initial states. However, the directional soil moisture influences on atmosphere are very hard to determine with confidence, because the atmospheric processes also exert influences on soil moisture the other way around. For example, precipitation wets soil; meanwhile, increased soil moisture may promote precipitation. One more example, soil moisture depletion heats the atmosphere; while the increased atmosphere temperature depletes soil moisture in return. A correlation, even lagged,

can say tell little about the soil moisture influence on precipitation or atmosphere temperature.

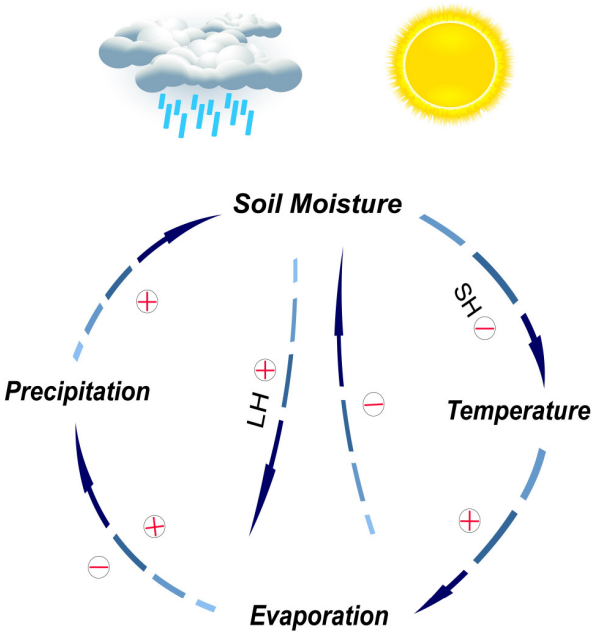

Figure 1.1: The hypothesized interactions between land and atmosphere. Red cross in circle indicates positive feedback and red dash in circle indicates negative feedback.

A correlation or regression analysis measures only the interdependence, and does not identify the statistical causality, that is, the statistical driver-response relationships. For sake of understanding the mechanisms of an interacting system, the latter is arguably more important. In this thesis, two sophisticated methods are used to detect the possible influences of land surface processes on overlaying atmosphere, i.e., the coupled (forced) manifold (*Navarra* et al., 2005) and the wavelet-based Granger causality (*Dhamala* et al., 2008). The coupled (forced) manifold offers a generalization of methods for variance analysis, such as the singular value decomposition (SVD) or canonical correlation analysis (CCA). Although defined in a loose way, it has the potential to detect the directed influences between climate fields by sophisticated data decomposition. The Granger causality has a strict definition of statistical causality based

on linear prediction theory (*Granger*., 1969). However, it should be noted as a caution that statistical results can not always be interpreted as directly reflecting the physics. Statistics can only be used to infer the possible physical mechanisms by combining it with the physical knowledge that already exists.

1.2 Thesis outline

The aim of this thesis is to assess the influence of soil moisture on the atmosphere over Europe from observational datasets, using a variety of rigorous statistics. The flow diagram of the content of this thesis is shown in Figure 1.2. This thesis consists of three parts, with respect to the time scales of interest and the statistical approaches employed. Soil moisture is an important agent in the land-atmosphere interactions. The memory scale of soil moisture is indicative of, at least partly, the persistence of the soil moisture feedback on atmosphere. The first part of this thesis, namely **Chapter 2**, deals with the memory scales as well as the statistical properties of soil moisture, which serve as the *priori* guiding the following signal interpretation and statistical modeling.

The second part consists of **Chapter 3-4**, where a rigorous spatio-temporal modeling approach, the coupled manifold, is used to deal with climate variability at interannual and interdecadal time scales. In **Chapter 3**, the functional relations between later winter soil moisture, proxied by accumulated precipitation, and the subsequent summer climate are investigated to understand if soil moisture can contribute to the interannual variability of summer climate. **Chapter 4** seeks for the causes of the marked variability of summer temperature at multidecadal time scales over Europe, suggesting a possible amplification of this signal by land surface processes.

The third part consists of **Chapter 5-6**. High-resolution daily datasets from satellite images and gridded in-situ observations are utilized to study the direction of the interactions between soil moisture and atmosphere (including precipitation and temperature) based on the statistical notion of Granger causality. The soil moisture-atmosphere interactions are expected to be nonstationary (time-varying) and occur in

mainly warm seasons; therefore this part focuses on the seasonal variability of the interactions. As a step of data preprocessing, the missing values of satellite soil moisture are estimated with a thin-spline smoother in **Chapter 5**. Alternatively, the Granger causality can be estimated with parametric or nonparametric approaches. **Chapter 6** uses a nonparametric approach based on spectral factorization of wavelet transforms to produce a time-frequency representation of the direction of the interactions between soil moisture and precipitation over Europe.

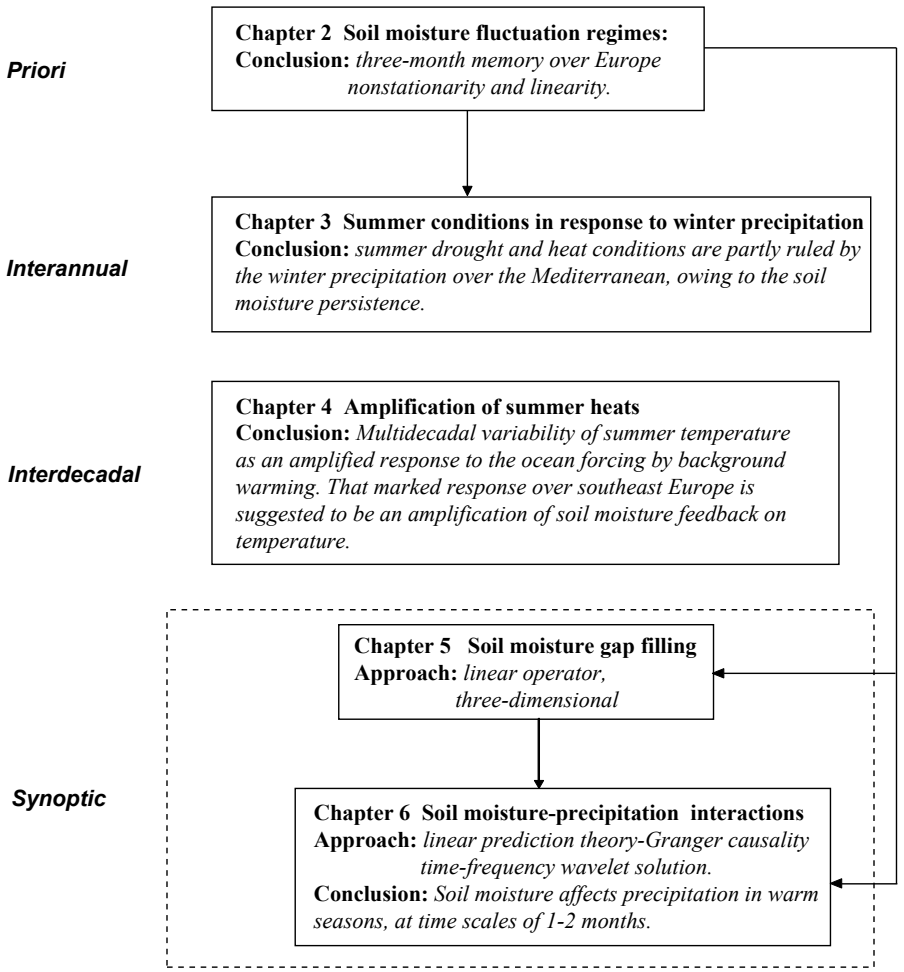

Figure 1.2: Flow diagram of the contents of this thesis, showing the interrelations between the chapters.

References

Cassou, C., L. Terray, and A. S. Phillips (2005), Tropical Atlantic influence on European heat waves, *J Climate*, *18*(15), 2805-2811.

Dhamala, M., G. Rangarajan, and M. Ding (2008), Estimating granger causality from fourier and wavelet transforms of time series data, *Phys Rev Lett*, *100*(1), 018701.

de Jeu, R. A. M. et al. Global Soil Moisture Patterns Observed by Space Borne Microwave Radiometers and Scatterometers (2008), *Surv Geophys 29*, 399-420.

Dirmeyer, P. A., Koster, R. D. & Guo, Z. C (2006), Do global models properly represent the feedback between land and atmosphere? *J Hydrometeorol*, 7, 1177-1198.

Dolman, A. J., and R. A. M. de Jeu (2010), Evaporation in focus, *Nat Geosci*, 3(5), 296-296.

Granger, C. W. J. (1969), Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, *37*(3), 414-438.

Koster, R. D. et al (2004), Regions of strong coupling between soil moisture and precipitation. *Science*, 305, 1138-1140.

Navarra, A., and J. Tribbia (2005), The coupled manifold, *Journal of the Atmospheric Sciences*, *62*(2), 310-330.

Owe, M., R. de Jeu, and T. Holmes (2008), Multisensor historical climatology of satellite-derived global land surface moisture, *J Geophys Res-Earth*, *113*, F01002, doi:10.1029/2007JF000769.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, *Earth-Sci Rev*, *99*(3-4), 125-161.

Seneviratne, S. I., Luthi, D., Litschi, M. & Schar, C (2006), Land-atmosphere coupling and climate change in Europe. *Nature 443*, 205-209.

Teuling, A. J., et al. (2010), Contrasting response of European forest and grassland energy exchange to heatwaves, *Nat Geosci*, *3*(10), 722-727.

Chapter 2

Soil Moisture Fluctuation Regimes in ERA-40 Re-analysis Data[1]

Abstract

Soil moisture variability is analyzed in the re-analysis data ERA-40 of the ECMWF (European Centre for Medium-Range Weather Forecasts) which includes four layers within 289 cm depth. Short-term correlations are characterized by an e-folding time scale assuming an exponential decay while the long-term memory is described by power-law decays with exponents determined by Detrended Fluctuation Analysis (DFA). On a global scale the short-term variability varies congruent with long-term memory in the surface layer. Key climatic regions (Europe, Amazon, and Sahara) reveal that soil moisture time series are nonstationary in arid regions and in deep layers within the time horizon of ERA40.

2.1 Introduction

Soil moisture has a key role in the hydrological and the energy cycles as well as the carbon cycle. The influence of soil moisture on precipitation and surface temperature has long been noticed and now has been drawing renewed attention in the recent years (e.g., *Delworth and Manabe*, 1988; *Hong and Kalnay*, 2000; *Koster and Suarez*, 2003; *Conil et al*., 2008).

[1]The content of this chapter has been published as Wang, G. J., Dolman, A. J., Blender, R. & Fraedrich, K (2010), Fluctuation regimes of soil moisture in ERA-40 re-analysis data. *Theor Appl Climatol*, **99**, 1-8.

The memory of soil moisture is particularly important to the seasonal prediction of precipitation, temperature, and other atmospheric variables, and various modeling studies (see for example *Dirmeyer et al*., 2000) have shown that there is predictability based on anomalies in land surface moisture. *Delworth and Manabe* (1988) pioneered the study of the temporal variability of soil moisture in the Geophysical Fluid Dynamics Laboratory (GFDL) GCM, and suggested soil moisture variations can be considered as a first-order Markov process. The autocorrelation function of the process decays exponentially

$$r(\tau) = \exp(-\lambda\tau) \tag{2.1}$$

where τ is the time lag and λ^{-1} is the *e*-folding time of the anomaly correlation in the absence of forcing. The e-folding time is generally referred to as the temporal scale of soil moisture, at which the soil integrates precipitation into a red soil moisture process. Here, time scales are considered for which precipitation is uncorrelated and spectrally white. The *e*-folding time was intensively studied from either observations (*Entin et al*., 2000*) or AGCMs (Atmosphere General Circulation Models*, *Wu and Dickinson* 2004) since soil moisture may have the ability to enhance extreme climate events, such as dry or wet spells in warm seasons.

Manabe and Delworth (1990) suggested that soil moisture has the potential to contribute substantially to low frequency atmospheric variability, as approximately half of the total variance of soil moisture process resides beyond a long 'separation time scale' in a first order Markov model (defined as $2\pi\lambda^{-1}$) implying a long-range correlation. Little work has been done to address this issue in more details. *Amenu et al*. (2005) reported 17-month, 34-month and 60-month low frequency modes in the Illinois soil moisture observations corresponding to ENSO signals; this gave an observational support to longer-term correlation beyond the classic *e*-folding time. Several studies have revealed that geophysical variables exhibit long-term correlations (equivalent to enhanced low frequency variability) (*Mandelbrot et al*., 1968), such as runoff *(Livina et al*., 2003; *Mudelsee et al*., 2007; *Wang et al*., 2008), and near surface temperature (*Fraedrich and Blender*, 2003; *Koscielny-Bunde et al*., 1998). However, as the major driver of hydrological cycle,

precipitation is spectrally white (*Kantelhardt et al*., 2006); *Blender and Fraedrich* (2006) suggest soil moisture as the major source of memory in hydrological cycle over land. Observed long-term correlations do not follow exponential decays but follow power-laws with exponents quantifying the strength of these correlations. The long-term memory of soil moisture may be relevant for the clustering of soil moisture deficiencies (*D'Andrea et al*., 2006) and the occurrence of heatwaves (*Seneviratne et al*., 2006).

Hasselmann (1976) introduced the concept of stochastic climate modes, and stated that climate variability in principle can be modeled by autoregressive (AR) processes. Mitchell (1964) also pointed out that '…persistence in meteorological data can ordinarily be described very well by a first order linear Markov model'. Based on this concept, *Delworth and Manabe* (1988) advanced the *e*-folding time model of soil moisture memory. However, many climatic variables behave nonlinearly, e.g., temperature (*Bartos and Jánosi*, 2006), suggesting that linear models cannot fully capture the statistical properties of such processes. To fully characterize the statistical properties of soil moisture and for a better understanding of the underlying dynamics, it is necessary to determine the degree of nonlinearity in soil moisture process time series. Linearity may be defined using the Fourier phases of the time series: if the statistical properties do not depend on the Fourier phases (this is calculated by randomly shuffling the phases), the time series is linear, and otherwise the series is considered to be nonlinear. This definition includes linear AR processes that are used by *Delworth and Manabe* (1988).

Ashkenazy et al. (2003) suggested a method to assess the nonlinearity in geophysical time series based on an analysis of the volatility time series, which is given by the absolute values of the increments. The main observation is that time series with long-term memory based on linear processes are characterized by an absence of long-term memory in the volatility time series, whereas for nonlinear processes the long-term memory in the volatility is preserved. Volatility correlation is found for example in river fluxes (*Livina et al., 2003*), and land surface temperatures (*Bartos and Jánosi*, 2006; *Govindan et al*., 2003). Similar properties of temperature are detected on in proxy records (*Ashkenazy et*

al., 2003) and in the abyssal equatorial Pacific (*Kalisky et al*., 2005). These studies suggest considerable nonlinearity and 'clustering' of magnitudes in these geophysical variables, that is, a large magnitude tends to follow a large magnitude while small magnitudes follow small volatilities.

The aim of this paper is to determine the temporal correlation properties of soil moisture on short and long time scales and to find possible relationships between both regimes. The degree of nonlinearity is addressed by a long-term memory analysis of the volatility time series. Since global high quality observations of soil moisture are sparse in space and time, we use the ECMWF re-analysis product ERA-40 in 1957 to 2002. This paper is organized as follows: in Section 2.2 the dataset and the analysis methods are described. In Section 2.3 global results on the short and long term memory properties are presented and Section 2.4 includes a nonlinearity analysis based on the volatility correlation properties. Section 2.5 concludes with a brief summary and discussion.

2.2 Data and methods

The variability analysis of soil moisture is based on global daily fields in the ECMWF re-analysis product ERA-40. To determine the short- and long-term variability correlations (which are related to memory) we use the Detrended Fluctuation Analysis (DFA). Nonlinearity of the underlying processes is determined by an analysis of the volatility of the soil moisture time series.

2.2.1 ERA-40 re-analysis data

We use land surface soil moisture data in the ERA-40 global re-analysis, produced by the European Centre for Medium-Range Weather Forecasts (ECMWF, *Uppala et al*., 2005). The land surface parameterization of ERA-40 (*van den Hurk et al*., 2000) models the soil-atmosphere and soil-vegetation interactions and delivers a daily surface water and energy balance at each grid cell on 4 prognostic layers for soil moisture with layer thicknesses of 7 cm, 21 cm, 72 cm and 189 cm during the entire period of 1957 to 2002. Both daily data and monthly averages are used in

this study. The seasonal cycle is removed at each grid point by subtracting the respective monthly and daily climatological means. Trends are not removed.

2.2.2 Detrended fluctuation analysis and power spectra

The Detrended Fluctuation Analysis (DFA, *Peng et al*., 1994) is a spectral method developed to detect long-term memory in stationary time series. First, the anomaly time series are determined by removing the climatological means from the original time series. As the first step of DFA, the anomaly series are integrated to the so-called profile. To determine the fluctuation function $F(\tau)$, the profile time series is partitioned into segments of duration τ, and linear fits are calculated separately for each segment. The fluctuation function $F(\tau)$ is the mean of the variances of the profile with respect to fits at the time scale τ. To obtain robust estimates overlapping windows are used. When the original time series shows polynomial trend of order k-1, polynomials of order k are fitted and subtracted in the segments (denoted as DFA*k*, note that the above mentioned DFA does not eliminate trends, *Bunde et al*., 2000). In this work, DFA is determined using the software matlab. Both DFA1 and DFA2 are performed in this work, and only the DFA1 results are reported since DFA2 produces the same results.

In the case of power-law in power spectrum, $S(f) \sim f^{-\beta}$, the fluctuation function $F(\tau)$ obeys a power-law, $F(\tau) \sim \tau^{\alpha}$, where α is the DFA exponent. This exponent α can be determined by the slope relating log $F(\tau)$ to log τ. The exponents are related by $\beta = 2\alpha - 1$. A long-term memory process is characterized by fluctuation exponents $\alpha > 0.5$ ($\beta > 0$, for low frequencies). An uncorrelated process (white noise) is given for $\alpha = 0.5$ ($\beta = 0$), and an anti-persistent process has $\alpha < 0.5$ ($\beta < 0$). Specifically, $\alpha = 1.5$ ($\beta = 2$) corresponds to Brownian noise, which can be regarded as the integration of white noise. Stationarity is violated for $\alpha > 1$ ($\beta > 1$), the threshold being given by $1/f$ (or flicker) noise.

2.2.3 Volatility analysis and nonlinearity

An empirical relationship between nonlinearity and volatility series was suggested by *Ashkenazy et al*. (2001). For a given a time series *u(i)* with

increments $\Delta u(i) = u(i+1)-u(i)$, the volatility series is defined as the absolute value of increments $vol(i) = |\Delta u(i)|$. It was found that long-range correlated linear series have uncorrelated volatility series, while long-range correlated nonlinear series have correlated volatility series. The detection of nonlinearity in time series is rather involved and requires long time series.

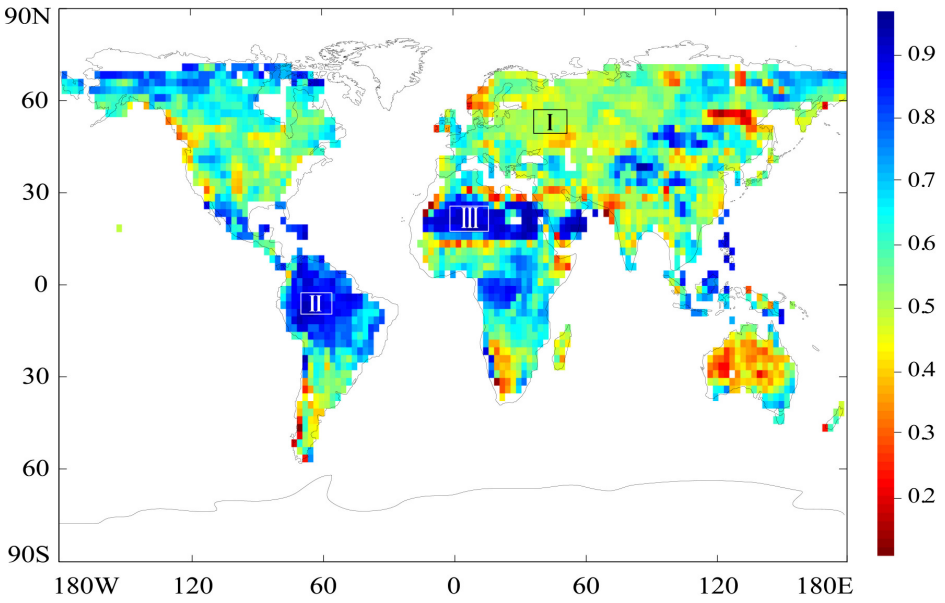

Figure 2.1 One-month-lag autocorrelation coefficients of monthly soil moisture anomalies in the top surface layer with 7 cm depth.

2.3 Temporal variability in soil moisture layers

The e-folding time is used to estimate the memory time scale of soil moisture assuming an exponential decay of the autocorrelation function of the time series. Using Eq. (2.1) we use the one-month lag autocorrelation values $r = 0.8$, 0.6, 0.4, and 0.2 to determine the e-folding times λ^{-1} of 4.5, 2.0, 1.1 and 0.6 months, respectively. The one-month lag autocorrelation of global soil moisture for the top surface layer (7 cm depth) in ERA40 data is shown in Figure 2.1. This result is based on monthly anomalies at each grid cell which are calculated by removing the

monthly climatological mean in the monthly time series. The substantial spatial variability of the e-folding times in the top surface layer show is consistent with previous studies (e.g. *Delworth and Manabe*, 1988; *Wu and Dickinson*, 2004). We find short values of the *e*-folding times in the Tropics and an increase with latitude as well as relatively higher values in arid and broadleaf forest regions. The processes involved in the dynamics of soil moisture (mainly precipitation, evapotranspiration and runoff) contribute to the variability, but the precise mechanisms are model dependent, and still not clearly understood (*Delworth and Manabe*, 1988; *Wu and Dickinson*, 2004).

To examine the fluctuation regimes, three regions are selected for the DFA: a mid-latitude region in Europe marked as region I with an *e*-folding time of 2-3 months, an evergreen broadleaf forest region marked as region II and an arid region marked as region III, both with an *e*-folding time of approximately one year. These three regions represent different types of fluctuations in global soil moisture. In these analyses, daily soil moisture data are used. First the daily values are averaged in each of the four layers for the whole re-analysis period, in each selected regions (this yields 12 time series in total). The DFA is performed to the anomalies obtained by removing the climatology means from daily soil moisture time series. The DFA in the three regions I, II, and III shows the following results:

Region I (Europe): A strong seasonality is present in the top three layers, while it is weak in the bottom layer (not shown). The DFA fluctuation functions show two distinct power-law regimes (Figure 2.2-I) with a crossover at around 3 months. The regimes corresponding to power spectra $S(f) \sim f^{-\beta}$, based on the relationship $\beta = 2\alpha - 1$ between the spectral exponent β and the fluctuation exponent α. In the first regime, the DFA fluctuation exponents in 1-90 days increase according $\alpha \approx 1.42$, 1.54, 1.70, and 1.82 from surface to bottom layer. Thus, in the two upper layers, the first regime (for shorter time) shows $\beta \approx 2$ what is consistent with the e-folding time scale obtained within the Markovian framework. The exact value $\beta = 2$ is found in the high frequency limit of the Lorentzian power spectrum $S(f) \sim 1/(\lambda^2 + f^2)$ which is obtained for an exponential decay, $\exp(-\lambda\tau)$, of the auto-correlation function. In the deepest layer $\alpha \approx 1.8$ is found which is related to $\beta \approx 2.6$; such 'redder'

spectra have been found in previous studies (*Wu and Dickinson*, 2004). The increase of α and β captures 'the redder spectra in deeper layers' related to an increasing memory. In the second regime, beyond 3 months, the exponents of the fluctuation functions of the four layers converge to the same value $\alpha \approx 0.8$ ($\beta \approx 0.6$). At such long time scales, the exchange between the soil layers is complete and the variability is coherent. Note that values $0<\beta<1$ indicate stationary long-term memory time series.

Region II (Amazon): This broadleaf forest region shows intense the seasonality of soil moisture in all layers. The crossover time scales is higher than in region I (Europe) and reaches roughly one year (Figure 2.2-II). Below one year the fluctuation function exponents α, in 1-300 days increase with depth, α=1.20, 1.26, 1.42, and 1.65 from top to bottom. Hence, the power spectra $S(f) \sim f^{-\beta}$ scale with the exponents β=1.4, 1.52, 1.84, and 2.3 and the increase of the memory with depth is similar to Europe. However, the short term memory is slightly weaker in this region than in Europe. In the long term regime above one year, α converges to α=1.28 (β=1.56). This value indicates nonstationarity of the anomaly time series up to the maximum time scale given by the duration of the ERA40 data. Therefore, averages determined in this data should be considered carefully; for climatological means longer time periods are necessary.

Region III (Sahara): In this region, seasonality is present only in the top two layers, and throughout the year the top layer is wetter than the second layer. Besides the rare rainfall events evaporation is the dominant process in the dynamics of soil moisture. Due to the small amounts involved and due to subsequent fast evaporation rapidly varying precipitation does not reach the deeper layers. The crossover times extend up to one year in the top two layers (Figure 2.2-III), beyond which DFA fluctuations converge to α=1.54 (β=2.08) determined in 1-300 days. The two bottom layers, which are not impacted by the high frequency components of the precipitation variability, show a unique variability for all time scales, with the same exponent, which is the limit of the two top layers for large time scales. Thus all spectra are red (β=2) in the whole frequency range accessible in the data set. The nonstationarity involved is even more vigorous than in the Amazon region.

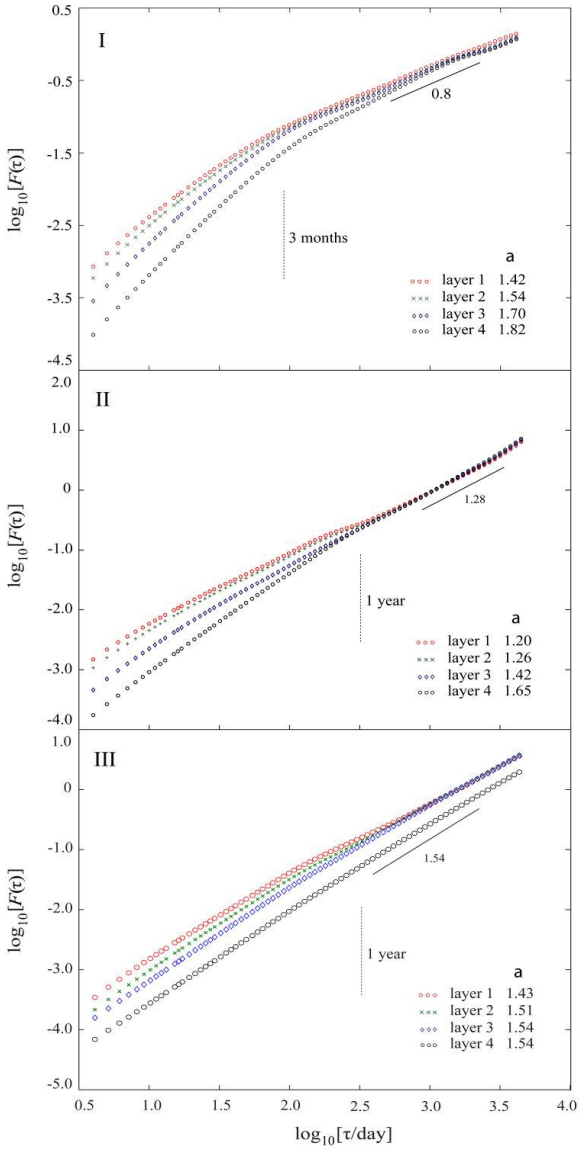

Figure 2.2 Log-log plots of DFA fluctuation functions $F(\tau) \sim \tau^{\alpha}$ from daily soil moisture anomalies in the three regions (I: Europe, II: Amazon, and III: Sahara) and the four layers from top to bottom (7 cm (●), 21 cm (x), 72 cm (◊), and 189 cm (○)). Crossover time scales are indicated by a vertical bar. The exponents α for short time scales are indicated (determined in 1-90 days in I and in 1-300 days II, III).

A global view of the interannual long-term memory determined for the top 7 cm-layer by a fit of the fluctuation exponent α in the time interval of 2-10 years (Figure 2.3). Since all of the four layers at each location have identical long-term correlation properties (see Figure 2.2), this map shows the long-term memory of the total soil moisture content. The similar spatial structure of the autocorrelation coefficients in Fig. 1 and the fluctuation exponents Figure 2.3 suggesting a relationship between the short- and the long-range correlations. Such a relationship is also found in the land surface temperature anomalies (*Kiraly* et al., 2006). The physical mechanisms leading to this interdependence need to be further clarified, possibly by simulations with coupled land-atmosphere climate models.

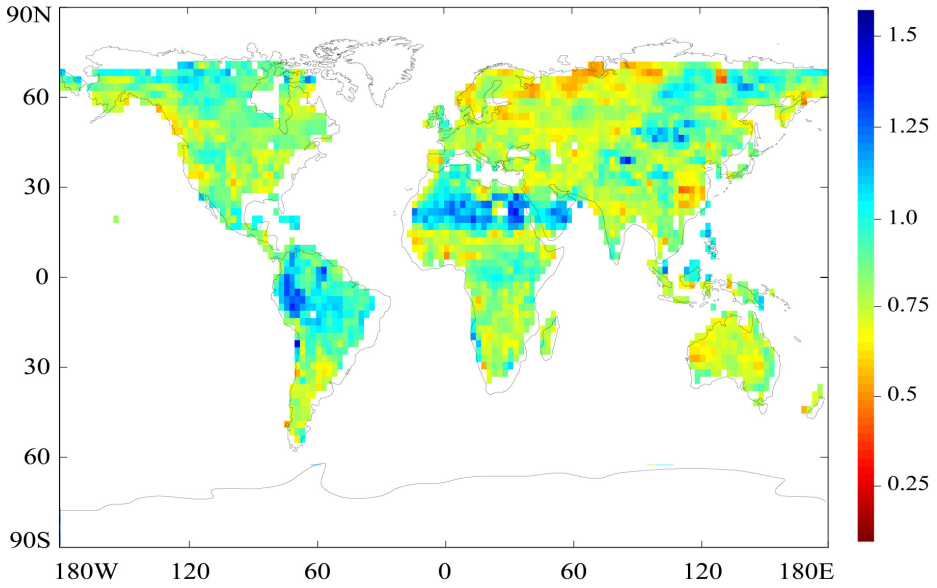

Figure 2.3 Long-term fluctuation exponent α of soil moisture variability determined in the top layer (7 cm) within 2-10 years.

2.4 Volatility correlations

Nonlinear geophysical processes can lead to a clustering of volatility which appears mainly as seasonality and long-range correlations (see Section 2.2.3). To determine the degree of nonlinearity involved in the dynamics of soil moisture, a long-term memory analysis is applied to the volatility time series $vol(i) = |u(i+1) - u(i)|$ (*Ashkenazy* et al., 2001). In the volatility time series the seasonal cycle is removed. The main observation is that time series with long-term memory based on linear processes are characterized by an absence of long-term memory in the volatility time series, whereas for nonlinear processes the long-term memory in the volatility is preserved. The numerical differentiation used to produce volatility series is known to strongly enhance the noise level inherent in the data (*Bartos and Jánosi*, 2006), and the typical volatility of fluctuations around local trends is extremely small, especially in arid regions, leading to noisy results.

In order to substantiate the analysis we apply a further test which is based on surrogate data for the soil moisture time series where the nonlinearity is destroyed while other statistical properties are preserved. To produce such surrogate data, *Schreiber and Schmitz* (2000) suggest the iterative amplitude-adjusted Fourier transform (iAAFT) method which preserves both the power spectrum and probability distribution of the series. The method proceeds as follows: (i) first the sequence of the original time series is shuffled, (ii) the shuffled series is Fourier-transformed, (iii) the power spectrum is adjusted to the power spectrum of the original series, and (iv) the inverse transform is applied to adjust the histogram to the histogram of the original series. Steps (ii)-(iv) are repeated until the result converges. This surrogate series has random Fourier phases and the nonlinearities stored in the phases are destroyed. By means of comparing the correlation properties of volatility series obtained from the original increment soil moisture series and the surrogate series, we can conclude whether the correlation in the volatility series is an indication of nonlinearity.

Before we perform a global analysis we consider the volatility anomalies series in the first two regions, Europe and Amazon (marked as I and II in

Figure 2.1). The long-term memory is determined by DFA (see Figure 2.4). The results of the analysis of the surrogate data and the volatility reveal (see Table 2.1):

(i) The soil moisture time series show long-term memory (α=0.8 in Europe, α=1.28 in the Amazon region).

(ii) The surrogate data (with nonlinearities eliminated) show the same long-term memory (this confirms the application of the iAAFT-method (*Schreiber and Schmitz, 2000*).

(iii) The volatility time series of soil moisture and of the surrogate time series show no long-term memory (α=0.5).

LTM Data	observations: nonlinear?		surrogate data: linearized	
Volatility	LTM	no LTM	LTM	no LTM
Result	nonlinear	<u>linear</u>	−	linear

Table 2.1 The table shows the decision process for the assessment of nonlinearity in LTM (long-term memory) data by volatility analysis and the comparison with linear surrogate data (the result in the present analysis is underlined).

The global distribution of the fluctuation exponent reveals no clear spatial pattern (not shown), therefore, we present scatter plots of the fluctuations exponents α obtained in the data grid points (Fig. 2.5). This figure confirms that the surrogate data has the same long-term memory as the soil moisture time series (see $\alpha_{original}$ vs. α_{surro}) with a wide distribution ranging from 0.3 to 1.8. The long-term memory is lost in the transformation to the volatility time series ($\alpha_{original}$ vs. α_{vol}) leading to the conclusion of linearity. The exponents of the volatility time series, α_{vol} vs. α_{surro_vol}, show a narrow distribution in α=0.4 to 0.6 centered at 0.5. Therefore, we conclude that the soil moisture time series is based on a linear physical process; the long-term memory is the same as in the linear surrogate data.

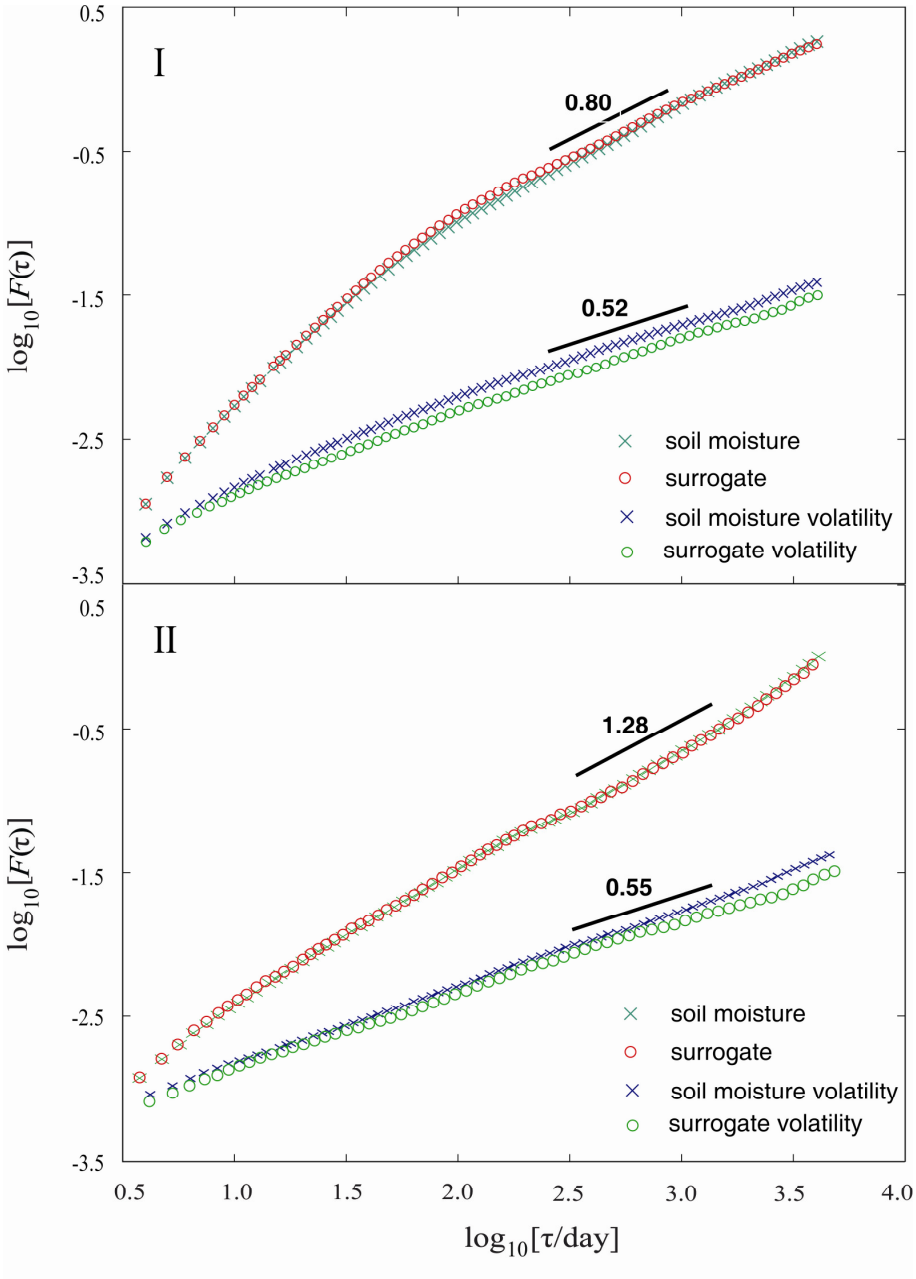

Figure 2.4 Volatility correlations for soil moisture and surrogate data for the regions in Fig. 2.1 (mid-latitude region I: Europe, broad leaf forest region II: Amazon)

Since the link between nonlinearity of the time series and the long-range volatility correlation is empirical (*Ashkenazy* et al., 2001), *Kalisky* et al. (2005) studied theoretical predictions of the relationship between the correlation exponent of a time series and its volatility. For a linear process obeying a power law, up to the value α=0.75, the volatility exponent is practically constant α_{vol} = 0.5, and then changes to an approximately linear increase. Unfortunately, this result is not confirmed in our analysis (see $\alpha_{original}$ vs. α_{vol}, for $\alpha_{original}$>0.75 in Figure 2.5).

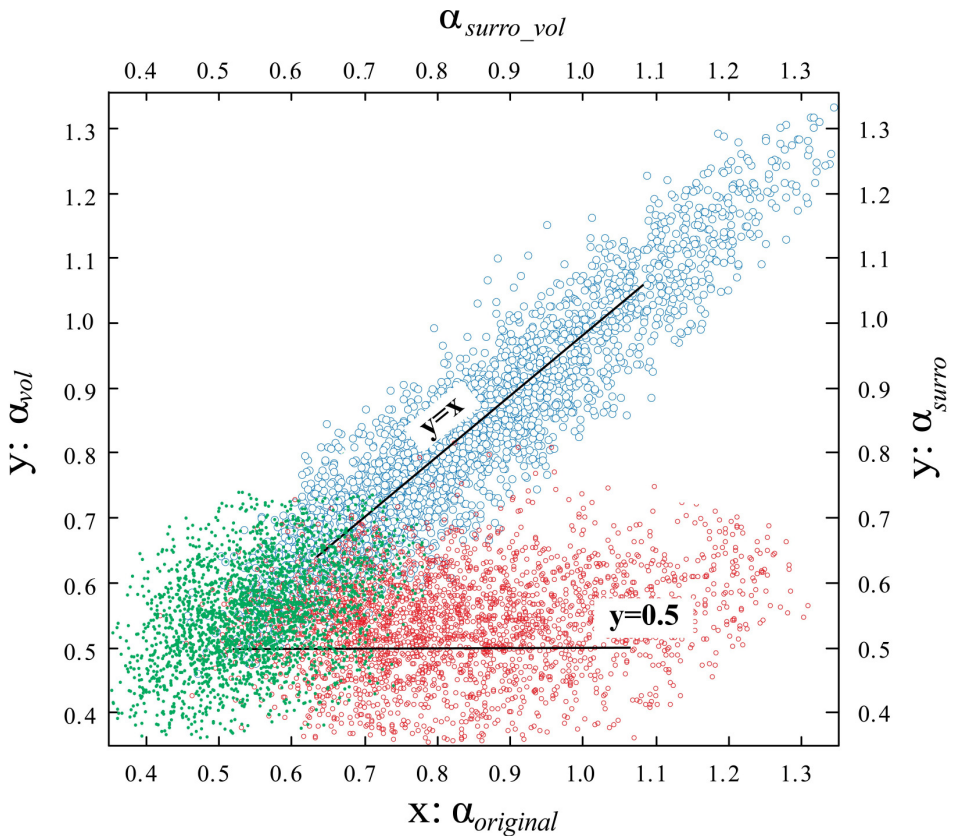

Figure 2.5 Combined scatter plot of the fluctuation exponents α for the original soil moisture time series, $\alpha_{original}$, surrogate series α_{surro}, volatility of the original time series $\alpha_{original\_vol}$, and volatility of the surrogate data α_{surro_vol}. The three plots are: $\alpha_{original}$ vs. α_{surro} (×), $\alpha_{original}$ vs. α_{vol} (o), and $\alpha_{original\_vol}$ vs. α_{surro_vol} (•).

2.5 Summary and discussion

Soil moisture is an important agent in land-atmosphere interactions, since it couples rapid precipitation fluctuations to storage with memory of the order of month to years. The memory time scale of soil moisture is widely characterized by an *e*-folding time within a Markovian framework. We study the power law fluctuations of soil moisture and find that there are two regimes, with either short- or long-rang correlations. The short-range correlation is equivalent to an e-folding time that is considered as the time scale at which soil moisture is integrated from uncorrelated precipitation. Beyond this time scale, the long-range correlations of soil moisture may contribute to low frequency climate variability. Within the short-range correlation regime, deep layers show a red spectrum, while within the long-range correlation regime, soil moisture shows an identical spectrum at different layers. There is a clear correspondence between the e-folding time and long-range correlations modified by local conditions. While the e-folding time is easily interpreted by simple first order autoregressive process, a simple model and the interpretation of the long-term power-law correlations are less clear and may originate in the complex interactions within the hydrological cycle (*Blender and Fraedrich*, 2006). In very dry and very wet as well as highly elevated regions, the fluctuation exponent for large interannual time scales is $\alpha>1$; this shows the nonstationarity of soil moisture process at very large time scale, and hampers the prediction by traditional statistics.

The volatility correlation of soil moisture, which is considered as an empirical indicator for nonlinearity reveals that soil moisture exhibits a white volatility spectrum. This suggests that the underlying processes of soil moisture in ERA-40 are linear, and that the statistical properties of soil moisture may be well approximated by linear models such as the Markovian model used by *Delworth and Manabe* (1988). This conclusion is substantiated by a surrogate data test. Further work is needed to clarify the physical processes and whether this corresponds to real linearity or it is in fact an artifact of the model used in the re-analysis.

The linearity found in the soil moisture time series does not necessarily imply a conflict with the expected 'nonlinear' interactions between soil moisture and other land surface water budget components, since the latter falls into a different definition of nonlinearity with respect to the dynamical equation (see *Ashkenazy,* 2003 for a review on the definition of nonlinearity), and actually there is no proven evidence between the nonlinearity found in the data and the nonlinearity in the governing dynamical equations (e.g. *Hsieh*, 2001).

References

Amenu, G. G., P. Kumar, and X. Z. Liang (2005), Interannual variability of deep-layer hydrologic memory and mechanisms of its influence on surface energy fluxes, *J Climate*, *18*(23), 5024-5045.

Ashkenazy, Y., D. R. Baker, H. Gildor, and S. Havlin (2003), Nonlinearity and multifractality of climate change in the past 420,000 years, *Geophys Res Lett*, *30*(22), 2146, doi:10.1029/2003GL018099.

Ashkenazy, Y., P. C. Ivanov, S. Havlin, C. K. Peng, A. L. Goldberger, and H. E. Stanley (2001), Magnitude and sign correlations in heartbeat fluctuations, *Phys Rev Lett*, *86*(9), 1900-1903.

Bartos, I., and I. M. Janosi (2006), Nonlinear correlations of daily temperature records over land, *Nonlinear Proc Geoph*, *13*(5), 571-576.

Blender, R., and K. Fraedrich (2006), Long-term memory of the hydrological cycle and river runoffs in China in a high-resolution climate model, *Int J Climatol*, *26*(12), 1547-1565.

Bunde, A., S. Havlin, J. W. Kantelhardt, T. Penzel, J. H. Peter, and K. Voigt (2000), Correlated and uncorrelated regions in heart-rate fluctuations during sleep, *Phys Rev Lett*, *85*(17), 3736-3739.

Conil, S., H. Douville, and S. Tyteca (2009), Contribution of realistic soil moisture initial conditions to boreal summer climate predictability, *Clim Dynam*, *32*(1), 75-93.

D'Andrea, F., A. Provenzale, R. Vautard, and N. De Noblet-Ducoudre (2007), Hot and cool summers: Multiple equilibria of the continental

water cycle, *Geophys Res Lett*, 34(3), L24807, doi:10.1029/2006GL027972.

Delworth, T., and S. Manabe (1989), The Influence of Soil Wetness on near-Surface Atmospheric Variability, *J Climate*, *2*(12), 1447-1462.

Dirmeyer, P. A. (2000), Using a global soil wetness dataset to improve seasonal climate simulation, *J Climate*, *13*(16), 2900-2922.

Entin, J. K., A. Robock, K. Y. Vinnikov, S. E. Hollinger, S. X. Liu, and A. Namkhai (2000), Temporal and spatial scales of observed soil moisture variations in the extratropics, *J Geophys Res-Atmos*, *105*(D9), 11865-11877.

Fraedrich, K., and R. Blender (2003), Scaling of atmosphere and ocean temperature correlations in observations and climate models, *Phys Rev Lett*, *90*(10), 1-4.

Govindan, R. B., A. Bunde, and S. Havlin (2003), Volatility in atmospheric temperature variability, *Physica A*, *318*(3-4), 529-536.

Hasselmann, K. (1976), Stochastic Climate Models .1. Theory, *Tellus*, *28*(6), 473-485.

Hong, S. Y., and E. Kalnay (2000), Role of sea surface temperature and soil-moisture feedback in the 1998 Oklahoma-Texas drought, *Nature*, *408*(6814), 842-844.

Hsieh, W. W. (2001), Nonlinear canonical correlation analysis of the tropical Pacific climate variability using a neural network approach, *J Climate*, *14*(12), 2528-2539.

Kalisky, T., Y. Ashkenazy, and S. Havlin (2005), Volatility of linear and nonlinear time series, *Phys Rev E*, *72*(1), 1-8.

Kantelhardt, J. W., E. Koscielny-Bunde, D. Rybski, P. Braun, A. Bunde, and S. Havlin (2006), Long-term persistence and multifractality of precipitation and river runoff records, *J Geophys Res-Atmos*, *111*(D1), 120-137.

Koster, R. D., and M. J. Suarez (2003), Impact of land surface initialization on seasonal precipitation and temperature prediction, *J Hydrometeorol*, *4*(2), 408-423.

Livina, V. N., Y. Ashkenazy, P. Braun, R. Monetti, A. Bunde, and S. Havlin (2003), Nonlinear volatility of river flux fluctuations, *Phys Rev E*, *67*(4), 042101.

Manabe, S., and T. Delworth (1990), The Temporal Variability of Soil Wetness and Its Impact on Climate, *Climatic Change*, *16*(2), 185-192.

Mandelbr.Bb, and J. R. Wallis (1969), Some Long-Run Properties of Geophysical Records, *Water Resour Res*, *5*(2), 321.

Mitchell, J. M. (1964), Further Remarks on the Power Spectrum of Red Noise, *J Atmos Sci*, *21*(4), 461-461.

Mudelsee, M. (2007), Long memory of rivers from spatial aggregation, *Water Resour Res*, *43*(1), 1-6.

Peng, C. K., S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger (1994), Mosaic Organization of DNA Nucleotides, *Phys Rev E*, 49(2), 1685-1689.

Schreiber, T., and A. Schmitz (2000), Surrogate time series, *Physica D*, 142(3-4), 346-382.

Seneviratne, S. I., D. Luthi, M. Litschi, and C. Schar (2006), Land-atmosphere coupling and climate change in Europe, *Nature*, *443*(7108), 205-209.

Uppala, S. M., et al. (2005), The ERA-40 re-analysis, *Q J Roy Meteor Soc*, *131*(612), 2961-3012.

Wang, G., T. Jiang, R. Blender, and K. Fraedrich (2008), Yangtze 1/f discharge variability and the interacting river-lake system, *J Hydrol*, *351*(1-2), 230-237.

Wu, W. R., and R. E. Dickinson (2004), Time scales of layered soil moisture memory in the context of land-atmosphere interaction, *J Climate*, *17*(14), 2752-2764.

Chapter 3

A Summer Climate Regime over Europe Modulated by the North Atlantic Oscillation[1]

Abstract

Recent summer heat waves in Europe were preceded by precipitation deficits in winter. Numerical studies suggest that these phenomena are dynamically linked by land-atmosphere interactions. However, there is still no clear evidence that connects summer climate variability to winter precipitation and the relevant circulation pattern so far. In this paper, we investigate the functional responses of summer mean as well as maximum temperature variability (Jun.–Aug., T_{mean} and T_{max}) to preceding winter precipitation (Jan.-Mar., P_{JFM}) for the period 1901-2005. There appear distinctive T_{mean} and T_{max} responses to P_{JFM} over the Mediterranean, where it is most sensitive to land-atmosphere interactions. An analysis of a soil moisture proxy (self-calibrating Palmer drought severity index, scPDSI) shows that the P_{JFM} seems to influence summer temperature via soil moisture, and therefore the T_{mean} and T_{max} responses we present here are very likely to be physical hints of water cycle interactions with temperature. We estimate that roughly 10~20% of

[1]The content of this chapter has been published as Wang, G., A. J. Dolman, and A. Alessandri (2011), A summer climate regime over Europe modulated by the North Atlantic Oscillation, *Hydrol Earth Syst Sci*, 15(1), 57-64.

the interannual variability of T_{max} and T_{mean} over the Mediterranean is statistically forced by P_{JFM}; for the scPDSI this value amount to 20~25%. Further analysis shows that these responses are highly correlated to the North Atlantic Oscillation (NAO) regime over the Mediterranean. Therefore we suggest that NAO modulates European summer temperature via controlling precipitation that initializes the moisture states of water cycle interactions with temperature. This picture of relations between European summer climate and NAO precipitation suggests potential for improved seasonal prediction of summer climate in particular extreme events.

3.1 Introduction

The recent European climate is characterized by an increasing frequency of summer heat waves with substantial societical and ecological impacts, e.g. the record-breaking heat wave in 2003. Climate projections point towards even higher-frequent and longer-lasting heat waves under increased greenhouse gas emission scenarios (*Scherrer* et al., 2005; *Pal et al*., 2004; *Stott et al*., 2004; *Meehl* et al., 2004). These past and projected heat waves highlighted the importance of a detailed understanding of the mechanisms that contribute to the initialization and persistence of extreme heat conditions. Hot and dry summers in Europe are generally associated with a specific, large-scale anticyclonic atmosphere circulation regime (*Cassou et al*., 2005; *Fischer* et al., 2007). It's also found that most of hot and dry summers over Europe were preceded by pronounced deficits of precipitation in winter and early spring (*Della-Marta et al*., 2007; *Vautard et al*., 2007). However, precipitation is nearly a white noise process with very limited memory; and winter precipitation cannot persist into summer solely through its atmospheric memory. With numerical experiments, *Vautard et al*. (2007) showed that the observed winter precipitation deficit and summer heat wave are dynamically linked via the feedback loops between land and atmosphere, wherein soil moisture plays a crucial role. The deficit of precipitation and subsequent drier soils resulted in reduced latent cooling and thereby an increase of air temperature, in agreement with other numerical experiments (e.g. *Seneviratne et al*., 2006; *Fischer et al.*, 2007; *Zampieri* et al., 2009).

These investigations of individual heat waves have highlighted the role of land-atmosphere interactions, and also pointed to the importance of circulation patterns, in the generation of summer heat waves. An immediate question that arises is whether this land surface feedback mechanism exists only for extraordinary hot summers or more systematically. *Schär et al*. (2004) underlined that an increase of interannual temperature variability in response to greenhouse-gas forcing might be an alternative causal mechanism for the occurrence of European summer heat waves; and numerical analysis by *Seneviratne et al*. (2006) suggested further that the increased interannual temperature variability is strongly related to the land-atmosphere interactions.

However, there exists as yet no clear analysis of observational evidence connecting the interannual variability of summer temperature to winter precipitation. The present paper aims to fill this gap in our understanding by investigating the relations of summer mean as well as maximum temperature and winter precipitation using long-term observations. Furthermore, a soil moisture analysis is also presented. The paper is organized as follows: in section 3.2 the observational datasets used are described and the statistical technique is briefly introduced. Section 3.3 is dedicated to the results and, finally, section 3.4 contains a discussion and the conclusions of this study.

3.2 Datasets and methods

3.2.1 Observational datasets

We use long-term observations of accumulated precipitation in Jan-Mar (P_{JFM}) and averaged daily mean as well as maximum temperature in Jun.-Aug. (T_{mean} and T_{max} respectively) for the period 1901-2005. Due to the sparseness of in situ soil moisture observations, the averaged self-calibrating Palmer drought severity index in Jun-Aug (scPDSI, *Wells et al*., 2004), is used as a proxy of soil moisture. The scPDSI is based on soil water content in a rather complex water budget model involving water cycle interactions with temperature; therefore it is suitable for the purpose of this study. Ideally one would use remotely sensed soil moisture observations (e.g. *de Jeu et al*., 2008) but the datasets are

unfortunately not yet sufficiently long in time. The scPDSI dataset obtained from CRU spans 1901-2002 on a monthly basis and range from -4 to +4 in the case of extremely dry and extremely wet conditions respectively (*van der Schrier et al*., 2006). These datasets, gridded at a horizontal resolution of 0.5º×0.5º, are derived from University of East Anglia Climatic Research Unit (CRU; *Mitchell et al*., 2005). The serial mean values over each pixel are removed to obtain anomalies. Data values over mountainous Scandinavia are not included in this study.

3.2.2 Coupled manifold technique

We use a technique, the Coupled Manifold Technique (CMT) recently proposed by *Navarra and Tribbia* (2005), to investigate the functional relations between fields of interest. Let $\boldsymbol{S}$ and $\boldsymbol{Z}$ stand for two fields and suppose they are linked by a linear relation. Their relation may then be expressed in terms of data matrices, as

$$\boldsymbol{Z}=\boldsymbol{Z}_{\text{for}}+\boldsymbol{Z}_{\text{free}}=\boldsymbol{AS}+\boldsymbol{Z}_{\text{free}}$$

$$\boldsymbol{S}=\boldsymbol{S}_{for}+\boldsymbol{S}_{free}=\boldsymbol{BZ}+\boldsymbol{S}_{free}. \qquad (3.1)$$

$\boldsymbol{A}$ and $\boldsymbol{B}$ are two linear operators that express the relation between $\boldsymbol{S}$ and $\boldsymbol{Z}$. Using the Procrustes method (*Richman* et al., 1993), the CMT seeks for $\boldsymbol{A}$ and $\boldsymbol{B}$ with

$$\boldsymbol{A}=\boldsymbol{ZS'}\,(\boldsymbol{SS'})^{-1}$$

$$\boldsymbol{B}=\boldsymbol{SZ'}(\boldsymbol{ZZ'})^{-1} \qquad (3.2)$$

where the primes denotes a matrix transpose operation. They are generally not equivalent: $\boldsymbol{A}$ expresses the influence of $\boldsymbol{S}$ on $\boldsymbol{Z}$, and $\boldsymbol{B}$ expresses the influence of $\boldsymbol{Z}$ on $\boldsymbol{S}$. Hence, the $\boldsymbol{Z}$ field can be separated into two parts using $\boldsymbol{A}$:

$$\boldsymbol{Z}_{\text{for}}=\boldsymbol{AS}$$

$$\boldsymbol{Z}_{free}=\boldsymbol{Z}-\boldsymbol{AS}. \qquad (3.3)$$

The $\boldsymbol{Z}_{for}$ part is the portion of $\boldsymbol{Z}$ variability forced by $\boldsymbol{S}$ (henceforth "forced manifold"), and $\boldsymbol{Z}_{free}$ is the portion independent from $\boldsymbol{S}$ (henceforth "free manifold"). The same equation can be solved for $\boldsymbol{S}_{for}$ and $\boldsymbol{S}_{free}$ using $\boldsymbol{B}$:

$$\boldsymbol{S}_{\text{for}} = \boldsymbol{BZ}$$

$$\boldsymbol{S}_{free} = \boldsymbol{S} - \boldsymbol{BZ}. \qquad (3.4)$$

The $\boldsymbol{Z}_{for}$ and $\boldsymbol{S}_{for}$ portions can be further decomposed by writing Equation (3.3) and (3.4) into the right-hand side of Equation (3.1):

$$\boldsymbol{Z} = \boldsymbol{A}(\boldsymbol{BZ} + \boldsymbol{S}_{free}) + \boldsymbol{Z}_{free} = \boldsymbol{ABZ} + \boldsymbol{AS}_{free} + \boldsymbol{Z}_{free}$$

$$\boldsymbol{S} = \boldsymbol{B}(\boldsymbol{AS} + \boldsymbol{Z}_{free}) + \boldsymbol{S}_{free} = \boldsymbol{BAS} + \boldsymbol{BZ}_{free} + \boldsymbol{S}_{free}. \qquad (3.5)$$

The $\boldsymbol{ABZ}$ and $\boldsymbol{BAS}$ represent the fully coupled manifolds of $\boldsymbol{Z}$ and $\boldsymbol{S}$ fields. From a viewpoint of physics, there are two possible mechanisms to generate these fully coupled manifolds. One is the external mechanism influencing both $\boldsymbol{S}$ and $\boldsymbol{Z}$ fields simultaneously. The other is the reciprocal feedbacks between $\boldsymbol{S}$ and $\boldsymbol{Z}$. Since $\boldsymbol{S}_{free}$ is free from $\boldsymbol{Z}$, the $\boldsymbol{AS}_{free}$ portion thus represents the $\boldsymbol{Z}$ variability purely forced by $\boldsymbol{S}$, that is, the directional influence. Similarly, the $\boldsymbol{BZ}_{free}$ portion represents the directional influence of $\boldsymbol{Z}$ on $\boldsymbol{S}$.

In our analysis, the CMT is applied to the EOF coefficients of fields of interest to simplify the computation, as suggested by *Navarra and Tribbia* (2005), with 99% of the total variance of each field retained. To obtain only significant relations, each elements in matrix A and B is tested against the null hypothesis of being equal to zero at the 1% significance level using the Student *t* distribution described by *Cherchi et al* (2007). When the forced manifold is obtained, a further significance test of the forced variance is performed to make the result robust using a Monte Carlo approach. Then *S* and the forced manifold Z_{for}, containing now only the variability in *Z* forced by *S*, are then subjected to traditional MCA to obtain the forcing and forced patterns as well as time coefficient series of interest.

3.3 Results

3.3.1 Responses of T_{mean} and Tmax to P_{JFM}

Figure 3.1a shows the percentage of T_{mean} variance forced by the P_{JFM} variability, and that for T_{max} is shown in Figure 3.1d. These values are derived from the ratio of the forced T_{mean} (T_{max}) manifold to the original T_{mean} (T_{max}) fields. We tested where the percentage values are significant different from zero at the 0.10 level. For each grid point, we tested the null hypothesis of getting as high or higher variance fractions through a Monte Carlo bootstrap method (10000 repetitions of the CMT) by randomizing the order number of P_{JFM} values on each grid. The largest values are found over southern Europe for both T_{mean} and T_{max} where it is most sensitive to land-atmosphere interactions (*Seneviratne* et al., 2006; *Fischer* et al, 2007; *Zampieri* et al., 2009), while little forcing (low values) is observed over northern Europe of 50ºN. Up to 5~15% of the summer T_{mean} variance over southern Europe appears to be forced by P_{JFM}. The forced T_{mean} variance by P_{JFM} is up to 8% over Western Europe, averaged within the green rectangle in Figure 3.1a, which doesn't pass the significance test. Over Eastern Europe, this value increases to 11% averaged within the red rectangle in Figure 3.1a, passing the significance test. This implies that summer T_{mean} over Eastern Europe is more sensitive to P_{JFM}. These values for T_{max} are a bit higher. The forced T_{max} variance is up to 10% over Western Europe and that value over Eastern Europe is up to 14%, averaged within the green as well as the red rectangles respectively in Figure 3.1d. Low values for both T_{mean} and T_{max} over North of 50ºN indicate little influence from P_{JFM}.

The MCA analysis was originally designed for detecting cross-correlation. In our study it is conducted to the P_{JFM} field and the T_{mean} (T_{max}) manifold forced by the P_{JFM} variability, and therefore what it detects is the forcing P_{JFM} pattern and the T_{mean} (T_{max}) response. Derived from the first MCA mode, Figure 3.1b and c shows the 1st pair of forcing P_{JFM} pattern and its T_{mean} response, containing 95% of the total squared covariance. This MCA mode exhibits unit correlated time coefficients ($r>0.999$), suggesting the derived forcing-forced relationship is very robust. The time coefficient series are shown as blue lines in Figure 3.3.

We note that the unit correlation derived here is due to the data preprocessing with CMT, which constructs only the T_{mean} variability forced by P_{JFM} at significance level of 0.01. The time coefficient series of the 1st MCA mode without CMT exhibit a correlation of 0.40 (not shown), which is clearly insufficient to conclude a significant linkage. The same situation also holds in the following analysis of T_{max} as well as soil moisture proxy of scPDSI.

Shown in Figure 3.1b and c, there exists only one significant P_{JFM} anomaly over the Mediterranean, with opposite sign of the T_{mean} response largely northward and eastward extended to 50 ºN compared to the P_{JFM} anomaly. This suggests that T_{mean} in summer fluctuates in correspondence to the anomalous states of P_{JFM} via the cooling effect of the surface energy balance. Precipitation is spectrally white with very limited memory up to two weeks due to the chaotic nature of atmosphere (*Wang* et al., 2010); therefore the extended memory of P_{JFM} is probably sustained by soil moisture feedbacks on precipitation. One may question the existence of forced T_{mean} anomalies in the opposite sign over north of 50 ºN. It appears not a physical response to P_{JFM} since there exists no forcing anomaly in P_{JFM} in the very location. Therefore we attribute it to be a statistical coexistence with no physical implication. Furthermore, this anomaly accounts for a very low percentage of the forced T_{mean} variance over north of 50 ºN (Figure 3.1a) and didn't pass our significance test.

Shown in Figure 3.1e and f is the 1st leading pair of the forcing P_{JFM} pattern and its T_{max} response, which contains 96% of the total squared covariance with unit correlated time coefficient series. Comparing Figure 3.1b and e, we can see clearly the T_{max} anomaly is forced by almost the same P_{JFM} anomaly as that forces T_{mean}; and the time coefficient series exhibit unit correlation ($r>0.999$). Furthermore, these time coefficient series are nearly unit correlated with those derived from the $P_{JFM} \sim T_{mean}$ association ($r>0.999$). These statistical properties suggest that the derived linkages between P_{JFM} and T_{mean} as well as T_{max} are very likely to be driven by the same climate dynamics. The forced T_{max} (Figure 3.1f) exhibits very similar dipole pattern as T_{mean} (Figure 3.1c), however the anomalies over north of 50°N are small again.

An important question regarding land-climate interactions is whether they lead to amplified variability of climate extremes, such as heat waves, particularly in the context of climate change (*Seneviratne* et al., 2009). Over south of 50 ºN, the percentage of T_{max} variance forced by P_{JFM} appears to be more homogenized than that of T_{mean}. Furthermore, the robust relations derived from MCA analysis after CMT enable us to compare the magnitudes of T_{mean} and T_{max} responses to P_{JFM}, where the magnitude of the T_{max} response appears to twice that of T_{mean}. Therefore P_{JFM} exerts to some extend larger influence on T_{max} than that on T_{mean} over south of 50 ºN, possibly through water cycle interactions.

3.3.2 The role of soil moisture

So far we showed that summer temperature fluctuates in relation to fluctuations in winter precipitation over Mediterranean. It is plausible to hypothesize that these responses are modulated by interactions between the water cycle and temperature with soil moisture playing a critical mediating role. An analysis of soil moisture would help to support this hypothesis. For this purpose the same analytic framework as above is conducted to P_{JFM} and summer scPDSI as a soil moisture proxy. Analysis of this field is expected to clarify the role of soil moisture in the forced T_{max} and T_{mean} responses to P_{JFM}. This analysis is restricted to south of 55 ºN where P_{JFM} has distinctive expressions in the T_{max} and T_{mean} fields.

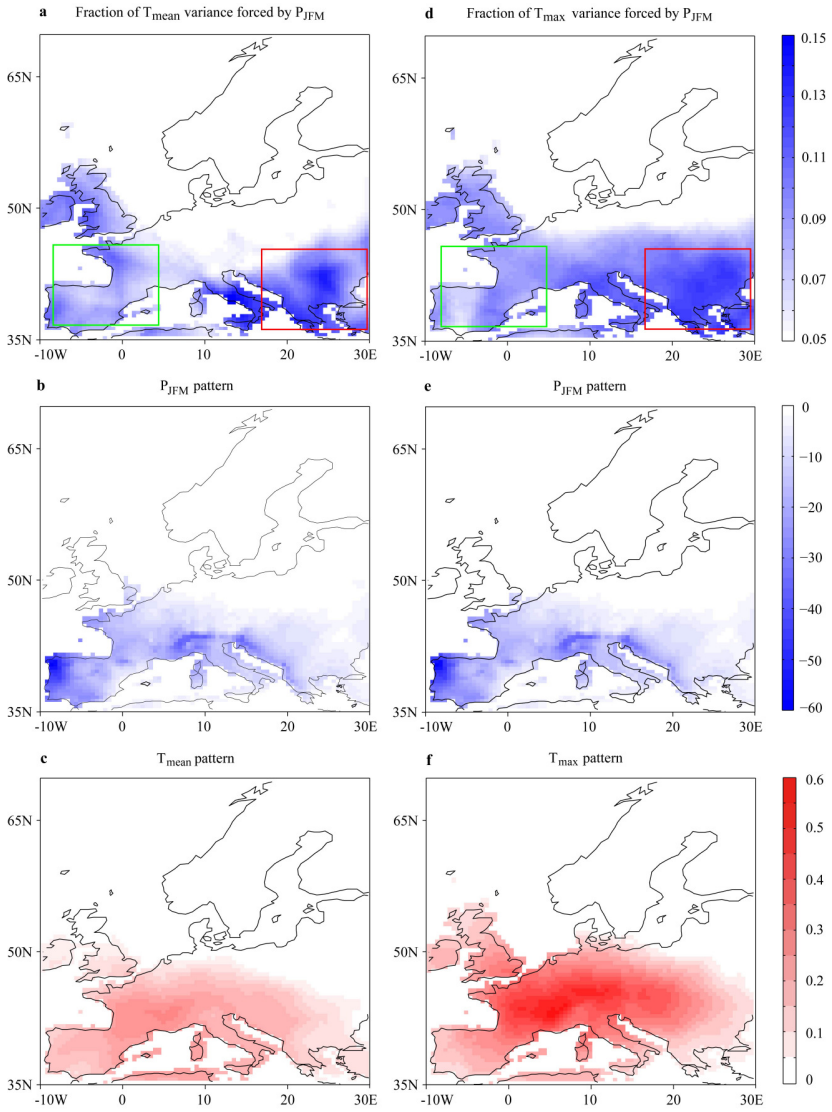

Figure 3.1 T_{mean} as well as T_{max} variability forced by P_{JFM}. a) Percentage of T_{mean} variance forced by P_{JFM} (sig=0.10 in the red rectangle). b) The forcing P_{JFM} pattern and c) its T_{mean} response. d) Percentage of T_{max} variance forced by P_{JFM} (sig=0.10 in the red rectangle). e) The forcing P_{JFM} pattern and f) its T_{max} response. All the relevant time coefficient series mutually exhibit unit correlation (r>0.999), shown in Figure 3.3. Units are K for T_{mean} as well as T_{max} and mm for P_{JFM}.

The scPDSI variability forced by P_{JFM} is shown in Figure 3.2. Shaded values in Figure 3.2a indicate the percentage of scPDSI variance forced by P_{JFM} that can pass significance test at 0.01 level, with the largest values of 20~25% existing in the west Mediterranean. The 1st MCA mode contains 80% of the total square covariance with unit correlated time coefficient series ($r>0.999$, green lines in Figure 3.3). The forcing P_{JFM} pattern exhibits a distinctive anomaly over Mediterranean (Figure 3.2b), very similar to that P_{JFM} patterns forcing T_{mean} and T_{max} (Figure 3.1a, d). The scPDSI response is of the same sign but largely northward and eastward extended (Figure 3.2c) compared to the forcing P_{JFM} anomalies. Of particular interest is that the time coefficient series are highly correlated with those from temperature analyses in section 3.3.1, with correlation coefficient $r>0.999$. The MCA analyses are summarized in Table 3.1. Therefore the responses of scPDSI, T_{mean} and T_{max} to the P_{JFM} variability we present here appear to be driven by the same climate dynamics, and P_{JFM} is very likely to influence T_{mean} and T_{max} via soil moisture. That is, a negative precipitation anomaly in winter is supposed to result in summer heating due to reduced latent cooling from soil moisture. The reverse relationship also holds, where a positive precipitation anomaly implies cooling.

		T_{mean}	T_{max}	scPDSI
1st MCA mode	explained squared covariance (%)	95	96	80
	correlation coeficient	>0.99	>0.99	>0.99

Table 3.1 A summery of the MCA analyses between P_{JFM} and the forced manifolds

These observational relations corroborate the interactions between water cycle and temperature established in previous numerical work, e.g., Seneviratne et al. (2006). Note that if we perform the same set of statistics to the winter precipitation and summer minimum temperature, we do not obtain the same relations. This is physically reasonable because the minimum temperature is highly constrained by external forcings, such as atmospheric circulation and sea surface temperature, rather than internal feedbacks (*Alfaro* et al., 2006; *Zhang* et al. 2008).

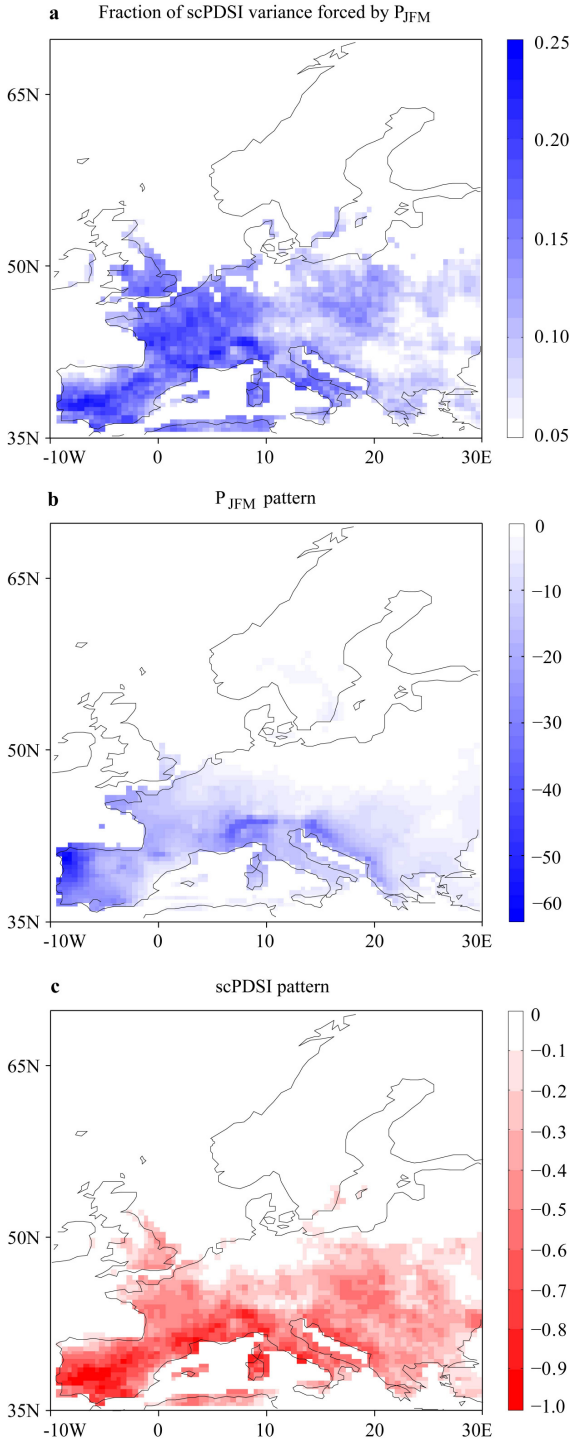

Figure 3.2 The scPDSI variability forced by P_{JFM}. a) Percentage of scPDSI variance forced by P_{JFM} (sig=0.01). b) The forcing P_{JFM} pattern and c) its scPDSI response, containing 80% of the total squared covariance. The MCA time coefficient series have unit correlation, as shown in Figure 3.3.

3.3.3 Link to the North Atlantic Oscillation

The North Atlantic Oscillation (NAO) is the dominating large-scale atmospheric circulation over the Atlantic-Europe sector in winter, with marked influence on winter climate. In recent years, the NAO is also observed to influence summer climate over Europe, in a weak but significant way. For example, *Qian et al* (2003) showed with observations that European summer temperature has positive correlation with the NAO index in previous January and February; *Kettlewell et al.* (2003) discovered a negative correlation between winter NAO and summer precipitation over Europe. However, the mechanism that links these phenomena remains still a puzzle. The derived forcing P_{JFM} patterns on T_{mean} and T_{max} as well as scPDSI in our analysis appear to resemble the NAO regime over Mediterranean, suggesting a plausible hypothesis that the NAO variability modulates summer climate over Europe through controlling winter precipitation that subsequently initializes the moisture states of water cycle interactions with temperature.

To further clarify the role of NAO in these processes, we compared the winter NAO index and the time coefficient series derived from MCA analysis. We use the averaged values of NAO index in Jan.-Mar. for the period 1901-2005, based on the difference of normalized sea level pressures between Gibraltar, Azores and SW Iceland. Shown in Figure 3.3, the derived time coefficient series from T_{mean}, T_{max} and scPDSI analyses have a high correlation with the NAO index with $r=0.65$ ($p<0.05$), suggesting a significant relation between the NAO variability and summer climate. The NAO variability is a north-south shift (or vice versa) in the track of storms and depressions across the North Atlantic Ocean and into Europe. The Atlantic storms that travel into Europe result in a dry Mediterranean Europe during a high NAO winter and the opposite during a low NAO winter (*Hurrell* et al., 2001). Based on the above analysis, we suggest that the NAO regime over the Mediterranean modulates European summer climate via initialization of the winter land surface moisture.

3.4 Discussion and conclusion

The importance of soil moisture initialization in winter and early spring for the seasonal prediction of heat and drought waves in European summer has been demonstrated in recent years (e.g., *Vautard* et al., 2007; *Fischer* et al, 2007; *Zampieri* et al., 2009; *Seneviratne* et al., 2006; *Ferranti* et al, 2006). Although soil moisture is closely related to precipitation, a clear picture of the relations between summer climate and preceding winter precipitation has not yet been demonstrated observationally. This is largely because the expected signal is very weak in the fields of interest, and traditional techniques for cross-correlation, such as MCA and CCA, are not capable of generating robust relations from this strong background noise.

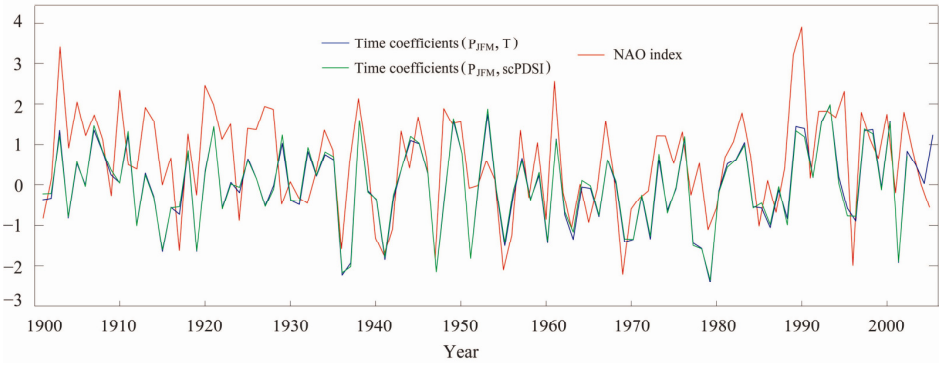

Figure 3.3 The MCA time coefficient series and the NAO index. Green line indicates the time coefficients series for the MCA modes of P_{JFM} and T_{mean} as well as T_{max}. Four time coefficient series mutually have correlation r>0.999, therefore shown with only one green line. Blue line indicates the unit correlated time coefficient series for the MCA mode of P_{JFM} and scPDSI. The red line indicates the averaged NAO index in Jan.-Mar.

Using the newly developed CMT technique that detects directional influence between climatic fields, we present in this paper robust

responses of summer T_{mean} and T_{max} as well as scPDSI to previous winter precipitation. Distinctive responses exist only over the Mediterranean area, where the temperature response is most sensitive to land-atmosphere interactions in regional climate models (*Schär et al*., 1999; *Seneviratne et al*., 2006). The P_{JFM} variability accounts for up to 10~15% of the total T_{mean} and T_{max} variance respectively for the period of 1901-2005; for the scPDSI this value amounts to 20~25% over Western Mediterranean. The P_{JFM} appears to influence T_{mean} and T_{max} via scPDSI, agreeing very well with our recent understanding of the water cycle dynamics over land (see *Seneviratne et al*., 2010 for a review). Therefore our findings are very likely to be physical of origin, although there is always a risk to infer physics from statistics. We also note that we are not addressing the full picture of land-atmosphere feedback processes but only that part that is related to Jan.-Mar. precipitation.

The extension of responses towards north and east is also observed in numerical experiments. *Vautard et al*. (2007) suggested that the northward propagation may be due to the southerly wind episodes carrying moisture northward. The eastward propagation is probably due to the heat low response over Central Europe, blocking the inflow of moist maritime air from the Atlantic and reinforcing the northward extension dynamically, addressed by *Haarsma et al*. (2008). Using a moisture tracer model, *Bisselink* and *Dolman* (2009) also found that advection is the most important contributor to precipitation over central Europe. It's notable that the largest anomalies of T_{mean} and T_{max} responses to P_{JFM} (Figure 3.1c, f) exist in central Europe, while the largest T_{mean} and T_{max} variance forced by P_{JFM} exists in southeast Europe (Figure 3.1a, d). This is because the interannual variability of T_{mean} as well as T_{max} over central Europe is much stronger than that over southeast Europe. We inferred that the forced variance can be considered as the forcing strength or land-atmosphere coupling strength, while the forced anomalies cannot. It can also be noticed that the P_{JFM} variability forces large T_{mean} and T_{max} variances over southeast Europe (Figure 3.1a, d), but that the variance of scPDSI is very limited there (Figure 3.2a). This is possibly because the derived T_{mean} and T_{max} variability over southeast Europe is closely related to the eastward extended heat response (*Haarsma et al*., 2009), while the soil moisture availability is very small there (*Bisselink and Dolman*, 2009).

We suggest that the NAO regime over the Mediterranean modulates summer climate over Europe through controlling winter precipitation that then initializes water cycle interactions with temperature. A positive phase of NAO tends to cause a hot and dry summer, or vice versa. This suggests there is scope for improved seasonal prediction of heat and drought waves from the pressure pattern of winter NAO. A remarkable feature of the NAO is its prolonged positive phases in the past 40 years, possibly related to anthropogenic warming (*Shindell et al*., 1999). This NAO dry pattern over the Mediterranean may have contributed to the increased frequency of heat and drought waves since then through modulating the water interactions over the Mediterranean.

References

Alessandri, A., and A. Navarra (2008), On the coupling between vegetation and rainfall inter-annual anomalies: Possible contributions to seasonal rainfall predictability over land areas, *Geophys Res Lett*, *35*(2), L02718, doi:10.1029/2007GL032415.

Alfaro, E. J., A. Gershunov, and D. Cayan (2006), Prediction of summer maximum and minimum temperature over the central and western United States: The roles of soil moisture and sea surface temperature, *J Climate*, *19*(8), 1407-1421.

Bisselink, B., and A. J. Dolman (2009), Recycling of moisture in Europe: contribution of evaporation to variability in very wet and dry years, *Hydrol Earth Syst Sc*, *13*(9), 1685-1697.

Cassou, C., L. Terray, and A. S. Phillips (2005), Tropical Atlantic influence on European heat waves, *J Climate*, *18*(15), 2805-2811.

Cherchi, A., S. Gualdi, S. Behera, J. J. Luo, S. Masson, T. Yamagata, and A. Navarra (2007), The influence of tropical Indian Ocean SST on the Indian summer monsoon, *J Climate*, *20*(13), 3083-3105.

de Jeu, R. A. M., W. Wagner, T. R. H. Holmes, A. J. Dolman, N. C. van de Giesen, and J. Friesen (2008), Global Soil Moisture Patterns Observed by Space Borne Microwave Radiometers and Scatterometers, *Surv Geophys*, *29*(4-5), 399-420.

Della-Marta, P. M., J. Luterbacher, H. von Weissenfluh, E. Xoplaki, M. Brunet, and H. Wanner (2007), Summer heat waves over western Europe 1880-2003, their relationship to large-scale forcings and predictability, *Clim Dynam*, *29*(2-3), 251-275.

Ferranti, L., and P. Viterbo (2006), The European summer of 2003: Sensitivity to soil water initial conditions, *J Climate*, *19*(15), 3659-3680.

Fischer, E. M., S. I. Seneviratne, D. Luthi, and C. Schar (2007), Contribution of land-atmosphere coupling to recent European summer heat waves, *Geophys Res Lett*, *34*(6), -.

Haarsma, R. J., F. Selten, B. V. Hurk, W. Hazeleger, and X. L. Wang (2009), Drier Mediterranean soils due to greenhouse warming bring easterly winds over summertime central Europe, *Geophys Res Lett*, 36, L04705, doi:10.1029/2008GL036617.

Hurrell, J. W., Y. Kushnir, and M. Visbeck (2001), The North Atlantic oscillation, *Science*, *291*(5504), 603.

Kettlewell, P. S., Stephenson, D. B., Atkinson, M. D. and Hollins, P. D. (2003), Summer rainfall and wheat grain quality: relationships with the North Atlantic Oscillation. *Weather*, 58, 155–164.

Koster, R. D., S. D. Schubert, and M. J. Suarez (2009), Analyzing the Concurrence of Meteorological Droughts and Warm Periods, with Implications for the Determination of Evaporative Regime, *J Climate*, *22*(12), 3331-3341.

Meehl, G. A., and C. Tebaldi (2004), More intense, more frequent, and longer lasting heat waves in the 21st century, *Science*, *305*(5686), 994-997.

Mitchell, T. D., and P. D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int J Climatol*, *25*(6), 693-712.

Navarra, A., and J. Tribbia (2005), The coupled manifold, *J Atmos Sci*, *62*(2), 310-330.

Pal, J. S., F. Giorgi, and X. Q. Bi (2004), Consistency of recent European summer precipitation trends and extremes with future regional climate projections, *Geophys Res Lett*, *31*(13), 31, L13202, doi:10.1029/2004GL019836.

Qian, B., and Saunders, M.A.: Summer U.K (2005), Temperature and Its Links to Preceding Eurasian Snow Cover, North Atlantic SSTs, and the NAO. *J. Climate*., *16*, 4108–4120.

Richman, M.B., and Vermette, S. J. (1993), The use of Procrustes target analysis to discriminate dominant source regions of fine sulfur in the western USA. *Atmos. Environ*., *27*A, 475–481.

Schar, C., D. Luthi, U. Beyerle, and E. Heise (1999), The soil-precipitation feedback: A process study with a regional climate model, *J Climate*, *12*(3), 722-741.

Schar, C., P. L. Vidale, D. Luthi, C. Frei, C. Haberli, M. A. Liniger, and C. Appenzeller (2004), The role of increasing temperature variability in European summer heatwaves, *Nature*, *427*(6972), 332-336.

Scherrer, S. C., C. Appenzeller, M. A. Liniger, and C. Schar (2005), European temperature distribution changes in observations and climate change scenarios, *Geophys Res Lett*, *32*(19), 32, L19705, doi:10.1029/2005GL024108.

Seneviratne, S. I., D. Luthi, M. Litschi, and C. Schar (2006), Land-atmosphere coupling and climate change in Europe, *Nature*, *443*(7108), 205-209.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, *Earth-Sci Rev*, *99*(3-4), 125-161.

Shindell, D. T., R. L. Miller, G. A. Schmidt, and L. Pandolfo (1999), Simulation of recent northern winter climate trends by greenhouse-gas forcing, *Nature*, *399*(6735), 452-455.

Stott, P. A., D. A. Stone, and M. R. Allen (2005), Human contribution to the European heatwave of 2003, *Nature*, 432, 610-614.

van der Schrier, G., K. R. Briffa, P. D. Jones, and T. J. Osborn (2006), Summer moisture variability across Europe, *J Climate*, *19*(12), 2818-2834.

Vautard, R., P. Yiou, F. D'Andrea, N. de Noblet, N. Viovy, C. Cassou, J. Polcher, P. Ciais, M. Kageyama, and Y. Fan (2007), Summertime

European heat and drought waves induced by wintertime Mediterranean rainfall deficit, *Geophys Res Lett*, *34*(7), -.

Wang, G. J., A. J. Dolman, R. Blender, and K. Fraedrich (2010), Fluctuation regimes of soil moisture in ERA-40 re-analysis data, *Theor Appl Climatol*, *99*(1-2), 1-8.

Wells, N., S. Goddard, and M. J. Hayes (2004), A self-calibrating Palmer Drought Severity Index, *J Climate*, *17*(12), 2335-2351.

Zampieri, M., F. D'Andrea, R. Vautard, P. Ciais, N. de Noblet-Ducoudre, and P. Yiou (2009), Hot European Summers and the Role of Soil Moisture in the Propagation of Mediterranean Drought, *J Climate*, *22*(18), 4747-4758.

Zhang, J. Y., W. C. Wang, and L. R. Leung (2008), Contribution of land-atmosphere coupling to summer climate variability over the contiguous United States, *J Geophys Res-Atmos*, *113*, D22109, doi:10.1029/2008JD010136.

Chapter 4

Amplified Summer Temperature Response to the Atlantic Multidecadal Oscillation over the Mediterranean Europe[1]

Abstract

The surface air temperature (SAT) over Mediterranean Europe exhibits marked variability at multidecadal scales, which is widely thought to be a response to the Atlantic Multidecadal Oscillation (AMO) of Sea Surface Temperature. Here, we analyze, using new rigorous spatio-temporal statistics, the relation between North Atlantic SST and European SAT in boreal summer (Jun.-Aug.) based on the instrumental records for the period 1901-2005. We show that the AMO variability of North Atlantic SST explains only half magnitude of the multidecadal SAT variability over the Mediterranean Europe. Its full magnitude can only be adequately explained as an amplified response in the presence of a long-term warming trend. We conclude that the multidecadal SAT variability over the Mediterranean Europe is not a direct response to the oceanic AMO variability, but contains a large component of amplification.

[1]The content of this chapter has been submitted as "Amplified summer temperature response to the Atlantic Multidecadal Oscillation over the Mediterranean Europe".

4.1 Introduction

The surface air temperature (SAT) over the Mediterranean Europe exhibits marked variability at multidecadal time scales (*Della-Marta et al*., 2007; *Shabalova and Weber*, 1999). With ongoing climate change, understanding the cause of this multidecadal variability is critical for realistic climate projections (*Keenlyside et al.*, 2008; *Sutton and Hodson*, 2005). It is particularly so for developing the decadal scale climate predictions and adaptation scenarios(*Collins et al.*, 2006; *Griffies and Bryan*, 1997; *Latif et al.*, 2006; *Meehl et al.*, 2009; *Sutton and Hodson*, 2005). This multidecadal SAT variability is widely considered to be a response to the basin-scale Sea Surface Temperature (SST) fluctuation over the North Atlantic Ocean (*Delworth and Mann*, 2000; *Knight et al.*, 2006; *Knight et al.*, 2005; *Sutton and Hodson*, 2005; *Trenberth and Shea*, 2006), known as the Atlantic Multidecadal Oscillation (AMO) (*Keenlyside et al.*, 2008; *Kerr*, 2000; *Knight et al.*, 2005). Such a linkage has been indeed demonstrated in modeling studies; however, there exist considerable uncertainties regarding the precise magnitude among these model outcomes (*Collins et al.*, 2006; *Delworth and Mann*, 2000; *Guan and Nigam*, 2009; *Sutton and Hodson*, 2005; *Zhang et al.*, 2007). Since the AMO variability of North Atlantic SST is thought to be driven by the Atlantic thermohaline circulation, it may be predictable once ocean circulation models have been improved (*Keenlyside et al.*, 2008; *Knight et al.*, 2005). This raises the hope and potential for decadal scale prediction of European temperatures (*Collins et al.*, 2006; *Keenlyside et al.*, 2008; *Latif et al.*, 2006; *Meehl et al.*, 2009).

The AMO states are generally represented by a serial index derived from detrended SST anomalies averaged over the North Atlantic (*Delworth and Mann*, 2000; *Knight et al.*, 2006; *Trenberth and Shea*, 2006). Alternatively, this index can also be obtained through the Empirical Orthogonal Function (EOF) of the SST field (*Parker et al.*, 2007). A correlation analysis is then often used to infer its climatic consequences (*Della-Marta et al.*, 2007; *Delworth and Mann*, 2000; *Huss et al.*, 2010; *Knight et al.*, 2006). This inference with a simple AMO index assumes spatially uniform capacity of the SSTs to affect continental climate, which is unlikely to be the case in reality. It is more plausible that SSTs at different locations have different capacities to affect the climate over

continents. This motivates our current attempt to quantify the European SAT response to the AMO variability that takes into account not only the temporal but also the spatial variations of SST. A rigorous statistical technique, the forced manifold (*Navarra and Tribbia*, 2005), is used. It is specially designed for analyzing the functional relations between climate fields, and has recently been used to study the coupling between land, ocean and atmosphere (*Alessandri and Navarra, 2008*; *Cherchi et al*., 2007; *Navarra and Tribbia*, 2005; *Wang et al*., 2011).

4.2 Dataset and method

4.2.1 Datasets

The data of North Atlantic SST is derived from the Met Office Hadley Centre (HadISST1) (*Rayner et al.*, 2003). This product is gridded at 1 degree resolution (from http://www.metoffice.gov.uk/hadobs/hadisst/). The European SAT data is obtained from the product (CRU TS3) of the Climatic Research Unit, School of Environmental Sciences, University of East Anglia, UK (*Mitchell and Jones*, 2005). This product is gridded at 0.5 degree, and is available from http://www.cru.uea.ac.uk/cru/data/. The SSTs used here are restricted to the extra tropical North Atlantic (60°W~40°E, 30°N~70°N), where most of the influence of temperatures in Europe resides (Sutton and Hodson, 2005; Zhang et al., 2007). The SST and SAT anomalies are obtained by removing the pixel-wise serial means. The SST anomalies are then subjected a running-window low-pass filter with window length of seven years to obtain its low frequency components (SST_L hereon).

4.2.2 Method

Given two climate fields **S** and **Z**, the forced manifold technique seeks for the solution of **Z** as a linear function of **S** in terms of data matrices (*Navarra and Tribbia*, 2005):

$$\mathbf{Z} = \mathbf{AS} + \mathbf{Z}_{free} \tag{4.1}$$

where

$$\mathbf{Z}= [z\,(1), z\,(2), \ldots , z(t)] \text{ and}$$

$$\mathbf{S}= [s\,(1), s\,(2), \ldots , s(t)] \tag{4.2}$$

are the data matrices at fixed time points. **A** is a matrix operator that expresses the influence of **S** on **Z**, satisfying

$$\min||\mathbf{Z}-\mathbf{AS}||^2, \tag{4.3}$$

where || || is the Frobenius norm. The solution is found by applying the minimization problem, which is known as the "Procrustes problem'. Owning to the applied low-pass filtering in our case, the matrix operator **A** might be not full rank**.** In this case, the **A** can be written as

$$\mathbf{A}=\mathbf{ZS'}(\mathbf{SS'})^{-1}, \tag{4.4}$$

with

$$(\mathbf{SS}')^{-1} = \sum_{i=1}^{K} u_i \sigma_i^{-2} u_i^{'} \ , \tag{4.5}$$

where u_i and σ_i are the left singular vectors and the singular values of the matrix S respectively, and the prime represents transpose. The summation extends over all its non-zeros singular values, and the modes that can not contribute to the variance of **S** are excluded from the inverse. The described method is applied to the Empirical Orthogonal Functions (EOFs) coefficients of **Z** and **S** fields, as *Navara* et al (2005) suggested, to reduce the mathematical dimension.

The data matrices of **Z** and **S** are first scaled with the square root of their respective covariance, i.e., $\mathbf{(ZZ')}^{-1/2}$ and $\mathbf{(SS')}^{-1/2}$, so that the rescaled data have unit covariance. Under such scaling each element of **A** is a correlation coefficient; and each element is then tested for statistical significance against the null hypothesis the correlation is zero (*Cherchi et al*., 2007). This test is based on student *t* distribution and uses *n*-2 degree of freedom. Only the coefficients that can pass a 1% significance criterion have been retrained in the operator **A**. Using **A**, we can decompose the field **Z** into **AS** and $\mathbf{Z}_{free}$ as described in (4.1). The derived **AS** is the portion of **Z** variability that is functionally forced by **S,** termed "the forced manifold"; and $\mathbf{Z}_{free}$ is the portion of **Z** variability independent of **S**. After rescaling back, the **AS** is further subjected to the EOF analysis to derive the principal spatio-temporal modes of interest.

4.3 Results

We first demonstrate the multidecadal variability in the SAT records. The SAT anomalies are first low-pass filtered with a running window of seven years, and then linearly detrended ($deSAT_L$ hereon). We then apply an EOF analysis to the $deSAT_L$. The derived first EOF mode, containing 58% of the total variance of the $deSAT_L$ field, represents the multidecadal variability of summer SAT over Europe, as shown in Figure 4.1. Similarly, the field of SST_L anomalies is first linearly detrended, to which ($deSST_L$ hereon) an EOF analysis is then applied. Shown in Figure 4.2, the 1st EOF mode of $deSST_L$, containing 60% of the total variance of the $deSST_L$ field, captures the basin-scale AMO variability. Hereon we refer to the principal component (PC) time series of its first EOF mode as the AMO index. This index agrees very well with that derived from the area average $deSST_L$ (Fig 4.2b). The PC time series of the first EOF mode of $deSAT_L$ resembles the AMO index, suggesting this EOF mode represents the SAT response to the AMO variability. High amplitudes of this response exist over Southeast Europe. Over Northern Europe, the amplitudes of this response are relatively low. We now continue to show with a focus on the Mediterranean Europe that the seeming high-amplitude response has largely aliased into the existing background warming.

We first compute the portion of SAT variability that is functionally forced by $deSST_L$. This part is called the forced manifold (*Navarra and Tribbia*, 2005). For this computation, the SAT anomalies are neither low-pass filtered nor detrended. To obtain reliable results, only the computed relations that can pass a 1% significance test (*Cherchi et al.*, 2007) are used to reconstruct the forced SAT manifold (see Methods). An EOF analysis is then applied to the derived SAT manifold. The spatial pattern and the associated PC time series of the derived first EOF mode are shown as Figure 4.3a and the PC_a in Figure 4.3c respectively. The AMO index is also shown in Figure 4.3c for comparison. The spatial pattern shows the same sign over the entire Europe, with the highest amplitudes occurring over the northern part and very low amplitudes over the Mediterranean part. Of particular significance is the fact that the derived PC time series (PC_a in Fig 4.3c) is nearly identical to the AMO index

(Fig 4.3c), which suggests this derived first EOF mode is an exact response to the AMO variability of North Atlantic SSTs.

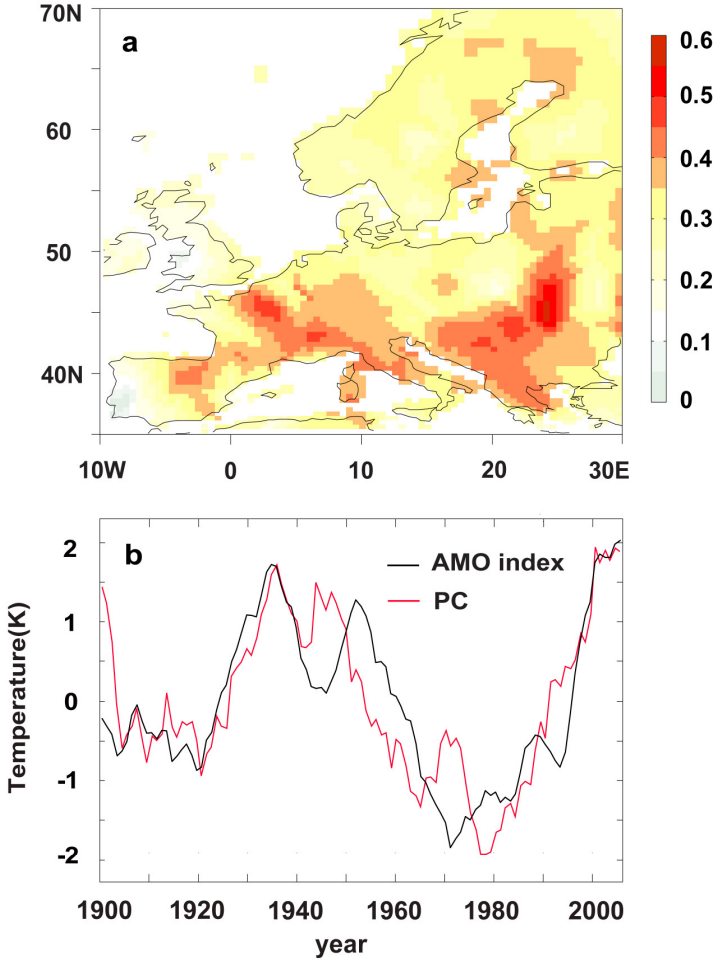

Figure 4.1 The first EOF mode of the deSAT$_L$ anomalies over Europe. (a) The spatial pattern (unit: K per standard deviation) and (b) the associated principal component (PC) time series with unit standard deviation. Also shown in (b) is the normalized AMO index.

In contrast with Figure 4.1, this result implies that the AMO has rather limited impact on the summer SAT over the South Europe where there actually appears a marked AMO signature in the instrumental records. It

is important to note that this SAT response exhibits quite similar pattern (Fig 4.3a) as those in response to the variations of Atlantic thermohaline circulation that drives the AMO (*Laurian et al.*, 2010; *Pohlmann et al.*, 2006). This supplies a strong indication that our result (Fig 4.3a) is very likely to be a realistic representation of the SAT response, if the AMO is indeed driven by the Atlantic thermohaline circulation.

Given this limited AMO impact in the above analysis, we suspect that the marked AMO signature in the summer SAT records over the Southern Europe (Fig 4.1) contains a large component of amplification, possibly owing to the existing background warming. Such amplification has previously been found in Arctic and Antarctic temperature records (*Chylek et al.*, 2010; *Chylek et al.*, 2009). To understand this more fully and, importantly, to quantify the possible amplification, we now compute the portion of the SAT variability that is forced by SST_L, that is, by the non-detrended SST_L anomalies. Again, the SAT anomalies are neither low-pass filtered nor detrended for this computation. The derived portion of SAT variability is then linearly detrended pixel by pixel ("deSAT manifold" hereon). If the suspected amplification effect indeed exists, the signal is now expected to be contained in the deSAT manifold. Thus by comparing with the SAT variability forced by $deSST_L$, we can determine the magnitudes of this amplification.

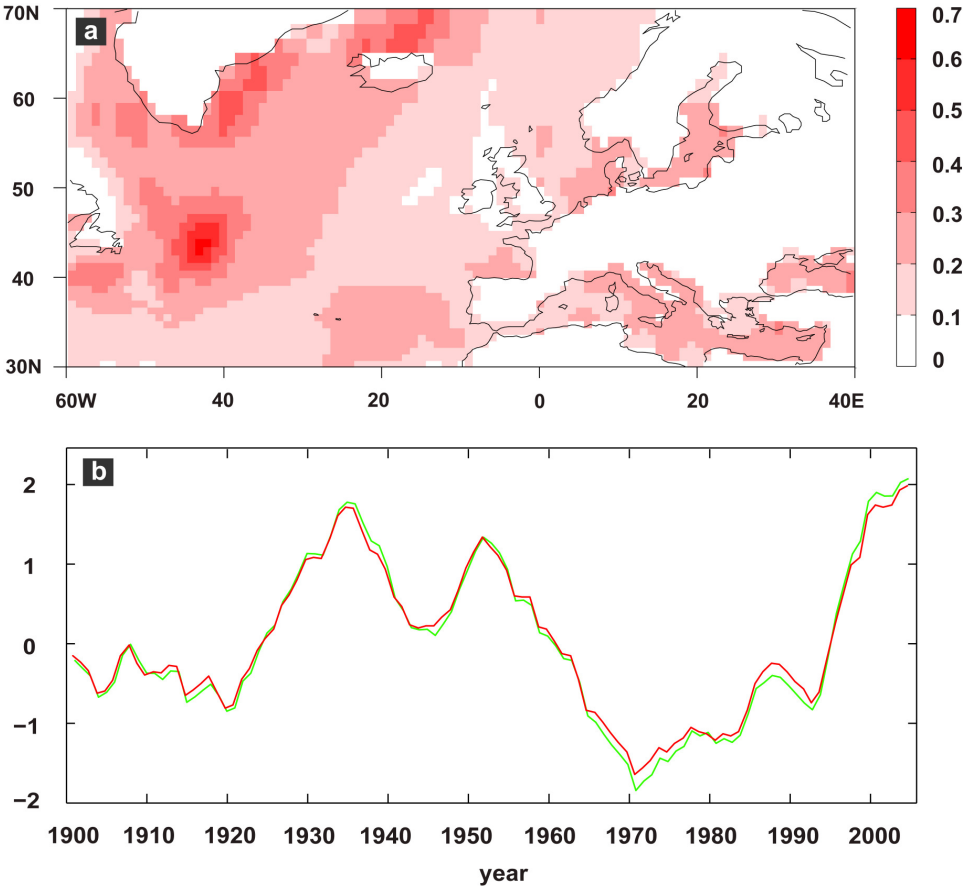

Figure 4.2 The basin-scale fluctuation of Atlantic Multidecadal Oscillation derived from the EOF analysis. (a) The spatial pattern of the 1st EOF mode of $deSST_L$, with corresponding PC time series shown as green line in (b). This EOF mode captures the basin-scale SST anomalies with persistent warm (1930s–1950s) and cool (1900s–1920s, 1960s–1980s) phases, and the onset of a warm phase in the 1990s. This resembles the AMO variability. The red line in (b) shows the area averaged $deSST_L$, which is also used as the AMO index at times. Both AMO indices are standardized to have unit deviation. The unit for spatial anomalies in (a) is K per standard deviation.

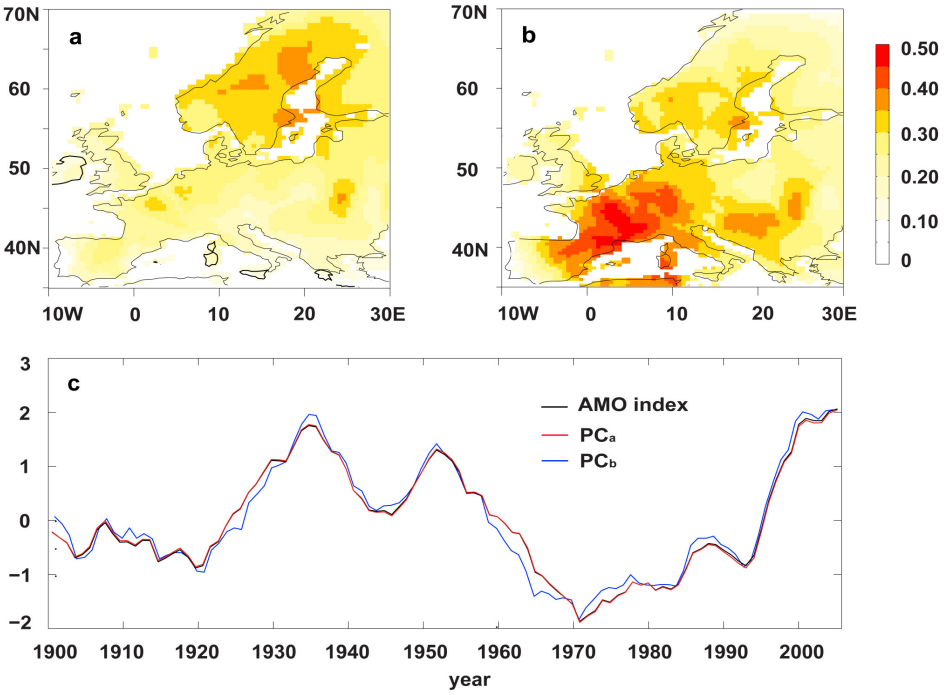

Figure 4.3 The summer SAT variability forced by $deSST_L$ and SST_L anomalies respectively. The first EOF mode of SAT manifold forced by $deSST_L$ consists of the spatial pattern in (a) and the PC time series, PC_a, in (c). The first EOF mode of deSAT manifold forced by SST_L consists of the spatial pattern in (b) and the PC time series, PC_b, in (c). The unit of spatial anomalies in (a) and (b) is K per standard deviation. Also shown in (c) is the AMO index. The PC_a , PC_b and AMO index in (c) all have unit standard deviations.

The derived deSAT manifold is now subjected to further EOF analysis. The spatial pattern and the associated PC time series of the derived first EOF mode are shown as Figure 4.3b and the PC_c in Figure 4.3c respectively. The spatial pattern shows the same sign over the entire Europe but with highest amplitudes over the Mediterranean part (Fig 4.3b). The associated PC time series (PC_b in Fig 4.3c) is very close to the AMO index, suggesting the derived EOF mode to be again a response to the AMO but, now in the presence of background warming. In contrast

with that forced by $deSST_L$, this portion of SAT variability has now dramatically increased over the Mediterranean Europe. This strongly suggests that the marked AMO signature in the summer SAT records over the Mediterranean Europe contains indeed a large component of amplification owing to the existing background warming.

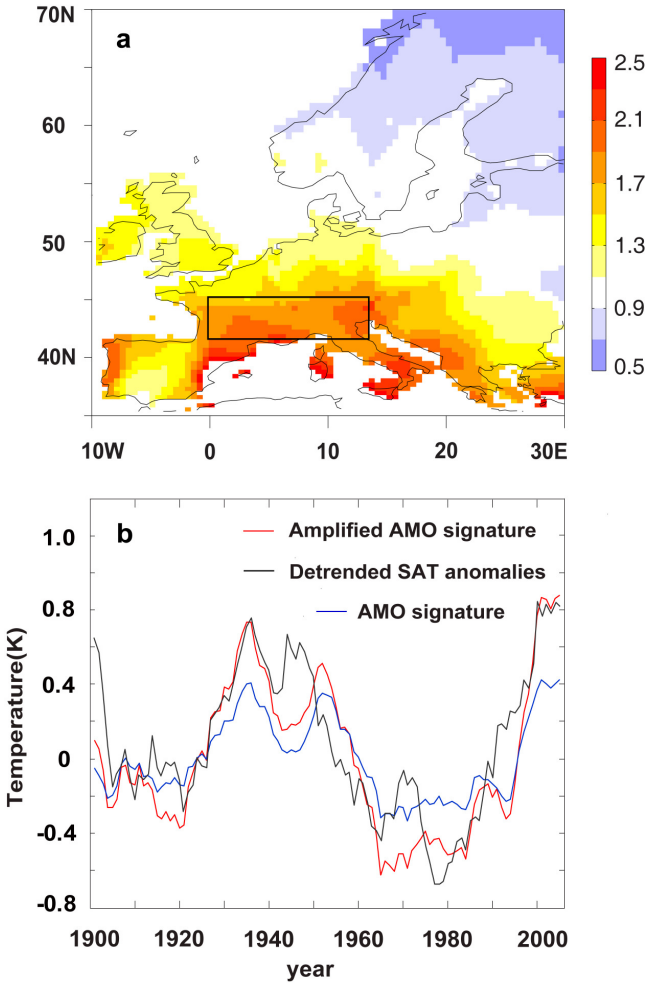

Figure 4.4 The magnitudes of the amplification. (a) The spatial distribution of the amplification factors. (b) The realistic AMO signature and its amplification over Mediterranean Europe, as with the realistic SAT variability at multidecadal scales observed directly from instrumental records.

To determine the magnitude of this amplification, we divide the amplified SAT response in the background warming by the SAT response purely to the AMO, both of which are derived as the EOF modes in the above analysis. Because they share nearly identical PC time series, the magnitudes of the amplification can be readily obtained by the amplitudes of the spatial pattern in Figure 4.3b divided by those in Figure 4.3a. The spatial distribution of these amplification factors is shown in Figure 4.4a. Not surprisingly, the largest amplification values exist over the Mediterranean Europe, where the background warming amplifies the SAT response to the AMO by a factor of roughly 2. This implies that roughly half magnitude of the AMO signature in the summer SAT records has aliased into the effect of the background warming. We further show in Figure 4.4b the time series of SAT anomalies in response to the AMO and those amplified in the background warming, representing the realistic and amplified AMO signature in the instrumental SAT records. These anomalies in response are averaged within the black square as marked in Figure 4.4a. Importantly, the validity of quantifying the amplification effect depends largely on whether the amplified AMO signature agrees with the realistic multidecadal SAT variability with respect to the amplitudes. To show this validity of our analysis, we also plot in Figure 4.4b the realistic multidecadal SAT variability in the same spatial domain, as determined by the first EOF mode of the $deSAT_L$ (Fig 4.1). It appears that the amplified AMO signature can adequately explain the magnitude of the realistic multidecadal SAT variability. This agreement of magnitudes strongly confirms the validity of our analysis, precluding the possible statistical artifact in our analysis. Shown in Figure 4.4b, in recent decades the SAT anomalies up to ~0.5 K can be decomposed from the direct AMO impact and attributed to the aliasing of the background warming. This helps to better understand the magnitude of recent climate change over Europe.

4.4 Conclusion and discussion

We have provided evidence from spatial-temporal analysis that the marked multidecadal variability of summer SAT over the Mediterranean Europe contains a large component of amplification in the background

warming. It is not just a direct response to the AMO variability of North Atlantic SST. Since the AMO is the most important source of decadal scale climate predictability (*Delworth and Mann*, 2000; *Keenlyside et al.*, 2008; *Latif et al.*, 2006; *Meehl et al.*, 2009; *Sutton and Hodson*, 2005), our result has clear implications for the prediction of future climate over Europe. The AMO is projected to cool down to its long-term mean state in the next few decades and this is expected to temporally offset the temperature increase over Europe (*Keenlyside et al.*, 2008; *Knight et al.*, 2005). This present study suggests a much reduced possibility for such a temperature relaxation to occur than previously expected.

An active debate exists on the precise shape of the AMO variability so far. Defining the AMO as a residual SST pattern after linear detrending may not fully reveal its true states (*Guan and Nigam*, 2009; *Mann and Emanuel*, 2006; *Ottera et al.*, 2010; *Ting et al.*, 2009; *Trenberth and Shea*, 2006). While this is possibly a caveat of our study, we have shown that the use of spatio-temporal statistics, rather than a simple index, does allow for a more realistic representation of relationships between the AMO and European temperature. Our analysis raises the warning that inferring the climatic consequences from a single AMO index may lead to overinterpretation over the Mediterranean Europe because of the aliasing of the background warming into the raw signal (*Mann and Emanuel*, 2006). However, such amplification is not so surprising, given the complicated interactions among different components of the climate system. We suggest there are physical processes, associated with the background warming, that induce positive feedbacks in the air-sea interactions that may lead to the amplified SAT response to the AMO forcing. However, our statistical analysis from instrumental records cannot provide the full unambiguous picture. Further modeling studies are needed to identify the precise feedback mechanisms. The strong AMO expression over the Southeast Europe (Fig 4.1) disappears partly in our spatio-temporal analysis (Fig 4.3). This may suggest the existence of amplification related to other processes rather than the background warming. The heat-low response over the Southeast Europe, caused by soil moisture feedbacks on temperature (*Haarsma et al.*, 2009; *Hirschi et al.*, 2011) and triggered by the AMO conditions at multidecadal scales, may be such a process.

References

Alessandri, A., and A. Navarra (2008), On the coupling between vegetation and rainfall inter-annual anomalies: Possible contributions to seasonal rainfall predictability over land areas, *Geophys. Res. Lett.*, *35*, L02718, doi:10.1029/2007GL032415.

Cherchi, A., S. Gualdi, S. Behera, J. J. Luo, S. Masson, T. Yamagata, and A. Navarra (2007), The influence of tropical Indian Ocean SST on the Indian summer monsoon, *J Climate*, *20*(13), 3083-3105.

Chylek, P., C. K. Folland, G. Lesins, and M. K. Dubey (2010), Twentieth century bipolar seesaw of the Arctic and Antarctic surface air temperatures, *Geophys. Res. Lett.*,, *37*, L08703.

Chylek, P., C. K. Folland, G. Lesins, M. K. Dubey, and M. Wang (2009), Arctic air temperature change amplification and the Atlantic Multidecadal Oscillation, *Geophys. Res. Lett.*, *36*, L14801.

Collins, M., et al. (2006), Interannual to decadal climate predictability in the North Atlantic: A multimodel-ensemble study, *J Climate*, *19*(7), 1195-1203.

Della-Marta, P. M., J. Luterbacher, H. von Weissenfluh, E. Xoplaki, M. Brunet, and H. Wanner (2007), Summer heat waves over western Europe 1880-2003, their relationship to large-scale forcings and predictability, *Climate Dynamics*, *29*(2-3), 251-275.

Delworth, T. L., and M. E. Mann (2000), Observed and simulated multidecadal variability in the Northern Hemisphere, *Climate Dynamics*, *16*(9), 661-676.

Griffies, S. M., and K. Bryan (1997), Predictability of North Atlantic multidecadal climate variability, *Science*, *275*(5297), 181-184.

Guan, B., and S. Nigam (2009), Analysis of Atlantic SST Variability Factoring Interbasin Links and the Secular Trend: Clarified Structure of the Atlantic Multidecadal Oscillation, *J Climate*, *22*(15), 4228-4240.

Haarsma, R. J., F. Selten, B. V. Hurk, W. Hazeleger, and X. L. Wang (2009), Drier Mediterranean soils due to greenhouse warming bring easterly winds over summertime central Europe, *Geophys. Res. Lett.*, 36, L04705, doi:10.1029/2008GL036617.

Hirschi, M., S. I. Seneviratne, V. Alexandrov, F. Boberg, C. Boroneant, O. B. Christensen, H. Formayer, B. Orlowsky, and P. Stepanek (2011), Observational evidence for soil-moisture impact on hot extremes in southeastern Europe, *Nature Geoscience*, *4*(1), 17-21.

Huss, M., R. Hock, A. Bauder, and M. Funk (2010), 100-year mass changes in the Swiss Alps linked to the Atlantic Multidecadal Oscillation, *Geophys. Res. Lett.*, 37, L10501, doi:10.1029/2010GL042616.

Keenlyside, N. S., M. Latif, J. Jungclaus, L. Kornblueh, and E. Roeckner (2008), Advancing decadal-scale climate prediction in the North Atlantic sector, *Nature*, *453*(7191), 84-88.

Kerr, R. A. (2000), A North Atlantic climate pacemaker for the centuries, *Science*, *288*(5473), 1984-1986.

Knight, J. R., C. K. Folland, and A. A. Scaife (2006), Climate impacts of the Atlantic Multidecadal Oscillation, *Geophys. Res. Lett.*, *33*, L17706, doi:10.1029/2006GL026242.

Knight, J. R., R. J. Allan, C. K. Folland, M. Vellinga, and M. E. Mann (2005), A signature of persistent natural thermohaline circulation cycles in observed climate, *Geophys. Res. Lett., 32*, L20708, doi:10.1029/2005GL024233.

Latif, M., M. Collins, H. Pohlmann, and N. Keenlyside (2006), A review of predictability studies of Atlantic sector climate on decadal time scales, *J Climate*, *19*(23), 5971-5987.

Laurian, A., S. S. Drijfhout, W. Hazeleger, and B. van den Hurk (2010), Response of the Western European climate to a collapse of the thermohaline circulation, *Climate Dynamics*, *34*(5), 689-697.

Mann, M. E., and K. A. Emanuel (2006), Atlantic hurricane trends linked to climate change, *Eos Trans. AGU*, *87*(24), 238-241.

Meehl, G. A., et al. (2009), Decadal Prediction Can It Be Skillful?, *Bulletin of the American Meteorological Society*, *90*(10), 1467.

Mitchell, T. D., and P. D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *International Journal of Climatology*, *25*(6), 693-712.

Navarra, A., and J. Tribbia (2005), The coupled manifold, *Journal of the Atmospheric Sciences*, *62*(2), 310-330.

Ottera, O. H., M. Bentsen, H. Drange, and L. L. Suo (2010), External forcing as a metronome for Atlantic multidecadal variability, *Nature Geoscience*, *3*(10), 688-694.

Parker, D., C. Folland, A. Scaife, J. Knight, A. Colman, P. Baines, and B. W. Dong (2007), Decadal to multidecadal variability and the climate change background, *J. Geophys. Res.*, *112*, D18115, doi:10.1029/2007JD008411.

Pohlmann, H., F. Sienz, and M. Latif (2006), Influence of the multidecadal Atlantic meridional overturning circulation variability on European climate, *J Climate*, *19*(23), 6062-6067.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*, 4407, doi:10.1029/2002JD002670.

Shabalova, M. V., and S. L. Weber (1999), Patterns of temperature variability on multidecadal to centennial timescales, *Journal of Geophysical Research-Atmospheres*, *104*(D24), 31023-31041.

Sutton, R. T., and D. L. R. Hodson (2005), Atlantic Ocean forcing of North American and European summer climate, *Science*, *309*(5731), 115-118.

Ting, M. F., Y. Kushnir, R. Seager, and C. H. Li (2009), Forced and Internal Twentieth-Century SST Trends in the North Atlantic, *J Climate*, *22*(6), 1469-1481.

Trenberth, K. E., and D. J. Shea (2006), Atlantic hurricanes and natural variability in 2005, *Geophys. Res. Lett.*, *33*, L12704, doi:10.1029/2006GL026894.

Wang, G., A. J. Dolman, and A. Alessandri (2011), A summer climate regime over Europe modulated by the North Atlantic Oscillation, *Hydrology and Earth System Sciences*, *15*(1), 57-64.

Zhang, R., T. L. Delworth, and I. M. Held (2007), Can the Atlantic Ocean drive the observed multidecadal variability in Northern

Hemisphere mean temperature?, *Geophys. Res. Lett.*, *34*, L02709, doi:10.1029/2006GL028683.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*, 4407, doi:10.1029/2002JD002670.

Chapter 5

Filling Gaps in the Large Soil Moisture Dataset from Satellite Images[1]

Abstract

The presence of data gaps is a common concern in geophysical records, creating not only difficulty, but more importantly, a large source of uncertainty in data analysis. Filling the data gaps is a necessity in the perspective of statistical modeling. There are numerous approaches for this purpose; however we face particular challenges regarding large spatio-temporal dataset such as Earth observations from satellites. Here we introduce an efficient three-dimensional method based on discrete cosine transforms, which explicitly utilizes information from both time and space to predict the missing values. To analyze its performance, this method was applied to a global soil moisture product derived from satellite images. We executed a validation by introducing synthetic gaps. It is shown this method is capable of filling data gaps in the global soil moisture dataset with high accuracy.

[1]The content of this chapter has been submitted as "Filling gaps in large soil moisture dataset from satellite images" to Environmental Modeling and Software.

5.1 Introduction

The presence of data gaps is a common concern in geophysical datasets, which presents a large source of uncertainty for data analysis. This concern is of particular importance when analyzing the spatio-temporal variability of large datasets, e.g., the large-scale satellite observations. In the last two decades satellite observations have demonstrated the potential to become a major tool for observing the properties of the Earth's land surface and atmosphere, such as soil moisture, temperature, aerosols and more recently green house gases. The data gaps in satellite datasets are inherent, primarily due to the satellite orbits. Other specific reasons such as clouds contamination or instrumental failure etc can also create data gaps. The rapidly growing volume and diversity of satellite datasets requires an efficient method for filling the data gaps.

Several methods for this purpose have emerged in recent years, among which the most promising ones are based on the empirical orthogonal function (EOF) of spatial variability (*Beckers et al*., 2003; *Alvera-Azcárate et al*., 2007) or the singular spectrum analysis (SSA) of temporal variability (*Kondrashov et al*., 2006; *Hocke et al*., 2008). These methods use a few spatial or temporal optimal modes occurring at low frequencies to predict the missing values. With the other components discarded as noise, these methods may lead to reduced accuracy of the statistical models fitted to the existing values and consequently the predicted missing values from these models. More importantly, for large spatio-temporal datasets it is of critical importance to utilize information from both spatial and temporal variability to predict the missing values. This demands a method that explicitly takes into account the full three-dimensionality (2-D spatial + time) of the spatio-temporal dataset. However, such a method is still not yet reported to date.

Here we introduce a penalized least square method based on three-dimensional discrete cosine transforms, for the purpose of filling data gaps in large spatio-temporal dataset. To show its performance, we apply it to a global soil moisture product derived from satellite images. There are two reasons to choose soil moisture dataset as a primary example. First, soil moisture is one important climate player, which affects the drought and heat conditions of lower atmosphere through partitioning the

available net radiation into latent heat for evaporation and sensible heat for temperature increase (*Koster* et al., 2004; *Seneviratne et al*., 2010). Complete soil moisture datasets are nowadays urgently needed for better understanding of the soil moisture-climate interactions, from which improved prediction of climate extremes is expected. Secondly, soil moisture exhibits temporally a red spectrum (*Wang et al*., 2010). This provides a special challenge to the existing gap filling methods utilizing only optimal modes at low frequencies (*Kondrashov et al*., 2006). It is worth noting that there exist some methods specially designed for filling data gaps in high-resolution soil moisture time series, as *Dumedah* has recently reviewed (2011); however, those are not applicable to spatio-temporal satellite product with coarse resolution.

5.2 Data and method

5.2.1 Global soil moisture product

We use the VU University-NASA (VUA-NASA) global volumetric soil moisture product (m3·m-3) derived from the Advanced Microwave Scanning Radiometer-Earth Observing System (*Owe et al*., 2008). This sensor is mounted on NASA's Aqua satellite and has daily ascending (13:30 equatorial local crossing time) and descending (01:30) overpasses. The surface moisture was retrieved with the Land Parameter Retrieval Model (LPRM) that solves simultaneously for the surface soil moisture and vegetation optical depth (*Owe et al*., 2008). The LPRM is based on a microwave radiative transfer model for passive microwave images that links surface geophysical variables, i.e. soil moisture, vegetation optimal depth and soil/canopy temperature, to the observed brightness temperatures. The C-band (6.9 GHz) channel is generally used to retrieve soil moisture; and it switches to X-band (10 GHz) when the C-band is contaminated by Radio frequency interference (RFI) (*Njoku et al*., 2005). This daily product has been validated extensively over a large variety of land surfaces of sparse to moderate vegetations, showing good agreement with in situ observations (*De Jeu et al*., 2008; *Wagner et al*., 2007). It has been shown that the VUA-NASA product outperforms other AMSR-E soil moisture product over various land cover types (*Draper et al*., 2009; *Rüdiger et al*., 2009). We apply the gap filling method to both the

ascending and descending products for the period 2003-2009, which are gridded at 0.50 degree resolution. Here we show only the results from the ascending product.

5.2.2 Gap filling method

The method to introduce is a panelized least square (PLS) regression based on three-dimensional discrete cosine transform (DCT), acronymed DCT-PLS. The DCT-PLS was originally proposed for automatic smoothing of multidimensional incomplete data (*Garcia*, 2010a*; Garcia*, 2010b), and we adapt it here for the purpose of filling data gaps of spatio-temporal geophysical datasets. The PLS regression is a thin-plate spline smoother for generally one-dimensional data array, which trades off fidelity to the data versus roughness of the mean function. Recently, *Garcia* (2010a) has demonstrated that the PLS regression can be formulated by the DCT, which expresses the data in terms of a sum of cosine functions oscillating at different frequencies. Since the DCT can be multidimensional, the DCT-based PLS regression can thus be immediately extended to multidimensional datasets. Following on we give a brief introduction of the DCT-PLS algorithm, and refer the mathematical details to *Garcia* (2010a).

Let X stand for a spatio-temporal dataset with gaps, and W a binary array of the same size indicating whether or not the values are missing. The DCT-PLS seeks for $\hat{X}$ that minimizes

$$F(\hat{X}) = \left\| W^{1/2} \circ (\hat{X} - X) \right\|^2 + s \left\| \nabla^2 \hat{X} \right\|^2; \tag{5.1}$$

i.e., it seeks $\hat{X}$ so that

$$\frac{\partial F}{\partial \hat{X}} = 0. \tag{5.2}$$

where || || is the Euclidean norm, ∇^2 and $\circ$ stand for the Laplace operator and the Schur (element-wise) product, respectively. The s is a positive scalar that controls the degree of smoothing: as s increases, the

smoothness of $\hat{X}$ also increases. The $\hat{X}$ can be easily achieved by rewriting Eq (1) with the type II DCT and its inverse (IDCT), which forms

$$\hat{X} = \mathrm{IDCT}(\Gamma \circ \mathrm{DCT}(W \circ (X - \hat{X}) + \hat{X})) . \tag{5.3}$$

Here, the Γ is a three-dimensional filtering tensor defined by

$$\Gamma_{i_1,i_2,i_3} = (1 + s(\sum_{j=1}^{3} (2 - \cos\frac{(i_j - 1)\pi}{n_j}))^2)^{-1}, \tag{5.4}$$

where i_j denotes the ith element along the jth dimension, and n_j denotes the size of X along this dimension.

In Eq (5.3) and (5.4), the DCT-PLS modeling relies only on the choice of the smoothing parameter s. For the purpose of filling data gaps, this parameter must have an infinitesimal value (≈ 0) to reduce the effect of smoothing. A high s value leads to the loss of high frequency components. For a specific s value, the fitness of the derived DCT-PLS model to the existing values can be evaluated by the reconstruction error. We define it as the normalized error between original existing values and their reconstructions, which falls into (0, 1):

$$\left\| W^{1/2} \circ (\hat{X} - X) \right\| / \left\| W^{1/2} \circ X \right\| . \tag{5.5}$$

Then the model with satisfying reconstruction error can then be used to predict the missing values.

5.3 Results

5.3.1 Gap filling results

Figure 5.1 shows the fraction of data gaps existing in the soil moisture product for the study period. Besides the systematic data gaps due to satellite orbit, the other major reasons creating data gaps in this dataset include dense vegetation, frozen soil (snow), waterbody etc. As a polar

orbiting satellite, the AQUA satellite gives better coverage over the high latitudes. However, the data gaps amount to 60-90% over north of 45ºN because of frozen soil. The same situation also exists for high elevation regions like in the Tibetan Plateau. Over regions of tropical rainforest, the vegetation is too dense to retrieve soil moisture. This product has the best coverage over Europe, with only 10-30% missing values.

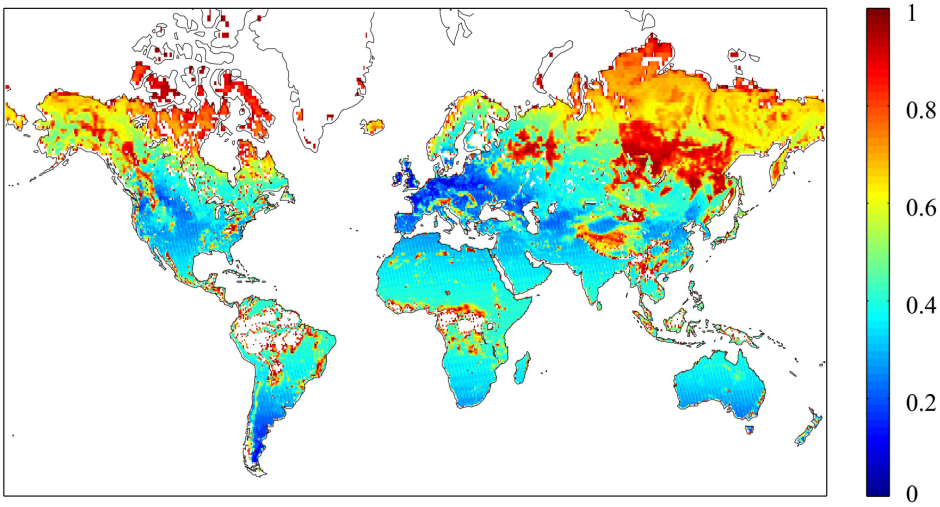

Figure 5.1 Fraction of data gaps in the ascending AMSR-E product for the period 2003-2009. White areas contain no data at all.

Using the DCT-PLS, the fitness of the derived model to the existing values is completely controlled by the smoothing parameter s, which can be any positive value. For the purpose of gap filling rather than smoothing, we consider here only s values smaller than 1. We apply the DCT-PLS to the global soil moisture product given s values of 10^{-N} with N=0, 1, … , 6 respectively. The global average reconstruction errors for each s value are calculated according to Eq (5.5), shown in Figure 5.2. Not surprisingly, a larger s value has resulted in a larger global error. When $s = 10^{-6}$ is used, the global error has already reached a very small value of 5×10^{-5}. This small error indicates that the derived DCT-PLS model approximates very well the existing values of the global soil

moisture dataset; and thus this model can be used to predict the missing values.

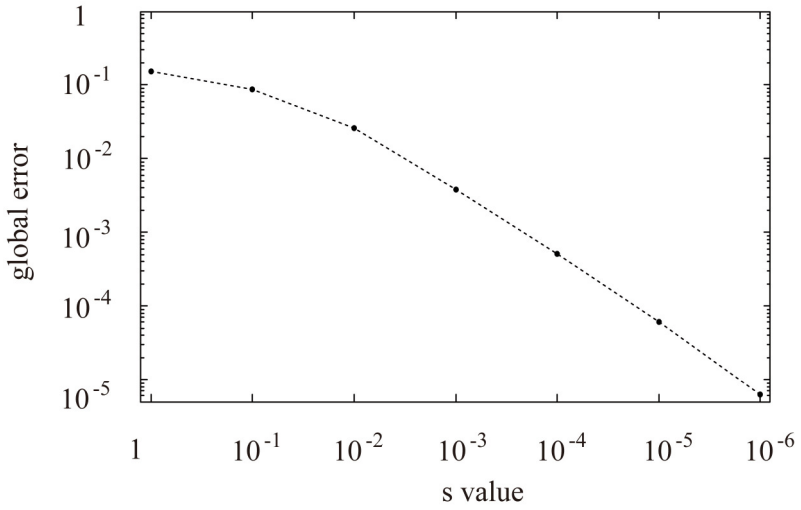

Figure 5.2 The errors averaged over globe for given s values.

Hereafter we demonstrate the gap filling result from the DCT-PLS with $s=10^{-6}$. We note that the data gaps in the entire dataset are filled by the three-dimensional DCT-PLS simultaneously. The data image and time series shown below are extracted from respectively the original and the gap-filled spatio-temporal datasets. Fig. 3 shows the data image on Jun. 5, 2003 prior to its model result. It appears the missing values are well filled not only between the satellite bypasses but also over the tropical rainforest regions where there are rare observations. In Figure 5.4, we show three time series with small to intermediate fraction of data gaps as well as their corresponding model outputs. For a clear presentation, only the data series for 2009 are shown. The upper panel shows the time series extracted from one pixel over Europe (47ºN, 2.5ºE), with 10% missing values in the original time series. The middle panel shows those from central US (35.5ºN, 99ºW), with 27% data gaps in the original series. The bottom panel shows those from equatorial Africa (11ºN, 0ºE), with 43% missing values in the original time series. In all three cases, the original values almost completely overlap their reconstructions by the DCT-PLS model, which is however not surprising because of the small

reconstruction error. It is noticeable that the extreme values existing in the original dataset are also well captured by the model, e.g., those emphasized with arrows in Figure 5.4. This indicates the predicted missing values from the used DCT-PLS model might be reliable; however, further validation is made in section 3.2.

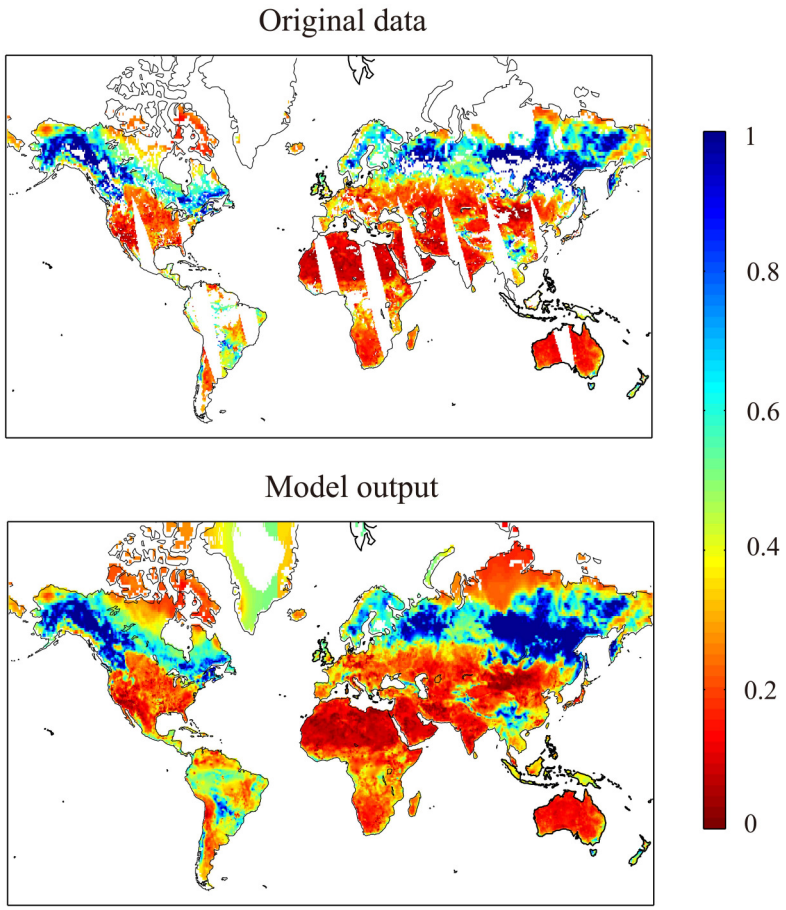

Figure 5.3 The data image on Jun. 5, 2003 prior to its model result. Unit: $m^3 \cdot m^{-3}$.

With conventional methods, the hardest part is to fill the continuous gaps. In spatio-temporal dataset the spatially continuous gaps can however be temporally intermittent, or visa versa, such as those between the satellite

bypasses. Owing to the three-dimensionality, the DCT-PLS method can easily cope with data gaps of this type. However, we need to give special care to data gaps of large spatio-temporal size, e.g., those over the tropical rainforest regions where the vegetation is too dense to retrieve soil moisture. In this case, the missing values are predicted mainly using the low frequency components of the dataset, leading to reduced reliability of filled-in high frequency components. A large portion of data gaps of this global soil moisture dataset is due to frozen soil, in which case the filled-in soil moisture values are physically not realistic, and we need to mask them out.

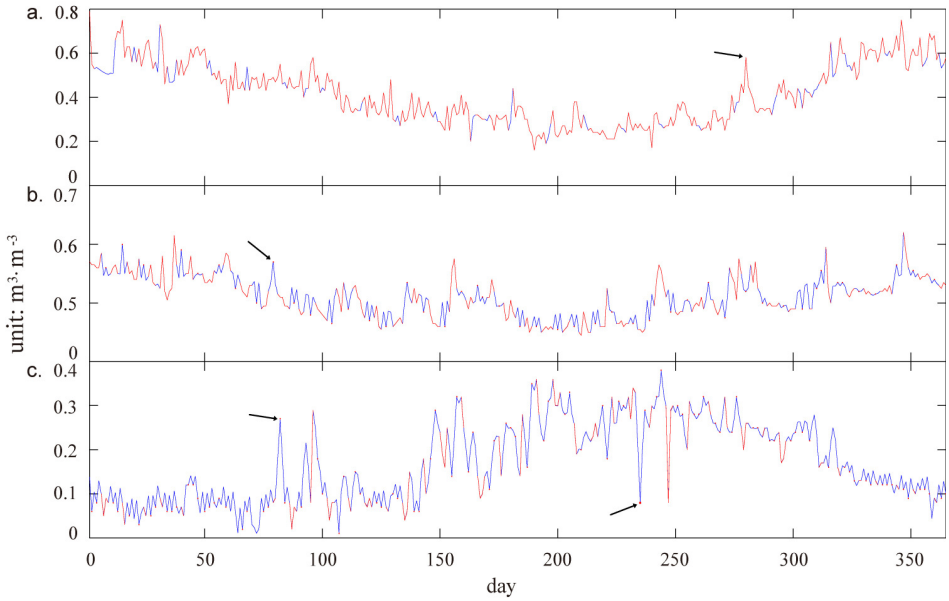

Figure 5.4 Original values (red) and the corresponding model reconstructions (blue) for the year 2009 from the pixels over **a.** Europe (47ºN, 2.5ºE), **b.** US (35.5ºN, 99ºW) and **c.** Africa (11ºN, 0ºE). Note that the original values are almost completely overlapped by the reconstructed values, due to the very small reconstruction errors. Emphasized with arrows are some extreme values.

5.3.2 Synthetic validation

Sometimes perfect fitting does not necessarily imply good prediction skill, for example, when overfitting problem occurs. Thus the prediction skill needs to be further validated, for which a general approach is to introduce synthetic gaps. To validate the prediction skill of the DCT-PLS method, we introduce synthetic gaps in addition into the original soil moisture dataset (2003-2009) by randomly removing 10% of the existing values over each pixel. Then the DCT-PLS gap filling process is applied to the new dataset with $s=10^{-6}$. In the synthetic gaps, we calculate the correlation coefficient (*Corr*) between the original values and their corresponding DCT-PLS predictions. The reconstruction error in the synthetic gaps can be alternatively used as the skill metric of prediction; however, it contains no more information than *Corr*, and we show here only *Corr*. The pixel-wise *Corr* ($p<0.05$) is shown in Figure 5. It appears that 85% of the validated pixels have higher *Corr* than 0.80, and those pixels with higher *Corr* than 0.90 amount to 64%. Specifically, the *Corr* values for the representative cases in Figure 4 are 0.97 (Europe), 0.95 (US) and 0.97 (Africa) respectively. This result suggests very good prediction skill of the DCT-PLS for the filling the data gaps of spatio-temporal soil moisture dataset.

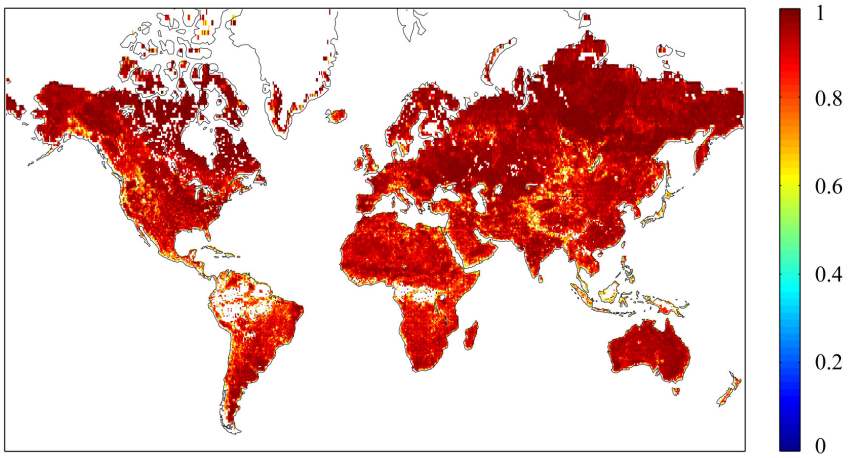

Figure 5.5 Pixel-wise Correlation ($p<0.05$) surface for the synthetic validation over globe.

5.4 Discussions

In this paper, we introduce an efficient DCT-PLS method for filling the data gaps in large spatio-temporal dataset; and we recommend it is of particular interest for the rapid growing volume and diversity of satellite observations in environmental sciences. Using a global satellite soil moisture dataset as a primary example and challenging case, we have demonstrated very good skill of this method for gap filling purpose.

This DCT-PLS method has some novel features with respect to other gap filling methods. It is a method of full three-dimensionality, and thus explicitly utilizes both spatial and temporal information of the dataset to derive the statistical model and predict the missing values. Instinctively, this strategy is rather sound for spatio-temporal dataset than using only spatial or temporal modeling. The statistical modeling process is completely controlled by one smoothing parameter which is easy to specify, eliminating the need of complicated model parameterization. Furthermore, with a small smoothing parameter the DCT-PLS method has the potential to reliably fill in the high frequency components.

However, we do not suggest "the smaller smoothing parameter, the better gap filling result". In case the geophysical dataset spatially have very large magnitude difference, an overfitting problem could occur with an extremely small smoothing parameter, leading to poor prediction performance. For example, in our soil moisture case the minor fluctuations in the dataset are indeed observed to be exaggerated over some regions, when a smoothing parameter smaller than 10^{-7} is used. There are alternative ways to avoid the overfitting problem regarding whether or not the dataset contains continuous spatio-temporal gaps of large size. For dataset with such gaps, the best choice is probably to introduce cross validation for better generalization as *Garcia* (2010) suggested. Yet, this may lead to underfiting and errorous prediction where there exist data gaps of large spatio-temporal size, in our experience. In this case we suggest a post-validation by introducing synthetic gaps to ensure the reliability of the predicted values.

References

Alvera-Azcarate, A., A. Barth, J. M. Beckers, and R. H. Weisberg (2007), Multivariate reconstruction of missing data in sea surface temperature, chlorophyll, and wind satellite fields, *J. Geophys. Res*., 112, C03008, doi:10.1029/2006JC003660.

Beckers, J. M., and M. Rixen (2003), EOF calculations and data filling from incomplete oceanographic datasets, *J Atmos Ocean Tech*, *20*(12), 1839-1856.

de Jeu, R. A. M., W. Wagner, T. R. H. Holmes, A. J. Dolman, N. C. van de Giesen, and J. Friesen (2008), Global Soil Moisture Patterns Observed by Space Borne Microwave Radiometers and Scatterometers, *Surv Geophys*, *29*(4-5), 399-420.

Draper, C. S., J. P. Walker, P. J. Steinle, R. A. M. de Jeu, and T. R. H. Holmes (2009), An evaluation of AMSR-E derived soil moisture over Australia, *Remote Sens Environ*, *113*(4), 703-710.

Dumedah, G., and P. Coulibaly (2011), Evaluation of statistical methods for infilling missing values in high-resolution soil moisture data, *J Hydrol*, *400*(1-2), 95-102.

Garcia, D. (2010), Robust smoothing of gridded data in one and higher dimensions with missing values, *Comput Stat Data An*, *54*(4), 1167-1178.

Garcia, D. (2011), A fast all-in-one method for automated post-processing of PIV data, *Exp Fluids*, *50*(5), 1247-1259.

Hocke, K., and N. Kampfer (2009), Gap filling and noise reduction of unevenly sampled data by means of the Lomb-Scargle periodogram, *Atmos Chem Phys*, *9*(12), 4197-4206.

Kondrashov, D., and M. Ghil (2006), Spatio-temporal filling of missing points in geophysical data sets, *Nonlinear Proc Geoph*, *13*(2), 151-159.

Koster, R. D., et al. (2004), Regions of strong coupling between soil moisture and precipitation, *Science*, *305*(5687), 1138-1140.

Njoku, E. G., P. Ashcroft, T. K. Chan, and L. Li (2005), Global survey and statistics of radio-frequency interference in AMSR-E land observations, *Ieee T Geosci Remote*, *43*(5), 938-947.

Owe, M., R. de Jeu, and T. Holmes (2008), Multisensor historical climatology of satellite-derived global land surface moisture, *J Geophys Res-Earth*, *113*(F1), -.

Rudiger, C., J. C. Calvet, C. Gruhier, T. R. H. Holmes, R. A. M. de Jeu, and W. Wagner (2009), An Intercomparison of ERS-Scat and AMSR-E Soil Moisture Observations with Model Simulations over France, *J Hydrometeorol*, *10*(2), 431-447.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, *Earth-Sci Rev*, *99*(3-4), 125-161.

Wagner, W., V. Naeimi, K. Scipal, R. de Jeu, and J. Martinez-Fernandez (2007), Soil moisture from operational meteorological satellites, *Hydrogeol J*, *15*(1), 121-131.

Wang, G. J., A. J. Dolman, R. Blender, and K. Fraedrich (2010), Fluctuation regimes of soil moisture in ERA-40 re-analysis data, *Theor Appl Climatol*, *99*(1-2), 1-8.

Chapter 6

The Observed Soil Moisture Feedback on Precipitation from Satellite Datasets over Europe

Abstract

Modeling studies have postulated the possibility of improved prediction of precipitation from soil moisture. This has raised the hope of early warning of summer droughts. However there exist considerable discrepancies among model outcomes. So far, conclusive evidence of soil moisture impact on precipitation has not been shown directly from observations over Europe. Using satellite observations over southern and central Europe, we disentangle and quantify the direction of the interactions between soil moisture and precipitation in the time-frequency space based on the statistical concept of Granger causality. We find a significant signal of soil moisture impact on summer precipitation over Europe. Soil moisture memory is identified as the mechanism that sustains this impact at time scales of 1-2 months. This result sheds light on the soil moisture-precipitation interactions as found in the real world rather than in modeling results.

6.1 Introduction

Precipitation originating from the surface moisture flux is critically important in the terrestrial water cycle and may play a major role in the generation of hydrological extremes, such as summer droughts and floods (*Koster et al.*, 2004; *Seneviratne et al.*, 2006; *Vautard et al.*, 2007). Modeling studies have provided evidence for the existence of such a soil moisture-precipitation feedback, albeit only over over specific regions. However, there remains a high uncertainty about the magnitude and even the sign of this feedback (*Cook et al.*, 2006; *Hohenegger et al.*, 2009; *Koster et al.*, 2004; *R. D. Koster et al.*, 2006; *Schar et al.*, 1999; *van den Hurk et al.*, 2010; *Vautard et al.*, 2007). With such high uncertainty our understanding of the soil moisture impact on precipitation remains necessarily limited (*Dirmeyer et al.*, 2006). One of the reasons for this lack of understanding is that until now empirical evidence to support the modeling studies directly from observations of soil moisture that could act as a benchmark is very rare. Substantial progress in retrieving soil moisture from space has now made such a more empirically based analysis possible (*de Jeu et al.*, 2008).

The challenge is then to identify the directions of soil moisture-precipitation interaction, i.e. the precise driver-response relationships. While this is straightforward in climate models by perturbing the system and analyzing the responses, disentangling these relationships in the real world, from observations, is much more difficult. Here we make such an attempt with soil moisture and precipitation observations using the statistical concept of Granger causality (GC) (*Granger*, 1969). The GC method is based on linear prediction theory that the cause occurs before the effect and, importantly, that knowledge of the cause helps the prediction of the effect (*Geweke*, 1982; *Granger*, 1969). It has been extensively used in many disciplines, e.g., econometrics (*Geweke*, 1982; *Granger*, 1969), neuroscience (*Brovelli et al.*, 2004) and also in climate studies (*Kaufmann and Stern*, 1997).

Formulated in terms of vector autoregression (VAR), the GC method assumes covariance stationarity (*Geweke*, 1982; *Granger*, 1969). This implies that, in a in statistical sense, the causal relationships are time-invariant. However, the impact of soil moisture on precipitation is

expected to vary with season (*Koster et al*., 2004; *Seneviratne et al.,* 2010). Identifiable timescales of this impact have appeared in a few numerical studies, partly as a function of soil moisture memory (e.g., *Dirmeyer et al*., 2009). It is thus plausible to establish a time-frequency representation of the impact of soil moisture on precipitation for understanding its time-varying properties and the time scales at which this impact operates. For this purpose, we obtain the GC results by combining Geweke's definition of spectral GC (1982) with the factorization of a wavelet spectral matrix (*Dhamala et al.*, 2008). An advantage of this wavelet-based method is that it eliminates the explicit need of parametric VAR modeling, which is impossible for soil moisture and precipitation due to their unbalanced variance.

6.2 Data and method

6.2.1 Remote sensing dataset

We use the daily VU University Amsterdam-NASA (VUA-NASA) product of volumetric soil moisture ($m^3 \cdot m^{-3}$). It is derived from satellite images of the Advanced Microwave Scanning Radiometer-Earth Observing System, using a Land Parameter Retrieval Model (LPRM) that solves simultaneously for the surface soil moisture and vegetation optical depth (*de Jeu et al*., 2008; *Owe et al*., 2008). The soil moisture retrievals from images of the descending bypasses (1:30 am, local time) are used. There exist 10-30% missing values in the soil moisture dataset, mainly due to satellite orbit and dense vegetation. We applied a three-dimensional thin-plate spline algorithm to interpolate the dataset, which generates full and consistent time series (*Garcia*, 2010), as shown in Chapter 5. The daily precipitation dataset is derived from the multiple satellite precipitation product of the Global Precipitation Climatology Project (*Huffman et al*., 2001).

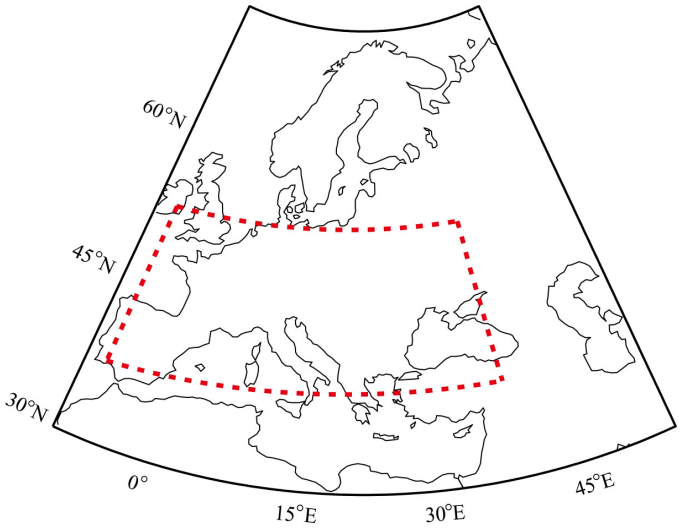

Figure 6.1 The analysis domain of this study.

Both daily datasets are resampled to have a horizontal resolution of one degree. The current study is restricted to the period of 2003-2007 where the datasets temporally overlap. The probability density function of soil moisture and precipitation are skewed. Before analyzing the interactions, the time series over each pixel are forced to have normal probability density function (PDF) with zero mean and unit variance using the lookup table method (*Jevrejeva et al.*, 2003). We note that, owing to existence of many zero values, the precipitation time series cannot be perfectly normalized; however, this is expected to have negligible influences on the results. This study is focused on southern and central Europe, shown as the red square in Figure 6.1, where soil moisture has the most potential to affect atmosphere in climate models. Within this spatial domain, there are 459 pixels with complete soil moisture and precipitation time series.

6.2.2 Wavelet-based Granger causality

This wavelet-based GC method combines *Geweke*'s (1982) spectral GC definition and the factorization of wavelet spectral matrix. Suppose we

have too stationary time series $X_1(t)$ and $X_2(t)$, and each admits a vector autoregressive (VAR) representation:

$$\begin{aligned} X_1(t) &= \sum_{\eta=1}^{p} A_{11,\eta} X_1(t-\eta) + \varepsilon_1(t) \\ X_2(t) &= \sum_{\eta=1}^{p} A_{22,\eta} X_2(t-\eta) + \varepsilon_2(t) \end{aligned}, \tag{6.1}$$

where $\eta = 1,..., p$ is the order of the VAR model, $\mathbf{A}_\eta = \begin{pmatrix} A_{11} \\ A_{22} \end{pmatrix}$ is the transfer coefficient of the model that indicates the contribution of the lagged values to the predicted values, $\varepsilon_1(t)$ and $\varepsilon_2(t)$ are models residuals (prediction errors). Jointly, they can be represented as a bivariate VAR process in the form

$$\begin{aligned} X_1(t) &= \sum_{\eta=1}^{p} A_{11,\eta} X_1(t-\eta) + \sum_{\eta=1}^{p} A_{12,\eta} X_2(t-\eta) + \varepsilon_{1|2}(t) \\ X_2(t) &= \sum_{\eta=1}^{p} A_{21,\eta} X_1(t-\eta) + \sum_{\eta=1}^{p} A_{22,\eta} X_2(t-\eta) + \varepsilon_{2|1}(t) \end{aligned}, \tag{6.2}$$

Here, $\varepsilon_{1|2}(t)$ and $\varepsilon_{2|1}(t)$ are the error items of the joint model, which are modeled as uncorrelated white noise. If $var(\varepsilon_{1|2}(t)) < var(\varepsilon_1(t))$, it is said $X_2(t)$ statistically has a causal influence on $X_1(t)$, which can be quantified in the time domain as (*Granger*, 1969)

$$GC_{2->1} = \ln \frac{\mathrm{var}(\varepsilon_1(t))}{\mathrm{var}(\varepsilon_{1|2}(t))}. \tag{6.3}$$

The can be understood as the improvement of the predictability of $X_1(t)$ by incorporating p-step history of $X_2(t)$. The statistical causality from $X_1(t)$ to $X_2(t)$ can be defined similarly.

Further interest lies in the frequency at which these influences operate for better understanding the underlying mechanisms. This can be achieved by the spectral decomposition of Granger causality as in (6.3), which is in time domain, using the formulation of *Geweke* (1982). A compact format of (6.2) and (6.3) forms

$$\mathrm{X}(t)=\sum_{\eta=1}^{p}\mathrm{A}_{\eta}\,\mathrm{X}(t-\eta)+\mathrm{E}(t) \tag{6.4}$$

where $\mathbf{A}_{\eta}=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ is the transfer coefficient matrix, and $\mathbf{E}(t)$ is the residual errors with covariance matrix denoted as $\Sigma=\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$.

Once the VAR model of (6.4) is fitted to the data, its spectral representation by Fourier transformation forms (*Geweke*, 1982)

$$\mathbf{S}(f)=\mathbf{H}(f)\,\boldsymbol{\Sigma}\,\mathbf{H}^{*}(f) \tag{6.5}$$

in which the asterisk denotes matrix conjugation, and $\mathbf{H}$ is a spectral representation of $\mathbf{A}_k$. Then The GC values from X_{2t} to X_{1t} ($GC_{2\to1}$) at frequency f become (*Geweke*, 1982)

$$GC_{2\to1}(f)=\ln\frac{S_{11}(f)}{S_{11}(f)-(\Sigma_{22}-\Sigma_{12}^{2}/\Sigma_{11})\left|H_{12(f)}\right|^{2}} \tag{6.6}$$

and, those from X_{1t} to X_{2t} ($GC_{1\to2}$) become

$$GC_{1\to2}(f)=\ln\frac{S_{22}(f)}{S_{22}(f)-(\Sigma_{11}-\Sigma_{21}^{2}/\Sigma_{22})\left|H_{21(f)}\right|^{2}}. \tag{6.7}$$

The GC method conceptually determines the improvement of the linear predictability of one variable by incorporating p-step history of the other variable. The spectral GC values can be viewed as the fraction of variance of one variable explained by the p-step history of the other, which can be obtained as (*Brovelli et al.*, 2004):

$$1-\mathrm{e}^{-GC(f)}. \tag{6.8}$$

The procedure of (6.1)-(6.7) assumes covariance stationarity, that is, the relationships between analyzed signals are time-invariant. Recently, Dhamala et al. (2008) proposed a nonparametric extension of the spectral GC method from the wavelet transforms, which eliminates the assumption of covariance stationarity and the explicit need of VAR modeling. This makes it possible to assess the time-varying properties of soil moisture-precipitation interactions. This wavelet-based GC method eliminates also the explicit need of VAR modeling, thus making the statistical assessment of the directed interactions between soil moisture and precipitation possible. However, the accuracy of spectral GC measurements relies heavily on the number of analyzed realizations; thus we subject this to gridded datasets. Suppose $X(t) = [X_1(t), X_2(t)]$ forms a realization of a bivariate process, and we have r realizations of this bivariate process. The spectral density matrix of $X(t)$ can be defined as

$$\mathbf{S} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \tag{6. 9}$$

which can be derived from the wavelet transforms of time series $X_1(t)$ and $X_2(t)$. We consider the continuous Morlet wavelet transform of a time series as W (*Torrence and Compo*, 1998), and derive the inner matrices of **S** by

$$S_{lm} = \langle W_{X_l}(t,f) \cdot W_{X_m}(t,f)^* \rangle_r \text{, with } l=1, 2 \text{ and } m=1, 2. \tag{6. 10}$$

Here, * indicates the matrix conjugate, and $\langle\rangle_r$ indicates ensemble averaging of r realizations. Using Wilson's spectral factorization algorithm (*Wilson*, 1972), the spectral matrix **S** can be factorized into a transfer matrix **H** and a noise covariance matrix $\sum$ that satisfies Equation (6.5) at each time step t. Defining the spectral factorization is complex and we refer the detailed mathematics to Dhamala et al. (2008). Then the derived **H** and $\sum$ at each time step t can be put into (6.6) and (6.7) to produce the time-varying spectral GC maps. We represent the result of our analysis using time scales, which is the inverse of frequency.

6.3 Interdependences and causal interactions

We consider each pair of soil moisture and precipitation time series over the same pixel and in the same year as one realization of the underlying bivariate stochastic process. The entire dataset constitute thus an ensemble of $5 \times 459 = 2295$ realizations. By means of ensemble averaging of these multiple realizations, the derived values can be interpreted as the averaged pixel-wise local relationships of soil moisture and precipitation in the study domain. Importantly, this ensemble averaging also reduces errors arising from possible biases of satellite datasets. In Figure 6.2, we show the ensemble averaged cross wavelet spectrum of soil moisture and precipitation (*Torrence and Compo*, 1998). A significance test is not performed in this step, as all the derived quantities have to be later subjected to spectral factorization. High spectral power is seen at timescales of 4-5 months and at timescales of 1-2 months in the summer season, indicating there are interdependences between soil moisture and precipitation at these timescales. However, these tell us little of soil moisture impact on precipitation, because high power of interdependence may occur from precipitation wetting the soil only (*Koster*, 2011).

Using the wavelet-based GC method (*Dhamala et al.*, 2008), we then decompose these interdependences into two-way interactions. The time-frequency GC maps resulting from this analysis are shown in Figure 6.3, with Figure 6.3a showing the precipitation impact on soil moisture ($GC_{P->SM}$) and Figure 6.3b the reverse soil moisture impact on precipitation ($GC_{SM->P}$). It is important to note that the GC measures only part of the statistical causality between soil moisture and precipitation. As indicated in Equation (6.4), by definition, the GC determines only the lagged predictability with model order $p>=1$. That is to say, the instantaneous influences, e.g., precipitation wetting of soil on the same day, are not captured in the GC analysis. We establish the statistical significance of the derived GC results by a random permutation approach. Surrogate realizations of the bivariate process are generated randomly using time series of soil moisture and precipitation from different original realizations, which are then subjected to the computation of spectral GC. This computation is repeated 1000 times to establish the statistical significance at 1% level. This procedure assumes spatial independence of

different pixels; thus our result reflects principally the local impacts. Importantly, this procedure preserves the spectral structure of the analyzed time series, thus reducing the possibility of obtaining high significance erroneously as a result of the serial autocorrelation.

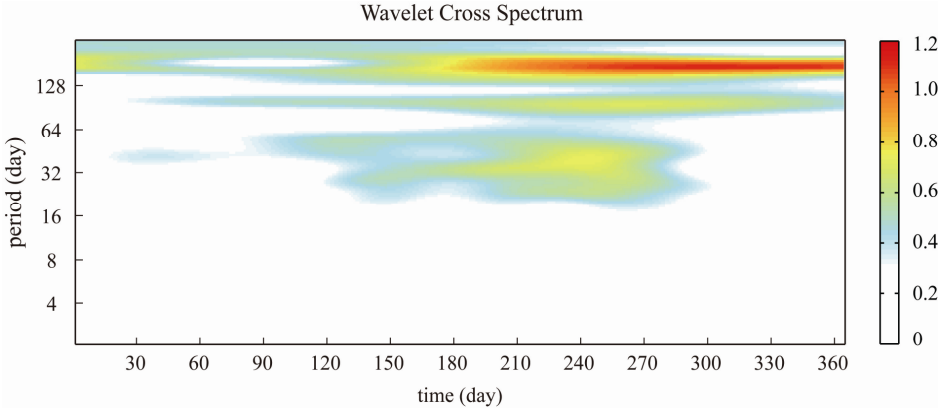

Figure 6.2 The cross wavelet spectrum of soil moisture and precipitation in a climatological year. The spectral power is obtained by ensemble averaging of the 2295 realizations in the analysis study domain.

It appears that the power peaks of the $GC_{P->SM}$ and the $GC_{SM->P}$ are located at different time scales of 4-5 months and 1-2 months respectively, although both in summer. The different locations of these power peaks imply that they are generated by different mechanisms. This suggests that we have been successful in decomposing the interdependences into two-way interactions using the GC method. It is notable that the power peak of the $GC_{SM->P}$ at time scales of 1-2 months resembles a subset of the $GC_{P->SM}$, with the latter less stronger. This raises the concern whether or not such signals in pair are generated by the same external forcing. Indeed, Orlowsky et al. (2010), using an equilibrium feedback analysis (EFA), illustrated that statistical diagnostics of the impact of soil moisture on precipitation, are often better attributed to the precipitation persistence that results from the influence of for instance sea surface temperatures (SST). The EFA method uses both the lagged cross-correlation and autocorrelation to assess the instantaneous feedbacks. It assumes a shorter persistence time of precipitation than the employed time lags which are generally of the

order of one month. This assumption can hardly be satisfied because of the external influence of SSTs. The GC method we use, although it conceptually also incorporates the lagged cross-correlation and autocorrelation, is based on the error items of the prediction. This essentially eliminates the possible erroneous signal of the impact of soil moisture on precipitation from precipitation persistence. In the Appendix of this chapter, a numerical experiment is performed to compare the performances of the EFA method and the GC method.

The significant $GC_{P->SM}$ power signifies the wetting of soils caused by persistent precipitation regimes. The strongest $GC_{P->SM}$ power appears in the summer at time scales of 4-5 months, for which the long persistence of precipitation may be best attributed to precipitable moisture advected from oceans. However, the $GC_{P->SM}$ power at time scales of 1-2 months is suggested to be of relevance to the precipitable moisture originating from surface flux. Shown in Figure 6.3b, the $GC_{SM->P}$ shows also significant power at time scales of 1-2 months in summer, indicating a soil moisture impact on subsequent precipitation. This persistent regime of soil moisture impact generates a corresponding persistent precipitation regime that keeps the soil wet in return. A signal can thus be expected to occur at the similar locations in the $GC_{P->SM}$ map if it is strong enough. It is noticeable that there exists little signal at shorter time scales for $GC_{P->SM}$ and $GC_{SM->P}$, as well as for the cross wavelet spectrum. We suggest that the ensemble averaging, as a low-pass filter, has essentially smoothed out the chaotic phenomena of rainfall which generally have smaller spatial and temporal scales, e.g., the processes associated with an individual event.

The derived $GC_{SM->P}$ suggests that soil moisture is likely to affect summer precipitation in Europe, sustained by a persistent soil moisture regime of 1-2 months. It is important to note that the "time scale" we have derived is not equivalent to the predictive time scale used in model experiments. While model experiments express straightforwardly the time leads at which soil moisture assists in predicting subsequent precipitation (*R. D. Koster et al.*, 2010; *van den Hurk et al.*, 2010), ours indicates the persistent regime of soil moisture that sustains an impact on precipitation. A schematic depiction of this result is shown in Figure 6.4 for better understanding. The GC method measures the improvement of

the precipitation predictability from incorporating backward soil moisture observations of p steps, which are segmentations of the soil moisture time scales. However, the steps of p remain hidden here because the parametric VAR modeling was eliminated. The derived spectral GC measurements can be conceptually understood as the fraction of precipitation variance explained by the p-step history of soil moisture at each time scale. Although significant in summer, the soil moisture impact on subsequent precipitation appears to be subtle, as it explains only 10-20% precipitation variance according to (6.8). One uncertainty in this aspect, which we have to point out, is that an increased number of analyzed realizations leads to reduced variance to explain.

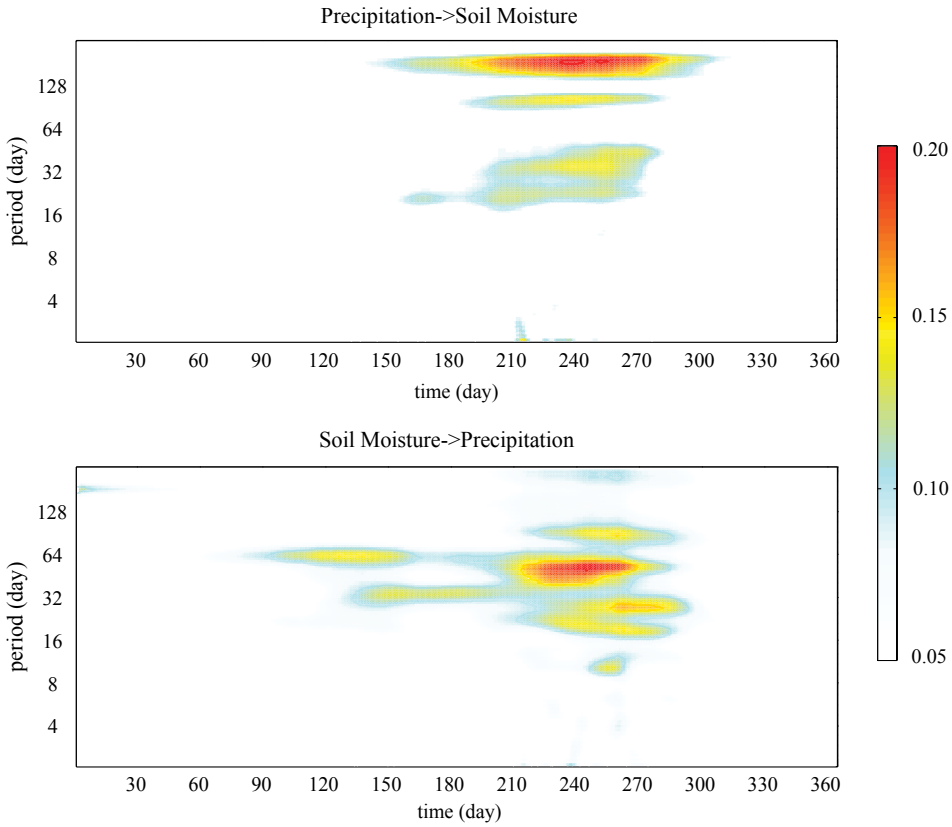

Figure 6.3 The two-way GC relations between soil moisture and precipitation over the Southern Europe. Shaded values are at 5% significance level**.**

The timing of this impact, i.e., its onset and dissipation, is suggested to relate to the evaporation process that connects soil moisture with precipitation. Evaporation over Europe is driven largely by energy constraints (*Miralles et al.*, 2011; *Teuling et al.*, 2009). In cold seasons, soils are generally saturated but the amount of evaporation is too small to sufficiently change atmospheric humidity and subsequent precipitation (*Bisselink and Dolman*, 2009). In warm seasons, the amount of evaporation becomes sufficiently large to impact precipitation. At the same time the drying soil starts to restrict evaporation. In such conditions, the variations in precipitation are likely to be partly guided by variations in soil moisture.

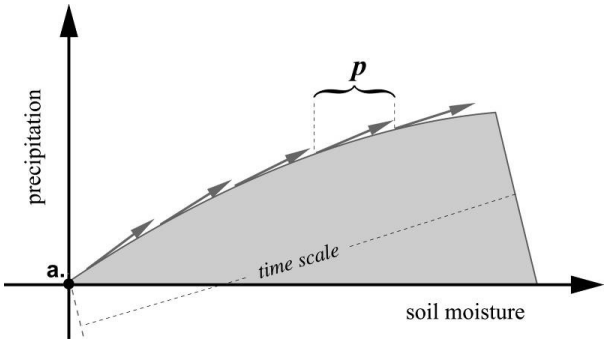

Figure 6.4 The schematic depiction of the derived soil moisture impact on subsequent precipitation. **a**. indicates one soil moisture state; x-axis indicates the evolution of soil moisture and y-axis indicates the precipitation predictability rooted in soil moisture. The shaded area indicates the soil moisture impact from soil moisture persistence of state **a.** in conjunction with the soil moisture evolution independent of that.

The underlying mechanisms of soil moisture impact on subsequent precipitation are of considerable concern (*Taylor et al.*, 2011; *D'Odorico and Porporato*, 2004; *Findell et al.*, 2011). Several studies have reported that the surface moisture flux modifies rather the probability of precipitation than the amount through the triggering mechanism, i.e., altering the boundary layer properties for precipitation generation (*D'Odorico and Porporato*, 2004; *Findell et al.*, 2011). This triggering

mechanism may also exist over Europe (*Schar et al.*, 1999) and our study certainly does not preclude it. Our analysis does support an impact of soil moisture on the precipitation amount rather than on the probability of its occurrence.

The persistent regime of soil moisture impact can be plausibly attributed to the memory scales of soil moisture, which have been estimated to be a few months over Europe (*Dirmeyer et al.*, 2009; *Wang et al.*, 2010). Soil moisture is a red noise process with strong memory. Although empirical evidence still remains rare, this strong memory is believed to be the mechanism that sustains the initial soil moisture into subsequent precipitation months later, from which the potential of subseasonal prediction is constituted (*D'Odorico and Porporato*, 2004; *Dirmeyer et al*., 2009; *Seneviratne et al.*, 2010). In our analysis, the soil moisture memory cannot be directly interpreted as the predictive time. Nonetheless, our result gives some hints that the memory of soil moisture plays indeed a role in sustaining its impact on precipitation, and thus the initial state of soil moisture has the potential to be remembered in future precipitation. Statistically speaking, the hidden p-step history constitutes the exact predictive time of precipitation that is rooted in soil moisture. However, there is also fair possibility that the expected soil moisture impact becomes too subtle, and, while still active physically, not statistically strong enough to be determined backward beyond p steps.

6.4 Conclusion

Our analysis using satellite observed soil moisture presents evidence that soil moisture affects the summer precipitation over Europe. This is important as it shows for the first time how in the real-world soil moisture-precipitation interactions take place. Soil moisture memory is thought to play an important role by sustaining the initial mpact at time scales of 1-2 months. It is well known that the current generation of climate models has difficulties in predicting convective precipitation over Europe. The comprehensive multiple-model Global Land-Atmosphere Coupling Experiment, and its second phase with realistic soil moisture initializations, has demonstrated little soil moisture impact on subsequent precipitation over Europe (*Koster et al.*, 2004; *van den Hurk et al.*, 2010).

Under such conditions of poor understanding, our observational analysis provides a valuable metric for benchmarking climate models and validating model-based analyses of soil moisture-precipitation interactions. This is of value also for understanding the predictive capability for hydroclimatic extremes, such as the summer droughts in the present and future climate over Europe (*Pal et al.*, 2004; *Seneviratne et al.*, 2006).

Appendix

Based on the stochastic climate theory, the EFA method was designed to assess the instantaneous influence of a slow-varying variable (e.g., SST) on a fast atmospheric variable (*Frankignoul et al*., 1977). Recently it is widely used in assessing land-atmosphere interaction (e.g., *Notaro et al*., 2008; *Notaro et al*., 2006; *Zhang et al*., 2008; *Liu et al*., 2006). Below with a numerical experiment, we demonstrate that the EFA tends to produce statistical pitfalls of the expected feedbacks when the atmospheric variable exhibits persistence; however, the spectral GC method we have used in this study does not show this behavior.

The EFA method assumes an atmospheric variable a to consist of two parts:

$$a(t+dt)=\lambda b(t)+\varepsilon(t+dt), \tag{6.11}$$

where $\lambda b(t)$ represents the atmospheric response to a change in the low-varying variable b after time dt, and $\varepsilon(t+dt)$ represents the noise generated by atmospheric internal variability. By taking the covariance of both sides with $b(t-\tau)$, Equation (6.9) becomes

$$\lambda = \frac{\mathrm{cov}[b(t-\tau), a(t+dt)]}{\mathrm{cov}[b(t-\tau), b(t)]}, \tag{6.12}$$

where λ is the feedback parameter. The persistence time of the atmospheric internal variability is usually shorter than one month. Therefore, dt can be neglected if monthly data are used, leading to a simplified estimator which is generally used

$$\lambda = \frac{\mathrm{cov}[b(t-\tau), a(t)]}{\mathrm{cov}[b(t-\tau), b(t)]}, \tag{6.13}$$

An apparent limitation of this EFA estimator is that it neglects the persistent regimes of atmospheric variables that are due to external forcings, which leads to statistical pitfalls when assessing land surface feedbacks on overlying atmosphere.

We consider a bivariate autoregressive process of X=[x_1, x_2], following *Dhamala* et al. (2008), with

$$\begin{aligned} x_1(t) &= 0.55x_1(t-1) - 0.8x_1(t-2) + c_{2\to 1}(t)\,x_2(t-1) + \varepsilon_1(t) \\ x_2(t) &= 0.55x_2(t-1) - 0.8x_2(t-2) + c_{1\to 2}(t)\,x_1(t-1) + \varepsilon_2(t) \end{aligned}. \qquad (6.14)$$

Here, t is an index of 900 time points, $\varepsilon_1(t)$ and $\varepsilon_2(t)$ are two different white noise with zeros means and unit variance, $c_{2->1}(t)$ is the feedback strength of x_2 on x_1, and $c_{1->2}(t)$ is the feedback strength of x_1 on x_2. The $c_{1->2}(t)$ is set to be 0.25 for $0<t<425$ and 0 for $425<t<900$; the $c_{2->1}(t)$ is set to be 0 for $0<t<425$ and 0.25 for $425<t<900$. These designed coupling strengths are shown in Figure 6.5a. By numerical control, x_1 and x_2 both have a persistent mode (period) of 5 data points. We generate 5000 realizations of this bivariate process.

We first estimate the EFA feedback parameters, as defined in (6.12). To get the time-varying estimates, the EFA estimator is applied to the bivariate process (x_1 and x_2) with a sliding window of 50 data points. The time lag is selected to be τ=1. Note that, x_1 and x_2 both have a persistent mode of 5 data points. This is to create an analog of the EFA practice in assessing land-atmosphere coupling. Generally, monthly data and a time lag of one month are used in such studies. However, the atmospheric variables are very likely to have persistence time of 3-5 months owing to SST forcings. In our case, as shown in Figure 6.2 and 6.3, the precipitation shows indeed a persistence time up to 4-5 months. The computation is done realization by realization, and the averaged EFA estimates of 5000 realizations are shown in Figure 6.5b. It appears that the EFA estimates yield correctly the designed x_1 influence on x_2 for the first half analyzed period and the x_2 influence on x_1 for the second half of the analyzed period. However, it appears also clearly that the EFA estimates yield pitfall of x_2 influence on x_1 for the first half analyzed period and x_1 influence on x_2 for the second half analyzed period. Note that zero EFA values indicate the absence of feedbacks.

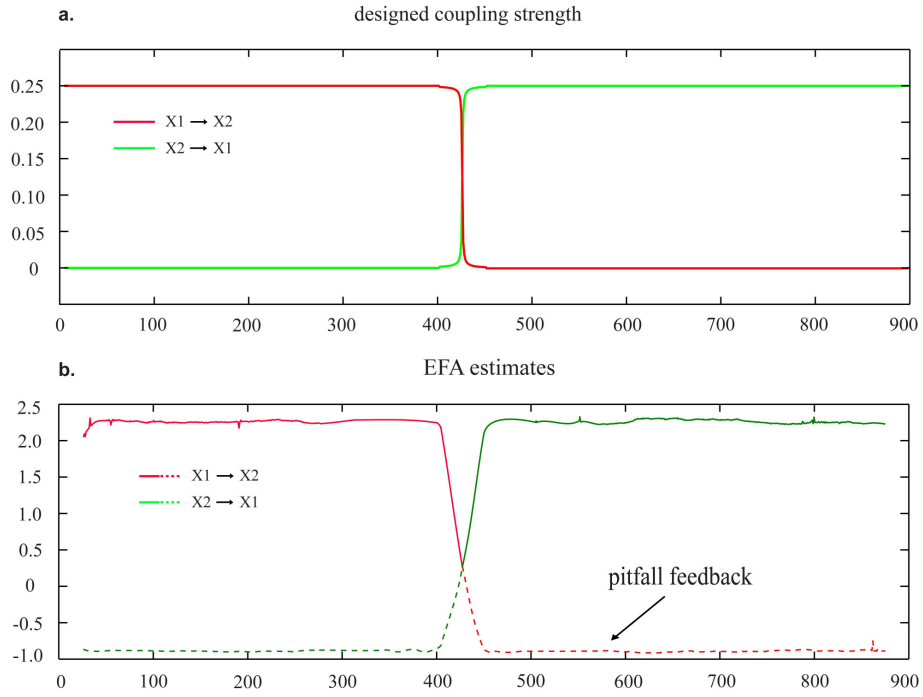

Figure 6.5 The designed coupling strength and the averaged EFA estimates of 5000 realizations. The dashed lines in **b**. indicate statistical pitfall of the feedbacks.

Consider now the analog of the soil moisture feedback on precipitation. The influence of precipitation on soil moisture is obvious, with possible persistence times longer than months owing to SSTs. However, the soil moisture influence on precipitation may not exist or be very subtle in observations. On such circumstance, statistical pitfall of soil moisture feedback on precipitation can be easily created by the EFA method.

Next, we subject the 5000 realizations of the designed bivariate process to the wavelet-based GC method. The derived results are shown in Figure 6.6. To demonstrate the details, these results are not subjected to significance tests. It appears that the GC estimates have correctly yielded the designed the two-way coupling as shown in Figure 6.5a, without statistical pitfall. This implies the wavelet-based GC method is exempt from the precipitation persistence, and thus valid for assessing the soil moisture feedback on subsequent precipitation.

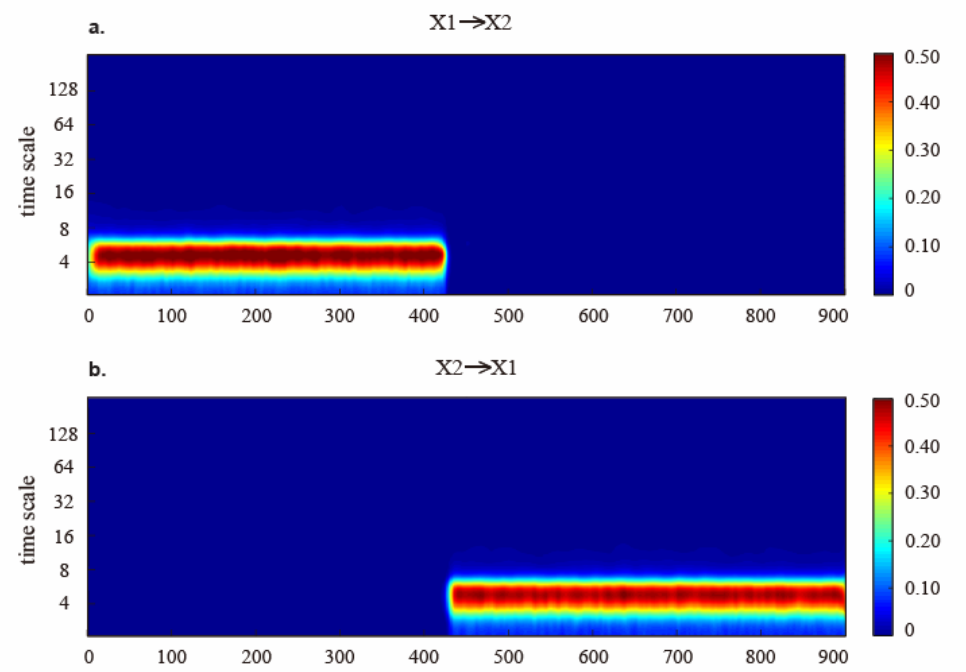

Figure 6.6 The directed couplings estimated from wavelet-based GC method. There is no significance test applied in this numerical example.

References

Bisselink, B., and A. J. Dolman (2009), Recycling of moisture in Europe: contribution of evaporation to variability in very wet and dry years, *Hydrol Earth Syst Sc*, *13*(9), 1685-1697.

Brovelli, A., M. Z. Ding, A. Ledberg, Y. H. Chen, R. Nakamura, and S. L. Bressler (2004), Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality, *P Natl Acad Sci USA*, *101*(26), 9849-9854.

Christopher M. Taylor, Amanda Gounou, Françoise Guichard, Phil P. Harris, Richard J. Ellis, F. Couvreux, and M. D. Kauwe (2011), Frequency of Sahelian storm initiation enhanced over mesoscale soil-moisture patterns, *Nature Geoscience*, *4*, 430-433.

Cook, B. I., G. B. Bonan, and S. Levis (2006), Soil moisture feedbacks to precipitation in southern Africa, *J Climate*, *19*(17), 4198-4206.

D'Odorico, P., and A. Porporato (2004), Preferential states in soil moisture and climate dynamics, *P Natl Acad Sci USA*, *101*(24), 8848-8851.

de Jeu, R. A. M., W. Wagner, T. R. H. Holmes, A. J. Dolman, N. C. van de Giesen, and J. Friesen (2008), Global Soil Moisture Patterns Observed by Space Borne Microwave Radiometers and Scatterometers, *Surv Geophys*, *29*(4-5), 399-420.

Dhamala, M., G. Rangarajan, and M. Ding (2008), Estimating granger causality from fourier and wavelet transforms of time series data, *Phys Rev Lett*, *100*(1), 018701.

Dhamala, M., Rangarajan, G., & Ding, M. (2008), Analyzing information flow in brain networks with nonparametric Granger causality. *NeuroImage*, *41*, 354–362.

Dirmeyer, P. A., R. D. Koster, and Z. C. Guo (2006), Do global models properly represent the feedback between land and atmosphere?, *J Hydrometeorol*, *7*(6), 1177-1198.

Dirmeyer, P. A., C. A. Schlosser, and K. L. Brubaker (2009), Precipitation, Recycling, and Land Memory: An Integrated Analysis, *J Hydrometeorol*, *10*(1), 278-288.

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

Garcia, D. (2010), Robust smoothing of gridded data in one and higher dimensions with missing values, *Comput Stat Data An*, *54*(4), 1167-1178.

Geweke, J. (1982), Measurement of linear-dependence and feedback between multiple time-series, *J Am Stat Assoc*, *77*(378), 304-313.

Granger, C. W. J. (1969), Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, *37*(3), 414-438.

Hohenegger, C., P. Brockhaus, C. S. Bretherton, and C. Schar (2009), The Soil Moisture-Precipitation Feedback in Simulations with Explicit and Parameterized Convection, *J Climate*, *22*(19), 5003-5020.

Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind (2001), Global precipitation at one-degree daily resolution from multisatellite observations, *J Hydrometeorol*, *2*(1), 36-50.

Jevrejeva, S., J. C. Moore, and A. Grinsted (2003), Influence of the arctic oscillation and El Nino-Southern Oscillation (ENSO) on ice conditions in the Baltic Sea: The wavelet approach, *J Geophys Res-Atmos*, *108*(D21), 4677, doi:10.1029/2003JD003417.

Kaufmann, R. K., and D. I. Stern (1997), Evidence for human influence on climate from hemispheric temperature relations, *Nature*, *388*(6637), 39-44.

Kirsten L. Findell, Pierre 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.

Koster, R. D. (2011), Climate science: Storm instigation from below, *Nat Geosci*. 4, 427-428.

Koster, R. D., et al. (2010), Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment, *Geophys Res Lett*, *37*, L02402, doi:10.1029/2009GL041677.

Koster, R. D., et al. (2004), Regions of strong coupling between soil moisture and precipitation, *Science*, *305*(5687), 1138-1140.

Koster, R. D., et al. (2006), GLACE: The Global Land-Atmosphere Coupling Experiment. Part I: Overview, *J Hydrometeorol*, *7*(4), 590-610.

Liu, Z., M. Notaro, J. Kutzbach, and N. Liu (2006), Assessing global vegetation-climate feedbacks from observations, *J. Clim*., 19, 787–814, doi:10.1175/JCLI3658.1.

Miralles, D. G., R. A. M. De Jeu, J. H. Gash, T. R. H. Holmes, and A. J. Dolman (2011), Magnitude and variability of land evaporation and its components at the global scale, *Hydrol Earth Syst Sc*, *15*(3), 967-981.

Notaro, M. (2008), Statistical identification of global hot spots in soil moisture feedbacks among IPCC AR4 models, *J. Geophys. Res.*, 113, D09101, doi:10.1029/2007JD009199.

Notaro, M., Z. Liu, and J. W. Williams (2006), Observed vegetation-climate feedbacks in the United States, *J. Clim., 19*, 763–786, doi:10.1175/JCLI3657.1.

Orlowsky, Boris, Sonia I. Seneviratne (2010), Statistical Analyses of Land–Atmosphere Feedbacks and Their Possible Pitfalls, *J. Clim., 23*, 3918–3932.

Owe, M., R. de Jeu, and T. Holmes (2008), Multisensor historical climatology of satellite-derived global land surface moisture, *J Geophys Res-Earth*, *113*(F1), F01002, doi:10.1029/2007JF000769.

Pal, J. S., F. Giorgi, and X. Q. Bi (2004), Consistency of recent European summer precipitation trends and extremes with future regional climate projections, *Geophys Res Lett*, *31*(13), L13202, doi:10.1029/2004GL019836.

Schar, C., D. Luthi, U. Beyerle, and E. Heise (1999), The soil-precipitation feedback: A process study with a regional climate model, *J Climate*, *12*(3), 722-741.

Seneviratne, S. I., D. Luthi, M. Litschi, and C. Schar (2006), Land-atmosphere coupling and climate change in Europe, *Nature*, *443*(7108), 205-209.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, *Earth-Sci Rev*, *99*(3-4), 125-161.

Teuling, A. J., et al. (2009), A regional perspective on trends in continental evaporation, *Geophys Res Lett*, *36*, L02404, doi:10.1029/2008GL036584.

Torrence, C., and G. P. Compo (1998), A practical guide to wavelet analysis, *B Am Meteorol Soc*, *79*(1), 61-78.

van den Hurk, B., F. Doblas-Reyes, G. Balsamo, R. Koster, S. Seneviratne, and H. Camargo (2010), Soil moisture effects on seasonal

temperature and precipitation forecast scores in Europe, *Clim Dynam*, 1-14, 10.1007/s00382-010-0956-2.

Vautard, R., P. Yiou, F. D'Andrea, N. de Noblet, N. Viovy, C. Cassou, J. Polcher, P. Ciais, M. Kageyama, and Y. Fan (2007), Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit, *Geophys Res Lett*, *34*(7), L07711, doi:10.1029/2006GL028001.

Wang, G. J., A. J. Dolman, R. Blender, and K. Fraedrich (2010), Fluctuation regimes of soil moisture in ERA-40 re-analysis data, *Theor Appl Climatol*, *99*(1-2), 1-8.

Wilson, G. T. (1972), Factorization of Matricial Spectral Densities, *Siam J Appl Math*, *23*(4), 420-426.

Chapter 7

Summary and Perspectives

7.1 Summary

Recently Mediterranean Europe has suffered from increasing occurrences of summer droughts and heat waves that have caused substantial societal and ecological impacts (*Pal et al.*, 2004; *van Oldenborgh et al.*, 2009; *Vautard et al.*, 2007). Although such extreme conditions are generally associated with specific large-scale regimes, model simulations have suggested the possible enhancement by soil moisture depletion (*Fischer et al.*, 2007; *Seneviratne et al.*, 2006; *Teuling et al.*, 2010; *Vautard et al.*, 2007). Given the severe scarcity of empirical evidence, this thesis has attempted to detect signals of variations in the overlying atmosphere that can be attributed to soil moisture feedback, using observational datasets. This feedback may exist over a variety of time scales. The focus was thus to understand whether or not the land surface processes, with soil moisture as the main agent, can contribute to subseasonal, interannual as well as interdecadal climate variability over Europe.

Chapter 2 describes the statistical properties of soil moisture time series. This is of importance for knowing the memory of soil moisture, and thus the possible persistence of soil moisture impact on atmosphere. Furthermore, these statistical features of the soil moisture dataset serve as *a priori* guidance to the signal interpretation and statistical modeling in the following analysis. The soil moisture memory over Europe is estimated to be roughly 1-3 months; within this regime, soil moisture exhibits a red spectrum and nonstationarity fluctuation properties. However, the soil moisture fluctuation in volatility exhibits a white spectrum, indicating that soil moisture is statistically a linear process.

This linearity of soil moisture indicates that the soil moisture time series can be well approximated by linear models.

Chapter 3 studies the interannual variability of summer temperature and drought in response to initial soil moisture states, as proxied by the accumulated precipitation, of later winter (Jan.-Mar.). Significant responses are found over the Mediterranean Europe. We estimate that roughly 10 to 5% of the interannual variability of summer temperature over the Mediterranean Europe can be explained by initial soil moisture states. The value for summer drought increases to 10-25%. In agreement with some numerical experiments (*Fischer et al.*, 2007; *Seneviratne et al.*, 2006), these results suggest seasonal predictability of temperature and drought conditions of Mediterranean summer.

Chapter 4 studies the possible causes of the marked summer temperature variability over Europe at multidecadal time scales. This multidecadal variability is generally considered to be a response to the Atlantic Multidecadal Oscillation (AMO) (*Huss et al.*, 2010; *Keenlyside et al.*, 2008; *Kerr*, 2000; *Knight et al.*, 2006; *Knight et al.*, 2005; *Sutton and Hodson*, 2005). In our rigorous statistical analysis, the oceanic AMO variability is found to explain only half of the magnitude of the observed magnitude of the AMO-like variability over the Mediterranean Europe. The full magnitude can be explained only as amplification in the background warming; and the amplification effect is quantified to be a factor of roughly 2. Over southeast Europe, where there exists the strongest multidecadal variability, the background warming doesn't show the amplification effect. This implies that the multidecadal variability over southeast Europe is more likely an amplification by local processes. There is numerical and observational evidence that soil moisture can have very strong impact on temperature over southeast Europe, where drier soil can induce a self-stimulating heat low response (*Haarsma et al.*, 2009; *Hirschi et al.*, 2011). Thus we suggest that such a response may be initialized by the AMO condition at the multidecadal frequency band.

In **Chapter 6**, attempts are made to quantify the direction of the interactions between soil moisture and precipitation as well as temperature using gridded daily observations (*de Jeu et al.*, 2008; *Owe et al.*, 2008). The soil moisture dataset from AMSR-E satellite images suffers from the inherent data gaps. As a step of data preprocessing,

Chapter 5 introduces a thin-spline smoother of triple dimensionality to fill the data gaps (*Garcia*, 2010). **Chapter 6** quantifies the directions of the interaction between soil moisture and precipitation. A nonparametric approach based on spectral factorization of wavelet transforms is used to achieve a time-frequency presentation of the Granger causality (*Dhamala et al.*, 2008; *Geweke*, 1982; *Granger*, 1969). The interdependences between soil moisture and precipitation are decomposed into two-way causal interactions. We find a significant signal of the impact of soil moisture on summer precipitation over Europe. Soil moisture memory is identified as the mechanism that sustains this impact at time scales of 1-2 months. This result sheds light on the soil moisture-precipitation interactions as found in the real world rather than in modeling results.

7.2 Research perspectives

The climate system is full of interactions, and the interactions at the land-atmosphere interface are a key player in the changing climate (*Seneviratne et al.*, 2006; *Seneviratne et al.*, 2010). Modeling studies addressing land-atmosphere feedbacks often display a large divergence between models, without being able to explain exactly why the models have these different responses. For instance, the comprehensive multiple-model Global Land-Atmosphere Coupling Experiment, and its second phase with realistic soil moisture initializations, has not demonstrated any significant soil moisture feedback on subsequent precipitation over Europe (*Koster et al.*, 2004; *van den Hurk et al.*, 2010). In contrast, observational evidence tends to be hard to find largely because soil moisture observations are very few, but when available they may be made into benchmark datasets that allow us to validate model-based studies. Both modeling and empirical studies have limitations, and have to be employed in conjunction for improved understanding of the coupled land-atmosphere system.

Substantial progress in retrieving land surface properties, especially soil moisture (*de Jeu et al.*, 2008; *Owe et al.*, 2008) and evaporation (*Miralles et al.*, 2011a; *Miralles et al.*, 2011b), from space has now made an empirically based analysis possible. However, due to the different directions of the interactions complications easily arise (e.g Figure 1.1). Correlation or coherence analysis unfortunately indicates little about the

impact of soil moisture on the lower atmosphere, as precipitation simply wets or temperature simply dries the soil in both directions (*Koster* 2011). Because of the availability of new satellite datasets, it is now timely to conceive a statistical strategy for disentangling the directions of interactions from observations, and thus looking more directly into the processes itself. The statistical concept of Granger causality is an excellent tool for such a strategy (Geweke, 1982; *Granger*, 1969).

However difficulties exist in making this happen, particularly in studies of land-atmosphere interactions using daily datasets, because of the nonstationarity involved in the time-solving interactions. By means of a spectral factorization of a wavelet matrix, **Chapter 6** demonstrates an application in quantifying the direction of the interactions between soil moisture and precipitation. However, the accuracy of this approach relies heavily on the number of realizations to analyze, which limits it use to gridded datasets. This is a condition that can hardly be satisfied by observational datasets. For observations with only one realization, ultimately we have to go to parametric approaches for estimating the Granger causality, for example, when using Fluxnet observations. To analyse the nonstationary (time-varying) properties of the land-atmosphere interactions, alternatively we need to perform adaptive VAR modeling, for which Kalman filtering can be used (*Havlicek et al.*, 2010; *Milde et al.*, 2010). However, regarding the detection of soil moisture-atmosphere, there remains a special challenge in parametric modeling. The daily soil moisture time series is a red noise process, leading to severe problems of unbalance variances with atmospheric variables. This problem may cause model misspecification and the variable with higher variance can be easily misinterpreted as the causal source. Further efforts are needed to find a solution of the VAR modeling that is immune to this unbalanced variance problem. This forms a promising research direction.

While empirical evidence is necessary for validating model-based analysis, it can also be easily misleading, particularly with respect to the magnitude of the relationships. In **Chapter 4**, the causes of the multidecadal variability of summer temperature over Southern Europe are studied statistically. This temperature variability is generally thought to result from a response to the Atlantic Multidecadal Oscillation (AMO) of the sea surface temperatures. However, we determined that the oceanic AMO forcing explains only half magnitude of the multidecadal

variability over the Mediterranean Europe. Its full magnitude can be explained only as an amplified response in presence of the background warming. The data is thus contaminated. Under such circumstances, the standard empirical analysis may overestimate the oceanic forcing, aliasing the background warming. This may occur with many other processes where positive feedbacks are involved.

The amplification over the Mediterranean Europe is likely to result from positive feedbacks associated with the background warming related to air-sea interactions. It is not yet fully known whether a soil moisture feedback plays a role in this amplification. The amplification over Southeast Europe may well be related to the soil moisture feedback, as there exists modeling evidence that such a feedback may cause the heat low response of temperature (*Haarsma et al.*, 2009; *Hirschi et al.*, 2011). If this hypothesis is true, it may improve our understanding how soil moisture feedbacks evantually contribute to climate variability at multidecadal time scales. However, statistical analysis from instrumental records alone, unfortunately, cannot provide a fully unambiguous picture. Further modeling studies are needed to identify the precise feedback mechanisms involved in this amplification.

References

de Jeu, R. A. M., W. Wagner, T. R. H. Holmes, A. J. Dolman, N. C. van de Giesen, and J. Friesen (2008), Global Soil Moisture Patterns Observed by Space Borne Microwave Radiometers and Scatterometers, *Surv Geophys*, *29*(4-5), 399-420.

Dhamala, M., G. Rangarajan, and M. Ding (2008), Estimating granger causality from fourier and wavelet transforms of time series data, *Phys Rev Lett*, *100*(1), 018701.

Fischer, E. M., S. I. Seneviratne, D. Luthi, and C. Schar (2007), Contribution of land-atmosphere coupling to recent European summer heat waves, *Geophys Res Lett*, 34(6), L06707, doi:10.1029/ 2006GL029068.

Garcia, D. (2010), Robust smoothing of gridded data in one and higher dimensions with missing values, *Comput Stat Data An*, *54*(4), 1167-1178.

Geweke, J. (1982), Measurement of Linear-Dependence and Feedback between Multiple Time-Series, *J Am Stat Assoc*, *77*(378), 304-313.

Granger, C. W. J. (1969), Investigating Causal Relations by Econometric Models and Cross-Spectral Methods, *Econometrica*, *37*(3), 414-438.

Haarsma, R. J., F. Selten, B. V. Hurk, W. Hazeleger, and X. L. Wang (2009), Drier Mediterranean soils due to greenhouse warming bring easterly winds over summertime central Europe, *Geophys Res Lett*, 36, L04705, doi:10.1029/2008GL036617.

Havlicek, M., J. Jan, M. Brazdil, and V. D. Calhoun (2010), Dynamic Granger causality based on Kalman filter for evaluation of functional network connectivity in fMRI data, *Neuroimage*, *53*(1), 65-77.

Hirschi, M., S. I. Seneviratne, V. Alexandrov, F. Boberg, C. Boroneant, O. B. Christensen, H. Formayer, B. Orlowsky, and P. Stepanek (2011), Observational evidence for soil-moisture impact on hot extremes in southeastern Europe, *Nat Geosci*, *4*(1), 17-21.

Huss, M., R. Hock, A. Bauder, and M. Funk (2010), 100-year mass changes in the Swiss Alps linked to the Atlantic Multidecadal Oscillation, *Geophys Res Lett*, *37*, L10501, doi:10.1029/2010GL042616.

Keenlyside, N. S., M. Latif, J. Jungclaus, L. Kornblueh, and E. Roeckner (2008), Advancing decadal-scale climate prediction in the North Atlantic sector, *Nature*, *453*(7191), 84-88.

Kerr, R. A. (2000), A North Atlantic climate pacemaker for the centuries, *Science*, *288*(5473), 1984-1986.

Knight, J. R., C. K. Folland, and A. A. Scaife (2006), Climate impacts of the Atlantic Multidecadal Oscillation, *Geophys Res Lett*, *33*(17), L17706, doi:10.1029/2006GL026242.

Knight, J. R., R. J. Allan, C. K. Folland, M. Vellinga, and M. E. Mann (2005), A signature of persistent natural thermohaline circulation cycles in observed climate, *Geophys Res Lett*, *32*(20), 32, L20708, doi:10.1029/2005GL024233.

Koster, R. D. (2011), Climate science: Storm instigation from below, *Nat Geosci*. 4, 427-428.

Milde, T., L. Leistritz, L. Astolfi, W. H. R. Miltner, T. Weiss, F. Babiloni, and H. Witte (2010), A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate AR models and its application in analysis of laser-evoked brain potentials, *Neuroimage*, *50*(3), 960-969.

Miralles, D. G., R. A. M. De Jeu, J. H. Gash, T. R. H. Holmes, and A. J. Dolman (2011a), Magnitude and variability of land evaporation and its components at the global scale, *Hydrol Earth Syst Sc*, *15*(3), 967-981.

Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman (2011b), Global land-surface evaporation estimated from satellite-based observations, *Hydrol Earth Syst Sc*, *15*(2), 453-469.

Owe, M., R. de Jeu, and T. Holmes (2008), Multisensor historical climatology of satellite-derived global land surface moisture, *J Geophys Res-Earth*, 113, F01002, doi:10.1029/2007JF000769.

Pal, J. S., F. Giorgi, and X. Q. Bi (2004), Consistency of recent European summer precipitation trends and extremes with future regional climate projections, *Geophys Res Lett*, *31*(13), L13202, doi:10.1029/2004GL019836.

Seneviratne, S. I., D. Luthi, M. Litschi, and C. Schar (2006), Land-atmosphere coupling and climate change in Europe, *Nature*, *443*(7108), 205-209.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling (2010), Investigating soil moisture-climate interactions in a changing climate: A review, *Earth-Sci Rev*, *99*(3-4), 125-161.

Sutton, R. T., and D. L. R. Hodson (2005), Atlantic Ocean forcing of North American and European summer climate, *Science*, *309*(5731), 115-118.

Teuling, A. J., et al. (2010), Contrasting response of European forest and grassland energy exchange to heatwaves, *Nat Geosci*, *3*(10), 722-727.

van Oldenborgh, G. J., S. Drijfhout, A. van Ulden, R. Haarsma, A. Sterl, C. Severijns, W. Hazeleger, and H. Dijkstra (2009), Western Europe is warming much faster than expected, *Clim Past*, *5*(1), 1-12.

Vautard, R., P. Yiou, F. D'Andrea, N. de Noblet, N. Viovy, C. Cassou, J. Polcher, P. Ciais, M. Kageyama, et al. (2007), Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit, *Geophys Res Lett*, *34*(7), L07711, doi:10.1029/2006GL028001.

Wang, G. J., A. J. Dolman, R. Blender, and K. Fraedrich (2010), Fluctuation regimes of soil moisture in ERA-40 re-analysis data, *Theor Appl Climatol*, *99*(1-2), 1-8.

Chapter 8

Samenvatting en verdere perspectieven voor onderzoek

8.1 Samenvatting

De laatste tien jaar heeft het Mediterrane gebied in Europe geleden onder periodes van grote droogte en hittegolven. Deze droogteperioden en hittegolven hebben een grote invloed op het functioneren van het maatschappelijk leven en op de natuur. Meestal worden deze bijzondere omstandigheden toegeschreven aan de aanwezigheid specifieke grootschalige circulatie patronen. Simulaties met meteorologische modellen suggereren echter dat bodemvocht ook een rol kan spelen bij het ontstaan en handhaven van deze weerscondities.

Dit proefschrift poogt, ondanks de vaak gebrekkige beschikbaarheid van gegevens, juist die signalen te bepalen, die te herleiden zijn op de invloed van bodemvocht op de variaties van temperatuur in de onderste lagen van de atmosfeer. Deze terugkoppeling kan bestaan op verschillende tijdsschalen. Onze grootste aandacht gaat uit naar het begrijpen hoe landoppervlakteprocessen, specifiek bodemvocht, bijdragen aan sub-seizoenale, jaarlijkse en decadale klimaatvariabiliteit in Europa.

Hoofdstuk 2 beschrijft de statistiek van tijdseries van bodemvocht in de ERA-40 her-analyse. Dit is van belang om het "geheugen" van bodemvocht goed te kunnen bepalen en de mogelijke doorwerking van dat geheugen op de atmosfeer. Bovendien geeft deze analyse ons een eerste inzicht in, en ondersteuning voor onze verdere statistische analyses. Het geheugen van bodemvocht kent een karakteristiek tijdschaal van ongeveer 1-3 maanden. Bodemvocht kenmerkt zich hierbij als een "rode ruis" proces met niet-stationaire fluctuaties. De fluctuaties

in het volatiliteit deel laten echter een “witte ruis” zien. Dit suggereert dat bodemvocht zich hier gedraagt als een lineair proces dat gemodelleerd kan worden met lineair statistische modellen.

Hoofdstuk 3 bestudeert de invloed van bodemvocht, voorgesteld door de som van de neerslag in de late winter (jan-maart), op de jaarlijkse variabiliteit van zomer temperatuur en droogte. We vinden hierbij een statistisch significante relatie over Mediterraan Europa die 10-15% van de jaarlijkse variabiliteit in zomer temperatuur kan verklaren. Bij droogte vinden we een sterkere relatie, die 10-25% kan verklaren. Deze resultaten zijn in overeenstemming met model experimenten die voorspelbaarheid op seizoenschaal aantonen voor met name zomertemperatuur en droogte in het Europese Mediterrane gebied.

Hoofdstuk 4 bestudeert de mogelijke oorzaken van de aantoonbare variatie in zomertemperaturen op multidecadale tijdschaal (orde >10 jaar). Men neemt aan dat deze variabiliteit een gevolg is van de Atlantische Multidecadale Oscillatie (AMO). In onze nieuwe, robuuste statische analyse vinden we echter dat de AMO variatie maar de helft van de waargenomen variatie in het Mediterrane gebied kan verklaren die doorgaans geheel aan de AMO wordt toegeschreven. De volledige grootte van het signaal kan enkel worden verklaard als een versterking van het originele signaal tegen de achtergrond van de algehele opwarming van de aarde. Deze versterking maakt het oorspronkelijke signaal groter met een factor twee. In Zuid Oost Europa, waar we de sterkste multidecadale variabiliteit aantreffen zien we deze algemene versterking niet, maar lijkt het erop dat lokale processen, zoals bodemvocht, een rol spelen. Model experimenten hebben voor dit gebied ook laten zien dat bodemvocht een grote rol speelt bij de totstandkoming en handhaving van lage druk gebieden. Het mechanisme waarbij dit gebeurt is dat een uitdrogende bodem tot hogere temperaturen leidt, die zo de uitdroging verder kunnen versterken. Onze analyse ondersteunt de analyse dat zo’n situatie geinitialiseerd kan worden door specifieke AMO periodes en dat deze zichtbaar wordt in de multidecadale frequenties van temperatuurfluctuaties.

In Hoofdstuk 6 ondernemen we een poging de richting van de interactie tussen bodemvocht en neerslag en temperatuur te bepalen. We gebruiken

hiervoor dagelijkse waarnemingen van bodemvocht afkomstig van satellietwaarnemingen. De bodemvocht data set is gebaseerd op AMSR-E metingen. Helaas missen in deze data set vele waarneming als gevolg van, onder andere, problemen met de satelliet. Die ontbrekende data moeten we ook berekenen en daarvoor is een "thin-spline smoother" ontwikkeld die in drie dimensies (lengte, breedte, tijd) de ontbrekende data kan berekenen. Hoofdstuk 6 berekent dan de richting van de interactie tussen bodemvocht en neerslag. We gebruiken hiervoor een niet-parametrische methode, die gebaseerd is op de spectrale omvorming van een wavelet. Dit geeft ons een beschrijving van de zogenoemde "Granger Causality" in het tijd-frequentie domein. De wederzijdse afhankelijkheid van bodemvocht en neerslag kan zo ontrafeld worden in twee aparte richtingen en we kunnen de grootte van die causaliteit dan berekenen. We vinden een duidelijk signaal van het bodemvocht op de zomer neerslag. Het bodemvochtgeheugen speelt hierbij een belangrijke rol om het signaal door te geven op tijdschalen van 1-2 maanden. Deze resultaten laten voor het eerst zien hoe bodemvocht-neerslag interacties zich in de realiteit afspelen.

8.2 Perspectief voor verder onderzoek

Het klimaatsysteem kent vele interacties. Model studies hebben laten zien dat juist de interacties tussen het land en de atmosfeer spelen een belangrijke rol in een veranderend klimaat. Er bestaan echter grote verschillen tussen die modellen en het blijkt vaak moeilijk precies aan te geven wat de oorzaak is van de gevonden verschillen. Als voorbeeld, het Global Land-Atmosphere Coupling Experiment (GLACE-II) heeft met realistische initialisatie van bodemvocht niet kunnen aantonen dat er een significante relatie bestond tussen bodemvocht en neerslag in Europa. Helaas is bewijs voor dat soort relaties moeilijk te krijgen als gevolg van gebrekkige waarnemingsreeksen. Zouden die datasets er wel komen, dan kunnen ze als "benchmarks" van modellen gebruikt worden. Hoe het ook zij, zowel model als op waarneming gebaseerde studies hebben zo hun beperkingen en voorlopig valt de meeste voortgang te boeken door ze allebei, liefst in samenhang, te gebruiken.

Er is de laatste veel vooruitgang geboekt bij het berekenen van bodemvocht en verdamping met behulp van satellietwaarnemingen. Dit heeft het mogelijk gemaakt om onze statische analyses te maken. Helaas is deze analyse niet altijd even eenduidig, omdat bodemvocht en neerslag elkaar wederzijds beïnvloeden. Correlatie, of coherentie analyse zijn daarbij van weinig nut, omdat ze geen onderscheid maken in de richting van beïnvloeding. Gelukkig kunnen we wel dankzij de nieuwe satellietwaarnemingen een strategie uitstippelen voor een statische aanpak. Als het lukt om de richting te bepalen van de interacties krijgen we immers meer zicht op de processen van die beïnvloeding dan voorheen. Het door ons gebruikte concept van "Granger Causality" is hierbij een uitstekend hulpmiddel.

De toepassing van deze concepten is echter niet makkelijk noch eenduidig. Het gebruik van dagelijkse waarnemingen leidt er toe dat de tijdreeksen niet meer stationair zijn omdat er dan in de reeksen seizoen en andere variaties op treden. Door middel van spectrale ontbinding in factoren van een wavelet matrix, is het ons gelukt in Hoofdstuk 6 de richting van de interacties te bepalen. Helaas is de nauwkeurigheid van die methode beperkt door het aantal realisaties dat we kunnen maken, en is de methode dan ook beperkt tot zogenaamde op grids gebaseerde datasets (zoals afkomstig uit meteorlogische of klimaat modellen). Voor waarnemingen met enkel één realisatie, zoals bijvoorbeeld van flux meetmasten (Fluxnet) hebben we parametrische methoden nodig. Om de niet stationaire eigenschappen van land atmosfeer interacties te kunnen bepalen kunnen we mogelijk "adaptive" VAR (Vector AutoRegression) technieken met Kalman filters toepassen.

Voor het bepalen van bodemvocht atmosfeer interacties blijven parametrische technieken een grote uitdaging. Immers, de dagelijkse variaties in bodemvocht laten zich kenmerken als rode ruis, en dit leidt tot grote problemen in niet-gebalanceerde varianties van de data reeksen. Dit kan op zijn beurt leiden tot model fouten en de variabele met de grotere variantie kan dan onterecht worden aangewezen als de veroorzaker van het effect. Er is meer onderzoek nodig naar VAR methoden die niet gevoelig zijn voor dit probleem. Dit lijkt een veelbelovende onderzoeksrichting.

Waarnemingen blijven noodzakelijk voor modelvalidatie. Echter, zoals we hebben laten zien in Hoofdstuk 4, kan analyse van datareeksen ook leiden tot verkeerde conclusies. In hoofdstuk 4 hebben we de oorzaken bestudeerd van de multidecadale variatie in zomertemperatuur in Zuid Europa. Hoewel, algemeen wordt aangenomen dat deze een gevolg is van de AMO, hebben we laten zien dat slechts de helft van de variatie daar echt op kan worden herleid. De andere helft van dat signaal is een versterking die optreedt als gevolg van de algehele opwarming van de aarde. Het lijkt er dus op dat de datareeks "vervuild" is en dan kan een empirische analyse tot overschatting van de relatie komen. Dit fenomeen kan vaker optreden in complexe (klimaat) situaties waarbij meerdere positieve terugkoppelingen op elkaar inwerken.

De versterking in Mediterraan Europa vindt vermoedelijke zijn oorsprong in positieve terugkoppelingen als gevolg van de opwarming van de aarde. Het is nog onvoldoende bekend of bodemvocht hierbij een rol speelt. Maar, er is ondersteuning voor deze vanuit modelstudies, die suggereren dat een dergelijke respons van uitdroging versterkt wordt door grootschalige lage drukgebieden die op hun beurt weer in stand worden gehouden door hogere landtemperaturen. Mocht dit het geval blijken te zijn dan is onze kennis over bodemvocht klimaatvariabiliteit weer een stap verder. Statistische analyse van tijdreeksen, hoe waardevol ook, kan nooit alleen het gehele beeld van variabiliteit verklaren. Hierbij zijn modelstudies, die precies de processen kunnen identificeren die bijvoorbeeld een rol kunnen spelen in de versterking van signalen zoals van de AMO, evenzeer noodzakelijk.

List of figures

List of tables

List of publications

Wang, G., Dolman, A. J., Richard de Jeu. The Observed Soil Moisture Feedback on Precipitation from Satellite Datasets over Europe. To be submitted.

Wang, **G**., Dolman, A. J. Amplified summer temperature response to the Atlantic Multidecadal Oscillation over the Mediterranean Europe. To be submitted.

Wang, **G**., Damien G., Yi Liu, Richard de Jeu, Dolman, A. J.. A three-dimensional gap filling method for large geophysical datasets: application to global satellite soil moisture observations. Submitted to *Environmental Modelling & Software*.

Wang, **G**., Dolman, A. J., Alessandri, A. A summer climate regime over Europe modulated by the North Atlantic Oscillation. *Hydrol. Earth Syst. Sci.*, **15**, 57-64, 2011.

Wang, **G**., Dolman, A. J., et al.. Fluctuation regimes of soil moisture in ERA40 re-analysis dataset. *Theor Appl Climatol*, **99**, 1–8, 2010.

Wang, **G**., Buda SU, Zbigniew W. KUNDZEWICZ, Jiang Tong. Linear and nonlinear scaling of the Yangtze River flow. *Hydrol. Processes*, **22**(10), 1532-1536, 2008.

Wang, **G**., and Tong Jiang, Richard Blender, and Klaus Fraedrich. Yangtze 1/f discharge variability and the interacting river-lake system. *Journal of Hydrology*, 351, 230-237, 2008.

Wang, G; Liao Shangang. The Spatial Heterogeneity of Human Disturbance on Land Use. *Chinese Journal of applied ecology*, **17**(04): 611-614, 2006. (In Chinese)

Wang, G; Jiang Tong. Structure and long memory of Yangtze River discharge. *ACTA Geographica Sinica,* 61(1), 47-56, 2006. (In Chinese)

Marco G., **Wang, G***., Jiang Tong. Dynamic Risk Mapping of Inundation and the Evaluation of Potential Loss in Honghu Flood Diversion Area. *J. Lake Sci.*, 18(5), 464-469, 2006. (In Chinese, *corresponding)