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

Validation of CMIP5 multimodel ensembles through the smoothness of climate variables

By MYOUNGJI LEE[1], MIKYOUNG JUN[2]* and MARC G. GENTON[3], [1]*Institute for Applied Mathematics and Computational Science, Texas A&M University, College Station, TX, USA*; [2]*Department of Statistics, Texas A&M University, College Station, TX, USA*; [3]*CEMSE Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia*

ABSTRACT

Smoothness is an important characteristic of a spatial process that measures local variability. If climate model outputs are realistic, then not only the values at each grid pixel but also the relative variation over nearby pixels should represent the true climate. We estimate the smoothness of long-term averages for land surface temperature anomalies in the Coupled Model Intercomparison Project Phase 5 (CMIP5), and compare them by climate regions and seasons. We also compare the estimated smoothness of the climate outputs in CMIP5 with those of reanalysis data. The estimation is done through the composite likelihood approach for locally self-similar processes. The composite likelihood that we consider is a product of conditional likelihoods of neighbouring observations. We find that the smoothness of the surface temperature anomalies in CMIP5 depends primarily on the modelling institution and on the climate region. The seasonal difference in the smoothness is generally small, except for some climate regions where the average temperature is extremely high or low.

Keywords: composite likelihood, Gaussian process, NCEP/NCAR reanalysis, restricted likelihood, surface temperature anomaly, uncertainty quantification, variogram

1. Introduction

Smoothness is an important characteristic of a spatial process that measures local variability of the process. Due to the presence of strong spatial correlation in most climate variables, the values of a climate variable at nearby locations are considered simultaneously in many studies. Realistic climate models are expected to produce plausible values of climate variables, not only at each grid pixel, but also at its nearby grid pixels, in order to accurately describe spatial variation of the true process. Therefore, the smoothness is an important measure to validate climate models in terms of their ability to simulate fine scale spatial variability of climate variables. The aim of this paper is twofold: first, we seek to validate the spatial smoothness of a (temporal) long-term average climate variable, by season and by climate regions; and second, we compare the smoothness of climate model outputs from the Coupled Model Intercomparison Project Phase 5 (CMIP5) and the meteorological reanalysis data. The spatial smoothness is estimated by assuming isotropy within each climate region, and we assess the nonstationarity of the climate model outputs by assuming that the smoothness varies by region. We also assess the similarities in the smoothness across climate model ensembles. If all climate ensembles represent the same true phenomenon but deviate from the truth by random errors, the difference in the smoothness of ensemble realisations and reanalysis data should be negligible without any patterns across the climate ensembles.

We consider multidecadal averages of land surface temperatures in CMIP5 experiments. CMIP5 comprises a standard set of coordinated climate change experiments of 60 deterministic climate models. It is processed by the Working Group on Coupled Modelling of the World Climate Research Programme, who has gathered around 20 climate modelling groups from across the world. Outputs are archived in a common format and can be downloaded from the Program for Climate Model Diagnosis and Intercomparison web site (PCMDI, www-pcmdi.llnl.gov/). Taylor et al. (2012a) presented an overview of the experimental designs in CMIP5 and Knutti and Sedláček (2013) described detailed characteristics of climate model projections in CMIP5. The reanalysis data that we consider are from the National Centers for Environmental Prediction

*Corresponding author.
email: mjun@stat.tamu.edu

Citation: Tellus A 2015, **67**, 23880, http://dx.doi.org/10.3402/tellusa.v67.23880

and the National Center for Atmospheric Research (NCEP/NCAR). The NCEP/NCAR reanalysis data set is a gridded data set representing the state of the Earth's atmosphere, incorporating observations and numerical weather prediction model output. It is provided by the Earth System Research Laboratory in the National Oceanic and Atmospheric Administration (www.esrl.noaa.gov/psd/). Kalnay et al. (1996) and Kistler et al. (2001) presented the details of the NCEP/NCAR reanalysis.

We introduce the concept of a locally self-similar process, and we model long-term average land surface temperature anomalies as a Gaussian locally self-similar process. The local self-similarity is a weaker (and thus more general) assumption than the Matérn covariance function (a widely used covariance model that enables modelling the spatial smoothness of a stochastic process) and is also capable of modelling spatial smoothness. The estimation of smoothness is done by optimising the composite restricted likelihood for the smoothness parameter of a locally self-similar process (Stein et al., 2004; Lee, 2012). The composite likelihood is a general term for any product of marginal or conditional likelihoods (Varin et al., 2011). It has been widely used as a substitute of likelihood, when the likelihood calculation is difficult. This paper considers a product of conditional likelihoods of neighbouring observations. Since nearby observations contain most information on the local behaviour of a process, our approach balances statistical and computational efficiency in estimating the smoothness of the process. In addition to the composite likelihood approach, Gaussian Markov Random Fields form another approximation approach applicable (Rue and Held, 2005). However, we chose the composite likelihood approach since its implementation is closer to traditional statistical practice and may be more familiar to the climate science community.

The remainder of this paper is organised as follows: Section 2 describes the statistical methodology used in the estimation of the smoothness parameter; Section 3 analyses long-term average near-surface air temperature anomalies in CMIP5 and NCEP/NCAR reanalysis by region and season; and Section 4 summarises the key findings and proposes related future work. Details on the statistical methodologies are provided in the Appendix.

2. Methodology

2.1. Locally self-similar process

Suppose that we observe a mean zero and isotropic Gaussian process, $\{Z(\mathbf{s}),\ \mathbf{s}\in\mathcal{E}\}$, for $\mathcal{E}\subset\mathbb{R}^3$. We assume that the semi-variogram of Z, $\gamma(\cdot)$, follows a power function around the origin. That is, for all $\mathbf{s}$ and $\mathbf{u}\in\mathcal{E}$, with the Euclidean norm, $\|\cdot\|$:

$$\begin{aligned}\gamma(\|\mathbf{s}-\mathbf{u}\|) &= \frac{1}{2}\mathrm{E}\{Z(\mathbf{s})-Z(\mathbf{u})\}^2 \\ &= C\|\mathbf{s}-\mathbf{u}\|^{2H}+o(\|\mathbf{s}-\mathbf{u}\|^{2H}), \\ &\text{as } \|\mathbf{s}-\mathbf{u}\|\to 0,\end{aligned} \tag{1}$$

where $\boldsymbol{\theta}=(C,\ H)$, for the *scale parameter*, $C>0$, and the *smoothness parameter*, $H\in(0,1)$ (Gneiting et al., 2012). The notation $o(||\mathbf{s}-\mathbf{u}||^{2H})$ means that $\{\gamma(||\mathbf{s}-\mathbf{u}||)-C||\mathbf{s}-\mathbf{u}||^{2H}\}/||\mathbf{s}-\mathbf{u}||^{2H}\to 0$, as $||\mathbf{s}-\mathbf{u}||\to 0$. The scale parameter controls the overall size of the variation of the process. And the larger the smoothness parameter is, the smoother the realised surfaces of Z. A process that satisfies eq. (1) is called *locally self-similar*, since a self-similar process with index H satisfies $\gamma(||\mathbf{s}-\mathbf{u}||)=C||\mathbf{s}-\mathbf{u}||^{2H}$ for all $\mathbf{s}$ and $\mathbf{u}\in\mathbb{R}^3$ (Samorodnitsky and Taqqu, 1994; Genton et al., 2007). The index H determines the smoothness of the self-similar process, and it has been widely used as a measure of surface roughness for various natural phenomena, such as the surface of soil, surface height measurements of computer chips, etc. [see Mandelbrot and Wallis (1969) and Adler (1981) for some application examples]. Note that the quantity, $2H$, is essentially the fractal index for the mean zero isotropic Gaussian process, Z (Gneiting and Schlather, 2004; Gneiting et al., 2012).

A locally self-similar process is more general than a self-similar process, in the sense that it does not fully specify the variogram but only in a region around the origin. Indeed, mean zero Gaussian processes with many widely used parametric covariance functions are locally self-similar. These include powered exponential, generalised Cauchy or Matérn covariances (Gneiting et al., 2012, Table 1). Among these, the most widely used, the isotropic Matérn covariance, takes the following form:

$$\mathrm{Cov}\{Z(\mathbf{s}+\mathbf{h}),Z(\mathbf{s})\}=\sigma^2\frac{\|\mathbf{h}/\phi\|^{\nu}\mathcal{K}_{\nu}(\|\mathbf{h}/\phi\|)}{2^{\nu-1}\Gamma(\nu)}, \tag{2}$$

for all $\mathbf{s}$ and $\mathbf{h}$, where $\Gamma(\cdot)$ is the gamma function and $\mathcal{K}_{\nu}$ is a Bessel function of the second kind of order ν. The parameters of the Matérn covariance function are the partial sill, $\sigma^2>0$, the range, $\phi>0$, and the smoothness parameter, $\nu>0$. The range parameter controls the rate of correlation decay with distance and the partial sill measures the size of variation of the process. A mean zero Gaussian process with the Matérn covariance satisfies eq. (1) with $H=\nu$ and $C=\sigma^2 2^{-1-\nu}\phi^{-2\nu}\Gamma(1-\nu)/\Gamma(1+\nu)$ when $0<\nu<1$. Note that our variogram model in eq. (1) is more general than the Matérn model as we only specify the variogram near the origin.

In this paper, we restrict our attention to the case when $0<H<1$, under which eq. (1) becomes a statistically valid variogram. Many natural phenomena satisfy this condition.

Table 1. The NCEP/NCAR reanalysis and the list of the climate models that comprise historical runs (Experiment 3.2) of CMIP5, with their modelling institution, official institution ID, country of modelling institution, grid resolution and number of ensemble realisations available in this paper

Modelling centre (or group)	Institute ID	Country	Model number	Model name	Resolution	# of replicates
National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR)	NCEP/NCAR	USA		Reanalysis	192 × 94	n/a
Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM)	CSIRO-BOM	Australia	33	ACCESS1.0	192 × 145	1
			32	ACCESS1.3	192 × 145	3
College of Global Change and Earth System Science, Beijing Normal University	GCESS	China	9	BNU-ESM	128 × 64	1
National Center for Atmospheric Research	NCAR	USA	46	CCSM4	288 × 192	6
Community Earth System Model Contributors	NSF-DOE-NCAR	USA	45	CESM1(BGC)	288 × 192	1
			22	CESM1(CAM5.1,FV2)	144 × 96	4
			43	CESM1(CAM5)	288 × 192	3
			44	CESM1(FASTCHEM)	288 × 192	3
			21	CESM1(WACCM)	144 × 96	4
Centro Euro-Mediterraneo per I Cambiamenti Climatici	CMCC	Europe	1	CMCC-CESM	96 × 48	1
			25	CMCC-CMS	192 × 96	1
			47	CMCC-CM	480 × 240	1
Centre National de Recherches Météorologiques/Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	CNRM-CERFACS	France	37	CNRM-CM5	256 × 128	10
Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	CSIRO-QCCCE	Australia	26	CSIRO-Mk3.6.0	192 × 96	10
Canadian Centre for Climate Modelling and Analysis	CCCMA	Canada	5	CanCM4	128 × 64	10
			4	CanESM2	128 × 64	5
EC-EARTH consortium	EC-EARTH	Europe	41	EC-EARTH	320 × 160	11
LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences and CESS, Tsinghua University	LASG-CESS	China	3	FGOALS-g2	128 × 60	5
The First Institute of Oceanography, SOA	FIO	China	10	FIO-ESM	128 × 64	1
National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory	NOAA GFDL	USA	20	GFDL-CM2.1	144 × 90	10
			19	GFDL-CM3	144 × 90	5
			18	GFDL-ESM2G	144 × 90	1
			17	GFDL-ESM2M	144 × 90	1
NASA Goddard Institute for Space Studies	NASA GISS	USA	13	GISS-E2-H-CC	144 × 90	1
			14	GISS-E2-H	144 × 90	1
			15	GISS-E2-R-CC	144 × 90	1
			16	GISS-E2-R	144 × 90	25
Met Office Hadley Centre (additional HadGEM2-ES realisations contributed by Instituto Nacional de Pesquisas Espaciais)	MOHC	UK	2	HadCM3	96 × 73	10
			34	HadGEM2-CC	192 × 145	3
	MOHC/INPE		36	HadGEM2-ES	192 × 145	4

Table 1 (*Continued*)

Modelling centre (or group)	Institute ID	Country	Model number	Model name	Resolution	# of replicates
National Institute of Meteorological Research/Korea Meteorological Administration	NIMR/KMA	Korea	35	HadGEM2-AO	192 × 145	1
Institut Pierre-Simon Laplace	IPSL	France	11	IPSL-CM5A-LR	96 × 96	6
			30	IPSL-CM5A-MR	144 × 143	3
			12	IPSL-CM5B-LR	96 × 96	1
Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	MIROC	Japan	6	MIROC-ESM-CHEM	128 × 64	1
			7	MIROC-ESM	128 × 64	3
Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	MIROC	Japan	48	MIROC4h	640 × 320	3
			38	MIROC5	256 × 128	5
Max-Planck-Institut für Meteorologie (Max Planck Institute for Meteorology)	MPI-M	Germany	27	MPI-ESM-LR	192 × 96	3
			29	MPI-ESM-MR	192 × 96	3
			28	MPI-ESM-P	192 × 96	2
Meteorological Research Institute	MRI	Japan	39	MRI-CGCM3	320 × 160	5
			40	MRI-ESM1	320 × 160	1
Norwegian Climate Centre	NCC	Norway	24	NorESM1-ME	144 × 96	1
			23	NorESM1-M	144 × 96	3
Beijing Climate Center, China Meteorological Administration	BCC	China	42	BCC-CSM1.1(m)	320 × 160	3
			8	BCC-CSM1.1	128 × 64	3
Institute for Numerical Mathematics	INM	Russia	31	INM-CM4	180 × 120	1

For example, Tuck (2008, p. 14, 41) studied atmospheric variability and observed that temperature is smoother than wind speed. The scaling exponent, which is the same as $2H$ of eq. (1), of the temperature was shown to be close to but less than unity. Lovejoy and Schertzer (1985, p. 1235) pointed out empirically that $0 < H < 1$ in the rate of energy transfer, buoyancy, velocity, temperature fluctuations, radar reflectivity and cloud drop volumes. North et al. (2011) found that the spatial covariance of temperature fields based on simple energy balance climate models follows the Matérn covariance with $\nu = 1$, and that $\nu < 1$ is expected due to rough landscapes. Sun et al. (2015) mentioned that precipitation amounts become smoother when summed over longer periods and they showed numerically that the smoothness of long-term precipitation amounts is less than $\nu = 0.5$. We determine that the smoothness of multidecadal average near-surface air temperature anomalies is between zero and one in Section 3.

One thing to note is that the estimated smoothness may depend on the grid resolution of the climate models. In the estimation procedure described in Section 2.2, the relationship, eq. (1), is applied to the number ($k = 3,\ldots,10$) of neighbouring observations. As shown in Table 1, climate models in CMIP5 have various grid resolutions. In Section 3, we check the effect of spatial grid resolution on the estimated smoothness.

2.2. Composite likelihood

To estimate the scale and smoothness parameters of a locally self-similar process, we consider the composite restricted likelihood of $\boldsymbol{\theta}$. We briefly introduce the idea of composite likelihood as opposed to the likelihood method in this section. Further details on how to calculate composite restricted likelihoods are given in the Appendix.

The idea of restricted likelihood is used to estimate variogram parameters without estimating nuisance parameters such as $\mathrm{E}\{Z(\cdot)\}$ or $\mathrm{Var}\{Z(\cdot)\}$ (Kitanidis, 1983). It is a marginal likelihood associated with any $N-1$ linearly independent error contrasts, mean zero linear combination of the observations. Since a locally self-similar process does not fully specify the variogram, we have neither the exact likelihood nor the restricted likelihood of $\boldsymbol{\theta}$. Therefore, we approximate the restricted likelihood of $\boldsymbol{\theta}$ by the composite restricted likelihood, similarly to Stein et al. (2004) and Lee (2012).

Let us first sketch the idea to obtain a composite likelihood. Suppose that $Z(\cdot)$ is observed at N locations, $\{\mathbf{s}_1,\ldots,\mathbf{s}_N\}$. Let $p(\cdot;\boldsymbol{\theta})$ indicate a generic probability density function, possibly conditional density. We order the observation locations by starting from a random location, $\mathbf{s}_1$, then selecting $\mathbf{s}_i$ to be the nearest location to any of $\{\mathbf{s}_1,\ldots,\mathbf{s}_{i-1}\}$ among the remaining locations, for $i \geq 2$. If there are two or more locations at equal distance from the set $\{\mathbf{s}_1,\ldots,\mathbf{s}_{i-1}\}$, we choose one randomly. The likelihood of $\boldsymbol{\theta}$ is

$$\begin{aligned} &p(Z(\mathbf{s}_1),\ldots,Z(\mathbf{s}_N);\ \boldsymbol{\theta}) \\ &\quad = p(Z(\mathbf{s}_1),\ldots,Z(\mathbf{s}_k);\ \boldsymbol{\theta}) \\ &\qquad \times \prod_{i=k+1}^{N} p(Z(\mathbf{s}_i)|Z(\mathbf{s}_1),\ldots,Z(\mathbf{s}_{i-1});\ \boldsymbol{\theta}). \end{aligned} \tag{3}$$

Now, in order to define a composite likelihood, for each $\mathbf{s}_i$, define k locations in proximity of $\mathbf{s}_i$, among the previously selected locations as $\{\mathbf{s}_{i,1},\ldots,\mathbf{s}_{i,k}\} \subset \{\mathbf{s}_1,\ldots,\mathbf{s}_{i-1}\}$, for $i > k$. Since closely located observations are highly correlated and informative about the smoothness of the process, the *composite likelihood* approximates eq. (3) by conditioning on $\{\mathbf{s}_{i,1},\ldots,\mathbf{s}_{i,k}\}$ only:

$$p(Z(\mathbf{s}_1),\ldots,Z(\mathbf{s}_k);\ \boldsymbol{\theta}) \prod_{i=k+1}^{N} p(Z(\mathbf{s}_i)|Z(\mathbf{s}_{i,1}),\ldots,Z(\mathbf{s}_{i,k});\ \boldsymbol{\theta}). \tag{4}$$

Call $\{Z(\mathbf{s}_{i,1}),\ldots,Z(\mathbf{s}_{i,k})\}$ the conditioning set of the composite likelihood, where k denotes the size of the conditioning set. The composite likelihood, eq. (4), is associated with the statistical optimal property if Z follows a Gaussian process. For a Gaussian probability density, p, $p(Z(\mathbf{s}_i)|Z(\mathbf{s}_{i,1}),\ldots,Z(\mathbf{s}_{i,k});\boldsymbol{\theta})$ is the density of the error of the best linear predictor of $Z(\mathbf{s}_i)$ based on $Z(\mathbf{s}_{i,1}),\ldots,Z(\mathbf{s}_{i,k})$. Also, the approximation in eq. (4) requires $O(k^3N)$ operations while the likelihood requires $O(N^3)$ operations. It is especially beneficial for large irregularly spaced observations where the likelihood calculation is computationally demanding.

The composite restricted log-likelihood, $\widetilde{rl}_k(\boldsymbol{\theta})$, provided in the Appendix, is defined similarly by applying the idea of the composite likelihood to the logarithm of the restricted likelihood. Our estimator, $\widehat{\boldsymbol{\theta}}$, is then defined as a value that maximises the composite restricted log-likelihood. We consider the conditioning set of size $k = 3,\ldots,10$ in Section 3. We assess the variance of $\widehat{\boldsymbol{\theta}}$ by the sandwich estimator, a widely used measure of the variance of estimators from an estimating equation, $\nabla\widetilde{rl}_k(\boldsymbol{\theta}) = 0$. Here, ∇ denotes the vector of partial derivatives with respect to $\boldsymbol{\theta}$. Then we have $\widehat{\boldsymbol{\theta}}$ is asymptotically normal with asymptotic covariance matrix

$$\{J_n(\boldsymbol{\theta})V_n^{-1}(\boldsymbol{\theta})J_n(\boldsymbol{\theta})\}^{-1},\ \text{where}$$

$$J_n(\boldsymbol{\theta}) = E\{-\nabla^2\widetilde{rl}_k(\boldsymbol{\theta})\}\ \text{and}\ V_n(\boldsymbol{\theta}) = \mathrm{Var}\{\nabla\widetilde{rl}_k(\boldsymbol{\theta})\}.$$

See Lindsay (1988) and Godambe and Heyde (2010) for more details.

3. Analysis

3.1. Data

Climate model outputs from CMIP5 consist of 3 and 6 hourly, daily, monthly and annual mean values of more than 404 ocean, land and atmosphere related climate variables for decadal hindcasts and predictions. The NCEP/NCAR reanalysis data consist of 6 hourly, daily and monthly mean values of atmospheric variables from January 1948 to the most recent month. In this paper, we analyse the long-term average near-surface air temperatures measured at 2 m above ground at gridded locations on the Earth from 1979 to 2005, the time period common to all climate models in CMIP5 and the NCEP/NCAR reanalysis.

We analyse 191 ensemble runs from the 48 climate models in CMIP5 (experiment 3.2). Each climate model has 1–25 ensemble replicates that are initialised under different or the same initial conditions but produced by different perturbed versions of the same model (Taylor et al., 2012b). Ensemble replicates are treated and interpreted independently from each other, and their spatial resolutions vary from ensemble to ensemble. Table 1 lists the climate models in CMIP5 and the NCEP/NCAR reanalysis data set used in this paper, with their grid resolutions and the numbers of ensemble replicates. The climate models are numbered in ascending order of the number of grid pixels. The model number thus represents the rank of the spatial resolution of the climate model. For the climate models with the same spatial resolutions, lower model numbers are given to the ones with smaller average estimated smoothness over the regions.

We focus on the mean surface temperatures in Boreal winter (December, January, February; DJF) and summer (June, July, August; JJA), averaged over 27 yr. That is, at each location, we use multidecadal averages of land surface air temperatures during DJF and JJA. Also, we divide the land area except for Antarctica into the 21 climate regions that are used in Giorgi and Francisco (2000). There are two main reasons for dividing the land areas into climate regions. It is common that the smoothness varies spatially in climate variables. Also, the distance between grid points becomes smaller in regions at higher latitudes. Since the estimated smoothness parameter depends on the resolution of the observed process, dividing regions where observations are separated by similar spacing is reasonable. The climate regions are shown in Fig. 1. The sizes of the regions vary from 807 to 6735 km in the north-south and east-west directions. Each region contains from 12 to 7649 grid pixels of the ensemble outputs from CMIP5, depending on the grid resolutions of the ensembles. The minimum spacing between grid locations at the equator ranges from 83 to 417 km.

3.2. Models

Denote the entire study region as $\mathcal{D}$. Then, partition $\mathcal{D}$ into the climate regions, $\mathcal{D} = \cup_{r=1}^{21} \mathcal{D}_r$. Let $T_{ijl}(\mathbf{s})$ be a multidecadal average of near-surface air temperature at grid location $\mathbf{s} \in \mathcal{D}$ for climate model $j = 1,\ldots,48$, ensemble

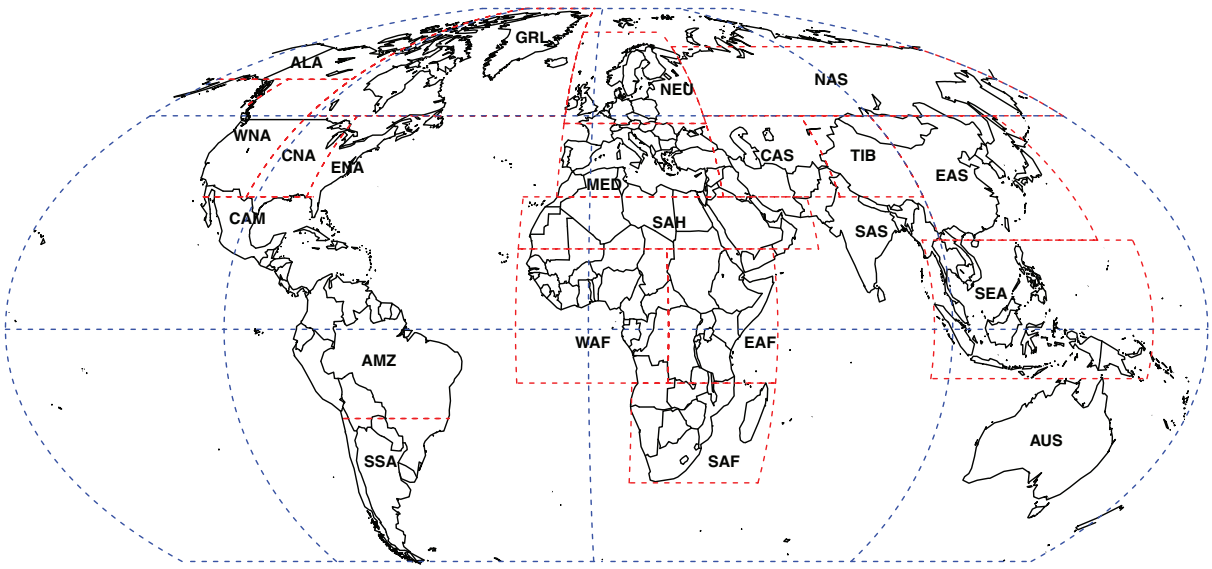

Fig. 1. Twenty-one land regions used in the study: Australia (AUS), Amazon Basin (AMZ), Southern South America (SSA), Central America (CAM), Western North America (WNA), Central North America (CNA), Eastern North America (ENA), Alaska (ALA), Greenland (GRL), Mediterranean Basin (MED), Northern Europe (NEU), Western Africa (WAF), Eastern Africa (EAF), Southern Africa (SAF), Sahara (SAH), Southeast Asia (SEA), East Asia (EAS), South Asia (SAS), Central Asia (CAS), Tibet (TIB) and North Asia (NAS).

replicate l, during DJF and JJA, for $i=1$ and $i=2$, respectively. The number of ensemble replicates varies by climate model. Let $\mu_{ijl}(\mathbf{s}) = \mathrm{E}\{T_{ijl}(\mathbf{s})\}$ be the mean of the multidecadal average and $\varepsilon_{ijl}(\mathbf{s})$ be the anomaly (residual) at location $\mathbf{s} \in \mathcal{D}$, such that

$$T_{ijl}(\mathbf{s}) = \mu_{ijl}(\mathbf{s}) + \varepsilon_{ijl}(\mathbf{s}). \tag{5}$$

Since we focus on modelling the smoothness of the temperature anomalies, ε_{ijl}, we first filter the data to estimate the mean, μ_{ijl}, and make the anomaly field close to mean zero. Spherical harmonics, $\{P_n^m(\sin L)\cos(ml), P_n^m(\sin L)\sin(ml)|\ n = 0, 1, 2, \ldots, m = 0, \ldots, \min(3, n)\}$, where $-\pi/2 \le L \le \pi/2$ is the latitude, $-\pi < l \le \pi$ is the longitude, and P_n^m is the Legendre polynomial of degree n and order m, provide a natural basis for capturing large-scale spatial patterns (Stein, 2007). Because surface temperatures are closely related to altitude, we estimate μ_{ijl} by regressing on the altitude from the sea level in addition to spherical harmonics for $n=12$, for each climate ensemble realisation in CMIP5 and the NCEP/NCAR reanalysis. The choice of $n=12$ is made following the literature dealing with similar data sets (Jun and Stein, 2008; Stein, 2008; Jun, 2011, 2014).

After the mean filtering through regression, we assume that ε_{ijl} in eq. (5) is a mean zero, locally self-similar Gaussian process that satisfies for $\mathbf{s}$ and $\mathbf{u} \in \mathcal{D}_r$, $\frac{1}{2}\mathrm{E}\{\varepsilon_{ijl}(\mathbf{s}) - \varepsilon_{ijl}(\mathbf{u})\}^2 = C_{ijrl}||\mathbf{s}-\mathbf{u}||^{2H_{ijrl}} + o(||\mathbf{s}-\mathbf{u}||^{2H_{ijrl}})$, as $||\mathbf{s}-\mathbf{u}|| \to 0$, for $r = 1, \ldots, 21$. The smoothness of the temperature anomalies in the NCEP/NCAR reanalysis is defined similarly. Since $\varepsilon_{ijl}(\mathbf{s})$ is a multidecadal average of temperature anomalies, its distribution may be close to a Gaussian distribution.

The top panels in Figs. 2 and 3 show the multidecadal average near-surface air temperature, T_{ijl}, the estimated mean, μ_{ijl}, and the anomaly, ε_{ijl}, in the reanalysis and GFDL-CM3 data, by season. The spherical harmonics terms and the altitude capture most of the patterns in the mean, and the anomalies do not have noticeable large-scale spatial patterns. Figure 4 compares the minimum, median and maximum values of the anomalies, shown in the bottom panel of Fig. 2, by climate region and season. In all regions, the medians of the anomalies are around zero and the ranges of the anomalies are similar regardless of season and region, except for ALA and GRL. The spatial patterns of the mean and residuals displayed in Figs. 2 and 4 are similar to patterns created by other ensemble models.

3.3. *Estimation of the smoothness*

We estimate the smoothness parameter, H, of the anomalies of multidecadal average land surface temperature in the NCEP/NCAR reanalysis and CMIP5 by maximising the composite restricted likelihood with a conditioning set of

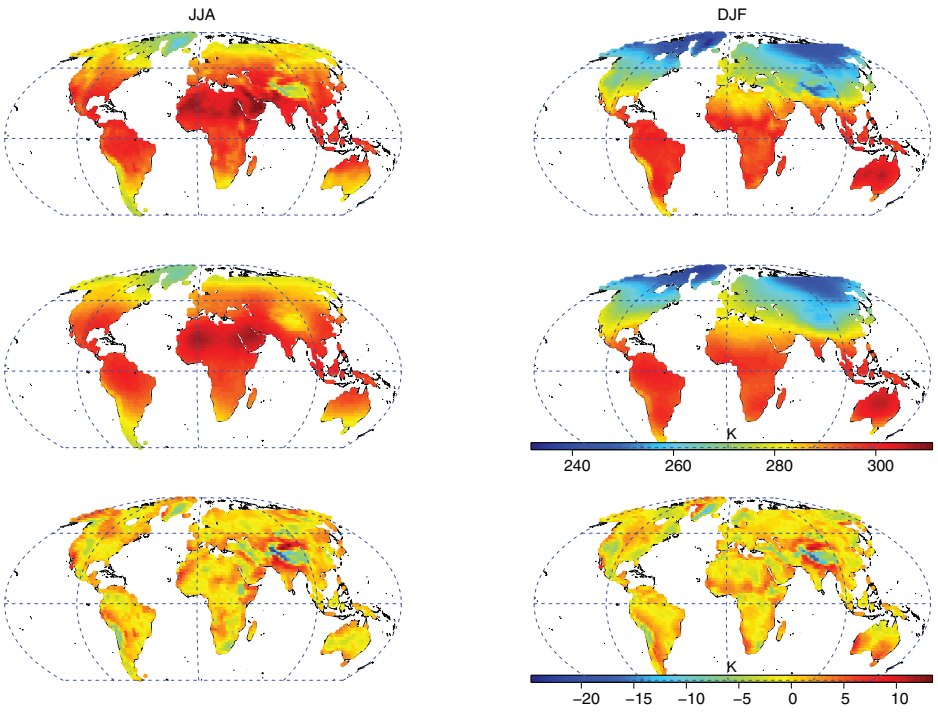

Fig. 2. The multidecadal average land surface temperature (top panel), its estimated mean (middle panel) and the residuals (bottom panel) for GFDL-CM3. The left and right panels are during JJA and DJF, respectively.

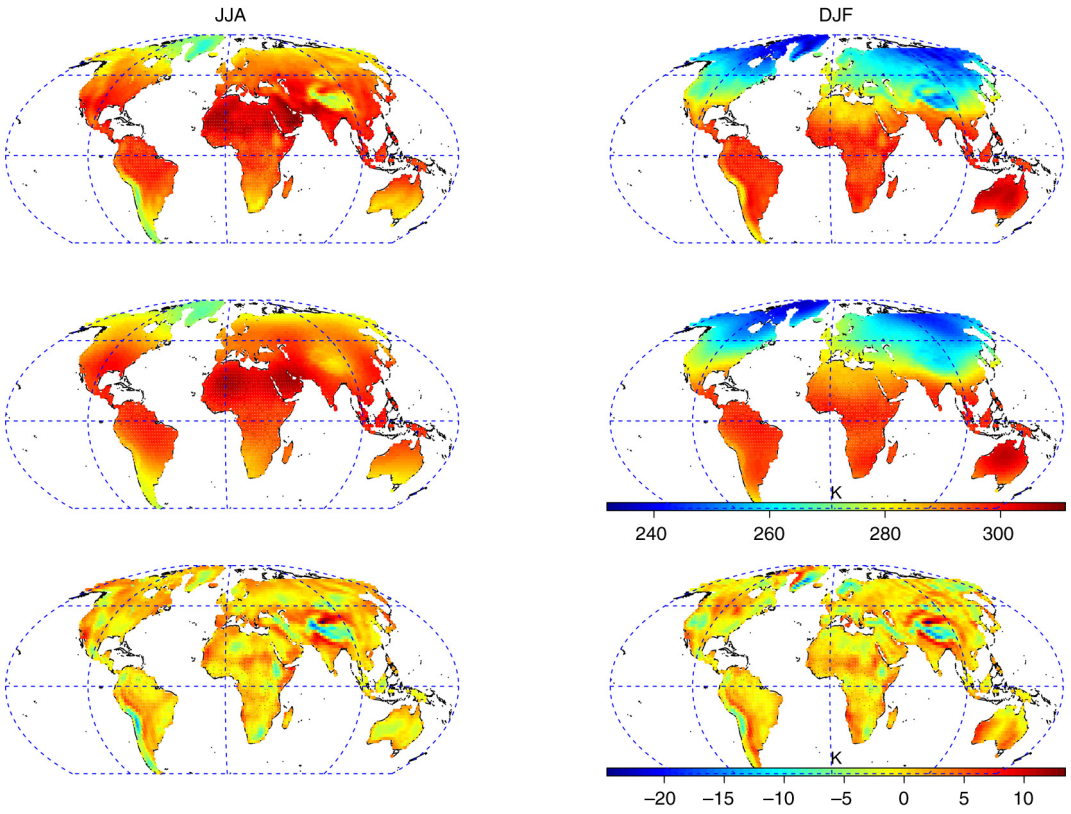

Fig. 3. The same as Fig. 2 but for the land surface temperatures in the NCEP/NCAR reanalysis.

size k, $k = 3,\ldots,10$. Figure 5 shows the changes in the smoothness estimates by increasing k. Each plot represents a climate model in CMIP5 and each curve represents a climate region in Fig. 1. Among the climate regions, WNA, SAH, NAS, AMZ and TIB are coloured. We explain the reason for choosing these specific regions after showing the estimated smoothness of the NCEP/NCAR reanalysis. The smoothness estimates become quickly stabilised as k increases. The estimated smoothness is always less than unity, regardless of the size of the conditioning set. Hereafter, we present the estimated smoothness using the composite restricted likelihood with a conditioning set of size $k = 5$.

Figure 6 maps the smoothness estimates of the temperature anomalies in the NCEP/NCAR reanalysis of the corresponding climate regions during DJF and JJA. Similarly, a smoothness map is drawn for each of the climate

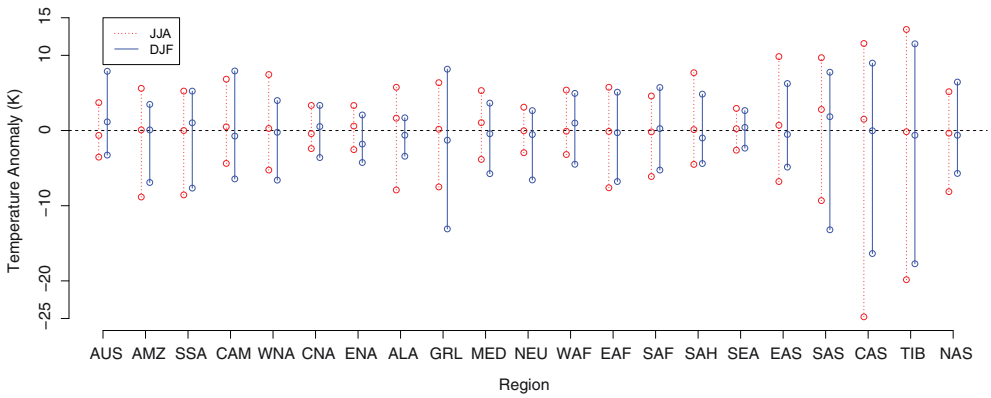

Fig. 4. Minimums, medians and maximums of the anomalies by season in GFDL-CM3.

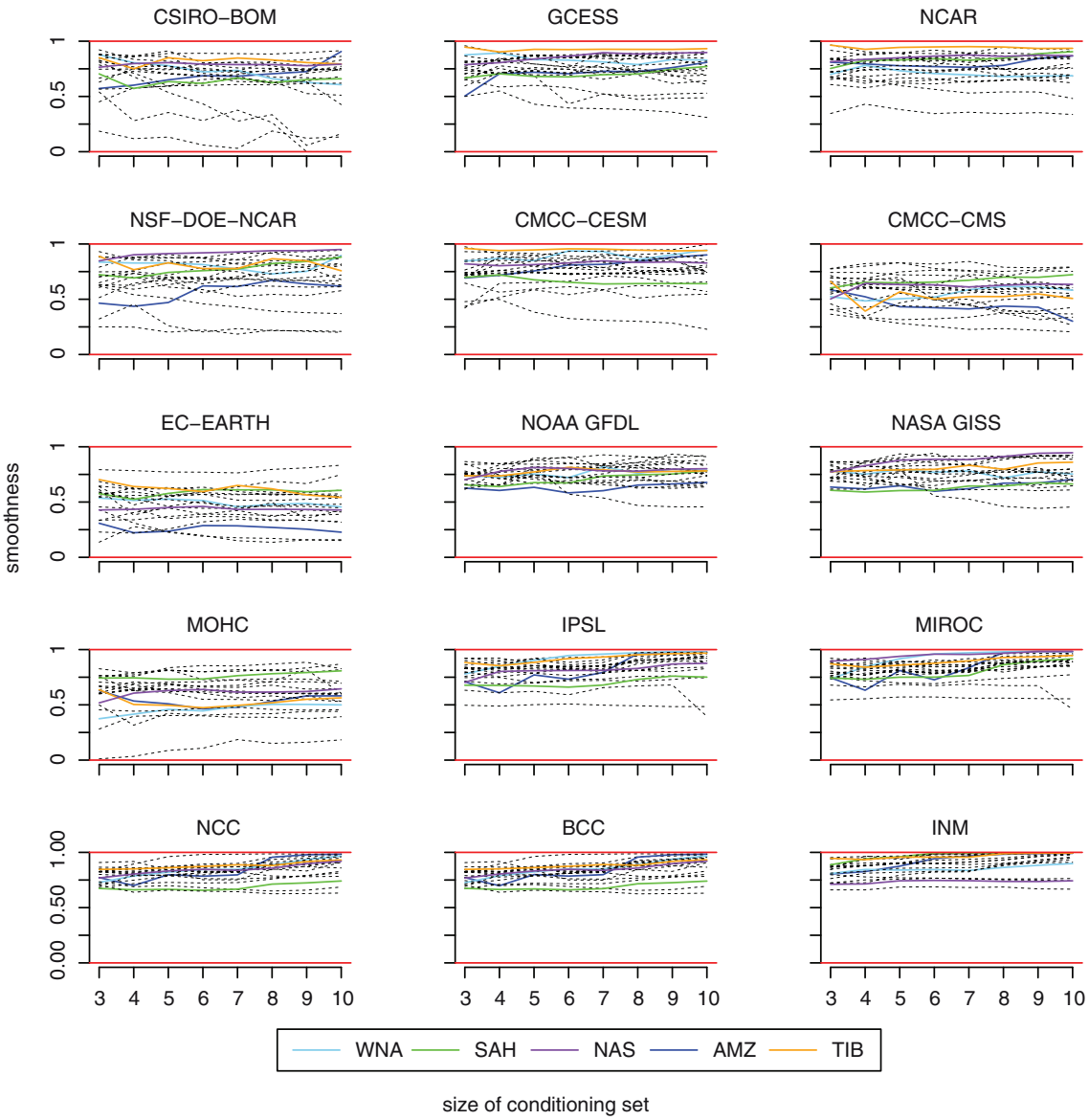

Fig. 5. Smoothness of the multidecadal average near-surface air temperature anomalies in CMIP5, by increasing the size of the conditioning set in the conditional composite restricted likelihood. Each plot corresponds to a climate model in CMIP5 and has 21 curves of the estimated smoothness during JJA, one for each climate region. The curves that correspond to the regions WNA, SAH, NAS, AMZ and TIB are coloured.

model ensembles in CMIP5. The ensemble replicates of the same climate model have almost the same smoothness in all climate regions. The climate model outputs generated from the same modelling institution also have similar smoothness in all land regions. Therefore, in order to save space, Figs. 7 and 8 show the smoothness maps of 15 climate models that have distinct patterns, for DJF and JJA, respectively.

The estimated smoothness for the NCEP/NCAR reanalysis is larger than 0.5 in all regions except CNA ($\widehat{H} = 0.041$)

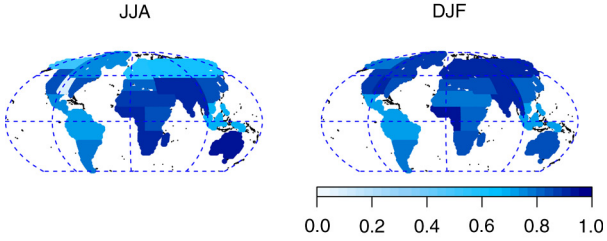

Fig. 6. Smoothness of the multidecadal average near-surface air temperature anomalies in the NCEP/NCAR reanalysis by season.

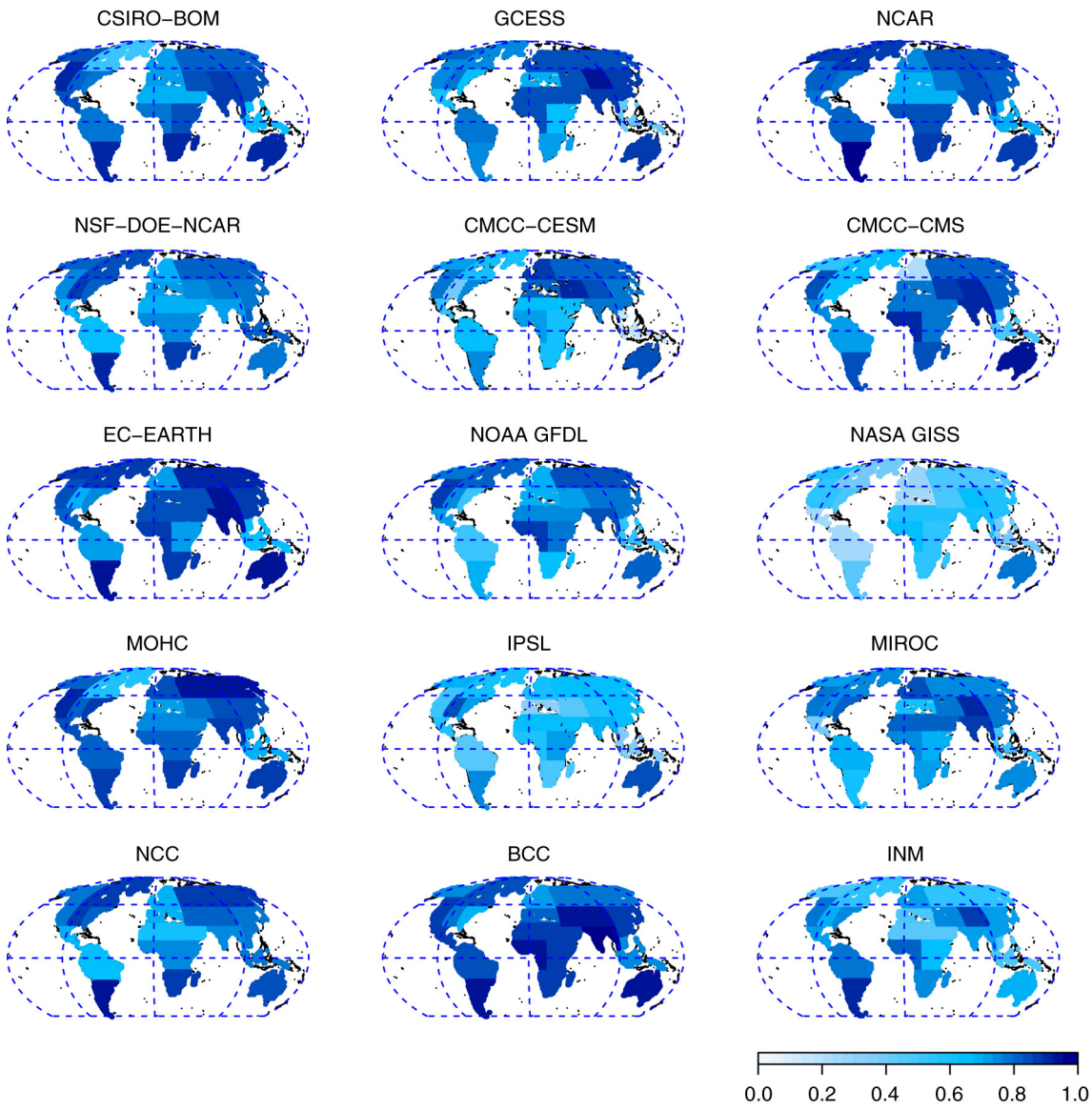

Fig. 7. Smoothness of the multidecadal average near-surface air temperature anomalies in CMIP5 during JJA.

during JJA. The regions near the North Pole (WNA, NAS, NEU, GRL, ALA) have smoother surface temperature anomaly fields during DJF than during JJA, while the regions near the equator (SAH, SAS) have smoother fields during JJA. The seasonal difference in the smoothness is small in other regions. This is the reason why we coloured WNA, SAH, NAS, AMZ and TIB in Fig. 5. Among these, the first three and the last two represent the climate regions with large and small seasonal differences, respectively. This pattern appears in GCESS, CMCC-CESM, CMCC-CMS and BCC in CMIP5. The rest of the climate models in CMIP5 exhibit similar smoothness for JJA and DJF.

Figures 6–8 show quite a lot of regional variation in each smoothness map. The smoothness maps from CMIP5 also vary across climate models. The climate modelling institution and the climate region are the main factors that determine the smoothness of the surface temperature anomalies. The relative regional characteristics, however, do not change across climate models. Figure 9 plots the smoothness of the temperature anomalies in CMIP5 against the climate model number, i.e. the rank of the spatial

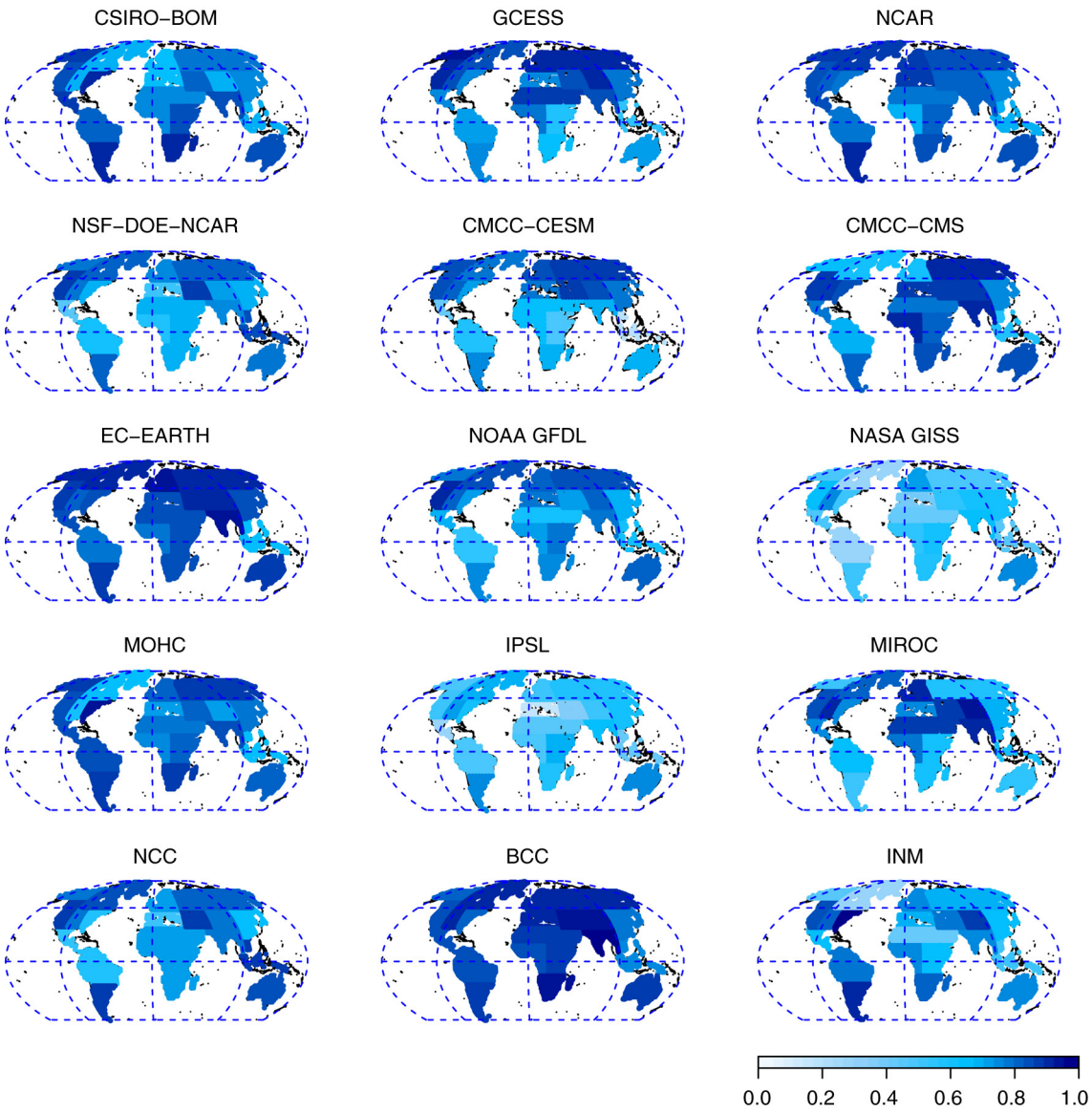

Fig. 8. Smoothness of the multidecadal average near-surface air temperature anomalies in CMIP5 during DJF.

resolution of the climate model. There are 21 curves, each of which represents a climate region. Again, the curves that represent WNA, SAH, NAS, AMZ and TIB regions are coloured. Generally, all the curves in Fig. 9 resemble each other. This implies that the climate model is the primary factor that determines the estimates for the smoothness, and the spatial resolution of the model has a weaker effect than the climate model on the estimated smoothness. Each climate model generates the relative regional characteristics well, while the average level of smoothness differs for each climate model. The climate models developed by NASA GISS, IPSL, and MOHC produce rougher temperature anomaly fields than do the other climate models over all climate regions during both seasons. The crosses at model number 22 indicate the smoothness of the NCEP/NCAR reanalysis, as the resolution of the reanalysis is similar to the resolution of climate model 22.

Some smoothness estimates are near the boundaries of the range of the smoothness parameter, $H \approx 0$, suggesting that the estimation failed. For the NCEP/NCAR reanalysis data, we fail to estimate the smoothness over the region CNA during JJA. In CNA, there were seven pairs of neighbouring observations out of 90 observations, which temperature anomalies differ significantly. Furthermore, the failure of

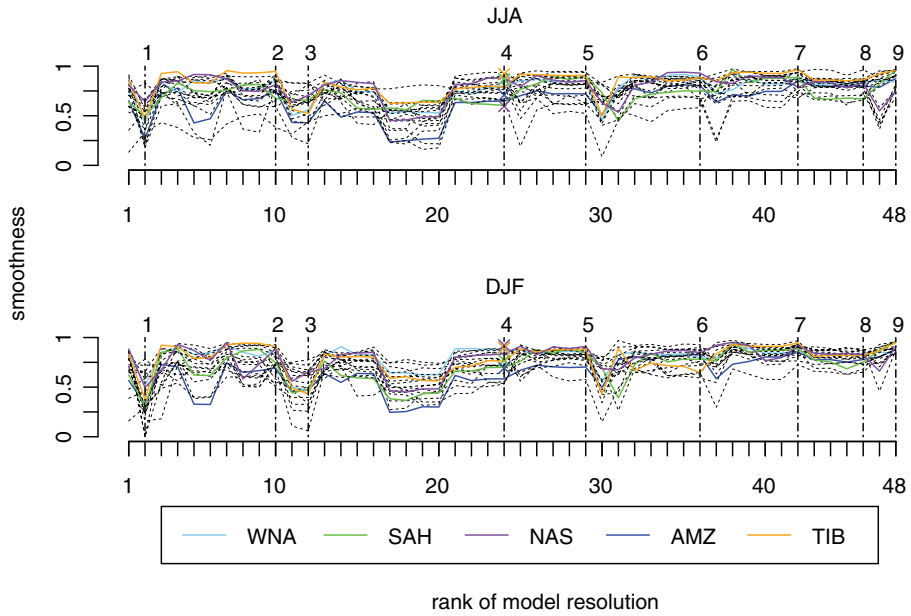

Fig. 9. Comparison of the smoothness of the multidecadal average near-surface air temperature anomalies in CMIP5, by climate region and season. Each curve corresponds to a climate region, among which WNA, SAH, NAS, AMZ, and TIB are coloured. Vertical lines differentiate resolutions (1: 96 × 73, 2: 128 × 64, 3: 96 × 96, 4: 144 × 96, 5: 192 × 96, 6: 192 × 145, 7: 320 × 160, 8: 288 × 192, 9: 640 × 320).

the estimation of CMIP5 happens when we do not have sufficient number of data, for the models with coarse grid resolution. Climate models 1 (CMCC-CESM) and 2 (HadCM3) of CMIP5 fail to estimate the smoothness of CAM, CNA, ALA or NEU during JJA or DJF. Climate model 1 has a spatial resolution of 96 × 48 on the Earth, and there are only 12 and 25 observations for CAM and CNA, respectively.

Figures 10–12 show the estimated scale parameters of the multidecadal average near-surface air temperature anomalies in reanalysis and CMIP5 during JJA and DJF. The values are plotted on a logarithmic scale due to wide range of scale parameter estimates for some climate models and/or climate regions (this is for the display purposes only). The climate models developed by NASA GISS and IPSL give large estimates of the scale parameters for MED, CAS and TIB, while NASA GISS gives large estimates for AMZ, CAM, GRL and NEU. Relatively large estimates of the scale parameters occur together with relatively little smoothness. NASA GISS and IPSL produce rougher temperature anomalies in those regions. In other models and regions, the estimates of scale parameters range from 0.0002 to 1.42 (in logarithmic scale). The standard errors of the smoothness and scale parameter estimates were mostly small, except for the climate models developed by NASA GISS, MOHC, and IPSL that produce rough land surface temperature anomaly fields (not shown).

4. Discussion and conclusion

The smoothness of a spatial process is one of the important measures of spatial dependence. This paper estimates the spatial smoothness of multidecadal averages of land surface temperature anomalies in the NCEP/NCAR reanalysis and the CMIP5 multimodel ensembles by climate region and season. The temperature anomaly field of the NCEP/NCAR reanalysis becomes smoother if the average temperature of the field is extremely high or low. This pattern

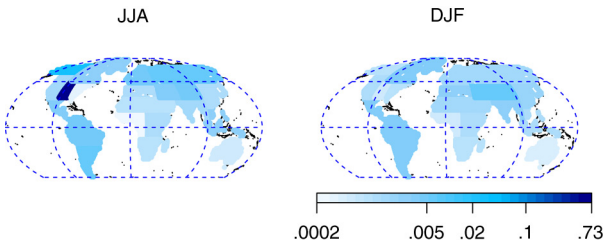

Fig. 10. Scale parameters of the multidecadal average near-surface air temperature anomalies in the NCEP/NCAR reanalysis data by season. The values are plotted on a logarithmic scale.

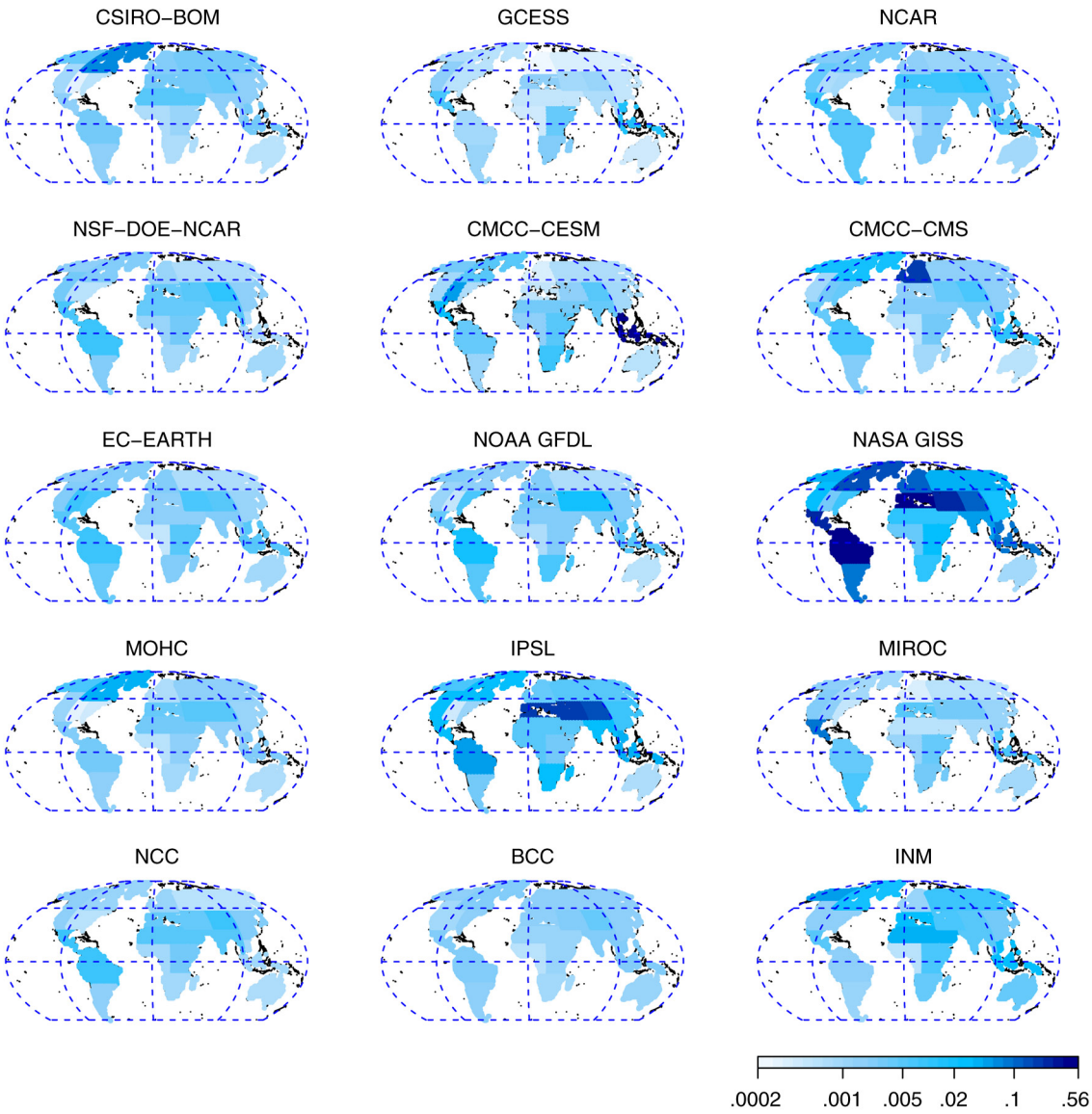

Fig. 11. Scale parameters of the multidecadal average near-surface air temperature anomalies in CMIP5 during JJA. The values are plotted on a logarithmic scale.

appears in some climate models from CMIP5, while the others exhibit similar smoothness during JJA and DJF. The smoothness of the multidecadal average near-surface air temperature anomalies in CMIP5 depends primarily on the modelling institution and the region. Interestingly, there are strong similarities in the smoothness between the climate models generated from the same institution, which supports observations that have been reported frequently in the literature (Tebaldi and Knutti, 2007; Jun et al., 2008a, b; Knutti et al., 2010).

In the future, we plan to examine the smoothness of various climate variables, such as precipitation amount, pressure, wind speed, etc. This can give us more insights into the characteristics of climate regions and climate models. Also, CMIP5 is an archive of well-designed experiments of vast climate models. This paper analyses historical runs (experiment 3.2) of CMIP5 under which all forcings are implemented. Experiments 5.1–5.5 of CMIP5 are composed of simulation runs of the same climate models as in experiment 3.2 but with emissions forcings with fixed or different scenarios of the carbon cycle. The carbon cycle is an important factor that affects surface temperature, as do the presence of sulphate, clouds, interactive aerosols and greenhouse gas emissions. By comparing the smoothness

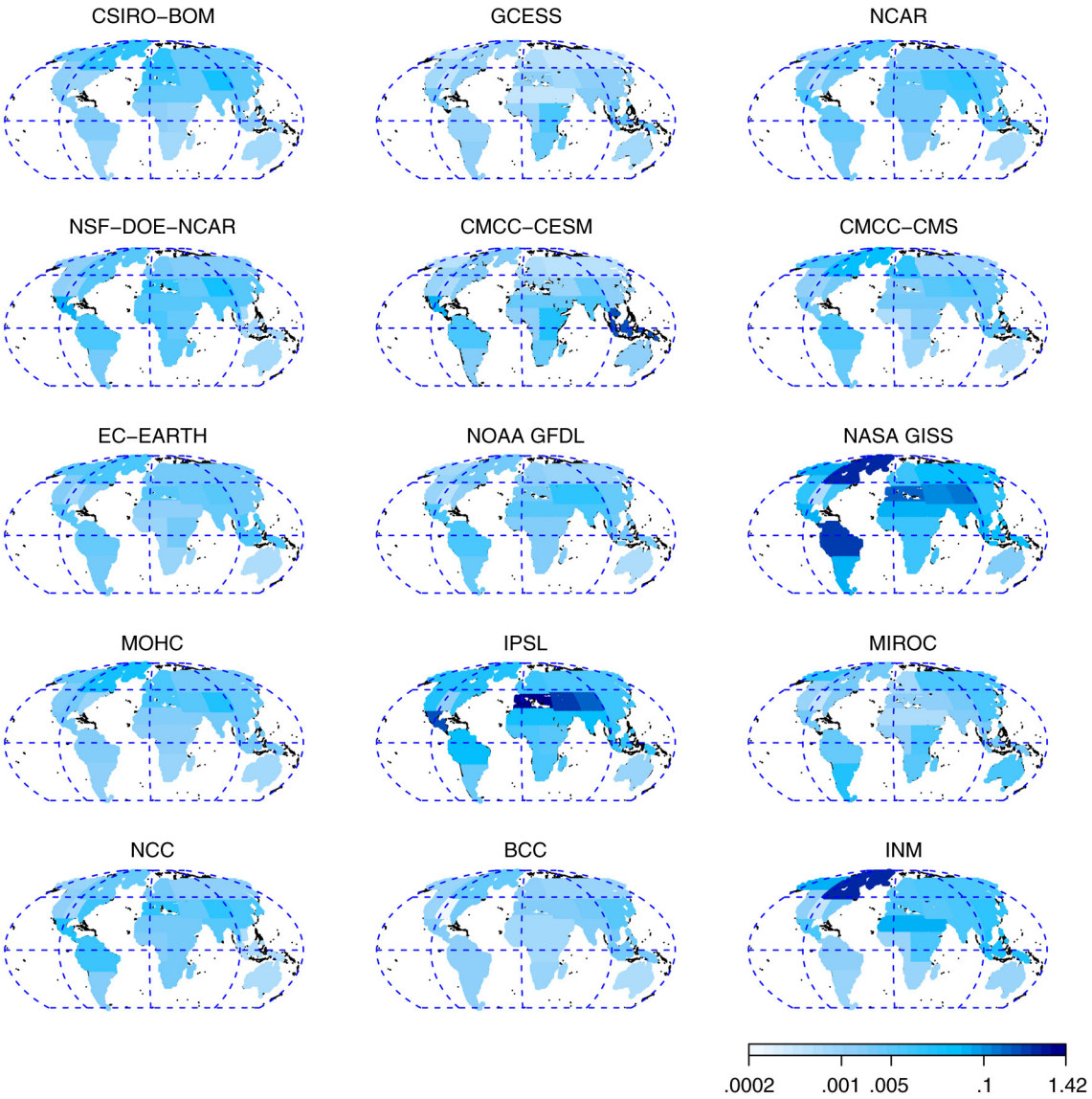

Fig. 12. Scale parameters of the multidecadal average near-surface air temperature anomalies in CMIP5 during DJF. The values are plotted on a logarithmic scale.

between the climate models with or without forcings, we may test the effect of forcings on the local variation of the surface temperature anomalies.

5. Acknowledgements

The authors thank the reviewer for their detailed comments that greatly helped improving the presentation of the paper. The authors also thank Dr. Michael Stein, Dr. Ramalingam Saravanan, Dr. Kenneth Bowman and Dr. Francis Zwiers for their very constructive and helpful comments that led to a number of improvements in this work. This publication is based in part on work supported by Award No. KUS-C1-016-04, made by King Abdullah University of Science and Technology (KAUST). Also, Mikyoung Jun's research was supported by NSF grant DMS-1208421. The authors acknowledge the modelling groups (listed in Table 1 of this paper) for making their simulations available for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the CMIP5 model output, and the World Climate Research Programme (WCRP)'s Working Group on the Coupled Modelling (WGCM) for organising the model data analysis activity. The WCRP CMIP5 multimodel data set is supported by the Office of Science, U.S. Department of Energy. The authors also acknowledge the Physical Sciences

Division of the Earth System Research Laboratory in the National Oceanic and Atmospheric Administration for providing the NCEP/NCAR Reanalysis.

6. Appendix

A.1. Details on composite restricted likelihood

Recall that the composite likelihood, eq. (4), factorises the densities of the error of the best linear predictor of $Z(\mathbf{s}_i)$ based on a few neighbouring observations preceding $\mathbf{s}_i$ in the ordering of the observation locations. Analogously, the composite restricted likelihood factorises the densities of the error of the best linear unbiased predictor (BLUP) of $Z(\mathbf{s}_i)$ based on the few preceding observations in a neighbourhood of $\mathbf{s}_i$. More specifically, let $\mathbf{Z} = (Z(\mathbf{s}_1),\ldots,Z(\mathbf{s}_N))^T$. For $i = k+1,\ldots,N$, let $\mathbf{B}_{i,k}(\boldsymbol{\theta})$ be a vector of length N so that $W_{i,k}(\boldsymbol{\theta}) = \mathbf{B}_{i,k}(\boldsymbol{\theta})^{\mathrm{T}}\,\mathbf{Z}$ is the error of the BLUP of $Z(\mathbf{s}_i)$ based on $Z(\mathbf{s}_{i,1}),\ldots,Z(\mathbf{s}_{i,k})$. For $i = k$, take $\mathbf{B}_{i,k}(\boldsymbol{\theta})$ to be a fixed matrix (independent of $\boldsymbol{\theta}$) of size $N \times (k-1)$ with rank $k-1$ so that $W_{i,k}(\boldsymbol{\theta})$ is a set of contrasts of $Z(\mathbf{s}_1),\ldots, Z(\mathbf{s}_k)$. Then, $W_{i,k}(\boldsymbol{\theta}) \sim \mathcal{N}(0, V_{i,k}(\boldsymbol{\theta}))$, for $V_{i,k}(\boldsymbol{\theta}) = \mathbf{B}_{i,k}(\boldsymbol{\theta})^T \mathrm{Var}(\mathbf{Z})\mathbf{B}_{i,k}(\boldsymbol{\theta})$, $i \geq k$. The *composite restricted log-likelihood* is: $\widetilde{rl}_k(\boldsymbol{\theta}) = -\frac{N-1}{2}\log(2\pi) - \frac{1}{2}\sum_{i=k}^{N}[\log\{\det(V_{i,k}(\boldsymbol{\theta}))\} + W_{i,k}(\boldsymbol{\theta})^T V_{i,k}(\boldsymbol{\theta})^{-1} W_{i,k}(\boldsymbol{\theta})]$. Refer to Stein et al. (2004, Appendix B) for the equations of $\mathbf{B}_{i,k}(\boldsymbol{\theta})$, $\nabla\widetilde{rl}_k(\boldsymbol{\theta})$ and $\nabla^2\widetilde{rl}_k(\boldsymbol{\theta})$ that are required in the sandwich estimator of the variance of the composite restricted likelihood estimator of $\boldsymbol{\theta}$. In the analysis in Section 3, we consider $\widetilde{rl}_k(\boldsymbol{\theta})$ with $k = 3,\ldots,10$. The estimation of $\boldsymbol{\theta}$ is done by profiling out C from $\widetilde{rl}_k(\boldsymbol{\theta})$ and maximising the profiled equation for H over (0,1), by a combination of golden section search and successive parabolic interpolation.

References

Adler, R. J. 1981. *The Geometry of Random Fields*. Wiley, New York.

Genton, M. G., Perrin, O. and Taqqu, M. S. 2007. Self-similarity and Lamperti transformation for random fields. *Stochastic Models* **23**(3), 397–411.

Giorgi, F. and Francisco, R. 2000. Uncertainties in regional climate change prediction: a regional analysis of ensemble simulations with the HADCM2 coupled AOGCM. *Clim Dyn*. **16**(2–3), 169–182.

Gneiting, T. and Schlather, M. 2004. Stochastic models that separate fractal dimension and the Hurst effect. *SIAM Rev*. **46**(2), 269–282. DOI: 10.1137/S0036144501394387.

Gneiting, T., Ševčíková, H. and Percival, D. B. 2012. Estimators of fractal dimension: assessing the smoothness of time series and spatial data. *Stat. Sci*. **27**, 247–277.

Godambe, V. and Heyde, C. 2010. Selected works of C.C. Heyde. In: *Chapter Quasi-Likelihood and Optimal Estimation* (eds. R. Maller, I. Basawa, P. Hall, and E. Seneta). Springer-Verlag, New York, pp. 386–399.

Jun, M. 2011. Nonstationary cross-covariance models for multivariate processes on a globe. *Scand. J. Stat*. **38**, 726–747.

Jun, M. 2014. Matérn-based nonstationary cross-covariance models for global processes. *J. Multivar. Anal*. **128**, 134–146.

Jun, M., Knutti, R. and Nychka, D. W. 2008a. Spatial analysis to quantify numerical model bias and dependence: how many climate models are there? *J. Am. Stat. Assoc*. **103**(483), 934–947.

Jun, M., Knutti, R. and Nychka, D. W. 2008b. Local eigenvalue analysis of CMIP3 climate model errors. *Tellus* **60A**(5), 992–1000.

Jun, M. and Stein, M. L. 2008. Nonstationary covariance models for global data. *Ann. Appl. Stat*. **2**(4), 1271–1289.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D. and co-authors. 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc*. **77**, 437–471.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G. and co-authors. 2001. The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation. *Bull. Am. Meteorol. Soc*. **82**, 247–267.

Kitanidis, P. K. 1983. Statistical estimation of polynomial generalized covariance functions and hydrologic applications. *Water Resour. Res*. **19**(4), 909–921.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J. and Meehl, G. A. 2010. Challenges in combining projections from multiple climate models. *J. Clim*. **23**, 2739–2758.

Knutti, R. and Sedláček, J. 2013. Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Clim. Change* **3**, 369–373.

Lee, M. 2012. Local properties of irregularly observed Gaussian fields. Ph.D. Dissertation, The University of Chicago.

Lindsay, B. G. 1988. Composite likelihood methods. *Contemp. Math*. **80**, 221–239.

Lovejoy, S. and Schertzer, D. 1985. Generalized scale invariance in the atmosphere and fractal models of rain. *Water Resour. Res*. **21**(8), 1233–1250.

Mandelbrot, B. B. and Wallis, J. R. 1969. Robustness of the rescaled range r/s in the measurement of noncyclic long run statistical dependence. *Water Resour. Res*. **5**(5), 967–988. DOI: 10.1029/WR005i005p00967.

North, G. R., Wang, J. and Genton, M. G. 2011. Correlation models for temperature fields. *J. Clim*. **24**, 5850–5862.

Rue, H. and Held, L. 2005. *Gaussian Markov Random Fields: Theory and Applications*. Volume 104 of Monographs on Statistics and Applied Probability. Chapman & Hall, London.

Samorodnitsky, G. and Taqqu, M. 1994. *Stable non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Stochastic Modeling Series. Chapman & Hall, New York.

Stein, M. L. 2007. Spatial variation of total column ozone on a global scale. *Ann. Appl. Stat*. **1**(1), 191–210.

Stein, M. L. 2008. A modelling approach for large spatial datasets. *J. Korean Stat. Soc*. **37**(1), 3–10.

Stein, M. L., Chi, Z. and Welty, L. 2004. Approximating likelihoods for large spatial data sets. *J. R. Stat. Soc. Series B* **66**(2), 275–296.

Sun, Y., Bowman, K., Genton, M. G. and Tokay, A. 2015. A Matérn model of the spatial covariance structure of point rain rates. *Stoch. Environ. Res. Risk Assess*. **29**, 411–416.

Taylor, K., Stouffer, R. and Meehl, G. 2012a. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**(4), 485–498.

Taylor, K. E., Balaji, V., Hankin, S., Juckes, M. and Lawrence, B. 2012b. CMIP5 data reference syntax (DRS) and controlled vocabularies. Online at: http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax.pdf

Tebaldi, C. and Knutti, R. 2007. The use of the multi-model ensemble in probabilistic climate projections. *Philos. Trans. R. Soc. A* **365**(1857), 2053–2075.

Tuck, A. 2008. *Atmospheric Turbulence: A Molecular Dynamics Perspective*. Oxford University Press, Oxford.

Varin, C., Reid, N. and Firth, D. 2011. An overview of composite likelihood methods. *Stat. Sin.* **21**, 5–42.