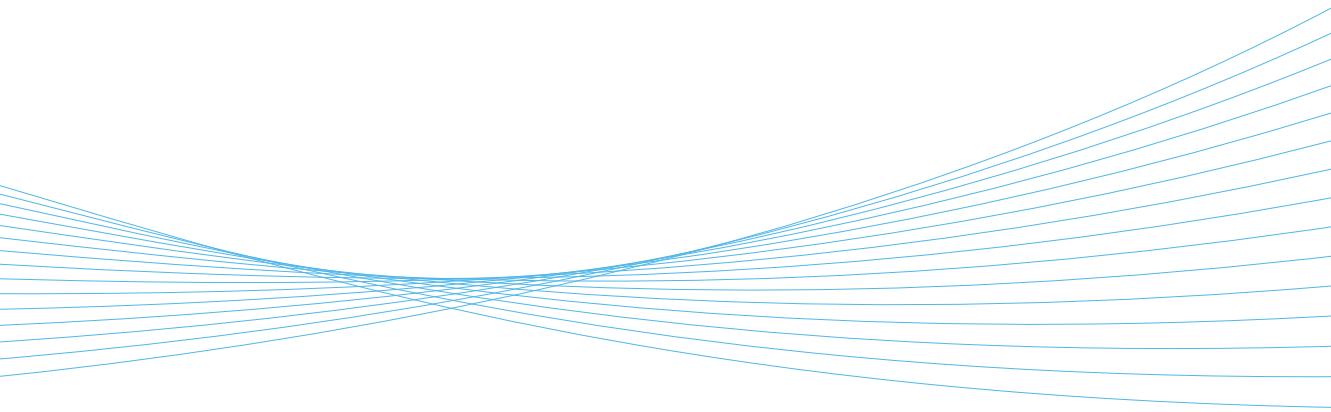
ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

126
CONTRIBUTIONS

DECOMPOSITION OF THE 20TH CENTURY CLIMATE VARIABILITY

TEIJA SEITOLA



FINNISH METEOROLOGICAL INSTITUTE
CONTRIBUTIONS

No. 126

DECOMPOSITION OF THE 20th CENTURY CLIMATE VARIABILITY

Teija Seitola

Department of Physics
Faculty of Science
University of Helsinki
Helsinki, Finland

ACADEMIC DISSERTATION in meteorology

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in auditorium Physicum D101 (Gustaf Hällströmin katu 2a) on December 9th, 2016, at 12 o'clock noon.

Finnish Meteorological Institute
Helsinki, 2016

Supervisors Professor Heikki Järvinen
Department of Physics
University of Helsinki, Finland

Dr. Johan Silén
Climate Research unit
Finnish Meteorological Institute, Finland

Reviewers Associate Professor Nedjeljka Žagar
Department of Physics
University of Ljubljana, Slovenia

Associate Professor Abdel Hannachi
Department of Meteorology
Stockholm University, Sweden

Custos Professor Heikki Järvinen
Department of Physics
University of Helsinki, Finland

Opponent Dr. Robert Vautard
Laboratoire des Sciences du Climat et
de l'Environnement, France

ISBN 978-951-697-897-3 (paperback)
ISBN 978-951-697-898-0 (pdf)
ISSN 0782-6117

Erweko
Helsinki, 2016

FINNISH METEOROLOGICAL INSTITUTE

Published by Finnish Meteorological Institute
(Erik Palménin aukio 1), P.O. Box 503
FIN-00101 Helsinki, Finland

Series title, number and report code of publication
Finnish Meteorological Institute
Contributions 126, FMI-CONT-126
Date
November 2016

Author
Teija Seitola

Title
Decomposition of the 20th century climate variability

Abstract

The ability of climate models to simulate the climate variability is of great importance when considering the reliability of, for instance, multi-annual or longer-term predictions. The aim of this thesis is to study the 20th century low-frequency variability patterns in the Earth system and how these patterns are represented by the current modelling systems. Another, equally important objective is to enable efficient spatio-temporal analysis of high-dimensional climate data sets. Decadal scale variability and predictability, from the point of view of the Nordic region, is also summarised in this study.

The work is based on the near-surface temperature of two 20th century reanalyses, obtained from the NOAA/OAR/ESRL and ECMWF, and historical climate model simulations from the coupled model intercomparison project 5 (CMIP5) data archive. In addition, a millennial Earth system model simulation is analysed. The analysis relies on a powerful dimensionality reduction method, called random projection (RP), which is introduced as a preprocessing for high-dimensional climate data sets to enhance or enable the analysis. The spectral decomposition of the data sets is based on randomised multi-channel singular spectrum analysis (RMSSA), which is one of the main achievements of this thesis.

It is shown that dimensionality reduction obtained by RP preserves the main spatial and temporal patterns with high accuracy. In addition, RMSSA is shown to provide an efficient tool for identifying different variability modes in high-dimensional climate data sets. This study shows that the 20th century variability patterns in the two reanalysis data sets are very similar. It is also shown that none of the studied climate models can closely reproduce all the variability modes identified in the reanalyses, although many aspects are simulated well.

Taking into account the rapidly accumulating amount of data and increasing dimensionality of data sets, RP is a promising method for dimensionality reduction. The results of the model evaluation can be useful in model development due better understanding of the deficiencies in representing the low-frequency modes. In addition to near-surface temperature, it would be a natural extension to include more variables in the analysis, especially because RP allows efficient data compression.

Publishing unit
Finnish Meteorological Institute

Classification (UDC)		Keywords
	55.581.1 551.583.2	climate model simulations, reanalysis, low-frequency variability, dimensionality reduction, random projection

ISSN and series title
0782-6117 Finnish Meteorological Institute Contributions

ISBN	Language	Pages
978-951-697-897-3 (paperback), 978-951-697-898-0 (pdf)	English	147

ILMATIETEEN LAITOS

Julkaisija
Ilmatieteen laitos
(Erik Palménin aukio 1), PL 503
00101 Helsinki

Julkaisun sarja, numero ja raporttikoodi
Finnish Meteorological Institute
Contributions 126, FMI-CONT-126
Julkaisuaika
Marraskuu 2016

Tekijä
Teija Seitola

Nimike
1900-luvun ilmastovaihtelun matalataajuiset komponentit

Tiivistelmä

On tärkeää selvittää, kuinka hyvin ilmastomallit simuloivat hitaita ilmastovaihteluita, jotta voidaan arvioida esimerkiksi monivuotisten ja sitä pidemmälle ulottuvien ennusteiden luotettavuutta. Tämän väitöskirjatyön tavoitteena on tutkia ilmastojärjestelmän matalataajuista vaihtelua 1900-luvulla sekä sitä, miten nykyiset ilmastomallit kuvaavat tätä vaihtelua. Toinen tärkeä tavoite on tehostaa aikaan ja paikkaan liittyvää tietoa sisältävien korkeaulotteisten aineistojen analyysimenetelmiä. Tässä työssä esitetään myös yhteenveto vuosikymmenjakson ilmastovaihteluista sekä niiden ennustettavuudesta Pohjoismaiden alueella.

Työssä on käytetty kahden 1900-luvun uusanalyysin sekä historiallisten ilmastomallisimulaatioiden tuottamaa pintailman lämpötila-aineistoa. Uusanalyysiaineistojen lähteenä ovat ECMWF sekä NOAA/OAR/ESRL. Historialliset ilmastomallisimulaatiot ovat peräisin CMIP5 (coupled model intercomparison project 5) -hankkeen tietoarkistosta. Työssä on myös hyödynnetty tuhatvuotista ilmastomallisimulaatiota. Aineistojen analysointi perustuu satunnaisprojektio-menetelmään (random projection, RP), jolla ensin pienennetään merkittävästi käsiteltävän aineiston kokoa. Aineistot hajotetaan taajuuskomponenteiksi käyttämällä satunnaistettua moniulotteista singulaarispektrianalyysiä (randomised multi-channel singular spectrum analysis, RMSSA).

Työssä osoitetaan, että satunnaisprojektiot säilyttävät aineistojen aikaan ja paikkaan liittyvien vaihteluiden rakenteet suurella tarkkuudella. Lisäksi näytetään, että RMSSA-menetelmän avulla voidaan korkeaulotteisista ilmastoaineistoista tunnistaa tehokkaasti erilaisia vaihtelurakenteita. Tutkimuksessa osoitetaan, että 1900-luvun matalataajuinen lämpötilavaihtelu on hyvin samankaltaista kummassakin uusanalyysiaineistossa. Lisäksi todetaan, että yksikään tutkituista ilmastomalleista ei tuota aivan samanlaisia vaihtelurakenteita kuin uusanalyysit, vaikka useita piirteitä mallit simuloivat hyvin.

Kun huomioidaan nopealla tahdilla kertyvät tietomäärät sekä samanaikainen aineistojen ulottuvuuksien kasvaminen, satunnaisprojektiot on lupaava menetelmä aineistojen pienentämiseen. Ilmastomallien ja uusanalyysien vertailun avulla voidaan tunnistaa mallien puutteet matalataajuisten vaihtelurakenteiden simuloinnissa ja tuloksia voidaan hyödyntää mallien kehitystyössä. Jatkossa voitaisiin tutkia pintailman lämpötilan lisäksi myös muita ilmastoa kuvaavia suureita, etenkin kun satunnaisprojektiot mahdollistavat suurten aineistomäärien tehokkaan analysoinnin.

Julkaisijayksikkö
Ilmatieteen laitos

Luokitus (UDK)		Asiasanat
	55.581.1 551.583.2	ilmastomallisimulaatiot, uusanalyysit, matalataajuinen vaihtelu dimensioreduktio, satunnaisprojektio

ISSN ja avainnimike
0782-6117 Finnish Meteorological Institute Contributions

ISBN	Kieli	Sivumäärä
978-951-697-897-3 (nid.), 978-951-697-898-0 (pdf)	englanti	147

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Heikki Järvinen and Johan Silén for all the help and guidance throughout the work. I also want to express my gratitude to Antti Mäkelä and Hilppa Gregow for giving me the opportunity to concentrate on my thesis and for being supportive and flexible group leaders. Many thanks to Jouni Räisänen for helping especially with the last article of this thesis work. I really want to thank all the colleagues and friends at Finnish Meteorological Institute (FMI), especially for the nice conversations and company during coffee breaks and lunches.

FMI has provided great facilities and resources for working and I'm thankful for that. Academy of Finland and Fortum Foundation are greatly acknowledged for their financial support.

Finally, I want to thank my parents, Taina and Tuomo, for all the help and support during my whole life. I also want to thank my sister Suvi and her family for all the nice moments we have spent together. I'm also grateful to my parents-in-law, Hannele and Jaakko, who have always offered their help when needed, especially in taking care of our children. Most of all, I want to thank Jarno, Inari and Lilja for being the most important part of my life.

Teija Seitola
Helsinki, October 2016

Contents

List of original publications

This thesis consists of an introductory part, followed by four research articles. In the introductory part, these papers are cited according to their roman numerals.

I **Seitola, T.**, Järvinen, H., 2014. Decadal climate variability and potential predictability in the Nordic region: a review. *Boreal Env. Res.*, **19**, 387–407.

II **Seitola, T.**, Mikkola, V., Silén, J., Järvinen, H., 2014. Random projections in reducing the dimensionality of climate simulation data. *Tellus A*, **66**, 25274, http://dx.doi.org/10.3402/tellusa.v66.25274

III **Seitola, T.**, Silén, J., and Järvinen, H. 2015. Randomised multichannel singular spectrum analysis of the 20th century climate data. *Tellus A*, **67**, 28876, http://dx.doi.org/10.3402/tellusa.v67.28876

IV Järvinen, H., **Seitola, T.**, Silén, J., and Räisänen, J., 2016. Multi-annual modes in the 20^{th} century temperature variability in reanalyses and CMIP5 models. *Geosci. Model Dev. Discuss.*, doi:10.5194/gmd-2016-61, accepted, in press.

1 Introduction

All life on Earth is regulated by some key climate variables, such as temperature and precipitation. Those variables are characterized by their long term mean values and the range of fluctuation around the mean. The extreme phases of the fluctuations, e.g. severe drought or flooding, have major, sometimes devastating, impacts on the ecosystems and societies. Thus, it is important to understand the behaviour of the climate system, including its components and their interaction.

It is generally easier to quantify the mean value of a climate variable, such as temperature, than the phenomena associated to its fluctuations. The objective of climate research is to understand and quantify the observed variability in the Earth system and to estimate its predictability. The ultimate goal is to exploit this understanding in reliable simulations of the future climate which is again relevant for societal decision making.

As climate is changing, it is of great importance to separate the contribution of the natural variability and anthropogenic forcing to the global temperature change. According to Flato et al. (2013) and Fyfe et al. (2016), most climate model simulations are not able to produce the slow-down in the warming trend of surface temperature in the early 2000s. A major contributor to this discrepancy to observations is thought to be the models having deficiencies in simulating the internal climate variability. On the other hand, there are several studies arguing against the existence of the slow-down in the observational surface temperature trend (e.g. Karl et al., 2015; Foster and Abraham, 2015; Lewandowsky et al., 2016). This, however, does not remove the discrepancy between the climate models and observations in the early 2000s.

The debate around the early 2000s surface temperature warming slow-down highlights the importance of quantifying the climate signals associated with natural internal variability, natural external forcing, and anthropogenic forcing. However, compared to the length of the longest time scales of climate variability, the direct observational record is quite short and sparse, especially in the beginning of the observational era. For example, the longest set of instrumental temperature observations is the Central England Temperature (Parker et al., 1992) which is available since 1659.

Considering the short direct observational record, the evaluation of the ability of climate models to simulate low-frequency (e.g. multi-decadal) climate variability is challenging. This is further complicated by the fact that the internal and external processes in driving the climate variations and the non-linear interactions between these mechanisms are not fully understood. Non-linearity means that there is no simple proportional

relation between cause and effect and the forcing may be amplified, dampened or delayed because of the complex feedback mechanisms in the Earth system.

As computing power is continuously increasing, the complexity and resolution of climate models is advancing accordingly. This is of course desirable, but poses a challenge for post-processing and analysing the high-dimensional output of complex models. Advanced spatio-temporal data-analysis is extremely useful in studying the climate signatures associated with internal variability and external forcing, but computation requires substantial amounts of memory and time in case of high-dimensional data.

The motivation of this thesis is two-fold: Firstly, the aim has been to develop efficient methods for studying high-dimensional spatio-temporal data, and secondly, to study the 20^{th} century low-frequency variability patterns in the Earth system and how these patterns are represented by the current modelling systems.

The main research problems are:

- What is the level of knowledge on the decadal climate variability and predictability in the Nordic region?
- How to handle high-dimensional data sets in advanced spatio-temporal data-analysis?
- What are the current capabilities of modelling the inter-annual to multi-decadal climate variability in the Earth system?

This thesis is organized as follows: climate variability on inter-annual to multi-decadal scales is introduced in Section 2, Section 3 explains the methods and Section 4 introduces the data sets used in this thesis. The main results of the Papers I-IV are presented in Section 5, and finally discussed in Section 6.

2 Climate variability

Climate variability refers to the fluctuations in key climate variables that are due to *internal* natural processes within the climate system, or to *external* forcing, that has either natural (such as volcanic eruptions and solar activity) or anthropogenic origin (e.g. changes in greenhouse gas emissions). Figure 1 shows a diagram of climate variability scales and processes.

Climate variability is usually described with anomalies, which are differences between momentary states of the climate system and its longer-term climatology. Climatology is the mean state computed over some time interval, such as months, years or decades (Hurrell and Deser, 2009).

Climate variability occurs at practically all conceivable time scales. Short time scale variability (monthly to inter-annual) are likely attributed to the atmospheric processes, whereas oceans have a crucial role on decadal and longer term climate variability (up to centuries or even millennia) due to their large heat capacity. In this thesis the focus is on the inter-annual to multi-decadal variability, and the related processes are reviewed in Paper I.

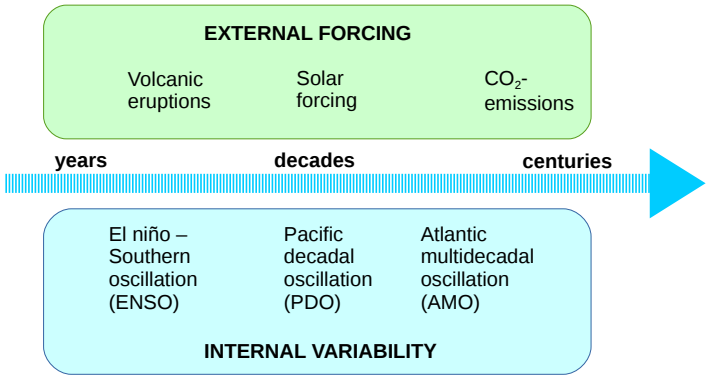


Figure 1: Climate variability time scales and processes.

2.1 Variability on inter-annual to multi-decadal scales

Internal climate variations such as the North Atlantic oscillation (NAO), the El Niño/ Southern oscillation (ENSO), the Pacific decadal oscillation (PDO) and the Atlantic multi-decadal oscillation (AMO) have major contribution on the longer-term climate variations. These phenomena are briefly presented in the following.

The internal variability on the multi-decadal scale is prominently related to the ocean dynamics. The Atlantic multi-decadal oscillation (AMO) is a major mode of variability manifested as a fluctuation of sea surface temperatures (SSTs) in the Atlantic Ocean. It is estimated to have periods of about 50–70 years (e.g. Kushnir, 1994; Delworth and Mann, 2000), but there is some controversy regarding its amplitude. AMO has support in the historical observations, such as in the longest instrumental record, Central England Temperature (Tung and Zhou, 2013). AMO has mostly been explained to be driven by the changes in ocean circulation (O'Reilly et al., 2016), especially the Atlantic Meridional Overturning Circulation (AMOC) (e.g. Delworth et al., 1993; Delworth and Mann, 2000; Ba et al., 2014). It has also been suggested that instead of being driven by the ocean circulation variability, AMO is the response to forcing from the mid-latitude atmospheric circulation (Clement et al., 2015).

Whereas the AMO is the leading mode of internal variability in the North Atlantic SSTs, the Pacific decadal oscillation (PDO) is a leading pattern of North Pacific internal SST variability. It has a roughly 20–30 yr period, and it is manifested as positive or negative SST anomalies in the Tropical Pacific and opposite anomalies in the western extra-tropical North and South Pacific. North Pacific SST variability has also a multi-decadal signal with a 50–70 yr period, which may partly be related to the AMO (Steinman et al., 2015). PDO has been associated with variations in surface temperature and precipitation in the land areas surrounding the Pacific Ocean, as well as with variability of the Pacific marine ecosystem and the Indian monsoon (Keenlyside and Ba, 2010).

On the inter-annual to decadal scale the El Niño Southern Oscillation (ENSO) is a prominent phenomenon having profound effects on the global weather and climate. It is related to coupled atmosphere-ocean variations: warming (cooling) of eastern tropical Pacific SST and high (low) surface pressure in the western tropical Pacific (Trenberth and Caron, 2000). ENSO is a quasi-periodic oscillation with a 2–7 yr period, but has a highest spectral density around 4 years. the anomalous warming of the tropical Pacific SSTs is known as El Niño and the opposite cooling phase is called La Niña. ENSO diversity (Capotondi et al., 2015) refers to the different ENSO types, with emphasis on the warm El Niño phase. For example, Kao and Yu (2009) have contrasted an eastern-Pacific (EP) type and a central-Pacific (CP) type having distinct spatial patterns and related atmospheric, surface and subsurface characteristics.

The North-Atlantic Oscillation (NAO) has variability on sub-seasonal to multi-decadal scales (Delworth and Zeng, 2016). It is measured by the difference in sea-level pressure between the subtropical (Azores) high and the subpolar (Island) low. The NAO is

primarily an atmospheric phenomenon, that produces changes in the large-scale atmospheric circulation and associated changes in temperature, precipitation and winds over the Atlantic as well as over North America and Europe (Trigo et al., 2002; Scaife et al., 2008; Hurrell and Deser, 2009). Positive phase of the NAO is associated with anomalous low pressure in the subarctic and high pressure in subtropics with stronger westerly winds and enhanced flow of warm and moist air across the North Atlantic and Europe (Hurrell, 1995).

Inter-annual to multi-decadal climate variations may partly be induced by processes that are external to the climate system. These are variations in solar activity, volcanic eruptions and anthropogenically forced changes in greenhouse gas concentrations and aerosols. The ∼11-yr sunspot cycle is quite well known but its climate effects are much debated (e.g. Rind et al., 2008). Over the past millennia, the solar forcing effect is deemed small on the Northern Hemisphere climate (Schurer et al., 2014), while e.g. Shindell et al. (2001) and Ineson et al. (2011) suggest connection between the low solar activity and negative phase of the NAO leading to colder temperatures over the Northern Hemisphere continents. Strong volcanic eruptions have climate effects that can persist for about a decade (Latif and Keenlyside, 2011). Anthropogenic changes in greenhouse gases and aerosols are an important forcing for climate on longer time scales.

The relative roles of internal and external processes in driving the climate variations are not well understood and there is a need for more precise quantification (Solomon et al., 2011). This is a challenging task, which is further complicated by the non-linear interactions between these mechanisms.

2.2 Predictability

Climate predictability refers to its *ability to be predicted* rather than to *ability to predict it* (Boer et al., 2013). In other words, predictability of a climate system is a measure of the extent to which it can be predicted in idealized conditions. Two main types of predictability studies can be found in the literature and these are *potential (or diagnostic)* and *classical (or prognostic)*.

Potential predictability is the upper limit of the forecast skill and it can be defined as the ratio of the potentially predictable variance to the total variance. Internal and externally forced climate variability are both important sources of potential predictability (Boer, 2011).

Prognostic predictability studies are conducted by performing ensemble experiments of perturbed initial conditions with a single model and the predictability is given by the

ratio of the actual signal variance to the ensemble variance.

Predictability is usually estimated in modelling studies, which can only approximate the predictability of the actual variability in the real climate system. The studies of the climate predictability must presume that the modern climate models are sufficiently similar to the actual climate system. Otherwise the predictability information provided by the model studies would be useless. In this respect it is important to understand the behavior of the current climate models and their capability to produce realistic climate variability.

3 Methods for studying climate variability and the problem of high dimensionality

Climate variability can be studied based on observations and climate model simulations. Since direct observational record is relatively short and sparse, especially over the oceans, modelling studies are often used for studying climate fluctuations. Time series (generated by observations or models) can be studied in time-domain or in frequency-domain. In the time domain the analysis is conducted with respect to time (continuous or discrete), whereas in frequency-domain with respect to frequency. Frequency domain approach includes spectral methods, which are motivated by the observation that the most regular behavior of a time series is to be periodic. Spectral analysis deals with determining the periodic components in the time series by computing periods, amplitudes and phases (Ghil et al., 2002). Spectral analysis includes a wide selection of methods, such as Fourier transform -based ones, Wavelet analysis, Principal component analysis (PCA), Singular spectrum analysis (SSA) and its multivariate version MSSA.

As climate simulation data are often high-dimensional, with thousands of time steps and grid points representing the state variables, some dimensionality reduction would be desirable before performing any complex data analysis. Averaging in time or space is of course one solution reducing the computational cost and allowing the use of well-established spectral analysis methods, such as Fourier-analysis. On the other hand, the averaging may lose some important aspects of the variability patterns. Another frequently-used method for dimension reduction is PCA, which retains most of the variability of the original data set in a small set of principal components. The drawback of PCA is that it might not be applicable with large data sets, since its computational complexity increases notably with increasing data dimension.

This section introduces the methods used in this thesis. First, PCA and MSSA are

briefly explained, and second, Random projections (RP) are introduced as a method for reducing the dimensionality and enabling analysis of high-dimensional data sets. Finally, a randomized version of the MSSA algorithm is presented.

3.1 Principal component analysis

In climate science, PCA is a widely-used method to extract the dominant spatio-temporal signals from multi-dimensional data sets and to reduce the dimensionality of the data (Von Storch and Zwiers, 2001; Hannachi et al., 2007). The idea of PCA is to find an orthogonal basis (i.e. the eigenvectors, or empirical orthogonal functions (EOFs), of the covariance matrix) to represent the original data set. By projecting the original data set onto the basis, the data set can be represented by uncorrelated linear combinations of the original variables which are called the principal components (PCs). PCA also enables dimensionality reduction, as most of the variance in the data set can be explained by only a small subset of PCs. An efficient technique for solving the eigenvectors and -values is singular value decomposition (SVD).

Let's say we have a data matrix $\mathbf{X}_{n\times d}$, where n represents the number of samples and d is the sample dimension. In case of gridded climate data, n is the number of time steps and d is the number of gridpoints. The singular value decomposition of $\mathbf{X}$ is

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T \tag{1}$$

The vectors of $\mathbf{U}$ are the eigenvectors of $\mathbf{Z} = \frac{1}{d}\mathbf{X}\mathbf{X}^T$ and $\mathbf{V}$ contains the eigenvectors of $\mathbf{C} = \frac{1}{n}\mathbf{X}^T\mathbf{X}$. The vectors of $\mathbf{V}$ are also known as EOFs. Diagonal elements of $\mathbf{D}$ are the singular values of $\mathbf{C}$ or $\mathbf{Z}$. The PCs ($\mathbf{S}$) can be calculated as follows:

$$\mathbf{S} = \mathbf{X}\mathbf{V} = \mathbf{U}\mathbf{D}\mathbf{V}^T\mathbf{V} = \mathbf{U}\mathbf{D} \tag{2}$$

Although PCA is widely used, it is not an ideal tool for extracting and illustrating spatio-temporal eigenmodes in climate data. Because of the orthogonality constraint, the PCs do not necessarily correspond to any physical phenomena or patterns (Demšar et al., 2013). In addition, the PCs may be a mixture of different physical phenomena, because of the constraint for the successive components to explain the maximum remaining variance (Aires et al., 2000). Nevertheless, PCA has been used in Paper II to demonstrate the structural similarity of an original data set and its compressed version.

In this respect, there are also other options for finding the spatio-temporal patterns. The Multi-Channel Singular Spectrum Analysis (MSSA; Broomhead and King, 1986a,b)

takes also into account the temporal autocorrelation in the original data set and provides a deeper insight into the dynamics of the system that generated the data set (Vautard and Ghil, 1989).

3.2 Multi-channel singular spectrum analysis (MSSA)

Singular spectrum analysis (SSA) and its multivariate extension Multi-Channel SSA (MSSA) were introduced into the study of dynamical systems by Broomhead and King (1986a,b). (In this connection, multivariate is a synonyme for multi-channel.) SSA and MSSA have similarities to PCA where spatial correlations are used in determining the patterns that explain most of the variability in a data set. The main difference to PCA is that MSSA finds the spatially and temporally coherent patterns that maximize the *lagged* covariance of the data set. As an analogue to PCA, MSSA eigenvectors are often called space-time EOFs (ST-EOFs), and the projections of the data set onto those ST-EOFs are called space-time principal components (ST-PCs).

In MSSA, an augmented data matrix $\mathbf{A}$ is constructed, containing M lagged copies of each column (or channel) in $\mathbf{X}_{n\times d}$. M represents the lag window. $\mathbf{A}$ has Md columns and $n' = n - M + 1$ rows, and SVD of $\mathbf{A}$ is calculated as in eq. 1 to obtain ST-EOFs and ST-PCs.

It is not trivial to choose the the lag window in MSSA. Large lag window enhances the spectral resolution, i.e. the number of different frequencies that can be identified, but at the same time the variance is distributed on a larger set of components. Because of the lag window, ST-PCs have reduced length (n') and they cannot be located into the same index space with the original time series. Instead, they can be represented in the original coordinate system by the reconstructed components, RCs (Plaut and Vautard, 1994; Ghil et al., 2002).

Similarly to PCA, the ST-PCs/ST-EOFs of MSSA do not necessarily correspond to any physical phenomena, but can be generated by some noise processes, such as first-order autoregressive (AR(1)) noise, so called 'red noise'. A significance test called Monte-Carlo MSSA (MC-MSSA) was formulated by Allen and Robertson (1996) to distinguish the 'true' oscillations from noise. In the test, the MSSA components are tested against a null-hypothesis of the data being generated by red noise, which is typical for geophysical processes.

The computational burden of MSSA becomes soon prohibitively high if the original data set is high-dimensional and lag window is chosen to be large. This is typically the situation in studies of low-frequency variability in climate data sets. Traditionally,

the dimensionality reduction has been obtained by calculating first a conventional PCA and retaining a set of dominant PCs for the following MSSA (e.g. Plaut and Vautard, 1994; Moron et al., 2012). Transformation to conventional PCs is a useful preprocessing step before MSSA, but according to Groth and Ghil (2015), its implications to signal detection are rather complex. For example, the compression of the data set into a small set of leading PCs may interfere with the detection of weak but significant signals. With high-dimensional data sets it may be the case that even PCA is not applicable. Clearly there is a need for a computationally more reasonable method for dimensionality reduction.

3.3 Random projections in dimensionality reduction

Random projection (RP) as a dimensionality reduction method is studied in Paper II. Before application to climate data, it has been successfully applied, for example, in image processing (Bingham and Mannila, 2001; Goel et al., 2005; Qi and Hughes, 2012) and for text data (Bingham and Mannila, 2001).

The core idea for random projections emerges from the Johnson-Lindenstrauss lemma (Johnson and Lindenstrauss, 1984).

Suppose we have an arbitrary matrix $\mathbf{X} \in R^{n \times d}$*. Given any* $\epsilon > 0$*, there is a mapping* $f : R^d \to R^k$ *, for any* $k \geq O\frac{\log n}{\epsilon^2}$ *, such that, for any two rows* $\mathbf{x_i}$*,* $\mathbf{x_j} \in \mathbf{X}$*, we have*

$$(1-\epsilon)||\mathbf{x_i} - \mathbf{x_j}||^2 \leq ||f(\mathbf{x_i}) - f(\mathbf{x_j})||^2 \leq (1+\epsilon)||\mathbf{x_i} - \mathbf{x_j}||^2 \tag{3}$$

In the lemma it is stated that the data points in d-dimensional space can be embedded into a k-dimensional subspace in such a way that the pairwise euclidean distances between the data points are approximately preserved with a factor of $1 \pm \epsilon$.

In the experiments of Papers II-IV, a commonly-used Gaussian mapping has been employed. Elements of $\mathbf{R}$ are $r_{ij} \sim N(0,1)$ and the row/column vectors of the random matrix are normalized to unit length. There are also other random distributions that satisfy the lemma (3). Those are presented for example in Achlioptas (2003).

In the data matrix $\mathbf{X}_{n \times d}$, n represents the number of samples and d is the sample dimension. In case of gridded climate data, n is the number of time steps and d is the number of gridpoints. The dimension reduction is performed in two steps: 1) generate a random matrix $\mathbf{R}_{d \times k}$ and 2) project $\mathbf{X}$ onto $\mathbf{R}$:

$$\mathbf{P}_{n \times k} = \mathbf{X}_{n \times d} \mathbf{R}_{d \times k}, \tag{4}$$

where $k \ll d$. In the projection, the number of samples are preserved but the dimension is reduced from d to k.

In the literature there are some estimates of a sufficient value for k (e.g. Frankl and Maehara, 1988; Dasgupta and Gupta, 2003). According to Johnson and Lindenstrauss (1984), the lower bound for k is of the order of $O(\log n/\epsilon^2)$, as stated in the lemma 3. There has also been some attempts to reveal an explicit formula, for instance, Dasgupta and Gupta (2003) showed that $k \geq 4(\epsilon^2/2 - \epsilon^3/3)^{-1} \log n$ is enough. It is notable that in these estimates the subdimension k does not depend on d, but on the number of samples n and the error rate ϵ.

It should be noted that these theoretical lower bounds for k are conservative estimates and usually much lower values for k still give good results, retaining most of the information of the original data set (e.g. Bingham and Mannila, 2001). This was also observed in Paper II. In practice, the value for k is usually chosen in an adaptive manner, according to the desired size for lower-dimensional approximation and by monitoring the associated error rate.

Figure 2 shows the error (in %) produced by RP as a function of retained dimensions (in % of the original dimensions). The original data set is the monthly mean near-surface temperature from the 20th century reanalysis (Compo et al., 2011). In this data set $n = 1704$ and $d = 18048$. The error is measured by the difference in euclidean distance between 100 pairs of data vectors in the original and dimensionality reduced space. Figure also shows the 95% confidence interval for the error, calculated over 500 realisations of RP with different random numbers. It can be seen that even with very low dimensions the error produced by RP is quite low, although the error confidence interval increases with decreasing number of retained dimensions.

RP is powerful, since it can be used in constructing a much lower-dimensional (ϵ-approximate) version of any algorithm depending only on the geometry of the data (i.e. the distances between the data points). RP is also easy to implement and can reduce complexity of algorithms with small costs. It is linear and indifferent to the data used, subdimension k does not depend on the dimensionality d of the original data, and it preserves the distances. RP can also be used in constructing efficient parallel implementations of existing algorithms.

RP has been applied in several fields or computational methods. It allows randomized matrix factorisations, such as randomised SVD (Halko et al., 2011). Some other applications include for example nearest-neighbour (e.g. Deegalla and Boström, 2006) and clustering (e.g. Fern and Brodley, 2003) algorithms.

Of course, RP is not a lossless method and some accuracy may be lost especially in

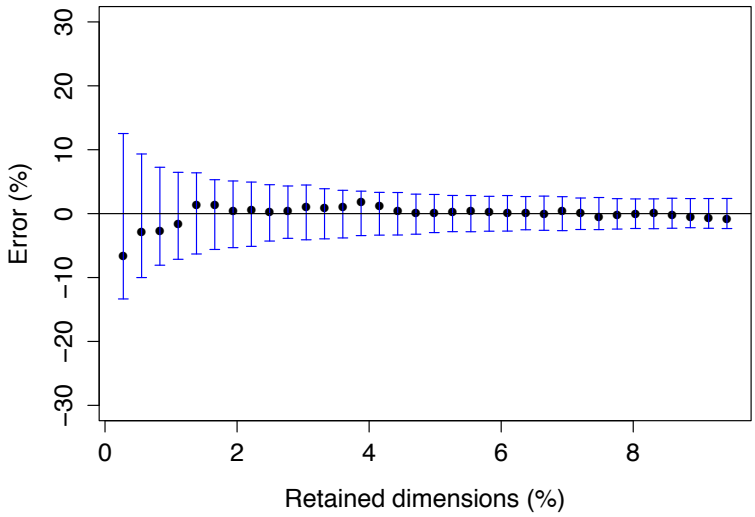


Figure 2: Error (in %) produced by RP and 95% confidence intervals, as a function of retained dimensions (%).

very low dimensions (Fig. 2), but on the other hand, computational and data storage expense is reduced. It should also be emphasized, that RP alone does not provide physical interpretation of the climate data, but is aimed to be used in conjunction with other methods, such as PCA or MSSA.

3.4 Randomised multi-channel singular spectrum analysis

The main achievement of Paper III was to introduce a randomised version of the MSSA algorithm, called RMSSA. This algorithm was motivated by the increasing computational complexity of MSSA with the increasing data dimension. The RMSSA-algorithm 1) reduces the dimension of the original data set by RP, 2) decomposes the data set by calculating standard MSSA steps in a reduced space, and 3) reconstructs the components in the original high-dimensional space.

- Step 1) is straightforward and is implemented as described in the previous section. The lower-dimensional matrix $\mathbf{P}_{n\times k}$ is obtained by eq. 4.

- Step 2) follows the conventional MSSA procedure as described in section 3.2, but in much lower dimensional space (k) compared to the original data dimensions (d). The augmented matrix $\mathbf{A}_{RP}$ is constructed from $\mathbf{P}$ and SVD is calculated:

$$\mathbf{A}_{RP} = \mathbf{U}_{RP}\mathbf{D}_{RP}\mathbf{V}_{RP}^{T} \tag{5}$$

- Step 3) requires calculating the eigenvectors (ST-EOFs) in the original d-dimensional space in order to represent the ST-PCs in the original coordinate system:

$$\mathbf{V}_{A} \approx \mathbf{A}^{T}\mathbf{U}_{RP}(\mathbf{D}_{RP})^{-1} \tag{6}$$

  The calculation can be limited only to the eigenmodes that are of interest.

Significance test of MSSA components requires solving conventional PCs of the original data set. The RMSSA implementation presented in Paper III also contains another version of the algorithm where the PCs are solved in the dimension-reduced space. This makes the computations feasible and affordable even in very high-dimensional problems.

In summary, RMSSA-algorithm is powerful when the dimensions of the data sets become prohibitively large. It allows a computationally efficient way of decomposing a data set into its spatio-temporal patterns.

4 Data sets

The monthly mean near-surface air temperature fields from the reanalysis data sets and climate model simulations were analysed in this thesis. Surface temperature was chosen, because it is routinely examined variable in atmospheric models and many processes must be adequately represented in models to realistically capture the observed temperature distribution (Flato et al., 2013). In the following subsections the reanalysis and model simulation data sets are described in more detail.

4.1 The 20th century reanalyses

Two 20th century reanalysis data sets were analysed in Papers III and IV: the 20th Century Reanalysis V2 data (hereafter 20CR) provided by the NOAA/OAR/ ESRL PSD (Compo et al., 2011), and ERA-20C data provided by ECMWF (Poli et al., 2013). These analyses provide a means to study the 20th century climate variability.

In 20CR the surface and sea level pressure observations are combined with a short term forecast to produce an ensemble of perturbed reanalyses, and the final data set

corresponds to the ensemble mean. The observed monthly sea-surface temperature and sea-ice distributions from HadISST1.1 (Rayner et al., 2003) are used as boundary conditions, and the reanalysis is forced by historical record of changes in climate forcing factors (greenhouse gases (CO2), volcanic aerosols and solar variations). Analysis is performed with an Ensemble Kalman Filter to produce an estimate of the complete state of the atmosphere and its uncertainty (Compo et al., 2011). 20CR has ∼2.0 degree horizontal resolution (approximately 210 km) and the gaussian gridded (192×94) data from 3-hour forecast values is used. The vertical resolution is 28 levels. The data set spans from 1871 to 2012.

ERA-20C is the first 20th century atmospheric reanalysis of ECMWF. In ERA-20C, observations of surface pressure and surface winds over the oceans are assimilated (Poli et al., 2013). ERA-20C is forced by historical time-varying changes in sea-surface temperature and sea-ice fraction, as well as climate forcing factors. Compared to 20CR, a more recent sea-surface temperature and sea ice cover from HadISST2 (Rayner et al., 2006) are used. ERA-20C uses a 24-hour four-dimensional variational (4D-Var) data assimilation scheme. The horizontal resolution of ERA-20C is approximately 125 km (T159) in a grid of 360×181 points and the vertical resolution is 91 levels. The data set covers the time sequence from 1900 to 2010. Thus, ERA-20C is shorter, but has finer resolution compared to 20CR.

Both reanalyses are affected by changes in the observing system and coverage of observations. They also omit the upper-air and satellite observations which means that they are not the best estimates beginning from those years when these observations have become available (Poli and NCAR Staff (Eds.), 2016). On the other hand, the 20th century data sets provide a means to study long time scale climate processes.

4.2 The climate model simulations

A monthly surface temperature data set from a millennial full-forcing Earth system model simulation (Jungclaus, 2008) was used in the experiments of Paper II. Purpose of the Millennium Earth System Model (M-ESM) has been to simulate the full Earth system over periods of hundreds to thousands of years. The simulations are forced by volcanoes, variations of solar irradiance, and land use changes. The ESM has four main components: the atmosphere model ECHAM, the land model JSBACH, the Max-Planck-Ocean-Model MPI-OM and the ocean-biogeochemistry-model HAMOCC (Budich et al., 2010). The data set used in Paper II has a resolution of 96 points in longitude and 48 points in latitude. Purpose of using this data set was to demonstrate the structure

preservation properties of RP.

The historical (1901–2005) simulations from the coupled model intercomparison project 5 (CMIP5) data archive, following the CMIP5 experimental protocol (Taylor et al., 2012), were analysed in Papers III and IV. In the 20th Century simulations the historical record of climate forcing factors are used. The simulations are produced by Atmosphere-ocean general circulation models (AOGCMs) or ESMs.

AOGCMs include atmosphere, ocean, land and sea ice components. They are primarily used for studying the dynamics of the climate system, and for making projections based on future greenhouse gas and aerosol forcing (Flato et al., 2013). AOGCMs are still extensively used in applications where the biogeochemical feedbacks are not critical (e.g. seasonal and decadal predictions). ESMs include also the biogeochemical cycles, which play an important role in simulating the response of the climate system to external forcing (Flato et al., 2013).

The CMIP5 simulations that were analysed in Papers III and IV have originally different resolutions, but all the model data sets were interpolated into a common grid of 144×73 points. A single ensemble member of each model was used in the analysis. In selecting the models, a major principle was to use only one model per institution. Furthermore, all the chosen models have undergone a long (generally several generations of) history of development, suggesting that the selected models collectively represent the state-of-the-art. The models that were used, are listed in Table 1.

Table 1: CMIP5 climate models used in this thesis.

Model name	**Modeling center**	**Country**
CanESM2	Canadian Centre for Climate Modelling and Analysis (CCCMA)	Canada
CESM1(CAM5)	Community Earth System Model Contributors (NSF-DOE-NCAR)	USA
CNRM-CM5-2	Centre National de Recherches Mtorologiques / Centre Europen de Recherche et Formation Avance en Calcul Scientifique (CNRM-CERFACS)	France
CSIRO-Mk3.6.0	Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence (CSIRO-QCCCE)	Australia
GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory (NOAA GFDL)	USA
GISS-E2-R	NASA Goddard Institute for Space Studies (NASA GISS)	USA
HadGEM2-ES	Met Office Hadley Centre (MOHC)	UK
INM-CM4	Institute for Numerical Mathematics (INM)	Russia
IPSL-CM5B-LR	Institut Pierre-Simon Laplace (IPSL)	France
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies (JAMSTEC/AORI/NIES)	Japan
MPI-ESM-MR	Max Planck Institute for Meteorology (MPI-M)	Germany
MRI-CGCM3	Meteorological Research Institute (MRI/JMA)	Japan

5 Main results

This section summarises the main results of the Papers I–IV.

5.1 Decadal variability and predictability in the Nordic region

Paper I reviewed the decadal (to multi-decadal) climate variability and predictability with emphasis on the Nordic region. The published studies indicate that the relative roles of internal and external mechanisms driving the long-term climate variability are not well understood. Decadal variability and predictability is found predominately over mid- to high-latitude oceans, especially in the North-Atlantic (NA) sector. The most prominent internal variability mechanism contributing to the decadal variability in the North-Atlantic sector is the NAO and the AMOC. Furthermore, the NA predictability is mainly due to the AMOC-variability, but over land areas predictability is deemed to be low.

Based on the review, the potential predictability of decadal scale variations in the Nordic region is highly uncertain. Some results indicate that the closeness to the North-Atlantic might imply some predictability in the coastal areas. On the other hand, some studies indicate that the potential decadal predictability may be generally reduced because of global warming.

The published papers on the decadal variability and predictability indicate that the climate variability patterns and their mutual interaction calls for more study. This subject is addressed in the following Papers II-IV by refining methods for studying the variability patterns, and finally comparing the low-frequency variability in reanalyses and contemporary climate models.

5.2 Random projections and climate data

Paper II introduced RP as a dimensionality reduction method applied to climate data sets. The structure-preservation properties of RP were demonstrated by applying PCA on the original and dimensionality reduced data sets. Experiments with lower-dimensional subspaces of 10% and 1% of the original data dimensions showed that even at 1 % of the original dimensions the main spatial and temporal patterns of the original surface temperature data set were approximately preserved. Figure 3 compares the eigenvectors 1–8 of the original and dimensionality reduced data sets and additionally, Figure 4 shows the correlation of the eigenvectors 1–15.

With a subspace of 10% of the original dimensions the PCs explaining 96 % of the

variance in the original data set were recovered, and with 1% the recovery was still successful until the PCs explaining 94 % of the original variance. Large part of the variance can be attributed to the annual cycle that was not removed from the data set in the experiments of Paper II.

The stability of the obtained results was also investigated by projecting the original data matrix onto a set of different realisations of random matrices. The PCA of each, slightly different projection was calculated, which allowed approximating confidence limits for the eigenvalues, i.e. the amount of variance explained by each PC. The results showed that some differences in the results can occur due to different random matrices, especially when the subspace is very small compared to the original size of the data set. The orthogonalisation of the random matrix may enhance the stability of the results, but this was not covered in Paper II.

Paper II further demonstrated the application of RP + PCA on a higher-dimensional atmospheric temperature data set including the vertical component. This allowed investigating the temperature patterns in three dimensions. A signal reminiscent of the ENSO was identified in the analysis and the spatial patterns related to this signal were studied in more detail. The three-dimensional analysis revealed, for instance, that the spatial pattern of the ENSO-related temperature signal is in an opposite phase in the upper atmosphere compared to the lower levels. However, one must be aware of the limitations of PCA in providing a physical interpretation of the results. Also, the characteristics of the data set must be considered, i.e. the ENSO representation of the Millennium simulations (Jungclaus et al., 2006). The main idea of the experiment was to show that RP can be applied as a preprocessing of high-dimensional data sets, reducing computational burden of further analysis, or even enabling it.

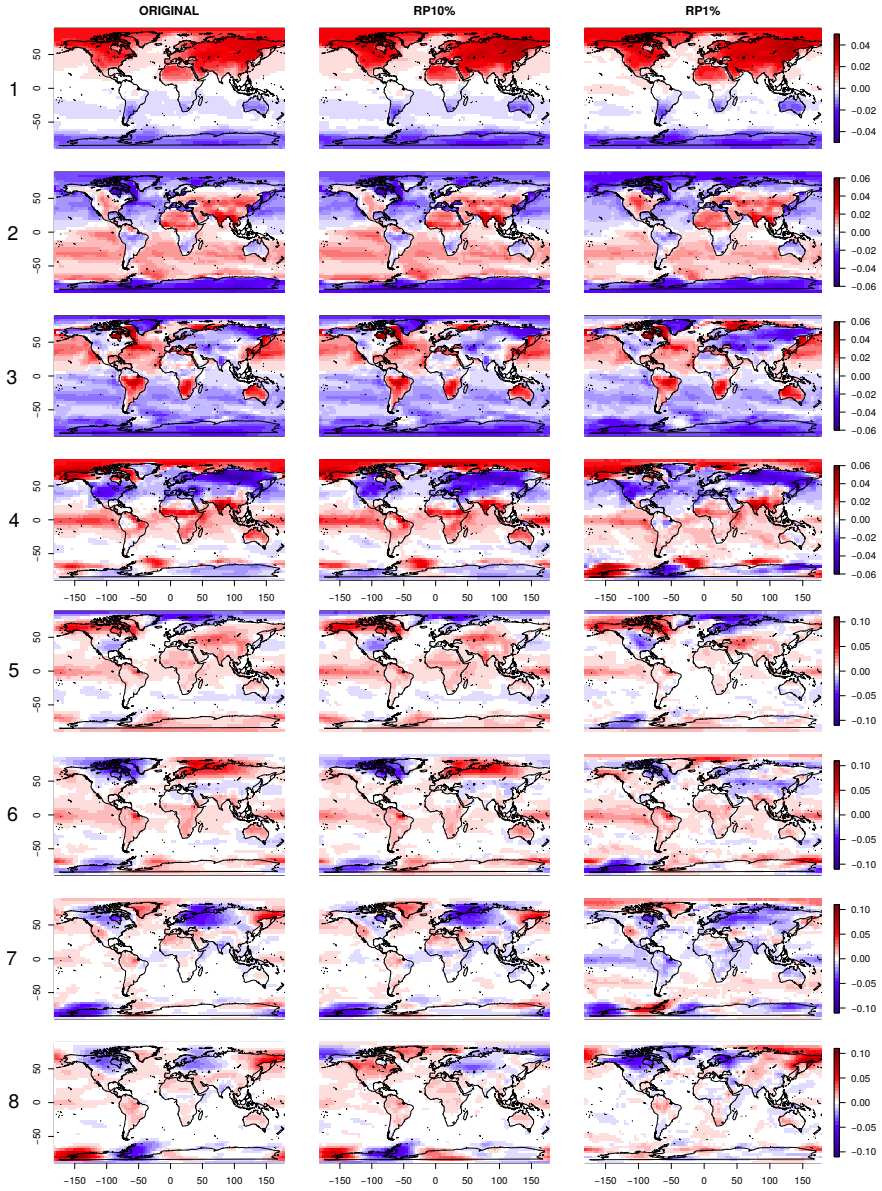


Figure 3: Comparison of eigenvectors 1–8 computed from the original and dimensionality reduced data sets (RP10% and RP1%). The unit of the colour scale is arbitrary. (Reproduced from Fig. 4 of Paper II)

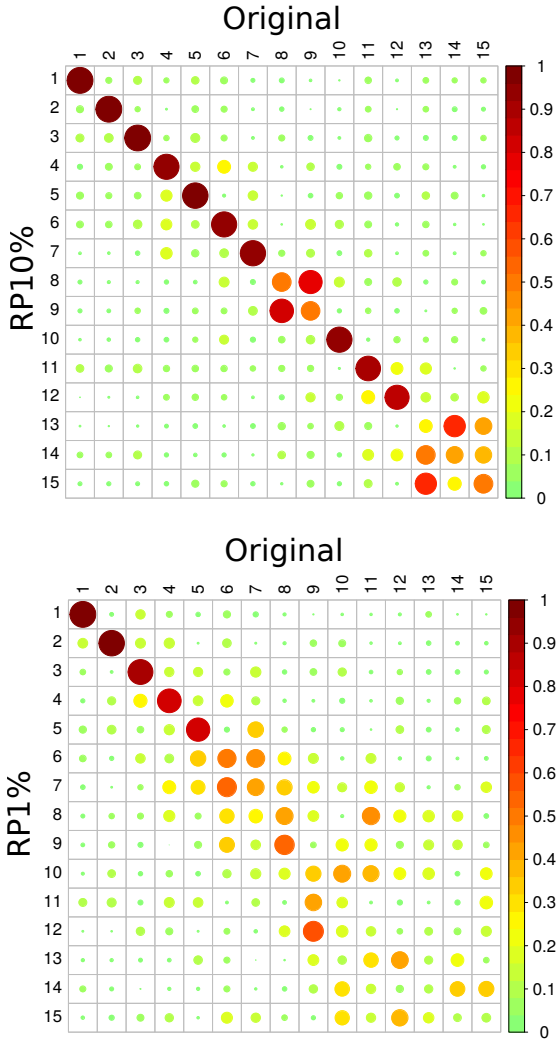


Figure 4: Correlation of eigenvectors 1–15 computed from the original and dimensionality reduced data sets. 'Original' refers to the eigenvectors of the original data set, 'RP10%' and 'RP1%' to the eigenvectors of the dimensionality reduced data sets. (Reproduced from Fig. 6 of Paper II)

5.3 The 20th century near-surface temperature variability in reanalyses and climate model data sets

In Papers III and IV the variability patterns of the reanalysis and climate model data sets extracted from the CMIP5 data archive were analysed and compared. The main goal of Paper III was to introduce the RMSSA -algorithm, and the use of the algorithm was demonstrated on the 20CR data set as well as on two historical climate model simulations (HadGEM-ES and MPI-ESM-MR). Paper IV extended the analysis by another reanalysis data set, ERA-20C, and 12 climate model simulations from the CMIP5 data archive. The data sets were standardised to avoid overweighting the high-latitude variance. Furthermore, the data sets were detrended and the dominating annual cycle was removed.

In Paper IV The spectral characteristics of the data sets were further studied by conducting a more detailed analysis on the oscillatory components (ST-PCs) extracted from the data sets. The spectral densities of the ST-PCs were estimated and summed up to obtain so called total spectrum for each data set. The total spectrum of each data set summarises the spectra of its components and facilitates the comparison of the climate model and reanalysis data sets. In addition, the statistical significance of the identified oscillatory modes was studied by MC-MSSA. Because one century covered by the reanalysis data sets is very short for analysing the decadal to multi-decadal variability, Paper IV concentrates on the multi-annual variability modes.

5.3.1 Comparison of variability modes in the two reanalyses

RMSSA of the 20CR and ERA-20C revealed that the decomposition of the two reanalyses data sets is very similar: the variance is distributed in a similar way to the components representing the different oscillatory modes.

In both data sets so-called trend components with multi-decadal scale period explain largest fraction of the variance compared to the following components. The multi-decadal components have relatively somewhat more explanatory power in 20CR compared to ERA-20C. One has to bear in mind, though, that the length of the time series (105 years) restricts the analysis of multi-decadal oscillations.

The multi-annual modes, explaining together the second largest fraction of the variance in both data sets, have periods of about 3–4 years and around 5 years. These modes may be related to the El Niño-Southern Oscillation (ENSO) which is a prominent phenomenon on those time scales. After the ENSO-type components, some differences between 20CR and ERA-20C start to occur, but the overview of the spectra in both data

sets is similar. This can also be seen in Figure 5 a, showing the total spectra of both data sets. The only clear difference is that the spectral power in ERA-20C is systematically slightly higher than in 20CR. This is most likely due to generally higher temperature variance in ERA-20C compared to 20CR. In addition, the 3–4 yr and 5 yr spectral peaks are relatively more pronounced in 20CR than in ERA-20C.

The statistical significance testing of the components shows that approximately the same multi-annual periods (in the range of 3.5–5.7 years) are significant in both reanalysis data sets (Figure 5 b–c). The annual cycle is removed from the data sets and therefore the eigenvalues corresponding to that mode are very low.

The representation of climate variability in 20CR data set has been studied in Compo et al. (2011). 20CR represents the longer time-scale variability fairly well, as measured by a few climate indices (the NAO, the Pacific Walker Circulation index (PWC) and the Pacific-North American Pattern (PNA)). The variability has been compared to other reanalysis data sets (ERA-Interim, NCEP-NCAR Reanalysis, ERA-40) and the correlation of indices between various estimates is very high.

In Poli et al. (2016), the climate fidelity of ERA-20C was also studied by investigating a selection of common climate indices calculated from the monthly mean data (the Niño 3.4 index, the Southern Oscillation index (SOI), the NAO index and the PNA index). These four monthly climate indices show excellent agreement for ERA-20C with other reanalysis products (20CRv2c, JRA-55, and ERA-Interim) especially after 1980, but there are more discrepancies at earlier times and regions where observation coverage is low.

Otherwise the literature evaluating the long-term climate variability in 20CR and ERA-20C data sets is scarce. The work in Paper IV adds to this literature by comparing the oscillatory modes identified in the two 20th century reanalysis products.

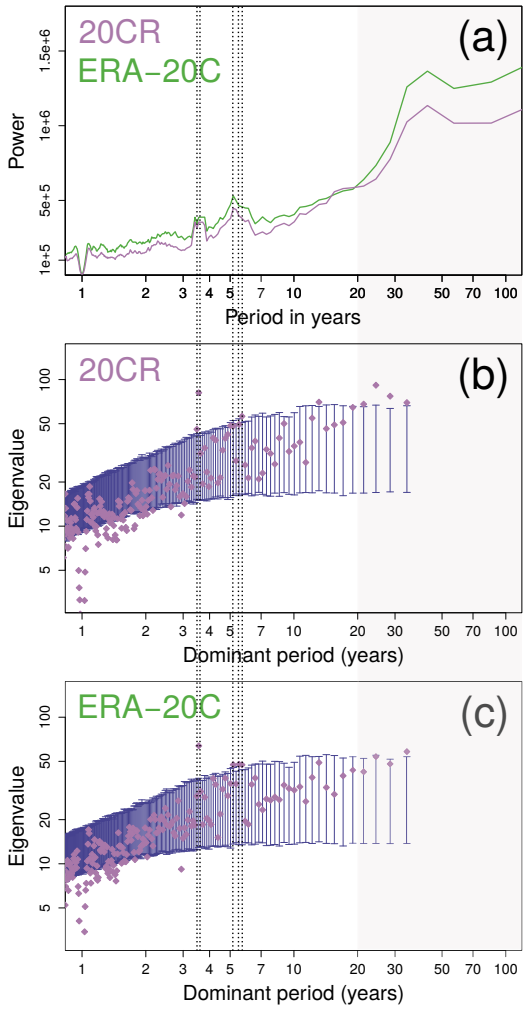


Figure 5: (a) Total spectrum of 20CR and ERA-20C near-surface temperature. (b) Significance test of the near-surface temperature variability in 20CR. Colored squares show the data eigenvalues plotted against the dominant frequency of the ST-PC corresponding to each eigenvalue. The vertical bars show the 95% confidence intervals for the eigenvalue distribution. The ST-PCs that correspond to eigenvalues rising above the 97.5^{th} percentiles are considered significant at the 5 % level. (c) Same as (b), but for ERA-20C. The grey shaded area denotes the frequencies beyond the lag window (20 yr). (Reproduced from Fig. 4 of Paper IV)

5.3.2 Comparison of variability modes in the reanalyses and CMIP5 models

The ability of a model to simulate the climate variability, in addition to the mean state and externally forced change, is crucial in determining the model performance. To study this aspect, 12 climate model simulations for period 1901–2005 were analysed in exactly the same way as the reanalysis data sets. Figure 6 shows the total spectra for the climate models. The statistically significant (at the 5% level) multi-annual modes with period less than 7 yrs are denoted by dashed vertical lines. The total spectra of the reanalyses are plotted in the background as a reference. Although the models cannot be simply ranked based on how different or similar the model spectra are from the reference, a comparison of the simulated and the reanalysis spectra provides useful hints of the strengths and weaknesses of the models.

The analysis of Paper IV shows that there are significant multi-annual (2–7 yr) variability patterns in most of the climate model data sets. However, the level of variability varies a lot among the models. For example, there is a group of models (a, b, d and e in Fig. 6) that are overactive on multi-annual scales. In most of the other models, the multi-annual variability is relatively less prominent than in the reanalyses.

Paper IV concentrates on the multi-annual scale, but it is also noted that the level of decadal scale variability (10–20 yr) is quite close to the reanalyses in majority of the models. However, some models, such as HadGEM2-ES (Fig. 6g), overestimate it. Also, some of the climate models seem to underestimate the level of multi-decadal variability ($>$ 20 yr) but the shortness of the time series (105 yrs) constrains the analysis.

Results of Paper IV indicate that the number of statistically significant periods (at 5% level) is larger in several models, in comparison to the reanalyses. This is explained, at least partly, by the fact that the modes have irregular periods captured by a range of adjacent frequencies. In addition, some models have several significant and distinct periods between 2 and 7 yrs, which are not detected in the reanalyses. Models a, i, j and k in Fig. 6 seem to be somewhat closer to the reanalysis in terms of number of significant periods.

In Paper IV the spatial pattern related to a 3–4 yr oscillatory mode was also analysed. This mode was identified as significant in 20CR and ERA-20C, and most of the climate model data sets. Phase composites, following the procedure of Plaut and Vautard (1994), were constructed from the 3–4 yr mode. These are illustrated in detail in the Supplement (S3) of Paper IV.

The 3–4 yr mode has a typical ENSO related temperature anomaly pattern in both reanalysis data sets, and climate model anomaly patterns are similar to the reanalyses in

many areas. However, some differences also exist (see the Supplement S3 of Paper IV). Especially some overestimation of the anomalies related to the 3–4 yr pattern is seen in several models. Furthermore, the equatorial Pacific anomalies tend to extend too west in about half of the models. The anomaly pattern in the northwestern North-America is present in all the models to some extent, but in most of them it is either somewhat misplaced or extends to the adjacent sea areas and the Eurasian continent.

Representation of inter-annual to multi-decadal climate variability in CMIP5 models has been analysed in a wide range of studies (e.g. Bellenger et al., 2014; Knutson et al., 2013; Ba et al., 2014 and Fredriksen and Rypdal, 2016). In Flato et al. (2013) it is stated that the ENSO representation in CMIP5 models has improved since CMIP3 and most CMIP5 models have variability maximum at the observed time scale (2–7 years). However, models still have biases in ENSO amplitude, period and spatial pattern, identified both in CMIP3 and CMIP5 simulations (e.g. Guilyardi et al., 2009; Bellenger et al., 2014).

Longer-term variability is also biased in some of the climate models. Many studies (e.g. Kumar et al., 2013; Ba et al., 2014) indicate that the Atlantic multi-decadal variability is weaker than observed in CMIP5 models. In Knutson et al. (2013) it is shown that on average, the CMIP5 models tend to overestimate the low-frequency surface temperature variability ($>$ 10 years) in high-latitude regions of the Northern Hemisphere, but underestimate it over much of the remaining lower latitude regions. Although the results of Paper IV also indicated that some of the models are underestimating the power at lower frequencies, the relatively short temporal coverage of the data sets restricts the analysis of decadal to multi-decadal variability. In addition, the choice of the lag-window (20 years) in RMSSA has also effect on the identified frequencies. The frequencies within the lag window are emphasized while the frequencies longer than 20 yr have relatively less power because those are not covered by the lag window.

5.3.3 Erratum to Paper III

In Paper III, Fig. 8, the latitudinal climate model temperature anomalies (HadGEM2 and MPI-ESM) have been plotted mistakenly in reverse order.

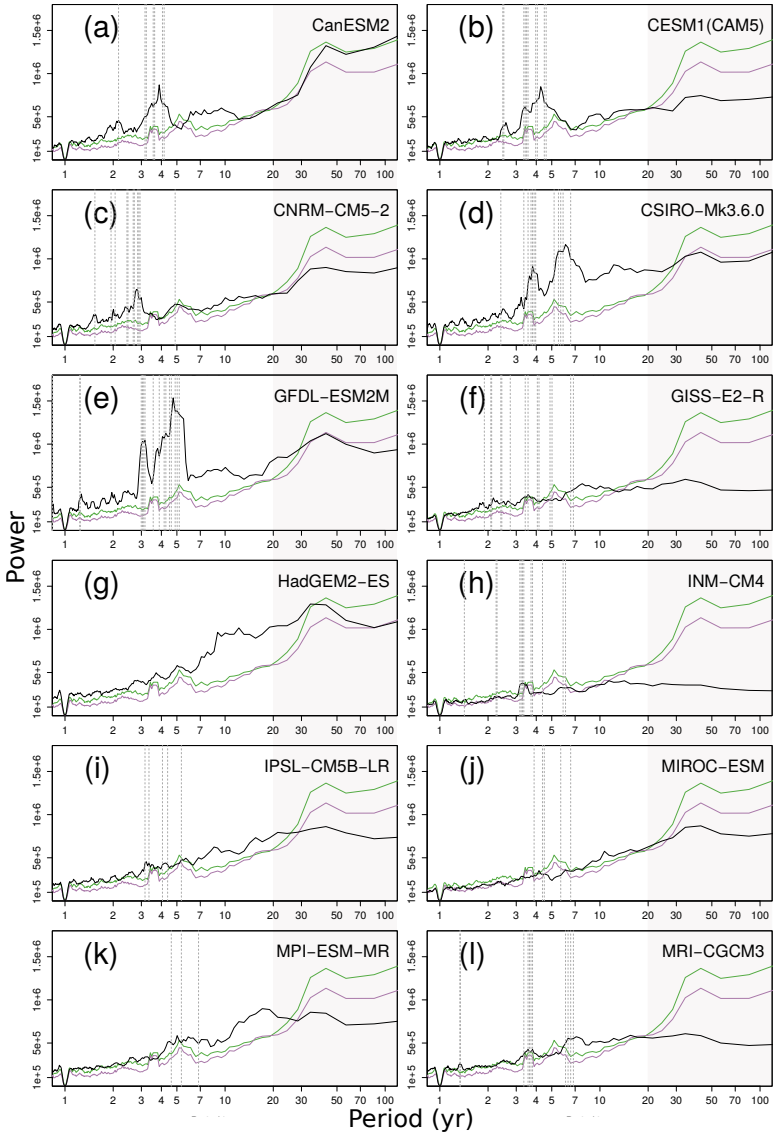


Figure 6: Total spectrum of the near-surface temperature in climate model data sets. The dashed vertical lines indicate the climate model multi-annual periods significant at 5% level. The total spectra of the reanalysis data sets are plotted with green and red lines. The grey shaded area denotes the frequencies beyond the lag window (20 yr). (Reproduced from Fig. 5 of Paper IV)

6 Summary and discussion

Understanding the climate model performance is of great importance when considering the reliability of, for instance, decadal or longer-term predictions or assessing the decadal predictability. The evaluation of the low-frequency variability in models is challenging for many reasons. First, the direct observational record, especially in the oceans, is relatively short and sparse for comparing low-frequency variability in models and observations. Secondly, internal and external processes in driving the climate variability and the non-linear interactions between these mechanisms are not fully understood. In addition, the increasing complexity and resolution of models hinders the analysis of the model output.

The main contribution of this thesis has been in applying efficient dimensionality reduction to the climate data sets, and refining decomposition methods (PCA, MSSA) to enable analysis of high-dimensional spatio-temporal data sets. In addition, the focus has been on extracting and comparing the low-frequency variability patterns of the 20th century near-surface temperature in reanalyses and current modelling systems. The present knowledge on the decadal variability and predictability, with a focus on the Nordic region, has also been summarised in this study. In the following, the research questions are self-assessed based on the obtained research results and available literature.

- *What is the level of knowledge on the decadal climate variability and predictability in the Nordic region?*

Paper I reviewed the existing knowledge on decadal climate variability and predictability, with emphasis on the Nordic region. It was found out that the internal variability mechanisms in the North Atlantic associated with decadal variability (such as AMOC and NAO) also affect the climate variability in the Nordic region. On the other hand, the decadal predictability in this area is low, although closeness to the North-Atlantic sector may also contribute to predictability in the Nordic region. Although the focus of Paper IV was on the multi-annual variability modes of the near-surface temperature, it was also noted that the level of the decadal scale variability is quite close to the reanalyses or overestimated in the studied CMIP5 models.

The findings in Paper I indicated that there is a need for better understanding of the variability patterns in the climate system and how those patterns are captured by the climate models. Paper I served as a background for the following studies, guiding to focus on the global scale. It also acted as a motivation to extract the inter-annual to multi-decadal modes of climate data sets and to study their spatial and temporal signatures.

- *How to handle high-dimensional data sets in advanced spatio-temporal data-analysis?*

This question was addressed in Papers II and III. Paper II introduced random projections as a powerfull, distance preserving solution for dealing with high-dimensional problems. The experiments showed that at notably lower dimensions the main spatial and temporal patterns of the original data set were preserved with high accuracy. It was also shown that random projections are very easy to implement, involving only random number generation and matrix multiplication. One question that was still left a bit open, was the lower bound for the dimensionality reduction. As noted, the theoretical lower bounds were much higher than the ones that were actually implemented in the experiments of Paper II. This question would deserve more attention in the future studies.

Taking into account the rapidly accumulating amount of data and increasing dimensionality of data sets, the results of Paper II are encouraging. Apart from enabling heavy data-analysis, random projections could also have other applications, for instance, it might be useful in reducing the data storage costs. Modelling results could be stored in a low-dimensional form and then recovered back to the original dimension when needed. However, this requires more study and is not covered by this thesis.

Paper III further developed the idea of Paper II and combined random projections with an effective spectral analysis tool, MSSA. The main achievement of Paper III was a randomised version of the MSSA algorithm, called RMSSA, which was shown to be efficient in finding spatially and temporally coherent patterns in high-dimensional problems.

- *What are the current capabilities of modelling the inter-annual to multi-decadal climate variability in the Earth system?*

Paper IV further showed the usefulness of the RMSSA-algorithm presented in Paper III, and compared the 20th century near-surface temperature variability patterns in the reanalysis and climate model data sets. Although the decadal to multi-decadal variability was also of interest, the Paper IV concentrated on the multi-annual modes of variability. This was constrained by the temporal coverage of the analysed data sets. In paper IV it was shown that the total spectra of the two reanalysis data sets (20CR and ERA-20C) are very similar on almost all time scales, the only difference being that the spectral power of ERA-20C is systematically slightly higher than in 20CR. It was also shown that the 3.5 and 5 yr oscillations were the prominent multi-annual variability modes in the reanalysis data sets. The literature comparing the low-frequency variability in the

two reanalysis data sets was found to be scarce and these results add to this literature. It was also shown that there are significant multi-annual (2–7 yr) variability patterns in the majority of the studied climate models. None of the studied models closely reproduce all aspects of the reanalysis spectra, although many aspects are represented well.

The results of Paper IV are aimed at providing guidance for model development by pointing towards the deficiencies in simulating the multi-annual temperature variability. RMSSA is efficient in identifying the relative power of different oscillatory modes in each model and analysing the corresponding spatial signatures. Different versions of existing models could be studied to identify the impact of, for example, stochastic parameterisations on the variability patterns. Total spectra of the data sets were calculated to enable comparisons between the reanalysed and simulated modes of variability. However, good agreement with the reference spectra might occasionally result from compensating errors in model processes. The weakness of this study is also the limited temporal coverage of the data sets, which restricted the analysis of the decadal to multi-decadal variability patterns. In this connection, one must note that there is a wide archive of proxy records, covering the last millennium and even longer. These records provide a means to asses the ability of state-of-the-art climate models to simulate the variability upto centennial time scales (Otto-Bliesner et al., 2016).

The future studies on this subject would also cover the inclusion of several variables in the analysis, not just near-surface temperature. The multivariate analysis could reveal the common oscillatory patterns among the different variables and give a deeper insight into the underlying dynamics. This thesis has contributed in answering the *what* and *how* -questions, but the question *why* is still unanswered and requires more study.

Summaries of the original publications

I **Seitola, T.**, Järvinen, H., 2014. Decadal climate variability and potential predictability in the Nordic region: a review. *Boreal Env. Res.*, **19**, 387–407.

Paper I reviews decadal climate variability and predictability with emphasis on the Nordic region. In the published studies, the decadal variability and predictability is found predominately over mid- to high-latitude oceans, especially in the North Atlantic (NA) sector. The most prominent internal mechanisms explaining the variability in the NA are the North Atlantic oscillation (NAO) and the Atlantic meridional overturning circulation (AMOC). The conclusions regarding the variability in the Nordic region are uncertain at the moment, despite the fact that new knowledge is rapidly accumulating. In general, the published studies indicate that the relative roles of internal and external mechanisms driving the long-term variability and their mutual interactions are not sufficiently understood.

I was responsible for all the analysis of the published studies and major part of the writing.

II **Seitola, T.**, Mikkola, V., Silén, J., Järvinen, H., 2014. Random projections in reducing the dimensionality of climate simulation data. *Tellus A*, **66**, 25274, http://dx.doi.org/10.3402/tellusa.v66.25274.

Paper II introduces Random projection (RP) as a dimensionality reduction method for climate data. In the experiments, RP is applied to simulated global surface temperature data set, and principal component analysis (PCA) is utilized to analyse how the structures are preserved in the lower dimensional data space of 10% or 1% of the original volume. The experiments show that even at 1% of original dimensions, the main spatial patterns and temporal signatures can be recovered.

I was responsible for all the computations, and for major part of the analysis and writing.

III **Seitola, T.**, Silén, J., and Järvinen, H. 2015. Randomised multichannel singular spectrum analysis of the 20th century climate data. Tellus A, 67, 28876, http://dx.doi.org/10.3402/tellusa.v67.28876.

In Paper III, a new algorithm called Randomized Multi-Channel Singular Spectrum Analysis (RMSSA) is introduced. RMSSA is a generalization of the traditional MSSA into problems of arbitrarily large dimension. RMSSA is applied to decompose the 20th Century global monthly mean near-surface temperature of a reanalysis

data set and two climate model simulations. The decomposition into low-frequency patterns reveals, for example, that the 2–6 year variability centered in the Pacific Ocean is captured by all the data sets with some differences in statistical significance and spatial patterns.

I was responsible for all the computations, and for major part of the analysis and writing.

IV Järvinen, H., **Seitola, T.**, Silén, J., and Räisänen, J., 2016. Multi-annual modes in the 20th century temperature variability in reanalyses and CMIP5 models. *Geosci. Model Dev. Discuss.*, doi:10.5194/gmd-2016-61, accepted, in press.

Paper IV compares the multi-annual near-surface temperature variability modes in 12 CMIP5 model simulations and two reanalysis data sets using the randomised multi-channel singular spectrum analysis (RMSSA). The two reanalysis data sets are very similar on all time scales, except that the spectral power in ERA-20C is slightly higher than in 20CR. None of the climate models closely reproduce all aspects of the reanalysis data sets, although some models represent many aspects well.

I was responsible for all the computations and wrote the data and method descriptions. I also participated in the analysis and writing of the results.

References

Achlioptas, D. Database-friendly random projections: Johnson-Lindenstrauss with binary coins. *Journal of computer and System Sciences*, 66(4):671–687, 2003.

Aires, F., Chédin, A., and Nadal, J.-P. Independent component analysis of multivariate time series - Application to the tropical SST variability. *Journal of Geophysical Research*, 105(D13):17–437, 2000.

Allen, M. and Robertson, A. Distinguishing modulated oscillations from coloured noise in multivariate datasets. *Climate Dynamics*, 12(11):775–784, 1996.

Ba, J., Keenlyside, N. S., Latif, M., Park, W., Ding, H., Lohmann, K., Mignot, J., Menary, M., Otterå, O. H., Wouters, B., Salas y Melia, D., Oka, A., Bellucci, A., and Volodin, E. A multi-model comparison of Atlantic multidecadal variability. *Climate Dynamics*, 43(9):2333–2348, doi:10.1007/s00382-014-2056-1, 2014.

Bellenger, H., Guilyardi, E., Leloup, J., Lengaigne, M., and Vialard, J. ENSO representation in climate models: from CMIP3 to CMIP5. *Climate Dynamics*, 42(7):1999–2018, doi:10.1007/s00382-013-1783-z, 2014.

Bingham, E. and Mannila, H. Random projection in dimensionality reduction: applications to image and text data. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 245–250. ACM, 2001.

Boer, G. J. Decadal potential predictability of twenty-first century climate. *Climate dynamics*, 36(5-6):1119–1133, 2011.

Boer, G. J., Kharin, V. V., and Merryfield, W. J. Decadal predictability and forecast skill. *Climate Dynamics*, 41(7):1817–1833, doi:10.1007/s00382-013-1705-0, 2013.

Broomhead, D. S. and King, G. P. Extracting qualitative dynamics from experimental data. *Physica D*, 20:217–236, 1986a.

Broomhead, D. S. and King, G. P. On the qualitative analysis of experimental dynamical systems. In Sarkar, S., editor, *Nonlinear Phenomena and Chaos*, pages 113–144. Adam Hilger, Bristol, 1986b.

Budich, R. G., Giorgetta, M. A., Jungclaus, J. H., Redler, R., and Reick, C. H. The MPI-M Millennium Earth System Model: an assembling guide for the COSMOS configuration, 2010.

Capotondi, A., Wittenberg, A. T., Newman, M., Lorenzo, E. D., Yu, J.-Y., Braconnot, P., Cole, J., Dewitte, B., Giese, B., Guilyardi, E., Jin, F.-F., Karnauskas, K., Kirtman, B., Lee, T., Schneider, N., Xue, Y., and Yeh, S.-W. Understanding ENSO diversity. *Bulletin of the American Meteorological Society*, 96(6):921–938, doi:10.1175/BAMS-D-13-00117.1, 2015.

Clement, A., Bellomo, K., Murphy, L. N., Cane, M. A., Mauritsen, T., Rädel, G., and Stevens, B. The Atlantic Multidecadal Oscillation without a role for ocean circulation. *Science*, 350(6258):320–324, doi:10.1126/science.aab3980, 2015.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R., Rutledge, G., Bessemoulin, P., et al. The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, 137(654):1–28, 2011.

Dasgupta, S. and Gupta, A. An elementary proof of a theorem of Johnson and Lindenstrauss. *Random Structures and Algorithms*, 22(1):60–65, doi:10.1002/rsa.10073, 2003.

Deegalla, S. and Boström, H. Reducing high-dimensional data by principal component analysis vs. random projection for nearest neighbor classification. In *Machine Learning and Applications, 2006. ICMLA'06. 5th International Conference on*, pages 245–250. IEEE, 2006.

Delworth, L. T. and Mann, E. M. Observed and simulated multidecadal variability in the Northern Hemisphere. *Climate Dynamics*, 16(9):661–676, doi:10.1007/s003820000075, 2000.

Delworth, T., Manabe, S., and Stouffer, R. J. Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model. *Journal of Climate*, 6(11):1993–2011, doi:10.1175/1520-0442(1993)006¡1993:IVOTTC¿2.0.CO;2, 1993.

Delworth, T. L. and Zeng, F. The impact of the North Atlantic Oscillation on climate through its influence on the Atlantic Meridional Overturning Circulation. *Journal of Climate*, 29(3):941–962, doi:10.1175/JCLI-D-15-0396.1, 2016.

Demšar, U., Harris, P., Brunsdon, C., Fotheringham, A. S., and McLoone, S. Principal component analysis on spatial data: an overview. *Annals of the Association of American Geographers*, 103(1):106–128, 2013.

Fern, X. Z. and Brodley, C. E. Random projection for high dimensional data clustering: A cluster ensemble approach. In *ICML*, volume 3, pages 186–193, 2003.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., et al. Evaluation of climate models. In *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 741–866. Cambridge University Press, 2013.

Foster, G. and Abraham, J. Lack of evidence for a slowdown in global temperature. *Variations*, 13(3). [Available online at `https://usclivar.org/sites/default/files/documents/2015/Variations2015Summer.pdf`], 2015.

Frankl, P. and Maehara, H. The Johnson-Lindenstrauss lemma and the sphericity of some graphs. *Journal of Combinatorial Theory, Series B*, 44(3):355–362, 1988.

Fredriksen, H.-B. and Rypdal, K. Spectral characteristics of instrumental and climate model surface temperatures. *Journal of Climate*, 29(4):1253–1268, doi:10.1175/JCLI-D-15-0457.1, 2016.

Fyfe, J. C., Meehl, G. A., England, M. H., Mann, M. E., Santer, B. D., Flato, G. M., Hawkins, E., Gillett, N. P., Xie, S.-P., Kosaka, Y., et al. Making sense of the early-2000s warming slowdown. *Nature Climate Change*, 6(3):224–228, 2016.

Ghil, M., Allen, M., Dettinger, M., Ide, K., Kondrashov, D., Mann, M., Robertson, A. W., Saunders, A., Tian, Y., Varadi, F., et al. Advanced spectral methods for climatic time series. *Reviews of geophysics*, 40(1), 2002.

Goel, N., Bebis, G., and Nefian, A. Face recognition experiments with random projection. In *Defense and Security*, pages 426–437. International Society for Optics and Photonics, 2005.

Groth, A. and Ghil, M. Monte Carlo Singular Spectrum Analysis (SSA) revisited: Detecting oscillator clusters in multivariate datasets. *Journal of Climate*, 28(19):7873–7893, 2015.

Guilyardi, E., Wittenberg, A., Fedorov, A., Collins, M., Wang, C., Capotondi, A., van Oldenborgh, G. J., and Stockdale, T. Understanding El Niño in ocean-atmosphere general circulation models: Progress and challenges. *Bulletin of the American Meteorological Society*, 90(3):325–340, doi:10.1175/2008BAMS2387.1, 2009.

Halko, N., Martinsson, P.-G., and Tropp, J. A. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM review*, 53(2):217–288, 2011.

Hannachi, A., Jolliffe, I. T., and Stephenson, D. B. Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, 27(9):1119–1152, doi:10.1002/joc.1499, 2007.

Hurrell, J. W. Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, 269(5224):676–679, doi:10.1126/science.269.5224.676, 1995.

Hurrell, J. W. and Deser, C. North Atlantic climate variability: The role of the North Atlantic Oscillation. *Journal of Marine Systems*, 78(1):28 – 41, doi:http://dx.doi.org/10.1016/j.jmarsys.2008.11.026, 2009.

Ineson, S., Scaife, A. A., Knight, J. R., Manners, J. C., Dunstone, N. J., Gray, L. J., and Haigh, J. D. Solar forcing of winter climate variability in the Northern Hemisphere. *Nature Geoscience*, 4(11):753–757, 2011.

Johnson, W. B. and Lindenstrauss, J. Extensions of Lipschitz mappings into a Hilbert space. *Contemporary mathematics*, 26(189-206):1, 1984.

Jungclaus, J. MPI-M Earth System Modelling framework: millennium full forcing experiment (ensemble member 1). `http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=mil0010`, 2008.

Jungclaus, J., Keenlyside, N., Botzet, M., Haak, H., Luo, J.-J., Latif, M., Marotzke, J., Mikolajewicz, U., and Roeckner, E. Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM. *Journal of Climate*, 19(16):3952–3972, 2006.

Kao, H.-Y. and Yu, J.-Y. Contrasting Eastern-Pacific and Central-Pacific types of ENSO. *Journal of Climate*, 22(3):615–632, doi:10.1175/2008JCLI2309.1, 2009.

Karl, T. R., Arguez, A., Huang, B., Lawrimore, J. H., McMahon, J. R., Menne, M. J., Peterson, T. C., Vose, R. S., and Zhang, H.-M. Possible artifacts of data biases in the recent global surface warming hiatus. *Science*, 348(6242):1469–1472, doi:10.1126/science.aaa5632, 2015.

Keenlyside, N. S. and Ba, J. Prospects for decadal climate prediction. *Wiley Interdisciplinary Reviews: Climate Change*, 1(5):627–635, doi:10.1002/wcc.69, 2010.

Knutson, T. R., Zeng, F., and Wittenberg, A. T. Multimodel assessment of regional surface temperature trends: CMIP3 and CMIP5 twentieth-century simulations. *Journal of Climate*, 26(22):8709–8743, doi:10.1175/JCLI-D-12-00567.1, 2013.

Kumar, S., III, J. K., Dirmeyer, P. A., Pan, Z., and Adams, J. Multidecadal climate variability and the warming hole in North America: Results from CMIP5 twentieth- and twenty-first-century climate simulations. *Journal of Climate*, 26(11):3511–3527, doi:10.1175/JCLI-D-12-00535.1, 2013.

Kushnir, Y. Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *Journal of Climate*, 7(1):141–157, 1994.

Latif, M. and Keenlyside, N. S. A perspective on decadal climate variability and predictability. *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(17):1880–1894, 2011.

Lewandowsky, S., Risbey, J. S., and Oreskes, N. The pause in global warming: Turning a routine fluctuation into a problem for science. *Bulletin of the American Meteorological Society*, 97(5):723–733, doi:10.1175/BAMS-D-14-00106.1, 2016.

Moron, V., Robertson, A. W., and Ghil, M. Impact of the modulated annual cycle and intraseasonal oscillation on daily-to-interannual rainfall variability across monsoonal India. *Climate dynamics*, 38(11-12):2409–2435, 2012.

O'Reilly, C. H., Huber, M., Woollings, T., and Zanna, L. The signature of low-frequency oceanic forcing in the Atlantic Multidecadal Oscillation. *Geophysical Research Letters*, 43(6):2810–2818, doi:10.1002/2016GL067925. 2016GL067925, 2016.

Otto-Bliesner, B. L., Brady, E. C., Fasullo, J., Jahn, A., Landrum, L., Stevenson, S., Rosenbloom, N., Mai, A., and Strand, G. Climate variability and change since 850 CE: An ensemble approach with the Community Earth System Model. *Bulletin of the American Meteorological Society*, 97(5):735–754, doi:10.1175/BAMS-D-14-00233.1, 2016.

Parker, D. E., Legg, T. P., and Folland, C. K. A new daily central England temperature series, 1772–1991. *International Journal of Climatology*, 12(4):317–342, 1992.

Plaut, G. and Vautard, R. Spells of low-frequency oscillations and weather regimes in the Northern Hemisphere. *Journal of the Atmospheric Sciences*, 51(2):210–236, 1994.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thpaut, J.-N., Trmolet, Y., Hlm, E. V., Bonavita, M., Isaksen, L., and Fisher, M. ERA-20C: An atmospheric reanalysis of the 20th century. *Journal of Climate*, doi:10.1175/JCLI-D-15-0556.1, 2016.

Poli, P., Hersbach, H., Tan, D., Dee, D., Thepaut, J.-N., Simmons, A., Peubey, C., Laloyaux, P., Komori, T., Berrisford, P., Dragani, R., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M. The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C). Era report series, ECMWF, 2013.

Poli, P. and NCAR Staff (Eds.). The climate data guide: ERA-20C: ECMWF's atmospheric reanalysis of the 20th century (and comparisons with NOAA's 20CR). `https://climatedataguide.ucar.edu/climate-data/era-20c-ecmwfs-atmospheric-reanalysis-20th-century-and-comparisons-noaas-20cr`. [Online; accessed 11-May-2016], 2016.

Qi, H. and Hughes, S. M. Invariance of principal components under low-dimensional random projection of the data. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 937–940. IEEE, 2012.

Rayner, N., Brohan, P., Parker, D., Folland, C., Kennedy, J., Vanicek, M., Ansell, T., and Tett, S. Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The HadSST2 dataset. *Journal of Climate*, 19(3):446–469, 2006.

Rayner, N., Parker, D. E., Horton, E., Folland, C., Alexander, L., Rowell, D., Kent, E., and Kaplan, A. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, 108(D14), 2003.

Rind, D., Lean, J., Lerner, J., Lonergan, P., and Leboissitier, A. Exploring the stratospheric/tropospheric response to solar forcing. *Journal of Geophysical Research: Atmospheres*, 113(D24), doi:10.1029/2008JD010114. D24103, 2008.

Scaife, A. A., Folland, C. K., Alexander, L. V., Moberg, A., and Knight, J. R. European climate extremes and the North Atlantic Oscillation. *Journal of Climate*, 21(1):72–83, doi:10.1175/2007JCLI1631.1, 2008.

Schurer, A. P., Tett, S. F., and Hegerl, G. C. Small influence of solar variability on climate over the past millennium. *Nature Geoscience*, 7(2):104–108, 2014.

Shindell, D. T., Schmidt, G. A., Mann, M. E., Rind, D., and Waple, A. Solar forcing of regional climate change during the Maunder Minimum. *Science*, 294(5549):2149–2152, doi:10.1126/science.1064363, 2001.

Solomon, A., Goddard, L., Kumar, A., Carton, J., Deser, C., Fukumori, I., Greene, A. M., Hegerl, G., Kirtman, B., Kushnir, Y., Newman, M., Smith, D., Vimont, D., Delworth, T., Meehl, G. A., and Stockdale, T. Distinguishing the roles of natural and anthropogenically forced decadal climate variability. *Bulletin of the American Meteorological Society*, 92(2):141–156, doi:10.1175/2010BAMS2962.1, 2011.

Steinman, B. A., Mann, M. E., and Miller, S. K. Atlantic and Pacific multidecadal oscillations and Northern Hemisphere temperatures. *Science*, 347(6225):988–991, doi:10.1126/science.1257856, 2015.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498, 2012.

Trenberth, K. E. and Caron, J. M. The Southern Oscillation revisited: Sea level pressures, surface temperatures, and precipitation. *Journal of Climate*, 13(24):4358–4365, doi:10.1175/1520-0442(2000)013¡4358:TSORSL¿2.0.CO;2, 2000.

Trigo, R. M., Osborn, T. J., Corte-Real, J. M., et al. The North Atlantic Oscillation influence on Europe: climate impacts and associated physical mechanisms. *Climate Research*, 20(1):9–17, 2002.

Tung, K.-K. and Zhou, J. Using data to attribute episodes of warming and cooling in instrumental records. *Proceedings of the National Academy of Sciences*, 110(6):2058–2063, doi:10.1073/pnas.1212471110, 2013.

Vautard, R. and Ghil, M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D: Nonlinear Phenomena*, 35(3):395–424, 1989.

Von Storch, H. and Zwiers, F. W. *Statistical analysis in climate research*. Cambridge university press, 2001.

I

BOREAL ENVIRONMENT RESEARCH 19: 387–407
ISSN 1239-6095 (print) ISSN 1797-2469 (online)

Helsinki 28 November 2014

Decadal climate variability and potential predictability in the Nordic region: a review

Teija Seitola[1)2)] and Heikki Järvinen[2)]

[1)] *Finnish Meteorological Institute, P.O. Box 503, FI-00101 Helsinki, Finland*
[2)] *Department of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland*

Received 29 Apr. 2013, final version received 30 Jan. 2014, accepted 27 Jan. 2014

Seitola, T. & Järvinen, H. 2014: Decadal climate variability and potential predictability in the Nordic region: a review. *Boreal Env. Res.* 19: 387–407.

This paper reviews decadal climate variability and predictability and its potential implications for adaptation decisions with emphasis on the Nordic region. In the North Atlantic sector, there is strong decadal to multi-decadal climate variability. The most prominent internal mechanisms explaining the variability are the North Atlantic oscillation (NAO) and the Atlantic meridional overturning circulation (AMOC). These affect also the climate variability in the Nordic region, but their impacts vary depending on local and regional conditions. The published studies also indicate that there appears to be potential for predictability of the decadal climate variations in the North Atlantic sector, mainly due to the AMOC variations. This also contributes to the predictability in the Nordic region, especially in the coastal areas adjacent to North Atlantic. The conclusions are uncertain at the moment, despite the fact that new knowledge is rapidly accumulating. Potential decadal predictability may generally be reduced due to global warming which is the largest over the high latitude oceans. For instance, weakening of the AMOC is generally noted in warmer world simulations. This may have consequences also on climate in the Nordic region, although the response is still uncertain.

Introduction

Adaptation to anthropogenic climate change has a typical time perspective of 10 to 30 years into the future. This is too far considering the current capabilities of weather centres regarding seasonal to inter-annual forecasting. At the same time, it is too close for the long-term climate change projections of climate service centres. An entirely new field of Earth science — decadal climate prediction — is thus emerging to bridge this gap and to provide guidance for planning and decision making.

Worldwide climate records contain plenty of evidence of climate variations at decadal time-scales. Of direct relevance to society, decadal to inter-decadal fluctuations are found in atmospheric circulation patterns, precipitation, and climate extremes (Keenlyside and Ba 2010). One extreme example is the Sahel drought in the 1980s which had profound effects on ecosystems and societies. In fact, the 1980s drought is just the latest one in a sequence of recurring events. It seems plausible that past and future drought events in the Sahel region are linked with the atmosphere–ocean–biosphere coupling at multi-decadal time-scales, driven by the low-frequency ocean fluctuations (Held *et al*. 2005, Zhang and

Editor in charge of this article: Veli-Matti Kerminen

Delworth 2006, Shanahan *et al*. 2009). Another example is the inter-decadal variation associated with North Atlantic oscillation such as strong changes in wintertime storminess, and European and North American surface temperature and precipitation (Hurrell *et al*. 2003). European temperature extremes also exhibit multi-decadal variations (Beniston and Stephenson 2004). Consequently, these events could in principle be predicted if the ocean state was known and the coupled Earth system models were initialized with faithful representations of the true Earth system state (Keenlyside *et al*. 2008). This is in sharp contrast to the common practice in long-term climate simulations which are initialized without sophisticated Earth system initial states. In these simulations, decadal scale climate variations appear somewhat realistic but are not in close correspondence with the actual variations.

Extensive resources are currently used to build decadal climate prediction capabilities in U.S., Europe and elsewhere. These aim at using climate observations in the context of Earth system models to explore the limits of predictability, formulated as initial value problems (Meehl *et al*. 2009). The initialization methodologies bear close resemblance to the data assimilation techniques used in numerical weather prediction. These industrial-scale research and development lines are beyond capacities of individual small nations. By clever networking we can however take full advantage of these international efforts, and feedback meaningful contributions in selected areas.

Decadal climate prediction technology is still in its infancy. While worldwide developments are taking place in key areas of predictive capabilities, basic research is needed to assess the level decadal predictability in the Nordic region, and its possible implications for preparation of adaptation decision. This article will review the current knowledge on decadal climate variability and potential predictability. The emphasis will be on how to interpret this knowledge from Nordic region's viewpoint.

Decadal climate variability

Climate variability can be described with anomalies, which are differences between momentary states of the climate system and the longer-term climatology. Climatology is the mean state computed over months, years, decades or centuries (Hurrell and Deser 2009). Climate variability may be due to natural internal processes within the climate system (internal variability), or to variations in natural (volcanic eruptions, solar activity) or anthropogenic external forcing.

Climate variations occur at practically all conceivable time-scales. Oceans play a crucial role in decadal and longer-term climate variability because the effect of the annual cycle and month-to-month variability in the atmospheric circulation decays rapidly with depth (Hurrell and Deser 2009). The mechanisms behind the decadal-to-multi-decadal variability are not well understood, but there is some consensus that the longer-term variability is driven by internal climate variations like the North Atlantic Oscillation (NAO), the El Niño/Southern Oscillation (ENSO), the Pacific Decadal Variability (PDV), the Atlantic Multi-decadal Variability (AMV) and the Atlantic Meridional Overturning Circulation (AMOC).

Decadal climate variability can be studied based on observations and climate model simulations. Since direct observational records are relatively short and sparse, especially over the oceans, modelling studies are often used for studying decadal-scale climate fluctuations.

Observed mechanisms of decadal climate variability

Changes in naturally-occurring patterns of atmospheric and oceanic climate variability affect large-scale variations in weather and climate globally at inter-annual and longer time-scales (Hurrell and Deser 2009). The Atlantic multi-decadal variability (AMV) or oscillation (AMO) is a mode of variability which occurs in the Atlantic Ocean and is mainly manifested as sea surface temperature (SST) anomalies (Fig. 1a). AMO has been linked to changes in Sahel, North American and European precipitation (Sutton and Hodson 2005), Atlantic hurricane activity and northern hemisphere (NH) surface temperature (Zhang *et al*. 2007). AMV has some support in historical

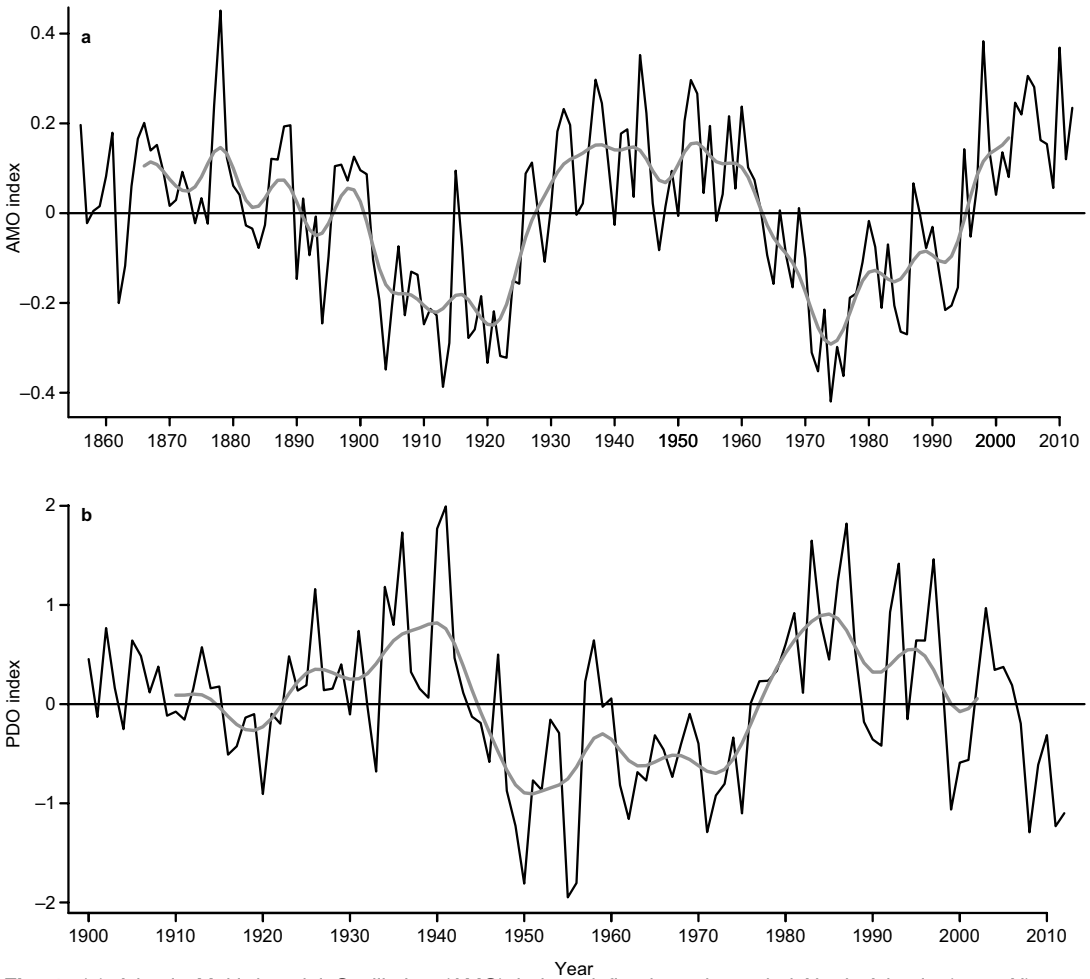


Fig. 1. (**a**) Atlantic Multi-decadal Oscillation (AMO) index, defined as detrended North Atlantic (0–70°N) area weighted average SST anomalies, and (**b**) Pacific Decadal Oscillation (PDO) index, derived as the leading PC of monthly detrended SST anomalies in the North Pacific Ocean, poleward of 20°N. Thin lines indicate annual mean and thick grey lines give smoothed annual values (smoothed with 21-point binomial filter). The AMO index data were obtained from http://www.esrl.noaa.gov/psd/data/correlation/amon.us.long.data and PDO index data from http://jisao.washington.edu/pdo/PDO.latest.

observations but there is controversy regarding its amplitude and it is estimated to have periods of about 40–70 years (e.g. Kushnir 1994, Delworth and Mann 2000).

The most prominent mechanism associated with AMV is the Atlantic Meridional Overturning Circulation (AMOC). It is a giant conveyor belt that brings warm water northwards into the North Atlantic, releases its heat to the atmosphere, and returns the cooled water to the south (Wood 2008). It consists of a wind-driven part and the thermohaline circulation (THC) (Pohlmann *et al*. 2006). There is evidence that the strength of this circulation can fluctuate naturally over periods of decades and it has the potential to influence North Atlantic and European climate (e.g. Pohlmann *et al*. 2006, Shaffrey and Sutton 2006). Since AMOC can affect the climate at multiple time-scales, there has been increasing interest in understanding the mechanisms behind the AMOC variability.

The North Atlantic Oscillation (NAO) is a leading pattern of weather and climate variability over the northern hemisphere. NAO is measured by an index which is defined as a difference in sea-level pressure between the Azores

high and the Iceland low (Hurrell 1995). NAO fluctuates from one phase to other producing large changes in surface air temperature, winds, storminess and precipitation over the Atlantic as well as the surrounding continents (Hurrell and Deser 2009). A positive phase of NAO is associated with anomalous low pressure in the subarctic and high pressure in subtropics with stronger westerly winds and enhanced flow of warm and moist air across the North Atlantic and Europe (Hurrel 1995).

According to Hurrel (1995), NAO exhibits quite strong inter-annual variability, but also some considerable decadal to multi-decadal variability. It has been shown that decadal to multi-decadal variations coherent with those in NAO can be also observed in the ocean (Curry *et al*. 1998), which indicates that there exists some kind of atmosphere–ocean interaction. However, Hurrel and Deser (2009) argued that there is little evidence for NAO to vary at any preferred time-scale: large changes can occur from one winter to the next, as well as from one decade to the next. In some studies decadal to multi-decadal changes in NAO have been linked to SSTs in the tropical Atlantic (Okumura *et al*. 2001) and Indo-Pacific region (e.g. Hoerling *et al*. 2001). Rodwell *et al*. (1999) suggested that NAO variability is mainly dependent on North Atlantic SST.

It is not clear which parts of the World Ocean drive the low frequency variations of NAO. Therefore it is important to take a global view on patterns of climate variability when considering the possible mechanisms of decadal climate variability in the North Atlantic sector. Similar decadal to multi-decadal variability as in the North Atlantic sector is seen in the North Pacific. Pacific Decadal Variability (PDV) or Pacific Decadal Oscillation (PDO) is a pattern of Pacific climate variability that is manifested as positive or negative SST anomalies in the tropical Pacific and opposite anomalies in the western extra-tropical North and South Pacific. The PDO index is derived from the monthly detrended SST anomalies in the North Pacific (Fig. 1b). PDO has been associated with variations in surface temperature and precipitation in the land areas at the rim of the Pacific, the Pacific marine ecosystem, and the Indian monsoon (Keenlyside and Ba 2010).

It has also been shown by Fraedrich and Müller (1992) and by Merkel and Latif (2002) that there is a significant response of the atmosphere over the North Atlantic to ENSO-related variations in tropical Pacific SST. ENSO is a climate pattern that is related to coupled atmosphere–ocean variations: warming (cooling) of eastern tropical Pacific SST and high (low) surface pressure in the western tropical Pacific (Trenberth *et al*. 2007). ENSO has an average period of five years but since it can influence the global circulation patterns far away from the tropical Pacific through the atmospheric bridge (Liu and Alexander 2007), it is also interesting from decadal climate variation point of view.

Inter-decadal climate variations may partly result from processes that are external to the climate system. These are variations in solar activity, volcanic eruptions and anthropogenically forced changes in greenhouse gas concentrations and aerosols. The amplitude of past variations in solar forcing is much debated but the irradiance variations over the 11-year sunspot cycles are quite well known, as they can be calibrated against satellite measurements since 1979 (Gouirand *et al*. 2007). The climate effects of strong volcanic eruptions can persist for about a decade (Latif and Keenlyside 2011). However, volcanic eruptions cannot be predicted but because of strong effects, they should be considered in decadal predictions. Anthropogenic changes in greenhouse gases and aerosols are an important forcing for climate at longer time-scales and should be taken into account when analysing multi-decadal variability. The role of uncertainties in forcing of anthropogenic emissions is anyhow likely to be relatively small at decadal time-scales (Meehl *et al*. 2009).

There is a controversy on how internal variability and external forcing affect the decadal variability. According to some studies natural decadal to multi-decadal climate variability has a potential to mask or enhance anthropogenic climate change, particularly at a regional level (e.g. Meehl *et al*. 2009, Keenlyside and Ba 2010). On the other hand, unpredictable external forcing through explosive volcanic eruptions and anomalous solar radiation may offset the internal variations (Latif and Keenlyside 2011). The relative roles of internal and external processes in

driving decadal and multi-decadal climate variations are not well understood and more precise quantification is needed. In addition, apart from separating the effects of external forcing and internal variability on inter-decadal variability, nonlinear interaction between these mechanisms should also be considered.

Decadal climate variability in modelling studies

Climate models can produce climate variability to some extent and therefore decadal climate variability can also be estimated based on climate model simulations. There are several studies that have investigated the contribution of AMOC to climate variability. For example, Latif *et al*. (2006a) investigated AMOC by analysing relationship between AMOC and SST found in global climate models. The strength of AMOC was defined as the SST difference between North and South Atlantic. Their results indicate that the AMOC variations are driven by the low-frequency variations of NAO through changes in the Labrador Sea convection and lag the corresponding variations of NAO by about a decade. In a more recent study, Ortega *et al*. (2012) analysed the AMOC variability in an unforced present-day control run, two forced runs for the last millennium, and two IPCC scenarios with ECHO-G atmosphere–ocean general circulation model. They suggest that at low frequencies (decadal to multi-decadal time-scales) AMOC is largely controlled by convection activity south of Greenland (Labrador and Irminger Seas) and the influence of NAO on AMOC through convection changes in this area is also identified. These results are in line with the findings in Latif *et al*. (2006a). In addition to Latif *et al*. (2006a) and Ortega *et al*. (2012), several other modelling studies suggest that AMOC may contribute to climate variability at inter-annual and decadal to multi-decadal time-scales (e.g. Delworth *et al*. 1993, Delworth and Mann 2000, Sutton and Hodson 2005). However, some observation-based studies indicate that the NAO–AMO/AMOC relation calls for further research. For example Walter and Graf (2002) identified a non-stationary relation between NAO and AMO: during the negative phase of AMO, North Atlantic SST is strongly correlated with the NAO index, but during the positive phase the correlation is weak. Vukcevic (2011) also showed the complexity of AMO–NAO multi-decadal relationship.

Modelling studies also indicate that external forcing has considerable effect on climate variations at multiple time-scales. For example Bauer *et al*. (2003) estimated the effects of natural and anthropogenic external forcing on climate variability for the past millennium. Their results indicate that the pre-industrial variations in the northern hemisphere (NH) temperature at annual to multi-centennial scales are predominantly caused by solar and volcanic activity. In the industrial period, increasing greenhouse gases and deforestation additionally affect temperature variability. However, Bauer *et al*. (2003) did not take any stand on what is the role of natural internal processes on NH temperature variations.

Ineson *et al*. (2011) investigated solar forcing of winter climate variability in NH. An ocean–atmosphere climate model was driven with ultraviolet variations estimated from satellite observations of solar variability. Their modelling results show that the solar minimum is connected to pressure and surface temperature patterns that resemble the negative phase of NAO and Arctic Oscillation (AO). Ineson *et al*. (2011) suggested that this result could have important implications in decadal prediction of the NAO.

Nonlinear interaction between external forcing and internal variability and its effect on inter-decadal variability is also estimated in modelling studies: e.g., in Dunstone *et al*. (2013) it is shown that decadal variability in tropical storm frequency is well reproduced through aerosol-induced north–south shifts in the Hadley circulation and only after incorporating aerosol effects in the model. In addition, the sensitivity of AMOC to external forcing was investigated by Ortega *et al*. (2012). Their results show that starting from the industrial era, increasing greenhouse gases have a major impact on AMOC weakening. There is also a weak but significant signal of AMOC strengthening because of major volcanic eruptions. This is due to the fact that volcanic eruptions produce colder and saltier surface conditions over the main convection

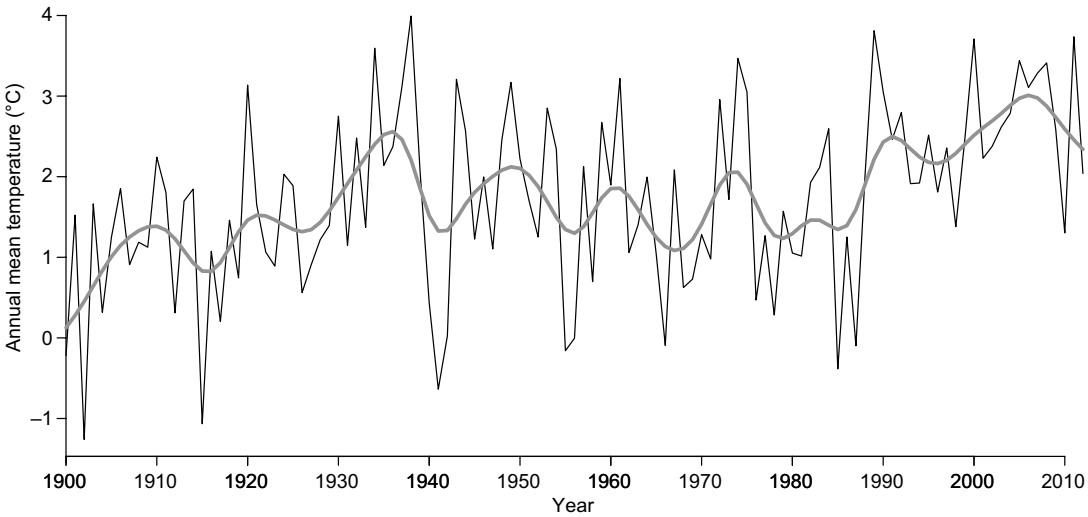


Fig. 2. Annual mean temperature (°C) of Finland 1900–2012 based on spatially interpolated monthly mean temperature records. Black line indicates annual values and grey line gives smoothed annual values (smoothed with 21-point binomial filter). Time series was extended at the end points before filtering to make the filtered time series cover the whole time range. After Tietäväinen *et al.* (2010).

regions driving AMOC. The impact of solar forcing on AMOC is deemed small.

Decadal climate variability in the Nordic region

At northern high latitudes, climate is characterized by large inter-annual and inter-decadal variability. For example, Tietäväinen *et al*. (2010) studied the annual and seasonal mean temperature climatology in 1847–2008 in Finland and showed that there is a distinct division into periods of cold and warm years with decadal-scale fluctuations (Fig. 2). Long-term climate variability in the Nordic region can be estimated based on observational data and simulations. However, the observational record is relatively short for estimating decadal-scale variability. According to Gouirand *et al*. (2007) there are few proxy-based temperature reconstructions for Scandinavia available for the entire millennium but they reflect only a fraction of the true climate variations and only for certain parts of the year.

The most prominent internal mechanisms affecting the climate variability in the Nordic region are NAO and AMOC. NAO has a significant influence on wintertime temperatures and precipitation in the Nordic region. Winters with positive the NAO index are associated with warmer than normal surface air temperatures and wetter than normal conditions over northern Europe (e.g. Hurrell and van Loon 1997, Serreze *et al*. 1997, Alexandersson *et al*. 1998, Visbeck *et al*. 2003). However, impacts of atmospheric circulation pattern depend crucially on local or regional details (Hurrell and Deser 2009) and according to Blenckner *et al*. (2004) it appears as if the influence of NAO on the local climate is less pronounced north of 65°N.

Modelling studies show that AMOC fluctuations have also the potential to influence the climate in the Nordic region. Persechino *et al*. (2013) studied the regional impact of AMOC variability at the decadal time-scale with the IPSL-CM5A-LR model. Their study results show that the AMOC impact on surface temperature at the decadal time-scale is dominant over the North Atlantic. Impact is much weaker over land but some marine influenced regions of western Europe show weak signal including parts of Scandinavia. The signal of the impact of AMOC variability on precipitation at the decadal time-scale largely resembles the corresponding signal of surface temperature. As for temperature, the impact over land can be seen in areas close to the ocean.

For the Nordic region, the study indicates that in the case of strong AMOC there might be drier in Norway and wetter in northern Finland. An earlier study by Pohlmann *et al*. (2006) investigated the influence of AMOC on European surface air temperature (SAT) by calculating the probability density functions (PDFs) of European SAT for strong and weak overturning conditions and using Atmosphere–Ocean General Circulation Model (AOGCM) ECHAM5/MPI-OM. In the case of weak AMOC conditions, SAT averaged over Europe is colder than in the case of strong AMOC conditions, and vice versa. The difference of mean SAT between years with strong and weak AMOC conditions increases from southern to northern Europe. According to Pohlmann *et al*. (2006) the difference of the mean precipitation between strong and weak AMOC displays also an enhancement over northern Europe.

In addition to internal variability, the external mechanisms, such as changes in radiative forcing caused by variations in greenhouse gases, solar irradiation and volcanic aerosols can affect the climate in the Nordic region. According to simulations of Gouirand *et al*. (2007), decadal and multi-decadal deviations from the centennial cooling–warming pattern in Scandinavia are the result of different causes. Some cold intervals can be explained with temporary decreases in solar radiation and sequences with strong volcanic eruption events. Negative phase in NAO can also explain especially low winter temperatures through a weakened westerly flow, but also cold summer temperature because of the large heat capacity of the nearby ocean.

Decadal potential predictability

Decadal predictability has been estimated based mainly on modelling studies. Most of the predictability studies are concentrated at the global scale and decadal predictability is found predominately over the mid- to high-latitude oceans (e.g. Boer 2004, Pohlmann *et al*. 2004, Boer and Lambert 2008). In addition to North Atlantic, the Southern Ocean has been identified to be the most prominent region in decadal predictability studies (e.g. Boer and Lambert 2008, Koenigk *et al*. 2011). Climate model studies indicate that in these regions the potential predictability of decadal variations is due to variations in the ocean circulation and heat storage (Pohlmann *et al*. 2004). Several studies agree with the idea that predictability at decadal time-scales resides in the ocean, where information can be stored and later transferred to the atmosphere (Latif *et al*. 2006a). According to Boer (2010) internal and externally forced variability are both important sources of potential predictability in global-scale projections. However, at regional level, relative importance of these factors varies substantially.

Methods for measuring decadal predictability

In the literature usually two types of predictability studies are described: potential and classical. Decadal potential predictability can be defined as the ratio of the variance at decadal time-scales to the total variance (Latif *et al*. 2009).

Potential or diagnostic predictability studies (e.g. Boer 2004, Boer and Lambert 2008, Boer 2010, Persechino *et al*. 2013) try to quantify the fraction of long-term variability from the internally generated natural variability, which is not predictable at long time-scales and considered noise. The long-term variability signal that may be distinguished from this noise is thought to arise from potentially predictable processes in the physical system (Latif and Keenlyside 2011).

In classical or prognostic predictability studies (e.g. Pohlmann *et al*. 2004, Koenigk *et al*. 2011, Branstator *et al*. 2012, Persechino *et al*. 2013) ensemble experiments are performed with a single coupled model by perturbing the initial conditions. Predictability of a variable is given by the ratio of the actual signal variance to the ensemble variance. This method assumes a perfect model and often near-perfect initial conditions that usually gives an upper limit of predictability (Latif and Keenlyside 2011).

A method called ‘ocean dynamics approach’ (Park and Latif 2005) has also been used. This method compares the variability simulated with and without the ocean–sea ice dynamics and identifies those regions in which ocean dynamics are important in generating the decadal-scale variability. Those regions are believed to be the

regions of high decadal predictability potential.

Persechino *et al*. (2013) studied decadal predictability of AMOC with the IPSL-CM5A-LR model using both diagnostic and prognostic potential predictability measures. Their results showed that both diagnostic and prognostic approaches generally brought out the same main features concerning both temperature and precipitation predictability. According to Latif *et al*. (2006b), all the three methods (diagnostic, prognostic and ocean dynamics approach) yield similar patterns of decadal predictability.

Decadal potential predictability in the North Atlantic and Nordic Region

Decadal predictability studies done so far were mostly concentrated at the global scale. The aim was to point out the areas that have most potential for decadal predictions. Decadal predictability studies concentrated especially on the Nordic region were not found during this review study, but there are several studies that present regional information on predictability over ocean and land areas in the North Atlantic sector (e.g. Collins *et al*. 2006, Boer 2009, Boer 2010, Hermanson and Sutton 2010, Koenigk *et al*. 2011, Persechino *et al*. 2013). A summary of the studies is presented in Table 1, with emphasis on the Nordic region.

Potential of decadal predictability appears to be quite large in the North Atlantic sector. The most prominent mechanism driving some of the decadal-scale variability seems to be AMOC which is a focus of many recent predictability studies (e.g. Msadek *et al*. 2010, Ortega *et al*. 2011, Tulloch and Marshall 2012, Persechino *et al*. 2013). Early analysis of Delworth *et al*. (1993) showed that there is a broad resemblance between simulated and observed multi-decadal SST variability patterns in the North Atlantic and that is usually associated with AMOC. Based on these results variability of AMOC may be predictable at decadal or longer time-scales. A multi-model-ensemble study of Collins *et al*. (2006) indicated potential predictability of inter-annual–decadal AMOC variations for one to two decades into the future. Persechino *et al*. (2013) showed that modelled AMOC has an average predictive skill of eight years. Studies of Collins and Sinha (2003), Sutton and Hodson (2005) and Pohlmann *et al*. (2006) showed that multi-decadal AMOC predictability in the HadCM3 and ECHAM5/MPI-OM models leads to some predictability of European climate. More recently Ortega *et al*. (2011) studied the processes that influence predictability of decadal variability in AMOC with the ECHO-G coupled climate model. They identified two predictors of AMOC variability: the anomalous heat flux averaged over a region in the Eastern Labrador Sea and an anomalous ocean density in a region of the Western Irminger Sea. These predictors together account for over 80% of the inter-annual variance of AMOC (Ortega *et al*. 2011). Thus, most state-of-the-art climate models seem to indicate that AMOC variations are predictable at decadal scales although there are still major uncertainties regarding the level and extent of predictability of different oceanic and atmospheric variables (Latif *et al*. 2006b).

Boer (2010) estimated the potential predictability of temperature and precipitation and its forced and internal components for the first part of the 21st century based on simulation data from a collection of coupled climate model results in the CMIP3 data archive. He used two measures of potential predictability. First, the multi-decadal view considered the forced component to be the difference from the beginning of the century. Second, the next-decade view considered the change in the forced component from the previous decade, thus putting emphasis on the change from the present rather than from an earlier period. Results of Boer (2010) show that in case of temperature, the forced component of potential predictability is generally largest over tropical oceans and declines with latitude being relatively low over mid- to high-latitude land. In contrast, internally-generated decadal potential predictability for temperature is largest over mid- to high-latitude oceans. It seems that internally-generated decadal potential predictability in the Nordic region is quite weak (2%–10%), but it is still slightly higher as compared with that for other land areas (2%–5%) (Table 1). However, it should be kept in mind that over land, the long time-scale internally-generated variability in temperature (the "signal") is masked by the rela-

Table 1. Summary of decadal predictability studies and implications for the Nordic region. The notation: ppvf stands for potential predictability variance fraction.

Study	Aim of the study	Model/Data	Method	Results in the Nordic region
Boer 2009	Internally generated variability of the unforced climate was compared with simulations with the B1 and A1B climate change scenarios	Multimodel simulation data from the CMIP3 archive	Diagnostic	Temperature: decadal potential predictability 0%–10% (unforced preindustrial control simulation) Precipitation: no predictability
Boer 2010	Potential predictability of temperature and precipitation and its forced and internal components were estimated for the first part of the 21st century	Multimodel simulation data from the CMIP3 archive, SRES B1 scenario	Diagnostic	Temperature: internally generated potential predictability 2%–10% (ppvf), forced ppvfs comparatively small in middle to high latitude areas Precipitation: internally generated and forced potential predictability very weak
Hermanson and Sutton 2010	The predictability of ocean and climate variables were investigated using a perfect model-based case study approach	Model: HadCM3	Prognostic	No significant longer time-scale predictability for temperature or precipitation
Koenigk *et al.* 2011	Study analyzed the upper limit of climate predictability on decadal time scales and its dependency on sea ice albedo parameterization	Model: EC-Earth, two perfect ensemble experiments	Prognostic	Temperature: significant prognostic potential predictability 0.5–0.8 (on a scale from 0 to 1) over Nordic region, close to high predictability over the sea Precipitation: no clear signal for predictability
Persechino *et al.* 2012	The decadal potential predictability of the AMOC and the predictability of associated oceanic and atmospheric fields were explored	Model: IPSL-CM5A-LR	Diagnostic/ prognostic	Temperature: internally generated decadal potential predictability 10%–20% in the coastal areas close to North Atlantic and 5%–10% elsewhere Precipitation: potential decadal predictability weak

tively strong short time-scale climate variability (the "noise"). Decadal potential predictability for precipitation for the unforced control climate is very weak. There is some, although relatively small potential predictability for precipitation due to the forced component mainly at middle to high latitude land areas (Boer 2010).

In line with previous studies, Persechino *et al*. (2013) found that potential predictability of surface temperature over land is less significant than over the ocean. Predictability over the coastal areas is found to be close to that of some of the potentially predictable oceanic regions and it is linked with AMOC fluctuations. In the Nordic region, internally-generated decadal potential predictability of surface temperature is 10%–20% in the coastal areas close to North Atlantic, and 5%–10% elsewhere (Table 1). These results are somewhat more positive than those of Boer (2010). Potential decadal predictability of precipitation is clearly smaller than for surface temperature and link to decadal AMOC fluctuations is less clear. However, the Nordic Seas are the most prominent regions where precipitation seems predictable at decadal time-scales. Persechino *et al*. (2013) also found convincing evidence that extreme changes in AMOC might be potentially predictable up to two decades ahead from the monitoring of its high-latitude Atlantic precursors (Sea Surface Salinity in the Labrador sea and the East Greenland Current (EGC) index). In line with the earlier study of Collins *et al*. (2006), results of Persechino *et al*. (2013) also show that the initial state corresponding to an anomalously strong AMOC is more predictable than those corresponding to weak AMOC.

A study of Koenigk *et al*. (2011) used prognostic methods for analysing the upper limit of climate predictability at decadal time-scales and its dependency on sea ice albedo parameterization with two perfect ensemble experiments with the global coupled climate model EC-Earth. Compared with experiment 1, in experiment 2, the sea-ice albedo was reduced by 0.03. Their results show that AMOC is highly predictable in both experiments and governs most of decadal climate predictability in the northern hemisphere. They found highest potential predictability for 2-m air temperature (T2m) over the northern North Atlantic and the southern South Atlantic. Also sea surface salinity and sea surface temperature show high predictability in these regions. Over most land regions, prognostic potential predictability of T2m is quite small and not significant. However, both experiments show a significant predictability of air temperature over northwestern Europe and most of the high potential predictability areas over land are located close to high predictability over sea (Table 1). In both experiments, precipitation shows largest decadal potential predictability in the northeastern North Atlantic and in the Barents Sea region as well as in the Labrador Sea. Compared with previous studies, results of Koenigk *et al*. (2011) indicate higher decadal predictability over land regions. The authors hypothesize that this might be due to higher resolution in EC-Earth as compared with that in the models used in most of the previous studies.

Hermanson and Sutton (2010) took a perfect model-based case study approach to investigate predictability of ocean and climate variables. They used the Hadley Centre HadCM3 coupled atmosphere–ocean model. Their results indicate that large-scale ocean variables such as volume-integrated ocean heat content, salinity or AMOC generally show significant predictability for several years or more. On the other hand, predictability of surface annual-mean climate variables is generally limited to two years at the most. Their results also indicate that there is no significant longer time-scale predictability for temperature or precipitation in the Nordic region. However, Hermanson and Sutton (2010) admitted that a single climate model of modest resolution they used and a small number of cases is the limitation of their study.

As mentioned before, some studies have identified connections between NAO and AMOC. According to Latif *et al*. (2006b) there is some evidence from observations of the last century and from forced ocean model simulations that the future state of AMOC may be predictable from past low-frequency variations of NAO. However, when considering predictability in the Atlantic sector, a global approach is needed because forcing from the other climate patterns in the tropics and extra tropics should also be considered.

Apart from potential predictability, actual near-term prediction skill is estimated for example in a recent study of Doblas-Reyes *et al.* (2013). Their study illustrates the forecast skill of initialized regional near-term climate predictions conducted as a part of the Fifth Coupled Model Intercomparison Project (CMIP5). The main result is that the climate forecast systems have a substantial skill in predicting multi-annual near-surface temperature anomalies at regional scales and most of the skill is due to changes in atmospheric composition, but also partly due to the initialization of predictions. In more detail, their results show significant skill in the North Atlantic for near-surface temperature predictions up to 6–9 years. In the Nordic region there is also some positive forecast skill but it is not statistically significant. The skill for land precipitation is much lower than for near-surface temperature, but there is some positive, although not statistically significant skill for predictions up to 6–9 years especially in the northern hemisphere and also in the Nordic region.

In summary it can be concluded that a potential for decadal predictability appears to be quite large in the North Atlantic sector and predictability is based on the variations of AMOC. This potential decadal-scale predictability of AMOC might also contribute to predictability in the Nordic region, especially the coastal areas close to North Atlantic, but any definite conclusions cannot be made yet. State-of-the-art climate prediction systems also show a substantial skill in predicting near-surface temperature up to 6–9 years in the North Atlantic. However, it should be kept in mind that the models used, the initial states employed and the measures of predictability differ among studies.

Decadal predictability under global warming

There are some studies that consider decadal climate variability and predictability under global warming (e.g. Parker *et al.* 2007, Boer 2009, Boer 2010). For example Parker *et al.* (2007) reviewed the most prominent modes of climate variability (e.g. PDO, ENSO, NAO, AMO) in the instrumental record and compared these with background signal of global warming. Their results show that regional climate variations result from these natural modes of decadal to inter-decadal variability as well as from anthropogenically-induced climate change in these modes. For example, the increase in NAO during 1965–1995 was partly naturally-induced but simulations also indicate that anthropogenic forcing was affecting the increase (Parker *et al.* 2007). According to Hurrell and Deser (2009) significant part of global warming in recent decades is attributed to decadal changes in two dominant climate patterns, NAO and ENSO. In addition, Corti *et al.* (1999) argued that the spatial pattern of the response to anthropogenic forcing may project principally onto these dominant modes of natural climate variability. Natural decadal to multi-decadal climate variability may also mask anthropogenic climate change (Latif *et al.* 2006b). Distinguishing the roles of natural internal and anthropogenically-forced variability is actually one of the major challenges in assessing decadal predictability and making regional decadal predictions (e.g. Solomon *et al.* 2011).

Boer (2009) compared the internally generated variability of the unforced climate with that of the warmer conditions for simulations with the B1 and A1B climate change scenarios. He investigated the changes in the variability of annual mean temperature and precipitation and in the variability of decadal potential predictability based on the collection of coupled climate model simulations in the Coupled Model Intercomparison Project phase 3 (CMIP3) data archive. According to the results of Boer (2009), global warming may induce a general decrease in decadal potential predictability for temperature and the decrease seems to be largest over the high-latitude oceans. Potential predictability of precipitation also decreases although it is already small in the beginning. In the Nordic region, decadal potential predictability of annual mean temperature is 0%–10% and there is no predictability for precipitation (unforced preindustrial control simulation). In a warmer world (B1 and A1B scenarios) decadal potential predictability of temperature and precipitation does not change or decreases slightly in the Nordic region (Table 1). According to Boer (2009) the overall decrease in decadal potential predict-

ability in future, warmer climate indicates that decadal potential predictability of the internally generated component might decrease. Results of Boer (2009) also show that long-term variability indices (e.g. AMOC, AMO, ENSO and NAO) may change because of global warming.

Weakening of AMOC and associated changes in heat transports are noted as a general result in warmer world simulations. For example, the simulations of Ortega *et al.* (2012), covering the Industrial Era and continuing in the future scenarios, show AMOC decreasing finally up to 40% when compared with the pre-industrial average. This final weakening is associated with a reduced meridional density gradient and with decreased convection in the North Atlantic. Therefore, the anthropogenic climate change may influence especially the Atlantic sector by inducing strong changes in the strength of the AMOC which in turn has direct consequences to North American and European climates (Latif *et al.* 2006b). However, there are still large uncertainties concerning the response of AMOC to global warming.

Decay of ice sheets and associated freshwater flux should also be considered in decadal predictability studies. According to Vizcaíno *et al.* (2010), ice sheets can modify atmospheric conditions via changes in e.g. albedo and orography, and indirectly via changes in ocean circulation. For example, Vizcaíno *et al.* (2010) studied the future evolution of global ice sheets under anthropogenic greenhouse forcing and its impact on the climate system with an Earth system model consisting of a coupled atmosphere–ocean general circulation model, a dynamic vegetation model and an ice sheet model. In their study, the North Atlantic meridional overturning circulation (NAMOC) weakens substantially in just 100 years in all the simulations. Their results show that the freshwater fluxes are dominated by increased precipitation over the ocean and increased river runoff. The freshwater flux from the Greenland ice sheet has a minor role. However, the modification of ocean density by the increased freshwater flux from the Greenland ice sheet seems to play an important role in hindering the recovery of the ocean circulation (Vizcaino *et al.* 2010). The experiments of Koenigk *et al.* (2011) also show that decadal variations are substantially smaller in the simulations with reduced ice albedo, which can be explained by reduced sea-ice thickness in these simulations. Koenig *et al.* (2011) hypothesized that reduced decadal-scale variations in the Arctic sea-ice volume reduces sea surface temperature and salinity variations in the Labrador Sea which in turn reduces the decadal variability of AMOC. This reduces temperature variations in mid- and high-latitude northern hemisphere regions. According to Latif and Keenlyside (2011) virtually all climate models considerably underestimate the observed Arctic sea-ice decline during the recent decades in the so-called 20th century integrations with prescribed (known natural and anthropogenic) observed forcing. This indicates that the simulations of future changes in ice sheets are still uncertain.

Decadal prediction

Seasonal prediction is considered an initial value problem (the evolution of the atmosphere-ocean system is largely determined by the initial condition) unlike centennial projection, which is a boundary value problem (the system evolution depends on the external forcing and formulation of boundary condition) (e.g. Palmer *et al.* 2004, IPCC 2007). Prediction of climate on decadal time-scales is somewhere between seasonal and centennial scales and produces both an initial and boundary value problems (Fig. 3).

In decadal prediction, initialization of climate models offers the potential to make predictions of internal variability in addition to external forcing. One of the major issues is the initialization technique used in predictions. The two main approaches are full-field initialization (in which an estimate of the observed climate state is used to initialize the model), and anomaly initialization (which uses estimates of observed ocean and sea ice anomalies on top of the model climatology) (e.g. Hazeleger *et al.* 2013). A comparison of the two methods shows that full-field initialization provides more skilful predictions at the seasonal time-scale (e.g. Smith *et al.* 2013a) but at the decadal time-scales the two methods show similar prediction skill (e.g. Hazeleger *et al.* 2013, Smith *et al.* 2013a). However, accord-

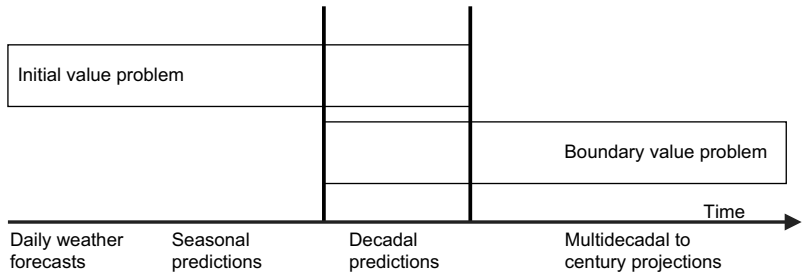


Fig. 3. Illustration of initial value problems with daily weather forecasts at one end, and multi-decadal to century projections as a boundary value problem at the other, with seasonal and decadal predictions in between. Adapted from Meehl *et al.* (2009).

ing to Hazeleger *et al*. (2013), anomaly initialization shows poorer skill in some regions (e.g. North Atlantic).

Initialized predictions should better quantify the uncertainty range in the near future by taking into account internal variability and the mean forced response. However, climate models are not perfect and when initialized with observations, they tend to drift towards their own and biased climatology (Meehl *et al*. 2009). Pioneering studies of Smith *et al*. (2007), Keenlyside *et al*. (2008) and Pohlmann *et al*. (2009) examined the impact of initial conditions on decadal the prediction skill with a coupled GCM. In general, these studies indicate that initialization improves the decadal prediction skill of climate variables. Results of Keenlyside *et al*. (2008) even indicate that initialization leads to a significant enhancement in the skill in the Nordic region. In contrast, Smith *et al*. (2007) does not suggest significant impact of initialization on the prediction skill for annual mean temperature in the Nordic region and results even show that in some regions initialization can lead to a decrease in the skill. Hermanson and Sutton (2010) showed that based on knowledge of initial conditions, climate variables are generally not predictable more than two years ahead, and only rarely predictable more than one year ahead. This discrepancy suggests that an improved skill in initialized predictions and hindcasts may arise from removing biases that exist in uninitialized climate models (e.g. Hermanson and Sutton 2010, Solomon *et al*. 2011).

In a recent study, Matei *et al*. (2012) investigated how two different ocean initializations (GECCO ocean reanalysis and an ensemble of ocean-forced experiments) impact the quality of decadal hindcasts performed with the ECHAM5/MPI-OM coupled model. Results show that initialization considerably increases the predictive skill of SST up to a decade ahead over the North Atlantic, central North Pacific, and the Mediterranean region. They found also a predictive skill of land surface air temperature at a decadal timescale in several land areas including northwestern Europe. Branstator *et al*. (2012) also quantified the initial-value predictability properties of six AOGCMs to help determine the benefit from initializing decadal predictions with the observed state of the climate system. Their results show that with a typical model and typical initial conditions predictability in upper-ocean conditions resulting from initialization lasts for about a decade in the North Atlantic, and somewhat less in the North Pacific. In line with many other studies, their study indicates that resources should be devoted to development of initialization of decadal predictions. On the other hand, it is highly uncertain to quantify the added value of these investments since the modelling results vary substantially (Branstator *et al*. 2012).

There are still many open questions concerning initialization. For example, impact of many processes in decadal predictions is still unsolved (such as sea ice and ocean conditions under the sea ice, snow cover and its modeling, frozen soil, soil moisture, stratospheric processes, land surface and vegetation). Initialization of these factors may have potential to contribute to the predictive skill. It has also been suggested that the skill of decadal prediction may depend on the initial state (e.g. Collins *et al*. 2006, Koenigk *et al*. 2011, Persechino *et al*. 2013). For example, AMOC seems to be more predictable if initialized from anomalously strong *versus* weak phase. Initialization is also substantially hampered by the lack of subsurface ocean observa-

tions and imperfect initialisation may lead to degradation of the forecast skill (e.g. Solomon *et al*. 2011). Furthermore, it is not entirely clear which is the best method for initialization.

In decadal prediction it is important to construct ensemble forecasts to sample the possible outcomes consistent with uncertainties in initial states and the model (Meehl *et al*. 2009). Constructing ensembles from different available GCMs has been shown to provide improved estimates of uncertainty as compared with single-model ensembles using only perturbed initial conditions (Hagedorn *et al*. 2005). Stochastic-dynamic parameterization schemes have also been proposed in this context to provide uncertainty estimates in decadal climate predictions (Palmer *et al*. 2009). Stochastic-dynamic methods are based on the fact that the climate system has components with different internal time-scales: fast components are treated as stochastic processes and the slow ones evolve following dynamical equations with stochastic forcing (Hasselmann 1976). There are also some statistical methods (such as lagged correlations, linear inverse modelling, and constructed analogues) that are found to have significant skill in predicting the internal variability of Atlantic SSTs for a decade ahead (Hawkins *et al*. 2011).

There is a broad set of decadal experiments conducted as part of CMIP5 (Coupled Model Intercomparison project) (Table 2). There are two core experiments, 10 and 30 year hindcasts (i.e., a "prediction" of the observed climate history of the recent past), or predictions. Ten-year simulations are initialized at least in every five years starting from 1960 and these experiments are meant for assessing the model skill in forecasting climate change at time-scales when the initial conditions drive the future evolution (Taylor *et al*. 2009). These experiments also try to increase the knowledge on decadal predictability and the best ways to initialize models in decadal predictions. The other core experiments extend the simulations initialized in years 1960, 1980 and 2005 to 30 years. These 30-year simulations explore the predictability and prediction in a longer time-scale when the external forcing from increasing greenhouse gas concentrations should become more important (Taylor *et al*. 2009). In these core experiments, volcanic

Table 2. CMIP5 decadal experiments. Information is mainly based on the ESGF-data portal (pcmdi9.llnl.gov/esgf-web-fe/live) and few other references that are noted as a superscript after the name of each model. The sources are: [1]bcc.cma.gov.cn/bcccsm/web/channel-34.htm, [2]www.cesm.ucar.edu/events/ws.2012/Presentations/CVC1/teng.pdf, [3]pcmdi9.llnl.gov/esgf-web-fe/live, [4]www.cccma.ec.gc.ca/data/cgcm4/CanCM4/index.shtml, and [5]gmao.gsfc.nasa.gov/research/climate/decadal_pred_GEOS-5.

Model	10-year hindcasts/predictions	30-year hindcasts/predictions	Ensemble members	Other experiments	Time frequency
BCC-CSM1.1[1]	initialized in years 1960– 2005 (except years for 30 year runs)	initialized every 5 years 1960–2005	4	Hindcasts without volcanoes (noVolc) initialized 1960, 75, 80, 85, 90 Prediction with 2010 Pinatubo-like eruption (volcIn2010)	3 hr, 6 hr, daily, monthly
CCSM4[2]	initialized 1966, 71, 75, 76, 85, 86, 90, 91, 95, 96, 2000, 01, 02, 03, 04, 05	initialized 1961, 80, 81, 2006	10	–	monthly
CFSv2-2011[3]	initialized 1980, 81, 83, 85, 90, 93, 95, 96, 98, 2000, 03, 05, 06, 09, 10	–	4	–	3 hr, monthly

continued

Table 2. Continued.

Model	10-year hindcasts/predictions	30-year hindcasts/predictions	Ensemble members	Other experiments	Time frequency
CMCC-CM[3]	initialized 1965, 70, 75, 85, 90, 95	initialized 1960, 1980, 2005	3	–	3 hr, 6 hr, daily, monthly
CNRM-CM5[3]	initialized 1964, 65, 69, 70, 74, 75, 84, 85, 89, 90, 94, 95, 99, 2000	initialized 1959, 60,79, 80, 2004, 05	10	volcIn2010	daily, monthly
CanCM4[4]	initialized 1961–2011 (except years for 30 year runs)	initialized 1960, 80, 2005	10	–	6 hr, daily, monthly
EC-EARTH[3]	initialized 1965, 70, 85, 90, 95, 2000, 05	initialized 1960, 80	10	–	3 hr, 6 hr, daily, monthly
FGOALS-g2[3]	initialized 1965, 70, 75, 85, 90, 95, 2000	initialized 1960, 80, 2005	3	–	3 hr, 6 hr, daily, monthly
FGOALS-s2[3]	initialized 1965, 70, 75, 80, 85, 95, 2000	initialized 1960, 90, 2005	3	–	monthly
GEOS-5[5]	initialized 1960–2009 (except years 1981–1984)	–	3	–	monthly
GFDL-CM2.1[3]	initialized 1961–2012	–	10	–	monthly
HadCM3[3]	initialized 1961–2009 (except years for 30 year runs)	initialized 1960, 80, 2005	10	–	daily, monthly
IPSL-CM5A-LR[3]	initialized 1965, 70, 75, 85, 90, 95, 2000	initialized 1960, 80, 2005	6	–	3 hr, 6 hr, daily, monthly, yearly
MIROC4h[3]	initialized 1965, 70, 75, 85, 90, 95, 2000	initialized 1960, 80, 2005	6	–	3 hr, 6 hr, daily, monthly
MIROC5[3]	initialized 1959–2010 (except years for 30 year runs)	initialized 1960, 80, 2005	6	–	3 hr, 6 hr, daily, monthly
MPI-ESM-LR[3]	initialized 1960–2010	–	10	noVolc: 1960, 75, 80, 85, 90 volcIn2010	6 hr, daily, monthly
MPI-ESM-MR[3]	initialized 1960, 65, 70, 75, 80, 85, 90, 95, 2000, 01, 02, 03, 04, 05, 06, 07, 08, 09, 10	–	3	–	6 hr, daily, monthly
MRI-CGCM3[3]	initialized 1965, 70, 75, 85, 90, 95, 2000, 10, 11	initialized 1960, 80, 2005	9	–	3 hr, 6 hr, daily, monthly

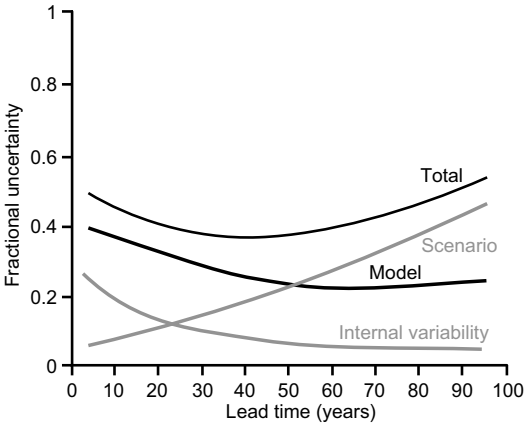


Fig. 4. The relative importance of different sources of uncertainty in decadal global mean surface temperature projections. Fractional uncertainty is the prediction uncertainty divided by the expected mean change of variable relative to 1971–2000. Adapted from Hawkins and Sutton (2009).

aerosol and solar cycle variability is prescribed using actual values for the past and assuming a climatological 11-year solar cycle and no volcanic eruptions in the future (Meehl *et al*. 2009). There are also few additional experiments that are hindcasts without volcanoes and predictions with the 2010 Pinatubo-like eruption.

Many of the forecasting centres that have produced decadal hindcasts for CMIP5 have also made experimental decadal predictions in real-time. There has been international activity to collect the predictions in a multi-model data set and results of these experimental decadal predictions are presented in Smith *et al*. (2013b). Predictions comprise 9 dynamical climate models and 3 empirical techniques. They are initialized in the year 2011 and made for 5-year periods 2012–2016 and 2016–2020. Predictions of AMV and PDV show no signal beyond climatology after the year 2015, but the Niño3 region is predicted to warm 0.5 °C during the coming decade. Results also show that initialized forecasts of globally averaged temperature are significantly cooler than uninitialized projections, consistent with the results of Meehl and Teng (2012). However, the global mean temperature is predicted to continue to rise with a 50% chance of every year after 2013 to exceed the current record. Uncertainties are still large for individual years and in most regions initialization has little impact after 4 years.

Challenges in decadal predictability and predictions

There are still many challenges and unsolved issues in decadal climate predictions and predictability. The uncertainties that are affecting climate projections for the 21st century are also affecting decadal climate predictions and predictability studies. These uncertainties arise from three sources: internal variability, model uncertainty and scenario uncertainty (Fig. 4). According to Hawkins and Sutton (2009), for lead times of the next few decades the main sources of uncertainty are internal variability and model uncertainty. For decadal time-scales and regional spatial scales the model uncertainty is more important. At longer lead times (more than 50 years) the emissions scenario uncertainty generally becomes dominant.

The main challenges have been already summed up in several studies and are now presented here following Meehl *et al*. (2009), Keenlyside and Ba (2010), Latif and Keenlyside (2011), Mehta *et al*. (2011) and Solomon *et al*. (2012):

i. Mechanisms of decadal climate variability are not well understood and the coherence among the climate models is limited. Interaction between natural and externally forced variability and sources of potential predictability should be identified.
ii. There are systematic errors in models that affect predictions and these errors should be identified, understood and corrected. Higher resolution is needed and parameterizations and coupling of additional climate subsystems should be improved.
iii. Lack of observations, especially in the ocean, are limiting forecast verification as well as development and testing of initialization and prediction systems. The instrumental record is short and properly covers only a few full cycles of decadal variability. It is crucial to maintain and enhance the existing observation systems.

iv. Long-lasting effects of solar and volcanic activity should be considered in decadal prediction models although they cannot be predicted.
v. Societal usefulness of decadal climate predictions should be elevated.

Despite all the challenges, some improvements have also occurred during the last decade such as the availability of enhanced ocean observations with the Argo array (www.argo.ucsd.edu). It is a global array of over 3000 free-drifting profiling floats that measure the temperature and salinity of the upper 2000 m of the ocean. This allows continuous *in situ* monitoring of temperature, salinity, and velocity of the upper ocean. It is likely that decadal climate predictions will benefit from the Argo data. On the other hand, the Argo data have only been available for approximately 10 years and it will take time until they can be used for verification of decadal predictions. However, there are already some results showing the benefits of the Argo data: for example in Zhang *et al*. (2007) AMOC is successfully reproduced only when the Argo observations are included in the simulations.

Summary and discussion

Decadal prediction is a new field of Earth science that is trying to bridge the gap between seasonal to inter-annual forecasting and climate change projections. Decadal predictions have a time perspective of 10–30 years into the future, which is a time-scale important for making societal adaptation decisions. This paper has reviewed the level of decadal predictability with emphasis on the Nordic region.

Understanding the mechanisms behind climate variability is important in making decadal predictions and assessing decadal predictability potential. Climate variability may be due to natural internal processes within the climate system, or to variations in natural or anthropogenic external forcing. The relative roles of internal and external processes in driving decadal and multi-decadal climate variations are not well understood and more precise quantification is needed. In addition to observations, modelling studies are often used for studying decadal scale climate fluctuations.

In the North Atlantic, there is strong decadal to multi-decadal variability. The most prominent internal mechanisms associated with this variability are the North Atlantic Oscillation and the Atlantic Meridional Overturning Circulation (AMOC). These internal mechanisms also affect climate variability in the Nordic region: a positive NAO index in winter and strong AMOC conditions are associated with higher surface temperatures and more precipitation than average. However, impacts of atmospheric circulation pattern may depend crucially on the local or regional details.

There is evidence that climate is predictable at decadal time-scales. Internal and externally forced variability are both important sources of potential predictability, but at a regional level, relative importance of these factors varies substantially (Boer 2010). Most of the decadal-predictability studies were carried out at the global scale, but decadal predictability is found predominately over mid- to high-latitude oceans (e.g. Boer and Lambert 2008). Over land predictability is usually low. The studies reviewed in this article indicate that potential of decadal predictability appears to be quite large in the North Atlantic sector and predictability is mainly based on the variations of AMOC. Some studies also indicate that AMOC predictability leads to some predictability of European climate (Collins and Sinha 2003, Sutton and Hodson 2005, Pohlmann *et al*. 2006). This potential decadal-scale predictability of AMOC might also contribute to predictability in the Nordic region, especially the coastal areas close to the North Atlantic (e.g. Koenigk *et al*. 2011, Persechino *et al*. 2013), but any robust conclusions cannot be made based on the current knowledge.

The prediction of climate at decadal time-scales is somewhere between seasonal and centennial scales and presents both an initial and boundary value problem. Initialization of climate models has been suggested to significantly increase the decadal prediction skill over the North Atlantic (e.g. Smith *et al*. 2007, Keenlyside *et al*. 2008). For example, a recent study of Doblas-Reyes *et al*. (2013) shows a significant skill in the North Atlantic for near-surface

temperature predictions up to 6–9 years. It has also been suggested that the skill of decadal prediction may depend on the initial state (e.g. Persechino *et al*. 2013) and AMOC seems to be more predictable if initialized from anomalously strong phase rather than from a weak phase. Based on the study results, it is not clear how initialisation affects the prediction skill in the Nordic region and there are still many open questions concerning initialization.

The main challenges associated with decadal predictions and predictability studies are poor understanding of mechanisms of decadal climate variability, systematic errors and need for improvements in models, lack of observations (especially in the ocean) and how to take into account the long-lasting effects of solar and volcanic activities. The societal usefulness of decadal predictions should also be elevated. Interactions between natural internal variability and anthropogenically-induced global warming are also important in assessing decadal predictability and making regional decadal predictions. According to Boer (2009) global warming may induce a general decrease in decadal potential predictability and the decrease seems to be largest over high latitude oceans. For example, weakening of AMOC is noted as a general result in warmer world simulations and this may have consequences also for climate in the Nordic region, although the response is still uncertain. Weakening of AMOC may result from decay of ice sheets and associated fresh-water flux (Vizcaíno *et al*. 2010) as well as reduced variations in sea surface temperature and salinity in Labrador sea (Koenigk *et al*. 2011). Natural decadal to multi-decadal climate variability can also mask anthropogenic climate change (Latif *et al*. 2006b).

Based on the current knowledge, there are still large uncertainties concerning decadal predictability in the Nordic region. However, closeness to the North Atlantic sector, which is the area of high potential decadal predictability, indicates that there might be some potential for making decadal predictions in this region.

Decadal predictions would offer valuable information for the society in making adaptation decisions. For example, the energy sector would benefit of guidance on climate variations in the following decades when making decisions on investments in energy production capacities. Climate conditions will also determine the usage of the carbon-intensive capacity, and therefore the decadal climate predictions are needed to produce the emission predictions from this sector.

As decadal predictability information in the Nordic region does not really exist or it is highly uncertain, there is a need for basic research in this field. For example data mining of the atmospheric data of the past millennium (CMIP5 and the COSMOS millennial-scale ESM simulations) could be conducted to find the predictable decadal climate signals for the Nordic region and to search for factors that are potentially related to predictable events. The topics for further research would also cover for example analysis of decadal climate variability and its relation to boreal biosphere as well as energy production and demand conditions.

Acknowledgements: The research has been funded by the Academy of Finland (project number 140771). This publication is contribution number 36 of the Nordic Centre of Excellence SVALI, Stability and Variations of Arctic Land Ice, funded by the Nordic Top-level Research Initiative. The authors thank an anonymous reviewer for the valuable comments that helped clarifying and improving the manuscript.

References

Alexandersson H., Schmith T., Iden K. & Tuomenvirta H. 1998. Long-term trend variations of the storm climate over NW Europe. *The Global Atmosphere and Ocean System*. 6: 97–120.

Bader J. & Latif M. 2003. The impact of decadal-scale Indian Ocean sea surface temperature anomalies on Sahelian rainfall and the North Atlantic Oscillation. *Geophys. Res. Lett*. 30, 2169, doi:10.1029/2003GL018426.

Bader J. & Latif M. 2005. North Atlantic Oscillation response to anomalous Indian Ocean SST in a coupled GCM. *J. Climate*. 18: 5382–5389.

Bauer E., Claussen M., Brovkin V. & Hünerbein A. 2003. Assessing climate forcings of the Earth system for the past millenium. *Geophys. Res. Lett*. 30: 1276–1279.

Beniston M. & Stephenson D.B. 2004. Extreme climatic events and their evolution under changing climatic conditions. *Global and Planetary Change*. 44: 1–9.

Blenckner T., Järvinen M. & Weyhenmeyer G.A. 2004. Atmospheric circulation and its impact on ice phenology in Scandinavia. *Boreal Env. Res*. 9: 371–380.

Boer G.J. 2004. Long-timescale potential predictability in an ensemble of coupled climate models. *Clim. Dyn*. 23: 29–44.

Boer G.J. 2009. Changes in interannual variability and decadal potential predictability under global warming. *J. Climate* 22: 3098–3109.

Boer G.J. 2010. Decadal potential predictability of twenty-first century climate. *Clim. Dyn.* 36: 1119–1133.

Boer G.J. & Lambert S.J. 2008. Multi-model decadal potential predictability of precipitation and temperature. *Geophys. Res. Lett.* 35, L05706, doi:10.1029/2008GL033234.

Branstator G., Teng H., Meehl G.A., Kimoto M., Knight J.R., Latif M. & Rosati A. 2012. Systematic estimates of initial-value decadal predictability for six AOGCMs. *J. Climate* 25: 1827–1846.

Collins M., Botzet A., Carril A.F., Drange H., Jouzeau A., Latif M., Masina S., Otteraa O.H., Pohlmann H., Sorteberg A., Sutton R. & Terray L. 2006. Interannual to decadal climate predictability in the North Atlantic: a multimodel-ensemble study. *J. Climate* 19: 1195–1203.

Collins M. & Sinha B. 2003. Predictability of decadal variations in the thermohaline circulation and climate. *Geophys. Res. Lett.* 30, 1306, doi:10.1029/2002GL016504.

Corti S., Molteni F. & Palmer T.N. 1999. Signature of recent climate change in frequencies of natural atmospheric circulation regimes. *Nature* 398: 799–802.

Curry R.G., McCartney M.S. & Joyce T.M. 1998. Oceanic transport of subpolar climate signals to mid-depth subtropical waters. *Nature*. 391: 575–577.

Delworth T.L., Manabe S. & Stouffer R.J. 1993. Interdecadal variations of the thermohaline circulation in a coupled ocean-atmosphere model. *J. Climate* 6: 1993–2011.

Delworth T.L. & Mann M.E. 2000. Observed and simulated multidecadal variability in the Northern Hemisphere. *Clim. Dyn.* 16: 661–676.

Doblas-Reyes F.J., Andreu-Burillo I., Chikamoto Y., García-Serrano J., Guemas V., Kimoto M., Mochizuki T., Rodrigues L.R.L. & van Oldenborgh G.J. 2013. Initialized near-term regional climate change prediction. *Nature communications*. 4, 1715, doi:10.1038/ncomms2704.

Dunstone N.J., Smith D.M., Booth B.B.B., Hermanson L. & Eade R. 2013. Anthropogenic aerosol forcing of Atlantic tropical storms. *Nature Geosci*. 6: 534–539.

Fraedrich K. & Müller K. 1992. Climate anomalies in Europe associated with ENSO extremes. *Int. J. Climatol.* 12: 25–31.

Gouirand I., Moberg A. & Zorita E. 2007. Climate variability in Scandinavia for the past millennium simulated by an atmosphere-ocean general circulation model. *Tellus* 59: 30–49.

Hagedorn R., Doblas-Reyes F.J. & Palmer T.N. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting — I. Basic concept. *Tellus* 57: 219–233.

Hasselmann K. 1976. Stochastic climate models. Part I: theory. *Tellus* 28: 473–485.

Hawkins E., Robson J., Sutton R., Smith D. & Keenlyside N. 2011. Evaluating the potential for statistical decadal predictions of sea surface temperatures with a perfect model approach. *Clim. Dyn.* 37: 2495–2509.

Hawkins E. & Sutton R. 2009. The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.* 90: 1095–1107.

Hazeleger W., Guemas V., Wouters B., Corti S., Andreu-Burillo I., Doblas-Reyes F.J., Wyser K. & Caian M. 2013. Multiyear climate predictions using two initialization strategies. *Geophys. Res. Lett.* 40: 1794–1798.

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

Hermanson L. & Sutton R.T. 2010. Case studies in interannual to decadal climate predictability. *Clim. Dyn.* 35: 1169–1189.

Hoerling M.P., Hurrell J.W. & Xu T.Y. 2001. Tropical origins for recent North Atlantic climate change. *Science* 292: 90–92.

Hurrell J.W. 1995. Decadal trends in the North-Atlantic-Oscillation regional temperatures and precipitation. *Science* 269: 676–679.

Hurrell J.W. & Deser C. 2009. North Atlantic climate variability: The role of the North Atlantic Oscillation. *J. Mar. Syst.* 78: 28–41.

Hurrell J.W., Kushnir Y., Ottersen G. & Visbeck M. 2003. An overview of the North Atlantic oscillation. In: Hurrell J.W., Kushnir Y., Ottersen G. & Visbeck M. (eds.), *The North Atlantic Oscillation: climate significance and environmental impact*, vol. 134, Geophysical Monograph Series, Washington DC: American Geophysical Union, pp. 1–35.

Hurrell J.W. & van Loon H. 1997. Decadal variations associated with the North Atlantic Oscillation. *Climatic Change* 36: 301–326.

Ineson S., Scaife A., Knight J., Manners J., Dunstone N., Gray L. & Haigh J. 2011. Solar forcing of winter climate variability in the northern hemisphere. *Nature Geosci*. 4: 753–757.

IPCC 2007. *Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Keenlyside N.S. & Ba J. 2010. Prospects for decadal climate prediction. *Wiley Interdisciplinary Reviews: Climate Change* 1: 627–635.

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

Koenigk T., Beatty C.K., Caian M., Döscher R. & Wyser K. 2011. Potential decadal predictability and its sensitivity to sea ice albedo parameterization in a global coupled model. *Clim. Dyn.* 38: 2389–2408.

Kushnir Y. 1994. Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Climate* 7: 141–157.

Latif M., Böning C., Willebrand J., Biastoch A., Dengg J., Keenlyside N., Schweckendiek U. & Madec G. 2006a. Is the thermohaline circulation changing? *J. Climate* 19: 4631–4637.

Latif M., Collins M., Pohlmann H. & Keenlyside N. 2006b. A review of predictability studies of Atlantic sector climate on decadal time scales. *J. Climate* 19: 5971–5987.

Latif M. & Keenlyside N.S. 2011. A perspective on decadal climate variability and predictability. *Deep Sea Research*

58: 1880–1894.

Latif M., Park W., Keenlyside N. & Ding H. 2009. Internal and external North Atlantic sector variability in the Kiel climate model. *Meteor. Zeitschrift* 18: 433–443.

Liu Z. & Alexander M. 2007. Atmospheric bridge, oceanic tunnel, and global climatic teleconnections. *Rev. Geophys*. 45: 1–34.

Matei D., Pohlmann H., Jungclaus J., Müller W., Haak H. & Marotzke J. 2012. Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model. *J. Climate* 25: 8502–8523.

Meehl G.A., Goddard L., Murphy J., Stouffer R.J., Boer G., Danabasoglu G., Dixon K., Giorgetta M.A., Greene A.M., Hawkins E., Hegerl G., Karoly D., Keenlyside N., Kimoto M., Kirtman B., Navarra A., Pulwarty R.S., Smith D., Stammer D. & Stockdale T. 2009. Decadal prediction: Can it be skillful? *Bull. Amer. Meteor. Soc*. 90: 1467–1485.

Meehl G.A. & Teng H. 2012. Case studies for initialized decadal hindcasts and predictions for the Pacific region. *Geophys. Res. Lett*. 39, *L22705, doi:*10.1029/2012GL053423.

Mehta V., Meehl G., Goddard L., Knight J., Kumar A., Latif M., Lee T., Rosati A. & Stammer D. 2011. Decadal climate predictability and prediction: where are we? *Bull. Amer. Meteor. Soc*. 92: 637–640.

Merkel U. & Latif M. 2002. A high-resolution AGCM study of the El Niño impact on the North Atlantic/European sector. *Geophys. Res. Lett*. 29, 1291, doi:10.1029/2001GL013726.

Msadek R., Dixon K.W., Delworth T.L. & Hurlin W. 2010. Assessing the predictability of the Atlantic meridional overturning circulation and associated fingerprints. *Geophys. Res. Lett*. 37, L19608, doi:10.1029/2010GL044517.

Okumura Y., Xie S.-P., Numaguti A. & Tanimoto Y. 2001. Tropical Atlantic air–sea interaction and its influence on the NAO. *Geophys. Res. Lett*. 28: 1507–1510.

Ortega P., Hawkins E. & Sutton R. 2011. Processes governing the predictability of the Atlantic meridional overturning circulation in a coupled GCM. *Clim. Dyn*. 37: 1771–1782.

Ortega P., Montoya M., Gonzáles-Rouco F., Mignot J. & Legutke S. 2012. Variability of the Atlantic meridional overturning circulation in the last millennium and two IPCC scenarios. *Clim. Dyn*. 38: 1925–1947.

Palmer T.N., Doblas-Reyes F.J., Weisheimer A., Shutts G.J., Berner J. & Murphy J.M. 2009. *Towards the probabilistic Earth-system model*. Available at http://arxiv.org/abs/0812.1074.

Palmer T.N., Andersen U., Cantelaube P., Davey M., Deque M., Doblas-Reyes F.J., Feddersen H., Graham R., Gualdi S., Gueremy J.-F., Hagedorn R., Hoshen M., Keenlyside N., Latif M., Lazar A., Maisonnave E., Marletto V., Morse A.P., Orfila B., Rogel P., Terres J.-M. & Thomson M.C. 2004. Development of a European multi-model ensemble system for seasonal to interannual prediction (DEMETER). *Bull. Amer. Meteorol. Soc*. 85: 853–872.

Park W. & Latif M. 2005. Ocean dynamics and the nature of air-sea interactions over the North Atlantic. *J. Climate*. 18: 982–995.

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

Persechino A., Mignot J., Swingedouw D., Labetoulle S. & Guilyardi E. 2013. Decadal predictability of the Atlantic meridional overturning circulation and climate in the IPSL-CM5A-LR model. *Clim. Dyn*. 40: 2359–2380.

Pohlmann H., Botzet M., Latif M., Roesch A., Wild M. & Tschuck P. 2004. Estimating the long-term predictability potential of a coupled AOGCM. *J. Climate*. 17: 4463–4472.

Pohlmann H. & Latif M. 2005. Atlantic versus Indo-Pacific influence on Atlantic–European climate. *Geophys. Res. Lett*. 32, L05707, doi:10.1029/2004GL021316.

Pohlmann H., Sienz F. & Latif M. 2006. Influence of the multidecadal Atlantic meridional overturning circulation variability on European climate. *J. Climate* 19: 6062–6067.

Pohlmann H., Jungclaus J., Koehl A., Stammer D. & Marotzke J. 2009. Initializing decadal climate predictions with the GECCO oceanic synthesis: effects on the North Atlantic. *J. Climate* 22: 3926–3938.

Rodwell M.J., Rowell D.P. & Folland C.K. 1999. Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature* 398: 320–323.

Serreze M.C., Carse F., Barry R.G. & Rogers J.C. 1997. Icelandic low activity, climatological features, linkages with the NAO, and relationships with recent changes in the northern hemisphere circulation. *J. Climate* 10: 453–464.

Shaffrey L. & Sutton R. 2006. Bjerknes compensation and the decadal variability of the energy transports in a coupled climate model. *J. Climate* 19: 1167–1181.

Shanahan T.M., Overpeck J.T., Anchukaitis K.J., Beck J.W., Cole J.E., Dettman D.L., Peck J.A., Scholz C.A. & King J.W. 2009. Atlantic Forcing of Persistent Drought in West Africa. *Science* 324: 377–380.

Smith D.M., Cusack S., Colman A.W., Folland C.K., Harris G.R. & Murphy J.M. 2007. Improved surface temperature prediction for the coming decade from a global climate model. *Science* 317: 796–799.

Smith D.M., Eade R. & Pohlmann H. 2013a. A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction. *Clim. Dyn*. 41: 3325–3338.

Smith D.M., Scaife A.A., Boer G.J., Caian M., Doblas-Reyes F.J., Guemas V., Hawkins E., Hazeleger W., Hermanson L., Ho C.K., Ishii M., Kharin V., Kimoto M., Kirtman B., Lean J., Matei D., Merryfield W.J., Muller W.A., Pohlmann H., Rosati A., Wouters B. & Wyser K. 2013b. Real-time multi-model decadal climate predictions. *Clim. Dyn*. 41: 2875–2888.

Solomon A., Goddard L., Kumar A., Carton J., Deser C., Fukumori I., Greene A.M., Hegerl G., Kirtman B., Kushnir Y., Newman M., Smith D., Vimont D.J., Delworth T., Meehl G.A. & Stockdale T. 2011. Distinguishing the roles of natural and anthropogenically forced decadal climate variability. *Bull. Amer. Meteor. Soc*. 92: 141–156.

Sutton R.W. & Hodson D.L.R. 2005. Atlantic Ocean forcing

of North American and European summer climate. *Science* 309: 115–118.

Taylor K.E., Stouffer R.J. & Meehl G.A. 2009. *A summary of the CMIP5 experiment design*. Available at http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf.

Tietäväinen H., Tuomenvirta H. & Venäläinen A. 2010. Annual and seasonal mean temperatures in Finland during the last 160 years based on gridded temperature data. *Int. J. Climatol.* 30: 2247–2256.

Trenberth K.E., Jones P.D., Ambenje P., Bojariu R., Easterling D., Klein Tank A., Parker D., Rahimzadeh F., Renwick J.A., Rusticucci M., Soden B. & Zhai P. 2007. Observations: Surface and atmospheric climate change. In: Solomon S., Qin D., Manning M., Chen Z., Marquis M., Averyt K.B., Tignor M. & Miller H.L. (eds.), *Climate Change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom, pp. 235–336.

Tulloch R. & Marshall J. 2012. Exploring mechanisms of variability and predictability of Atlantic meridional overturning circulation in two coupled climate models. *J. Climate* 25: 4067–4080.

Visbeck M., Chassignet E.P., Curry R.G., Delworth T.L., Dickson R.R. & Krahmann G. 2003. The ocean's response to North Atlantic Oscillation variability. In: Hurrell J.W., Kushnir Y., Ottersen G. & Visbeck M. (eds.), *The North Atlantic Oscillation, climatic significance and environmental impact*, AGU Geophysical Monograph Series 134, pp. 113–146.

Vizcaíno M., Mikolajewicz U., Jungclaus J. & Schurgers G. 2010. Climate modification by future ice sheet changes and consequences for ice sheet mass balance. *Clim. Dyn.* 34: 301–324.

Walter K. & Graf H.-F. 2002. On the changing nature of the regional connection between the North Atlantic Oscillation and sea surface temperature. *J. Geophys. Res.* 107, 4338, doi:10.1029/2001JD000850.

Wood R. 2008. Natural ups and downs. *Nature* 453: 43–45.

Vukcevic M.A. 2011. North Atlantic Oscillations I: An overview of the AMO (SST)–NAO data embedded relationship. Available at http://hal.archives-ouvertes.fr/docs/00/64/12/35/PDF/NorthAtlanticOscillations-I.pdf.

Zhang R. & Delworth T.L. 2006. Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophys. Res. Lett.* 33, L17712, doi:10.1029/2006GL026267.

Zhang R., Delworth T.L. & Held I.M. 2007. Can the Atlantic Ocean drive the observed multidecadal variability in Northern Hemisphere mean temperature? *Geophys. Res. Lett.* 34, L02709, doi:10.1029/2006GL028683.

Zhang S., Harrison M.J., Rosati A. & Wittenberg A. 2007. System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon. Wea. Rev.* 135: 3541–3564.

II

Tellus SERIES A DYNAMIC METEOROLOGY AND OCEANOGRAPHY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Random projections in reducing the dimensionality of climate simulation data

By TEIJA SEITOLA[1,2]*, VISA MIKKOLA[3], JOHAN SILEN[1] and HEIKKI JÄRVINEN[2], [1]*Finnish Meteorological Institute, Helsinki, Finland*; [2]*Department of Physics, University of Helsinki, Finland*; [3]*Department of Mathematics and Statistics, University of Helsinki, Finland*

(Manuscript received 24 June 2014; in final form 4 September 2014)

ABSTRACT

Random projection (RP) is a dimensionality reduction method that has been earlier applied to high-dimensional data sets, for instance, in image processing. This study presents experimental results of RP applied to simulated global surface temperature data. Principal component analysis (PCA) is utilised to analyse how RP preserves structures when the original data set is compressed down to 10% or 1% of its original volume. Our experiments show that, although information is naturally lost in RP, the main spatial patterns (the principal component loadings) and temporal signatures (spectra of the principal component scores) can nevertheless be recovered from the randomly projected low-dimensional subspaces. Our results imply that RP could be used as a pre-processing step before analysing the structure of high-dimensional climate data sets having many state variables, time steps and spatial locations.

Keywords: random projection, principal component analysis, dimensionality reduction, climate simulation data, El Niño – Southern Oscillation

1. Introduction

Climate simulation data are often high-dimensional, with thousands of time steps and grid points representing the state variables. High dimensionality is of course desirable, but it also presents a problem by making post-processing computations expensive and time-consuming. Data dimensionality reduction methods are therefore attractive, since they may enable the application of elaborate data analysis methods to otherwise prohibitively high-dimensional data sets.

Principal component analysis (PCA), also known as empirical orthogonal function (EOF) analysis (e.g. Rinne and Karhila, 1979; Von Storch and Zwiers, 1999), has been widely used in climate science in order to extract the dominant components of climate data time series. With large data sets, this method is computationally expensive, and rather soon becomes non-applicable unless the dimension of the original data set is significantly reduced. The use of time averaging, such as monthly or annual means instead of the original daily data, is an example of dimension reduction that sometimes enables PCAs use. This, however, significantly distorts the original information content of the data set: all temporal variability shorter than the averaging period is lost, and periods longer than the averaging period are affected. Thus, time averaging is not necessarily an optimal dimension reduction method.

This paper studies random projection (RP) as a dimensionality reduction method. It has been successfully applied in image processing (Bingham and Mannila, 2001; Goel et al., 2005; Qi and Hughes, 2012) and for text data (Bingham and Mannila, 2001). RPs fall into the theory of compressive sampling (CS), which has emerged as a novel paradigm in data sampling after the publications of Candès et al. (2006) and Donoho (2006). CS relies on the idea that most data have an inherent structure which can be viewed as sparsity. This means that, for example, a continuous signal in time may carry much less information than suggested by the difference between its upper and lower frequencies (Candès and Wakin, 2008; Bryan and Leise, 2013).

The aim of this paper is to introduce RP as a dimensionality reduction method in climate science. We will present the basic theory behind RP, and apply the method to climate data and show how the projected data preserve the essential structure of the original data. This is demonstrated by applying PCA to the original and randomly projected

*Corresponding author.
email: teija.seitola@fmi.fi
Responsible Editor: Abdel Hannachi, Stockholm University, Sweden.

Citation: Tellus A 2014, **66**, 25274, http://dx.doi.org/10.3402/tellusa.v66.25274

low-dimensional data sets to show that the leading principal components of the original data set can be recovered from the lower dimensional subspace. Section 2 presents the RP and PCA methods. In Section 3, we show some experimental results of applying RP and PCA to the original and dimensionality-reduced data sets. In addition, Section 4 demonstrates the application of the RP method to a very high-dimensional data set that represents multiple atmospheric model layers simultaneously.

2. Methods

2.1. Random projections

Random projection means that the $n \times d$ original data matrix ($\mathbf{X}$) of n d-dimensional observations is projected by a $d \times k$ random matrix ($\mathbf{R}$) (where $k < d$) to produce a lower dimensional subspace $\mathbf{P}$ of $n \times k$:

$$\mathbf{P}_{n\times k} = \mathbf{X}_{n\times d}\mathbf{R}_{d\times k} \tag{1}$$

In RP we are projecting our data set onto k random directions defined by the column vectors of $\mathbf{R}$. From these projections we can construct a lower dimensional representation of the original data set. The computational complexity of RP is of the order of $O(knd)$. Due to the simplicity of RP, involving only matrix multiplication, it can be applied to a wide range of data sets, even those with a very high number of dimensions. Figure 1 illustrates how the dimensionality of the data matrix is reduced by RP.

The idea of RPs stems from the Johnson–Lindenstrauss lemma (Johnson and Lindenstrauss, 1984):

Suppose we have an arbitrary matrix $\mathbf{X} \in \mathbb{R}^{n\times d}$. *Given any* $\varepsilon > 0$, *there is a mapping f:* $\mathbb{R}^d \to \mathbb{R}^k$, *for any* $k \geq O\frac{\log n}{\varepsilon^2}$, *such that, for any two rows* $\mathbf{x_i}$, $\mathbf{x_j} \in \mathbf{X}$, *we have*

$$(1-\varepsilon)||\mathbf{x_i} - \mathbf{x_j}||^2 \leq ||f(\mathbf{x_i}) - f(\mathbf{x_j})||^2 \leq (1+\varepsilon)||\mathbf{x_i} - \mathbf{x_j}||^2 \tag{2}$$

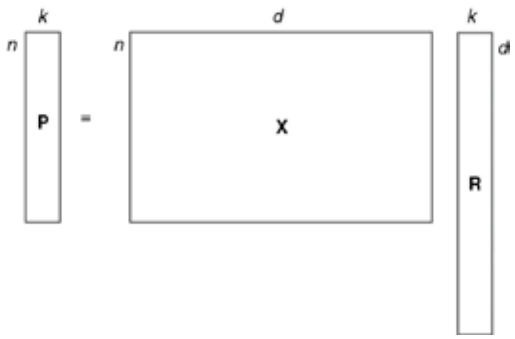


Fig. 1. Dimensionality reduction by random projection. Original data $\mathbf{X}$ is projected onto a random matrix $\mathbf{R}$ to have a lower dimensional subspace $\mathbf{P}$.

In the lemma it is stated that the data points in d-dimensional space can be embedded into a k-dimensional subspace in such a way that the pairwise Euclidean distances between the data points are approximately preserved with a factor of $1 \pm \varepsilon$. (See, e.g., Dasgupta and Gupta (2003) for proof of this result.)

Work has been done on finding suitable constructions of such mappings f (e.g. Frankl and Maehara, 1988; Achlioptas, 2003). In our experiments, we have employed a commonly used mapping ($\mathbf{R}$) which consists of the vectors of normally distributed $N(0,1)$ random numbers and the row vectors of the random matrix are scaled to have unit length. There are also other random distributions that satisfy the lemma [eq. (2)]. For example, Achlioptas (2003) has shown that a matrix of elements (r_{ij}) distributed as

$$r_{ij} = \sqrt{3} \times \begin{cases} +1 & \text{with probability } \frac{1}{6} \\ 0 & \text{with probability } \frac{2}{3} \\ -1 & \text{with probability } \frac{1}{6} \end{cases} \tag{3}$$

satisfies the requirements of a suitable mapping.

It should also be noted that in eq. (1) we are assuming an orthogonal projection, although the column vectors of $\mathbf{R}$ are not perfectly orthogonal. Here we can rely on a theorem of Hecht-Nielsen (1994) stating that as the dimension of the space increases, the number of *almost* orthogonal vectors increases. According to Bingham and Mannila (2001), the mean squared error between $\mathbf{R}\mathbf{R}^T$ and an identity matrix is about $1/k$ per matrix element. We can therefore assume that the vectors of $\mathbf{R}$ are sufficiently orthogonal for the projection to work. It is also possible to orthogonalise the vectors of $\mathbf{R}$, but it is computationally expensive.

We should also address the question of number of subdimensions (k) needed to get a representation of the original data set that is accurate enough. Some estimates can be found in the literature. Originally, Johnson and Lindenstrauss (1984) showed that the lower bound for k is of the order of $O(\log n/\varepsilon^2)$. There has also been some work on revealing an explicit formula for k. For example, Frankl and Maehara (1988) came up with the result that $k = \left\lceil 9(\varepsilon^2 - 2\varepsilon^3/3)^{-1} \log n \right\rceil + 1$ is sufficient to satisfy the Johnson–Lindenstrauss theorem, while Dasgupta and Gupta (2003) showed that $k \geq 4(\varepsilon^2/2 - \varepsilon^3/3)^{-1} \log n$ is enough. It is notable that the estimates of k depend only on the number of data points (observations) n, and are independent of d.

2.2. Principal component analysis

PCA is a widely used method to extract the dominant spatio-temporal signals from multidimensional data sets and to reduce the dimensionality of the data. In climate science, the principal component loadings are also known

as empirical orthogonal functions (e.g. Rinne and Karhila, 1979; Von Storch and Zwiers, 1999).

PCA is based on the idea of finding a basis to represent the original data set (Shlens, 2009). The aim is to find latent variables that explain most of the variance in the original data set via uncorrelated linear combinations of the original variables (Hannachi et al., 2007). This also enables dimensionality reduction, as most of the variance in the data set can be explained by only a small subset of principal components.

One of the techniques for finding the principal components of the data matrix is singular value decomposition (SVD). SVD is based on a theorem stating that any matrix $\mathbf{X}_{n\times d}$ can be broken down into orthogonal matrices $\mathbf{U}_{n\times n}$ and $\mathbf{V}_{d\times d}$ and a diagonal matrix $\mathbf{D}_{n\times d}$:

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T \tag{4}$$

where the columns of $\mathbf{V}$ are orthonormal eigenvectors of $\mathbf{C} = \mathbf{X}^T\mathbf{X}$ ($\mathbf{C}$ is the covariance matrix of $\mathbf{X}$), the columns of $\mathbf{U}$ are orthonormal eigenvectors of $\mathbf{Z} = \mathbf{X}\mathbf{X}^T$ and $\mathbf{D}$ is a diagonal matrix containing the square roots of the eigenvalues of $\mathbf{C}$ or $\mathbf{Z}$ in descending order. Since the column vectors of $\mathbf{V}$ are the eigenvectors of $\mathbf{C}$, SVD is a direct way of computing the PCA of the original data matrix $\mathbf{X}$. The column vectors of $\mathbf{V}$ are also known as the PC loading vectors, and the PC score matrix $\mathbf{S}$ can be calculated as follows:

$$\mathbf{S} = \mathbf{X}\mathbf{V} = \mathbf{U}\mathbf{D}\mathbf{V}^T\mathbf{V} = \mathbf{U}\mathbf{D} \tag{5}$$

As already mentioned, the PC loadings are also known as the EOFs (e.g. Rinne and Järvenoja, 1986) in which case the data set is often represented as a function of space (l) and time (t)

$$f(l,t) = fm(l) + \sum_{i=1}^{w} s_i(t)v_i(l) + \varepsilon(w,l,t) \tag{6}$$

where fm is a mean field, v_i is the spatial function of the i^{th} component (i.e. the PC loading) and the s_i are time-dependent coefficients associated with v_i. The number of EOFs is denoted by w. If the EOF series is truncated, that is, the data set is projected onto a subset of PC loadings, a residual term $\varepsilon(w,l,t)$ is included.

In PCA, it is generally recommended to use a mean-centred data matrix (Varmuza and Filzmoser, 2009). If the data matrix is not centred, then typically the PCs resulting from the PCA are not uncorrelated with each other and the eigenvalues do not indicate variance but rather the non-central second moments of the PCs (Cadima and Jolliffe, 2009). In uncentred PCA, it is often the case that the first eigenvector (PC loading) is close to the direction of the vector of column means of the data matrix.

The computational complexity of PCA (implemented by SVD) is of the order of $O(d^2n) + O(d^3)$, but there are also computationally less-expensive methods for finding only a certain number of eigenvalues and vectors (see e.g. Bingham and Mannila (2001) and references therein). The aim of this study is to compare the results of normal PCA (implemented by SVD) applied to the original and dimensionality-reduced (RP + PCA) data sets. The computational complexity of the latter can be expressed as $O(knd) + O(k^2n) + O(k^3)$. Now the original dimensions are reduced from d to k, which means computational savings in the PCA.

PCA has its own limitations in providing interpretability of the physical patterns. Because of spatial orthogonality and temporal uncorrelation, the PCs do not necessarily correspond to any physical phenomena or patterns (Demšar et al., 2013). The constraint in PCA for the successive components to explain the maximum remaining variance may lead to a mixing of physical phenomena in the extracted PCs (Aires et al., 2000). There are several methods to overcome these limitations, e.g. rotating the PC loadings. It has also been argued that the decorrelation assumption of PCA is not enough, and that the statistical independence of the extracted components is needed to analyse the dynamical complexity of physical phenomena (Aires et al., 2000). However, in this study we are more concerned with demonstrating the RP method with the aid of PCA, and therefore we only utilise the normal PCA without any rotations. The focus is more on the method than on the physical interpretation of the data.

3. Comparison of the original and the projected data

3.1. Data

A monthly surface temperature data set from a millennial full-forcing Earth system model simulation (Jungclaus, 2008) was used in this experiment. The original monthly archived simulation data set has 14 472 time steps, but we selected for our use only 4608 time steps (the dimension n) from the end of the data set. The simulation data set has a resolution of 96 points in longitude and 48 points in latitude, resulting in 4608 locations or grid points (the dimension d). The dimensions n and d were chosen to be of equal size so that they could be reduced with RP equivalently. The 4608×4608 data matrix is quite large, but it is still manageable when performing PCA on it for comparison with the projected lower dimensional subspaces. Surface temperature was chosen because it has some well-known global patterns (e.g. El Niño – Southern Oscillation, ENSO) that can be identified with PCA.

3.2. *Applying RP and PCA to the climate simulation data set*

RP was applied in two different ways: the original data matrix was arranged so that (1) the time steps n were in the rows and the spatial locations d (gridpoints) were in the columns and (2) the locations were in the rows and the time steps in the columns. In this way, it was possible to project the data matrix in order to correspondingly reduce either the spatial (case 1) or the temporal (case 2) dimension, since with RP we can only reduce one dimension at a time. The original data matrix was mean-centred before projection on the lower dimensional subspace.

When the PCA of the original data matrix is calculated, the PC loading vectors give us the spatial maps corresponding to the PC scores. The PC score vectors are the projections of the original data matrix onto the PC loadings, and the scores can be presented as time series. After the dimensionality of the data matrix is reduced by RP, we have then reduced either the temporal or the spatial dimension. Therefore it is not possible to get the corresponding PC scores and loadings when the other dimension has been reduced. Using SVD to find the PCs of the dimensionality-reduced data set $\mathbf{P}_{n\times k}$, where the spatial dimension d has been reduced, gives us

$$\mathbf{P}_{n\times k} = \mathbf{U}_{n\times n}\mathbf{D}_{n\times k}\mathbf{V}^{T}_{k\times k} \tag{7}$$

The loading vectors in $\mathbf{V}$ cannot be plotted on the original grid because we are now in $\mathbb{R}^k$ instead of $\mathbb{R}^d$ [see the Johnson–Lindenstrauss lemma; eq. (2)]. If the temporal dimension is reduced, we have $\mathbf{P}_{k\times d}$ and the score vectors cannot be presented as time series comparable to the original PC scores. However, in the Appendix we present a novel method whereby the loadings (or scores) can be approximated by calculating the matrices $\mathbf{U}$ (or $\mathbf{V}$) and $\mathbf{D}$ in the lower dimensional subspace and then multiplying these with the original data set. This method is applied in Section 4.

The number of subdimensions k needed for RP was discussed in Section 2.1. If we follow the bound given in Dasgupta and Gupta (2003), with an arbitrary value $\varepsilon = 0.2$ and $n = 4608$, the Johnson–Lindenstrauss theorem gives a limit of $k = 4(\varepsilon^2/2 - \varepsilon^3/3)^{-1}\log n \approx 1947$ (42% of the original dimensions) to make the projections with an accuracy of $1 \pm \varepsilon$. However, our experiments will show that, with our data set, a much smaller k still gives good results, recovering most of the information of the original data set. In this work, we are not looking for an exact lower bound for RPs applied to our data set but instead we are interested in demonstrating the method itself, keeping practical applications in mind. We therefore chose the dimensions for the RPs to be 10% and 1% of the original dimensions (4608). These percentages are equivalent to $k \approx 460$ (hereafter denoted as RP10%) and $k \approx 46$ (RP1%).

In order to investigate the stability of the results obtained by RPs, the original data matrix was projected onto 100 different realisations of RP matrices of the same k (where k is 46 or 460). For the uncertainty estimation, the original data matrix was arranged as in case 1. The PCA of each projection was calculated, making it possible to approximate the mean and the 95% confidence limits for the amount of variance explained by the PCs (Fig. 2). These confidence limits describe the uncertainties that arise from different projection matrices. From Fig. 2 we can see that the results can be somewhat different depending on what kind of RP matrix has been used. Some differences are to

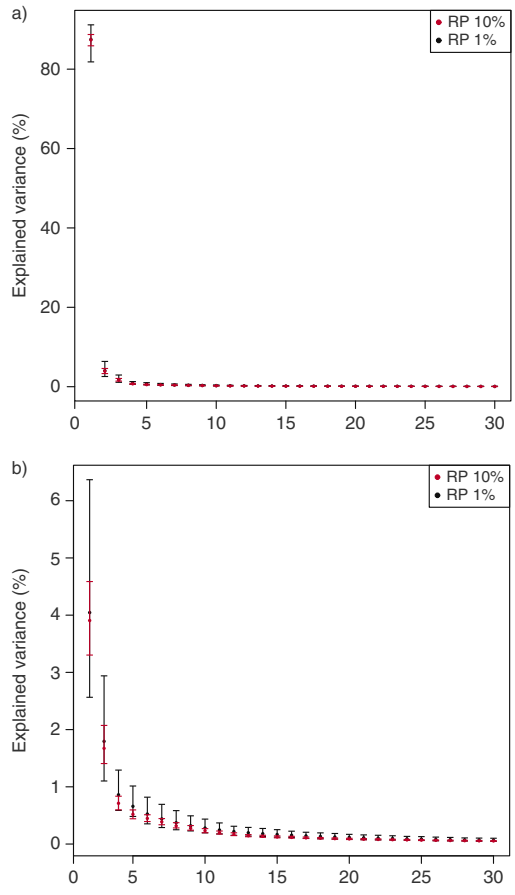


Fig. 2. Uncertainties of random projections. Mean and 95% confidence limits of the variance explained by the PCs (a) 1–30 and (b) 2–30 calculated from 100 realisations of projections of RP10% and RP1%. The explained variance of the first eigenvalue is excluded from subfigure (b) to show more details. In RP, the spatial dimension of the original data matrix is reduced.

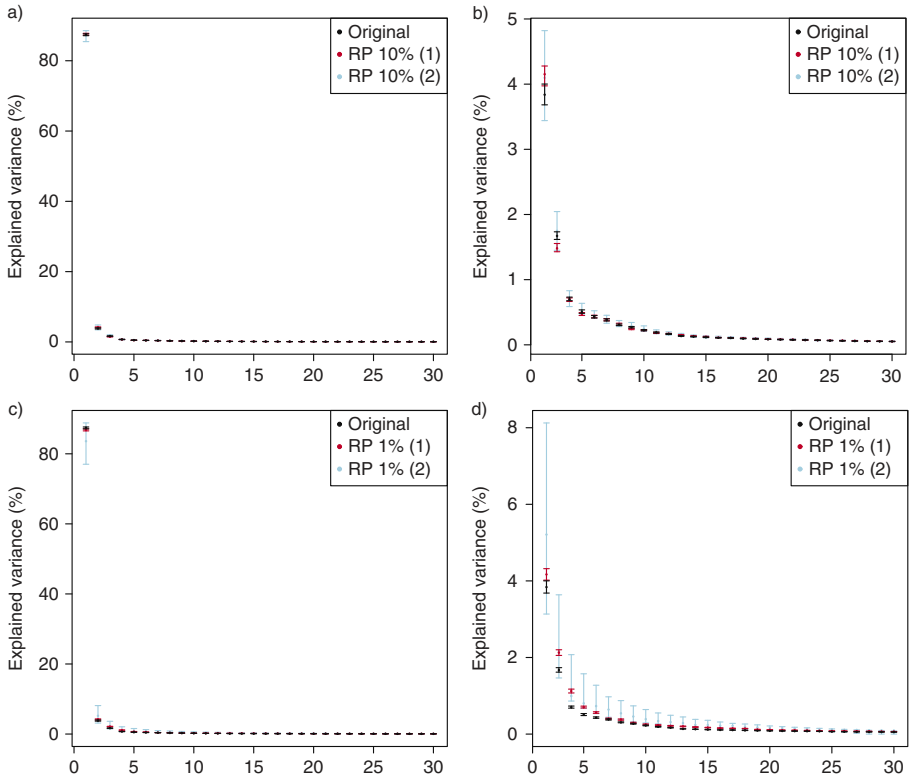


Fig. 3. Explained variance of the 30 first PCs with their 95% confidence limits. (a–b) Original and RP10%, (c–d) Original and RP1%. The explained variance of the first eigenvalue is excluded from subfigures (b) and (d) to show more details. In RP, the spatial (1) and temporal (2) dimensions are reduced. The confidence limits are obtained by re-sampling the original and projected data sets 100 times, and the PCA of each sample is calculated.

be expected, since the elements in the RP matrix are always different, although normally distributed $N(0,1)$. Furthermore, as the projection dimension k increases, the 95% confidence intervals (of the same k) become narrower.

3.3. *Results of PCA*

3.3.1. Explained variance of PCs. The eigenvalues of the data covariance matrix are in descending order and indicate the significance, that is, the amount of variance, of the principal components. An essential part of EOF studies (e.g. Hannachi, 2007) is to analyse the eigenvalues in the detection of the dominant signals or patterns in climate data.

Figure 3 shows the percentage of explained variance of the PCs with their 95% confidence limits approximated from the original and projected data sets (RP10% and RP1%). The confidence limits are based on bootstrapping where the original and projected data sets are re-sampled 100 times with replacement and the PCA of each bootstrap sample is calculated. The sampling is done with respect to the temporal dimension and the obtained samples are arranged in chronological order. In the case of the projected data sets, the variances of the PCs are obtained using one realisation of each projection (RP10% and RP1% and cases 1 and 2 of both). We have also re-sampled these realisations of projected data matrices to analyse the uncertainties related to these specific projections. Notice the difference to the previous section, where we estimated the uncertainties of RP due to regenerated random matrices. In Fig. 3, we can see that in case 2, in which the temporal dimension n is reduced, the 95% confidence intervals become wider, as can be expected. Otherwise the confidence intervals are quite narrow because of large n.

Figure 3 shows that the eigenvalues (illustrated as the percentage of the explained variance) decrease monotonically, and are quite similar in the cases of the original and projected data sets even when the dimension has been reduced to only 1% of the original dimensions. The eigenvalues of PCs 1–4 seem to be separated from the rest and also from each other, except in the case of RP1% with reduced temporal dimension n, where the 95% confidence limits of PCs 2 and 3, 3 and 4 as well as 4 and 5 overlap. The confidence limits of PCs 4 and 5 of RP10% with reduced n also overlap. PCs 1–4 explain almost 94% of the variance of the original and projected data sets. PC1 explains the majority (approximately 87%), PC2 4%, PC3 2% and PC4 approximately 1% of the variance. The rest of the eigenvalues decrease quite smoothly, which causes difficulties in distinguishing those small eigenvalues due to signal and those due to noise.

We saw that the eigenvalue spectra of the original and randomly projected data sets look quite similar. However, this only tells us that the amplitudes of the dominant signals are similar in both the original and the projected data sets. We also need to compare the PC loadings (i.e. the eigenvectors of the covariance matrix) and the PC scores to find out whether the spatio-temporal signatures have the same features.

3.3.2. PC loadings. The PC loadings, or the spatial patterns of the PCs of the original and dimensionality-reduced data sets, are shown in Figs. 4 and 5. Visual inspection shows that the original data and RP10% have very similar spatial patterns of PCs 1–12, with some differences however in PCs 8 and 9. RP1% PCs have mostly similar spatial patterns with the original PCs up to component 5, subsequent loadings of RP1% having more deviations. It should be noted that a PC loading vector has an arbitrary sign. To facilitate comparison, some of the RP10% and RP1% loading vectors were multiplied by -1 if they correlated negatively with the original PC loading vectors.

Spatial maps (especially PCs 4, 5, 6 and 11) show some features in surface temperature patterns that can be associated with the El Niño – Southern Oscillation (ENSO), e.g. distinct loadings in the Tropical Pacific and northwest/midwest North America (Trenberth and Caron, 2000). These same patterns can be found in the original, the RP10% and the RP1% maps and mostly in the same components.

The correlations of PC 1–20 loadings of the original and dimensionality-reduced (RP10% and RP1%) data sets are shown in Fig. 6. We can see that the RP10% loadings are strongly correlated with the original loadings until PC12 (correlation coefficient $r > 0.8$ and $r > 0.9$ up to PC7) and the RP1% loadings until PC5. PCs 1–5 already explain 94% of the variance of the data set and PCs 1–12 explain 96%. We can also see that some of the components of RP10%/RP1% have stronger correlations with adjacent ones of the original data set; for example, PC9 of RP10% has a stronger correlation with PC8 than with PC9 of the original data set. These adjacent components typically have similar variances.

Results are in line with the findings of Qi and Hughes (2012), where it is theoretically verified that, although RP disperses the energy of a PC in different directions, the original PC remains as the direction with the most energy. Due to this, oscillations with similar variance can be assigned to different, adjacent components, leading to some ambiguity in the indices. Another, or supplemental explanation for the switching of adjacent PCs is provided in Jolliffe (1989). According to that paper, it is a well-known fact that PCs whose variances (or eigenvalues) are nearly equal are unstable, but their joint subspace is stable. It has been shown that small changes in the variances in this subspace can lead to large changes in corresponding PC loading vectors, and this may lead to the switching of adjacent PCs. Thus it is more important to detect the same oscillations and patterns in the original and projected data sets, not in having them assigned to exactly the same components.

3.3.3. PC scores. The time series of PC scores were analysed with the Multitaper spectral analysis method (Thomson, 1982; Mann and Lees, 1996) to find the most powerful frequencies in these time series. The power spectra of the original and projected PC scores are shown in Fig. 7. Dominant features of the power spectra are the harmonic component frequencies which are integer multiples of the fundamental frequency. In the monthly surface temperature data set, the fundamental frequency is 1/12, which corresponds to a period of 1 yr and the harmonics clearly visible in the power spectra of PCs are 1/6, 1/4 and 1/3, corresponding to periods of 1/2, 1/3 and 1/4 yr that are related to intra-annual variations of surface temperature. The peaks at these frequencies are very similar in corresponding components of the original data, RP10% and RP1%. The peaks at the harmonics may also indicate that the orthogonality constraint of PCA is not suitable for this data set. The PCs are global and may have the same structure so that the first PC possesses the fundamental frequency while the following ones possess its harmonic frequencies (Aires et al., 2000).

Apart from the seasonal/harmonic frequencies, there are distinct peaks in the PC score spectra around the period of 3 yr. This might be related to ENSO which has a cycle of 2–6 yr. These peaks are clearly distinguishable in PCs 6 (original), 7 (RP10% and RP1%) and 11 (original and

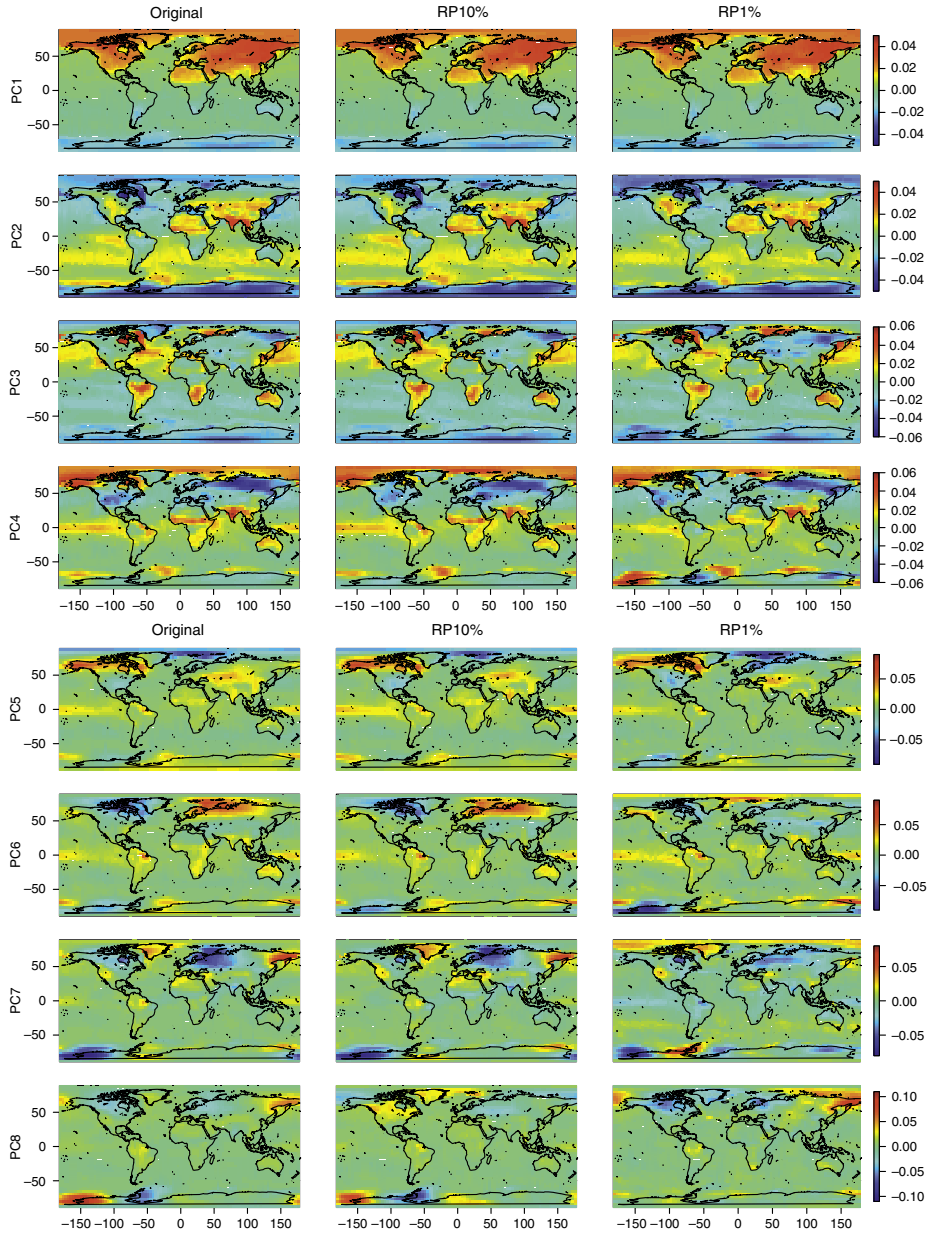


Fig. 4. Spatial patterns of PC1–PC8 loadings. Comparison of the original, RP10% and RP1% data sets. In RP, the temporal dimension is reduced.

RP1%) with some differences between the original and dimensionality-reduced data sets. We already identified some ENSO-related features in the spatial maps.

The correlations of the original and RP10%/RP1% PC scores (Fig. 8) are quite similar to the correlations in the loadings (Fig. 6). The RP10% correlations to the original

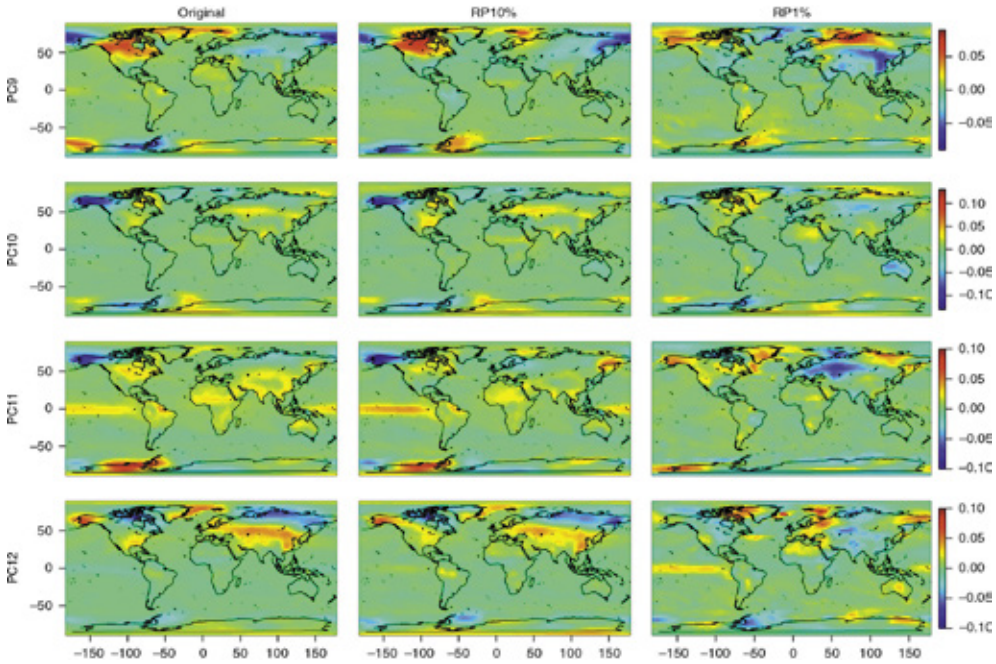


Fig. 5. Spatial patterns of PC9–PC12 loadings. Comparison of the original, RP10% and RP1% data sets. In RP, the temporal dimension is reduced.

scores are strong ($r > 0.8$) until about PC13, and PCs 6 and 7 are cross-correlated. The RP1% correlations with the original scores are also strong until PC5, although the correlation coefficient of PC5 is slightly less than 0.8.

4. Application of RP to a very high-dimensional data set

To demonstrate the application of the RP method to a very high-dimensional data set, we used a monthly temperature data set from a millennial full-forcing Earth system model simulation (Jungclaus, 2008) with a vertical resolution of 17 levels in the atmosphere between 1000 and 10 hPa. Inclusion of the vertical component increased the dimension d of the data matrix to $4608 \times 17 = 78\,336$. We extracted 3600 time steps (n) from the end of the data set. The increase of d from 4608 to 78 336 makes in our case PCA non-applicable (in a laptop computer), and thus we call the dimension 'very high'. Therefore the dimensionality of the data matrix was reduced by RP to make PCA applicable.

The original data matrix is $\mathbf{X}(n \times d)$ with $n = 3600$ and $d = 78\,336$, referring to time step and location, respectively. The dimensionality of the data matrix was reduced by projecting it onto a random matrix $\mathbf{R}(d \times k)$, where $k \approx 783$ is the subspace dimension (1% of the original dimensions d) [eq. (8)]. We then calculated the SVD of the lower dimensional data $\mathbf{P}(n \times k)$ to get the matrix $\mathbf{U}_{RP}(n \times k)$ [eq. (9)]. The PC loadings $\mathbf{V}(d \times k)$ were then approximated by multiplying the transpose of the original data matrix $\mathbf{X}(n \times d)$ with $\mathbf{U}_{RP}(n \times k)$ and the inverse of the diagonal matrix $\mathbf{D}_{RP}(k \times k)$ which we got from the SVD of $\mathbf{P}$ [eq. (10)] (see Appendix).

$$\mathbf{P} = \mathbf{XR} \tag{8}$$

$$\mathbf{P} = \mathbf{U}_{RP}\mathbf{D}_{RP}\mathbf{V}_{RP}^{T} \tag{9}$$

$$\mathbf{V} \approx \mathbf{X}^{T}\mathbf{U}_{RP}\mathbf{D}_{RP}^{-1} \tag{10}$$

The diagonal elements of $\mathbf{D}_{RP}(k \times k)$ are the square roots of the eigenvalues of the data covariance matrix indicating the significance of the PCs. Columns of $\mathbf{U}_{RP}(n \times k)$ multiplied by $\mathbf{D}_{RP}(k \times k)$ [see eq. (5)] are the PC scores: these are analysed with the Multitaper spectral analysis method (Thomson, 1982; Mann and Lees, 1996) as in Section 3. The columns of $\mathbf{V}(d \times k)$ are the PC loading vectors, that is, the spatial patterns corresponding to the PC scores. The elements of a loading vector contain the spatial patterns of a certain PC at 17 standard pressure levels of the atmosphere. The first 1–4608 elements correspond to level 1 (1000 hPa), elements 4609–9216 correspond to level 2 (925 hPa), and so on until 10 hPa.

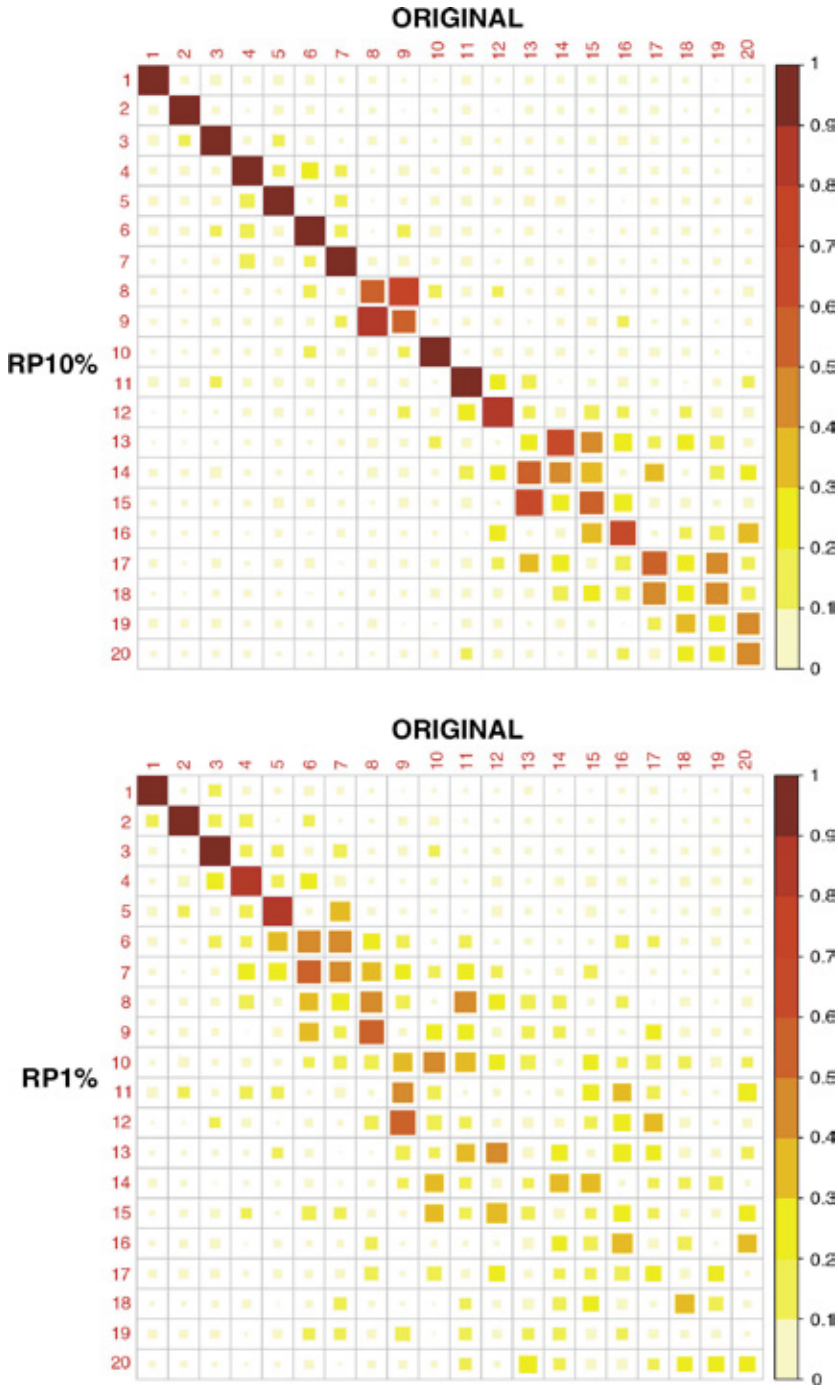


Fig. 6. Correlation of the original and projected (RP10% and RP1%) PC loadings. In RP, the temporal dimension is reduced.

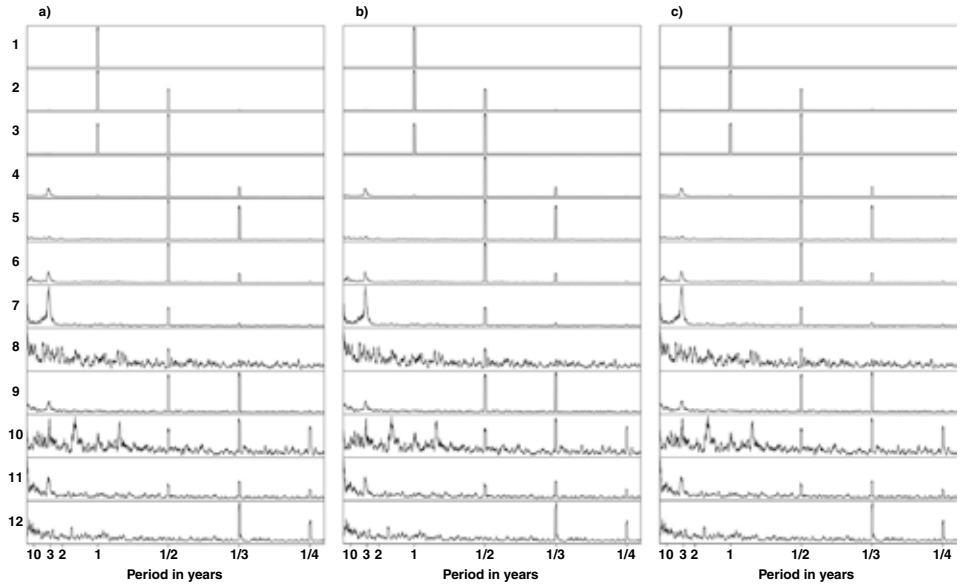


Fig. 7. Spectra of PC1–PC12 scores. (a) The original data set, (b) RP10% and (c) RP1%. In RP, the spatial dimension is reduced.

Figure 9 shows the percentage of the variance the PCs explain with their 95% confidence limits. The confidence limits were estimated by bootstrapping, as we did with the surface temperature data set in the previous section. PCs 1–3 are clearly separated from the rest and also from each other. PCs 4 and 5 (and maybe even PCs 6, 7 and 8 as their own subgroup) still seem to be distinguishable from the remaining eigenvalues which decrease quite smoothly. PC1 explains the majority of the variance in the data set (approximately 89%), PC2 explains 3.5%, PC3 approximately 1.5% and PCs 4 and 5 both explain approximately 0.7%. PCs 1–3 together account for 94% of the variance in the data set. The confidence intervals are narrow because of the relatively large sample size n.

Figure 10 shows that the dominant frequencies of the atmospheric temperature variation are those related to annual and intra-annual oscillations, which were also detected in the surface temperature data set in the previous section. There are also peaks in the PC score spectra around the period of 3 yr which might be related to ENSO. The most distinct ENSO-related component is PC5 and its spatial patterns at the 1000–30 hPa levels are shown in Figs. 11 and 12. At the lower atmospheric levels, especially 1000–925 hPa, temperature patterns related to ENSO can be identified in the Tropical Pacific and northwest/midwest North America. At the 850–600 hPa levels the positive loadings near the equator decrease but again increase at levels from 500 to 250 hPa and at the same time spread both north- and southwards, especially in the Pacific. The North American pattern attenuates little by little, but is still identifiable up to 400 hPa. At the upper levels the loadings around the tropics and subtropics become negative, meaning that the oscillation in the upper atmosphere is in an opposite phase compared to lower levels, where the pattern is clearly positive in the same areas.

Some caution is needed in the physical interpretation of these results. We already mentioned the limitations of PCA in Section 2.2. It should be noted that PC5 also has a distinct half-year peak, meaning that this component also carries an intra-annual signal. This is most likely to be related to the mixing problem of PCA. The ENSO representation of the model used in the simulations should also be considered (See, e.g., Jungclaus et al., 2006; Bellenger et al., 2014). Despite the limitations in the physical interpretation of the results, this experiment gives an example of how a large, multidimensional data set can be preprocessed with RP and then analysed efficiently to find, for example, the latent structures in the data set.

5. Summary and conclusions

The dimensionality of a simulated surface temperature data set was reduced by RP, and PCA was utilised to compare the structure of the original and projected data sets. Lower dimensional subspaces of 10% and 1% of the original data dimensions were investigated. The experiments showed that even at 1% of the original dimensions the main spatial and temporal patterns or principal components of the

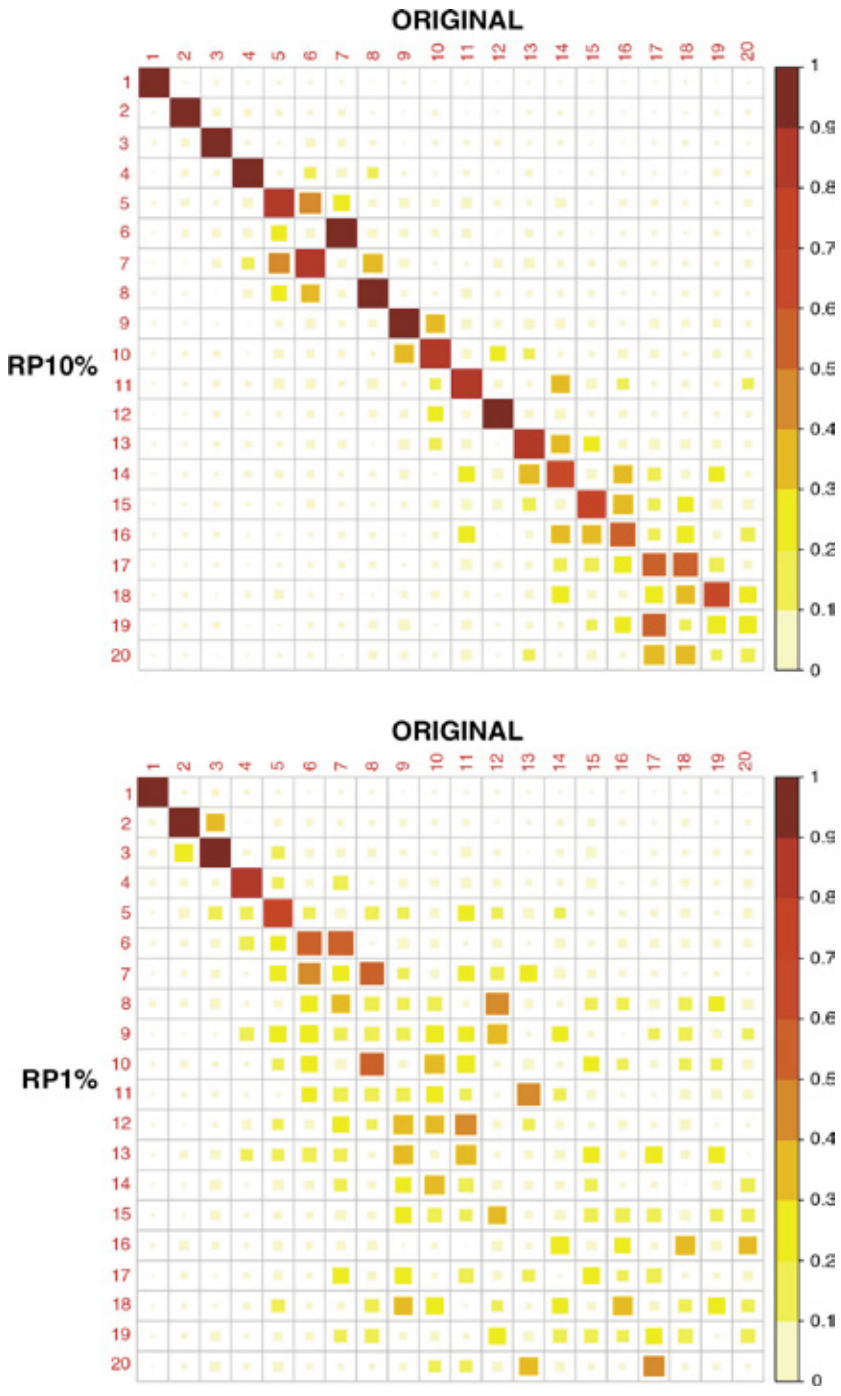


Fig. 8. Correlation of the original and projected (RP10% and RP1%) PC scores. In RP, the spatial dimension is reduced.

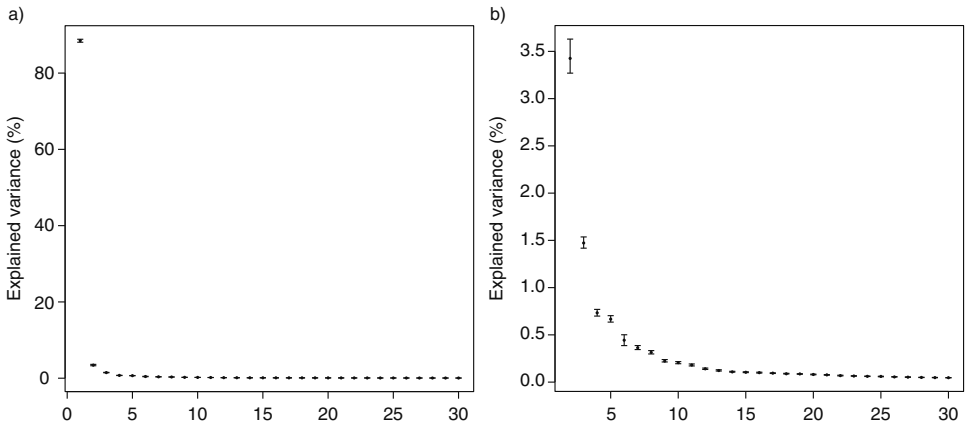


Fig. 9. Explained variance (%) of PCs with their 95% confidence limits estimated by bootstrapping. PCs (a) 1–30 and (b) 2–30 of the three-dimensional atmospheric temperature data set (the spatial dimension is reduced by RP) are shown. The explained variance of the first eigenvalue is excluded from subfigure (b) to show more details.

original surface temperature data set were approximately preserved. With a subspace of 10% of the original dimensions, we were able to recover the PCs explaining 96% of the variance in the original data set and with 1% we still could recover the PCs explaining 94% of the original variance.

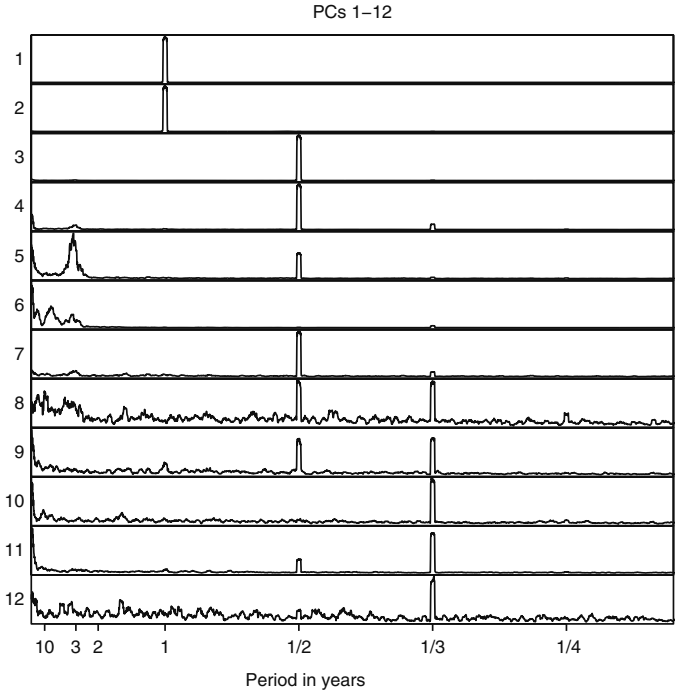


Fig. 10. Spectra of PC1–PC12 scores of the three-dimensional atmospheric temperature data set (the spatial dimension is reduced by RP).

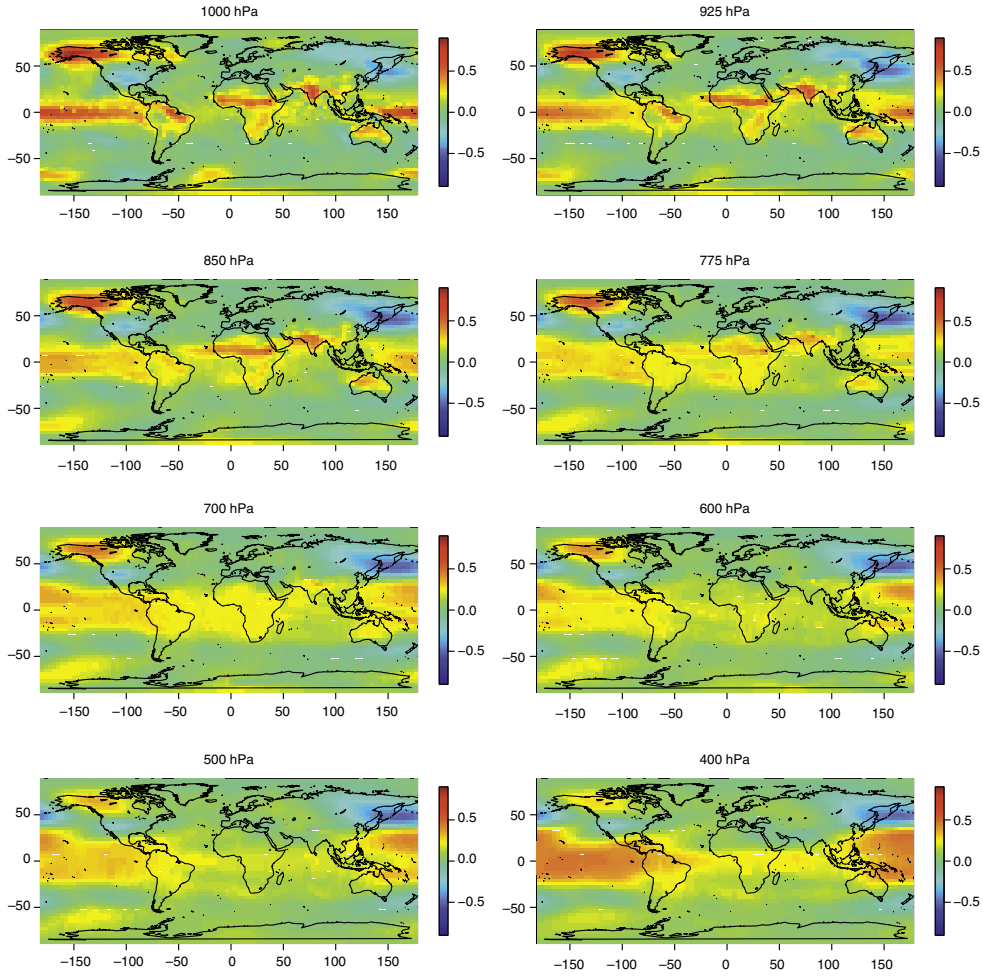


Fig. 11. The spatial patterns of the PC5 loadings of the atmospheric temperature data set (the spatial dimension is reduced by RP) between 1000 and 400 hPa. The spatial patterns are approximated using the method explained in the Appendix.

The findings of this work are supported by the results presented in Qi and Hughes (2012). In their paper, it is theoretically and experimentally shown that a normal PCA performed on low-dimensional random subspaces recovers the principal components of the original data set very well, and as the number of data samples n increases the principal components of the random subspace converge to the true original components.

RP is computationally fast compared to other methods for dimensionality reduction (e.g. PCA) since it involves only matrix multiplication. It can therefore be applied to very high-dimensional data sets. Based on our experiments, it seems to open new possibilities in reducing the dimensionality of climate data. One of the topics of our forthcoming research is to investigate the applicability of RP before the use of some other computationally heavy analysis methods for multivariate climate data, for example, multichannel singular spectrum analysis (e.g. Ghil et al., 2002).

As mentioned, there are some estimates available for the lowest bound for the reduced dimensions k. These estimates depend on the number of observations (dimension n) in the original data set and the desired accuracy of the projection (controlled by error ε). These estimated bounds seem to be much higher than the ones we used with good results. This suggests that the bounds for dimensionality reduction with RP should be investigated in more detail in

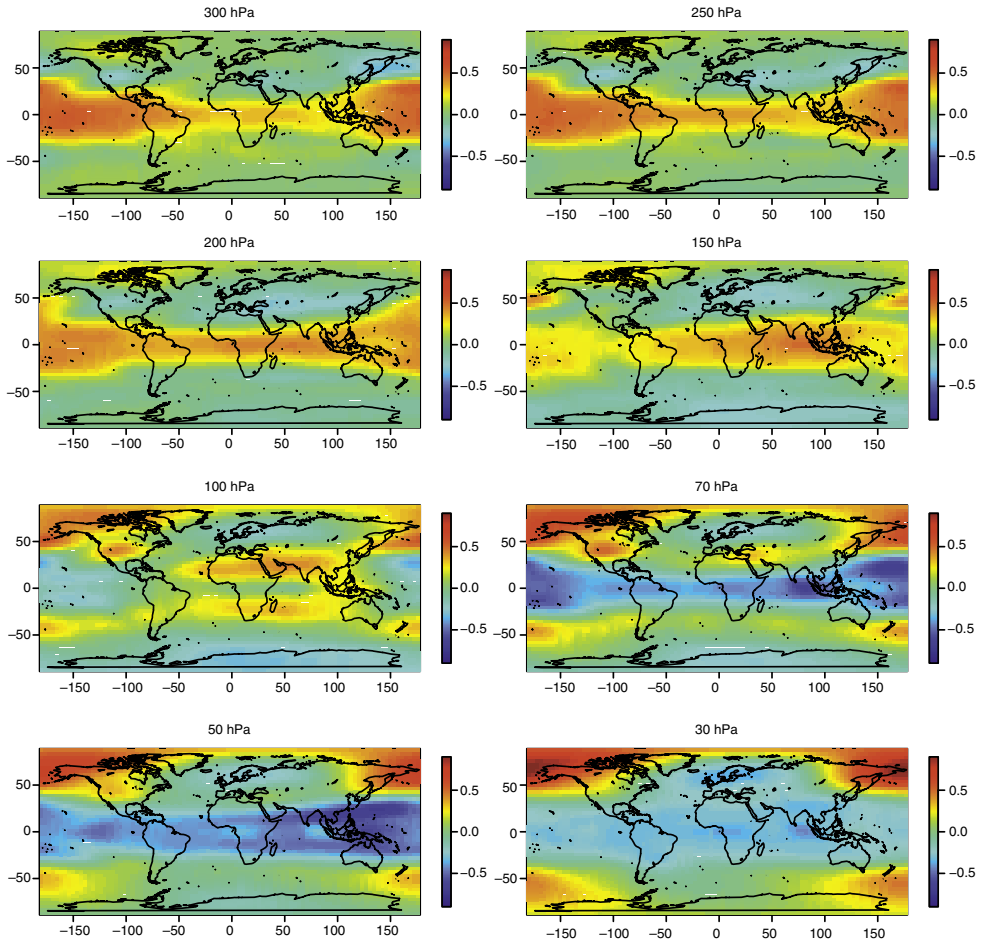


Fig. 12. The spatial patterns of the PC5 loadings of the atmospheric temperature data set (the spatial dimension is reduced by RP) between 300 and 30 hPa. The spatial patterns are approximated using the method explained in the Appendix.

the case of climate data. We would then also need to know what is the information content of our data set, that is, the signals that rise above the noise in the original data set.

We also demonstrated the application of the RP method to a very high-dimensional data set of the atmospheric temperature in three dimensions. Our results imply that RP could be used as a pre-processing step before analysing the structure of large data sets. This might allow an investigation of the dynamics of truly high-dimensional climate data sets of several state variables, time steps and spatial locations.

6. Acknowledgements

This research has been funded by the Academy of Finland (project number 140771). We want to thank the two anonymous reviewers for their valuable comments and suggestions that helped in substantially improving the manuscript.

7. Appendix

A.1. Random projection and the singular value decomposition

In this Appendix we explain the method used in Section 4. Let's say we have an original data $\mathbf{X}_{n\times d}$. The singular value decomposition (SVD) of $\mathbf{X}$ is:

$$\mathbf{X}_{n\times d} = \mathbf{U}_{n\times n}\mathbf{D}_{n\times d}\mathbf{V}^T_{d\times d} \tag{A1}$$

The covariance matrix of $\mathbf{X}$ is $\mathbf{C} = \mathbf{X}^T\mathbf{X}$ and the columns of $\mathbf{V}$ are the eigenvectors of $\mathbf{C}$. Also, the columns of $\mathbf{U}$

are the eigenvectors of $\mathbf{Z}=\mathbf{X}\mathbf{X}^{\boldsymbol{T}}$. $\mathbf{D}$ is a diagonal matrix containing the square roots of the eigenvalues of $\mathbf{C}$ or $\mathbf{Z}$ in descending order.

Since the random projection (RP) of $\mathbf{X}$ is $\mathbf{P}=\mathbf{X}\mathbf{R}$, where $\mathbf{R}_{d\times k}$ is the projection matrix, (the row vectors of $\mathbf{R}$ are scaled to have unit length), we can write:

$$\mathbf{C}_{RP}=(\mathbf{X}\mathbf{R})^T\mathbf{X}\mathbf{R}=\mathbf{R}^T\mathbf{X}^T\mathbf{X}\mathbf{R}=\mathbf{R}^T\mathbf{C}\mathbf{R} \tag{A2}$$

$$\mathbf{Z}_{RP}=\mathbf{X}\mathbf{R}(\mathbf{X}\mathbf{R})^T=\mathbf{X}\mathbf{R}\mathbf{R}^T\mathbf{X}^T\approx\mathbf{X}\mathbf{X}^T=\mathbf{Z} \tag{A3}$$

In the previous we have assumed that $\mathbf{R}\mathbf{R}^T\approx\mathbf{I}$, because the row vectors of $\mathbf{R}$ are nearly orthonormal. It is also possible to make the vectors of $\mathbf{R}$ strictly orthogonal, but this is computationally quite expensive.

Let's rewrite eq. (A1) as $\mathbf{X}_{n\times d}=\mathbf{U}_{n\times r}\mathbf{D}_{r\times r}\mathbf{V}^T_{r\times d}$, where r = rank($\mathbf{X}$). Now we can manipulate eq. (A1):

$$\begin{aligned}\mathbf{X}&=\mathbf{U}\mathbf{D}\mathbf{V}^T && (\mathbf{V}^T\mathbf{V}=\mathbf{I})\\ \mathbf{X}\mathbf{V}&=\mathbf{U}\mathbf{D} && (\mathbf{D}\mathbf{D}^{-1}=\mathbf{I})\\ \mathbf{U}&=\mathbf{X}\mathbf{V}\mathbf{D}^{-1}\end{aligned} \tag{A4}$$

or

$$\begin{aligned}\mathbf{X}&=\mathbf{U}\mathbf{D}\mathbf{V}^T && (\mathbf{U}^T\mathbf{U}=\mathbf{I})\\ \mathbf{U}^T\mathbf{X}&=\mathbf{D}\mathbf{V}^T && (\mathbf{D}^{-1}\mathbf{D}=\mathbf{I})\\ \mathbf{V}^T&=\mathbf{D}^{-1}\mathbf{U}^T\mathbf{X} && \text{transpose of both sides}\\ \mathbf{V}&=\mathbf{X}^T\mathbf{U}(\mathbf{D}^{-1})^T=\mathbf{X}^T\mathbf{U}\mathbf{D}^{-1}\end{aligned} \tag{A5}$$

Because $\mathbf{Z}\approx\mathbf{Z}_{RP}$, we can approximate

$$\begin{aligned}&\mathbf{U}\approx\mathbf{U}_{RP},\\ &\mathbf{D}\approx\mathbf{D}_{RP} && \text{and}\\ &\mathbf{V}\approx\mathbf{X}^T\mathbf{U}_{RP}\mathbf{D}^{-1}_{RP}\end{aligned} \tag{A6}$$

In the previous we have defined $\mathbf{U}_{RP}$ as $n\times k$ and $\mathbf{D}_{RP}$ as a $k\times k$ matrix, where k is the rank of matrix $\mathbf{P}_{n\times k}$.

If we have a very high-dimensional data set $\mathbf{X}$ we can first reduce the dimensionality of $\mathbf{X}$ by RP and then approximate $\mathbf{U}$ (or $\mathbf{V}$) and $\mathbf{D}$ in a lower dimensional subspace. We can then multiply the original data matrix with the approximated matrices $\mathbf{U}$ (or $\mathbf{V}$) and $\mathbf{D}$, finally getting the approximations of the PC scores or loadings depending on which dimension we have reduced in RP.

References

Achlioptas, D. 2003. Database-friendly random projections: Johnson–Lindenstrauss with binary coins. *J. Comput. System Sci.* **66**, 671–687.

Aires, F., Chedin, A. and Nadal, J. P. 2000. Independent component analysis of multivariate time series: application to tropical SST variability. *J. Geophys. Res.* **105**, 17437–17455.

Bellenger, H., Guilyardi, E., Leloup, J., Lengaigne, M. and Vialard, J. 2014. ENSO representation in climate models: from CMIP3 to CMIP5. *Clim. Dynam.* **42**, 1999–2018.

Bingham, E. and Mannila, H. 2001. Random projection in dimensionality reduction: applications to image and text data. In: *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '01*. ACM, New York, 245–250.

Bryan, K. and Leise, T. 2013. Making do with less: an introduction to compressed sensing. *SIAM Rev.* **55**, 547–566.

Cadima, J. and Jolliffe, I. 2009. On relationships between uncentered and column-centered principal component analysis. *Pak. J. Statist.* **25**, 473–503.

Candès, E. J., Romberg, J. and Tao, T. 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory.* **52**, 489–509.

Candès, E. J. and Wakin, M. B. 2008. An introduction to compressive sampling. *IEEE Signal Process. Mag.* **52**, 21–30.

Dasgupta, S. and Gupta, A. 2003. An elementary proof of a theorem of Johnson and Lindenstrauss. *Random. Struct. Algorithm.* **22**, 60–65.

Demšar, U., Harris, P., Brunsdon, C., Fotheringham, A. S. and McLoone, S. 2013. Principal component analysis on spatial data: an overview. *Ann. Assoc. Am. Geogr.* **103**, 106–128.

Donoho, D. 2006. Compressed sensing. *IEEE Trans. Inform. Theory.* **52**, 1289–1306.

Frankl, P. and Maehara, H. 1988. The Johnson–Lindenstrauss lemma and the sphericity of some graphs. *J. Combin. Theor. Series B.* **44**, 355–362.

Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K., Kondrashov, D. and co-authors. 2002. Advanced spectral methods for climatic time series. *Rev. Geophys.* **40**, 1–40.

Goel, N., Bebis, G. and Nefian, A. 2005. Face recognition experiments with random projection. In: *Defense and Security. International Society for Optics and Photonics*, **5779**, 426–437.

Hannachi, A. 2007. Pattern hunting in climate: a new method for finding trends in gridded climate data. *Int. J. Climatol.* **27**, 1–15.

Hannachi, A., Jolliffe, I. T. and Stephenson, D. B. 2007. Empirical orthogonal functions and related techniques in atmospheric science: a review. *Int. J. Climatol.* **27**, 1119–1152.

Hecht-Nielsen, R. 1994. Context vectors: general purpose approximate meaning representations self-organized from raw data. *Computational Intelligence: Imitating Life, IEEE Press*, 43–56.

Johnson, W. and Lindenstrauss, J. 1984. Extensions of Lipschitz mappings into a Hilbert space. In: *Conference in modern analysis and probability (New Haven, Conn., 1982)*, American Mathematical Society, Contemporary Mathematics, 26, 189–206.

Jolliffe, I. T. 1989. Rotation of Ill-defined principal components. *J. Roy. Stat. Soc. Series C (Applied Statistics)* **38**, 139–147.

Jungclaus, J. 2008. MPI-M earth system modelling framework: millennium full forcing experiment (ensemble member 1). World Data Center for Climate. CERA-DB "mil0010". Online at: http://cera-www.dkrz.de/WDCC/ui/Compact.jsp?acronym=mil0010.

Jungclaus, J. H., Keenlyside, N., Botzet, M., Haak, H., Luo, J.-J. and co-authors. 2006. Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM. *J. Clim.* **19**, 3952–3972.

Mann, M. E. and Lees, J. M. 1996. Robust estimation of background noise and signal detection in climatic time series. *Clim. Change*. **33**, 409–445.

Qi, H. and Hughes, S. M. 2012. Invariance of principal components under low-dimensional random projection of the data. In: *Proceedings of the 19th IEEE International Conference on Image Processing (ICIP)*, IEEE, 937–940.

Rinne, J. and Järvenoja, S. 1986. A rapid method of computing empirical orthogonal functions from a large dataset. *Mon. Weather. Rev*. **114**, 2571–2577.

Rinne, J. and Karhila, V. 1979. Empirical orthogonal functions of 500 mb height in the northern hemisphere determined from a large data sample. *Q. J. Roy. Meteorol. Soc*. **105**, 873–884.

Shlens, J. 2009. A tutorial on principal component analysis. Online at: http://www.snl.salk.edu/ ∼shlens/pca.pdf.

Thomson, D. J. 1982. Spectrum estimation and harmonic analysis. *Proc. IEEE*. **70**, 1055–1096.

Trenberth, K. and Caron, J. 2000. The Southern Oscillation revisited: sea level pressures, surface temperatures, and precipitation. *J. Clim*. **13**, 4358–4365.

Varmuza, K. and Filzmoser, P. 2009. *Introduction to Multivariate Statistical Analysis in Chemometrics*. CRC Press, Boca Raton.

Von Storch, H. and Zwiers, F. W. 1999. *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge.

III

Tellus SERIES A DYNAMIC METEOROLOGY AND OCEANOGRAPHY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Randomised multichannel singular spectrum analysis of the 20th century climate data

By TEIJA SEITOLA[1,2]*, JOHAN SILÉN[3] and HEIKKI JÄRVINEN[2], [1]*Climate Service Centre unit, Finnish Meteorological Institute, Helsinki, Finland*; [2]*Department of Physics, University of Helsinki, Helsinki, Finland*; [3]*Climate Research unit, Finnish Meteorological Institute, Helsinki, Finland*

(Manuscript received 24 June 2015; in final form 8 December 2015)

ABSTRACT

In this article, we introduce a new algorithm called randomised multichannel singular spectrum analysis (RMSSA), which is a generalisation of the traditional multichannel singular spectrum analysis (MSSA) into problems of arbitrarily large dimension. RMSSA consists of (1) a dimension reduction of the original data via random projections, (2) the standard MSSA step and (3) a recovery of the MSSA eigenmodes from the reduced space back to the original space. The RMSSA algorithm is presented in detail and additionally we show how to integrate it with a significance test based on a red noise null-hypothesis by Monte-Carlo simulation. Finally, RMSSA is applied to decompose the 20th century global monthly mean near-surface temperature variability into its low-frequency components. The decomposition of a reanalysis data set and two climate model simulations reveals, for instance, that the 2–6 yr variability centred in the Pacific Ocean is captured by all the data sets with some differences in statistical significance and spatial patterns.

Keywords: multichannel singular spectrum analysis, random projection, dimensionality reduction, El Niño – southern oscillation, 20th century reanalysis, HadGEM2-ES, MPI-ESM-MR

1. Introduction

Our motivation to focus on advanced spatio-temporal data analysis is to better understand the decadal climate variability in the Earth system and illuminate the capabilities of prediction tools to capture the associated signals (Meehl et al., 2014). Inter- and intra-decadal climate variability is inherent to the ocean–atmosphere system and is further coupled to other Earth system components, such as sea-ice and land surface (Meehl et al., 2009). The variability appears as complex four-dimensional (or spatio-temporal) structures in Earth system variables, such as wind, temperature and precipitation (Solomon et al., 2011).

These structures are embedded in extremely large-dimensional data sets gathered and generated in reanalysis of atmospheric and oceanic observations, and in massive simulation endeavours using Earth system models worldwide. Applicability of advanced data analysis tools is severely hampered by the very large dimensionality of the climate data.

Many common analysis methods, such as principal component analysis (PCA; Von Storch and Zwiers, 1999), involve eigen-problems, which become impossible to solve with increasing data dimension. Earlier we illustrated the use of random projections (RP) as a tool to tackle high-dimensional problems (Seitola et al., 2014). We demonstrated how PCA can be applied in three-dimensions to problems that are beyond practical computational limits without efficient dimension reduction. PCA is not an ideal tool, however, to extract and illustrate four-dimensional eigen-features in climate data. In this respect, the multichannel singular spectrum analysis (MSSA; Broomhead and King, 1986a, b) is a much more appealing method since the MSSA eigen-problem inherently contains the auto-covariance in the lagged copies of the original data vectors. The computational burden is, however, even larger than in PCA. We overcome this burden by a novel randomised version of MSSA, called RMSSA. To our knowledge, this approach has not been suggested before. We note that Oropeza and Sacchi (2011) incorporate a randomising operator into MSSA for noise attenuation in seismic data, but their algorithm is not aimed directly at large-dimensional problems. In RMSSA, RPs are used essentially to enable analysis of extremely large-dimensional data sets.

*Corresponding author.
email: teija.seitola@fmi.fi

Citation: Tellus A 2015, **67**, 28876, http://dx.doi.org/10.3402/tellusa.v67.28876

In this article, we present the RMSSA algorithm in detail and also include a test for the statistical significance of the results (Monte-Carlo MSSA; Allen and Robertson, 1996) in the algorithm. We demonstrate the use of RMSSA by decomposing the 20th century global monthly mean near-surface temperature variability into its low-frequency components. The data sources are described in Section 2.3.

2. Methods and Data

2.1. Multichannel singular spectrum analysis

MSSA was introduced into the study of dynamical systems by Broomhead and King (1986a, b). The method is equivalent to extended empirical orthogonal function (EEOF) analysis (Weare and Nasstrom, 1982), but there are differences in the choice of some important parameters and in the interpretation of the results (Plaut and Vautard, 1994).

In traditional PCA or EOF analysis (e.g. Rinne and Karhila, 1979), spatial correlations (in case of climatic data sets) are used in determining the patterns that explain most of the variability in the data set, but MSSA differs from this traditional method by also taking into account the temporal correlations. In other words, standard PCA decomposes a spatio-temporal field into spatial PC loading patterns (EOFs) and corresponding PC score time series (PCs), whereas MSSA also adds a temporal dimension to EOFs. MSSA PCs and EOFs are often called space-time PCs (ST-PCs) and space-time EOFs (ST-EOFs), and we have adopted this notation here. A more detailed description of MSSA is presented in Ghil et al. (2002) and in Appendix A.1 here.

2.1.1. Choice of the lag window. The idea of MSSA, in brief, is to find the patterns that maximise the lagged covariance of the data set $\mathbf{X}_{N\times L}$ within M lags. In case of a gridded climate data set, N represents the time steps and L is the number of grid points. The columns of the data matrix $\mathbf{X}$ are often called channels. The length of the lag window M is a user choice. For example, Elsner and Tsonis (1996) suggest that the results of MSSA do not change significantly with varying M as long as $N>>M$ and they recommend using $M=N/4$. Vautard and Ghil (1989) recommend to choose M no larger than approximately $N/3$. Clearly, if the number of channels L is large in the beginning, choosing large M would result in a very high-dimensional data matrix with $M\times L$ columns, including lagged copies of each channel in $\mathbf{X}$.

Determining the length of the lag window M is a trade-off between spectral resolution and statistical significance of the obtained components. The larger M is chosen, the more temporal information can be extracted but at the same time the variance is distributed on a larger set of components. If M is small, the statistical significance of the obtained components is enhanced. In this study, we used several values of M in order to test its effects on the results.

2.1.2. Assessing statistical significance with Monte-Carlo MSSA. ST-PCs/ST-EOFs often appear in pairs ('sinusoidal') that explain approximately the same variance and are $\pi/2$ out of phase with each other. These pairs are said to present stationary or propagating oscillatory modes of the data set (Plaut and Vautard, 1994). Modes with period less than or equal to M can be only presented by such pairs. However, existence of such a pair does not guarantee any physical oscillation, and according to Allen and Robertson (1996) such pairs can also be generated by non-oscillatory processes, such as first-order autoregressive noise.

This finding led Allen and Robertson (1996) to formulate a test for the statistical significance of MSSA components. The identified components are tested against a null-hypothesis of the data being generated by independent AR(1) processes (i.e. red noise) with the same variance and lag-1 autocorrelation as the original input time series. This procedure is called Monte-Carlo MSSA (MC-MSSA), and it is described in more detail in the original study of Allen and Robertson (1996) as well as in Appendix A.1 of this article.

2.1.3. Reconstructed components. ST-PCs cannot be compared to the original time series as such; instead, they can be represented in the original coordinate system by their reconstructed components, RCs (Plaut and Vautard, 1994; Ghil et al., 2002). In the reconstruction, the ST-PCs are projected back onto the eigenvectors (ST-EOFs) and each RC is a kind of filtered version of the original multivariate time series. Construction of RCs is illustrated in Fig. 1. Several ST-PCs/ST-EOFs can be used in the reconstructions, and if there is an oscillation that appears as a sinusoidal pair, both of these ST-PCs/ST-EOFs should be included in the reconstruction of that certain oscillatory mode. This is done by summing up the corresponding RCs. No information is lost in the reconstruction, and the original time series is a sum of all individual RCs.

2.2. Randomised algorithm for MSSA

As mentioned earlier, the computational burden of MSSA becomes soon prohibitively high if the original data set is high-dimensional and M is chosen to be large. This is typically the situation in studies of low-frequency variability in climate data sets. Traditionally, the dimensionality reduction has been obtained by calculating first a conventional PCA and retaining a set of dominant PCs for the MSSA

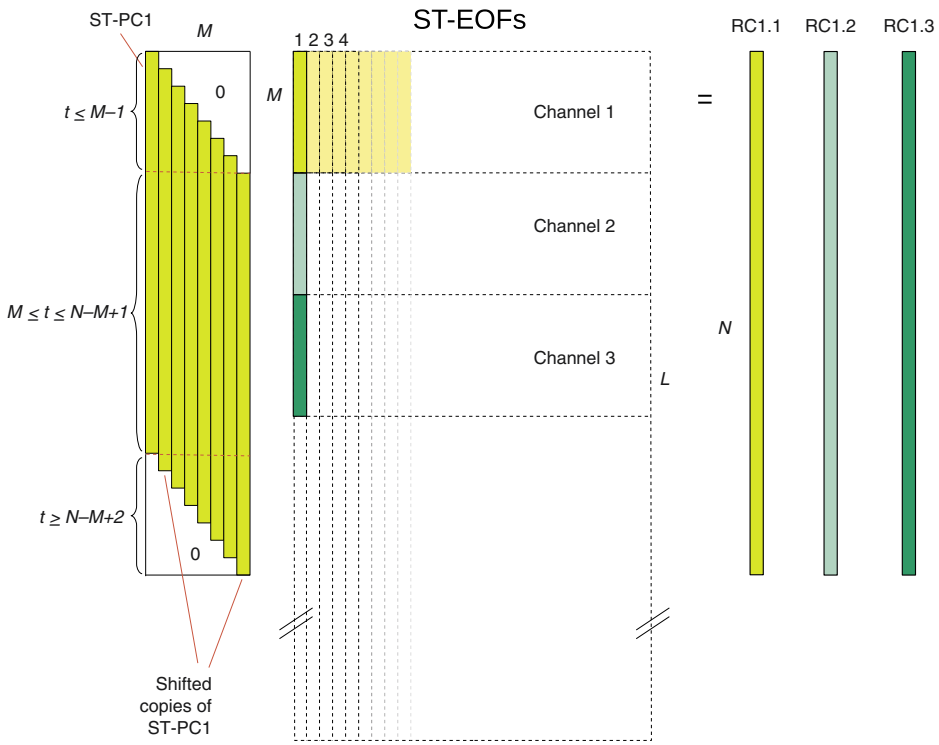


Fig. 1. Example calculation of the reconstructed components (RCs). A matrix of M shifted copies of a ST-PC (ST-PC1 in this example) is constructed to calculate reconstructions of that ST-PC in a time series of each channel (grid point). This matrix is then multiplied with that part of ST-EOF that corresponds to each channel. If $t \le M$, the elements of RC are divided by t, if $M \le t \le N-M+1$, divided by M, and if $t \ge N-M+2$, divided by $N-t+1$.

(see chapter 2.2.3). However, in this article we apply a different approach to dimensionality reduction. That is, we use RP to reduce the dimensionality of the original data set before performing MSSA.

In Halko et al. (2011), it is stated that randomised methods provide a powerful tool for constructing approximate matrix factorisations. Compared with standard deterministic algorithms, the randomised methods are often faster and more robust. Halko et al. (2011) present also numerical evidence that these algorithms succeed for real computational problems.

2.2.1. Description of RMSSA. In our approach, RP is applied to reduce the dimension of the original data matrix **X** after which the traditional MSSA calculation is performed in the lower-dimensional subspace. Finally, we reconstruct the ST-EOFs and RCs in the original space. We call this algorithm randomised multichannel singular spectrum analysis (RMSSA).

In RP, the original data set is projected onto a matrix **R** of Gaussian distributed (zero mean and unit variance) random numbers in order to construct a lower-dimensional representation **P** of the data set:

$$\mathbf{P}_{N\times k} = \mathbf{X}_{N\times L}\mathbf{R}_{L\times k} \tag{1}$$

In other words, we are projecting our data set onto k random directions determined by the column vectors of **R**. From these projections a lower-dimensional representation of the original data set can be constructed. Due to the simplicity of RP, involving only matrix multiplication, it can be applied to a wide range of data sets, even very high-dimensional ones.

RP has already been applied to climate data in Seitola et al. (2014) and it has been shown to preserve structures of the original data very well. In that article, the theoretical background of RP is presented in more detail with additional references.

The projected lower-dimensional data set **P** can be processed through MSSA where instead of original L channels we have now only k channels. This implies substantial computational savings (see Appendix A.2, algorithm 1). In the literature, there are some estimates of an appropriate value for k (e.g. Frankl and Maehara, 1988; Dasgupta and Gupta, 2003). However, these theoretical lower bounds for k are the worst case estimates and usually much lower values for k still give good results, retaining most of the information of the original data set (see e.g. Bingham and Mannila, 2001; Seitola et al., 2014). In practice, the value for k is usually chosen adaptively keeping the desired size for lower-dimensional approximation in mind.

A final step of the algorithm is to calculate the reconstructed components. This requires the recovery of the MSSA eigenmodes from the reduced space back to the original space, allowing the reconstruction of the original time series. This means that the eigenvectors (ST-EOFs) should be calculated in the original space instead of the reduced one. This part of the algorithm is also presented in Appendix A.2. Furthermore, in Appendix A.4 we explain how RP preserves the lagged covariance structure of the original data set.

2.2.2. Comparison of RMSSA to previous work. To our knowledge, the proposed RMSSA algorithm is unique. Some published work comes close to our approach but RMSSA has some important differences to the randomised MSSA algorithms used in seismic data processing (Oropeza and Sacchi, 2011; Chiu, 2013). The aim of Chiu (2013) was to introduce a new rank-based-reduction denoising algorithm to perform coherent and random noise filtering concurrently. Chiu (2013) named this algorithm, or rather filter, MSSARD (MSSA in the randomised domain). In MSSARD, the randomising operator randomly rearranges the order of the input data and reorganises the coherent noise into incoherent noise. The most important difference to our algorithm is in the randomising operator: In our case, we are using RP to reduce the dimensionality of the input data whereas Chiu's (2013) approach is to randomly rearrange the input data.

The technique of Oropeza and Sacchi (2011) was to embed a spatial data at a given temporal frequency into a block Hankel matrix after which a randomised singular value decomposition (SVD) was adopted to accelerate the rank reduction stage of the algorithm. Construction of a Hankel matrix corresponds to the construction of an augmented data matrix **A** in our algorithm (see Appendix A.1). Our algorithm is different in the sense that we apply RP on the original input data before construction of the augmented (or Hankel) matrix. This notably reduces the computational burden of MSSA because we are processing a much smaller data set already in the augmentation phase of the algorithm (see algorithm 1 in Appendix A.2).

In addition to these main differences, the above-mentioned seismological applications involve handling a data set where each time/frequency slice of spatial (x-y) data is processed separately through the algorithm. In our case, we are processing the whole time–longitude–latitude data set at once through the RMSSA algorithm.

2.2.3. Enhancing PC-based MSSA. In many studies, where MSSA is used as an analysis method (e.g. Plaut and Vautard, 1994; Moron et al., 2012), the dimension of the original data matrix has been reduced by calculating a conventional PCA of the original data matrix and then limiting MSSA into the dominant PCs. One has to bear in mind that the problem dimension may be prohibitive to contemplate solving even PCA, let alone MSSA. Nevertheless, the number of retained PCs is a somewhat arbitrary choice, but can be estimated by inspecting the eigenvalue spectrum and choosing the PCs that account for the majority of the variance and are separated from the rest of the spectrum. In geophysical datasets, however, the eigenvalue spectrum often decreases monotonically and it is difficult to distinguish the appropriate cut-off point. The aim of the study does also affect the choice of the PCs. For example, if the focus is on large-scale patterns, it might be more convenient to choose the low-frequency PCs for further analysis. Performing the calculations with different number of PCs and comparing the results can also help in finding the appropriate number of PCs. Importantly, RMSSA (Appendix A.2, algorithm 1) does not suffer from this problem because the lower-dimensional data set has essentially the same structure as the original high-dimensional data set.

PCA-based dimensionality reduction is, however, a preferred method if the oscillatory modes identified with MSSA are tested against a red noise null-hypothesis through Monte-Carlo simulation. According to Allen and Robertson (1996) the test is only useful if the channels in the data matrix are orthogonal or at least very nearly so. The PCs fulfil the orthogonality condition exactly. The randomised method can still accelerate – and in the case of a very-high-dimensional data set even enable – the calculation of the PCs (see Appendix A.2, algorithm 2).

This also raises the question as to whether the projected data set [i.e. matrix **P** in eq. (1)] could be used directly in MC-MSSA. Like the PCs, RP is also an orthogonal projection and the columns of **P** are also nearly orthogonal. However, this question is beyond the scope of this study and will not be discussed here any further.

2.3. Data

As an illustration of applying the RMSSA algorithm, we analysed the monthly mean near-surface air temperature

field from the 20th Century Reanalysis V2 data, hereafter 20CR, provided by the NOAA/OAR/ESRL PSD (Compo et al., 2011). In addition, we repeated the analysis for the historical 20th century simulations by Hadley Global Environment Model 2 – Earth System HadGEM2-ES (Collins et al., 2011), hereafter HadGEM2, and MPI Earth System Model (ESM) running on a medium resolution grid MPI-ESM-MR (Stevens et al., 2013), hereafter MPI-ESM. We extended the historical simulations (1901–2005) until 2012 using the rcp45 simulations. The historical and rcp45 simulations were extracted from the CMIP5 data archive and they follow the CMIP5 experimental protocol (Taylor et al., 2012). In the 20th century simulations, the historical record of climate forcing factors such as greenhouse gases, aerosols and natural forcings such as solar and volcanic changes is used. Rcp45 simulations follow the RCP4.5 greenhouse gas scenario. We used a single ensemble member of each model: r2i1p1 in case of HadGEM2 and r1i1p1 in case of MPI-ESM.

The 20CR data set is produced using an ensemble of perturbed reanalyses, and the final data set corresponds to the ensemble mean. Only surface pressure observations are assimilated, and the observed monthly sea-surface temperature and sea-ice distributions are used as boundary conditions to generate full three-dimensional estimates of the state of the troposphere (Compo et al., 2011). The 20CR data set is available from 1871 to 2012 but to be consistent with HadGEM2 and MPI-ESM simulations, the time sequence analysed here is 1901–2012 (1344 time steps). 20CR has ~ 2.0 degree horizontal resolution and we have used Gaussian gridded (192×94) data from 3-hour forecast values. HadGEM2 and MPI-ESM have both a global grid of 144×73 points. Thus, we have original data sets $\mathbf{X}_{N \times L}$ with $N = 1344$, $L = 18048$ (20CR) and $L = 10512$ (HadGEM2 and MPI-ESM).

As an illustrative example of the high-dimensionality of the MSSA problem, let's choose a lag window of $M = 240$ (months). In the case of the 20CR data set, this would result in an augmented matrix with $M \times L = 4331520$ columns. Clearly some kind of dimensionality reduction is needed in order to make the computations more efficient or even make them possible.

3. Results

3.1. Application of RMSSA to climatic data sets

In the previous section, we have introduced the RMSSA algorithm and the data sets to be analysed. Next we will proceed to the applications of the proposed method and discuss the results.

First, the original data sets were mean centred and RMSSA (algorithm 1 in Appendix A.2) was applied with $k = 500$. The first 1–30 ST-PCs of 20CR are shown in Fig. 2. In order to find the most powerful frequencies associated with the ST-PCs, the Multitaper spectral analysis method (Thomson, 1982; Mann and Lees, 1996) was applied. The power spectra of the ST-PCs are shown on the right in Fig. 2. The first pair of ST-PCs is clearly related to the annual cycle and this pair together explains the majority of the variance of the data set (almost 90%). The pairs of ST-PCs 3–4, 7–8 and 12–13 are the subharmonic frequencies of the annual cycle. The periods of ST-PCs 5, 6 and 11 as well as of ST-PCs 14, 17 and 18 fall outside the lag window length M and are the so-called trend components. ST-PC5 may be related to a centennial scale warming trend and ST-PC11 has a multi-decadal scale variability. ST-PCs 22 and 24 have clear spectral peaks on a 5–6 yr period and ST-PCs 29 and 30 are oscillating on a period of 3–4 yr. Those ST-PCs might be related to the El Niño-Southern Oscillation (ENSO) which is a prominent phenomenon on those time scales. ST-PCs 19–21 are related to a decadal scale variability, but the spectra of those components are quite broad on a 10–20 yr time scale.

The above analysis was also performed for the HadGEM2 and MPI-ESM data sets (figures not shown). As the annual cycle is too dominating in each data set, the analysis in the following sections will be repeated without the annual cycle. We also integrate a MC-MSSA step in the RMSSA algorithm (Appendix A.2, algorithm 2) in order to study the statistical significance of the obtained components.

3.1.1. Pre-processing the data for Monte-Carlo MSSA. Some pre-processing of the original data sets was crucial in order to assess the statistical significance of the low-frequency variability using MC-MSSA. First of all, the original data sets were standardised (i.e. the time series of each grid point was mean centred and divided by its standard deviation) in order to avoid overweighting the grid points with higher variance. Furthermore, the annual cycle of the time series of each grid point was estimated by STL (Loess based Seasonal-Trend Decomposition) and removed from the original data set. The STL method is a filtering procedure for decomposing a time series into trend, seasonal and remainder components. It includes some parameter choices controlling, for instance, how rapidly the trend and seasonal components can change. The method is described in detail in Cleveland et al. (1990) and we have followed their guidelines in choosing the related parameters. Without this procedure the annual cycle would dominate the results and starve the lower ranked MSSA components of power when tested against the red noise null-hypothesis. Linear trends were also fitted and removed from the data sets in order to avoid the dominance of the centennial scale trend.

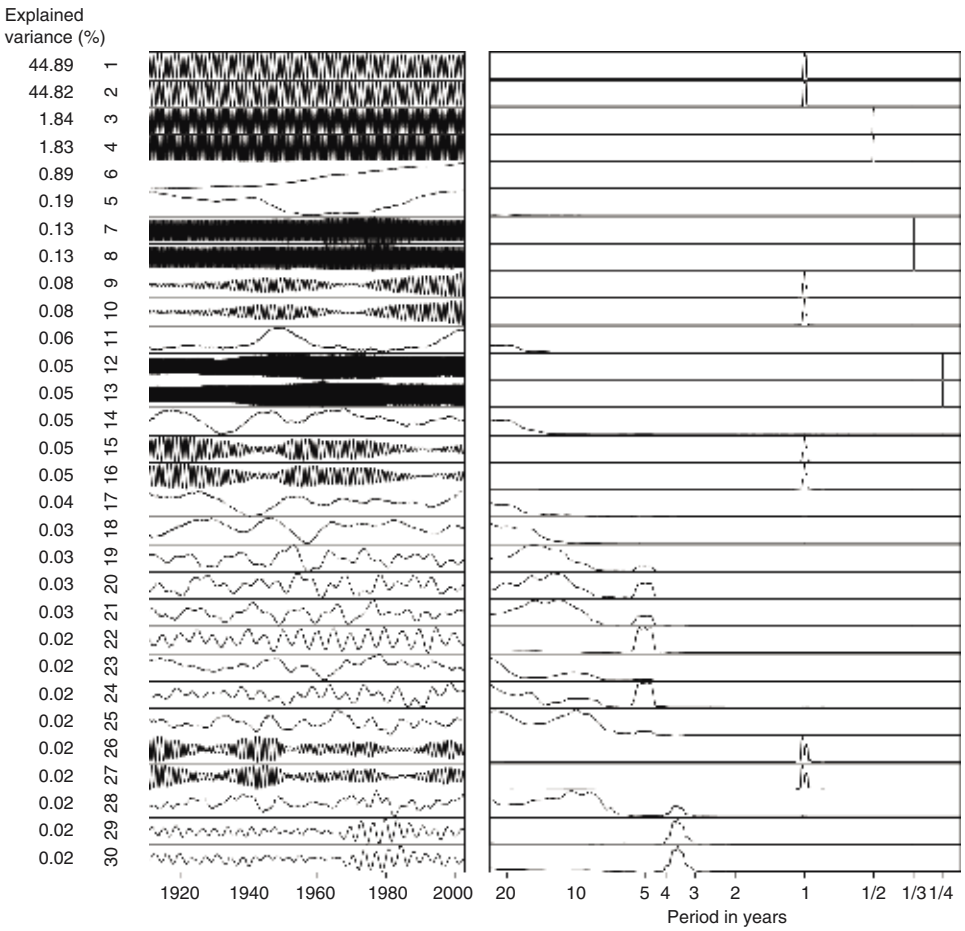


Fig. 2. ST-PCs 1–30 of 20CR monthly near-surface temperature 1901–2012 and their spectra. The lag window length M used in RMSSA is 20 yr (240 months). The data set is centred and algorithm 1 of Appendix A.2 is applied. The proportion of the variance explained by each component is also presented in the figure.

For the sake of comparison, the annual cycle was also estimated by calculating the mean values of each calendar month and those values were subtracted from the data to get monthly anomaly time series. However, determining the base for the anomaly calculation is not that straightforward and the choice of a base period may have severe impacts on the results (Kawale et al., 2011). Furthermore, the average annual cycle is only removed and if the annual cycle varies in the time series, the anomalies still contain a residual annual cycle.

The dimensions of the original data sets were reduced by applying RP with $k = 500$ to have a lower-dimensional approximation $\mathbf{P}_{N\times k}$ of each data set. To be able to perform MC-MSSA, we further calculated SVD of $\mathbf{P}$ and retained 30 first PCs of each data set, explaining approximately 72% (20CR), 67% (HadGEM2) and 64% (MPI-ESM) of the variance. Those 30 PCs were used as input channels in the MC-MSSA-step.

3.1.2. Decomposition of the pre-processed data sets. The ST-PCs 1–30 of each data set and their spectra are presented in Figs. 3–5. These figures show the results after applying the steps 1–8 of algorithm 2 in Appendix A.2 (note that the annual cycle and linear trend were removed from the original data sets). In 20CR (Fig. 3), the ST-PCs 1–2 are so-called trend components explaining together almost 9% of the variability of the data set. Pairs of

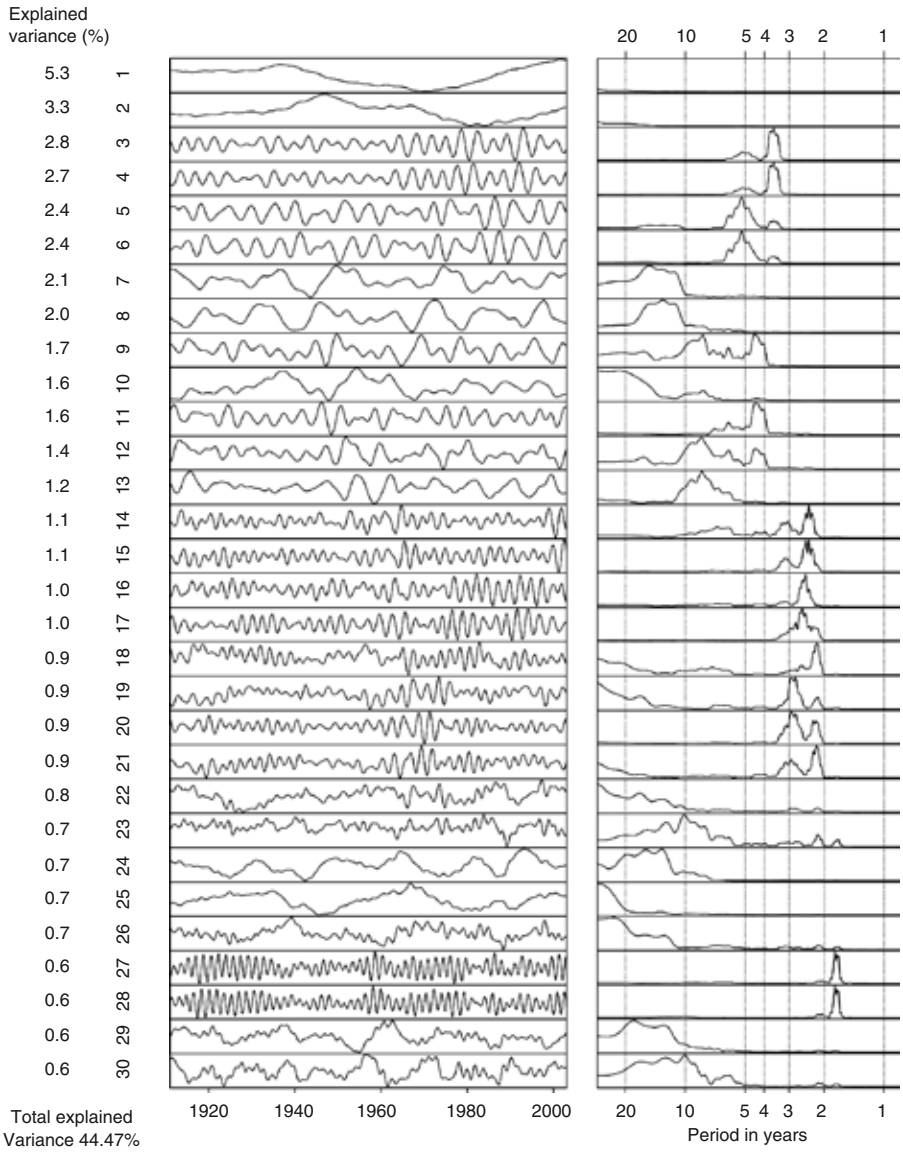


Fig. 3. ST-PCs 1–30 of 20CR monthly near-surface temperature 1901–2012 and their spectra. The lag window length M used in RMSSA is 20 yr (240 months). The annual cycle and linear trend are removed from the original data set and algorithm 2 of Appendix A.2 is applied. The proportion of the remaining variance explained by each component is also presented in the figure.

ST-PCs 3–4 and 5–6 in 20CR have clear peaks in frequencies corresponding to 3–4 yr and over 5 yr periods. In addition, 2–3 yr periodicities are distributed on several ST-PCs beginning from the 14th one. When the model simulations are compared to the 20CR components, the main differences are the prominent decadal scale components of HadGEM2 (ST-PCs 2–3, 9.3% of explained variance) and the 2–7 yr variability of MPI-ESM that is distributed on a large set of successive components. For more details, the readers are advised to study Figs. 3–5.

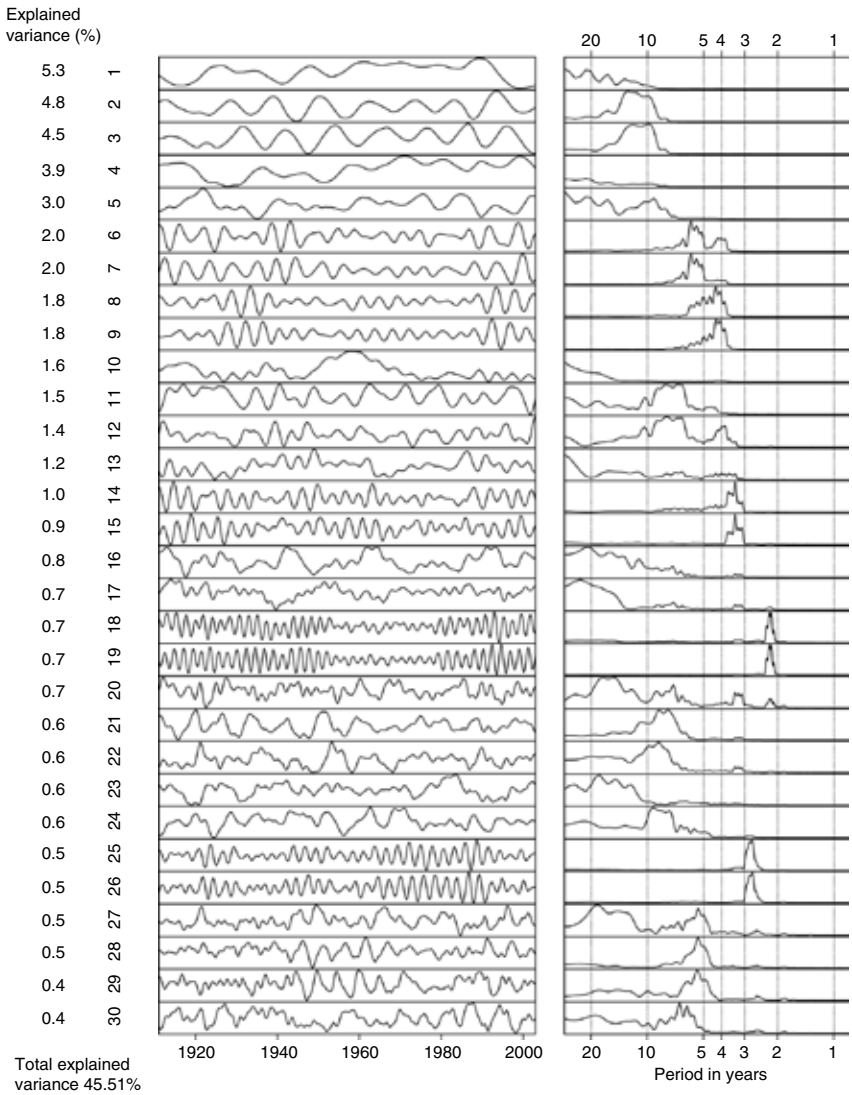


Fig. 4. ST-PCs 1–30 of HadGEM2 monthly near-surface temperature 1901–2012 and their spectra. The lag window length M used in RMSSA is 20 yr (240 months). The annual cycle and linear trend are removed from the original data set and algorithm 2 of Appendix A.2 is applied. The proportion of the remaining variance explained by each component is also presented in the figure.

3.2. Identifying significant oscillations

In MC-MSSA step, in total of 1000 realisations of red-noise surrogates were generated and the red-noise basis was used to estimate the 90, 95 and 99% confidence intervals for the eigenvalues generated by the noise model that consists of independent first-order autoregressive processes. Figure 6 shows the results of the Monte-Carlo significance test of 20CR, HadGEM2 and MPI-ESM with a 20 yr lag window ($M = 240$ months). In that figure, the data eigenvalues and 2.5th and 97.5th percentiles of the distribution of the surrogate eigenvalues are plotted against the dominant frequencies of the corresponding red-noise basis vectors (noise ST-EOFs). The dominant frequencies are estimated using fast

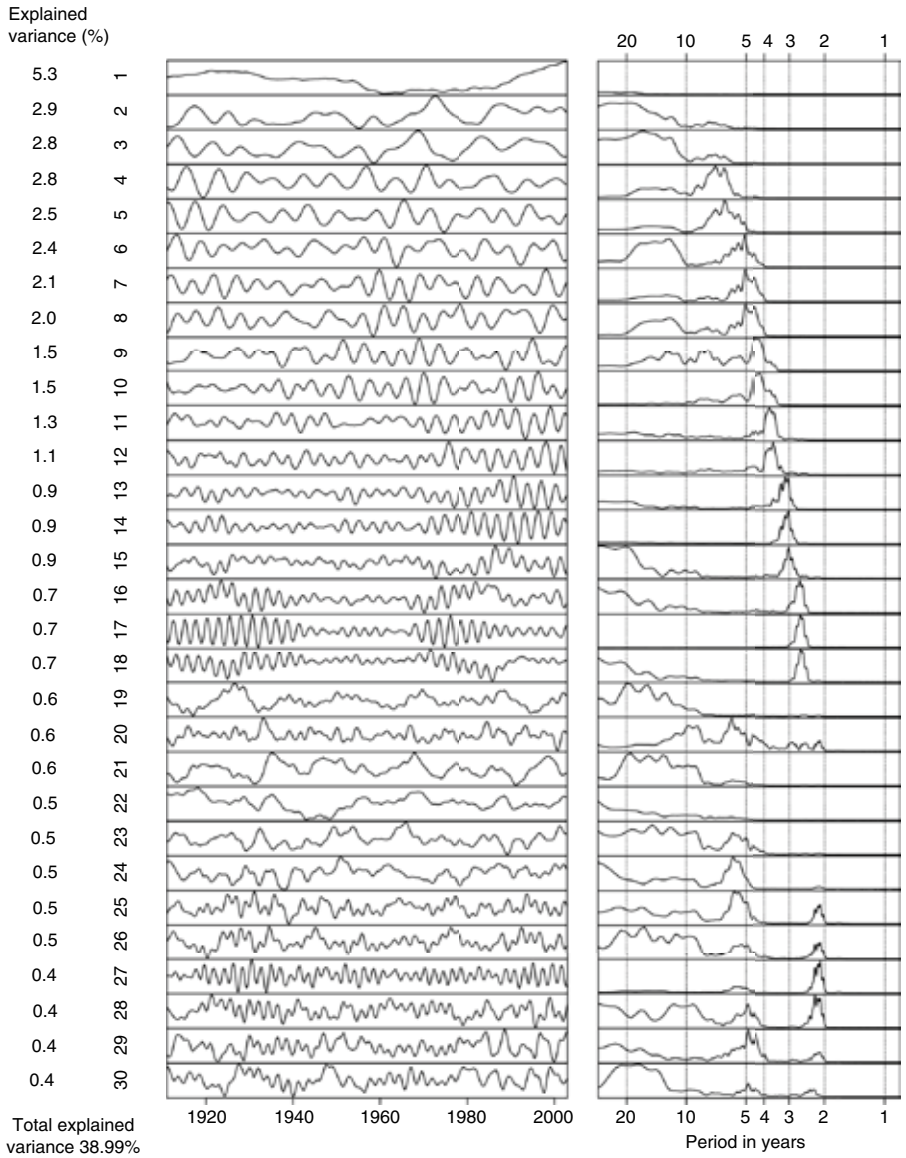


Fig. 5. ST-PCs 1–30 of MPI-ESM monthly near-surface temperature 1901–2012 and their spectra. The lag window length M used in RMSSA is 20 yr (240 months). The annual cycle and linear trend are removed from the original data set and algorithm 2 of Appendix A.2 is applied. The proportion of the remaining variance explained by each component is also presented in the figure.

Fourier transform (FFT). It should be noted, that the estimate of the dominant frequency of the noise ST-EOFs may not be exactly the same as the dominant frequency of the data ST-EOFs which may cause some small inaccuracies in the results.

The significant signals (at 5% significance level) in Fig. 6 are those whose data eigenvalues lie above the 97.5th percentiles of the surrogate eigenvalues. According to the test these signals have more variance than would be expected to have from a noise process. According to Plaut and

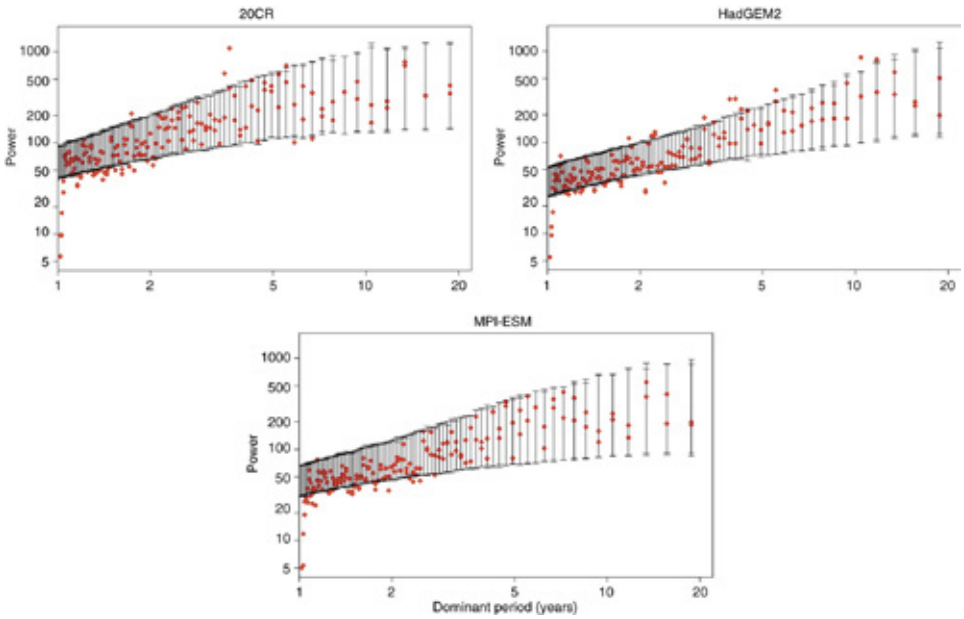


Fig. 6. MC-MSSA test of the monthly near-surface temperature variability in 20CR, HadGEM2 and MPI-ESM data sets 1901–2012. PCs 1–30 of RP + PCA (see Appendix A.2, algorithm 2) are used as input channels in the analysis and the lag window length M is 20 yr (240 months). In MC-MSSA, the red-noise basis is used. Red squares show the data eigenvalues plotted against the dominant frequency of the ST-PC corresponding to each eigenvalue. The vertical bars show the 2.5th and 97.5th percentiles of the eigenvalue distribution calculated from 1000 realisations of the red-noise surrogates. The ST-PCs that correspond to eigenvalues rising above the 97.5th percentiles are considered significant at the 5% level. Note the missing power at 1 yr due to the removal of the annual cycle.

Vautard (1994) the use of a lag window length M typically allows the distinction of oscillations with periods in the range $[M/5, M]$. Therefore we only show the significance test of the periodicities that are covered by the 20 yr lag window used in this example. From Fig. 6, we can see that in 20CR data set there are some significant periodicities (at 5% level) between 1.7 and 5.5 yr. HadGEM2 has somewhat more significant periodicities compared to 20CR, especially on 10 yr time scales, but MPI-ESM has hardly any eigenvalues lying above the 97.5th percentile.

3.2.1. Results with different lag window lengths. As noted earlier, the Monte-Carlo simulations were performed with varying lag window M to estimate its effect on the statistical significance of the oscillations. Spectral resolution increases with lag window length and oscillatory pairs with longer periodicity can be identified. However, at the same time the statistical significance of the identified oscillations may decline. We used the following values of M: 5 yr ($M = 60$ months), 10 yr ($M = 120$), 20 yr ($M = 240$), approx. 28 yr [$M = 340 \approx N/4$, following the recommendation of Elsner and Tsonis (1996)] and approx. 38 yr [$M = 450 \approx N/3$, following Vautard and Ghil (1989)].

The identified periodicities and their significance levels with increasing lag window are presented in Fig. 7. The numbers in Fig. 7 show the dominant periods associated with the oscillations. These dominant periods are estimated using FFT. From Fig. 7 we can see that in 20CR the significant periodicities are consistently found at 3.6, 2.3 and 1.7 yr, depending to some extent on M. Those periods are more or less visible in HadGEM2 and to a lesser extent in MPI-ESM. Significant 5–6 yr oscillations are identified in all the data sets and especially a $\sim$5.5 yr variability is found consistently.

2–6 yr oscillations are usually attributed to ENSO which is a globally dominating form of variability on annual to decadal time scales (e.g. Kleeman, 2008). It is a broadband phenomenon with several spectral peaks and the highest peak is around 4 yr. This can also be seen in our analysis of 20CR, HadGEM2 and MPI-ESM data sets because most of the significant oscillations are concentrated on 2–6 yr time scales. However, the spectra of MPI-ESM (Fig. 5) differs distinctly from the spectra of the other two data sets: the

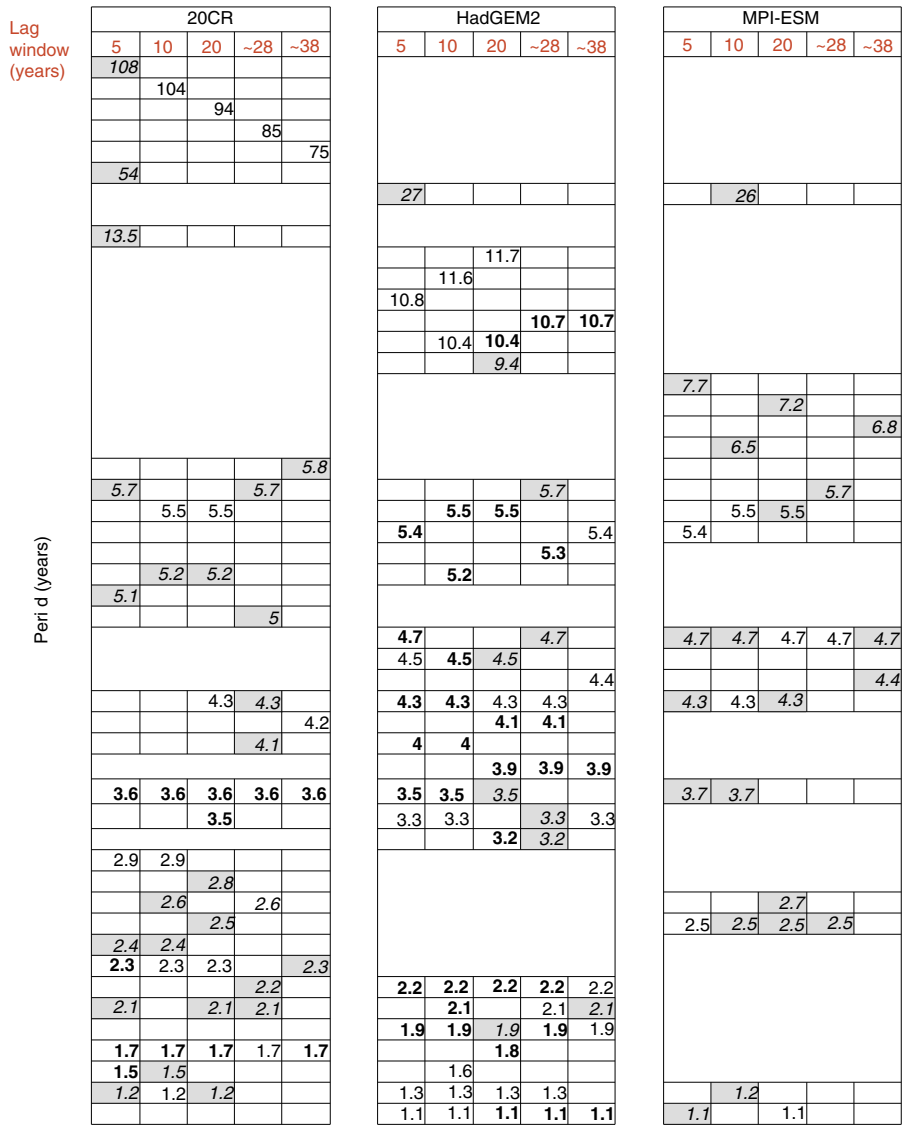


Fig. 7. Periodicities (years) detected by RMSSA/MC-MSSA with varying lag window length (years) for each data set (20CR, HadGEM2 and MPI-ESM). The similar periodicities among the data sets are aligned. Numbers in the figure are in bold if the significance level of a periodicity is 1%, and with grey background if 10%. Otherwise the significance level is 5%. Dominant frequencies of the oscillations are estimated using fast Fourier transform (FFT).

power on 2–8 yr time scales is distributed on a large set of components (especially ST-PCs 4–18) which also decreases the statistical significance of oscillations on those time scales.

In HadGEM2, significant decadal scale oscillations are identified with all lag window lengths. Dominant peak on the decadal time scales has been noted by Collins et al. (2008) and one of the possible reasons for this is in deficiencies of simulation of the ENSO phase-changing process in HadGEM2 (Martin et al., 2010).

There are also significant multi-decadal components in 20CR data set, but their period decreases with increasing lag window M. The time series to be analysed become

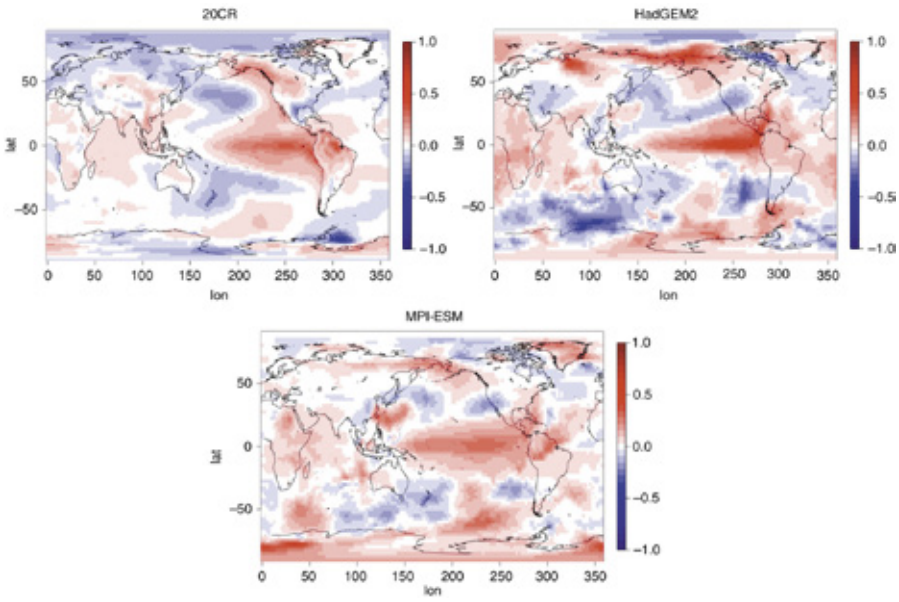


Fig. 8. Global patterns of $\sim$5.5 yr oscillation of the near-surface temperature anomaly (°C) in 20CR, HadGEM2 and MPI-ESM data sets 1901–2012. The patterns are calculated as composites of eight cases, when the oscillation is in its maximum positive phase in the equatorial Pacific. Those positive events are defined as an average of winter months (Nov–Mar). See the text for more details on the reconstruction procedure. The identified patterns have similarities to El Niño -phenomenon.

shorter with increasing M and this may have an effect on the identified period length. We did not find significant multi-decadal components in HadGEM2 and MPI-ESM data sets, although 27 yr and 26 yr periods are identified on 10% significance levels, but only with a single lag window length. However, the use of a lag window M typically allows only the distinction of oscillations with periods $\leq M$ and thus the interpretation of those multi-decadal components remains uncertain.

3.3. *Reconstruction of the significant oscillations*

The final step of our analysis is to reconstruct the decomposed signals in the original space. As an illustration, we have chosen to reconstruct the signal corresponding to approximately 5.5 yr variability, which was identified in all the data sets.

In order to see the time evolution of the $\sim$5.5 yr variability, we have reconstructed the time series in each gridpoint of the original data set with the ST-PCs corresponding to the signal of interest. I.e., in the reconstruction we have projected the original (centred) data set onto ST-PCs (calculated in the reduced space) to obtain ST-EOFs in the original space and then projected the ST-PCs onto those ST-EOFs (see Appendices A.2 and A.4 for more details). In order to see the global effects of the $\sim$5.5 yr cycle, the time series of each grid point has its original variance. The above calculations were completed for each data set using their own $\sim$5.5 yr patterns. ST-PCs 5 and 6 of the 20CR data set (Fig. 3), ST-PCs 6 and 7 of HadGEM2 (Fig. 4) and ST-PCs 4 and 5 of MPI-ESM (Fig. 5) were used in the reconstruction.

Once we have reconstructed the time series in each gridpoint we can plot the anomalies related to the signal month by month. These plots are presented as animations of each data set (20CR, HadGEM2 and MPI-ESM) for a time period of 1901–2012 (the animations are available at www.youtube.com/channel/UCRjwc6cI-TzbvtShONYZ7cg). In Fig. 8, we also show the global patterns of the $\sim$5.5 yr variability of near-surface temperature anomaly. The patterns are composites of eight cases, when the oscillation is in its positive phase in the equatorial Pacific. Positive events are defined as an average of winter months (November–March).

The temperature anomalies of 20CR have many similarities to global El Niño effects, such as above average temperatures in the central and eastern equatorial Pacific Ocean, in the western and northern parts of North-America and South-America as well as in South-East Asia, Australia and southern Africa. Below average anomalies are found in the south-east parts of North-America, in the north-west and south-west Pacific as well as in northern parts of Eurasia.

In 20CR, a typical north-south wave train is also seen, but the east-west patterns are weaker, except for the anomalies at the Amazonas.

HadGEM2 and MPI-ESM show similarities to 20CR, but differences can be seen, for example, in the Pacific forcing patterns. Especially in MPI-ESM the centre of the forcing pattern seems to be more western. In the model simulations, the negative anomaly near the west-coast of North-America extends to the continent, which is not detected in 20CR. The positive anomalies in HadGEM2 and MPI-ESM also extend into the northern Eurasia and there are anomaly patterns in the southern Indian Ocean which are absent in 20CR. MPI-ESM has a stronger positive anomaly in the coast of South-East Asia compared to the other two data sets. In addition, there is a strong anomaly near the Antarctic Peninsula in the Weddel Sea in the 20CR data set which is not detected in the model simulations. The anomaly patterns in the Atlantic Ocean are also weaker in 20CR compared to simulations.

The animations of the $\sim$5.5 yr pattern (available at www.youtube.com/channel/UCRjwc6cI-TzbvtShONYZ7cg) show some more features in addition to the ones seen in Fig. 8. For instance, in 20CR animation there is a quite strong anomaly pattern to the west of Ural Mountains. This pattern is not usually associated with ENSO, and its maximum negative and positive phases seem to occur at different times compared to the ENSO-related anomaly patterns in the Pacific. However, this pattern to the west of Ural might also reflect some other phenomenon, mixed with the ENSO patterns.

The animations also show that the variability has a more propagating character in 20CR data set whereas the anomaly patterns in the model simulations are more stationary. In the northern and southern Pacific Ocean, for example, the anomalies seem to propagate eastward in the 20CR animation.

Compared to 20CR, HadGEM2 and MPI-ESM show a richer structure in Fig. 8 and in the animations. One has to remember that the reanalysis data set is an ensemble mean whereas the analysis of the climate model simulations is conducted on a single ensemble member of each model. This may also contribute to the structure seen in the model simulations. Different, more or less real, phenomena may also be mixed in the variability patterns of the simulations.

4. Summary and Discussion

We have introduced an RMSSA algorithm, which allows the calculation of MSSA of extremely high-dimensional problems. The RMSSA algorithm first reduces the dimension of the original data set by RP, then decomposes the data set into components of different frequencies by calculating MSSA in a reduced space, and finally reconstructs the components in the original high-dimensional space.

We have applied the RMSSA algorithm to decompose the monthly mean near-surface air temperature of the 20th century reanalysis and the historical 20th century simulations of HadGEM2-ES and MPI-ESM-MR extracted from the CMIP5 data archives. We have also performed Monte-Carlo simulations in order to estimate the significance of the identified low-frequency components. Our analysis shows that 2–6 yr oscillations are present in all the data sets. Their statistical significance is highest in HadGEM2 while in MPI-ESM the power on those timescales is distributed on a large set of components decreasing their statistical significance.

2–6 yr oscillations are usually attributed to ENSO which is a globally dominating form of variability on annual to decadal time scales. Our global monthly animations of 5–6 yr near-surface temperature cycle match quite well with the known temperature anomalies related to ENSO. The reanalysis and the historical simulations have similar anomaly patterns in the central and eastern Pacific Ocean, around the northern part of Indian Ocean as well as in the north-west North-America, but also some notable differences in several areas, such as Eurasia. Also, our animations of the 5–6 yr cycle reveal a propagating structure in the near-surface temperature anomalies of 20CR, while the variability in HadGEM2 and MPI-ESM data sets is more stationary. The focus of this study was to introduce the RMSSA algorithm and the discussion on the possible causes for the differences in oscillatory patterns of the data sets is limited. However, this would be a subject for a further study with a larger set of climate model data sets included.

RMSSA algorithm is a powerful tool when the dimensions of the data sets become prohibitively large. It allows a computationally efficient way of decomposing a data set into its spatio-temporal patterns. Several climatic state variables can be incorporated in the RMSSA at the same time in order to find the co-varying signals and illustrate their propagation. RMSSA can also be used in studying the oscillations in three dimensions including data from several atmospheric levels in the analysis.

5. Acknowledgements

We thank Petteri Uotila and Mikko Alestalo from Finnish Meteorological Institute for interesting discussions at earlier phases of the work. We also thank Jouni Räisänen from University of Helsinki for preparing the HadGEM2-ES and MPI-ESM-MR data sets used in this study. In addition, we thank the reviewer for her/his valuable comments and suggestions that helped in improving the manuscript. This research has been funded by the Academy of Finland (project number 140771).

Appendix A

A.1. MSSA and Monte-Carlo MSSA

A.1.1. Multichannel singular spectrum analysis (MSSA)

The aim of MSSA is to identify spatially and temporally coherent patterns in a multivariate data set. In MSSA terminology, the columns of the original data matrix $\mathbf{X}_{N\times L}$ are called channels. In case of gridded data set, N represents the time steps and L is the number of grid points:

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,L} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,L} \end{bmatrix} \tag{A1}$$

The next step is to construct an augmented data matrix $\mathbf{A}$, which contains M lagged copies of each channel in $\mathbf{X}$:

$$\mathbf{Y_i} = \begin{bmatrix} x_{1,i} & x_{2,i} & \cdots & x_{M,i} \\ x_{2,i} & x_{3,i} & \cdots & x_{M+1,i} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N\prime,i} & x_{N\prime+1,i} & \cdots & x_{N,i} \end{bmatrix}, i = 1...L \tag{A2}$$

and

$$\mathbf{A} = [\mathbf{Y_1} \quad \mathbf{Y_2} \quad \cdots \quad \mathbf{Y_L}] \tag{3}$$

In MSSA, M represents the lag window. $\mathbf{A}$ has now ML columns and $N' = N - M + 1$ rows. The singular value decomposition (SVD) of $\mathbf{A}$ can now be calculated:

$$\mathbf{A} = \mathbf{U}_A \mathbf{D}_A^{1/2} \mathbf{V}_A^T, \tag{4}$$

The vectors of $\mathbf{U}_A$ are the eigenvectors of $\mathbf{Z}_A = \frac{1}{ML}\mathbf{A}\mathbf{A}^T$ and $\mathbf{V}_A^T$ contains the eigenvectors of $\mathbf{C}_A = \frac{1}{N\prime}\mathbf{A}^T\mathbf{A}$. These vectors are orthogonal and often called space-time principal components (ST-PCs) and space-time empirical orthogonal functions (ST-EOFs), respectively. Diagonal elements of $\mathbf{D}_A$ are the eigenvalues of $\mathbf{C}_A$ or $\mathbf{Z}_A$.

Optionally the lag-covariance matrix $\mathbf{C}_A$ (or $\mathbf{Z}_A$) and its eigendecomposition can be calculated to yield eigenvectors $\mathbf{V}_A^T$ (or $\mathbf{U}_A$) and eigenvalues (diagonal elements of matrix $\mathbf{D}_A = \mathbf{V}_A^T\mathbf{C}_A\mathbf{V}_A$or $\mathbf{D}_A = \mathbf{U}_A^T\mathbf{Z}_A\mathbf{U}_A$). Matrix $\mathbf{U}_A$ (or $\mathbf{V}_A^T$) can be obtained by projecting $\mathbf{A}$ onto $\mathbf{V}_A^T$(or $\mathbf{U}_A$). If $N' > ML$ (or $ML > N'$), it is more convenient to estimate $\mathbf{C}_A$ (or $\mathbf{Z}_A$) because it is smaller. See Allen and Robertson (1996) for details.

A.1.2. Monte-Carlo MSSA

The components obtained by MSSA can be tested against a null-hypothesis of the data being generated by independent AR(1) processes (i.e. red noise). The red noise model has the form:

$$u_{t+1,s} = \gamma_s u_{t,s} + \alpha_s w_{t,s}, \tag{A5}$$

where γ_s is the lag-1 autocorrelation of channel s (in the original data set), $\alpha_s = \sqrt{c_s(1-\gamma_s^2)}$ (c_s is the variance of channel s) and $W_{t,s}$ is Gaussian white noise. The data set generated by the model in (A5) is called the surrogate data set and it is subjected to MSSA in the same way as the original data set. Large number of surrogates are generated in order to estimate the confidence limits for the MSSA results of the original data set.

In the test of Allen and Robertson (1996), the lag-covariance matrices of the original data set and the red-noise surrogates are projected either onto the data-adaptive basis (i.e. $\mathbf{U}_A$ or $\mathbf{V}_A^T$) or the null-hypothesis basis. The null-hypothesis basis can be calculated from the expected lag-covariance matrix $\mathbf{C}_N$ of the red-noise surrogates. $\mathbf{C}_N$ can be estimated analytically by

$$[\mathbf{C}_N] = \frac{1}{ML}\sum_{s=1}^{ML} c_s \gamma_s^{|ii-jj|} \tag{A6}$$

Projection onto the red-noise basis is considered more reliable because the use of the data-adaptive basis assumes the existence of an oscillation even in a case where it is uncertain whether or not the oscillation is significant.

According to Allen and Robertson (1996), the input channels should be uncorrelated (or at least nearly) at zero lag for the test to be useful. In the case of a gridded data set, where all the grid point time series are used as input channels, the decorrelation condition is not valid. The test might still be useful if we are using grid points sufficiently far from each other as the input channels for the test (Ghil et al., 2002).

A.2. Randomised algorithms for MSSA

1: Original MSSA algorithm enhanced by RP

(1) construct the original data matrix $\mathbf{X}_{N\times L}$
(2) (pre-processing of $\mathbf{X}$, if needed)
(3) generate k L-dimensional vectors of Gaussian distributed random numbers to get matrix $\mathbf{R}$ (and optionally orthogonalise the random vectors)
(4) project the original data matrix onto random vectors: $\mathbf{P}_{N\times k} = \frac{1}{\sqrt{k}}\mathbf{X}_{N\times L}\mathbf{R}_{L\times k}$
(5) generate augmented matrix $\mathbf{A}_{RP}$ of $\mathbf{P}$
(6) calculate SVD: $\mathbf{A}_{RP} = \mathbf{U}_{RP}\mathbf{D}_{RP}^{1/2}\mathbf{V}_{RP}^T$ (or covariance matrix $\mathbf{C}_{RP}$ or $\mathbf{Z}_{RP}$ and its eigendecomposition)
(7) calculate ST-EOFs in the original space: $\mathbf{V}_A \approx \mathbf{A}^T\mathbf{U}_{RP}(\mathbf{D}_{RP}^{1/2})^{-1}$ (see Appendix A.4 for an explanation)
(8) calculate RCs using ST-EOFs of step 7.

2: PC-based MSSA algorithm enhanced by RP

(1) construct the original data matrix $\mathbf{X}_{N\times L}$
(2) (pre-processing of $\mathbf{X}$, if needed)
(3) generate k L-dimensional vectors of Gaussian distributed random numbers to get matrix $\mathbf{R}$ (and optionally orthogonalise the random vectors)
(4) project the original data matrix onto random vectors: $\mathbf{P}_{N\times k} = \frac{1}{\sqrt{k}}\mathbf{X}_{N\times L}\mathbf{R}_{L\times k}$
(5) calculate SVD of $\mathbf{P}$ (see Appendix A.3 for an explanation of how the covariance is preserved in RP + SVD)
(6) retain e.g. 30 first PCs of $\mathbf{P}$ to obtain reduced matrix $\mathbf{T}$
(7) generate augmented matrix $\mathbf{A}_{PC}$ of $\mathbf{T}$
(8) calculate SVD: $\mathbf{A}_{PC} = \mathbf{U}_{PC}\mathbf{D}_{PC}^{1/2}\mathbf{V}_{PC}^T$ (or covariance matrix $\mathbf{C}_{PC}$ or $\mathbf{Z}_{PC}$ and its eigendecomposition)
(9) (MC-MSSA step)
(10) calculate ST-EOFs in the original space: $\mathbf{V}_A \approx \mathbf{A}^T\mathbf{U}_{PC}(\mathbf{D}_{PC}^{1/2})^{-1}$ (see Appendix A.4 for an explanation)
(11) calculate RCs using ST-EOFs of step 10.

A.3. RP and SVD

The method to back-project from the reduced space to the original space in the case of RP + SVD is explained in Seitola et al. (2014) (Appendix A.1) but we also present it briefly here:

The SVD of the original data matrix $\mathbf{X}_{N\times L}$ is:

$$\mathbf{X}_{N\times L} = \mathbf{U}_{N\times N}\mathbf{D}_{N\times L}\mathbf{V}_{L\times L}^T \tag{A7}$$

$\mathbf{U}$ contains the eigenvectors of $\mathbf{Z} = \mathbf{X}\mathbf{X}^T$.

Random projection (RP) of $\mathbf{X}$ is $\mathbf{P} = \mathbf{X}\mathbf{R}$, where $\mathbf{R}_{L\times k}$ is the projection matrix and the row vectors of $\mathbf{R}$ are scaled to have unit length. Thus, we can write:

$$\mathbf{Z}_{RP} = \mathbf{X}\mathbf{R}(\mathbf{X}\mathbf{R})^T = \mathbf{X}\mathbf{R}\mathbf{R}^T\mathbf{X}^T \approx \mathbf{X}\mathbf{X}^T = \mathbf{Z} \tag{A8}$$

In the previous, we have assumed that $\mathbf{R}\mathbf{R}^T \approx \mathbf{I}$ because the row vectors of $\mathbf{R}$ are nearly orthonormal. It is also possible to make the vectors of $\mathbf{R}$ strictly orthonormal, in which case $\mathbf{R}\mathbf{R}^T = \mathbf{I}$. However, orthogonalisation is often not necessary, because the difference between the orthogonalised and non-orthogonalised random vectors is very small, especially in high-dimensions.

Let's rewrite (A7) as $\mathbf{X}_{N\times L} = \mathbf{U}_{N\times r}\mathbf{D}_{r\times r}\mathbf{V}_{r\times L}^T$, where $r = \mathrm{rank}(\mathbf{X})$. Now we can manipulate (A7):

$$\begin{aligned}
&\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T && (\mathbf{U}^T\mathbf{U} = \mathbf{I})\\
&\mathbf{U}^T\mathbf{X} = \mathbf{D}\mathbf{V}^T && (\mathbf{D}^{-1}\mathbf{D} = \mathbf{I})\\
&\mathbf{V}^T = \mathbf{D}^{-1}\mathbf{U}^T\mathbf{X} && \text{transpose of both sides}\\
&\mathbf{V} = \mathbf{X}^T\mathbf{U}(\mathbf{D}^{-1})^T = \mathbf{X}^T\mathbf{U}\mathbf{D}^{-1}
\end{aligned} \tag{A9}$$

Because $\mathbf{Z} \approx \mathbf{Z}_{RP}$ we can approximate

$$\begin{aligned}
&\mathbf{U} \approx \mathbf{U}_{RP},\\
&\mathbf{D} \approx \mathbf{D}_{RP} \quad \text{and}\\
&\mathbf{V} \approx \mathbf{X}^T\mathbf{U}_{RP}\mathbf{D}_{RP}^{-1}
\end{aligned} \tag{A10}$$

In the previous, we have defined $\mathbf{U}_{RP}$ as $N\times k$ and $\mathbf{D}_{RP}$ as a $k\times k$ matrix, where k is the rank of matrix $\mathbf{P}_{N\times k}$.

A.4. RP and MSSA

In this appendix, we will explain how to get from the reduced space back to the original space in the case of RP + MSSA.

Let's write the original data matrix $\mathbf{X}_{N\times L}$ as

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,L}\\ x_{2,1} & x_{2,2} & \cdots & x_{2,L}\\ \vdots & \vdots & \ddots & \vdots\\ x_{N,1} & x_{N,2} & \cdots & x_{N,L}\end{bmatrix} = \begin{bmatrix}\mathbf{x_1}\\ \mathbf{x_2}\\ \vdots\\ \mathbf{x_N}\end{bmatrix}, \tag{A11}$$

where $\mathbf{x_i}$ are the row vectors of $\mathbf{X}$.

The augmented matrix $\mathbf{A}$ of $\mathbf{X}$ is already defined in Appendix A.1. Now let's calculate $\mathbf{A}\mathbf{A}^T$.

$$\begin{aligned}
\mathbf{A}\mathbf{A}^T &= \begin{bmatrix}\mathbf{Y_1} & \mathbf{Y_2} & \cdots & \mathbf{Y_L}\end{bmatrix}\begin{bmatrix}\mathbf{Y_1^T}\\ \mathbf{Y_2^T}\\ \vdots\\ \mathbf{Y_L^T}\end{bmatrix}\\
&= \left[\mathbf{Y_1}\mathbf{Y_1^T} + \mathbf{Y_2}\mathbf{Y_2^T} + \cdots + \mathbf{Y_L}\mathbf{Y_L^T}\right]
\end{aligned} \tag{A12}$$

After some algebra we get

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} \mathbf{x}_1\mathbf{x}_1^T + \mathbf{x}_2\mathbf{x}_2^T + \cdots + \mathbf{x}_M\mathbf{x}_M^T & \mathbf{x}_1\mathbf{x}_2^T + \mathbf{x}_2\mathbf{x}_3^T + \cdots + \mathbf{x}_M\mathbf{x}_{M+1}^T & \cdots & \mathbf{x}_1\mathbf{x}_{N'}^T + \mathbf{x}_2\mathbf{x}_{N'+1}^T + \cdots + \mathbf{x}_M\mathbf{x}_N^T\\ \mathbf{x}_2\mathbf{x}_1^T + \mathbf{x}_3\mathbf{x}_2^T + \cdots + \mathbf{x}_{M+1}\mathbf{x}_M^T & \mathbf{x}_2\mathbf{x}_2^T + \mathbf{x}_3\mathbf{x}_3^T + \cdots + \mathbf{x}_{M+1}\mathbf{x}_{M+1}^T & \cdots & \mathbf{x}_2\mathbf{x}_{N'}^T + \mathbf{x}_3\mathbf{x}_{N'+1}^T + \cdots + \mathbf{x}_{M+1}\mathbf{x}_N^T\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{x}_{N'}\mathbf{x}_1^T + \mathbf{x}_{N'+1}\mathbf{x}_2^T + \cdots + \mathbf{x}_N\mathbf{x}_M^T & \mathbf{x}_{N'}\mathbf{x}_2^T + \mathbf{x}_{N'+1}\mathbf{x}_3^T + \cdots + \mathbf{x}_N\mathbf{x}_{M+1}^T & \cdots & \mathbf{x}_{N'}\mathbf{x}_{N'}^T + \mathbf{x}_{N'+1}\mathbf{x}_{N'+1}^T + \cdots + \mathbf{x}_N\mathbf{x}_N^T \end{bmatrix} \tag{A13}$$

Now let's calculate RP of $\mathbf{X}$:

$$\mathbf{XR} = \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ \vdots \\ \mathbf{x_N} \end{bmatrix} \mathbf{R} = \begin{bmatrix} \mathbf{x_1 R} \\ \mathbf{x_2 R} \\ \vdots \\ \mathbf{x_N R} \end{bmatrix} \tag{A14}$$

The augmented matrix of is $\mathbf{A}_{RP}$:

$$\mathbf{A}_{RP} = \begin{bmatrix} \mathbf{x_1 R} & \mathbf{x_2 R} & \cdots & \mathbf{x_M R} \\ \mathbf{x_2 R} & \mathbf{x_3 R} & \cdots & \mathbf{x_{M+1} R} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x_{N\prime} R} & \mathbf{x_{N\prime+1} R} & \cdots & \mathbf{x_N R} \end{bmatrix} \tag{A15}$$

Let's calculate $\mathbf{A}_{RP}\mathbf{A}_{RP}^T$:

$$\mathbf{A}_{RP}\mathbf{A}_{RP}^T = \begin{bmatrix} \mathbf{x_1 R} & \mathbf{x_2 R} & \cdots & \mathbf{x_M R} \\ \mathbf{x_2 R} & \mathbf{x_3 R} & \cdots & \mathbf{x_{M+1} R} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x_{N\prime} R} & \mathbf{x_{N\prime+1} R} & \cdots & \mathbf{x_N R} \end{bmatrix} \times \begin{bmatrix} \mathbf{R^T x_1^T} & \mathbf{R^T x_2^T} & \cdots & \mathbf{R^T x_{N\prime}^T} \\ \mathbf{R^T x_2^T} & \mathbf{R^T x_3^T} & \cdots & \mathbf{R^T x_{N\prime+1}^T} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R^T x_M^T} & \mathbf{R^T x_{M+1}^T} & \cdots & \mathbf{R^T x_N^T} \end{bmatrix} \tag{A16}$$

Because $\mathbf{RR}^T \approx \mathbf{I}$, the first element of $\mathbf{A}_{RP}\mathbf{A}_{RP}^T$ can be written as $\mathbf{x_1 R R^T x_1^T} + \mathbf{x_2 R R^T x_2^T} + \cdots + \mathbf{x_M R R^T x_M^T} \approx \mathbf{x_1 x_1^T} + \mathbf{x_2 x_2^T} + \cdots + \mathbf{x_M x_M^T}$

After calculating all the elements of $\mathbf{A}_{RP}\mathbf{A}_{RP}^T$ as above, we see that $\mathbf{AA}^T \approx \mathbf{A}_{RP}\mathbf{A}_{RP}^T$. Therefore, as in Appendix A.3, we can approximate

$$\begin{aligned} &\mathbf{U}_A \approx \mathbf{U}_{RP}, \\ &\mathbf{D}_A \approx \mathbf{D}_{RP} \quad \text{and} \\ &\mathbf{V}_A \approx \mathbf{A}^T \mathbf{U}_{RP} \mathbf{D}_{RP}^{-1} \end{aligned} \tag{A17}$$

Same kind of reasoning applies also when the PCs of the data set are used as channels in MSSA. We can write PCs as $\mathbf{U}_{N\times r}\mathbf{D}_{r\times r} = \mathbf{X}_{N\times L}\mathbf{V}_{L\times r}$, where $r = \mathrm{rank}(\mathbf{X})$. Vectors of $\mathbf{V}$ are orthonormal, so in the above calculations we can replace $\mathbf{R}$ with $\mathbf{V}$.

References

Allen, M. R. and Robertson, A. W. 1996. Distinguishing modulated oscillations from coloured noise in multivariate datasets. *Clim. Dyn.* **12**(11), 775–784.

Bingham, E. and Mannila, H. 2001. Random projection in dimensionality reduction: applications to image and text data. In: *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '01. New York, ACM, 245–250.

Broomhead, D. S. and King, G. P. 1986a. Extracting qualitative dynamics from experimental data. *Physica D* **20**, 217–236.

Broomhead, D. S. and King, G. P. 1986b. On the qualitative analysis of experimental dynamical systems. In: *Nonlinear Phenomena and Chaos* (ed. S. Sarkar). Adam Hilger, Bristol, pp. 113–144.

Chiu, S. K. 2013. Coherent and random noise attenuation via multichannel singular spectrum analysis in the randomized domain. *Geophys. Prospect.* **61**, 1–9.

Cleveland, R. B., Cleveland, W. S., McRae, J. E. and Terpenning, I. 1990. STL: a seasonal-trend decomposition procedure based on loess. *J Off. Stat.* **6**, 373.

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P. and co-authors. 2011. Development and evaluation of an Earth-System model HadGEM2. *Geosci. Model Dev.* **4**, 1051–1075.

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Hinton, T. and co-authors. 2008. Evaluation of the HadGEM2 model. Met Office Hadley Centre Technical Note no. HCTN 74, available from Met Office, FitzRoy Road, Exeter EX1 3PB. Online at: http://www.metoffice.gov.uk/publications/HCTN/index.html

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J. and co-authors. 2011. The twentieth century reanalysis project. *Quart. J. Roy. Meteorol. Soc.* **137**, 1–28.

Dasgupta, S. and Gupta, A. 2003. An elementary proof of a theorem of Johnson and Lindenstrauss. *Rand. Struct. Algo.* **22**, 60–65.

Elsner, J. B. and Tsonis, A. A. (eds.). 1996. *Singular Spectrum Analysis: A New Tool in Time Series Analysis*. Springer Science & Business Media, New York, NY, USA.

Frankl, P. and Maehara, H. 1988. The Johnson-Lindenstrauss lemma and the sphericity of some graphs. *J. Combin. Theory Ser. B* **44**, 355–362.

Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K., Kondrashov, D. and co-authors. 2002. Advanced spectral methods for climatic time series. *Rev. Geophys.* **40**(1), 1–41.

Halko, N., Martinsson, P. G. and Tropp, J. A. 2011. Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.* **53**(2), 217–288.

Kawale, J., Chatterjee, S., Kumar, A., Liess, S., Steinbach, M. and co-authors. 2011. Anomaly construction in climate data: issues and challenges. In: *Proceedings of the 2011 Conference on Intelligent Data Understanding*, CIDU, Mountain View, CA, USA, pp. 189–203.

Kleeman, R. 2008. Stochastic theories for the irregularity of ENSO. *Philos. Trans. Roy. Soc. A Math. Phys. Eng. Sci.* **366**(1875), 2509–2524.

Mann, M. E. and Lees, J. M. 1996. Robust estimation of background noise and signal detection in climatic time series. *Clim. Change* **33**, 409–445.

Martin, G. M., Milton, S. F., Senior, C. A., Brooks, M. E., Ineson, S. and co-authors. 2010. Analysis and reduction of systematic errors through a seamless approach to modeling weather and climate. *J. Clim.* **23**(22), 5933–5957.

Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G. and co-authors. 2014. Decadal climate prediction: an update from the trenches. *Bull. Amer. Meteor. Soc.* **95**, 243–267.

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G. and co-authors. 2009. Decadal prediction: can it be skillful? *Bull. Amer. Meteor. Soc.* **90**, 1467–1485.

Moron, V., Robertson, A. W. and Ghil, M. 2012. Impact of the modulated annual cycle and intraseasonal oscillation on

daily-to-interannual rainfall variability across monsoonal India. *Clim. Dyn.* **38**, 2409–2435.

Oropeza, V. and Sacchi, M. 2011. Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis. *Geophysics* **76**, V25–V32.

Plaut, G. and Vautard, R. 1994. Spells of low-frequency oscillations and weather regimes in the Northern Hemisphere. *J. Atmos. Sci.* **51**(2), 210–236.

Rinne, J. and Karhila, V. 1979. Empirical orthogonal functions of 500 mb height in the northern hemisphere determined from a large data sample. *Quart. J. Roy. Meteorol. Soc.* **105**, 873–884.

Seitola, T., Mikkola, V., Silen, J. and Järvinen, H. 2014. Random projections in reducing the dimensionality of climate simulation data. *Tellus A* **66**, 25274. DOI: http://dx.doi.org/10.3402/tellusa.v66.25274

Solomon, A., Goddard, L., Kumar, A., Carton, J., Deser, C. and co-authors. 2011. Distinguishing the roles of natural and anthropogenically forced decadal climate variability. *Bull. Amer. Meteor. Soc.* **92**, 141–156.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T. and co-authors. 2013. Atmospheric component of the MPIM Earth System Model: ECHAM6. *J. Adv. Model. Earth Syst.* **5**(2), 146–172.

Taylor, K. E., Stouffer, R. J. and Meehl, G. A. 2012. An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.* **93**, 485–498.

Thomson, D. J. 1982. Spectrum estimation and harmonic analysis. *Proc. IEEE.* **70**, 1055–1096.

Vautard, R. and Ghil, M. 1989. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Phys. D Nonlin. Phenom.* **35**(3), 395–424.

Von Storch, H. and Zwiers, F. W. 1999. *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge.

Weare, B. C. and Nasstrom, J. S. 1982. Examples of extended empirical orthogonal function analysis. *Mon. Weath. Rev.* **110**, 481–485.

IV

Multi-annual modes in the 20th century temperature variability in reanalyses and CMIP5 models

Heikki Järvinen[1], Teija Seitola[1,2], Johan Silén[2], and Jouni Räisänen[1]

[1]Department of Physics, University of Helsinki, Finland
[2]Finnish Meteorological Institute, Helsinki, Finland

Correspondence to: Heikki Järvinen (heikki.j.jarvinen@helsinki.fi)

Abstract. A performance expectation is that Earth system models simulate well the climate mean state and the climate variability. To test this expectation, we decompose two 20th century reanalysis data sets and 12 CMIP5 model simulations for years 1901 – 2005 of the monthly mean near-surface air temperature using Randomised Multi-Channel Singular Spectrum Analysis (RMSSA). Due to the relatively short time span, we concentrate on the representation of multi-annual variability which the RMSSA method effectively captures as separate and mutually orthogonal spatio-temporal components. This decomposition is a unique way to separate statistically significant quasi-periodic oscillations from one another in high-dimensional data sets.

The main results are as follows. First, the total spectra for the two reanalysis data sets are remarkably similar in all time scales, except that the spectral power in ERA-20C is systematically slightly higher than in 20CR. Apart from the slow components related to multi-decadal periodicities, ENSO oscillations with approximately 3.5 yr and 5 yr periods are the most prominent forms of variability in both reanalyses. In 20CR, these are relatively slightly more pronounced than in ERA-20C. Since about the 1970's, the amplitudes of the 3.5 yr and 5 yr oscillations have increased, presumably due to some combination of forced climate change, intrinsic low-frequency climate variability, or change in global observing network. Second, none of the 12 coupled climate models closely reproduce all aspects of the reanalysis spectra, although some models represent many aspects well. For instance, the GFDL-ESM2M model has two nicely separated ENSO periods although they are relatively too prominent as compared with the reanalyses. There is an extensive Supplement and Youtube videos to illustrate the multi-annual variability of the data sets.

Keywords: spatio-temporal modes, climate variability, climate model simulation, random projection, RMSSA algorithm, ENSO oscillation, Youtube video

1 Introduction

The ultimate goal in developing Earth system models (ESM) is to enable exploitation of the inherent Earth system predictability, and hence reduce weather and climate related uncertainties in our daily life, and guide societies in making sustainable choices (e.g., Slingo and Palmer 2011; Meehl et al. 2014). For the predictions to be useful and usable, the expectation is that the climate mean state and climate variability are well simulated by these tools. Due to the complexity of the models and the data they

produce, testing the expectation poses a challenge: many aspects of the model performance are gathered under the variability concept and no single diagnostic alone is sufficient to exhaust its all facets. Yet, understanding the discrepancies between the observed and simulated variability is crucial feedback for model development.

Representation of climate variability among models participating in climate model inter-comparisons, such as CMIP5, has been studied by e.g. Bellenger et al. (2014), Knutson et al. (2013), Ba et al. (2014), and Fredriksen and Rypdal (2016). We will add to this literature by interfacing a representative set of contemporary coupled climate models with reanalysis data focusing on spatio-temporal modes of climate variability. One century covered with global reanalysis data is naturally very short for this purpose and severely constrains inter-comparison studies (e. g. Wittenberg 2009 and Stevenson et al. 2010). First, time series should cover a sufficient number of recurring "events" for obtaining significance for the findings. Therefore, decadal-to-multi-decadal variability is of interest but not as informative as focusing on shorter cycles of variability. Second, the applied methods have to be very effective in extracting information from the short but high-dimensional data sets. For these reasons, we concentrate on the representation of multi-annual variability in reanalyses and coupled climate models applying Randomised Multi-Channel Singular Spectrum Analysis (RMSSA; Seitola et al. 2014, 2015) which effectively separates mutually orthogonal spatio-temporal components from our high-dimensional data sets.

The aim of this study is to decompose the 20th century climate variability into its multi-annual modes, and to assess how these modes are represented by the contemporary climate models. We hope this to provide guidance for model development due better understanding of the deficiencies in representing reanalysed modes of multi-annual climate variability. Ultimately, interpreting the hints about model deficiencies as development topics are due for the development teams themselves. Our role is to point towards the potential error sources. For reassuring the teams that high-dimensional time series analysis is possible today, we emphasise the methodological aspect of this study. RMSSA can, under very weak assumptions on the data, decompose high-dimensional data sets in a unique way and separate statistically significant quasi-periodic spatio-temporal oscillations from one another. This is in contrast to many other approaches which either make assumptions about the oscillation structures, such as Fourier or spherical decomposition, or resolve only either spatial or temporal aspects of variability. RMSSA can detect spatially evolving "chains of events" through resolving eigenmodes of spatio-temporal covariance data. This is a significant advantage, say, over PCA which only resolves eigenmodes of spatial covariances and often projects temporal evolution of an "event" onto a number of different eigenmodes. In addition, the novel data compression based on random projections enable here a vast increase in tractable problem size (i.e., data dimension) - even multi-variate decomposition is now possible, although not included here.

2 Methods and Data

2.1 Randomised multi-channel singular spectrum analysis

Multi-channel singular spectrum analysis (MSSA; Broomhead and King, 1986a,b) can be characterised as being a time series analysis method for high-dimensional problems. It effectively identifies spatially and temporally coherent patterns of a data set by decomposing a lag-covariance data matrix into its eigenvectors and eigenvalues (e.g., Ghil et al., 2002) using singular value

decomposition (SVD). The lag window in MSSA is a user choice, recommended typically to be shorter than approximately one third of the length of the time series (Vautard and Ghil, 1989). Long lag window enhances the spectral resolution, i.e., the number of frequencies that can be identified, but distributes the variance on a larger set of components. MSSA eigenvectors are called here space-time EOFs (ST-EOFs), and the projections of the data set onto those ST-EOFs space-time principal components (ST-PCs). Because of the lag window, ST-PCs have a reduced length and they cannot be located into the same index space with the original time series. However, they can be represented in the original coordinate system by the reconstructed components (RC; Plaut and Vautard, 1994).

MSSA is computationally expensive and practical limits are easily exceeded for large data sets and long lag windows. In order to overcome this limitation, a computationally more efficient variant, called Randomised MSSA (RMSSA; Seitola et al., 2015), is applied here. The RMSSA algorithm, in a nutshell, (1) reduces the dimension of the original data set by using so-called random projections (RP; Bingham and Mannila, 2001; Achlioptas, 2003), (2) decomposes the data set by calculating standard MSSA in the low-dimensional space, and (3) reconstructs the components in the original high-dimensional space.

In RP, the original data set is projected onto a matrix of Gaussian distributed random numbers (zero mean and unit variance) in order to construct a lower dimensional representation. In this study, we reduce the data volume to about 5 % of the original volume. Since the computational complexity of RP is low, involving only a matrix multiplication, it can be applied to very high-dimensional data sets. Although RP is not a lossless compression, it has the important property that the lower-dimensional data set has essentially the same structure as the original high-dimensional data set. This has been demonstrated for climate model data in Seitola et al. (2014). The RMSSA algorithm is briefly presented in the Appendix A.

2.2 Computation of spectra

The ST-PCs represent the different oscillatory modes extracted from the data set. In order to estimate the dominant frequencies associated with each ST-PC, the power spectrum is calculated with the Multitaper spectral analysis method (MTM) (Thomson, 1982; Mann and Lees, 1996). To further compare the variability modes and their intensities in different data sets, the power spectrum of all the ST-PCs of each data set is summed up to obtain so-called total spectrum. The ST-PCs are already weighted by their respective explanatory power, i.e. multiplied by the corresponding eigenvalue. Therefore the components with more explanatory power have also higher spectral densities compared to the ones that explain only a small fraction of the variance. Therefore no extra weighting is needed in this step.

The uncertainty related to the explanatory power of each ST-PC (i.e. the confidence interval of the respective eigenvalue) is estimated using the Norths rule of thumb for sampling errors (North et al., 1982). The sampling error (e_k) is given by $e_k \sim \lambda_k(2/N)$, where λ_k is the eigenvalue associated with the k^{th} ST-PC and N is the length of the time series. Thus, the confidence interval of the total spectrum describes the uncertainties related to the explanatory power of each ST-PC.

2.3 Statistical significance testing

In data sets of dynamical systems, ST-PCs/ST-EOFs of MSSA often appear as quadratic pairs that explain approximately the same variance and are $\pi/2$ out of phase with each other. However, existence of such a pair does not guarantee any physical

oscillation in the data set, and it may be due to some non-oscillatory processes, such as first-order autoregressive noise. Allen and Robertson (1996) formulated a test, where the oscillatory modes identified with MSSA are tested against a red noise null-hypothesis through Monte Carlo simulation.

Significance testing in MSSA requires solving conventional PCs of the original data set. In case of very high-dimensional problems this easily exceeds practical computational limits. The RMSSA implementation in Seitola et al. (2015) contains the Allen-Robertson test such that the PCs are solved in the dimension-reduced space, and is thus affordable even in very high-dimensional problems. The Appendix A also includes a short description of the significance test.

2.4 Data sources

The data consists of the monthly mean near-surface air temperature from the historical 20th Century simulations of 12 different climate models (Table 1). The selected models originate from different modelling centres, and thus do not have close common ancestor models. Furthermore, the selected models have undergone a long (generally several generations of) history of development, suggesting that the chosen models collectively represent the state-of-the-art. Near-surface temperature was chosen, because many processes must be adequately represented in coupled models to realistically capture the observed temperature distribution (Flato et al., 2013). These include processes in the Earth system component models (atmosphere, ocean, etc.) as well as in their mutual coupling models. Also, for the near-surface temperature, there are corresponding reanalysis data available.

The historical (1901–2005) simulations were extracted from the CMIP5 data archive and they follow the CMIP5 experimental protocol (Taylor et al. 2012). The 20th Century simulations use the historical record of climate forcing factors such as greenhouse gases, aerosols, solar variability, and volcanic eruptions. We used a single ensemble member of each model and the model data sets were interpolated into a common grid of 144×73 points.

As a reference, we used two reanalysis data sets: the 20th Century Reanalysis V2 data (hereafer 20CR) provided by the NOAA/OAR/ESRL PSD (Compo et al., 2011), and ERA-20C data provided by ECMWF (Poli et al., 2013). The data sets are produced using an ensemble of perturbed reanalyses, and the final data set corresponds to the ensemble mean. In 20CR, only surface pressure observations are assimilated, and the observed monthly sea-surface temperature and sea-ice distributions from HadISST1.1 (Rayner et al., 2003) are used as boundary conditions (Compo et al., 2011). In ERA-20C, observations of surface pressure and surface marine winds are assimilated (Poli et al., 2013). Unlike 20CR, it uses a more recent sea-surface temperature and sea ice cover analysis from HadISST2 (Rayner et al., 2006). Both 20CR and ERA-20C are forced by historical record of changes in climate forcing factors (greenhouse gases, volcanic aerosols and solar variations). In order to be consistent with the climate model simulations, the same time period is used (1901–2005, i.e., 1260 monthly mean fields) and the reanalysis data sets were interpolated into the same grid as the model simulations (144×73 points).

2.5 Data processing

Some pre-processing of the data was needed before applying RMSSA. At each grid point the data sets were processed as follows:

– linear trend was fitted and removed,

– annual cycle was estimated using Seasonal-Trend Decomposition (STL; Cleveland et al., 1990) and removed,

– resulting values were mean-centered and divided by the average standard deviation of all the data sets (see Figure 1). Average standard deviation was obtained after removal of the trend and the annual cycle.

The reanalysis and climate model data sets have different temperature standard deviations, which would impact the temperature variability from inter-annual to multi-decadal timescales (e.g., Thompson et al. 2015). To retain these differences, we have used a common normalization factor (i.e., the average standard deviation of all the data sets). This procedure reduces the weight of grid points with high variance, typically at higher latitudes, and hence adds weight on the lower latitude features. After the pre-processing, the dimension reduction step of RMSSA was applied so that approximately 5 % of the original data dimensions were retained. The lag window in the analysis was 20 yr (240 months). The total spectra were obtained from this analysis, and are comparable due to normalisation using the common standard deviation of the data sets.

The statistical significance test uses a red noise null hypothesis. In the test we have used data sets that are normalised by their own standard deviations. Using a common normalisation interferes with generating the red noise surrogates corresponding to each data set. The first 50 PCs of each data set were retained as input. Those PCs explain 79 % of the variability in 20CR, 75 % in ERA-20C, and 70 %–80 % in the climate model data sets. A total of 1000 realisations of red noise surrogate data sets were generated, and confidence interval (95 %) for the oscillatory modes were estimated. We note that transformation to PCs may interfere with the detection of weak signals, as demonstrated by Groth and Ghil (2015).

2.6 Data visualization

We used reconstructed components (RC; see Appendix A) for visualisation of the spatial patterns related to ST-PCs. For each grid point time series, we can calculate the RCs corresponding to the ST-PCs (or modes) of interest. These RC values, reflecting the contribution of each grid point to the mode, can be plotted on a map at each time step. We have used these maps to construct videos of the spatio-temporal modes. In Section 3.5, we have analysed RCs corresponding to 3–4 yr variability. The result is a time series of the data corresponding to the 3–4 yr mode in each grid point and according to its variance after detrending and removing the annual cycle. In the analysis we have neglected 5 yrs in the beginning and the end of the time series, because the reconstruction procedure may be biased there (see the Appendix, eq. A4). The videos can be found at our Youtube channel (https://www.youtube.com/channel/UCu1zJdwJfLaXvfvTqsKCLHw).

To summarise the animations, we have calculated composite maps of the modes. The compositing procedure follows the one described in Plaut and Vautard (1994). The idea is to choose the grid point time series (RC_l) for which the variance is largest, and calculate its time derivative ($RC_l^{'}$). The phase of the mode at each time step is determined by calculating the angle between the vector (RC_l, $RC_l^{'}$) and the vector $(0, 1)$. These phases, in the interval $(0, 2\pi)$, are then classified into eight equally populated categories. Composite maps are constructed from these categories.

3 Results

3.1 Reanalysis decompositions

The main outcome of the RMSSA method, the space-time principal components (ST-PCs) characterise both the spatial and temporal structure of the modes of variability. Sections 3.1 – 3.4 focus on their temporal aspects. The leading 30 ST-PC time series and the corresponding power spectra are displayed in Figure 2 for 20CR and ERA-20C, ordered according to the explained variance. We can see that

- components with predominantly multi-decadal periodicity (1, 2, 5, and 6) explain a total of 7.2 % and 5.9 % of the variance in 20CR and ERA-20C, respectively, with clear similarities in their time series and spectra
- multi-annual components (3, 4, 7, and 8) explain 4.2 % and 3.2 % of the variance in 20CR and ERA-20C, respectively
- there is a broad multi-annual peak centered at 5 yr and a narrower peak at 3.5 yr in both reanalyses; these are clearly separated in ERA-20C at the components 3 and 4 versus 7 and 8. This separation in 20CR is less clear
- there are many spectral peaks in the reanalyses at 2–3 year periods with little explained variance but some are well separated and distinct

The conclusion based on Figure 2 is that the leading sources of the near-surface air temperature variability at multi-decadal and multi-annual periods are well identifiable in the reanalysis data sets. 20CR and ERA-20C are composed of very similar components explaining the variance in the two data sets. This is of course expected but it is also reassuring from the methodological view point: despite its complexity, the RMSSA decomposition is consistent.

It is noteworthy in Figure 2 that the components 3, 4, 7, and 8 in both reanalyses have become more prominent with time. Since about the 1970's, the amplitudes of these 3.5 yr and 5 yr oscillations have been at a higher level, presumably due to some combination of forced climate change, intrinsic low-frequency climate variability, or changes in global observing network (the rather sudden increase in the amplitude seems to coincide with the onset of the modern era of satellite observations). This finding seems to be in support of e.g. Russell and Gnanadesikan (2014). In this connection it should be noted, however, that apparent low-frequency variations and changes in amplitude may simply arise from random fluctuations of the time series (Wunsch, 1999; Wittenberg, 2009). Back-projection of these components into the original grid representation (Figure 3), reveals that the components are indeed associated with the ENSO phenomenon and are geographically similar in 20CR and ERA-20C. In the snapshots from January 1987 and January 1998 (Figure 3), there is a typical El Niño pattern with positive anomalies in the equatorial Pacific Ocean, South-America, and northwestern North-America. These are associated with synchronous evolution of (i) a dipole structure in the western Antarctica with easterly motion, and (ii) a wave-train type pattern in the northernmost North-America with north-easterly motion. The components 3, 4, 7, and 8 thus represent a global phenomenon, with an increased amplitude in recent decades. These features are nicely depicted in our Youtube channel (https://www.youtube.com/watch?v=vehbT8fOHeM, https://www.youtube.com/watch?v=xG--SiUqqAI).

3.2 Reanalysis total spectra

Figure 4a shows the total spectrum for the reanalyses constructed from the ST-PCs, and their confidence intervals (dashed lines). As in the ST-PCs, there is most power in the slow modes. At periods of about 3.5 yr and 5 yr, there are the spectral peaks of the components 3, 4, 7, and 8. The dip at 1 yr reflects the removed annual cycle.

As Fig. 2 already suggests, the shape of the two spectra is remarkably similar in all time scales (Fig. 4a). This leaves hardly any doubt that the data assimilation systems of 20CR and ERA-20C extract observed information in a very similar manner. There are some differences, however. The spectral power in ERA-20C is systematically slightly higher than in 20CR. This difference is statistically significant at almost all time scales. This is most likely due to generally higher temperature variance in ERA-20C compared to 20CR, especially in the Southern Ocean and Arctic Ocean. Also, in 20CR, the 3.5 yr and 5 yr spectral peaks are relatively more pronounced than in ERA-20C.

Statistical significance tests are presented in Figs. 4b and 4c for 20CR and ERA-20C, respectively. The multi-annual periods (less than 7 yr) rising above the 95 % confidence interval (i.e., the red dots above the region covered by the vertical bars) are 3.5 yr, 3.6 yr, and 5.7 yr in 20CR and 3.6 yr, 5.2 yr, 5.5 yr, and 5.7 yr in ERA-20C. Thus, nearly the same periodicities rise above the red noise in the two data sets. It is logical that the frequency corresponding to the annual cycle is present in the red noise surrogates while it is absent from the data, and therefore the red dots fall far below their expected values. Interestingly, the period of 2.9 yr in 20CR and ERA-20C fall below the 95 % confidence interval. Our conclusion is therefore that the multi-annual climate variability in the near-surface air temperature is very similar in 20CR and ERA-20C.

3.3 CMIP5 model total spectra

The total spectra for the 12 CMIP5 model are shown in Fig. 5 (solid lines) with their 95 % confidence intervals (dashed envelopes) and the reanalysis spectra as a reference (thin lines). Statistically significant multi-annual modes (at 5 % level) are denoted by vertical dashed lines. As in the case of reanalyses, these spectra are unique expressions of the low-frequency variability present in the simulation data. A comparison between the simulated and the reanalysis spectra provides one means to assess the strengths and weaknesses of these models. However, one cannot simply rank the models based on how "far off" the model spectra are from the reference, because this comparison focuses on just one (although important) aspect of model performance and because seemingly good agreement with observations might occasionally result from compensating errors in model processes.

Here we will concentrate on the multi-annual aspects but note in passing that the level of multi-decadal variability (> 20 yr) is close to reanalyses in models a, c, d, e, and g. In the rest of the models, the level seems too low. In the decadal scale ($\sim$10 ... 20 yr), the level of variance is close to reanalyses in a, b, c, f, i, j, and l. Subjectively, the shape of the low-frequency end of the spectra appears most realistic in models a and c.

In multi-annual scales, the model performance varies a lot among the models. There is a group of models (a, b, d, and e) with high spectral density at about 3 – 7 yr periods. The models d and e have a bi-modal spectral structure, as in the reanalyses, while

models a and b have a broad unimodal peak. Decompositions (available in the Supplementary material, S1) partly explain the reasons leading to these total spectra.

In model a, for instance, there is a unimodal broad peak at 3.5 – 4 yr periods (Fig. 5a). The decomposition reveals that there are, in fact, two well separated component pairs at 3.5 yr and 4 yr generating one merged peak to the total spectrum (Fig. S1a in the Supplement). A development hint is thus to investigate these modes which can help to better understand some underlying modelling deficiencies, and to keep monitoring how this aspect of model performance evolves in the future model upgrades. An additional concern in model a is the excessive spectral density at about 2 yr and 7 – 10 yr periods.

In model e, there is a bimodal total spectrum (Fig. 5e), although far too pronounced as compared with the reanalyses. The decomposition (Fig. S1e in the Supplement) reveals that the ST-PC components 1 – 10 (except 7–8) are all multi-annual and peak strongly and well in isolation at 3 yr, 3.5 yr, 4 yr, and 5 yrs, explaining together no less than 13.9 % of the total variance. The development hint for model e is thus to investigate the mechanisms behind the components 1 – 10 and thereby obtain guidance for improving the realism of simulations.

In most other models, the multi-annual variability is less prominent than in the reanalyses. In model c (Fig. 5c), on one hand, the decomposition points out (Fig. S1c in the Supplement) that there are about 12 ST-PC components with periods between 1.5 – 3 yrs leading to a total spectrum with a broad peak of 2 – 3 yr periods. These components tend to have very regular cycles, remotely resembling a coupled harmonic oscillator and seemingly missing the "offbeats" or true quasi-periodicity of the reanalyses. The task seems to be to find out reasons why model c produces too rapid and regular multi-annual variability. In model g (Fig. 5g), on the other hand, the leading ST-PC components 1 – 9 are on either decadal or multi-decadal periods and these overwhelm the total spectrum. It should be important to find out the causes for this accentuated variability, especially on the decadal scale.

Finally, Fig. 5 casts light on models' overall level of variability compared to reanalyses. Clearly, this level in model h (Fig. 5h) is low. Curiously enough, the leading ST-PC component pair in model h explains only 1.4 % of variance and peaks at 3.2 yr. This corresponds to the isolated peak in the total spectrum.

3.4 Significance of multi-annual modes in CMIP5 models

In the reanalyses (Fig. 4), only a few multi-annual periods rise above the red noise (three in 20CR and four in ERA-20C). They are at approximately 3.5 yr and 5 yr periods. For the CMIP5 models, the test results are available in the Supplementary material (S2). In Fig. 5, the multi-annual modes with periods less than 7 yrs at the 5 % significance level are denoted by dashed vertical lines.

In summary, there are 5 – 15 statistically significant periods in the models, except model k (Fig. 5k) with three and model g (Fig. 5g) with zero periods. The large number of significant periods (models d and e, for instance) can be explained, at least partly by the fact that the modes are quasi-periodic and the spectral density therefore appears on a range of frequencies. This manifests as excursion of the red-noise threshold on several adjacent frequencies. This is typical for models with large spectral power on certain time scales. In model l (Fig. 5l), for instance, there are two broad and distinct spectral peaks at about 3.5 yr and 6 yr periods, and many significant periods are gathered at these and nearby frequencies. In contrast, models f and h (and

to some extent model c) have several significant and distinct periods between 2 yr and 7 yr. In terms of number of significant modes, models a, i, j, and k seem to be closest to the reanalyses.

3.5 Spatial patterns of the 3-4 yr mode

ST-PC components can be represented in the original coordinate system as so called reconstructed components (RC) that can be visualised. In this section, some visualisation results are presented and discussed.

In ERA-20C, there is a spectral peak at 3.5 yr period, which is significant at 5 % level (Fig. 4). This peak is due to the ST-PC components 7 and 8 with spectral density closely concentrated on this frequency (Fig. 2). Figure 6 depicts composite maps of each of the eight phases of the 3.5 yr mode in ERA-20C. Firstly, the mode is global with the largest temperature anomalies in the Pacific and North-America. Secondly, the mode contains tropical Pacific temperature anomalies, like in the ENSO phenomenon (e.g. Kleeman, 2008). The cold (warm) maximum is in phase 1 (5) with the anomalies extending to the South-American continent. Thirdly, there are traveling temperature anomalies at high latitudes on both hemispheres. These are described next.

In phase 1 (Fig. 6), there is a small warm temperature anomaly in the North-Pacific (lon 160°W, lat 30°N). This pattern slowly moves northeast reaching Alaska in phase 5 and then gradually dissipating over the northernmost North-America in phase 8 (and being visible still in phase 1). There is a very similar evolution of a cold anomaly starting in phase 5. At the same time, there is an oscillating temperature anomaly over the Eurasian continent in opposite phase. In Fig. 6, there is also a traveling temperature anomaly in the Southern Hemisphere. In phase 8 (Fig. 6), there is a cold anomaly over the Southern Ocean (lon 160°W). This strengthens, moves east, weakens, and crosses the Antarctic Peninsula in phase 4 and remains in the Weddell Sea until phase 7. Similarly, there is a warm anomaly in phase 4 (lon 160°W) with similar evolution as the cold one.

Next, 20CR and the CMIP5 model behaviour is studied. The 3.5 yr mode is significant in 20CR and ERA-20C. For the illustration, we have chosen component pairs from the model decompositions (Supplementary material Fig. S1) that have spectral peaks between 3 and 4 years and do not express substantial variability on other time scales. In most climate models, such a corresponding mode exists, except in models g and k. In model c this mode is not significant at 5 % level, but it is illustrated anyway. Supplementary material reveals how these modes are represented in different data sets (Fig. S3–S14). The format is the same as in Fig. 6. A short summary is presented next.

In 20CR (Fig. S3), the anomalies are weaker compared to ERA-20C (Fig. S4). This is mainly because the 3 – 4 yr mode is distributed on two component pairs in 20CR whereas in ERA-20C it is concentrated on one pair. Nevertheless, similar although weaker signal is evident in 20CR, such as the northeast propagation of the North-Pacific temperature anomaly. (Note that in Fig. 3, the combination of components 3, 4, 7, and 8 produce highly similar global patterns for 20CR and ERA-20C.) A prominent feature is also the opposite temperature anomalies in the northern Eurasia versus North-America. All models (Figs. S5–S14) produce a temperature anomaly to the equatorial Pacific Ocean (and South-America). The amplitude is larger and/or the area extends further to the west than in ERA-20C in six models (a, b, d, e, h, l). The anomaly pattern in the northwestern North-America is present in all the models to some extent. In the reanalyses, the anomaly is strictly confined to land areas but in most models, it is either somewhat misplaced or extends to the adjacent sea areas and the Eurasian continent. Models c, e,

and f produce the North-American pattern quite similar to reanalyses, and the northeast propagation is captured to some extent by models b, c, f, i, and l.

4 Discussion

We note that a substantial portion of variance at inter-annual to inter-decadal timescales can be attributed to "climate noise" associated with processes with timescales much shorter than the inter-annual scale (Wunsch 1999; Feldstein 2000). If the amplitude of the variability mode exceeds some noise threshold (such as red noise), then the variability mode is also likely driven by some process external to the atmosphere, in addition to the climate noise. For example, large part of the inter-annual atmospheric ENSO pattern is presumably driven by anomalies of tropical diabatic heating associated with sea surface temperature anomalies (Feldstein, 2000). We assume that for this reason the multi-annual patterns related to ENSO clearly exceed the noise threshold in the results of this study.

5 Conclusions

The aim of this study is to decompose the 20th century climate variability into its multi-annual modes, and to assess how these modes are represented by the contemporary climate models. To this end, two 20th century reanalysis data sets and 12 CMIP5 model simulations for years 1901–2005 of the monthly mean near-surface air temperature have been decomposed using Randomised Multi-Channel Singular Spectrum Analysis (RMSSA). The statistical significance of the identified modes has been estimated with Monte Carlo simulations. The main conclusions are as follows.

Spectral properties of the 20CR and ERA-20C reanalysis data appear remarkably similar. The most prominent forms of variability in both data sets are related to approximately 3.5 yr and 5 yr modes which are significant at 5 % level. The spectral power in ERA-20C is systematically slightly higher than in 20CR. The 3.5 yr mode is illustrated in more detail. In ERA-20C, the mode is associated with typical ENSO pattern of temperature anomalies in the equatorial Pacific Ocean, South-America, and northwestern North-America. On top of these, the mode also contains a northeast propagating temperature anomaly over the northernmost North-America, and another eastward propagating anomaly in the vicinity of western Antarctica. Since about the 1970's, the amplitude of this 3.5 yr global mode have increased.

None of the 12 coupled climate models closely reproduce all aspects of the reanalysis spectra, although some models represent many aspects well. For instance, the GFDL-ESM2M model has two nicely separated ENSO -related periods although they are relatively too prominent as compared to the reanalyses. Also, a number of models represent the propagating temperature anomalies at 3 – 4 yr time frame. Some suggestions are provided in the text for potential model development aspects.

There is an extensive Supplement available presenting the results in visual format for each reanalysis and model data set. In the future, relaxation of the uni-variate nature of the present study would seem a natural extension. This is now possible since the use of random projections allow efficient data structures preserving compression. Of special interest would be to study

behaviour of variables directly linked with atmosphere-ocean coupling processes, such as heat, momentum, and moisture fluxes over oceans.

6 Data and code availability

All data used in this study was downloaded from open sources. The RMSSA algorithm and the statistical significance testing are implemented using GNU licensed free software from the R Project for Statistical Computing (www.r-project.org). Our implementation is available on request. The animations of the 3–4 yr mode are available for all data sets at https://www.youtube.com/channel/UCu1zJdwJfLaXvfvTqsKCLHw.

Appendix A: Randomised multi-channel singular spectrum analysis (RMSSA)

The RMSSA algorithm and the significance test is briefly presented here. The original data matrix is $\mathbf{X}_{N\times L}$, where the columns are called *channels*. In case of gridded data set, N represents the time steps and L is the number of grid points. It is useful to think N as the time steps when the *sample* of dimension L is collected. The dimension reduction is a projection $\mathbf{X}_{N\times L} \rightarrow \mathbf{P}_{N\times k}$, where $L \gg k$. In other words, we preserve all samples but reduce the sample dimension from L to k. The dimension reduction is performed in two steps: (1) generate a random matrix $\mathbf{R}_{L\times k}$, where the matrix elements are $r_{ij} \sim N(0,1)$ and column vectors of $\mathbf{R}$ are normalised to unit length, and (2) project $\mathbf{X}$ onto $\mathbf{R}$:

$$\mathbf{P}_{N\times k} = \mathbf{X}_{N\times L}\mathbf{R}_{L\times k}. \tag{A1}$$

The next step is to construct an augmented data matrix $\mathbf{A}$, which contains M lagged copies of each channel in $\mathbf{P}$. In RMSSA, M represents the lag window. $\mathbf{A}$ now has Mk columns and $N' = N - M + 1$ rows. The singular value decomposition of $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{U}\mathbf{D}^{1/2}\mathbf{V}^T \tag{A2}$$

The vectors of $\mathbf{U}$ are the eigenvectors of $\mathbf{Z} = \frac{1}{Mk}\mathbf{A}\mathbf{A}^T$ and $\mathbf{V}^T$ contains the eigenvectors of $\mathbf{C} = \frac{1}{N'}\mathbf{A}^T\mathbf{A}$. These vectors are orthogonal and often called space-time principal components (ST-PCs) and space-time empirical orthogonal functions (ST-EOFs), respectively. Note that the ST-EOFs are now in reduced space k. Diagonal elements of $\mathbf{D}$ are the eigenvalues of $\mathbf{C}$ or $\mathbf{Z}$. Finally, the eigenvectors (ST-EOFs) are calculated in the original L-dimensional space by

$$\mathbf{V} \approx \mathbf{A}_o^T\mathbf{U}(\mathbf{D}^{1/2})^{-1}, \tag{A3}$$

where $\mathbf{A}_o$ is the augmented matrix of the original data matrix $\mathbf{X}$. Note that the calculation of ST-EOFs in Eq. (A3) can be limited only to the eigenmodes of interest.

The ST-PCs can be represented in the original coordinate system by the reconstructed components, RCs (Plaut and Vautard, 1994; Ghil et al., 2002). This transformation is given by

$$rc_{le}(n) = \frac{1}{M_n} \sum_{m=I_n}^{J_n} u_e(n-m+1) v_{le}(m), \tag{A4}$$

where u_e are the ST-PCs and v_{le} are the ST-EOFs calculated in Eq. (A3) (the part of ST-EOF corresponding to channel l). e is the index of the eigenmode that is calculated. The normalisation factor M_n and the summation bounds I_n and J_n are given in Ghil et al. (2002) and for the central part of the time series ($M \leq n \leq N - M + 1$) they are $(M, 1, M)$, respectively.

RMSSA with significance testing is briefly presented in the following. Testing the MSSA components against a red-noise null-hypothesis requires orthogonal input vectors, which are obtained by calculating first a conventional PCA and retaining a set of dominant PCs. Therefore some additional calculation steps are included in the RMSSA-algorithm:

SVD of lower dimensional matrix $\mathbf{P}$ is calculated to obtain the principal components (PCs, calculated as $\mathbf{UD}^{1/2}$). PCs fullfil the orthogonality constraint exactly. PCs, that explain large part of the variance of the data set (e.g. 50 first), are retained to obtain matrix $\mathbf{T}$, where the columns are the PCs. Next, the augmented matrix $\mathbf{A}_{PC}$ is constructed from $\mathbf{T}$ and SVD is calculated as in Eq. (A2).

Finally, a large number of red-noise processes (i.e. surrogate data sets) are generated, and the confidence limits for the MSSA eigenmodes are determined. This signicance test (Monte Carlo MSSA) is described in detail in Allen and Robertson (1996).

Author contributions. HJ suggested the study and mostly wrote the article. TS implemented the methods, performed all computations, and wrote data and method descriptions, JS supported the method development, and JR the climate model data analysis.

Acknowledgements. This research has been funded by the Academy of Finland (project number 140771), Academy of Finland Centre of Excellence Programme (project number 272041), and the Fortum Foundation (grant number 201500127).

References

Achlioptas, D.: Database-friendly random projections: Johnson-Lindenstrauss with binary coins, Journal of computer and System Sciences, 66(4), 671–687, 2003.

Allen, M. R. and Robertson, A. W.: Distinguishing modulated oscillations from coloured noise in multivariate datasets, Climate Dynamics, 12(11), 775–784, 1996.

Ba, J., Keenlyside, N. S., Latif, M., Park, W., Ding, H., Lohmann, K., Mignot, J., Menary, M., Otterå, O. H., Wouters, B., Salas y Melia, D., Oka, A., Bellucci, A. and Volodin, E.: A multi-model comparison of Atlantic multidecadal variability, Climate Dynamics, 9, 2333–2348, 2014.

Bellenger, H., Guilyardi, E., Leloup, J., Lengaigne, M. and Vialard, J.: ENSO representation in climate models: from CMIP3 to CMIP5, Climate Dynamics, 42, 1999–2018, 2014.

Bingham, E. and Mannila, H.: Random projection in dimensionality reduction: applications to image and text data, in: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '01, ACM, New York, 245–250, 2001.

Broomhead, D. S., and King, G. P.: Extracting qualitative dynamics from experimental data, Physica D, 20, 217–236, 1986a.

Broomhead, D. S., and King, G. P.: On the qualitative analysis of experimental dynamical systems, in: Nonlinear Phenomena and Chaos, S. Sarkar (Ed.), Adam Hilger, Bristol, England, 113–144, 1986b.

Clement, A., Bellomo, K., Murphy, L. N., Cane, M. A., Mauritsen, T., Rädel, G. and Stevens B.: The Atlantic Multidecadal Oscillation without a role for ocean circulation, Science, 350, 320–324, doi: 10.1126/science.aab3980, 2015.

Cleveland, R.b., Cleveland, W.S., McRae, J.E. and Terpenning, I.: STL: A Seasonal-Trend Decomposition Procedure Based on Loess, Journal of Official Statistics, 6, 3–73, 1990.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D. and Worley, S. J.: The Twentieth Century Reanalysis Project, Quarterly J. Roy. Meteorol. Soc., 137, 1–28, 2011.

Feldstein, S. B.: The timescale, power spectra, and climate noise properties of teleconnection patterns, J. Climate, 13, 4430-4440, 2000.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C. and Rummukainen, M: Evaluation of Climate Models, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, [Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P.M. (Eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Fredriksen, H.-B. and Rypdal, K.: Spectral Characteristics of Instrumental and Climate Model Surface Temperatures, J. Climate, 29, 1253–1268, doi: http://dx.doi.org/10.1175/JCLI-D-15-0457.1, 2016.

Ghil, M.: Natural climate variability, in: Encyclopedia of Global Environmental Change, T. Munn (Ed.), Vol. 1, J. Wiley & Sons, Chichester/New York, pp. 544–549, ISBN 0-471-97796-9, 2002.

Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K., Kondrashov, D., Mann, M. E., Robertson, A.W., Saunders, A., Tian, Y., Varadi, F. and Yiou, P.: Advanced spectral methods for climatic time series, Reviews of geophysics, 40(1), 1–41, 2002.

Groth, A. and Ghil, M.: Monte Carlo Singular Spectrum Analysis (SSA) Revisited: Detecting Oscillator Clusters in Multivariate Datasets, Journal of Climate, 28(19), 7873–7893, 2015.

Kleeman, R.: Stochastic theories for the irregularity of ENSO, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 366(1875), 2509–2524, 2008.

Knutson, T. R., Zeng, F. and Wittenberg, A. T.: Multimodel Assessment of Regional Surface Temperature Trends: CMIP3 and CMIP5 Twentieth-Century Simulations, J. Climate, 26, 870–8743, 2013.

Mann, M. E. and Lees J. M.: Robust estimation of background noise and signal detection in climatic time series, Clim. Change, 33, 409–445, 1996.

Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., Corti, S., Danabasoglu, G., Doblas-Reyes, F., Hawkins, E., Karspeck, A., Kimoto, M., Kumar, A., Matei, D., Mignot, J., Msadek, R., Navarra, A., Pohlmann, H., Rienecker, M., Rosati, T., Schneider, E., Smith, D., Sutton, R., Teng, H., van Oldenborgh, G., Vecchi, G. and Yeager, S.: Decadal Climate Prediction: An Update from the Trenches, Bull. Amer. Meteor. Soc., 95, 243–267, 2014.

North, G. R., Bell, T. L., Cahalan, R. F. and Moeng, F. J.: Sampling errors in the estimation of empirical orthogonal functions, Monthly Weather Review, 110(7), 699–706, 1982.

Plaut, G. and Vautard, R.: Spells of low-frequency oscillations and weather regimes in the Northern Hemisphere, Journal of the atmospheric sciences, 51(2), 210–236, 1994.

Poli, P., Hersbach, H., Tan, D., Dee, D., Thepaut, J. N., Simmons, A., Peubey, C., Laloyaux, P., Komori, T., Berrisford, P., Dragani, R., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L. and Fisher, M.: The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C), ERA Report Series, ECMWF, 2013.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, Journal of Geophysical Research: Atmospheres, 108(D14), 446–469, 2003.

Rayner, N. A., Brohan, P., Parker, D. E., Folland, C. K., Kennedy, J. J., Vanicek, M., Ansell, T. J., and Tett, S. F. B.: Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The HadSST2 dataset, Journal of Climate, 19(3), 446–469, 2006.

Russell, A. M. and Gnanadesikan, A.: Understanding Multidecadal Variability in ENSO Amplitude, Journal of Climate, 27, 4037–4051, doi:10.1175/JCLI-D-13-00147.1, 2014.

Seitola, T., Mikkola, V., Silen, J. and Järvinen, H.: Random projections in reducing the dimensionality of climate simulation data, Tellus A, 66, doi:http://dx.doi.org/10.3402/tellusa.v66.25274, 2014.

Seitola, T., Silén, J. and Järvinen, H.: Randomised multichannel singular spectrum analysis of the 20th century climate data, Tellus A, doi:http://dx.doi.org/10.3402/tellusa.v67.28876, 2015.

Slingo, J. and Palmer, T.: Uncertainty in weather and climate prediction, Phil. Trans. R. Soc. A., 369, 4751–4767, 2011.

Stevenson, S., Fox-Kemper, B., Jochum, M., Rajagopalan, B. and Yeager, S. G.: ENSO model validation using wavelet probability analysis. Journal of Climate, 23(20), 5540–5547, 2010.

Taylor, K. E., Stouffer, R. J. and Meehl, G. A.: An Overview of CMIP5 and the experiment design, Bull. Amer. Meteor. Soc., 93, 485–498, 2012.

Thompson, D. W. J., Barnes, E. A., Deser, C., Foust, W. E., and Phillips, A. S.: Quantifying the role of internal climate variability in future climate trends. J. Climate, 28, 6443-6456, 2015.

Thomson, D. J.: Spectrum estimation and harmonic analysis, Proc. IEEE, 70, 1055–1096, 1982.

Vautard, R. and Ghil, M.: Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series, Physica D: Nonlinear Phenomena, 35(3), 395–424, 1989.

Wittenberg, A. T.: Are historical records sufficient to constrain ENSO simulations?, Geophys. Res. Lett., 36, L12702, 2009.

Wunsch, C.: The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations, Bulletin of the American Meteorological Society, 80, 245-255, 1999.

Table 1. Climate models used in the study. For more details of the models, see Table 9.1. in Flato et al. (2013).

Model ID	Model name	Modeling center	Country
a	CanESM2	CCCMA	Canada
b	CESM1(CAM5)	NSF-DOE-NCAR	USA
c	CNRM-CM5-2	CNRM-CERFACS	France
d	CSIRO-Mk3.6.0	CSIRO-QCCCE	Australia
e	GFDL-ESM2M	NOAA GFDL	USA
f	GISS-E2-R	NASA GISS	USA
g	HadGEM2-ES	MOHC	UK
h	INM-CM4	INM	Russia
i	IPSL-CM5B-LR	IPSL	France
j	MIROC-ESM	JAMSTEC/AORI/NIES	Japan
k	MPI-ESM-MR	MPI-M	Germany
l	MRI-CGCM3	MRI/JMA	Japan

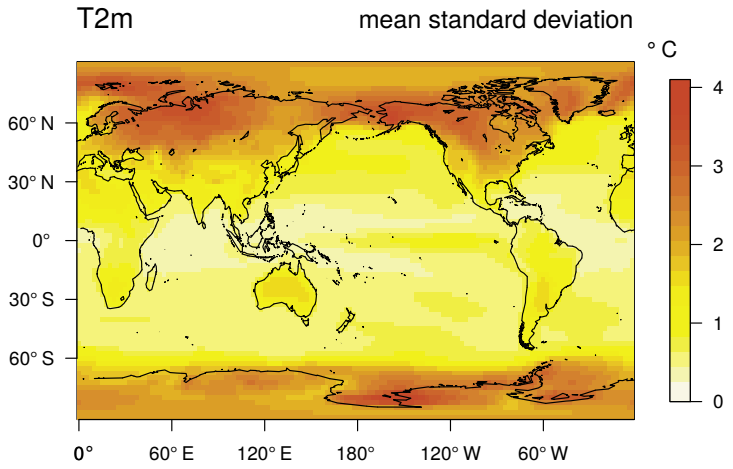


Figure 1. Map of the common normalisation factor. Shown is the mean standard deviation of 2 metre temperature (degC) across all the data sets.

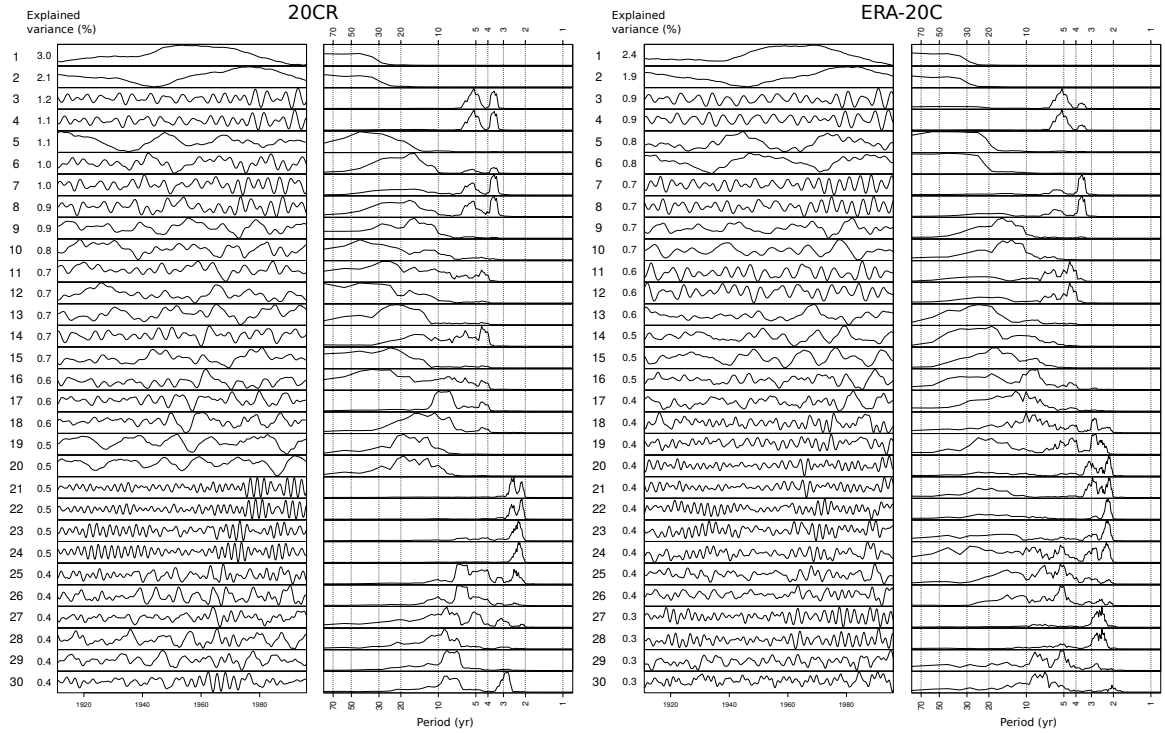


Figure 2. Reanalysis ST-PC time series (columns 1 and 3) of monthly near-surface temperature 1901–2005 and their spectra (columns 2 and 4) for 20CR and ERA-20C. The components are ordered according to the explained variance (%).

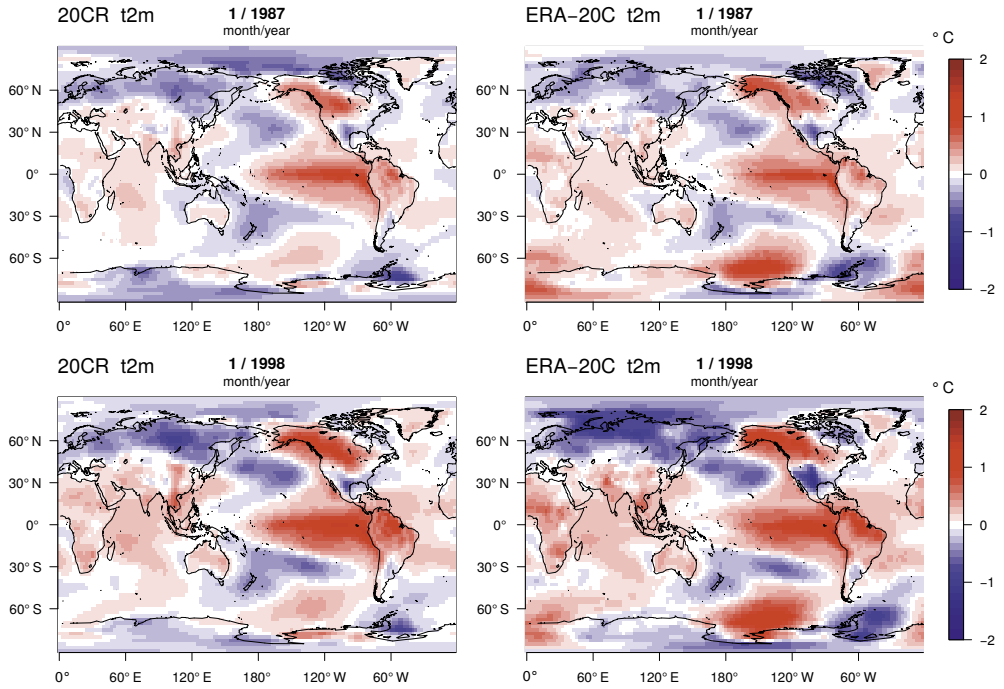


Figure 3. Global patterns of 2 metre temperature for the components 3, 4, 7 and 8 in 20CR (left column) and ERA-20C (right column). Snapshots are taken from Jan 1987 (top row) and Jan 1998 (bottom row). Unit degC.

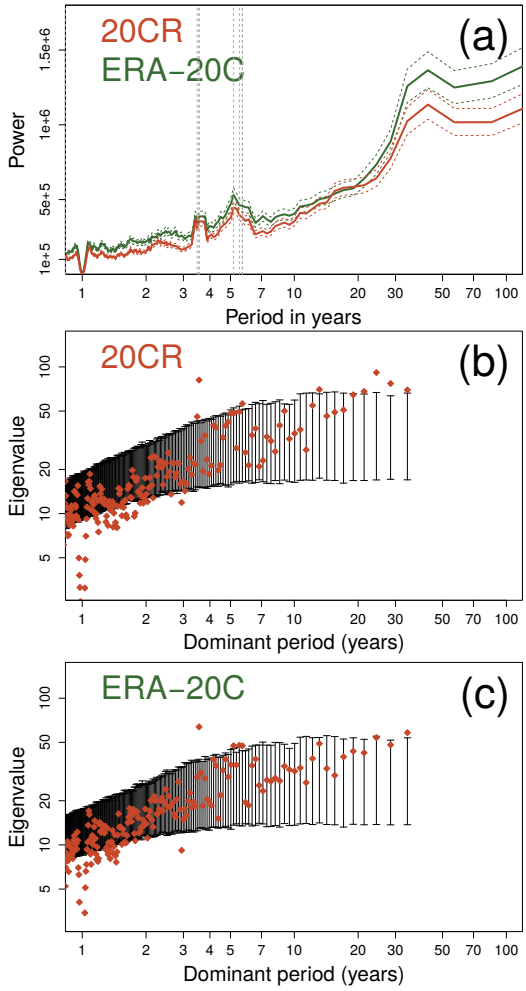


Figure 4. (a) Total spectrum of 20CR (red line) and ERA-20C (green line) with their min-max confidence intervals. The unit of the spectral density is arbitrary. (b) Significance of the 20CR periodicities against red-noise null-hypothesis. Shown are the data eigenvalues (red squares) and the 2.5^{th} and 97.5^{th} percentiles of the eigenvalue distribution of the red-noise surrogates (vertical bars). (c) Same as (b), but for the ERA-20C data set.

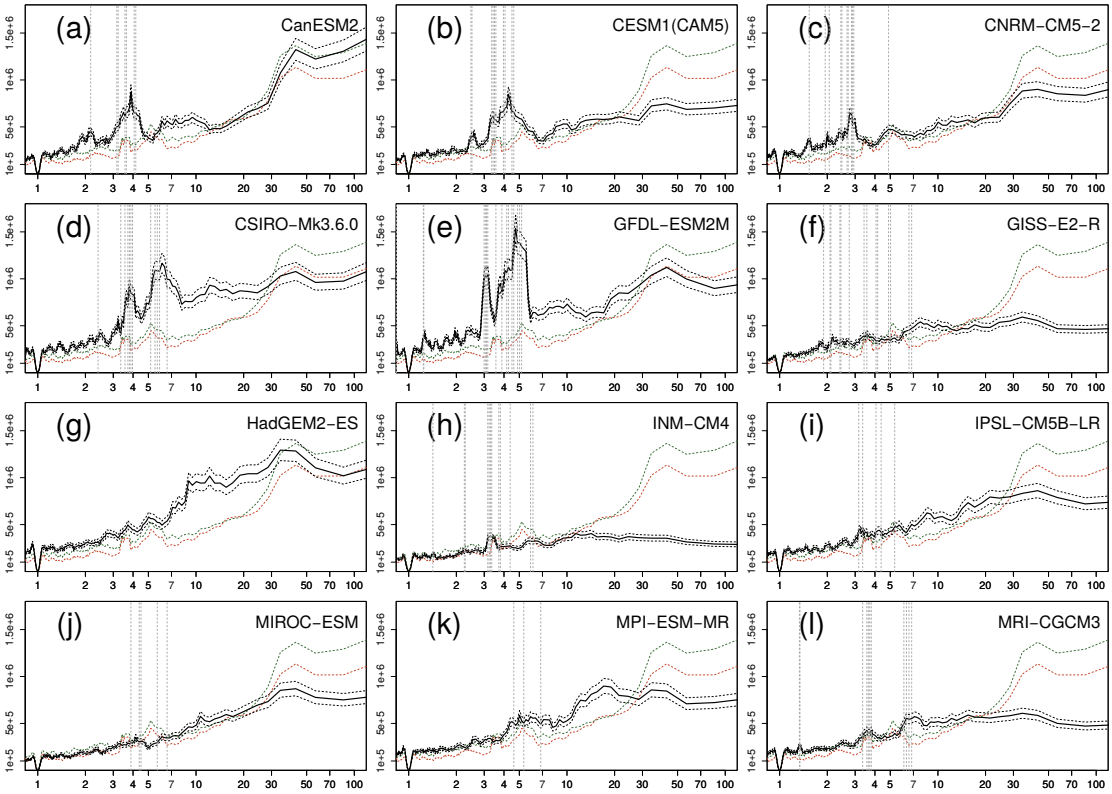


Figure 5. As Figure 4a but now for each climate model (black line). The reanalysis spectra are shown as a reference (dashed green and red lines). The dashed vertical lines indicate the climate model multi-annual periods significant at 5 % level.

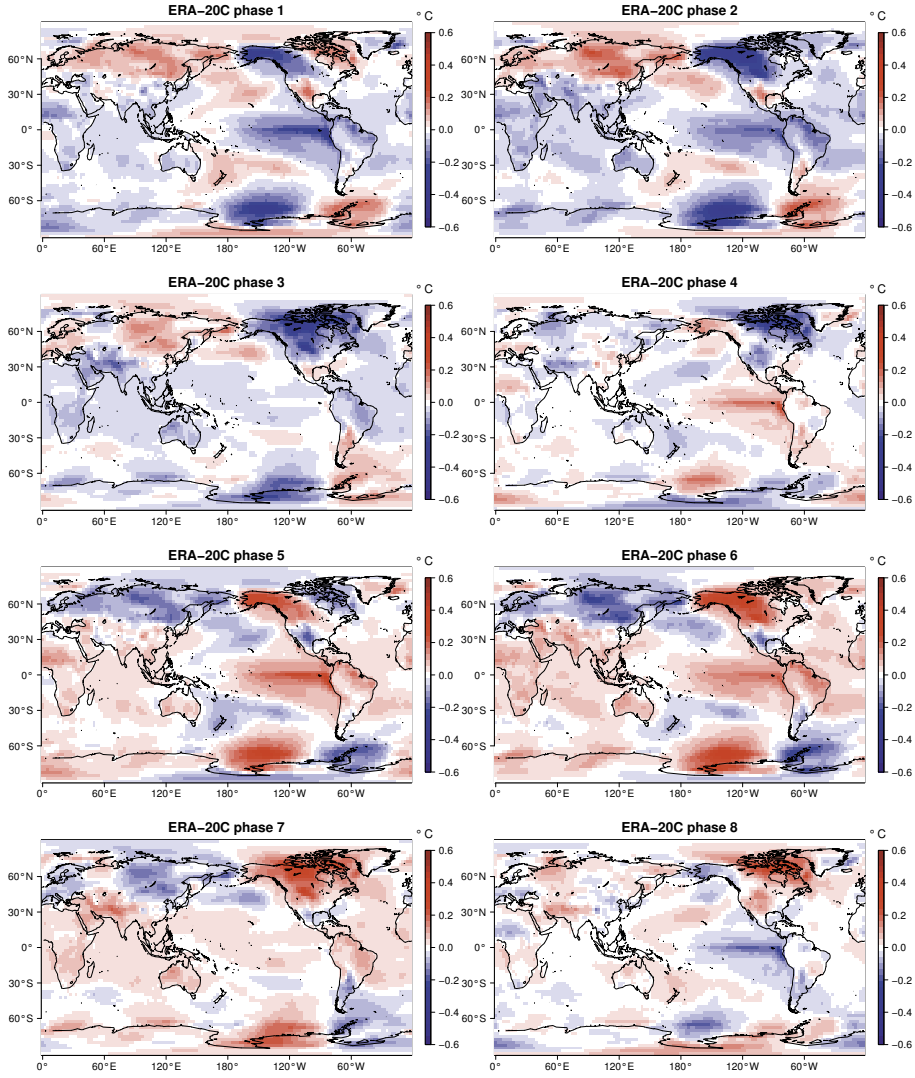


Figure 6. ERA-20C phase (1–8) composites of the 3–4 yr variability mode. Unit degC.

S1 Decompositions of models

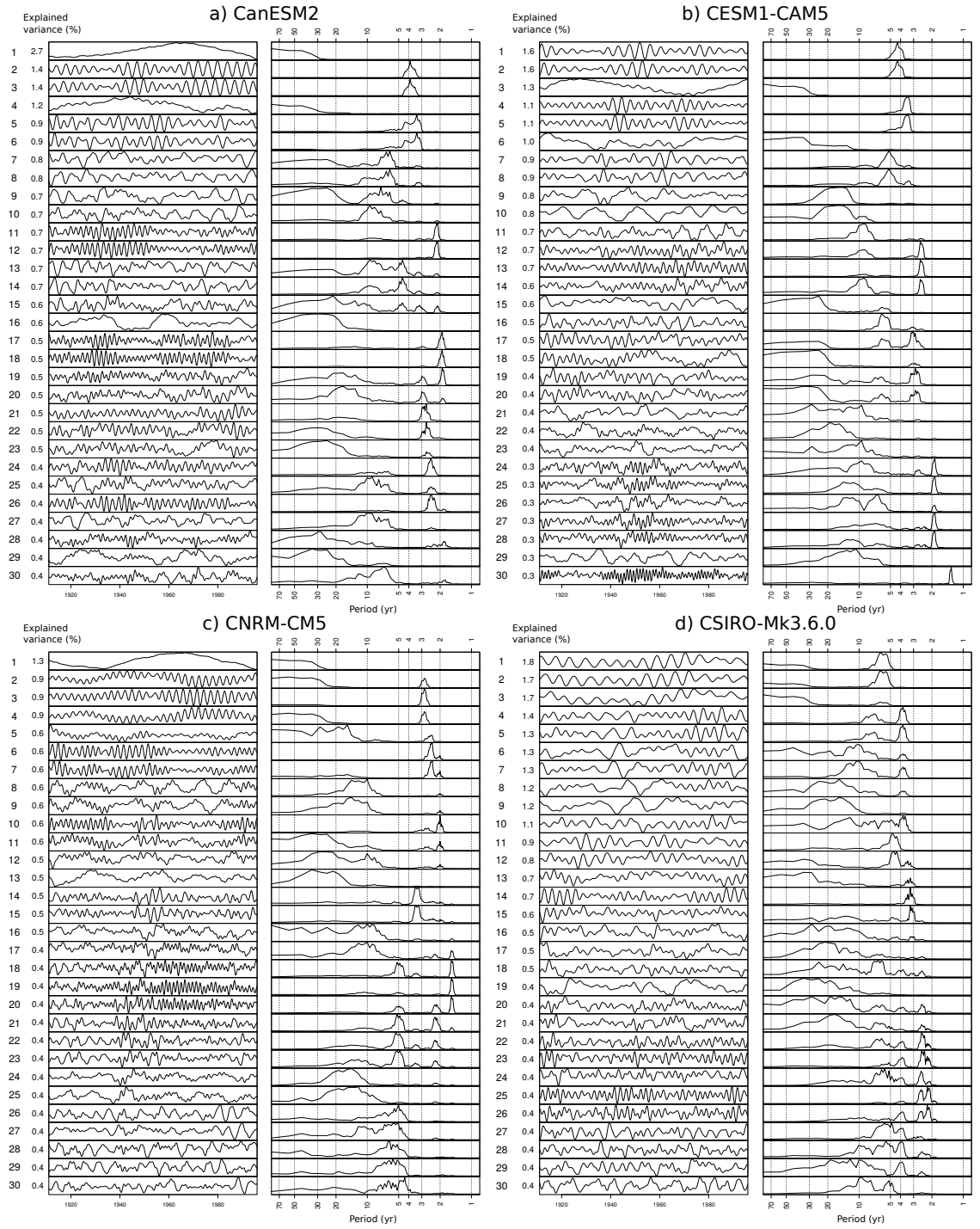


Figure S1: Models a, b, c and d: ST-PCs of monthly near-surface temperature 1901–2005 and their spectra. The components are ordered according to the explained variance (%).

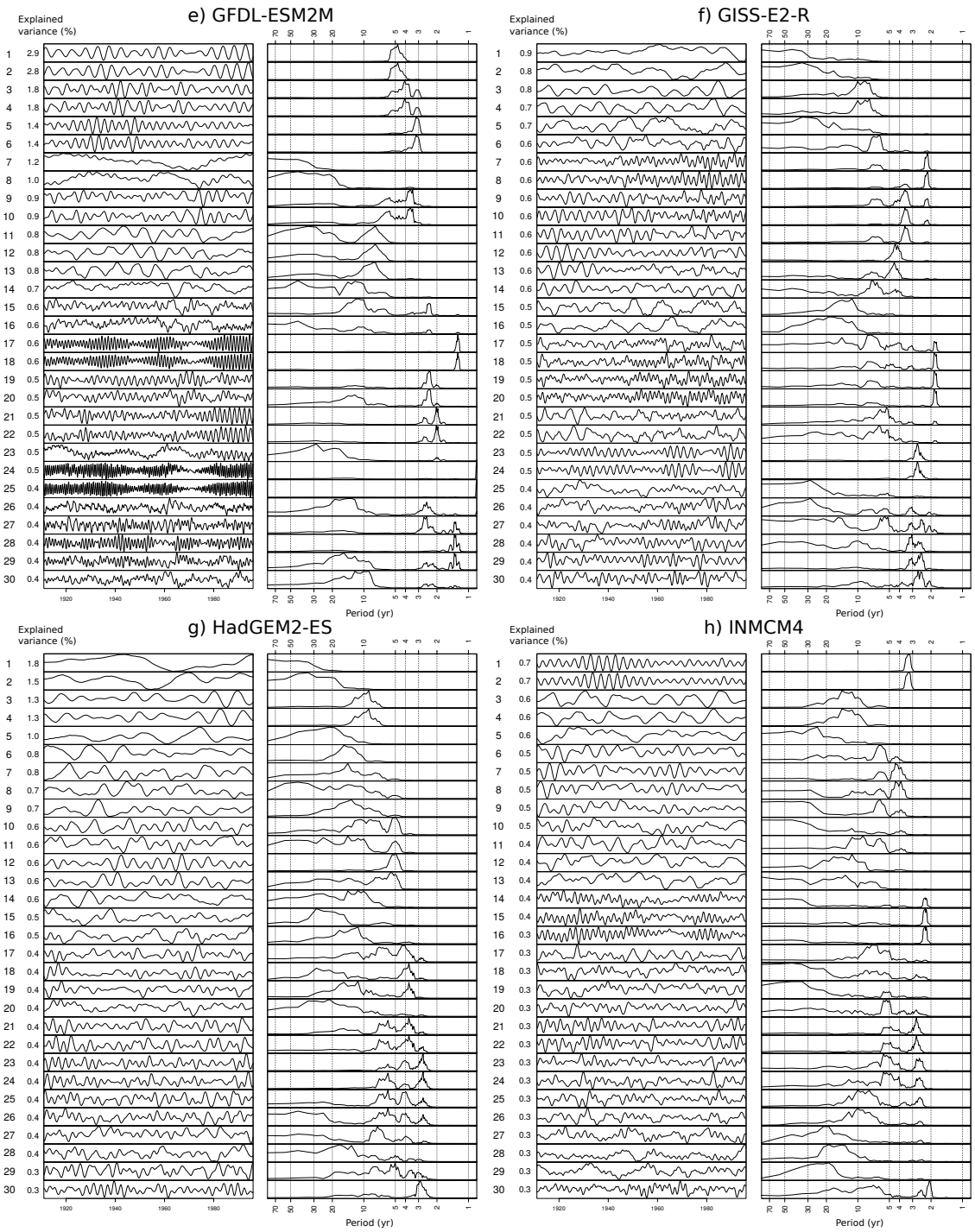


Figure S1 (continued): Models e, f, g and h.

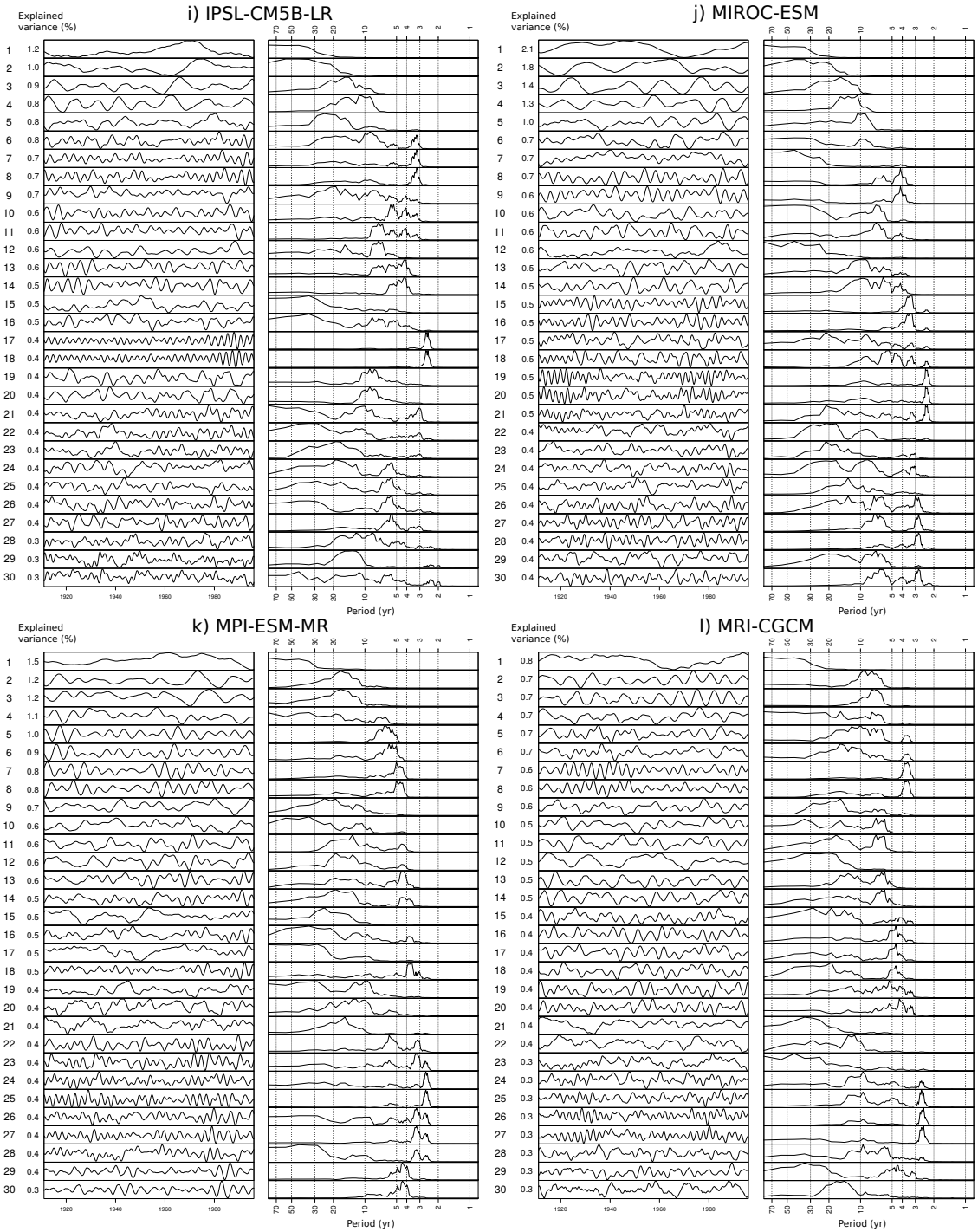


Figure S1 (continued): Models i, j, k and l.

S2 Significance tests of models

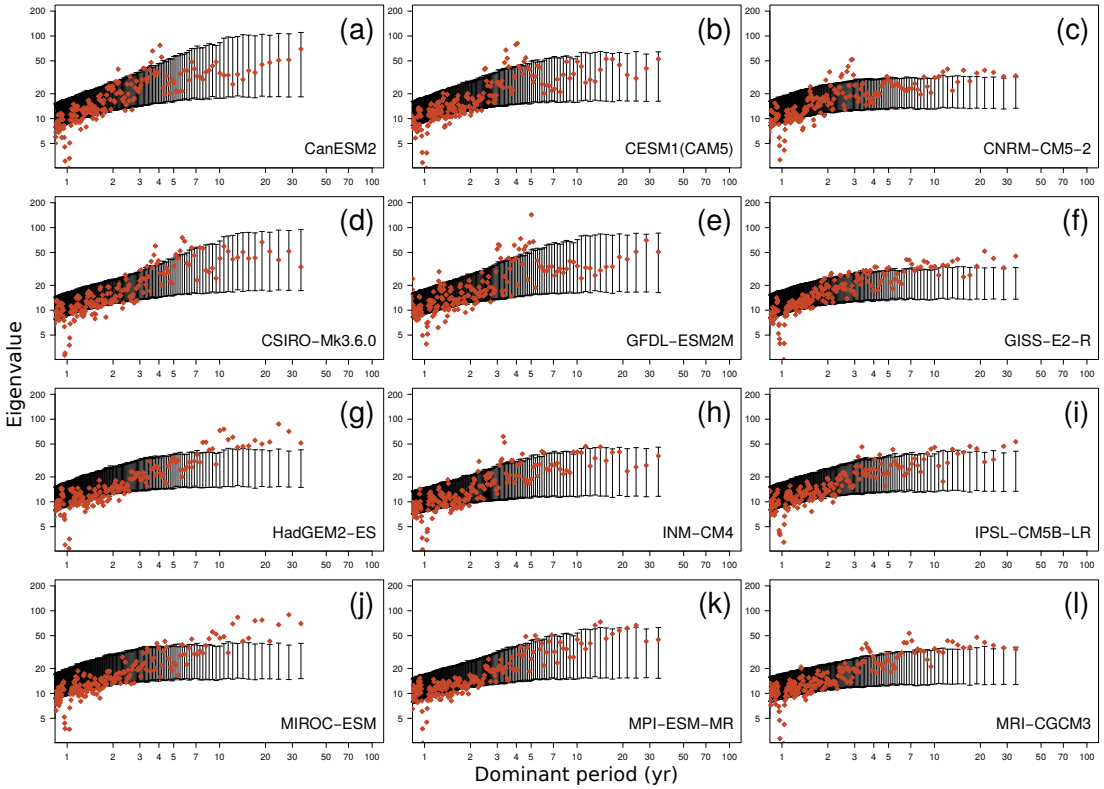


Figure S2: Significance test of the CMIP5 model eigenvalues against the red-noise null-hypothesis. Shown are the data eigenvalues (red squares) and the 2.5^{th} and 97.5^{th} percentiles of the eigenvalue distribution of the red-noise surrogates (vertical bars).

S3 Phase composites of the 3–4 yr variability modes

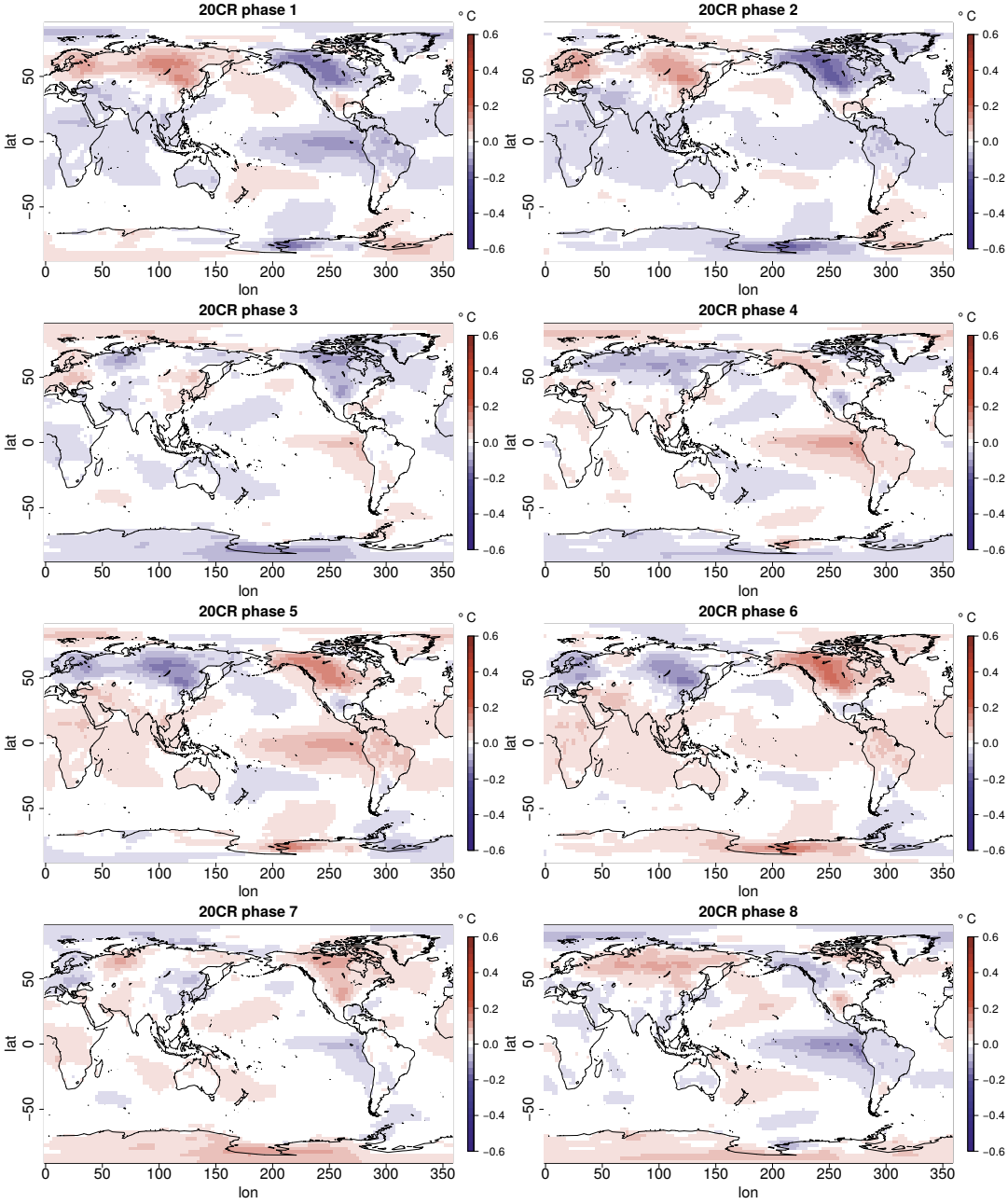


Figure S3: 20CR phase 1–8 composite figures of the 3–4 yr variability mode; ST-PC pair 7–8.

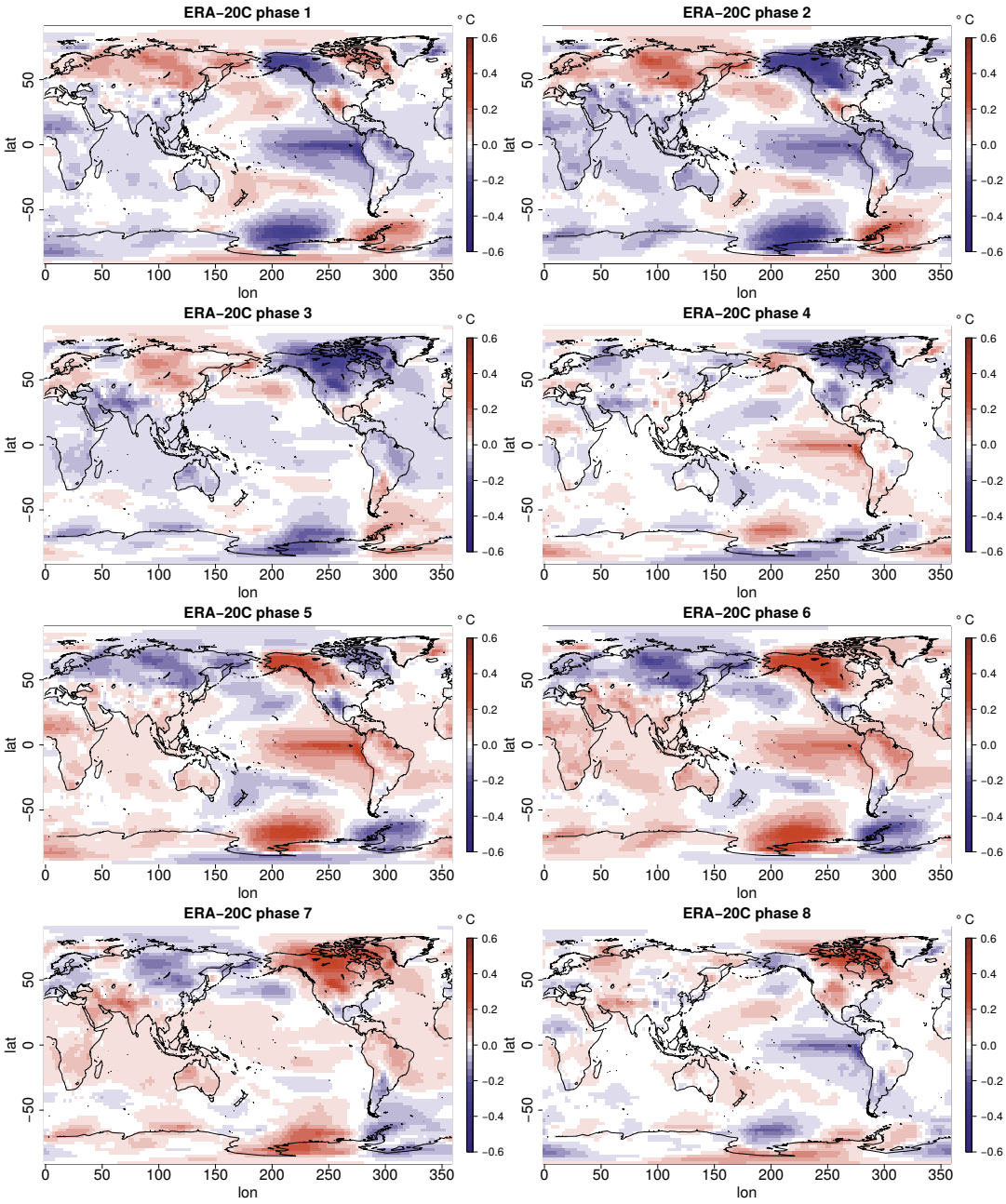


Figure S4: ERA-20C phase 1–8 composite figures of the 3–4 yr variability mode; ST-PC pair 7–8. (Same as Figure 6 in the article.)

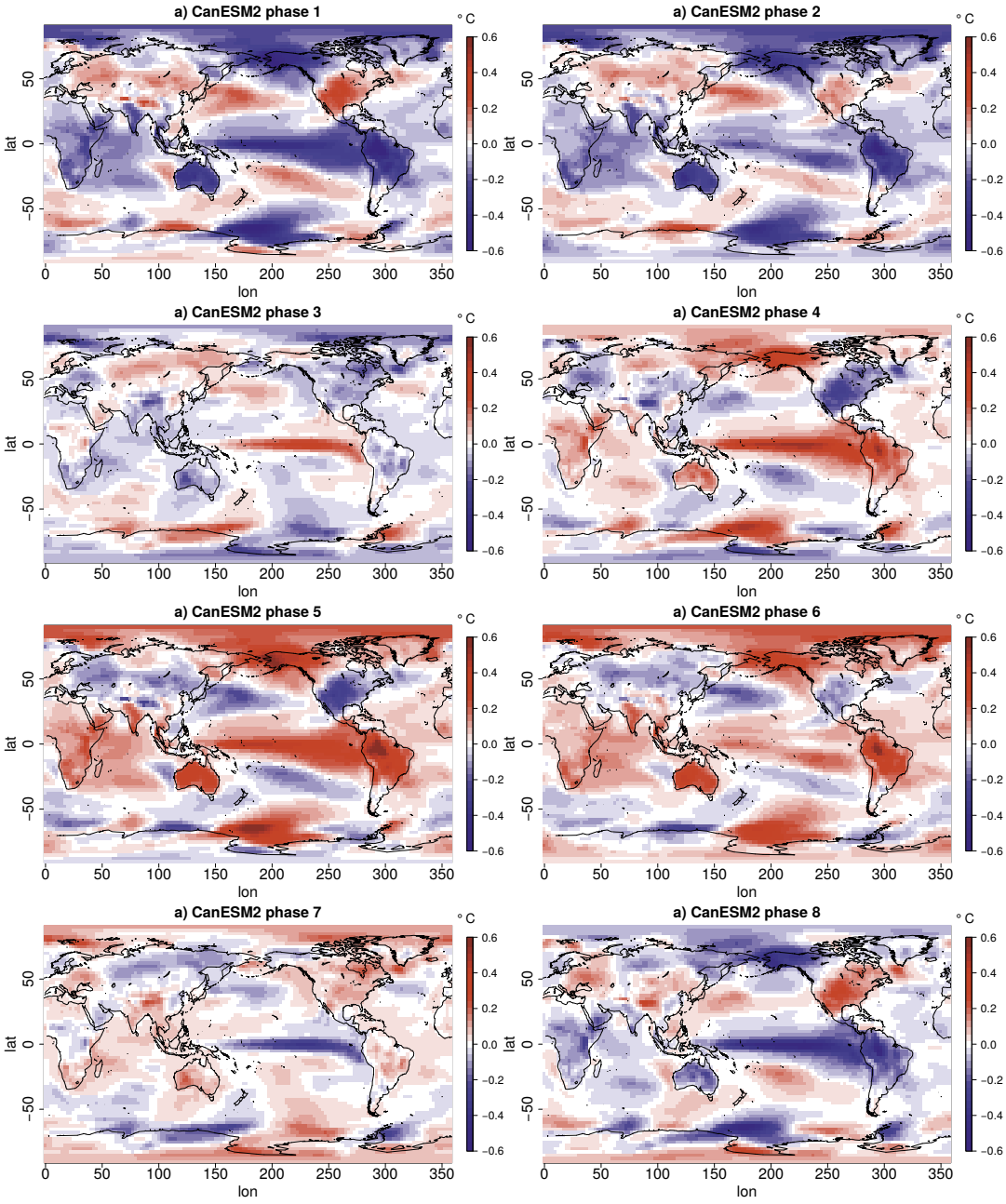


Figure S5: As Figure S3 but now for model a; ST-PC pair 2–3.

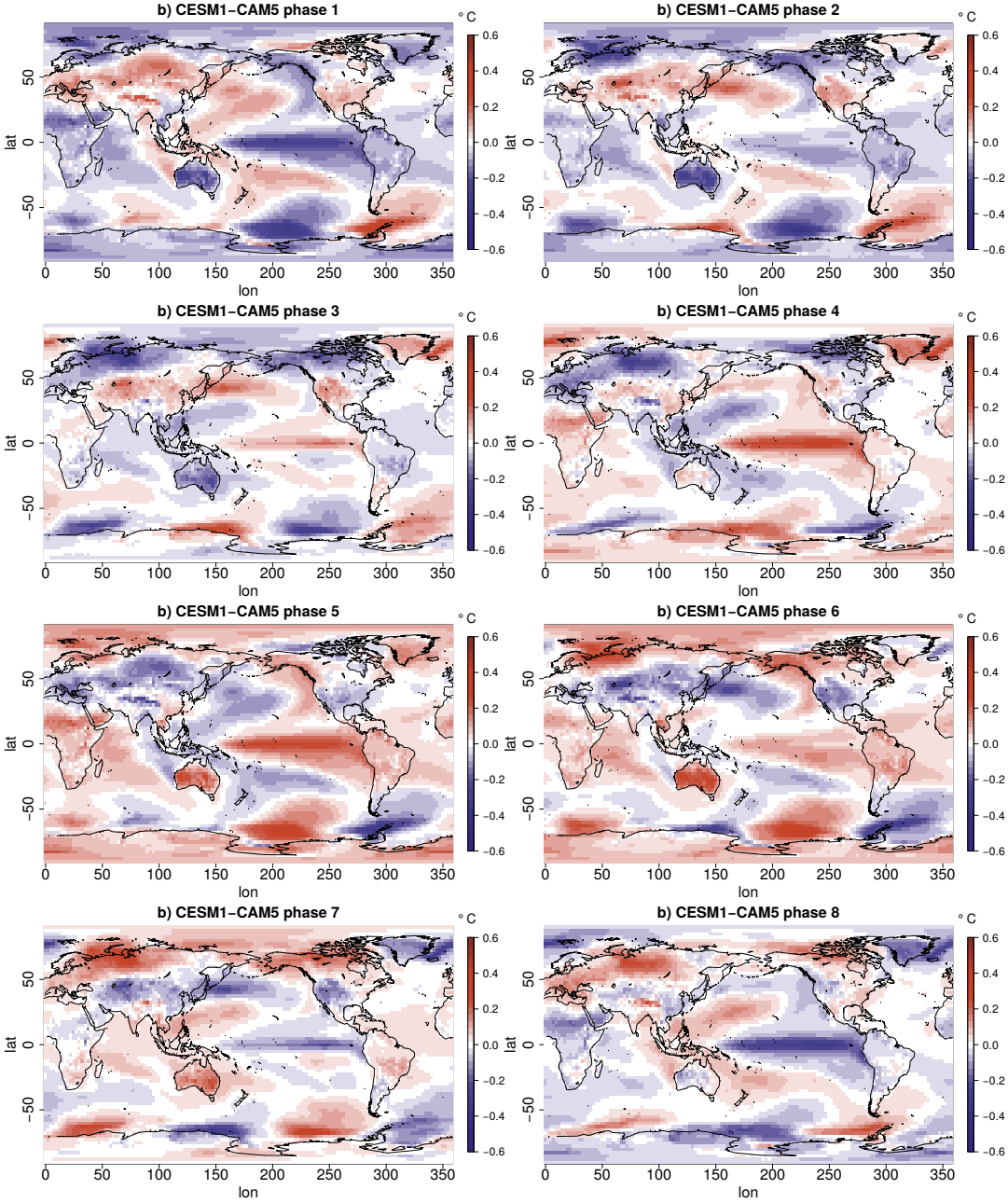


Figure S6: As Figure S3 but now for model b; ST-PC pair 4–5.

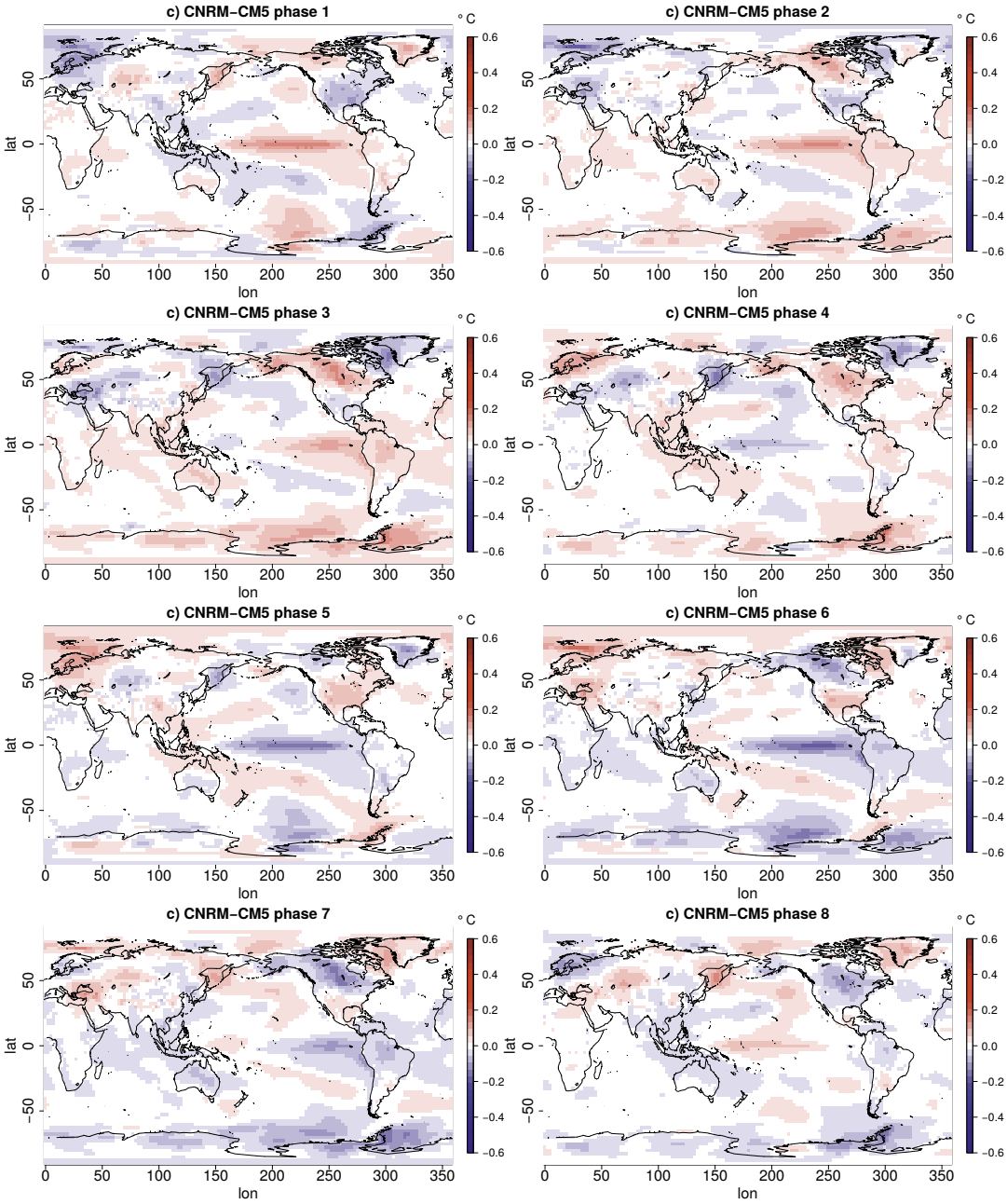


Figure S7: As Figure S3 but now for model c; ST-PC pair 14–15.

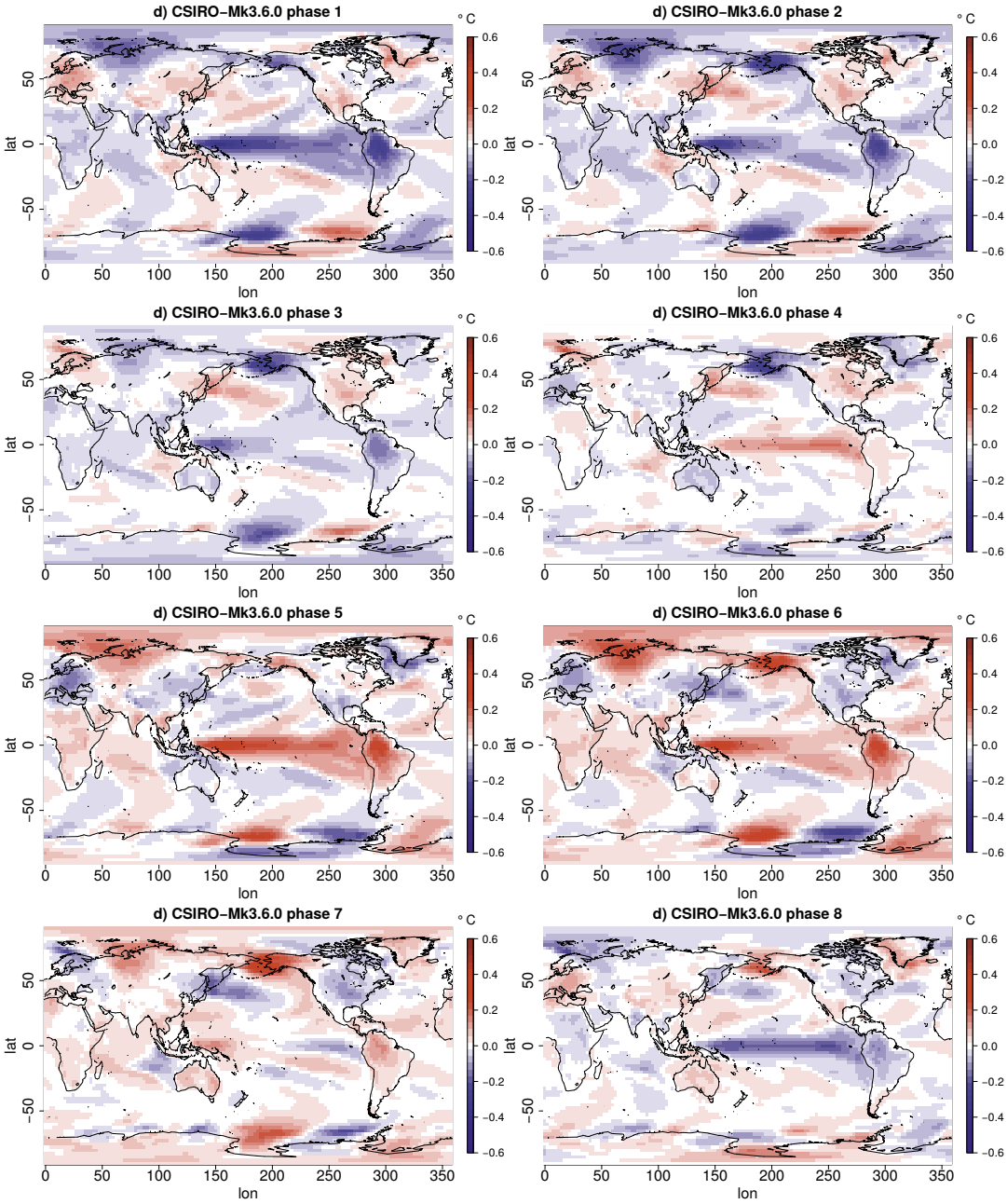


Figure S8: As Figure S3 but now for model d; ST-PC pair 4–5.

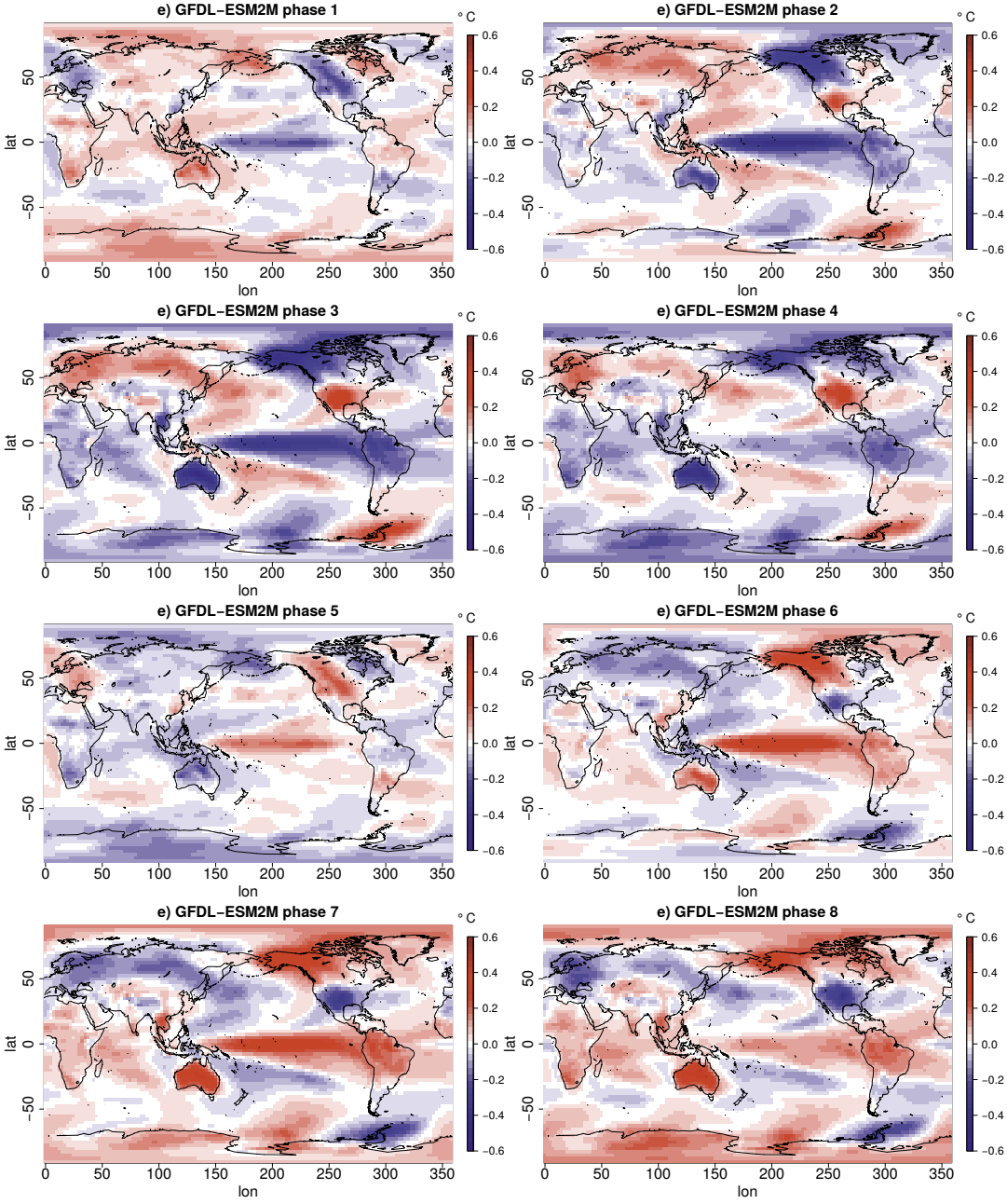


Figure S9: As Figure S3 but now for model e; ST-PC pair 5–6.

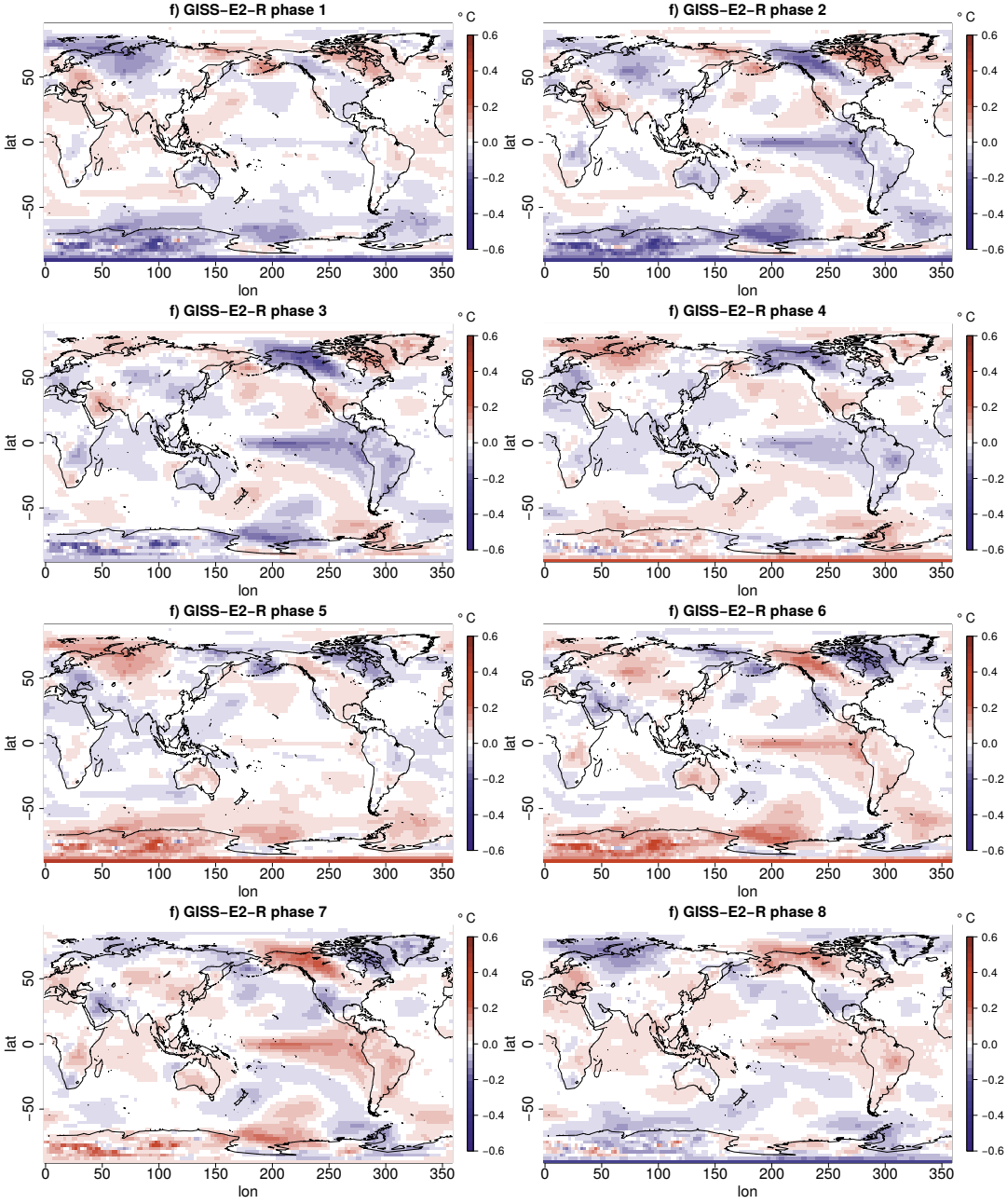


Figure S10: As Figure S3 but now for model f; ST-PC pair 10–11.

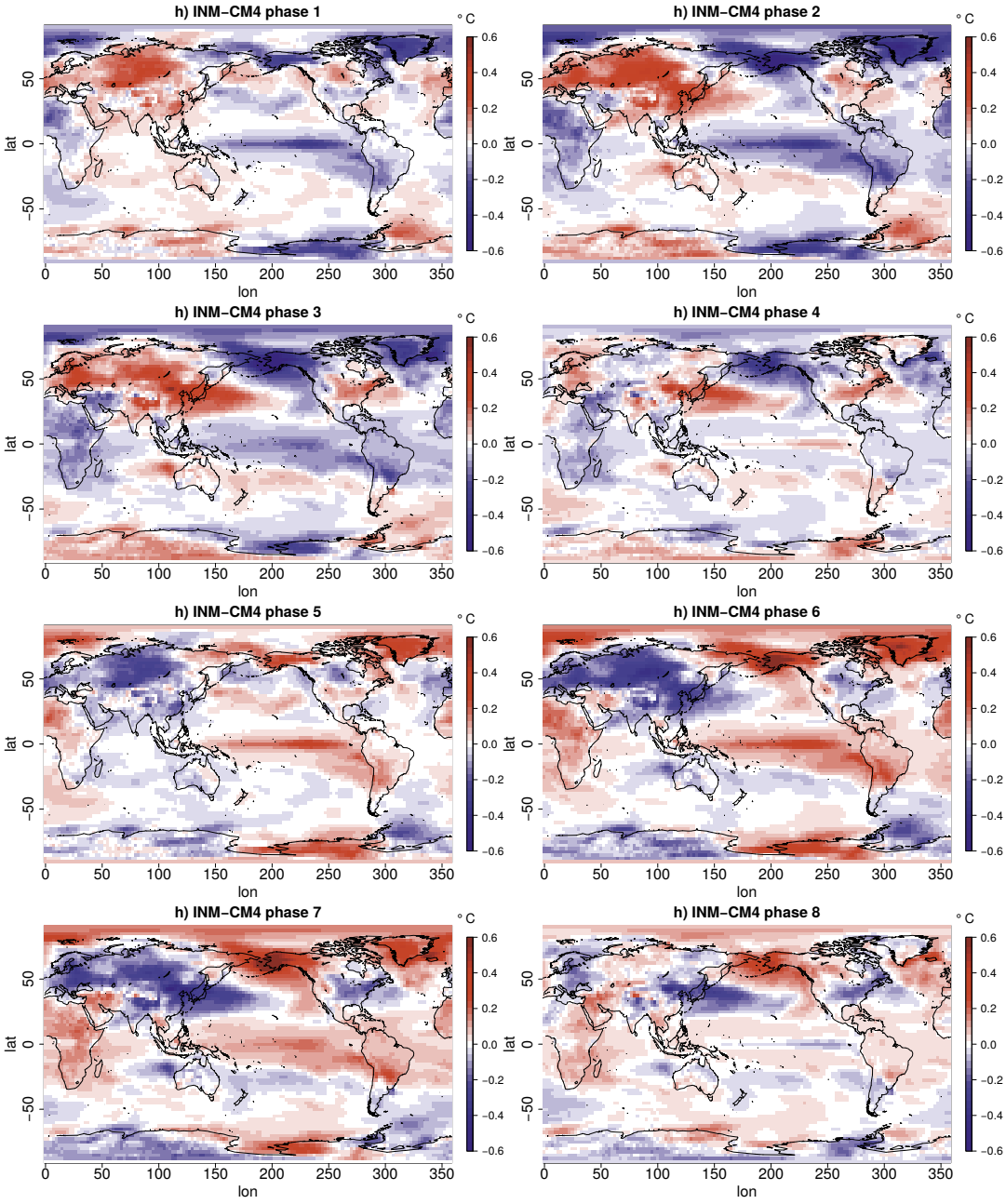


Figure S11: As Figure S3 but now for model h; ST-PC pair 1–2.

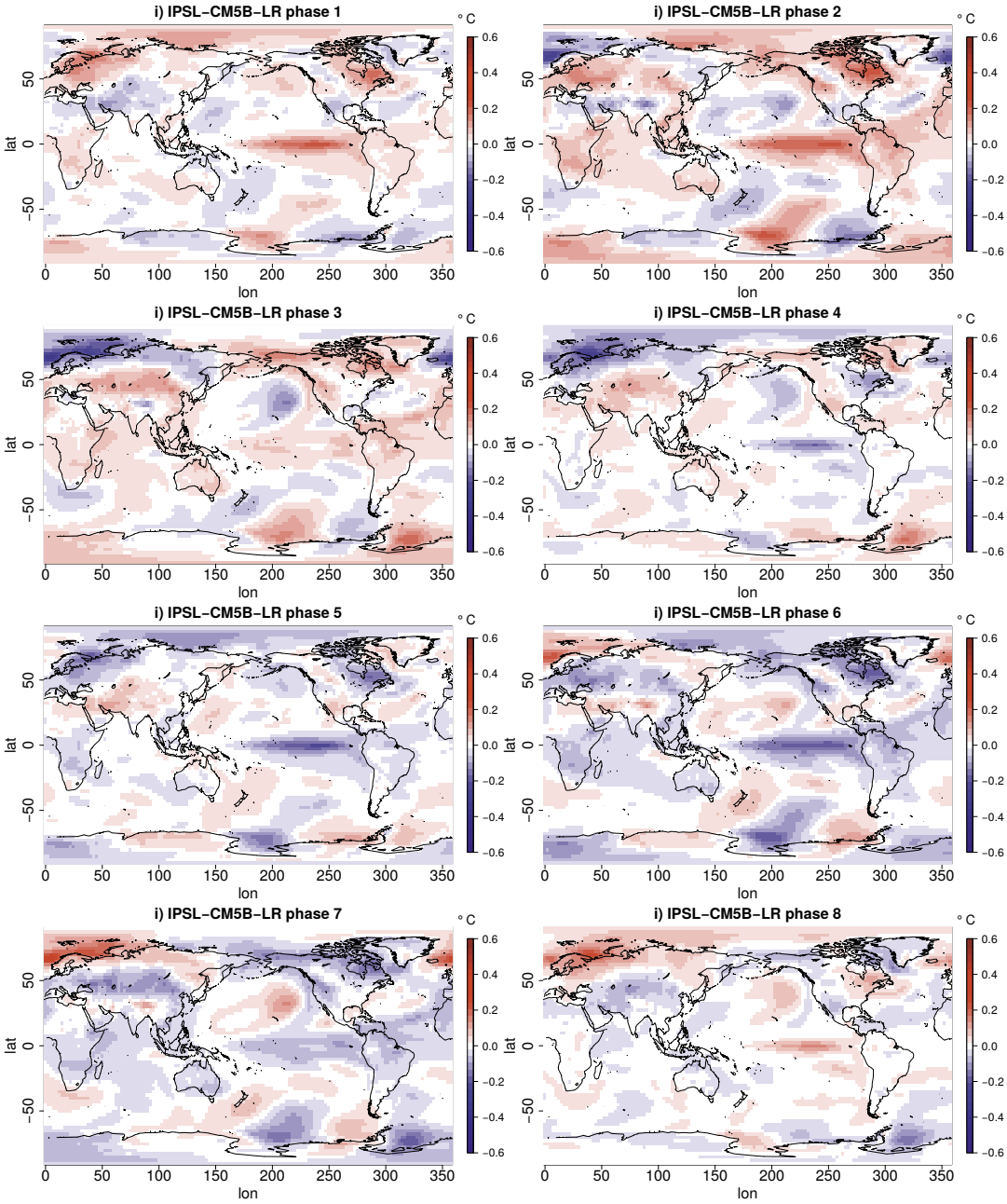


Figure S12: As Figure S3 but now for model i; ST-PC pair 7–8.

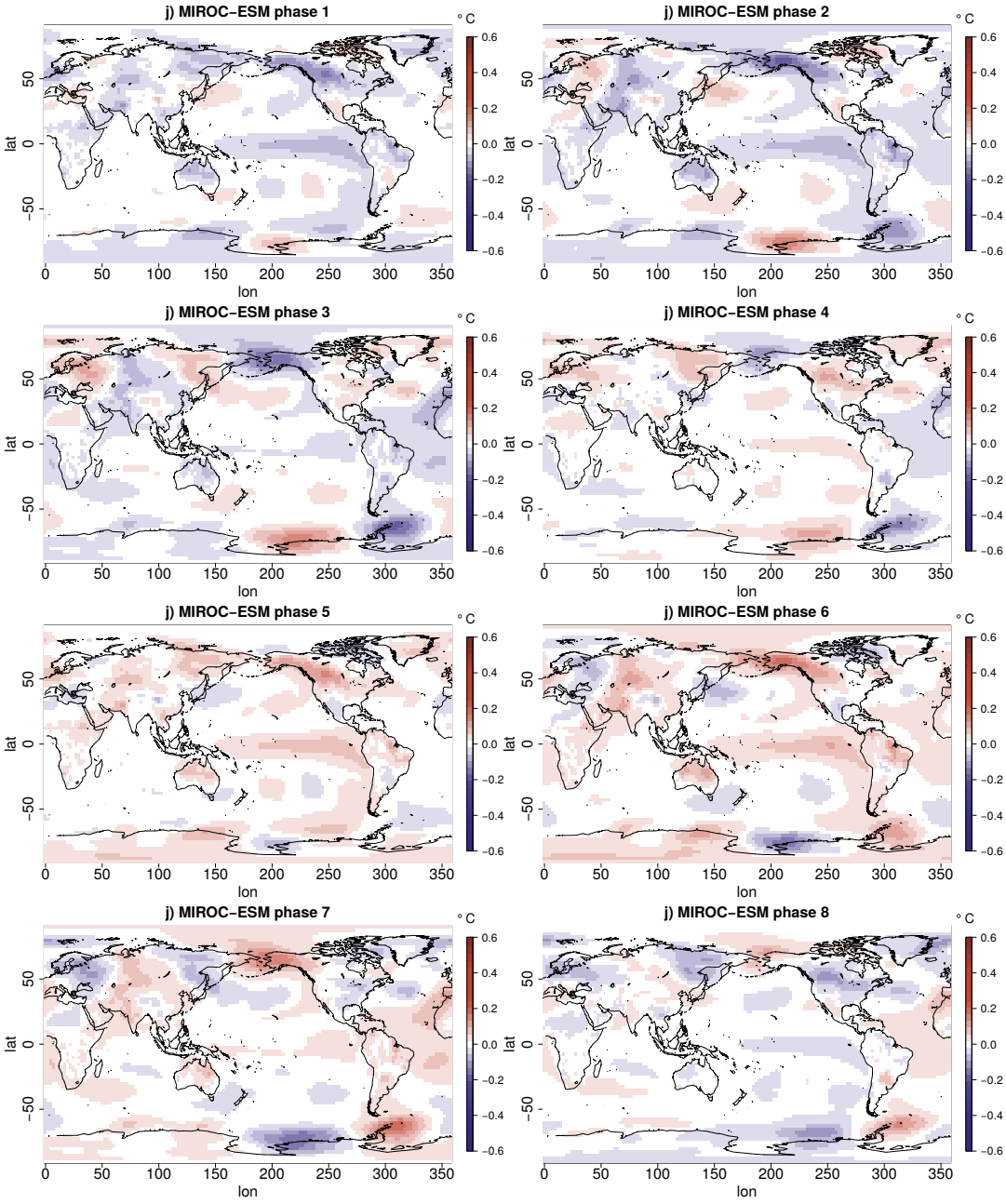


Figure S13: As Figure S3 but now for model j; ST-PC pair 15–16.

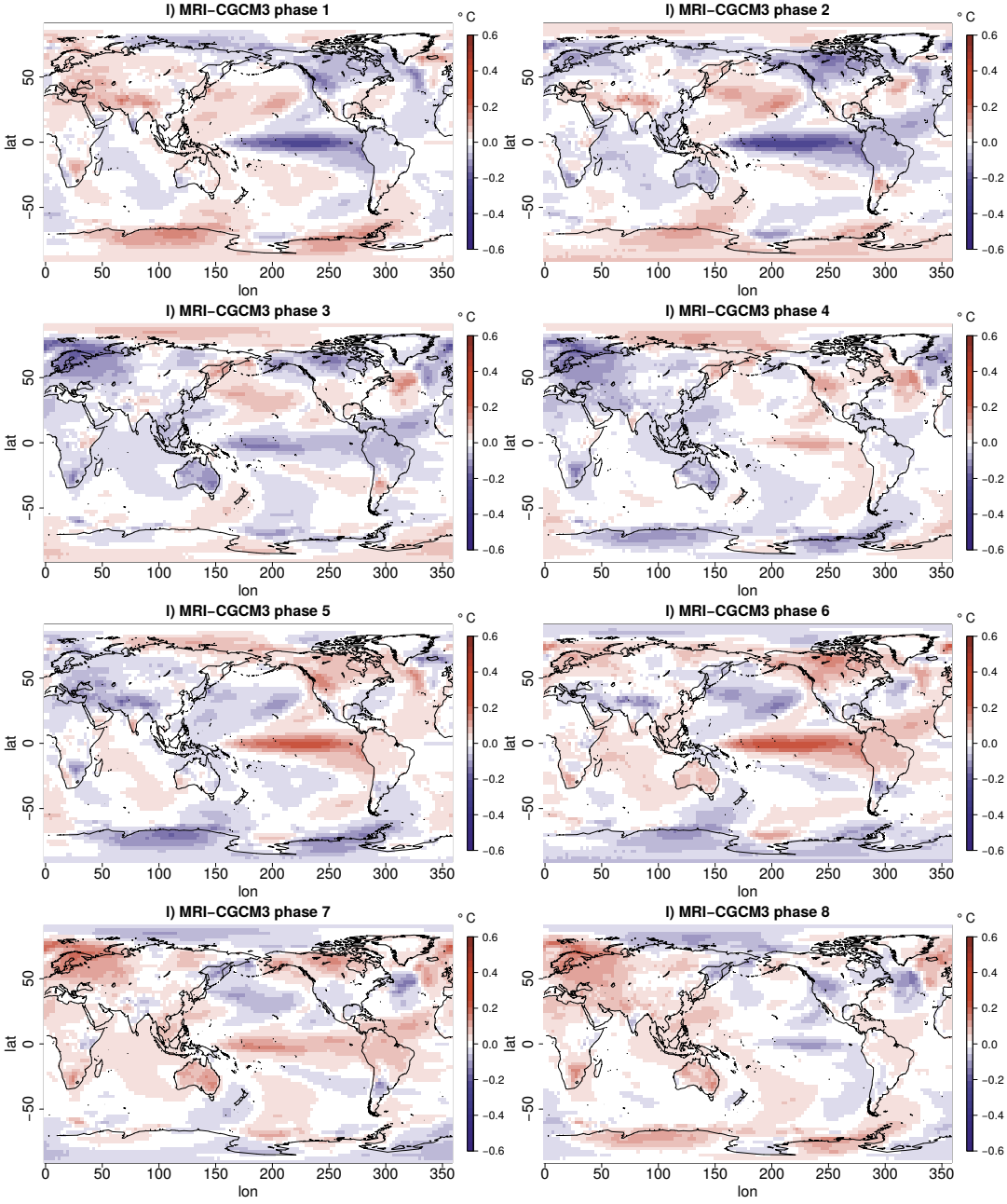


Figure S14: As Figure S3 but now for model l; ST-PC pair 7–8.

FINNISH METEOROLOGICAL INSTITUTE
Erik Palménin aukio 1
P.O.Box 503
FI-00101 HELSINKI
tel. +358 29 539 1000
WWW.FMI.FI

FINNISH METEOROLOGICAL INSTITUTE
CONTRIBUTIONS No. 126
ISBN 978-951-697-897-3 (paperback)
ISSN 0782-6117
Erweko
Helsinki 2016

ISBN 978-951-697-898-0 (pdf)
Helsinki 2016