

Effective Spatial Degrees of Freedom of Natural Temperature Variability as a Function of Frequency

TORBEN KUNZ^a AND THOMAS LAEPPLE^{a,b}

^a *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany*

^b *MARUM—Center for Marine Environmental Sciences, University of Bremen, Bremen, Germany*

(Manuscript received 6 February 2023, in final form 27 January 2024, accepted 31 January 2024)

ABSTRACT: A fundamental statistic of climate variability is its spatiotemporal correlation function. Its complex structure can be concisely summarized by a frequency-dependent measure of the effective spatial degrees of freedom (ESDOF). Here we present, for the first time, frequency-dependent ESDOF estimates of global natural surface temperature variability from purely instrumental measurements, using the HadCRUT4 dataset (1850–2014). The approach is based on a newly developed method for estimating the frequency-dependent spatial correlation function from gappy data fields. Results reveal a multicomponent structure of the spatial correlation function, including a large-amplitude short-distance component (with weak time scale dependence) and a small-amplitude long-distance component (with increasing relative amplitude toward the longer time scales). Two frequency-dependent ESDOF measures are applied, each responding mainly to either of the two components. Both measures exhibit a significant ESDOF reduction from monthly to multidecadal time scales, implying an increase of the effective spatial scale of natural surface temperature fluctuations. Moreover, it is found that a good approximation to the global number of equally spaced samples needed to estimate the variance of global mean temperature is given, at any frequency, by the greater one of the two ESDOF measures, decreasing from ~ 130 at monthly to ~ 30 at multidecadal time scales. Finally, the multicomponent structure of the correlation function together with the detected ESDOF scaling properties indicate that the ESDOF reduction toward the longer time scales cannot be explained simply by diffusion acting on stochastically driven anomalies, as it might be suggested from simple stochastic-diffusive energy balance models.

KEYWORDS: Atmosphere; Statistics; Climate variability

1. Introduction


Global natural surface temperature variability occurs over wide ranges of spatial and temporal scales, and it exhibits a complex spatiotemporal correlation structure. This complex structure can be concisely summarized by simple metrics, characterizing the space–time statistics of the variability. Inherent to such metrics is always a dimension reduction of the spatiotemporal domain.

A common approach to characterize the spatial correlation structure consists in applying a measure of the effective spatial degrees of freedom (ESDOF) to a time series of, for example, global temperature fields. Although various ESDOF measures of different complexity have been proposed in the literature (Livezey and Chen 1983; Smith et al. 1994; Jones et al. 1997; Wang and Shen 1999; Bretherton et al. 1999; Kunz and Laepple 2021, and references therein), each of these measures effectively condenses the entire correlation structure into a single number, which can be interpreted as the effective number of independent spatial samples.

To also include the time scale dependence of the spatial correlation structure, it is possible either to filter the time series before applying an ESDOF measure (Jones et al. 1997) or to apply an explicitly frequency-dependent ESDOF measure to the unfiltered time series (Kunz and Laepple 2021). The latter approach has the advantage that it directly yields ESDOF-frequency spectra, allowing for an evaluation of the ESDOF scaling properties across time scales.

There are various motivations for summarizing the space–time statistics of temperature variability by applying a frequency-dependent ESDOF measure. For example, frequency-dependent ESDOF estimates may provide information regarding the representative spatial scale of a local measurement and its dependence on time scale. Another application consists in determining the global number of samples needed to estimate the variance of global mean temperature at a given time scale. Furthermore, ESDOF-frequency spectra may serve as a simple diagnostic for comparing the space–time statistics between different climate models or between models and observations, and they may provide a basis for the formulation of simple stochastic models of global temperature variability.

In this study we present, for the first time, ESDOF-frequency spectra of global natural surface temperature variability, ranging from monthly to multidecadal time scales, and based exclusively on instrumental measurements. The datasets used are described in section 2, and the methods applied to them are provided by section 3, including the definitions of the frequency-dependent ESDOF measures. The results are presented in section 4, and a discussion and conclusions follow in section 5.

 Denotes content that is immediately available upon publication as open access.

Kunz's current affiliation: Institute of Meteorology, Freie Universität Berlin, Berlin, Germany.

Corresponding author: Torben Kunz, torben.kunz@fu-berlin.de

DOI: 10.1175/JCLI-D-23-0040.1

© 2024 American Meteorological Society. This published article is licensed under the terms of the default AMS reuse license. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

2. Data

a. Instrumental data: HadCRUT4

We use the global gridded (5° longitude \times 5° latitude) deseasonalized surface temperature dataset HadCRUT4 (Morice et al. 2012) that is exclusively based on instrumental measurements, combining ship-based sea surface temperature with land station air temperature data. For this study we select the time period 1850–2014 (165 years). Grid boxes without any observations in a given month are represented as data gaps. The average spatiotemporal coverage of the global dataset during the selected period is about 60%.

Since we are interested in natural temperature variability, we apply a nonlinear detrending procedure to remove the anthropogenic warming signal. Specifically, zonally averaged temperature is regressed, separately at each latitude, onto the global and annual mean time series of the total anthropogenic surface radiative forcing,¹ $F_{\log(\text{CO}_2\text{eq})}(t)$, using the logarithm of the CO_2 -equivalent concentration. The response to anthropogenic forcing is then defined as $b(\phi)F_{\log(\text{CO}_2\text{eq})}(t)$, where $b(\phi)$ is the latitude-dependent regression coefficient. This response is extended in longitude and subtracted from the global temperature fields.

We also investigate the sensitivity of our analysis to variations in the temporal and spatial structure of the calculated response to anthropogenic forcing. To investigate the sensitivity to the temporal structure, the full forcing time series $F_{\log(\text{CO}_2\text{eq})}(t)$ is decomposed into the CO_2 and the remaining non- CO_2 component, denoted by $F_{\log(\text{CO}_2)}(t)$ and $[F_{\log(\text{CO}_2\text{eq})}(t) - F_{\log(\text{CO}_2)}(t)]$, respectively. The latter one is then either increased or decreased by 50%, that is, zonally averaged temperature is now regressed onto $F_{\log(\text{CO}_2)}(t) + a[F_{\log(\text{CO}_2\text{eq})}(t) - F_{\log(\text{CO}_2)}(t)]$, with $a = 0.5$ or 1.5 . This approach is motivated by the fact that the CO_2 contribution to the total anthropogenic forcing is relatively certain, whereas the non- CO_2 contribution is rather uncertain, mainly caused by the uncertainties associated with anthropogenic aerosols. To investigate the sensitivity to the spatial structure, globally (rather than zonally) averaged temperature is regressed onto the full forcing time series, $F_{\log(\text{CO}_2\text{eq})}(t)$, which eliminates the latitudinal structure from the response.

The HadCRUT4 dataset is provided together with detailed error covariance estimates for each month and grid box, which we use to correct our spatial correlation and ESDOF metrics (defined in section 3). Because the errors are assumed to be independent of temperature, and our metrics are all based on second-moment statistics like variances, covariances, and power spectral densities, we can simply compute the same metrics from the HadCRUT4 temperature fields and from random realizations of the errors, and then subtract the latter from the former to obtain error corrected estimates of our metrics.

¹The annual mean time series is linearly interpolated to monthly resolution for the regression. The time series consists of historical data for the period 1850–2004 and is extended by the representative concentration pathway RCP4.5 time series for the period 2005–14 (Meinshausen et al. 2011). A visualization of the total anthropogenic surface radiative forcing time series is given by Fig. 8.18 (red line) in Myhre et al. (2013).

b. Reanalysis: NOAA20CRv3

We use the ensemble mean NOAA Twentieth Century Reanalysis, version 3, global surface temperature dataset (NOAA20CRv3 hereinafter; see Slivinski et al. 2019; selecting again the period 1850–2014, and from which we subtracted the climatological annual cycle, including its higher harmonics) to study the potential impact of the data gaps on our spatial correlation and ESDOF metrics. For this purpose, we interpolate the NOAA20CRv3 temperature fields onto the HadCRUT4 $5^\circ \times 5^\circ$ grid, using a second-order conservative remapping scheme, such that the HadCRUT4 data gaps can be imposed to the NOAA20CRv3 fields. This allows us to compute our metrics from both the complete and the gappy data fields, and to compare the obtained results. We also apply the same nonlinear detrending procedure to remove the anthropogenic warming signal as described in section 2a for HadCRUT4. Using the reanalysis has the advantage that it is based on the same trajectory of internal climate variability as the instrumental observations. Thus, it allows us to investigate the interaction between this trajectory and the specific spatiotemporal distribution of the HadCRUT4 data gaps.

c. Climate models: CMIP6

To investigate the estimation bias and uncertainty of our spatial correlation and ESDOF metrics, we use the global surface temperature fields from an ensemble of Coupled Model Intercomparison Project phase 6 (CMIP6) climate model simulations (Eyring et al. 2016) from which we subtracted the climatological annual cycle, including its higher harmonics. Specifically, we analyze simulations of length 165 years of the preindustrial control (CMIP6-piCtrl) experiment which includes no external forcings and, thus, generates only internal climate variability. We employ 27 climate models (listed in Table A1 in appendix A) from each of which we use 3 independent simulations, resulting in an ensemble of 81 members in total. As for NOAA20CRv3, all CMIP6 temperature fields are interpolated onto the HadCRUT4 $5^\circ \times 5^\circ$ grid, which allows us to impose the HadCRUT4 data gaps and, thus, to investigate the impact of the gaps on the results for the climate models in an ensemble mean sense, and where the trajectories of internal climate variability are independent of the observed trajectory.

3. Methods

In this study we use two different frequency-dependent ESDOF measures, introduced previously by Kunz and Laepple (2021). The first measure is defined as

$$D(f) = 1/M[R(\theta; f)], \quad (1)$$

where $R(\theta; f)$ denotes the frequency-dependent spatial correlation function, $\theta \in [0, \pi]$ is the angular distance between two locations on the globe (the angle between them as seen from the center of Earth), f is frequency, and the operator $M[x(\theta)] = [\int_0^\pi x(\theta) \sin\theta d\theta]/2$ represents the area-weighted global mean of any radial function $x(\theta)$. Specifically, $R(\theta; f) = C(\theta; f)/C(0; f)$, where $C(\theta; f)$ is the spatial covariance of surface temperature variability at frequency f , averaged over

all pairs of locations separated by an angular distance θ . The procedure to estimate $C(\theta; f)$ from a time series of global gridded temperature fields follows the approach of [Kunz and Laepple \(2021\)](#) that uses spherical harmonic and Fourier decompositions for the transformation from longitude, latitude, and time to angular distance θ and frequency f . Here we apply an advanced variant of that approach which is capable of dealing with gappy temperature fields, that is, with fields that include empty grid boxes due to missing observations (see [appendix B](#) for details).

It can be shown (see [Kunz and Laepple 2021](#)) that $M[R(\theta; f)] = M[C(\theta; f)]/C(0; f) = S_{\text{gib}}(f)/S_{\text{loc}}(f)$, where $S_{\text{gib}}(f)$ is the power spectral density of the global mean and $S_{\text{loc}}(f)$ is the global mean of the local power spectral density of surface temperature anomalies. Thus, the above frequency-dependent ESDOF measure (1) can also be expressed as [see [Kunz and Laepple 2021](#), their Eq. (20)]

$$D(f) = S_{\text{loc}}(f)/S_{\text{gib}}(f). \quad (2)$$

The basic interpretation of this ESDOF measure is as follows. If all grid boxes of the global temperature field are perfectly correlated (and have equal power spectral density), at a given frequency f , then $S_{\text{gib}}(f) = S_{\text{loc}}(f)$ and, thus, $D(f) = 1$. On the other hand, if there are N uncorrelated (and equally weighted) grid boxes, then $S_{\text{gib}}(f) = S_{\text{loc}}(f)/N$ and, thus, $D(f) = N$. In applications to global temperature fields, $D(f)$ typically attains values between 1 and N [for a detailed discussion of the measure, see [Kunz and Laepple \(2021\)](#)]. Note that a frequency-independent version of this measure can be defined that is identical to the ESDOF measure of [Jones et al. \(1997\)](#) according to their Eq. (10).²

The second frequency-dependent ESDOF measure, used in this study, is defined as [see [Kunz and Laepple 2021](#), their Eq. (26)]

$$D_{\text{fit}}(f) = 1/M[R_{\text{fit}}(\theta; f)], \quad (3)$$

where $R_{\text{fit}}(\theta; f) = \exp[-\theta/\theta_e(f)]$ is an exponential correlation function, the e -folding scale of which matches that of $R(\theta; f)$, that is, $R[\theta_e(f); f] = 1/e$. By analogy with our first ESDOF measure, a frequency-independent version can also be defined of our second measure, which corresponds to the second ESDOF measure of [Jones et al. \(1997\)](#) according to their Eq. (14),³ with the exception that they use a different normalization procedure to obtain the correlation function from which θ_e is determined. In summary, our first measure, $D(f)$, represents a summarizing metric of the entire radial correlation structure of the global temperature field, whereas our second measure, $D_{\text{fit}}(f)$, depends only on the e -folding scale $\theta_e(f)$.

² This frequency-independent measure is defined by $D = \sigma_{\text{loc}}^2/\sigma_{\text{gib}}^2$, where $\sigma_{\text{loc}}^2 = \int_{-\infty}^{\infty} S_{\text{loc}}(f) df$ is the global mean of the local variance and $\sigma_{\text{gib}}^2 = \int_{-\infty}^{\infty} S_{\text{gib}}(f) df$ is the variance of the global mean.

³ Note that from their Eq. (14), and realizing that the expression \bar{x}_0/R in their notation is equal to θ_e in our notation, we can rewrite our above frequency-dependent ESDOF measure in Eq. (3) as $D_{\text{fit}}[\theta_e(f)] = 2[1 + 1/\theta_e^2(f)]/[1 + \exp[-\pi/\theta_e(f)]]$.

To investigate the estimation bias and uncertainty of the two ESDOF measures we use the CMIP6 climate model ensemble [as the theoretical expressions for the estimation bias and uncertainty, derived by [Kunz and Laepple \(2021\)](#), are only valid for the first measure $D(f)$ and only if it is applied to complete data fields]. The ensemble allows us to define an unbiased and a biased ensemble mean estimator, the difference of which equals the expected estimation bias of an ESDOF estimate obtained from a single realization of temperature variability, as given by HadCRUT4. The ensemble is also used to quantify the expected estimation uncertainty by investigating the ensemble spread (see [appendix C](#) for details of the bias and uncertainty analysis).

Both ESDOF measures can be translated into an associated length scale, defined as the radius of a spherical cap, the area of which covers that fraction of the globe that is equal to the reciprocal of the ESDOF measure. It can be shown that this definition implies (see [Kunz and Laepple 2021](#), their appendix A) that

$$L(f) = r \arccos[1 - 2/D(f)] \quad (4)$$

and

$$L_{\text{fit}}(f) = r \arccos[1 - 2/D_{\text{fit}}(f)], \quad (5)$$

where r denotes the radius of Earth. These length scales can be interpreted as an effective correlation radius. Note, that the e -folding length, defined as

$$L_e(f) = r\theta_e(f), \quad (6)$$

which is simply the e -folding scale $\theta_e(f)$ expressed in units of length, is not equal to $L_{\text{fit}}(f)$ because of the spherical geometry of the spatial domain.

In addition to estimating the spatial correlation functions $R(\theta; f)$ and $R_{\text{fit}}(\theta; f)$ at each specific frequency f , they are also estimated for three different frequency bands,⁴ denoted as the multidecadal, interannual, and subannual band (defined in [Table 1](#)). From these frequency-band correlation functions, $R(\theta)$ and $R_{\text{fit}}(\theta)$, the corresponding frequency-band values of D , D_{fit} , L , L_{fit} , and L_e are computed by analogy with (1), (3), (4), (5), and (6), respectively.

One potential application of a global ESDOF measure consists in using its value, after rounding it to the nearest integer N , as the global number of equally spaced samples that is needed to estimate the variance of the global mean, σ_{gib}^2 . Likewise, the value of a frequency-dependent ESDOF measure can be used to estimate $S_{\text{gib}}(f)$. This application makes

⁴ Note that the order of operation is as follows: first, integration across the frequency band is performed on $C(\theta; f)$ to obtain the frequency-band covariance function $C(\theta)$ and, second, normalization is performed to obtain the frequency-band correlation function $R(\theta) = C(\theta)/C(0)$, from which θ_e is determined to obtain $R_{\text{fit}}(\theta)$, because the opposite order (normalizing before integration) would create an additional estimation bias in $R(\theta)$ and, consequently, in θ_e and $R_{\text{fit}}(\theta)$, caused by the scattering estimator $C(0)$ in the denominator of the integrand.

TABLE 1. Names of frequency bands, associated frequency ranges, and bias-corrected frequency-band values of the ESDOF measures D and D_{fit} and of the length scales L and L_{fit} , together with the estimation uncertainty intervals indicated in brackets, estimated from the nonlinearly detrended HadCRUT4 temperature fields.

f band	f range	D	D_{fit}	L (10^3 km)	L_{fit} (10^3 km)
Multidecadal	$1/(165 \text{ yr}) \leq f < 1/(10 \text{ yr})$	10.9 [7.27–16.2]	29.3 [24.9–34.5]	3.93 [3.20–4.84]	2.37 [2.18–2.57]
Interannual	$1/(10 \text{ yr}) < f < 1/(1 \text{ yr})$	25.5 [22.8–28.5]	34.4 [33.6–35.2]	2.54 [2.40–2.69]	2.19 [2.16–2.21]
Subannual	$1/(1 \text{ yr}) < f \leq 1/(2 \text{ mon})$	128 [121–135]	44.2 [43.8–44.7]	1.13 [1.10–1.16]	1.92 [1.91–1.93]

sense in situations where one is given only sparse spatial data, or even has to expensively collect the data first. For example, if the variance of the global mean in a specific frequency band is to be estimated from a past period where only data from paleoclimate proxies are available or have to be collected, then, given an ESDOF value (obtained from high-resolution instrumental data), its nearest integer N may be used as a first guess (or lower bound) for the global number of samples (proxy locations) needed to obtain a reasonable variance estimate. This approach works best if the underlying spatial fields have the structure of (discrete) white noise. For more complex correlation structures, as it is found for surface temperature or any other climate variable, however, ESDOF values may imply too small sample numbers N , leading to an overestimation of the variance of the global mean. It is, therefore, meaningful to investigate the extent of variance overestimation that has to be expected for the various ESDOF measures across frequencies, given the frequency-dependent spatial covariance function obtained from the HadCRUT4 instrumental dataset.

Given the relatively high spatial resolution of HadCRUT4, the frequency-dependent spatial covariance function $C(\theta; f)$ is estimated at a sufficiently high accuracy such that we can treat the power spectral density of the global mean, obtained as $S_{\text{glob}}(f) = M[C(\theta; f)]$, as its true value. In addition, we can use $C(\theta; f)$ to compute the expected power spectral density of the global mean, $S_{\text{glob},N}(f)$, if it were estimated from N equally spaced samples around the globe. Specifically, we obtain $S_{\text{glob},N}(f)$ by taking the mean over N equally spaced samples of the HadCRUT4 covariance function $C(\delta; f)$, using the coordinate $\delta = -\cos\theta \in [-1, 1]$ to account for area weighting. From this, the expected variance overestimation can be expressed as

$$\Delta S_{\text{glob},N}(f) = S_{\text{glob},N}(f) - S_{\text{glob}}(f). \quad (7)$$

Then, substituting N by $D(f)$ or $D_{\text{fit}}(f)$, rounded to an integer, yields the expected variance overestimation when using the ESDOF value as the global sample number. If this is expressed as the percentage of relative overestimation, $p_N(f) = \Delta S_{\text{glob},N}(f)/S_{\text{glob}}(f) \times 100$, and, additionally, a required maximum percentage of relative overestimation, p_0 , is set, it can be checked which ESDOF measure at which frequencies fulfills the required condition $p_N(f) < p_0$; with N again being substituted by a rounded ESDOF value.

Conversely, one may ask for the number of samples $N_{p_0}(f)$ that yields the required percentage of relative overestimation p_0 . To obtain $N_{p_0}(f)$, we first determine $p_N(f)$ for a suitable

range of integer values of N . Then linear interpolation between those N associated with the p values closest to the required value p_0 yields the (generally real) value $N_{p_0}(f)$.

4. Results

The frequency-dependent ESDOF measure $D(f)$, estimated from the nonlinearly detrended HadCRUT4 temperature fields, exhibits a notable reduction from monthly toward multidecadal time scales (Fig. 1a, red line). In terms of the frequency bands defined in Table 1, global natural surface temperature variability has more than 100 ESDOFs in the subannual frequency band and just above 10 ESDOFs in the multidecadal band. When the same measure is estimated from CMIP6-piCtrl temperature fields (with HadCRUT4 gaps imposed), the ensemble median ESDOF spectrum $D(f)$ exhibits a similar behavior (Fig. 1a, black line), but values are roughly 25% larger across the entire frequency range. This ESDOF spectrum appears as a superposition of two components, namely (i) an almost uniform power-law scaling across all frequencies, that is, following $D(f) \sim f^{\beta_D}$, with scaling exponent β_D , and (ii) a pronounced ENSO signature characterized by smaller ESDOF values at interannual time scales, reflecting large-scale coherent fluctuations associated with ENSO-related teleconnections. In terms of the 5% to 95% quantile range of the CMIP6-piCtrl climate model ensemble (Fig. 1a, gray shading), the HadCRUT4 ESDOF spectrum appears to be consistent with the climate models, although the superposition of the two components would be less discernible from the HadCRUT4 spectrum alone because of the estimation uncertainty.

The consistency between CMIP6-piCtrl and detrended HadCRUT4 justifies the use of the model ensemble spread and bias as an estimate of the estimation error of the HadCRUT4 ESDOF spectrum. The bias-corrected HadCRUT4 ESDOF spectrum and frequency-band values are shown in Fig. 1c (red lines), ranging from 128 ESDOFs in the subannual to 25.5 ESDOFs in the interannual to 10.9 ESDOFs in the multidecadal frequency band, corresponding to associated length scales L of 1.13×10^3 , 2.54×10^3 , and 3.93×10^3 km, respectively (Table 1). Figure 1c also indicates that the bias-corrected HadCRUT4 ESDOF spectrum follows roughly a power-law scaling with scaling exponent $\beta_D \approx 0.5$. The uncertainty intervals in D (and associated L), applied to the bias-corrected HadCRUT4 frequency-band values, are also indicated in Fig. 1c (vertical lines) and specified in Table 1. Note that, as shown in Fig. 1a (black whiskers), the estimation uncertainty alone is indeed smaller than the total uncertainty including the

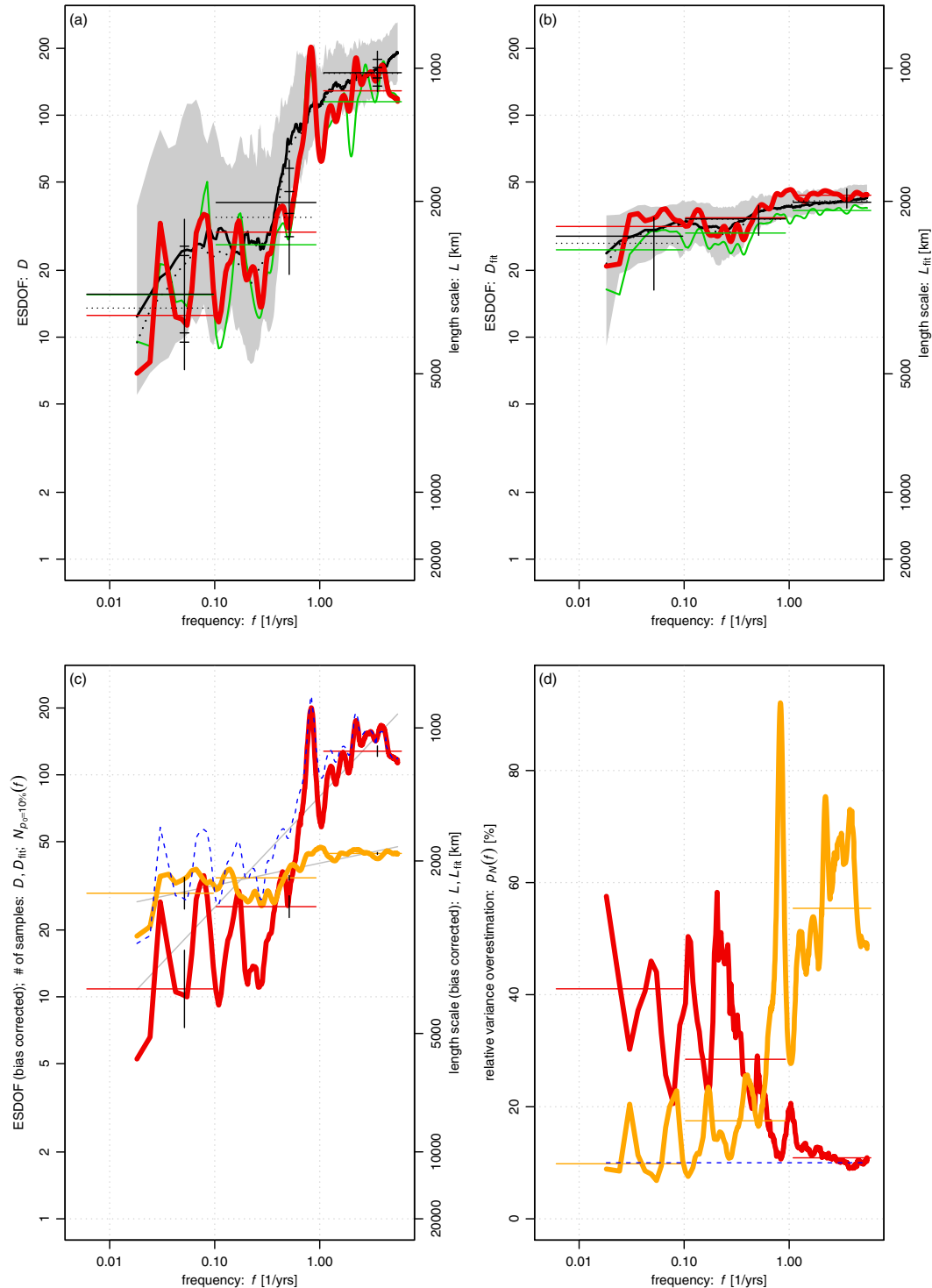


FIG. 1. (a) ESDOF measure $D(f)$: HadCRUT4 (red), NOAA20CRv3 (green), CMIP6-piCtrl biased (black solid) and unbiased (black dotted) ensemble estimator, and 5%–95% quantile range (gray shading). Also shown are frequency-band values (horizontal lines), with quantile ranges (vertical black lines) and uncertainty intervals (black whiskers—inner: estimation; outer: total). (b) As in (a), but for $D_{fit}(f)$, and without uncertainty intervals. (c) Bias-corrected HadCRUT4 $D(f)$ (red) and $D_{fit}(f)$ (orange), frequency-band estimation uncertainties (vertical black lines), sample number $N_{p_0=10\%}(f)$ (blue dashed), and power-law scaling with exponents $\beta = 0.1$ and $\beta = 0.5$ (gray lines). (d) Relative variance overestimation $p_N(f)$ using bias-corrected HadCRUT4 $D(f)$ (red) and $D_{fit}(f)$ (orange) as sample number, and $p_0 = 10\%$ (blue dashed). Spectra include a log-frequency smoothing.

intermodel spread, and that the relative difference between them is smallest in the multidecadal frequency band.

Since the ESDOF measure $D(f)$ is based on the spatial integral of the frequency-dependent spatial correlation function $R(\theta; f)$, inspection of the latter helps to understand the behavior of the former. The spatial correlation function estimated from the detrended HadCRUT4 temperature fields is shown in Figs. 2a–c for the three frequency bands. The structure of these correlation functions suggests that it consists of three components: (i) a strongly decaying short-distance component that dominates the correlation structure at short distances ($< 2 \times 10^3$ km), (ii) a weakly decaying long-distance component that dominates at larger distances, most clearly seen in the multidecadal frequency band, and (iii) an oscillatory component that reflects anticorrelated teleconnections, seen most clearly in the subannual band, with a wavelength of about 6.7×10^3 km. This HadCRUT4 correlation structure appears to be consistent with the CMIP6-piCtrl climate model ensemble in terms of the 5% to 95% quantile range.

To relate the changes of $D(f)$ across frequencies to the changes in the structure of $R(\theta; f)$, recall that the integral of the spatial correlation function involves the area-weighting factor $\sin\theta$ (see the operator $M[\cdot]$ in section 3). Accordingly, Figs. 2d and 2e show the difference of the area-weighted correlation function between the multidecadal and interannual, and between the interannual and subannual frequency bands, respectively. This difference illustrates how the contribution to the integral over the correlation function is distributed across angular distance θ . It turns out that the bulk contribution to the changes between frequency bands comes from large distances, where the large-distance component of the correlation function dominates over the short-distance component. Thus, the reduction of the ESDOF measure $D(f)$ toward the longer time scales is caused by an increase in relative amplitude of the weakly decaying long-distance component. Since, however, even in the multidecadal frequency band the relative amplitude of this weakly decaying component is small compared to the strongly decaying short-distance component, the e -folding length L_e of the full correlation function (indicated inside Figs. 2a–c) undergoes only little change across frequencies (less than 20% between the subannual and the multidecadal band).

This particular multicomponent structure of the spatial correlation function in the multidecadal frequency band, with strongly decaying short-distance correlations followed by a long tail of weak long-distance correlations, may lead to significant overestimation of the variance of global mean temperature, when estimated from a finite number of $D(f)$ equally spaced samples across the globe. The expected relative overestimation, $\Delta S_{\text{glob}}(f)/S_{\text{glob}}(f)$, defined by (7), is shown in Fig. 1d (red line). The variance overestimation is indeed largest in the multidecadal frequency band where it amounts to about 40%, and it decreases to only about 10% in the subannual band. This implies that the ESDOF measure $D(f)$ provides a suitable estimate of the number of samples needed, for estimating the variance of global mean temperature, only at subannual scale, but not at longer time scales where the

strongly decaying short-distance component of the correlation function is effectively undersampled.

Since this short-distance component of the correlation function is characterized by the e -folding length L_e even in the multidecadal frequency band, the above issue may be solved by using the alternative ESDOF measure $D_{\text{fit}}(f)$, which depends only on L_e but not on the long-distance component. The measure $D_{\text{fit}}(f)$, estimated from the detrended HadCRUT4 temperature fields, is shown in Fig. 1b (red line). As in the case of $D(f)$, the measure $D_{\text{fit}}(f)$ is again highly consistent with estimates from the CMIP6-piCtrl model ensemble (Fig. 1b, black line and gray shading), but its frequency dependence is much weaker than that of $D(f)$, in accordance with the weak frequency-dependence of L_e . Using the model ensemble spread and bias as before, bias-corrected HadCRUT4 ESDOF estimates in terms of $D_{\text{fit}}(f)$ are obtained (Fig. 1c, orange line), ranging from 44.2 ESDOFs in the subannual to 34.4 ESDOFs in the interannual to 29.3 ESDOFs in the multidecadal frequency band, corresponding to associated length scales L_{fit} of 1.92×10^3 , 2.19×10^3 , and 2.37×10^3 km, respectively. Thus, $D_{\text{fit}}(f)$ is roughly 3 times smaller than $D(f)$ in the subannual, and roughly 3 times larger than $D(f)$ in the multidecadal frequency band. Bias-corrected values and uncertainty intervals are specified in Table 1. Figure 1c also indicates that the bias-corrected HadCRUT4 ESDOF spectrum follows roughly a power-law scaling, $D_{\text{fit}}(f) \sim f^{\beta_{D_{\text{fit}}}}$, with scaling exponent $\beta_{D_{\text{fit}}} \approx 0.1$. Note that in the multidecadal frequency band the $\log(D_{\text{fit}})$ model ensemble distribution (Fig. 1b, vertical black line) is highly asymmetric with a skew toward lower ESDOF values, such that the corresponding uncertainty interval is only a rough guide and both the upper and lower limit must be expected to be too high.

The difference between $D(f)$ and $D_{\text{fit}}(f)$ can be related to the (area-weighted) difference between the underlying spatial correlation functions, $R(\theta; f)$ and $R_{\text{fit}}(\theta; f)$ (Figs. 2f–h). In the multidecadal frequency band (Fig. 2f, red line), this difference is largely due to the long-distance component of $R(\theta; f)$, whereas in the subannual band (Fig. 2h, red line) this difference is due to the oscillatory component of $R(\theta; f)$ (reflecting anticorrelated teleconnections), both of which are absent, by construction, in the exponential correlation function $R_{\text{fit}}(\theta; f)$. In the interannual frequency band (Fig. 2g, red line) these two opposite effects on the integral of the spatial correlation function and, thus, on $D(f)$ and $D_{\text{fit}}(f)$, cancel each other out to some extent, such that the difference between the ESDOF measures is relatively small in this band (Fig. 1c).

When using the measure $D_{\text{fit}}(f)$ as the number of equally spaced samples across the globe for estimating the variance of global mean temperature, the expected relative overestimation, $\Delta S_{\text{glob}}(f)/S_{\text{glob}}(f)$, again based on the full correlation function $R(\theta; f)$, is smallest in the multidecadal frequency band where it amounts to about 10%, and it increases to about 55% in the subannual band (Fig. 1d, orange line). It, therefore, exhibits the opposite behavior to the case of using $D(f)$ as the sample number. Hence, when accepting a variance overestimation of 10% to 20%, the measure $D_{\text{fit}}(f)$ provides a suitable estimate of the required sample number at frequencies below 1/(2 yr) (Fig. 1d), and the measure $D(f)$ provides a

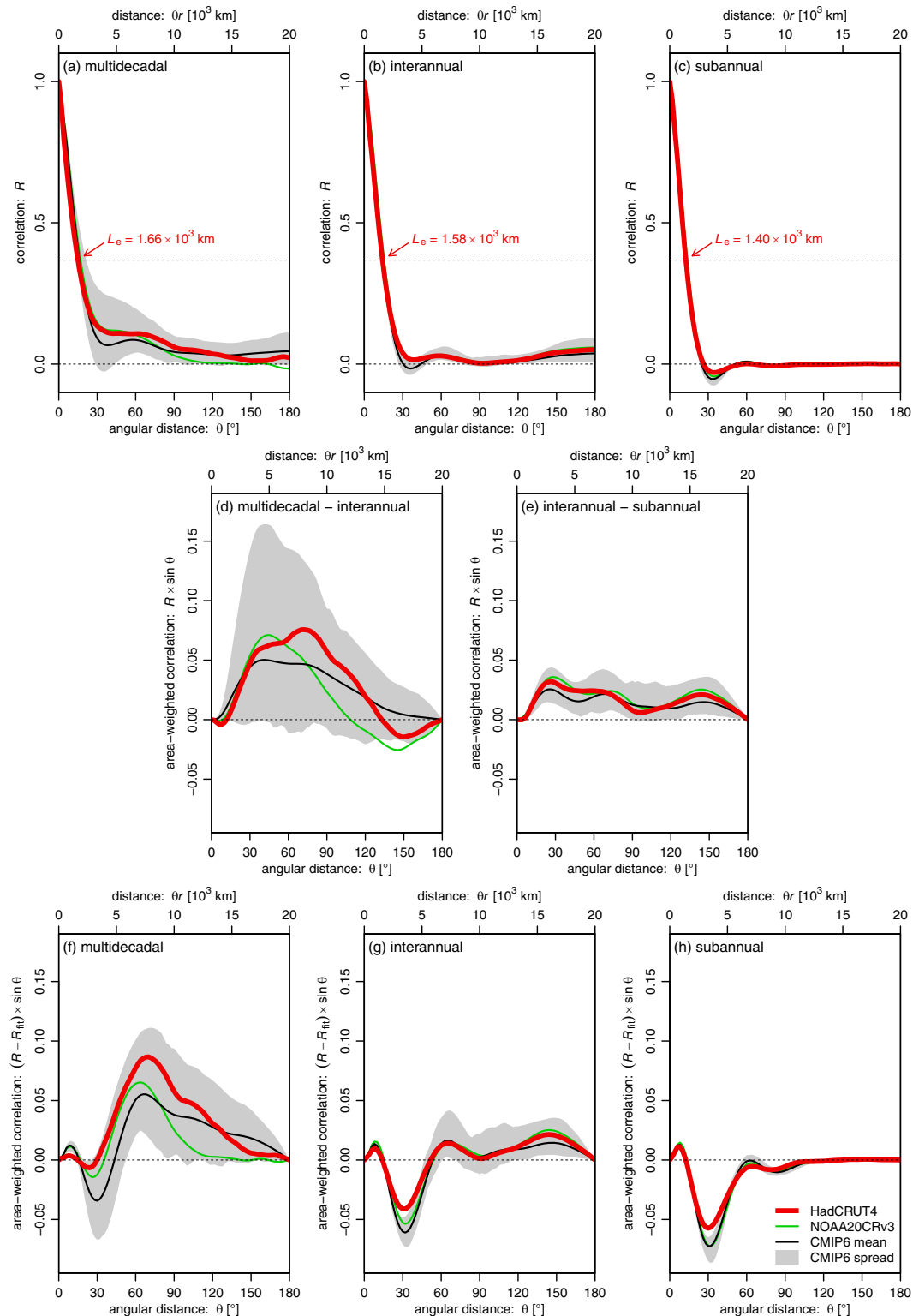


FIG. 2. Spatial correlation function $R(\theta)$ for the (a) multidecadal, (b) interannual, and (c) subannual frequency band: HadCRUT4 (red), NOAA20CRv3 (green), CMIP6-piCtrl biased ensemble estimator (black), and 5%–95% quantile range (gray shading). (d),(e) As in (a)–(c), but for frequency-band differences and area-weighted correlation functions. (f)–(h) As in (a)–(c), but showing the difference between $R(\theta)$ and $R_{\text{fit}}(\theta)$, and for area-weighted correlation functions. The HadCRUT4 e -folding length L_e is indicated by red labels in (a)–(c).

suitable estimate at the higher frequencies. Conversely, when requiring an overestimation of only 10% across all frequencies (Fig. 1d, blue dashed line), the implied number of equally spaced samples across the globe, $N_{p=10\%}$, increases from about 30 at multidecadal to about 130 at monthly time scales (Fig. 1c, blue dashed line). It turns out that at any frequency the greater one of the two measures, $D(f)$ and $D_{\text{fit}}(f)$, provides a suitable estimate of the sample number needed for estimating the variance of global mean temperature. At the low frequencies $D(f)$ provides a too small sample number because the short-distance component is effectively undersampled. At the high frequencies $D_{\text{fit}}(f)$ provides a too small sample number because the narrow θ -interval with negative correlations of the oscillatory component is undersampled.

The results presented in this section, obtained from the gappy HadCRUT4 temperature fields, were shown to be in close agreement with the corresponding results obtained from the CMIP6-piCtrl model ensemble with HadCRUT4 gaps imposed. Figures 1a, 1b, and 2 also include the results obtained from applying the same analysis to the detrended NOAA20CRv3 temperature fields with HadCRUT4 gaps imposed (green lines), and again good agreement with HadCRUT4 results is found. Therefore, both the model ensemble and the reanalysis can be used to investigate the impact of the data gaps on the results, by repeating the analysis for the complete temperature fields and comparing the results to those obtained from the gappy data fields. It is found (not shown) that the impact of the gaps is small in both cases in relation to the estimation uncertainty. Hence, it can be concluded that the data gaps, reflecting a lack of observations in certain months and grid boxes, do not significantly affect the results. Note, that the spectral peak of the ESDOF measure $D(f)$ at periods near 15 months, seen in both HadCRUT4 and NOAA20CRv3 with gaps imposed (Figs. 1a,c), is found to be absent in the corresponding ESDOF spectra obtained from the complete NOAA20CRv3 fields (not shown). Thus, this peak occurs by chance and is related to the additional scatter of the ESDOF estimator in the presence of data gaps (see appendix B for details).

The sensitivity of the results to the details of the nonlinear detrending procedure, intended to remove the anthropogenic warming signal, is also investigated. It is found (not shown) that neither the variation of the latitudinal nor of the temporal structure of the estimated response to anthropogenic forcing leads to a significant change of the results presented in this section.

To ensure the ESDOF estimates have converged toward their true value, given the finite spatial resolution of the HadCRUT4 5° longitude \times 5° latitude grid, we recomputed both ESDOF measures for various spatial resolutions. As demonstrated in appendix D, convergence is actually guaranteed for both measures in all frequency bands.

5. Discussion and conclusions

In section 4, ESDOF estimates of global natural surface temperature variability, as a function of frequency, have been presented, obtained from purely instrumental measurements

(HadCRUT4, using 1850–2014); and the estimates have been translated into an effective spatial scale (i.e., the effective correlation radius) of natural temperature fluctuations. Additionally, it has been shown how these ESDOF estimates can be used to determine the minimum global number of equally spaced samples needed to estimate the variance of global mean temperature. Since these results are based on the averaged spatial correlation function, although global temperature variability is spatially nonstationary, the derived minimum number of samples must, therefore, be understood as an expected sample number across all possible sets of locations. In practice, the minimum sample number will be sensitive to the specific set of locations chosen. Nonetheless, the derived estimates of the required minimum sample number provide an important benchmark for this quantity.

This study focuses on natural surface temperature variability from instrumental data, and the CMIP6-piCtrl model ensemble was used only to obtain bias and uncertainty estimates for the detrended HadCRUT4 ESDOFs, and to investigate the impact of the data gaps on the results. Nonetheless, it is noteworthy that the results from detrended HadCRUT4 and from CMIP6-piCtrl are largely consistent, in terms of both the ESDOFs and the radial correlation structure. Whereas CMIP6-piCtrl represents exclusively internal climate variability, the detrended HadCRUT4 temperature fields represent internal variability plus the responses to natural external forcings. This suggests that natural external forcings do not notably impact the space–time statistics of global surface temperature variability in the range from monthly to multidecadal time scales in the global mean sense. However, this does not exclude the possibility of forcing-induced changes in the space–time statistics at more regional scales, compensating between different regions.

The physical processes underlying the detected ESDOF reduction toward the lower frequencies are yet to be identified. Simple stochastic-diffusive energy balance models (EBMs; see, for example, North et al. 2011; Rypdal et al. 2015), which are sometimes used as paradigmatic models of global natural surface temperature variability, may suggest horizontal diffusion as the primary underlying physical process. Within the diffusive frequency regime, these EBMs exhibit power-law scalings at the local and the global scale, that is, $S_{\text{loc}}(f) \sim f^{\beta_{\text{loc}}}$ and $S_{\text{gib}}(f) \sim f^{\beta_{\text{gib}}}$ (with $\beta_{\text{loc}}, \beta_{\text{gib}} < 0$), such that $\beta_{\text{gib}} = 2\beta_{\text{loc}}$ (Rypdal et al. 2015, their Fig. 8). Together with definition (2), this implies that the ESDOF measure $D(f)$ exhibits a power-law scaling with exponent $\beta_D = \beta_{\text{loc}} - \beta_{\text{gib}} = -\beta_{\text{loc}}$. Since for detrended HadCRUT4, it is $\beta_D \approx 0.5$ (Fig. 1c) and $\beta_{\text{loc}} \approx -0.5$ (not shown), observed natural temperature variability may appear consistent with the EBM behavior. However, the detected multicomponent structure of the HadCRUT4 spatial correlation function (Fig. 2) is inconsistent with the EBMs, which exhibit a single-component correlation function [see the frequency-dependent correlation functions derived by Rypdal et al. (2015), illustrated by their Figs. 1 and 6]. One may then ask whether the short-distance component of the HadCRUT4 correlation function alone, approximated by the exponential correlation function underlying the ESDOF measure $D_{\text{fit}}(f)$, might be consistent with the diffusive EBM behavior. However, for HadCRUT4 it is $\beta_{D_{\text{fit}}} \approx 0.1$ (Fig. 1c)

which is inconsistent with $\beta_{\text{loc}} \approx -0.5$. Hence, the space–time statistics, in a global mean sense, of observed natural surface temperature variability cannot be explained simply by diffusion acting on stochastically driven anomalies. Future studies that systematically quantify the various components of the frequency-dependent spatial correlation function may help to reveal the underlying physical processes and to formulate suitable stochastic models of natural surface temperature variability.

Note, that we also computed the frequency-dependent spatial correlation function for various spatial subdomains of the globe, namely, the tropics (between 30°N and 30°S), the extratropics (poleward of 30°N and of 30°S), the global land, and the global sea areas. For each of the four subdomains the frequency-dependent spatial correlation function exhibits a multicomponent structure (not shown)⁵ similar to that obtained from the global analysis (Figs. 2a–c), particularly in the multidecadal frequency band. This indicates that this multicomponent structure does not simply result from a superposition of the different spatial correlation functions associated with the various subdomains, but that it rather reflects an intrinsic feature of the spatial correlation structure of natural surface temperature variability.

It is an interesting question how the ESDOF-frequency scaling may continue beyond multidecadal time scales resolved by the instrumental record. As long as the underlying physical processes are not clarified, it is unclear whether the observed ESDOF reduction toward the longer time scales can be expected to continue. The main climate drivers at centennial and longer time scales are probably related to natural external forcings like variations in greenhouse gas and volcanic aerosol concentrations. Because the response to such forcings can be expected to be of near-global scale, it is at least a plausible assumption that the ESDOFs do not increase again beyond multidecadal time scales. Hence, the ESDOF values at multidecadal time scales presented in this study (Table 1) are likely to represent upper bounds to the slower variations.

To investigate the ESDOF-frequency scaling at time scales longer than the instrumental period in future studies, the presented methodology may be extended and applied to collections of paleoclimate data such as PAGES2k (PAGES2k Consortium 2017). By providing statistical information on the

time scale dependence of the effective number of independent spatial samples on the globe, such studies are potentially useful for data assimilation efforts aiming at global paleoclimate field reconstruction. Note, however, that paleoclimate proxies are always associated with noise for which the ESDOF results had to be corrected, and that this noise has its own correlation structure (Kunz et al. 2020; Dolman et al. 2021) to be taken into account when correcting for noise.

Acknowledgments. This is a contribution to the SPACE ERC project; this project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement 716092). The work profited from discussions at the Climate Variability Across Scales (CVAS) working group of the Past Global Changes (PAGES) program. This project was supported by the Informationsinfrastrukturen Grant of the Helmholtz Association as part of the DataHub Earth and Environment. We also thank Igor Kröner for assistance in preprocessing the CMIP6 model data.

Data availability statement. The HadCRUT4 dataset (Morice et al. 2012) is available from the Met Office observations download page (<https://www.metoffice.gov.uk/hadobs/hadcrut4/data/current/download.html>), the NOAA20CRv3 dataset (Slivinski et al. 2019) is available from the NOAA Physical Sciences Laboratory reanalysis download page (https://psl.noaa.gov/data/gridded/data.20thC_ReanV3.html), the CMIP6 climate model data (Eyring et al. 2016) are available through the Earth System Grid Federation portal (<https://esgf-data.dkrz.de/search/cmip6-dkrz/>), and the anthropogenic surface radiative forcing time series (Meinshausen et al. 2011) are available online (<https://www.pik-potsdam.de/~mmalte/rcps/>).

APPENDIX A

List of CMIP6 Climate Models

The names of the 27 CMIP6 climate models employed in this study are listed in Table A1.

⁵ For any of the spatial subdomains, the frequency-dependent spatial correlation function is obtained by treating all grid boxes outside the respective subdomain as data gaps and then performing the same computation as for the global analysis.

TABLE A1. List of the CMIP6 climate models employed in this study. For CanESM5 the model physics variant (denoted by p1 and p2) is indicated in parentheses.

No.	Model	No.	Model	No.	Model
1	ACCESS-ESM1-5	10	CIESM	19	IPSL-CM6A-LR
2	BCC-CSM2-MR	11	CNRM-CM6-1	20	MIROC6
3	CanESM5 (p1)	12	CNRM-ESM2-1	21	MIROC-ES2L
4	CanESM5 (p2)	13	E3SM-1-0	22	MPI-ESM1-2-HR
5	CAS-ESM2-0	14	FGOALS-f3-L	23	MPI-ESM1-2-LR
6	CESM2-FV2	15	FIO-ESM-2-0	24	MRI-ESM2-0
7	CESM2	16	GISS-E2-1-G	25	NESM3
8	CESM2-WACCM-FV2	17	GISS-E2-1-H	26	NorCPM1
9	CESM2-WACCM	18	INM-CM5-0	27	UKESM1-0-LL

APPENDIX B

Frequency-Dependent Spatial Covariance Function from Gappy Data Fields

The frequency-dependent spatial covariance function $C(\theta; f)$ is obtained by taking the Fourier transform of the spatiotemporal covariance function $C(\theta; \tau)$, where τ denotes the time lag,

$$C(\theta; f) = \mathcal{F}[C(\theta; \tau)], \quad (\text{B1})$$

using the Wiener–Khinchine theorem (e.g., Priestley 1981), stating that the (cross-) covariance function and the (cross-) power spectral density are a Fourier transform pair. In (B1), for any fixed θ , $C(\theta; \tau)$ is a temporal covariance function and, thus, $C(\theta; f)$ is a spectral density. Only after integrating $C(\theta; f)$ over a (possibly narrow) frequency band does it have units of variance. Nonetheless, for simplicity, we refer to $C(\theta; f)$ as the frequency-dependent spatial covariance function hereinafter. By using (B1), our approach to estimating $C(\theta; f)$, from gappy data fields, reduces to estimating $C(\theta; \tau)$ which can be done without any interpolation across data gaps. Note that interpolation must be avoided here as it would potentially distort the spatiotemporal correlation structure.

In the following, in section a, the basic principle for estimating the mean covariance function from gappy data is illustrated for the simplified case of a one-dimensional domain. Subsequently, in section b, it is outlined how this basic principle is applied to the multidimensional case of estimating $C(\theta; \tau)$ from a time series of gappy global temperature fields. Since this approach involves a spherical harmonic transform, the temperature fields have to be remapped to a Gaussian latitude grid, and this negatively biases the local variance. It is demonstrated, however, in section c, that this bias can be sufficiently alleviated by using a higher Gaussian grid resolution.

a. Basic principle for estimating the mean covariance function from gappy data

The basic principle can be illustrated most easily for the case of a one-dimensional, zero-mean random process, X_t , defined on a discrete finite time domain with $t \in \{0, \dots, N-1\}$ and with cyclic boundary conditions (such that for any variable $x_{t+kN} = x_t$ for all $k \in \mathbb{Z}$). In addition, we define the

deterministic signal μ_t , serving as a mask, with $\mu_t = 1$ at those t where X_t is observed and $\mu_t = 0$ where X_t is not observed, that is, at the data gaps. Thus, any observed data series can be represented as a single realization of the masked process $\mu_t X_t$. By defining the covariance function of X_t as $C_X(t, t + \tau) = \langle X_t X_{t+\tau} \rangle$, where $\langle \cdot \rangle$ denotes the expected value operator, and the covariance function of μ_t as $C_\mu(t, t + \tau) = \mu_t \mu_{t+\tau}$, the covariance function of the masked process $\mu_t X_t$ is then given by $C_{\mu X}(t, t + \tau) = \langle \mu_t X_t \mu_{t+\tau} X_{t+\tau} \rangle = \mu_t \mu_{t+\tau} \langle X_t X_{t+\tau} \rangle = C_\mu(t, t + \tau) \times C_X(t, t + \tau)$.

The desired quantity is the (time) mean covariance function of X_t , defined as $\overline{C}_X(\tau) = N^{-1} \sum_{t=0}^{N-1} C_X(t, t + \tau)$. Known from observations, however, are only the mean covariance function of μ_t , given by $\overline{C}_\mu(\tau) = N^{-1} \sum_{t=0}^{N-1} \mu_t \mu_{t+\tau}$, and the mean covariance function of $\mu_t X_t$, given by

$$\overline{C}_{\mu X}(\tau) = N^{-1} \sum_{t=0}^{N-1} C_{\mu X}(t, t + \tau) \quad (\text{B2})$$

$$= N^{-1} \sum_{t=0}^{N-1} \mu_t \mu_{t+\tau} [\overline{C}_X(\tau) + C'_X(t, t + \tau)] \quad (\text{B3})$$

$$= \overline{C}_\mu(\tau) \overline{C}_X(\tau) + N^{-1} \sum_{t=0}^{N-1} \mu_t \mu_{t+\tau} C'_X(t, t + \tau), \quad (\text{B4})$$

where $C'_X(t, t + \tau) = C_X(t, t + \tau) - \overline{C}_X(\tau)$ denotes the nonstationary component of $C_X(t, t + \tau)$, which is equal to zero for all t and τ if the process X_t is stationary. Dividing $\overline{C}_{\mu X}(\tau)$ by $\overline{C}_\mu(\tau)$ yields $\overline{C}_{\mu X}^{(\mu)}(\tau)$, that is, the mean covariance function of $\mu_t X_t$ corrected for μ_t ,

$$\overline{C}_{\mu X}^{(\mu)}(\tau) = \overline{C}_{\mu X}(\tau) / \overline{C}_\mu(\tau) \quad (\text{B5})$$

$$= \overline{C}_X(\tau) + N_\tau^{-1} \sum_{t=0}^{N-1} \mu_t \mu_{t+\tau} C'_X(t, t + \tau), \quad (\text{B6})$$

where $N_\tau = \sum_{t=0}^{N-1} \mu_t \mu_{t+\tau}$. Hence, if X_t is stationary, then $\overline{C}_{\mu X}^{(\mu)}(\tau) = \overline{C}_X(\tau)$ as desired. If, however, X_t is nonstationary, there might be a bias, given by the second term of (B6), which is simply the average over the observed part of $C'_X(t, t + \tau)$. Note, that $\overline{C}_{\mu X}^{(\mu)}(\tau)$ exists only if, for all τ , it is $\overline{C}_\mu(\tau) > 0$ or,

equivalently, $N_\tau > 0$ [because $\bar{C}_\mu(\tau) = N_\tau/N$], that is, if for each time lag there is at least one pair of times, separated by a lag of τ , where X_t is observed.

By defining an estimator of $\bar{C}_{\mu X}(\tau)$, based on a single realization of $\mu_t X_t$, as $\hat{C}_{\mu X}(\tau) = N^{-1} \sum_{t=0}^{N-1} \mu_t X_t \mu_{t+\tau} X_{t+\tau}$, we obtain an estimator of $\bar{C}_{\mu X}^{(\mu)}(\tau)$, denoted by $\hat{C}(\tau)$, according to (B5),

$$\hat{C}(\tau) = \hat{C}_{\mu X}(\tau) / \bar{C}_\mu(\tau) \tag{B7}$$

$$= N_\tau^{-1} \sum_{t=0}^{N-1} \mu_t \mu_{t+\tau} X_t X_{t+\tau}. \tag{B8}$$

Because, as seen from (B8), $\hat{C}(\tau)$ is simply the arithmetic mean over all products $X_t X_{t+\tau}$ available from observations, the estimator itself is unbiased. Nonetheless, as shown above, there might be a bias due to a lack of information, that is, if the nonstationary component of the covariance function is sampled by the observations such that positive and negative contributions to the second term of (B6) do not average out. The scatter, however, of $\hat{C}(\tau)$ increases in the presence of data gaps because of the smaller number $N_\tau (< N)$ of available products involved in the arithmetic mean in (B8).

The above principle is equally applicable to spatial domains, with time and time lag being replaced by spatial position and distance, as well as to domains of higher dimension and with various geometries. In the following, it is applied to the specific case of time-dependent random fields on the sphere.

b. Spatiotemporal covariance function from gappy global temperature fields

Let $X_{i,j,t}$ denote a zero-mean, discrete, spatiotemporal random field on the sphere, where the indices $i \in \{0, \dots, I-1\}$, $j \in \{0, \dots, J-1\}$, and $t \in \{0, \dots, N-1\}$ represent longitude, latitude, and time, respectively, with $I (=2J)$ equidistant longitudes and J Gaussian latitudes, and with cyclic boundary conditions in time (such that for any variable $x_{i,j,t+kN} = x_{i,j,t}$ for all $k \in \mathbb{Z}$). As before we define a deterministic signal $\mu_{i,j,t}$ which is equal to 1 where $X_{i,j,t}$ is observed and equal to 0 at the data gaps. Any observed time series of gappy global temperature fields may then be represented as a single realization of the masked random field $Z_{i,j,t} = \mu_{i,j,t} X_{i,j,t}$.

The desired quantity is the (spatiotemporal) mean spatiotemporal covariance function of $X_{i,j,t}$, denoted by $\bar{C}_X(\theta; \tau)$. Given from observations, however, are only the mean spatiotemporal covariance function of $Z_{i,j,t}$, denoted by $\bar{C}_Z(\theta; \tau)$, and of the mask $\mu_{i,j,t}$, denoted by $\bar{C}_\mu(\theta; \tau)$. According to the basic principle, introduced in the previous section of this appendix, we can define the mean spatiotemporal covariance function of $Z_{i,j,t}$ corrected for $\mu_{i,j,t}$, by analogy with (B5), as

$$\bar{C}_Z^{(\mu)}(\theta; \tau) = \bar{C}_Z(\theta; \tau) / \bar{C}_\mu(\theta; \tau), \tag{B9}$$

which equals $\bar{C}_X(\theta; \tau)$, as desired, if the random field $X_{i,j,t}$ is stationary in space and time, but might be biased if the random field is nonstationary, depending on how the nonstationary

component of the spatiotemporal covariance function is sampled by the observations.

The approach to obtain the mean spatiotemporal covariance functions $\bar{C}_Z(\theta; \tau)$ and $\bar{C}_\mu(\theta; \tau)$ is based on Kunz and Laepple (2021), using spherical harmonic decomposition of the spatial fields. By projecting each realization of $Z_{i,j,t}$ onto the (n, m) th discrete spherical harmonic function $(Y_{nm})_{i,j}$, with total wavenumber n and zonal wavenumber m , we obtain the (n, m) th spherical harmonic component $\tilde{Z}_{n,m,t} = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} w_{i,j} Z_{i,j,t} (Y_{nm})_{i,j}$, which is again a three-dimensional random field, and where the grid box area weights $w_{i,j}$ are normalized according to $\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} w_{i,j} = 1$. The spherical harmonic functions are normalized to have unit power, $\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} w_{i,j} (Y_{nm})_{i,j}^2 = 1$, such that the spatial mean of the local variance is related to the spherical harmonic components by $\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} w_{i,j} \langle Z_{i,j,t}^2 \rangle = \sum_{n=0}^{n_T} \sum_{m=-n}^n \langle \tilde{Z}_{n,m,t}^2 \rangle$, where n_T is the truncation wavenumber according to the finite Gaussian grid resolution. We can then define the temporal covariance function of the (n, m) th spherical harmonic component, $C_{\tilde{Z}_{n,m}}(t, t + \tau) = \langle \tilde{Z}_{n,m,t} \tilde{Z}_{n,m,t+\tau} \rangle$, and its time mean temporal covariance function, $\bar{C}_{\tilde{Z}_{n,m}}(\tau) = N^{-1} \sum_{t=0}^{N-1} C_{\tilde{Z}_{n,m}}(t, t + \tau)$. From this, we finally obtain the mean spatiotemporal covariance function of $Z_{i,j,t}$ as an inverse Legendre integral transform,

$$\bar{C}_Z(\theta; \tau) = \sum_{n=0}^{n_T} P_n(\theta) c_n(\tau), \tag{B10}$$

with the n th Legendre polynomial $P_n(\theta)$, the n th Legendre coefficient $c_n(\tau) = \sum_{m=-n}^n \bar{C}_{\tilde{Z}_{n,m}}(\tau)$, and angular distance θ . The summation over m corresponds to the spatial averaging operator, and it reduces the spatial domain from two dimensions (longitude and latitude) to one dimension (angular distance). Note, that $\tau \in \{0, \dots, N-1\}$ is a discrete variable, whereas $\theta \in [0, \pi]$ is a continuous variable. Similarly, we can obtain the spherical harmonic components of the mask, $\tilde{\mu}_{n,m,t}$, and their time mean temporal covariance function $\bar{C}_{\tilde{\mu}_{n,m}}(\tau) = N^{-1} \sum_{t=0}^{N-1} \tilde{\mu}_{n,m,t} \tilde{\mu}_{n,m,t+\tau}$, from which, by analogy with (B10),

$$\bar{C}_\mu(\theta; \tau) = \sum_{n=0}^{n_T} P_n(\theta) d_n(\tau), \tag{B11}$$

with coefficients $d_n(\tau) = \sum_{m=-n}^n \bar{C}_{\tilde{\mu}_{n,m}}(\tau)$.

By defining an estimator of $\bar{C}_{\tilde{Z}_{n,m}}(\tau)$, based on a single realization of $Z_{i,j,t}$, as $\hat{C}_{\tilde{Z}_{n,m}}(\tau) = N^{-1} \sum_{t=0}^{N-1} \tilde{Z}_{n,m,t} \tilde{Z}_{n,m,t+\tau}$, we obtain an estimator of $\bar{C}_Z^{(\mu)}(\theta; \tau)$, denoted by $\hat{C}(\theta; \tau)$, according to (B9),

$$\hat{C}(\theta; \tau) = \hat{C}_Z(\theta; \tau) / \bar{C}_\mu(\theta; \tau), \tag{B12}$$

where, by analogy with (B10), $\hat{C}_Z(\theta; \tau) = \sum_{n=0}^{n_T} P_n(\theta) \hat{c}_n(\tau)$, with $\hat{c}_n(\tau) = \sum_{m=-n}^n \hat{C}_{\tilde{Z}_{n,m}}(\tau)$. By analogy with (B7), the estimator (B12) itself is unbiased, but there might be a bias due to a lack of information caused by the data gaps, and the scatter of the estimator increases the more data gaps there are. When applied to the HadCRUT4 temperature fields for the

period 1850–2014, however, this additional scatter is found to be small. It is also found that $\overline{C}_\mu(\theta; \tau) > 0$ for all combinations of θ and τ , which ensures the existence of the estimator over the entire domain. Finally, according to (B1), taking the discrete Fourier transform of $\hat{C}(\theta; \tau)$, at any fixed value of θ , yields an estimator of the frequency-dependent spatial covariance function,

$$\hat{C}(\theta; f) = \mathcal{F}[\hat{C}(\theta; \tau)]. \quad (\text{B13})$$

Note, that, since $P_n(0) = 1$ for all n , it follows from the normalization condition of the spherical harmonic functions that (B13) at $\theta = 0$ is an estimator of the mean local power spectral density,

$$\hat{S}_{\text{loc}}(f) = \hat{C}(0; f). \quad (\text{B14})$$

In the absence of data gaps, the spectral density estimator (B14) is χ^2 -distributed as usual. When estimated from gappy data, however, the additional scatter of the estimator (B12) translates into an additional scatter of (B14), which may then become negative at individual frequencies. Since, however, the additional scatter is small when applied to the HadCRUT4 temperature fields, spectral smoothing of (B14) over only a few discrete frequencies suffices to remove the negative spectral density estimates across all frequencies.

c. Effect of the remapping from equidistant to Gaussian latitudes

The above approach to estimating $C(\theta; f)$, according to (B13), requires data fields given on a Gaussian grid because of the involved spherical harmonic transform, whereas the HadCRUT4 grid is equidistant in latitude. Remapping the HadCRUT4 temperature fields to a Gaussian grid, however, biases the spatial covariance structure. By using a second-order conservative remapping scheme, the global mean of the fields is unaffected and the variance bias is largest at the local scale. To quantify this bias, we need to compare results obtained from (B14) to an alternative estimator of $S_{\text{loc}}(f)$ that can be computed from arbitrary latitude grids.

1) ESTIMATION OF THE MEAN LOCAL VARIANCE FROM ARBITRARY LATITUDE GRIDS

Assuming the same spatiotemporal random field on the sphere, $X_{i,j,t}$, and the same mask, $\mu_{i,j,t}$, as in the previous section of this appendix, we can define the (spatiotemporal) mean temporal covariance function of the masked random field $Z_{i,j,t} = \mu_{i,j,t}X_{i,j,t}$, as $\overline{C}_Z(\tau) = N^{-1} \sum_{t=0}^{N-1} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} w_{ij} \langle Z_{i,j,t} Z_{i,j,t+\tau} \rangle$, and of the mask, as $\overline{C}_\mu(\tau) = N^{-1} \sum_{t=0}^{N-1} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} w_{ij} \mu_{i,j,t} \mu_{i,j,t+\tau}$. From this, we obtain the mean temporal covariance function of $Z_{i,j,t}$ corrected for $\mu_{i,j,t}$, given by $\overline{C}_Z^{(\mu)}(\tau) = \overline{C}_Z(\tau) / \overline{C}_\mu(\tau)$, applying the same basic principle as before. By defining an estimator of $\overline{C}_Z(\tau)$, based on a single realization of $Z_{i,j,t}$, as $\hat{C}_Z(\tau) = N^{-1} \sum_{t=0}^{N-1} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} w_{ij} Z_{i,j,t} Z_{i,j,t+\tau}$, we obtain an estimator of $\overline{C}_Z^{(\mu)}(\tau)$,

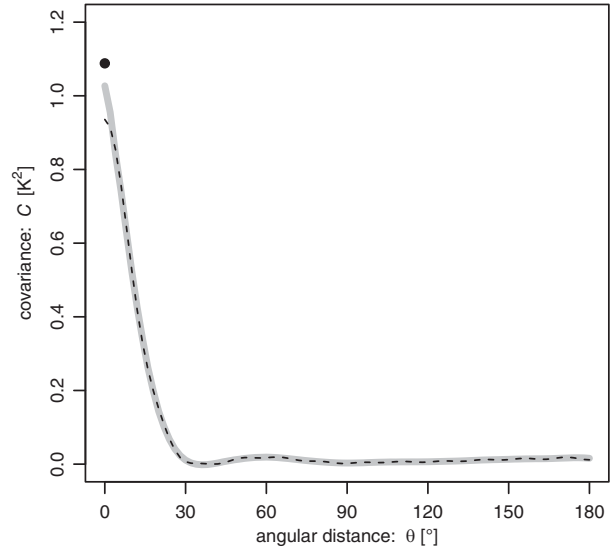


FIG. B1. Estimate of the mean spatial covariance function $\hat{C}(\theta; \tau = 0)$ obtained from the HadCRUT4 temperature fields remapped to a T23 (dashed black) and a T85 (solid gray) Gaussian grid. The black dot represents the unbiased estimate of the mean local variance $\hat{C}(\tau = 0)$ obtained from the original HadCRUT4 grid with equidistant latitudes.

$$\hat{C}(\tau) = \hat{C}_Z(\tau) / \overline{C}_\mu(\tau). \quad (\text{B15})$$

By analogy with the estimators in the previous sections of this appendix, the estimator (B15) itself is unbiased, but there might be a bias due to a lack of information caused by the data gaps, and the scatter of the estimator increases the more data gaps there are. When applied to the HadCRUT4 temperature fields for the period 1850–2014, it is found that $\overline{C}_\mu(\tau) > 0$ for all τ , which ensures the existence of the estimator. Taking the discrete Fourier transform of $\hat{C}(\tau)$ yields an estimator of the mean local power spectral density,

$$\hat{S}_{\text{loc}}(f) = \mathcal{F}[\hat{C}(\tau)]. \quad (\text{B16})$$

When applied to the same data fields on the same Gaussian grid, (B16) is identical to (B14). Since (B16), however, is not based on spherical harmonic transforms, it can be computed from fields on arbitrary latitude grids. In particular, it can be applied to the equidistant latitude grid of the HadCRUT4 temperature fields.

2) QUANTIFICATION OF THE MEAN LOCAL VARIANCE BIAS CAUSED BY THE LATITUDE REMAPPING

To quantify the mean local variance bias, we remap the nonlinearly detrended HadCRUT4 temperature fields from the original 5° longitude \times 5° latitude grid with 72 longitudes and 36 latitudes to (i) a T23 Gaussian grid with the same number of longitudes and latitudes (corresponding to a truncation wavenumber of $n_T = 23$), and (ii) a T85 Gaussian grid with 256 longitudes and 128 latitudes ($n_T = 85$). From each of the three grids we estimate the mean local

power spectral density according to (B16), and it is found that the relative bias in $\hat{S}_{\text{loc}}(f)$ is virtually independent of frequency (not shown). It is, therefore, sufficient to quantify the relative bias for the total mean local variance, given by the mean temporal covariance function at lag zero, $\hat{C}(\tau = 0)$, according to (B15). For the Gaussian grids this is identical to $\hat{C}(\theta = 0; \tau = 0)$, according to (B12). For the T23 Gaussian grid the relative bias in $\hat{C}(\tau = 0)$ amounts to -14% and for the T85 Gaussian grid to -6% , relative to the unbiased estimate obtained from the original HadCRUT4 grid. Figure B1 shows $\hat{C}(\theta; \tau = 0)$ for the Gaussian grids together with $\hat{C}(\tau = 0)$ for the original HadCRUT4 grid. This confirms that the bias is confined to the local scale. As the T85 relative bias seems to be sufficiently small in the context of this study, all results presented in the main text are based on data fields remapped to a T85 Gaussian grid.

APPENDIX C

Estimation Bias and Uncertainty

a. Bias

From the CMIP6 climate model ensemble, the expected estimation bias of the frequency-dependent correlation function, $R(\theta; f)$, can be obtained from the covariance function $C(\theta; f)$ as follows. Two estimators of $R(\theta; f)$ are defined: a biased estimator, defined as the ensemble mean of the ratio $C(\theta; f)/C(0; f)$, and a (virtually) unbiased estimator, defined as the ratio of the ensemble mean of $C(\theta; f)$ to the ensemble mean of $C(0; f)$. The latter one is (virtually) unbiased because after averaging over the 81 ensemble members the scatter of $C(0; f)$ (which causes the bias) is very small. The difference between the two estimators yields an estimate of the expected bias.

However, the additional scatter to $C(0; f)$ (of a single member), caused by the data gaps, brings its estimate close to zero at a few frequencies, leading to very large positive outliers in the estimate of the ratio $C(\theta; f)/C(0; f)$. Therefore, for the biased estimator of $R(\theta; f)$, we use the ensemble median instead of the mean. Since the median is less biased than the mean, the bias is somewhat underestimated, but its much smaller sensitivity to the outliers makes it the more practicable estimator.

From the two estimators of $R(\theta; f)$ we obtain two estimates of θ_e and, thus, of $R_{\text{fit}}(\theta)$. Hence, two estimators of $D(f)$ and of $D_{\text{fit}}(f)$ are obtained, which allows to compute the expected biases of the two ESDOF measures. Finally, the bias correction of the HadCRUT4 ESDOF estimates is performed in a relative sense because the bias scales proportionally with the absolute ESDOF value (see Kunz and Laepple 2021, their appendix B), that is, the biased HadCRUT4 ESDOF estimate is divided by the ratio of the biased to the unbiased CMIP6 ESDOF estimate.

b. Uncertainty

The estimation uncertainty of the ESDOF measures is also obtained from the CMIP6 climate model ensemble. The ESDOF estimates (at a given frequency) of all members are

first log-transformed because the distributions of the log-transformed ESDOF estimates are found to be largely symmetric (little skew), such that the scatter of the log-transformed estimates can be simply characterized by their variance σ^2 . Since our 81 member ensemble consists of $n_{\text{mod}} = 27$ models with $n_{\text{mem}} = 3$ members each, and the various models have different climates and space–time statistics, the total ensemble variance $\sigma^2 = \sigma_{\text{est}}^2 + \sigma_{\text{cli}}^2$ includes not only the variance due to the estimation uncertainty σ_{est}^2 but also the variance due to the intermodel spread σ_{cli}^2 . The variance σ_{est}^2 can be obtained by computing an unbiased estimate of the variance (dividing by $n_{\text{mem}} - 1$) across the $n_{\text{mem}} = 3$ members, separately for each model, and then averaging over these n_{mod} variance estimates. The limits of the uncertainty interval for the HadCRUT4 ESDOF estimates are then defined as the bias-corrected (log-transformed) HadCRUT4 ESDOF estimate plus or minus one standard deviation σ_{est} . Last, an inverse log-transform yields the asymmetric uncertainty intervals around the bias-corrected HadCRUT4 ESDOF estimates (as specified in Table 1).

APPENDIX D

Impact of the Spatial Resolution of the Gridded Surface Temperature Dataset

The HadCRUT4 5° longitude \times 5° latitude grid has 72 longitudes and 36 latitudes. Since our method to obtain the frequency-dependent spatial correlation function is based on a spherical harmonic decomposition, and a Gaussian grid with that number of longitudes and latitudes corresponds to a T23 spectral resolution, with truncation wavenumber $n_T = 23$, and the total number of spherical harmonic components is equal to $(n_T + 1)^2$, the HadCRUT4 grid is sufficient to represent a discrete global white noise with $24^2 = 576$ ESDOFs. By comparison, the largest ESDOF estimates obtained in our analysis lie between 100 and 200 (Figs. 1a,c), which serves as a first indication that the spatial resolution might be sufficient and the ESDOF estimates may have converged toward their true value.

To demonstrate this convergence explicitly, we recomputed both ESDOF measures, D and D_{fit} , for each of the three frequency bands, but with varying spectral resolution, that is, by varying the truncation wavenumber n_T . Specifically, this corresponds to varying the upper limit of the sum in (B10) and (B11). The result is shown in Fig. D1, and it turns out that both ESDOF measures in all frequency bands have largely converged already at $n_T < 23$.

The fact that the results shown in Fig. D1 extend up to $n_T = 85$ is due to the necessity to extrapolate the HadCRUT4 fields onto a T85 Gaussian grid before the spherical harmonic decomposition so as to minimize the local variance bias, as explained in appendix B, section c. All results shown here are based on the T85 Gaussian grid.

Recall that the HadCRUT4 dataset represents monthly temperature averages. When using, for example, daily temperature time series instead, resolving also smaller-scale synoptic variability, the 5° longitude \times 5° latitude grid resolution

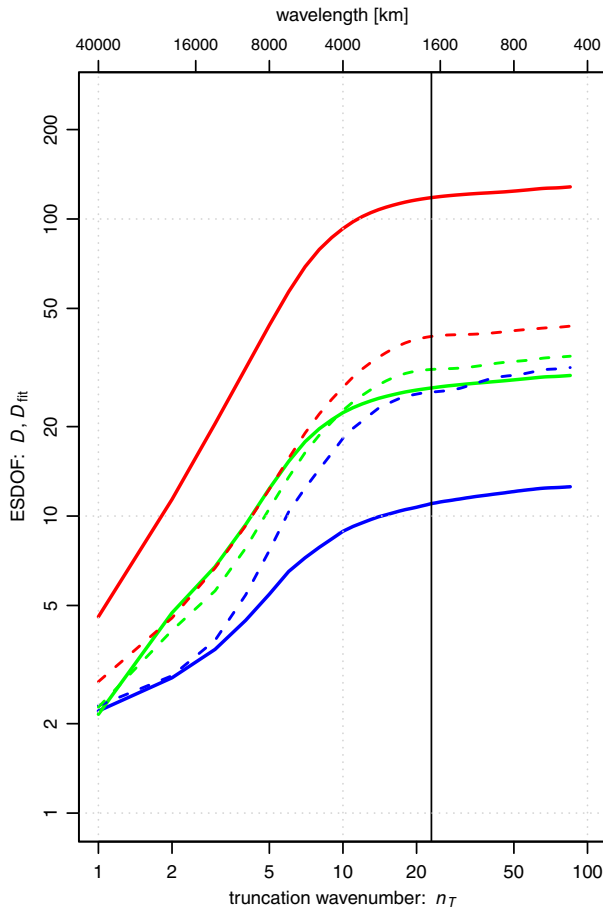


FIG. D1. Frequency-band values of the ESDOF measures D (solid lines) and D_{fit} (dashed lines), obtained from the HadCRUT4 temperature fields, as a function of the truncation wavenumber n_T , for the multidecadal (blue), interannual (green), and subannual (red) frequency band. The vertical black line is at $n_T = 23$.

might actually be insufficient and ESDOF estimates may not yet have converged at $n_T = 23$.

REFERENCES

- Bretherton, C. S., M. Widmann, V. P. Dymnikov, J. M. Wallace, and I. Bladé, 1999: The effective number of spatial degrees of freedom of a time-varying field. *J. Climate*, **12**, 1990–2009, [https://doi.org/10.1175/1520-0442\(1999\)012<1990:TENOSD>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2).
- Dolman, A. M., T. Kunz, J. Groeneveld, and T. Laepple, 2021: A spectral approach to estimating the timescale-dependent uncertainty of paleoclimate records—Part 2: Application and interpretation. *Climate Past*, **17**, 825–841, <https://doi.org/10.5194/cp-17-825-2021>.
- Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, **9**, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>.
- Jones, P. D., T. J. Osborn, and K. R. Briffa, 1997: Estimating sampling errors in large-scale temperature averages. *J. Climate*, **10**, 2548–2568, [https://doi.org/10.1175/1520-0442\(1997\)010<2548:ESEILS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<2548:ESEILS>2.0.CO;2).
- Kunz, T., and T. Laepple, 2021: Frequency-dependent estimation of effective spatial degrees of freedom. *J. Climate*, **34**, 7373–7388, <https://doi.org/10.1175/JCLI-D-20-0228.1>.
- , A. M. Dolman, and T. Laepple, 2020: A spectral approach to estimating the timescale-dependent uncertainty of paleoclimate records—Part 1: Theoretical concept. *Climate Past*, **16**, 1469–1492, <https://doi.org/10.5194/cp-16-1469-2020>.
- Livezey, R. E., and W. Y. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.*, **111**, 46–59, [https://doi.org/10.1175/1520-0493\(1983\)111<0046:SFSASD>2.0.CO;2](https://doi.org/10.1175/1520-0493(1983)111<0046:SFSASD>2.0.CO;2).
- Meinshausen, M., and Coauthors, 2011: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213, <https://doi.org/10.1007/s10584-011-0156-z>.
- Morice, C. P., J. J. Kennedy, N. A. Rayner, and P. D. Jones, 2012: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *J. Geophys. Res.*, **117**, D08101, <https://doi.org/10.1029/2011JD017187>.
- Myhre, G., and Coauthors, 2013: Anthropogenic and natural radiative forcing. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 659–740, <https://doi.org/10.1017/CBO9781107415324.018>.
- North, G. R., J. Wang, and M. G. Genton, 2011: Correlation models for temperature fields. *J. Climate*, **24**, 5850–5862, <https://doi.org/10.1175/2011JCLI4199.1>.
- PAGES2k Consortium, 2017: A global multiproxy database for temperature reconstructions of the Common Era. *Sci. Data*, **4**, 170088, <https://doi.org/10.1038/sdata.2017.88>.
- Priestley, M. B., 1981: *Spectral Analysis and Time Series*. Academic Press, 890 pp.
- Rypdal, K., M. Rypdal, and H.-B. Fredriksen, 2015: Spatiotemporal long-range persistence in Earth's temperature field: Analysis of stochastic–diffusive energy balance models. *J. Climate*, **28**, 8379–8395, <https://doi.org/10.1175/JCLI-D-15-0183.1>.
- Slivinski, L. C., and Coauthors, 2019: Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis System. *Quart. J. Roy. Meteor. Soc.*, **145**, 2876–2908, <https://doi.org/10.1002/qj.3598>.
- Smith, T. M., R. W. Reynolds, and C. F. Ropelewski, 1994: Optimal averaging of seasonal sea surface temperatures and associated confidence intervals (1860–1989). *J. Climate*, **7**, 949–964, [https://doi.org/10.1175/1520-0442\(1994\)007<0949:OAOSSS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<0949:OAOSSS>2.0.CO;2).
- Wang, X., and S. S. Shen, 1999: Estimation of spatial degrees of freedom of a climate field. *J. Climate*, **12**, 1280–1291, [https://doi.org/10.1175/1520-0442\(1999\)012<1280:EOSDOF>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<1280:EOSDOF>2.0.CO;2).