Taylor & Francis
Taylor & Francis Group

ⓐ OPEN ACCESS | Check for updates |

Efficient Stochastic Generators with Spherical Harmonic Transformation for High-Resolution Global Climate Simulations from CESM2-LENS2

Yan Song[a], Zubair Khalid[b], and Marc G. Genton[a]

[a]Statistics Program, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia; [b]School of Science and Engineering, Lahore University of Management Sciences, Lahore, Pakistan

ABSTRACT

Earth system models (ESMs) are fundamental for understanding Earth's complex climate system. However, the computational demands and storage requirements of ESM simulations limit their utility. For the newly published CESM2-LENS2 data, which suffer from this issue, we propose a novel stochastic generator (SG) as a practical complement to the CESM2, capable of rapidly producing emulations closely mirroring training simulations. Our SG leverages the spherical harmonic transformation (SHT) to shift from spatial to spectral domains, enabling efficient low-rank approximations that significantly reduce computational and storage costs. By accounting for axial symmetry and retaining distinct ranks for land and ocean regions, our SG captures intricate nonstationary spatial dependencies. Additionally, a modified Tukey g-and-h (TGH) transformation accommodates non-Gaussianity in high-temporal-resolution data. We apply the proposed SG to generate emulations for surface temperature simulations from the CESM2-LENS2 data across various scales, marking the first attempt of reproducing daily data. These emulations are then meticulously validated against training simulations. This work offers a promising complementary pathway for efficient climate modeling and analysis while overcoming computational and storage limitations. Supplementary materials for this article are available online, including a standardized description of the materials available for reproducing the work.

1. Introduction

Earth system models (ESMs) are complex mathematical equations that describe and simulate the transformation and interaction of energy and materials within the Earth's climate system. Their development is grounded in the fundamental laws of physics, fluid dynamics, and chemistry, necessitating the identification and quantification of physical processes represented through mathematical equations (National Oceanic and Atmospheric Administration (NOAA) 2022). ESMs are indispensable tools in both scientific research and government policymaking, enabling a deeper understanding and more accurate predictions of Earth's systems. For example, the Intergovernmental Panel on Climate Change (IPCC) used datasets from the Coupled Model Intercomparison Project's fifth (CMIP5) and sixth (CMIP6) phases (Taylor, Stouffer, and Meehl 2012; Eyring et al. 2016) to compile their fifth and sixth synthesis assessment reports on climate change.

Despite their significance, ESMs have limitations as approximations of the Earth's complex system. They are sensitive to changes in model inputs, such as physics parameters and emission scenarios, which need being assessed. Additionally, distinguishing between model errors and the effects of internal variability, which arises from atmospheric, oceanic, land, and cryospheric processes and their interactions, is challenging with

a limited number of simulations (Kay et al. 2015). This underscores the need for multiple climate simulations to comprehensively characterize Earth's system. The Community Earth System Model (CESM, Kay et al. 2015) is one such model designed to study climate change while accounting for internal climate variability.

Running ESMs is a time-consuming process, even with the support of powerful supercomputers. Scientists partition the planet into thousands of three-dimensional grids. Within each grid cell, they specify variables and conditions, employ computers to solve equations, propagate results to neighboring cells, and then iterate the above procedure with the updated variables and conditions. The whole process is repeated through time steps of different scales, which is the temporal resolution. The spatial resolution of the model is determined by the size of the grid cell, with smaller sizes yielding higher resolution but requiring more time. Generating multiple climate simulations demands weeks and even months of computational resources, exclusively available to several research institutes globally (Huang et al. 2023). Furthermore, while modern computational power allows for the generation of climate simulations, storing them is constrained by technological limitations, resource availability, and cost considerations. For instance, the storage requirement for CMIP5 data exceeds 2.5 Petabytes, incurring significant expenses for

---

institutions like the National Center for Atmospheric Research (NCAR), a prominent U.S. research center for Earth Systems science. Consequently, budget constraints may necessitate the abandonment of partial data or higher resolutions (Guinness and Hammerling 2018).

Emulators play a crucial role in providing rapid approximations for computationally demanding model outputs and serve as effective surrogates for the original model. Specifically, emulators are statistical models that undergo proper training using a set of existing simulations and are then employed to approximate simulations for unexplored inputs. Over the past few decades, emulators have found widespread application, including assessing the sensitivity of various physical parameters (Oakley and O'Hagan 2002, 2004) and performing model calibration (Kennedy and O'Hagan 2001; Chang et al. 2014). Within the climate research community, emulators have been extensively developed to evaluate the influence of physics parameters (Sacks, Schiller, and Welch 1989), emission scenarios (Castruccio and Stein 2013; Castruccio et al. 2014), and the internal variability (Jeong et al. 2019; Hu and Castruccio 2021) on climate model outputs. These emulators enable efficient exploration of model behavior and facilitate a deeper understanding of complex climate systems.

Stochastic generators (SGs, Jeong et al. 2018), as a type of emulator for quantifying internal variability uncertainty, align closely with our requirements. SGs are spatio-temporal models adeptly tailored with a limited number of climate simulations, capable of generating numerous annual and monthly simulations across extensive and high-resolution domains at any moment (Jeong et al. 2018, 2019; Castruccio et al. 2019; Tagle et al. 2020), yet daily scale remains elusive to this date. Their implementation is affordable and even efficient, albeit at the expense of detailed physical mechanisms. SGs require the storage of only model parameters, enabling them to generate simulations that capture the general behavior of the original climate simulations. It is essential to distinguish SGs from compression methods (Baker et al. 2017; Underwood et al. 2022), which effectively reduce data volume using compression algorithms. While compression methods store and recover compressed data, they are designed for reconstructing individual data rather than generating new simulations. Moreover, the storage requirements for compression methods tend to increase with additional simulations.

SGs typically use Gaussian processes (GPs) and autoregressive models to capture the intricate spatio-temporal dependencies among climate simulations driven by underlying physical mechanisms. While SGs offer several advantages, they face notable challenges when emulating ESMs. First, global climate simulations, which need the technique of SGs, involve extensive high-resolution data. Inference with GPs on such data can be computationally prohibitive, often requiring time on the order of data size cubed. Addressing this computational challenge can draw inspiration from existing techniques such as low-rank approximations (Cressie and Johannesson 2008; Banerjee et al. 2008), composite likelihoods (Vecchia 1988; Guinness 2021), stochastic partial differential equations (Rue, Martino, and Chopin 2009; Bolin, Simas, and Xiong 2023), covariance tapering (Kaufman, Schervish, and Nychka 2008) and their

combinations (Datta et al. 2016). Second, global climate simulations frequently exhibit nonstationary spatial dependence. The covariance structure can vary significantly with latitude while may remain relatively stationary along the longitude, which is called axial symmetry (Jones 1963) and has been noticed and discussed by Stein (2007), Stein (2008) and Jun and Stein (2008). Additionally, the smoothness and variance of data can change between land and ocean at the same latitude (Castruccio and Guinness 2017; Jeong et al. 2018). A well-defined covariance function based on geodesic distance rather than chordal distance is needed to model the nonstationary spatial dependence structure. The use of chordal distance may cause physically unrealistic distortions (Jeong, Jun, and Genton 2017). Lastly, global climate simulations with high temporal resolutions, such as monthly and daily scales, may deviate from Gaussian assumptions when studying temporal evolution. To address this, the Tukey g-and-h (TGH) transformation has been employed to Gaussianize the data, striking a balance between model flexibility and simplicity (Yan and Genton 2019; Jeong et al. 2019).

For global surface temperature simulations from the CESM2-LENS2 data, which are newly published in the climate community and face the above-mentioned challenges, we propose a novel SG. Our SG serves as a pragmatic complement to CESM2, offering the capability to rapidly generate an unlimited number of emulations closely resembling simulations. Central to our method is the utilization of the spherical harmonic transformation (SHT)—the well-known spherical counterpart of the Fourier transformation, allowing us to efficiently transition global simulations from the spatial to the spectral domain. Moreover, it unlocks a practical low-rank approximation strategy, using spherical harmonics as available multi-resolution basis functions that are naturally suitable for global data. The modeling in the spectral domain ushers in significant reductions in both computational and storage costs. By judiciously retaining distinct ranks for land and ocean regions and applying a general assumption of axial symmetry, our SG captures the intricate nonstationary spatial dependence among different latitudes and between land and ocean regions. Furthermore, the employment of a modified TGH transformation equips our SG to effectively model the temporal dependence of simulations, accommodating the non-Gaussianity inherent in high-temporal-resolution data. In our case study, we apply these principles to construct SGs and generate emulations for the surface temperature simulations, spanning annual, monthly, and even daily scales. We meticulously illustrate and compare the generated emulations against the training simulations, both through visual inspection and quantitative metrics. Notably, our work marks a significant advancement by being the first to generate, display, and evaluate emulations at a daily scale.

The remainder of the article is structured as follows. In Section 2, we provide an overview of temperature simulations across annual, monthly, and daily scales. Section 3 delves into the intricate procedures involved in constructing an SG using the SHT and generating emulations. The generated emulations, along with comparisons to the training simulations, are showcased in Section 4. Finally, Section 5 encapsulates our conclusions and outlines avenues for future research.

2. CESM2-LENS2 Data

The Community Earth System Model version 2 Large Ensembles (CESM2-LENS2), as detailed in Rodgers et al. (2021), is a recently released extensive ensemble of climate change projections. It originates from the CESM2 and serves as a valuable resource for studying the sensitivity of internal climate fluctuations to greenhouse warming. This ensemble comprises 100 individual members, operating at a spatial resolution of $0.9375° \times 1.25°$ (latitude×longitude), spanning the temporal domain from 1850 to 2100, and generated through a combination of diverse oceanic and atmospheric initial states. CESM2-LENS2 draws upon CMIP6 historical forcing, spanning the period from 1850 to 2014, and integrates the Shared Socioeconomic Pathways 370 (SSP370) future radiative forcing (RF, an influential factor in climate system describing the difference between incoming and outgoing sun radiation) from 2015 to 2100. The SSP scenarios, introduced in response to the sixth IPCC report, encompass a range of socioeconomic developments and atmospheric greenhouse gas concentration pathways (O'Neill et al. 2014; Riahi et al. 2017). Notably, SSP370 represents a medium-to-high emission scenario, projecting an additional RF of 7 Wm^{-2} by the year 2100 under the social development pathway of regional rivalry.
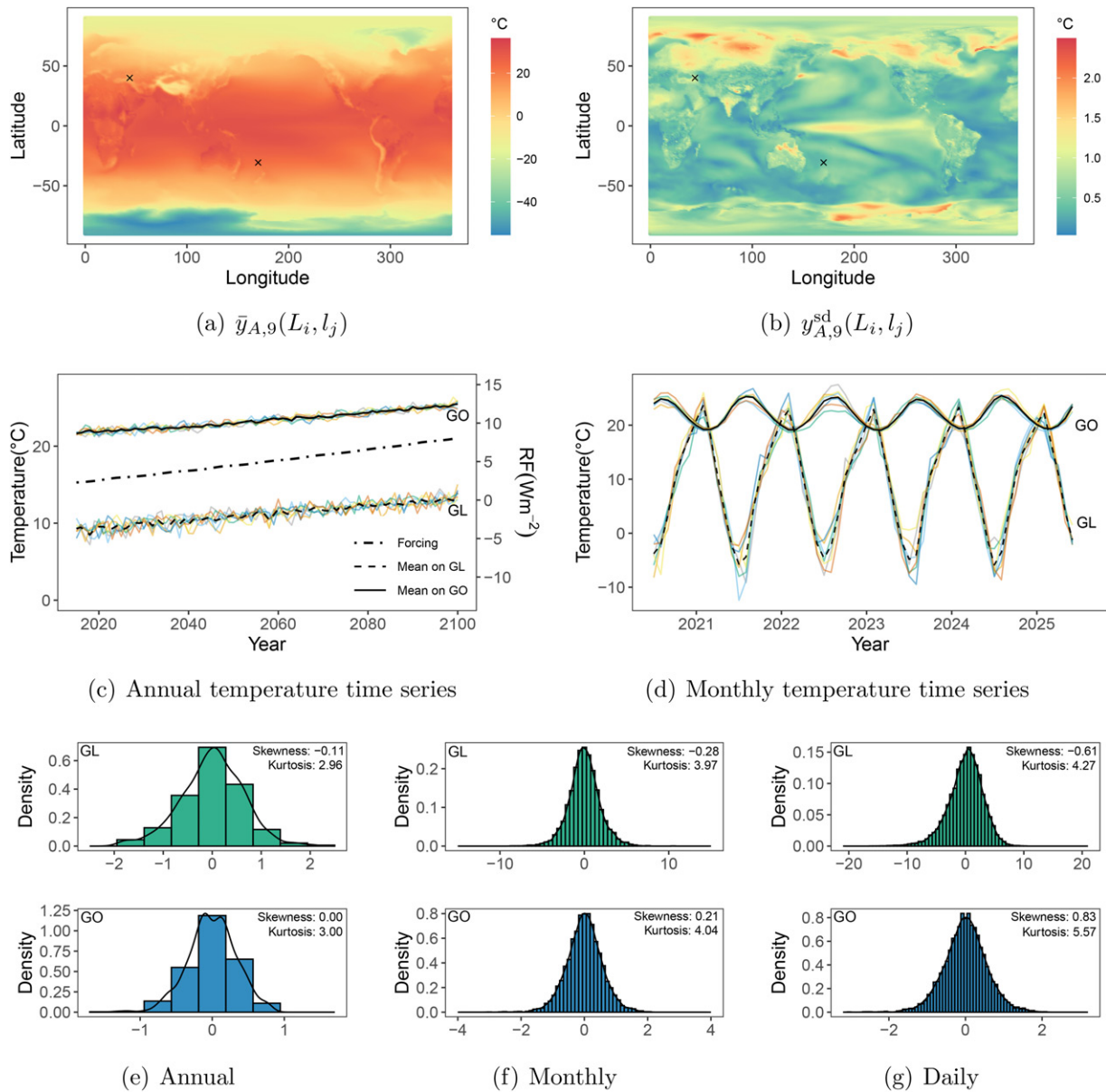
The choice of CESM2-LENS2 in our study is motivated by several factors. First, it provides comprehensive and credible global climate simulations at high resolutions, which often pose storage challenges. Second, its ample ensemble size allows for effective training and evaluation of our SG. Third, despite generating considerable interest within the climate community (Kay et al. 2022; Muñoz et al. 2023; Dong, Peings, and Magnusdottir 2023), CESM2-LENS2 has yet to gain widespread attention in the field of statistics. This article marks the first introduction of an SG, extending to daily scales, to complement CESM2-LENS2. Both CESM2-LENS2 and SSP370 RF are readily accessible from the NCAR and Zenodo websites, respectively. Our analysis focuses on global surface temperatures for the period 2015–2100, encompassing all $192 \times 288 = 55{,}296$ grid points in space. Existing literature (Jeong et al. 2018, 2019; Huang et al. 2023) has demonstrated that fewer than ten ensembles are sufficient to construct an efficient SG for a climate variable. Therefore, we select training simulations from a subset of 20 members micro-initialized from the year 1251, with each member created through random perturbations of the atmospheric potential temperature field. These members are assumed to be independent from each other. We aggregate the ensembles at annual, monthly, and daily scales, despite their original availability at a three-hourly resolution. Consequently, each annual, monthly, and daily aggregated temperature simulation comprises approximately 4.75 million ($192 \times 288 \times 86$), 57.07 million, and 1.73 billion data points, respectively.

Let $y_t^{(r)}(L_i, l_j)$ denote the temperature in Celsius at latitude L_i, longitude l_j, time point t after year 2014, and ensemble r, where $i = 1, \ldots, I$ ($I = 192$), $j = 1, \ldots, J$ ($J = 288$), $t = 1, \ldots, T$, and $r = 1, \ldots, R$. The value of T varies depending on the temporal resolution chosen, with possible values of 86, 1032, or 31,390. Similarly, let x_t represent the RF at time t for all ensembles, as they are generated with the same external forcing. We use members 11 to 17 (out of 20) to illustrate certain data

characteristics. The justification for the number of ensembles will be further discussed in Figure 3(a), using an index for goodness-of-fit. Let $y_{A,t}^{(r)}(L_i, l_j)$ be the annually aggregated temperature at grid point (L_i, l_j), year $t+2014$ and ensemble r. Consequently, the ensemble mean and standard deviation at (L_i, l_j) and $t+2014$ are denoted as $\bar{y}_{A,t}(L_i, l_j) = R^{-1} \sum_{r=1}^{R} y_{A,t}^{(r)}(L_i, l_j)$ and $y_{A,t}^{\mathrm{sd}}(L_i, l_j) = \{R^{-1} \sum_{r=1}^{R} (y_{A,t}^{(r)}(L_i, l_j) - \bar{y}_{A,t}(L_i, l_j))^2\}^{1/2}$, respectively. Figure 1(a) and (b) depict maps of $\bar{y}_{A,9}(L_i, l_j)$ and $y_{A,9}^{\mathrm{sd}}(L_i, l_j)$. The maps align with the intuitive understanding that temperature generally decreases with increasing latitude. The surface temperatures in the Himalayas and Andes ranges are lower than in surrounding regions. Numerical instabilities occur near the two pole regions. Additionally, the surface temperature at the tropical Pacific Ocean exhibits larger variation across ensembles due to the climate phenomenon called El Niño–Southern Oscillation.

Let $\{y_{A,t}^{(r)}(L_i, l_j)\}_{t=1}^{T}$ represent the annual temperature trajectory of ensemble r at (L_i, l_j). Figure 1(c) illustrates these time series for both GL (land) and GO (ocean) locations, that is, $\{y_{A,t}^{(r)}(\mathrm{GL})\}_{t=1}^{T}$ and $\{y_{A,t}^{(r)}(\mathrm{GO})\}_{t=1}^{T}$, alongside their respective ensemble means $\{\bar{y}_{A,t}(\mathrm{GL})\}_{t=1}^{T}$ and $\{\bar{y}_{A,t}(\mathrm{GO})\}_{t=1}^{T}$. Take $\{y_{A,t}^{(r)}(\mathrm{GL})\}_{t=1}^{T}$ as an example. All ensembles exhibit a shared increasing trend, influenced by the RF also depicted in Figure 1(c). Moreover, each ensemble exhibits its unique shape, signifying the presence of uncertainty among ensembles. These motivate us to develop an SG incorporating both deterministic and stochastic factors. Similarly, Figure 1(d) provides monthly temperature time series of years 2021–2025 at GL and GO, along with their respective ensemble means. They experience alternating peaks and valleys because they are situated in the northern and southern hemispheres, respectively. In comparison to annual data, Figure 1(d) exhibits a clear seasonality, which is a common feature in monthly and daily data. The seasonality of surface temperature varies across grid points and remains consistent among all ensembles at a fixed grid point. Notably, from Figure 1(c) and (d), both the uncertainty and the variation are more pronounced at GL.

Figure 1(e)–(g) demonstrate histograms, skewness and kurtosis of the detrended annual, monthly and daily temperature simulations at GL and GO, respectively. For example, the top panel of Figure 1(e) displays the histogram of $\{y_{A,t}^{(r)}(\mathrm{GL}) - R^{-1} \sum_{r=1}^{R} y_{A,t}^{(r)}(\mathrm{GL})\}_{r=1,\ldots,R;t=1,\ldots,T}$ and its skewness and kurtosis. The close-to-zero skewness and the close-to-three kurtosis enable a Gaussian assumption when we model R time series $\{y_{A,t}^{(r)}(\mathrm{GL}) - R^{-1} \sum_{r=1}^{R} y_{A,t}^{(r)}(\mathrm{GL})\}_{t=1}^{T}$, with an auto-regression. In contrast, the histograms in Figure 1(f) and (g) tend to be skewed and heavy-tailed. Figure S1 presents the skewness and kurtosis for time series across all global grid points. The degree of skewness and the presence of heavy tails become more pronounced with increasing temporal resolution. The non-Gaussianity is particularly severe in the band region below latitude $-60°$, termed as Band, and the North Pole region, where the surface temperature simulations exhibit numerical instabilities. The reason will be discussed in Section 4.2. These observations underscore the necessity for additional transformations and parameters to Gaussianize the monthly and daily simulations, as the first two moments

(a) $\bar{y}_{A,9}(L_i, l_j)$

(b) $y_{A,9}^{\text{sd}}(L_i, l_j)$

(c) Annual temperature time series

(d) Monthly temperature time series

(e) Annual

(f) Monthly

(g) Daily

Figure 1. Illustration of simulations with different scales. (a) and (b) are maps of the ensemble mean and standard deviation of the annually aggregated temperature in the year 2023. Two black "×" marks indicate grid points located at coordinates (40.05, 43.75) and (−30.63, 170.00), designating points on land (GL) and ocean (GO), respectively. (c) shows annual temperature time series at GL and GO, where each ensemble is represented by a distinct color. (d) shows the monthly temperature time series at GL and GO, where the black curves are ensemble means. (e)–(g) are histograms of annual, monthly and daily temperature simulations at GL and GO, which have been detrended by ensemble means. Only data of years 2020, 2040, 2060, 2080, and 2100 are used in daily data.

alone are insufficient to characterize data with greater temporal complexity.

3. Stochastic Generator Methodology

In this section, we detail the procedures of constructing an SG and generating emulations for the temperature simulations described in Section 2. Leveraging the SHT, our proposed SG offers efficient and adaptable capabilities for dealing with global climate simulations that exhibit: (a) a substantial size in terms of the number of observations in space and time; (b) nonstationary spatial dependencies among latitudes and land/ocean regions; and (c) non-Gaussian temporal trajectories.

The data characteristics outlined in Section 2 underscore the deterministic chaotic nature inherent in climate models (Lorenz 1963; Branstator and Teng 2010; Castruccio and Genton 2018),

which motivates us to decouple the data into deterministic and stochastic components as follows:

$$y_t^{(r)}(L_i, l_j) = m_t(L_i, l_j) + \sigma(L_i, l_j)Z_t^{(r)}(L_i, l_j). \quad (1)$$

Here, m_t and σ are deterministic functions responsible for the mean trend and standard error, respectively, and are shared across all ensembles. $Z_t^{(r)}(L_i, l_j)$ is the stochastic component at grid point (L_i, l_j), time point t, and ensemble r. Given the expansive parameter space, simultaneously inferring both the deterministic and stochastic components would be computationally infeasible. Therefore, we adopt a two-stage approach, commonly employed in existing work (Castruccio and Genton 2018; Jeong et al. 2019; Huang et al. 2023). First, we evaluate $m_t(L_i, l_j)$ and $\sigma(L_i, l_j)$ at each grid point, making the independence assumption regarding $\{Z_t^{(r)}(L_i, l_j)\}_{i=1,...,I;j=1,...,J;t=1,...,T;r=1,...,R}$.

Subsequently, we analyze the dependence structure of $Z_t^{(r)}(L_i, l_j)$ by detrending and rescaling $y_t^{(r)}(L_i, l_j)$ with the estimate of $m_t(L_i, l_j)$ and $\sigma(L_i, l_j)$, respectively.

3.1. Deterministic Component of SG

In the first stage of constructing the SG, we focus on determining the mean trend m_t. It is crucial that the mean trend is both simple enough to minimize the storage requirements for parameters and informative enough to capture the physical relationships between simulations and essential covariates. Let $y_t^{(r)}(L_i, l_j)$ represent the annually aggregated temperature data. Previous research (Castruccio et al. 2014; Huang et al. 2023) has shown its dependence on the RF trajectory and adopted an infinite distributed lag model:

$$m_t(L_i, l_j) = \beta_0(L_i, l_j) + \beta_1(L_i, l_j)x_t$$
$$+ \beta_2(L_i, l_j)\{1 - \rho(L_i, l_j)\} \sum_{s=1}^{\infty} \rho(L_i, l_j)^{s-1} x_{t-s},$$

where $\beta_0(L_i, l_j)$ is the intercept, $\beta_1(L_i, l_j)$ and $\beta_2(L_i, l_j)$ are slopes for the current and past RFs, respectively. Lag weights $\beta_2(L_i, l_j)\{1 - \rho(L_i, l_j)\}\rho(L_i, l_j)^{s-1}$ decrease the impact of past RFs exponentially by $\rho(L_i, l_j) \in [0, 1]$. Moreover, if $y_t^{(r)}(L_i, l_j)$ represents the monthly aggregated temperature data, additional harmonic terms $\{\cos(2\pi tk/12), \sin(2\pi tk/12)\}_{k=1}^{K_M}$ should be included to fit the interannual cycle as shown in Figure 1(d). That is,

$$m_t(L_i, l_j) = \beta_0(L_i, l_j) + \beta_1(L_i, l_j)x_{\lceil t/12 \rceil}$$
$$+ \beta_2(L_i, l_j)\{1 - \rho(L_i, l_j)\} \sum_{s=1}^{\infty} \rho(L_i, l_j)^{s-1} x_{\lceil t/12 \rceil - s}$$
$$+ \sum_{k=1}^{K_M} \left\{ a_k(L_i, l_j) \cos\left(\frac{2\pi tk}{12}\right) \right.$$
$$\left. + b_k(L_i, l_j) \sin\left(\frac{2\pi tk}{12}\right) \right\},$$

which would also be applied to the daily aggregated temperature by replacing K_M and $t/12$ with K_D and $t/365$, respectively. The larger the value of K_M (or K_D), the higher the frequency the harmonic terms can present. For the deterministic component, $\{\beta_0(L_i, l_j), \beta_1(L_i, l_j), \beta_2(L_i, l_j), \rho(L_i, l_j), \sigma(L_i, l_j), a_1(L_i, l_j), \ldots, a_K(L_i, l_j), b_1(L_i, l_j), \ldots, b_K(L_i, l_j)\}_{i=1,\ldots,I;j=1,\ldots,J}$ are all $(5 + 2K)IJ$ parameters to be evaluated and stored, where K can take values of 0, K_M, and K_D to represent the annual, monthly and daily temperature, respectively. With the independence assumption, we can efficiently estimate the parameters of the deterministic component in parallel for each grid point. The detailed inferential process is given in Section S3.1 of the supplementary materials.

3.2. Stochastic Component of SG

Now, we sequentially model the spatial and temporal dependence of the stochastic component $Z_t^{(r)}(L_i, l_j)$, which is a general principle for analyzing space-time data (Stein 2007; Castruccio and Genton 2018) to bypass the prohibitive computation.
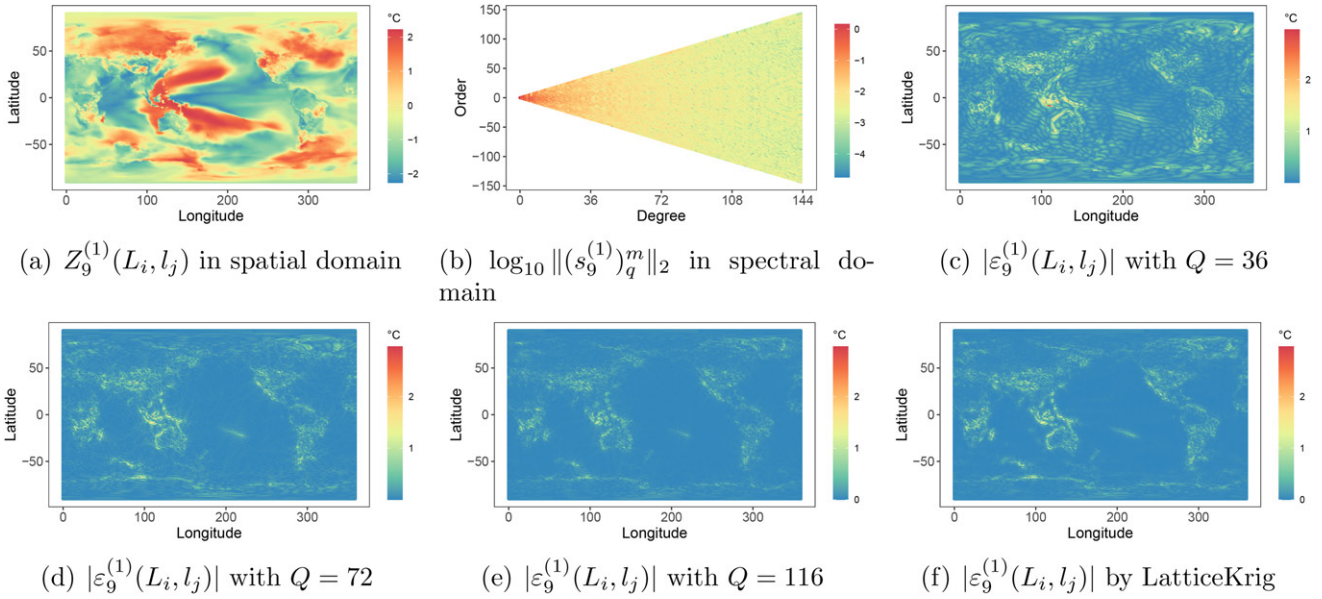
3.2.1. Modeling the Spatial Dependence

The inference of the spatial dependence should consider several crucial factors: (a) The simulations are extensive; (b) The simulations encompass the entire globe, making the use of chordal distance inappropriate (Jeong, Jun, and Genton 2017); and (c) The simulations exhibit nonstationarity among different latitudes and between land and ocean regions. Given these considerations, we propose using the SHT (Jones 1963; Stein 2007, 2008), which involves expanding the stochastic component $Z_t^{(r)}(L_i, l_j)$ with spherical harmonics:

$$Z_t^{(r)}(L_i, l_j) = \sum_{q=0}^{Q-1} \sum_{m=-q}^{q} (s_t^{(r)})_q^m H_q^m(L_i, l_j) + \varepsilon_t^{(r)}(L_i, l_j). \quad (2)$$

In (2), $\{H_q^m\}_{q=0,1,2,\ldots;m=-q,\ldots,q}$ is a complete set of orthonormal basis functions defined in $\mathcal{L}^2(\mathbb{S}^2)$ and indexed by integer degree q and order m, where $\mathcal{L}^2(\mathbb{S}^2)$ is the Hilbert space consisting of squared-integrable functions on the sphere $\mathbb{S}^2$. The spherical harmonic H_q^m is a complex-valued function with a closed form $H_q^m(L_i, l_j) = \sqrt{\frac{2q+1}{4\pi} \frac{(q-m)!}{(q+m)!}} P_q^m(\cos \theta_i) \exp(\iota m \psi_j)$, where $\iota^2 = -1$, $(\theta_i, \psi_j) = (-\pi L_i/180 + \pi/2, \pi l_j/180)$ is a reparameterization of (L_i, l_j) and P_q^m is the associated Legendre polynomial of degree q and order m, such that $H_q^{-m} = (-1)^m \overline{H_q^m}$. Corresponding to the H_q^m, $(s_t^{(r)})_q^m$ is the complex-valued spherical harmonic coefficient satisfying $(s_t^{(r)})_q^{-m} = (-1)^m \overline{(s_t^{(r)})_q^m}$. The norm of $(s_t^{(r)})_q^m$, that is, $\|(s_t^{(r)})_q^m\|_2 = \sqrt{\Re^2\{(s_t^{(r)})_q^m\} + \Im^2\{(s_t^{(r)})_q^m\}}$, indicates the energy gathered at degree q and order m. More details about spherical harmonics and their application on spectrum and differentiability analysis are given in Section S3.2.1 of the supplementary materials. Q is an integer ranging from 0 to $Q_{\max}$, with $Q_{\max} = \min\{I - 1, (J + 1)/2\}$ derived in Section S3.2.2 and determined by the spatial resolution of the data. The denser the grid points are, the higher the frequency they capture, and the larger the value of $Q_{\max}$ could be. For CESM2-LENS2, $Q_{\max} = 144$. The term $\varepsilon_t^{(r)}(L_i, l_j)$ accounts for the remaining information and is assumed to be independent and follow $\mathcal{N}(0, v^2(L_i, l_j))$. In a physical sense, (2) translates $\{Z_t^{(r)}(L_i, l_j)\}_{i=1,\ldots,I;j=1,\ldots,J}$ from the spatial domain to $\{(s_t^{(r)})_q^m\}_{q=0,\ldots,Q-1;l=-q,\ldots,q}$ in the spectral domain. In a statistical sense, with $Q \ll \sqrt{IJ}$, (2) provides a low-rank approximation for the stochastic process Z_t, where $(s_t)_q^m$ serves as a random coefficient and Q trades off the quality of approximation against the computational complexity.

We provide Figure 2 to facilitate the understanding of SHT. Figure 2(a) shows a set of stochastic components in the spatial domain, that is, $\{Z_9^{(1)}(L_i, l_j)\}_{i=1,\ldots,I;j=1,\ldots,J}$, which is obtained by detrending and rescaling the annually aggregated temperature at year 2023 and ensemble one with ensemble mean and standard error in Figure 1(a) and (b), respectively. After applying the SHT, Figure 2(b) presents the pattern of $\{\log_{10} \|(s_9^{(1)})_q^m\|_2\}_{q=0,1,\ldots,Q_{\max};m=-q,\ldots,q}$, which indicates the energy allocation in the spectral domain. The energy is more concentrated at lower degrees of lower frequencies, gradually decreasing as the degree increases. The decay rate reflects the differentiability property of the data, as detailed in Section S3.2.1 of

(a) $Z_9^{(1)}(L_i, l_j)$ in spatial domain

(b) $\log_{10} \|(s_9^{(1)})_q^m\|_2$ in spectral domain

(c) $|\varepsilon_9^{(1)}(L_i, l_j)|$ with $Q = 36$

(d) $|\varepsilon_9^{(1)}(L_i, l_j)|$ with $Q = 72$

(e) $|\varepsilon_9^{(1)}(L_i, l_j)|$ with $Q = 116$

(f) $|\varepsilon_9^{(1)}(L_i, l_j)|$ by LatticeKrig

Figure 2. Demonstration of SHT. (a) shows a set of stochastic component $\{Z_9^{(1)}(L_i, l_j)\}_{i=1,\dots,I;j=1,\dots,J}$ in the spatial domain. (b) indicates energy allocation of (a) in spectral domain. (c)–(e) depict the absolute values of $\{\varepsilon_9^{(1)}(L_i, l_j)\}_{i=1,\dots,I;j=1,\dots,J}$ in (2) with $Q = 36$, 72, and 116, respectively. (f) depicts the absolute values of $\{\varepsilon_9^{(1)}(L_i, l_j)\}_{i=1,\dots,I;j=1,\dots,J}$ obtained by using LatticeKrig with 13,658 bases of 6 levels, where $13{,}658 > 116^2 = 13{,}456$.

the supplementary materials. Figure 2(c)–(e) show the influence of Q on the loss of using SHT for a single set of stochastic components. Section S3.2.3 provides a more comprehensive illustration of the general approximation performance. In essence, a larger value of Q retains more information and provides a better approximation. Additionally, the spatial structure of surface temperature varies significantly between land and ocean.

We compare SHT with two popular bases used in low-rank approximations, empirical orthogonal functions (EOFs) and those in LatticeKrig (Nychka et al. 2016), as detailed in Section S3.2.4. EOFs, as data-driven basis functions, lack a closed form and must be stored. Their evaluation involves computing and eigen-decomposing an $IJ \times IJ$ matrix, leading to constraints on computational time and resources. Figure 2(e) and (f) illustrate comparable approximation performance of SHT and LatticeKrig on $Z_9^{(1)}(L_i, l_j)$ with similar numbers of bases. However, LatticeKrig entails more time for parameter tuning selection and coefficient calculation, along with greater storage requirements. For a more detailed and comprehensive comparison, readers are referred to Section S3.2.4. We emphasize the advantages of employing the spherical harmonics to represent substantial global climate simulations: (a) Natural basis functions. Spherical harmonics are eigenfunctions of the Laplace-Beltrami operator, making them well-suited for analyzing spherical data of various variables, as well as their spectrum; (b) Easily available multiresolution basis functions. Spherical harmonics have a closed form, which saves the procedures for getting bases, conserves the memory for storing them, and facilitates the derivation of special forms of covariance under certain assumptions. Their resolutions are automatically multiple, eliminating the need for researchers to manually allocate knots or tuning parameters. Instead, one can simply select a degree Q to control the level of detail; (c) Efficient coefficient calculation. As shown in Section S3.2.2, both the SHT and its inverse can be efficiently calculated with a computational time $O(Q^3)$. Moreover, this computational time can be further reduced to $O(Q^2)$ by parallel

computing. On a MacBook with a 3.2 GHz Apple M1 Pro processor with ten cores, performing the SHT with $Q = Q_{\max}$ in Figure 2 takes about 3 sec, and the inverse SHT with $Q = 36$, 72, and 144 take no more than 0.06, 0.3, and 2 sec, respectively.

Next, we model the spatial dependence structure, which is assumed to remain consistent over time. The covariance between $Z_t^{(r)}(L_i, l_j)$ and $Z_t^{(r)}(L_{i'}, l_{j'})$ is

$$c\{(L_i, l_j), (L_{i'}, l_{j'})\} = \mathbf{H}(L_i, l_j)^\top \mathbf{K} \overline{\mathbf{H}(L_{i'}, l_{j'})} + \delta_{i=i'}\delta_{j=j'} v^2(L_i, l_j), \tag{3}$$

where $\mathbf{K} = \mathbf{K}_t^{(r)} = E\{\mathbf{s}_t^{(r)}(\mathbf{s}_t^{(r)})^H\}$ is a $Q^2 \times Q^2$ covariance matrix of $\{(s_t^{(r)})_q^m\}_{q=0,1,\dots,Q;m=-q,\dots,q}$, H is the operator of conjugate transpose, $\mathbf{H}(L_i, l_j)$ and $\mathbf{s}_t^{(r)}$ are vectors consisting of basis functions and spherical harmonic coefficients, indexed by $(q^2 + q + m + 1)$. Specifically, the $(q^2 + q + m + 1)$th elements of $\mathbf{H}(L_i, l_j)$ and $\mathbf{s}_t^{(r)}$ are $H_q^m(L_i, l_j)$ and $(s_t^{(r)})_q^m$, respectively.

Further examining the structure of $\mathbf{K}$, it is evident that the stochastic component in Figure 2(a) exhibits strong heterogeneity across different latitudes, as observed in previous studies of global climate data (Stein 2007; Jeong et al. 2018; Huang et al. 2023). Consequently, the commonly used isotropic assumption for Euclidean data is no longer appropriate. Instead, we adopt the assumption of axial symmetry, which posits that only data on the same latitude are stationary (Jones 1963). This leads to the covariance $c\{(L_i, l_j), (L_{i'}, l_{j'})\}$ becoming a function of L_i (θ_i), $L_{i'}$ ($\theta_{i'}$), and $|l_j - l_{j'}|$ ($|\psi_j - \psi_{j'}|$, the central angle between points (L_i, l_j) and $(L_{i'}, l_{j'})$ expressed in chordal distance). Using a simple derivation outlined in Jones (1963), we have

$$c(L_i, L_{i'}, |l_j - l_{j'}|)$$

$$= \sum_{q=0}^{Q-1} \sum_{q'=0}^{Q-1} \sum_{m=-q}^{q} k_{qq'm} H_q^m(\theta_i, 0) H_{q'}^m(\theta_{i'}, 0) \exp(\iota m |\psi_j - \psi_{j'}|) \tag{4}$$

with $\mathrm{E}\{(s_t^{(r)})_q^m (s_t^{(r)})_{q'}^{m'}\} = \delta_{m=m'} k_{qq'm}$. It means that axial symmetry assumes dependence only on coefficients of the same order. Thus, $\mathbf{K}$ becomes a sparse matrix, consisting of $Q^2 + (Q-1)^2 + \cdots + 1^2 = Q^3/3 + Q^2/2 + Q/6$ nonzero elements.

However, from Figures 1 and 2, we still observe heterogeneity between land and ocean data on the same latitude. The stochastic components on land and ocean exhibit different spatial structures, requiring different numbers of basis functions to capture them adequately. Therefore, we propose replacing (2) by

$$Z_t^{(r)}(L_i, l_j) = \delta_{(L_i, l_j) \in \mathcal{S}_l} \sum_{q=0}^{Q_l-1} \sum_{m=-q}^{q} (s_t^{(r)})_q^m H_q^m(L_i, l_j)$$
$$+ \delta_{(L_i, l_j) \in \mathcal{S}_o} \sum_{q=0}^{Q_o-1} \sum_{m=-q}^{q} (s_t^{(r)})_q^m H_q^m(L_i, l_j) + \varepsilon_t^{(r)}(L_i, l_j),$$
(5)

where $\mathcal{S}_o$ ($\mathcal{S}_l$) represents the set of grid points over the ocean (land) and Q_o (Q_l) denotes the value of Q for ocean (land). We ignore the continuity of temperature changes between land and ocean, since (5) is formulated on grid points and our target is not to predict at new points. Accordingly, the covariance $c(L_i, L_i, |l_j - l_{j'}|)$ is given by plugging the Q values of (L_i, l_j) and $(L_{i'}, l_{j'})$ into (4). We present the process of choosing appropriate values for Q_o and Q_l in Section 4.

3.2.2. Modeling the Temporal Dependence

Next, we investigate the temporal dependence within the vector time series $\{\mathbf{s}_t^{(r)}\}_{t=1}^{T}$ in the spectral domain. We will employ an auto-regressive model, but with two additional operations: converting the coefficients to real-valued ones and Gaussianizing them using the TGH transformation. Specifically, for the complex-valued vector $\mathbf{s}_t^{(r)}$ with the special structure of $(s_t^{(r)})_q^{-m} = (-1)^m \overline{(s_t^{(r)})_q^m}$, we first linearly transform it into a real-valued vector:

$$\tilde{\mathbf{s}}_t^{(r)} = \mathbf{A}^{-1} \mathbf{s}_t^{(r)} = [(s_t^{(r)})_0^0, \Im\{(s_t^{(r)})_1^1\}, (s_t^{(r)})_1^0, \Re\{(s_t^{(r)})_1^1\}, \ldots,$$
$$\Im\{(s_t^{(r)})_{Q-1}^{Q-1}\}, \ldots, \Re\{(s_t^{(r)})_{Q-1}^{Q-1}\}]^\top,$$

where both $\mathbf{A}$ and its inverse are $Q^2 \times Q^2$ sparse matrices with known structures. For example, in the case of $Q = 2$, $\mathbf{A}$ and $\mathbf{A}^{-1}$ can be represented as

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \iota & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & \iota & 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\iota/2 & 0 & -\iota/2 \\ 0 & 0 & 1 & 0 \\ 0 & -1/2 & 0 & 1/2 \end{pmatrix},$$

respectively. For simplicity, we denote the $(q^2 + q + m + 1)$th element of $\tilde{\mathbf{s}}_t^{(r)}$ as $(\tilde{s}_t^{(r)})_q^m$. Consequently, we have $Z_t^{(r)}(L_i, l_j) = \mathbf{H}(L_i, l_j)^\top \mathbf{A} \tilde{\mathbf{s}}_t^{(r)} + \varepsilon_t^{(r)}(L_i, l_j)$, where $\mathbf{A}^\top \mathbf{H}(L_i, l_j)$ and $\tilde{\mathbf{s}}_t^{(r)}$ are vectors of bases and coefficients, respectively. Under the assumption of axial symmetry, the covariance of $\tilde{\mathbf{s}}_t^{(r)}$, denoted as $\tilde{\mathbf{K}}_0 = \mathrm{E}(\tilde{\mathbf{s}}_t^{(r)} \tilde{\mathbf{s}}_t^{(r)\top})$, is also a sparse matrix. Specifically, $\mathrm{E}\{(\tilde{s}_t^{(r)})_q^m (\tilde{s}_t^{(r)})_{q'}^{m'}\} = \mathrm{E}\{(\tilde{s}_t^{(r)})_q^{-m} (\tilde{s}_t^{(r)})_{q'}^{-m'}\} = \tilde{k}_{qq'm} \delta_{m=m'} \delta_{m \geq 0}$. The derivation is given in Section S3.3 of supplementary materials. The matrix $\tilde{\mathbf{K}}_0$ consists of $2Q^3/3 + Q/3$ nonzero elements and only $Q^3/3 + Q^2/2 + Q/3$ ones need to be stored. If $Z_t^{(r)}(L_i, l_j)$

represents the stochastic component of annually aggregated temperature, it is reasonable to assume the Gaussianity of $\{Z_t^{(r)}(L_i, l_j)\}_{t=1}^{T}$, and therefore, the Gaussianity of $\{\tilde{\mathbf{s}}_t^{(r)}\}_{t=1}^{T}$. This allows us to model it with a vector auto-regressive model of order P (VAR(P)): $\tilde{\mathbf{s}}_t^{(r)} = \sum_{p=1}^{P} \boldsymbol{\Phi}_p \tilde{\mathbf{s}}_{t-p}^{(r)} + \boldsymbol{\xi}_t^{(r)}$, where $\boldsymbol{\xi}_t^{(r)} \overset{iid}{\sim} \mathcal{N}_{Q^2}(\mathbf{0}, \mathbf{U})$ and $\mathbf{U} = \tilde{\mathbf{K}}_0 - \sum_{p=1}^{P} \boldsymbol{\Phi}_p \tilde{\mathbf{K}}_0 \boldsymbol{\Phi}_p^\top - \sum_{p=1}^{P} \sum_{p' \neq p}^{P} \boldsymbol{\Phi}_p \tilde{\mathbf{K}}_{|p'-p|} \boldsymbol{\Phi}_{p'}^\top$ with $\tilde{\mathbf{K}}_{|p'-p|} = \mathrm{E}(\tilde{\mathbf{s}}_{t-p}^{(r)} \tilde{\mathbf{s}}_{t-p'}^{(r)\top})$.

However, as in Sections 2 and S2, when $Z_t^{(r)}(L_i, l_j)$ is the stochastic component of the monthly or daily aggregated temperature, the Gaussianity assumption for time series $\{Z_t^{(r)}(L_i, l_j)\}_{t=1}^{T}$ at most grid points (L_i, l_j) may not hold. It implies that $\{(\tilde{s}_t^{(r)})_q^m\}_{t=1}^{T}$ at some (q, m) pairs are not Gaussian and motivates us to employ a modified TGH transformation. Denote by $\mathcal{S}_{gh}$ the set of (q, m) such that $\{(\tilde{s}_t^{(r)})_q^m\}_{t=1}^{T}$ rejects the Gaussianity. We model the coefficients with a TGH auto-regressive model: $\check{\mathbf{s}}_t^{(r)} = \sum_{p=1}^{P} \boldsymbol{\Phi}_p \check{\mathbf{s}}_{t-p}^{(r)} + \boldsymbol{\xi}_t^{(r)}$, where the $(q^2 + q + m + 1)$th element of $\check{\mathbf{s}}_t^{(r)}$ is

$$(\check{s}_t^{(r)})_q^m = \begin{cases} \lambda_q^m \tau_{g_q^m, h_q^m}^{-1}\{(\tilde{s}_t^{(r)})_q^m / \omega_q^m\}, & \text{if} \quad (q, m) \in \mathcal{S}_{gh}, \\ (\tilde{s}_t^{(r)})_q^m, & \text{if} \quad (q, m) \notin \mathcal{S}_{gh}, \end{cases}$$
(6)

with $\tau_{g_q^m, h_q^m}(s) = (g_q^m)^{-1}\{\exp(g_q^m s) - 1\} \exp(h_q^m s^2 / 2)$. Compared to the regular TGH, the one in (6) removes a location parameter and introduces a scale parameter λ_q^m. The former is due to the zero-mean $\{(\tilde{s}_t^{(r)})_q^m\}_{t=1,\ldots,T; r=1,\ldots,R}$. The later ensures the standard deviation of $\{(\check{s}_t^{(r)})_q^m\}_{t=1,\ldots,T; r=1,\ldots,R}$ to be equal to that of $\{(\tilde{s}_t^{(r)})_q^m\}_{t=1,\ldots,T; r=1,\ldots,R}$. Four additional parameters describing characteristics beyond the first two moments are to be stored. Based on (6), the covariance matrix of $\boldsymbol{\xi}_t^{(r)}$ is $\mathbf{U} = \check{\mathbf{K}}_0 - \sum_{p=1}^{P} \boldsymbol{\Phi}_p \check{\mathbf{K}}_0 \boldsymbol{\Phi}_p^\top - \sum_{p=1}^{P} \sum_{p' \neq p}^{P} \boldsymbol{\Phi}_p \check{\mathbf{K}}_{|p-p'|} \boldsymbol{\Phi}_{p'}^\top$, where both $\check{\mathbf{K}}_0 = \mathrm{E}(\check{\mathbf{s}}_t^{(r)} \check{\mathbf{s}}_t^{(r)\top})$ with $\mathrm{E}\{(\check{s}_t^{(r)})_q^m (\check{s}_t^{(r)})_{q'}^{m'}\} = \mathrm{E}\{(\check{s}_t^{(r)})_q^{-m} (\check{s}_t^{(r)})_{q'}^{-m'}\} = \check{k}_{qq'm} \delta_{m=m'} \delta_{m \geq 0}$ and $\check{\mathbf{K}}_{|p-p'|} = \mathrm{E}(\check{\mathbf{s}}_{t-p}^{(r)} \check{\mathbf{s}}_{t-p'}^{(r)\top})$ have sparse structures under the assumption of axial symmetry.

Furthermore, we assume that $\boldsymbol{\Phi}_p$ is a diagonal matrix with the $(q^2 + q + m + 1)$th diagonal element denoted as $(\phi_p)_q^m$, which neglects the cross dependence between $\{(\tilde{s}_t^{(r)})_q^m\}_{q=0,\ldots,Q-1; m=-q,\ldots,q}$. This assumption enables the independent and parallel evaluation of $\{(\phi_p)_q^m\}_{p=1}^{P}$ across q and m and ensures $\mathbf{U}$ to inherit the same sparse structure from $\check{\mathbf{K}}_0$ and $\check{\mathbf{K}}_{|p-p'|}$. We briefly validated it in Section S3.4 by testing the significance of the first temporal lag of the cross-correlation. The proportion of p-values less than 0.05 is nearly zero, indicating negligible cross-dependence. The details about the parameter estimation are given in Section S3.5.

3.3. Emulation with SG

The procedures of constructing an SG for the monthly and daily simulations from CESM2-LENS2 is summarized in Algorithm S1 of Section S3.6. We assume that the tuning parameters K, Q_l, Q_o, and P are known. More details about their selection will be provided in Section 4. In Stage 1, we calculate and store a total of $(5 + 2K)IJ$ parameters in the deterministic component.

In Step 1) of Stage 2, although we use different Q values for land and ocean data to evaluate $\nu(L_i, l_j)$, we still need coefficients up to $Q' = \max(Q_l, Q_o)$. The number of stored parameters in Stage 2 is $IJ + (4 + P)Q'^2 + (Q'^3/3 + Q'^2/2 + Q'/3)$. Therefore, the total number of parameters is of the order $O(IJ + Q'^3)$. For the annually aggregated temperature, there is no need to perform a TGH transformation in Step 3) of Stage 2. Consequently, $\mathcal{S}_{gh}$ is an empty set, and only $IJ + PQ'^2 + (Q'^3/3 + Q'^2/2 + Q'/3)$ parameters are needed for the stochastic component. In summary, we develop an SG by training on simulations of size $IJRT$ and store it for generating unlimited emulations using at most $(6 + 2K)IJ + (4 + P)Q'^2 + (Q'^3/3 + Q'^2/2 + Q'/3)$ memorized parameters.

Here, we contrast the storage requirements of our SG with that of another efficient SG proposed by Huang et al. (2023) (referred to as HCBG-SG) for surface temperature simulations from LENS1 (Kay et al. 2015). HCBG-SG was primarily developed within the spatial domain and required memorization of $(2K + P_H + 8)IJ + 6I + 2$ parameters. This consisted of $(5 + 2K)IJ$ parameters for the deterministic component, $P_H IJ$ parameters for evaluating the temporal dependence with P_H representing the selected order for the auto-regressive model, $3IJ$ parameters for TGH transformation, and $6I + 2$ parameters for assessing the nonstationary spatial dependence. The storage advantage of our SG stems from its modeling in the spectral domain rather than the spatial domain, with $Q'^2 \ll IJ$. Taking the case of $P = P_H = 1$ and $Q' = 70$ as an example, which maximizes storage requirements of our SG for all temporal resolutions, the daily HCBG-SG necessitates additional 25,747 parameters. However, for annual simulations, which do not need a TGH, HCBG-SG needs 120,541 fewer parameters. A further comparison of their performance is presented in Section 4.

In Algorithm 1, we outline the procedures for generating monthly and daily temperature emulations for CESM2-LENS2. Note that a Cholesky decomposition should be performed on $\mathbf{U}$ when generating $\check{\mathbf{s}}_t^{(r')}$, which may be time-consuming if Q' is not small. However, the sparse structure of $\mathbf{U}$ avoids the possible computational limitation. For the annually aggregated temperature, we can simply remove Step (b) in Stage 2 and replace $\tilde{\mathbf{s}}_{SG,t}^{(r')}$ in Step (c) with $\check{\mathbf{s}}_{SG,t}^{(r')}$.

4. Case Study: CESM2-LENS2

In this section, we develop emulators and generate emulations for surface temperature simulations from CESM2-LENS2 using the proposed algorithms. We delve into the implementation of our SG (referred to as SHT-SG in this section), covering the choice of tuning parameters and the modeling of the dependence structure, and the assessment of their performance. Additionally, we include the results of HCBG-SG for comparison. Discussion about the monthly SG is presented in Section S4.2 of the supplementary materials to save space.

The assessment of an SG relies on the specific purpose for which the emulator is intended (Castruccio et al. 2014). Our main goal is to generate emulations that closely resemble the given simulations. To achieve this, our SG should efficiently decouple and capture the deterministic and stochastic components of the data. First, the estimated mean trend $\hat{m}_t$ should

---

Algorithm 1 Emulation with SG

Input: $\rho(L_i, l_j)$, $\beta_0(L_i, l_j)$, $\beta_1(L_i, l_j)$, $\beta_2(L_i, l_j)$, $a_1(L_i, l_j)$,..., $a_K(L_i, l_j)$, $b_1(L_i, l_j)$,..., $b_K(L_i, l_j)$, $\sigma(L_i, l_j)$, $\nu(L_i, l_j)$, λ_q^m, ω_q^m, g_q^m, h_q^m, $(\phi_1)_q^m$,..., $(\phi_P)_q^m$, $i = 1, \ldots, I$, $j = 1, \ldots, J$, $q = 0, \ldots, Q_o - 1$, $m = -q, \ldots, q$, $\mathbf{U}$, Q_l, Q_o, K, P, $\mathbf{A} \in \mathbb{R}^{Q_o^2 \times Q_o^2}$, R'.

Stage 1: Preliminary
1) For each grid point (L_i, l_j), calculate $\{m_t(L_i, l_j)\}_{t=1}^T$ with given parameters of the deterministic component, $i = 1, \ldots, I$ and $j = 1, \ldots, J$.

Stage 2: Emulation
For $r' = 1, \ldots, R'$:
1) generate $\check{\mathbf{s}}_{SG,-(p-1)}^{(r')} \sim \mathcal{N}_{Q_o^2}(\mathbf{0}, \mathbf{U})$, $p = 1, \ldots, P$,
2) for $t = 1, \ldots, T$:
(a) generate $\boldsymbol{\xi}_{SG,t}^{(r')} \sim \mathcal{N}_{Q_o^2}(\mathbf{0}, \mathbf{U})$, and calculate $\check{\mathbf{s}}_{SG,t}^{(r')} = \sum_{p=1}^P \Phi_p \check{\mathbf{s}}_{SG,t-p}^{(r')} + \boldsymbol{\xi}_{SG,t}^{(r')}$,
(b) calculate $\tilde{\mathbf{s}}_{SG,t}^{(r')}$, where $(\tilde{s}_{SG,t}^{(r')})_q^m = \omega_q^m \tau_{g_q^m, h_q^m} \{(\check{s}_{SG,t}^{(r')})_q^m / \lambda_q^m\}$,
(c) calculate $\mathbf{s}_{SG,t}^{(r')} = \mathbf{A}\tilde{\mathbf{s}}_{SG,t}^{(r')}$, and perform the inverse SHT on $\mathbf{s}_{SG,t}^{(r')}$ with (5) to obtain $\{\text{invSHT}(\mathbf{s}_{SG,t}^{(r')})(L_i, l_j)\}_{i=1,\ldots,I;j=1,\ldots,J}$,
(d) generate $\varepsilon_{SG,t}^{(r')}(L_i, l_j) \sim \mathcal{N}(0, \hat{\nu}(L_i, l_j)^2)$, $i = 1, \ldots, I; j = 1, \ldots, J$;
(e) generate emulations

$$y_{SG,t}^{(r')}(L_i, l_j) = m_t(L_i, l_j) + \sigma(L_i, l_j)\{\text{invSHT}(\mathbf{s}_{SG,t}^{(r')})(L_i, l_j) + \varepsilon_{SG,t}^{(r')}(L_i, l_j)\}.$$

Output: $\{y_{SG,t}^{(r')}(L_i, l_j)\}_{i=1,\ldots,I;j=1,\ldots,J;r'=1,\ldots,R'}$.

---

mimic and extract the variation of the ensemble mean in the simulations. After adding the stochastic component, the variability of emulations should be similar to the internal variability observed in simulations. Therefore, we will discuss two key factors: goodness-of-fit and variability, quantified with

$$I_{\text{fit}}(L_i, l_j) = \frac{\sum_{r=1}^R \sum_{t=1}^T \{y_t^{(r)}(L_i, l_j) - \hat{m}_t(L_i, l_j)\}^2}{R/(R-1) \sum_{r=1}^R \sum_{t=1}^T \{y_t^{(r)}(L_i, l_j) - \bar{y}_t(L_i, l_j)\}^2}$$

and

$$I_{\text{uq}}(L_i, l_j) = \frac{\text{CRA}\{y_{SG,t}^{(1)}(L_i, l_j), \ldots, y_{SG,t}^{(R)}(L_i, l_j)\}}{\text{CRA}\{y_t^{(1)}(L_i, l_j), \ldots, y_t^{(R)}(L_i, l_j)\}},$$

respectively. Given a grid point (L_i, l_j), the index $I_{\text{fit}}(L_i, l_j)$ compares $\hat{m}_t(L_i, l_j)$ with $2K + 4$ parameters with the ensemble mean $\bar{y}_t(L_i, l_j) = R^{-1} \sum_{r=1}^R y_t^{(r)}(L_i, l_j)$ with T parameters among all ensembles and time points. The closer the I_{fit} value to 1, the better the ability of the SG to capture the variability of the ensemble mean. Index I_{uq}, a measure proposed in Huang et al. (2023), leverages the concept of functional data depths (López-Pintado and Romo 2009). In particular, we treat $\{y_{SG,t}^{(r')}(L_i, l_j)\}_{r'=1}^R$ as R functions of t at (L_i, l_j) and measure the interquartile range (Sun and Genton 2011) of these functions using the central

(a) $\mathrm{I}_{\mathrm{fit}}$ versus R

(b) $\hat{\sigma}(L_i, l_j)$

(c) Choice of Q

(d) $\hat{v}(L_i, l_j)$
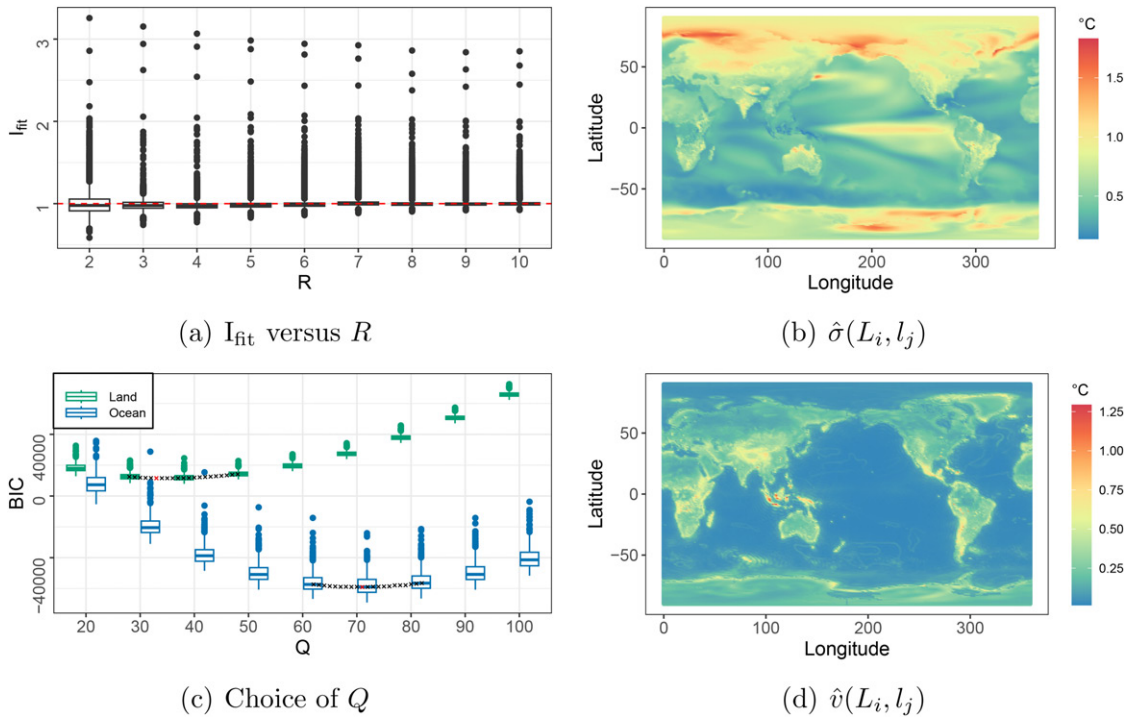
Figure 3. Inference process for the annual SG. (a) shows boxplots of $\{\mathrm{I}_{\mathrm{fit}}(L_i, l_j)\}_{i=1,...,I; j=1,...,J}$ for different values of R. (b) is the map of the estimated standard deviation $\{\hat{\sigma}(L_i, l_j)\}_{i=1,...,I; j=1,...,J}$. (c) shows boxplots of $\{\mathrm{BIC}_{*,t}^{(r)}(Q)\}_{r=1,...,R; t=1,...,T}$ for different values of Q, where $*$ takes "l" and "o". Points $\times$ show medians of BIC values. The red points indicate the minimum values. (d) is the map of $\{\hat{v}(L_i, l_j)\}_{i=1,...,I; j=1,...,J}$.

region area (CRA). Then, $\mathrm{I}_{\mathrm{uq}}(L_i, l_j)$ is the uncertainty qualification of $\{y_{SG,t}^{(r')}(L_i, l_j)\}_{r'=1}^R$ against that of $\{y_t^{(r)}(L_i, l_j)\}_{r=1}^R$. A value of I_{uq} close to 1 suggests that the variability of the generated emulations closely reflects the internal variability of the training simulations.

Moreover, we will visually and numerically compare various statistical characteristics of the emulations to those of simulations. This ensures that our emulations faithfully capture the essential features of the temperature data from CESM2-LENS2. For example, at each grid point, we compare the empirical distribution of $\{y_{SG,t}^{(r')}(L_i, l_j)\}_{t=1,...,T; r'=1,...,R}$ to that of $\{y_t^{(r)}(L_i, l_j)\}_{t=1,...,T; r=1,...,R}$ by the Wasserstein distance (Santambrogio 2015), which is denoted as $\mathrm{WD}_S(L_i, l_j)$. Similarly, we calculate the Wasserstein distance between the empirical distribution of $\{y_{SG,t}^{(r')}(L_i, l_j)\}_{i=1,...,I; j=1,...,J; r'=1,...,R}$ and $\{y_t^{(r)}(L_i, l_j)\}_{i=1,...,I; j=1,...,J; r=1,...,R}$ at each time point t, which is denoted as $\mathrm{WD}_T(t)$.
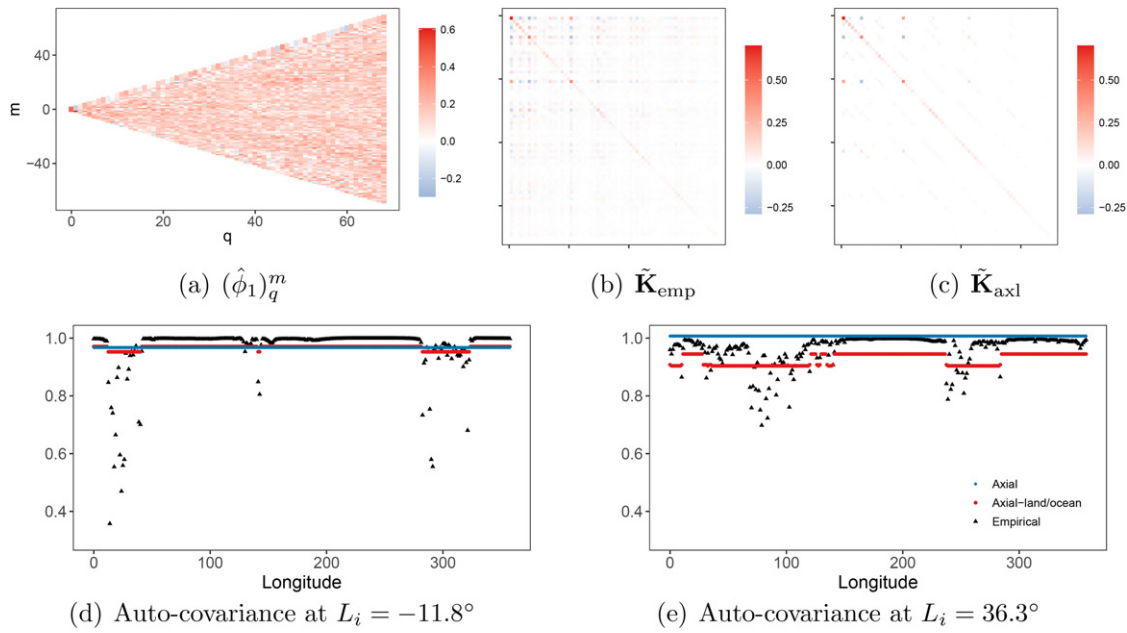
4.1. Annually Aggregated Temperature

First, we evaluate the deterministic component m_t and σ of the annually aggregated temperature simulations $\{y_t^{(r)}(L_i, l_j)\}_{i=1,...,I; j=1,...,J; t=1,...,T; r=1,...,R}$, with a total of $T = 86$ time points. Figure 3(a) presents boxplots of $\{\mathrm{I}_{\mathrm{fit}}(L_i, l_j)\}_{i=1,...,I; j=1,...,J}$ for different values of R. These values tend to get closer to 1 as R increases, and they stabilize after $R \geq 7$, which means that using seven training simulations is sufficient to stabilize the inference. Therefore, we select $R = 7$ for the subsequent analysis. The corresponding

$\{\mathrm{I}_{\mathrm{fit}}(L_i, l_j)\}_{i=1,...,I; j=1,...,J}$ values are shown in Figure S5(a), with a median value of 1.001. Temperature can be fitted well at most grid points. Furthermore, Figure 3(b) displays the map of $\hat{\sigma}(L_i, l_j)$, illustrating the nonstationarity of the data. It is evident that larger variations of surface temperatures are observed over land.

Then, we perform the SHT on the stochastic component, necessitating the selection of appropriate values for Q_l and Q_o. Based on the Gaussian assumption of $\varepsilon_t^{(r)}(L_i, l_j)$, we propose a Bayesian information criterion (BIC) for $\{Z_t^{(r)}(L_i, l_j)\}_{i=1,...,I; j=1,...,J}$ at time t of ensemble r:

$$\mathrm{BIC}_{*,t}^{(r)}(Q) = \log(n_*)Q^2 + n_* \log(2\pi)$$
$$+ \delta_{(L_i, l_j) \in \mathcal{S}_*} \sum_{i=1}^{I} \sum_{j=1}^{J} \log\{\hat{v}^2(L_i, l_j)\}$$
$$+ \delta_{(L_i, l_j) \in \mathcal{S}_*} \sum_{i=1}^{I} \sum_{j=1}^{J} \left\{ \frac{\varepsilon_t^{(r)}(L_i, l_j)}{\hat{v}(L_i, l_j)} \right\}^2,$$

where $\hat{v}^2(L_i, l_j) = (RT)^{-1} \sum_{r=1}^{R} \sum_{t=1}^{T} \{\varepsilon_t^{(r)}(L_i, l_j)\}^2$. When $*$ takes "o", the above BIC with $n_o = \sum_{i=1}^{I} \sum_{j=1}^{J} \delta_{(L_i, l_j)} \in \mathcal{S}_o$ helps select an appropriate Q_o for $Z_t^{(r)}(L_i, l_j)$ over the ocean. Similarly, when $*$ is "l", it leads to the determination of Q_l. Figure 3(c) illustrates boxplots of $\{\mathrm{BIC}_{*,t}^{(r)}\}_{r=1,...,R; t=1,...,T}$ against different values of Q, revealing a substantial discrepancy in preferred Q values between land and ocean. The stochastic components on land exhibit a preference for a smaller Q value compared to those on the ocean. This might seem counterintuitive, given that Figure 2(c) and (d) suggest that the stochastic component on

(a) $(\hat{\phi}_1)_q^m$

(b) $\tilde{\mathbf{K}}_{\mathrm{emp}}$

(c) $\tilde{\mathbf{K}}_{\mathrm{axl}}$

(d) Auto-covariance at $L_i = -11.8°$

(e) Auto-covariance at $L_i = 36.3°$

Figure 4. Temporal and spatial dependence structure. (a) is the map of the estimated $\{(\phi_1)_q^m\}_{q=0,\dots,Q_o-1;m=-q,\dots,q}$ for the annual simulations. (b) is the first 69 rows and columns of $\tilde{\mathbf{K}}_0$ empirically evaluated by $(RT)^{-1}\sum_{r=1}^R\sum_{t=1}^T \tilde{\mathbf{s}}_t^{(r)}\tilde{\mathbf{s}}_t^{(r)\top}$ and denoted as $\tilde{\mathbf{K}}_{\mathrm{emp}}$. (c) is the first 69 rows and columns of $\tilde{\mathbf{K}}_0$ evaluated under the axial symmetric assumption and denoted as $\tilde{\mathbf{K}}_{\mathrm{axl}}$. (d) and (e) show empirical and fitted auto-covariance at latitudes $-11.8°$ and $36.3°$, respectively.

land requires more harmonic terms for a better approximation. However, it is essential to note that the introduction of Q aims not only for accurate approximation but also to strike a balance between accuracy and the number of parameters, leading to improved model selection. Examining the boxplots for land, the improvement gained by including additional terms does not fully compensate for the associated drawbacks, such as increased storage requirements and accumulated errors from evaluation. The subsequent Figure 4(d) and (e) further validate this model selection result. Finally, the $\times$ points in Figure 3(c) suggest selecting $Q_l = 35$ and $Q_o = 69$ to minimize the medians of BIC for land and ocean, respectively.
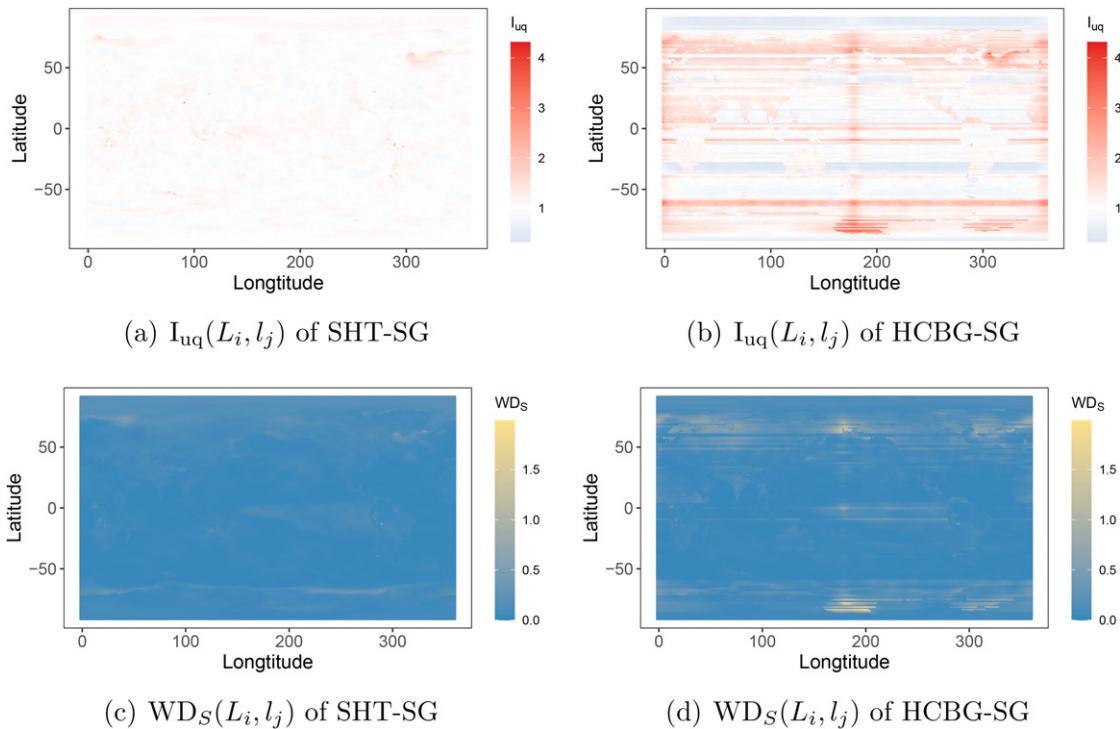
In Figure 3(d), we further analyze the values of $\hat{v}(L_i, l_j)$, which serves as an assessment of the low-rank approximation on (L_i, l_j). There are several regions on land that exhibit significantly larger $\hat{v}(L_i, l_j)$: (a) High-altitude regions such as the Himalayas and the Andes ranges with altitudes exceeding 3500 meters above sea level; and (b) Regions characterized by diverse topography, such as the Indonesian archipelago, encompassing islands, mountains, and coastal areas. These geographic factors contribute to a more complex spatial and temporal structure in surface temperature in these regions. To capture more details, one may consider including additional covariates in the deterministic component and choosing specific values of Q for these regions in the stochastic component. These considerations are avenues for future exploration.

Next, we assess the temporal dependence by fitting $\{(\tilde{s}_t^{(r)})_q^m\}_{t=1,\dots,T;r=1,\dots,R}$ with the VAR(P) model. For each (q, m), we suggest to choose P by minimizing $\mathrm{BIC}_q^m(P) = P\log\{(T-P)R\} + R(T-P)\{\log(2\pi)+1\} + R(T-P)\log((\hat{u}_q^m)^2)$, where $(\hat{u}_q^m)^2 = R^{-1}(T-P)^{-1}\sum_{r=1}^R\sum_{t=P+1}^T\{(\tilde{s}_t^{(r)})_q^m - \sum_{p=1}^P(\hat{\phi}_p)_q^m(\tilde{s}_{t-p}^{(r)})_q^m\}^2$. For briefness, we use the conditional log-likelihood (given $(\tilde{s}_1^{(r)})_q^m, \dots, (\tilde{s}_P^{(r)})_q^m$) rather than an exact

log-likelihood in the above BIC. The proportions of $P = 1,\dots,5$ are 99.3%, 0.2%, 0.2%, 0.1%, and 0.2%, respectively. Therefore, we opt for $P = 1$ for the annual case. Figure 4(a) further demonstrates the estimates of $\{(\phi_1)_q^m\}_{q=0,\dots,Q_o-1;m=-q,\dots,q}$.

Finally, we evaluate the spatial dependence by examining the matrix $\tilde{\mathbf{K}}_0$. Figure 4(b) and (c) show top-left corners of two evaluations of $\tilde{\mathbf{K}}_0$. $\tilde{\mathbf{K}}_{\mathrm{emp}}$ is the sample covariance matrix of $\tilde{\mathbf{s}}_t^{(r)}$, forming a dense $Q_o^2 \times Q_o^2$ matrix. In contrast, $\tilde{\mathbf{K}}_{\mathrm{axl}}$ is a sparse matrix. The sparsity of $\tilde{\mathbf{K}}_{\mathrm{axl}}$ would be inherited by $\mathbf{U}$, facilitating storage and enabling fast Cholesky decomposition using efficient algorithms (Furrer and Sain 2010). Figure 4(d) and (e) display empirical and fitted covariances for $Z(L_i, l_j)$ and $Z(L_i, l_{j+1})$, specifically the auto-covariance $\{\mathrm{cov}\{Z(L_i, l_j), Z(L_i, l_{j+1})\}\}_{j=1}^{J-1}$ at $L_i = -11.8°$ and $36.3°$. These plots highlight the distinction between the proposed separate SHT for land and ocean (5) (Axial-land/ocean) and the original SHT (2) (Axial). The empirical auto-covariance over the ocean appears almost flat, supporting the assumption of axial symmetry over the ocean. In contrast, the auto-covariance over land significantly differs from that over the ocean. This difference is not captured by the fitted covariance of Axial, which selects $Q = 77$ via BIC. Axial-land/ocean fits the covariance separately for land and ocean, resulting in segmented lines in Figure 4(d) and (e). However, land temperature exhibits a complex and nonstationary dependence structure, with covariance weakening as one moves farther inland. This intricacy warrants further investigation in future research.

Inputting the above-estimated parameters, we generate $R' = R = 7$ emulations with Algorithm 1. Each emulation takes approximately 30 sec to produce without parallel computation. Additionally, we include emulations generated by HCBG-SG for comparison, with detailed implementation procedures shown in the supplementary materials. Figure 5(a) and (b) compare I_{uq} values for the proposed SHT-SG and HCBG-SG, with median

(a) $\mathrm{I_{uq}}(L_i, l_j)$ of SHT-SG

(b) $\mathrm{I_{uq}}(L_i, l_j)$ of HCBG-SG

(c) $\mathrm{WD}_S(L_i, l_j)$ of SHT-SG

(d) $\mathrm{WD}_S(L_i, l_j)$ of HCBG-SG

Figure 5. (a) and (b) are maps of $\{\mathrm{I_{uq}}(L_i, l_j)\}_{i=1,\dots,I;j=1,\dots,J}$ for the annual temperature emulations. (c) and (d) are maps of $\{\mathrm{WD}_S(L_i, l_j)\}_{i=1,\dots,I;j=1,\dots,J}$ for the annual temperature emulations.

values of 1.013 and 1.085, respectively. Our emulations can successfully capture the variability of simulations at most grid locations, including polar regions. A high-$\mathrm{I_{uq}}$ region near the location $(50.00, 300.00)$ is caused by the poor evaluation of the mean trend, which is discussed in Section S4.1.1 of the supplementary materials. Figure 5(b) shows a poor spatial continuity for HCBG-SG, which may be caused by the independent evaluation of parameters at each latitude and a relatively weak model of dependence across latitudes. A similar pattern can be seen in Figure 5(d), which depicts the map of WD_S for HCBG-SG, with a median value of 0.076. The WD_S values for SHT-SG in Figure 5(c), which are closer to zero with a median value of 0.062, provide further evidence of the similarity between our emulations and the simulations. From Figure 5, geographic factors have no obvious impact on both two SGs regarding $\mathrm{I_{uq}}$ and WD_S values. The WD_T values for SHT-SG and HCBG-SG are comparable, with medians 0.126 and 0.108, respectively. In the supplementary materials, we further compare other basic statistical characteristics of our emulations to those of the simulations by replicating the procedures outlined in Figure 1.
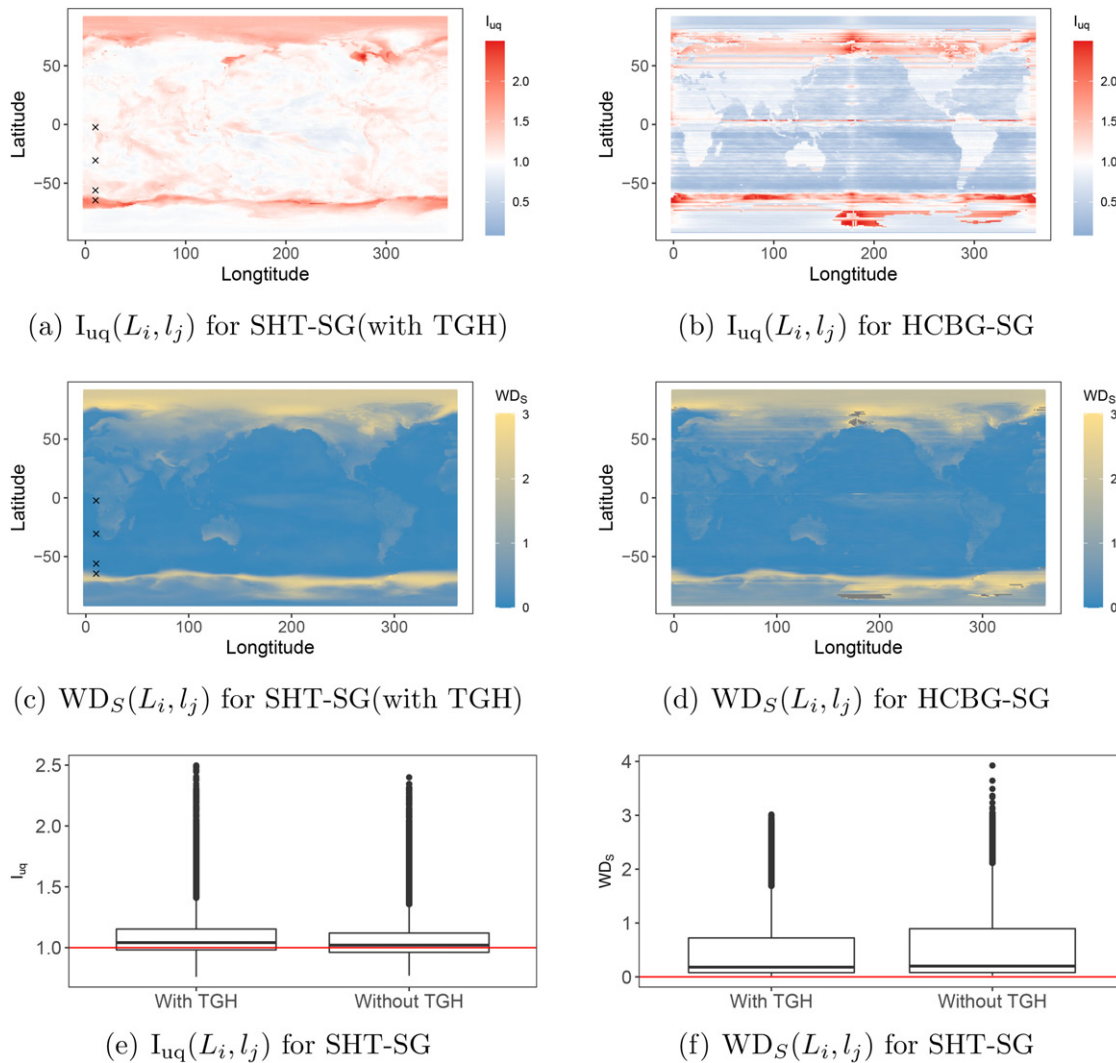
4.2. Daily Aggregated Temperature

Now, we develop an SG for the daily aggregated temperature simulations with $T = 31,390$. Analyzing such a vast amount of data presents a computational challenge. Therefore, we adopt a strategy similar to Huang et al. (2023) to perform inference only on data for the years 2020, 2040, 2060, 2080, and 2100. The procedure of inference closely follows that of the monthly case. Therefore, we put the details of development to Section S4.3 of the supplementary materials and focus on assessing the

performance of the emulations and a further exploration of data characteristics.

For the daily data, we choose $K_D = 4$, $Q_l = 36$, $Q_o = 68$, and $P = 1$. The choice of $K_D > K_M = 3$ reveals the presence of more complex fluctuations in daily time series, requiring harmonic terms with higher frequencies in the model. The more structured and higher $(\hat{\phi}_1)_q^m$ in Figure S14(f) indicates a stronger dependence among time points. The performance of our SG, both with and without the use of TGH (referred to as SHT-SG(with TGH) and SHT-SG(without TGH), respectively), along with HCBG-SG, is illustrated in Figure 6. For better demonstration, indices larger than maximum values in Figure 6(a) and (c) are excluded from Figure 6(b) and (d), respectively. Comparing emulations generated by HCBG-SG with those of SHT-SG(with TGH), the latter illustrates a closer resemblance to simulations, greater robustness in numerically unstable regions, and better spatial continuity. From Figure 6(f), the incorporation of TGH in our SG results in an improvement in WD_S. The empirical distribution of emulations from SHT-SG(with TGH) is closer to that of simulations, aligning with the correct assumption. This advantage may not be apparent in Figure 6(e), as $\mathrm{I_{uq}}$ is used mainly for assessing the variabilities of emulations.

As in the monthly case, the daily emulations exhibit challenges in accurately representing certain regions, notably the North Pole and the Band regions. These regions display larger $\mathrm{I_{uq}}$ values, indicating higher uncertainty. Additionally, they may deviate from the behavior observed in the simulations, as reflected by the larger WD_S values. Moreover, WD_S values are generally higher over land compared to ocean areas, which may be attributed to the complex spatial and temporal structures and larger variation of the temperature on land.

We further explore these by intuitively displaying and comparing simulations and emulations at four testing grid points

(a) $I_{uq}(L_i, l_j)$ for SHT-SG(with TGH)

(b) $I_{uq}(L_i, l_j)$ for HCBG-SG

(c) $WD_S(L_i, l_j)$ for SHT-SG(with TGH)

(d) $WD_S(L_i, l_j)$ for HCBG-SG

(e) $I_{uq}(L_i, l_j)$ for SHT-SG

(f) $WD_S(L_i, l_j)$ for SHT-SG

Figure 6. Performance assessment of $R' = 7$ daily emulations generated by various methods. (a) and (b) are maps of $\{I_{uq}(L_i, l_j)\}_{i=1,\ldots,I;j=1,\ldots,J}$ for the SHT-SG(with TGH) and HCBG-SG, respectively. (c) and (d) are maps of $\{WD_S(L_i, l_j)\}_{i=1,\ldots,I;j=1,\ldots,J}$ for SHT-SG(with TGH) and HCBG-SG, respectively. Points $\times$ represent four test-ing grid points TGL=$(-2.36, 10.00)$, TGO=$(-30.63, 10.00)$, TGN=$(-56.07, 10.00)$, and TGB=$(-64.55, 10.00)$. (e) and (f) are boxplots of $\{I_{uq}(L_i, l_j)\}_{i=1,\ldots,I;j=1,\ldots,J}$ and $\{WD_S(L_i, l_j)\}_{i=1,\ldots,I;j=1,\ldots,J}$ for comparing the SHT-SG(with TGH) and SHT-SG(without TGH).

selected from the southern hemisphere with the same longitude in Figure 7. TGL is a land point near the equator. In Figure 7(a), temperature simulations on TGL exhibit larger variability but can be well replicated by the emulations. TGO is an ocean point in the middle latitudes. From Figure 7(b), temperature simu-lations on TGO have smaller variability overall, with relatively larger variability at peaks and valleys of each year. These patterns can also be observed at emulations on TGO. TGN is also an ocean point in the middle latitudes, but very close to the Band region. In Figure 7(c) (and Figure S14(a)), some temperature simulations exhibit sudden drops at a few days in year 2020, which leads to an overestimate of $\hat{\sigma}$(TGN), and hence a rel-atively higher variability in the emulations. Despite this, with I_{uq}(TGN) = 1.200 and WD_S(TGN) = 0.078, the abnormal drops do not significantly affect these indices. Both I_{uq} and WD_S exhibit a certain tolerance to isolated "outliers". TGB is an ocean point within the Band region. Compared with the scenarios in Figure 7(c), temperature simulations at TGB display much more rapid and significant fluctuations at lots of time slots in years 2020, 2040, and 2060. In contrast, in years 2080 and 2100, the

temperature simulations behave as other grid points on ocean with small variability. As shown in Figure S14(b), simulations at the North Pole region exhibit similar patterns, that is, surface temperature experiences sharp fluctuations on numerous days. According to the explanation kindly provided by NCAR, these temperature fluctuations are caused by the way of data genera-tion. Specifically, the surface temperature in simulations is the spatial average of the surface temperature of whatever medium is in the grid box, whether it is water, land, or sea ice. Grid points at Band and North Pole regions undergo a dynamic transition, shifting back and forth between being partial sea ice and being open ocean. Consequently, sometimes we get the temperature of sea ice, and sometimes we get the temperature of the top layer of the ocean (or a mixture of the two). Combining this with Figure S14, the numerical instabilities exhibit four features: they do (a) not occur in each ensemble; (b) not happen during a fixed period; (c) not vary within a similar extent; and 4) not exhibit a consistent pattern across different grid points. There-fore, although the high I_{uq} and WD_S values at the North Pole and Band regions directly stem from the inadequate assumption
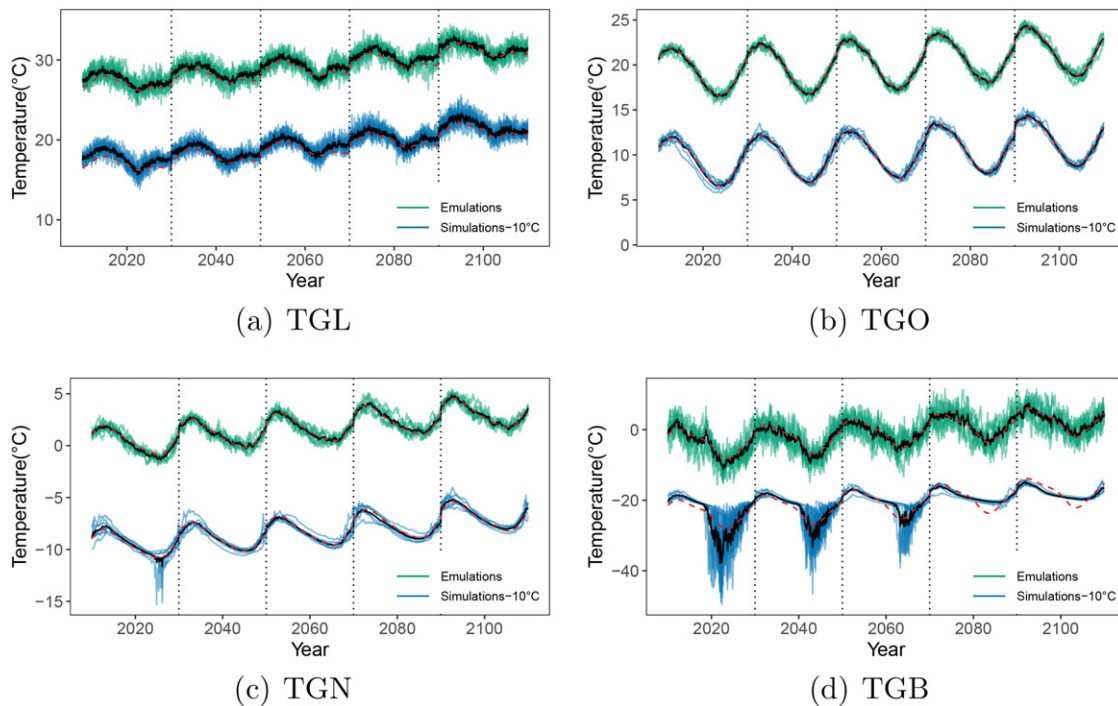
Figure 7. Comparison of daily emulations and simulations at four testing grid points. The black solid curves are ensemble means for simulations and emulations. The red dashed curves are $\hat{m}_t$s.

of a fixed $\hat{\sigma}$ over time, developing an SG that can accurately model the temperature fluctuations in these two regions is a challenge.

5. Conclusion and Future Work

In this article, we presented an efficient SG for global temperature simulations from the newly published CESM2-LENS2. With at most $O(IJ + Q_o^3)$ parameters to be stored, an unlimited number of emulations of size $O(IJT)$, even at a daily scale, can be generated to help with the investigation of climate internal variability. Such a saving of computational time and resources comes from the use of SHT, which expands data with $Q_o^2 \ll IJ$ spherical harmonics and represents a practical low-rank approximation on the sphere. By customizing Q values for grid points on land and ocean and leveraging the axial symmetry, the proposed SG can properly capture the complex nonstationary dependencies among different latitudes and land/ocean regions. To account for the non-Gaussian nature of the high-resolution time series, we introduced a modified TGH transformation into the SG. In our case study, we developed SGs based on $R = 7$ annually, monthly and daily aggregated simulations, respectively. We evaluated the efficiency and accuracy of the proposed SG by comparing the generated emulations to the original simulations and emulations generated by Huang et al. (2023) using various indexes and visual inspections.

While the proposed SGs have demonstrated commendable performance, there remains room for further enhancements: (a) Including altitude factors in both the deterministic and stochastic components. The influence of altitude has been evident in the inference process and the performance of emulations. Surface temperatures are sensitive to altitude (Castruccio and Genton 2016); (b) Improving the efficiency of SG on land, particularly for monthly and daily simulations. The larger variability observed in these cases suggests the need for more suitable models and additional factors to account for the complicated spatial and temporal structures on land; (c) Replacing the constant standard error with a temporal-varying one, denoted as σ_t. For simulations with higher temporal resolution, the assumption of a constant σ leads to an overestimate of standard error, further amplifying the variability of emulations, especially on land; and (d) Extending the current work to include hourly simulations. Hourly data may exhibit more intriguing characteristics, and exploring their emulation presents an opportunity for further investigation and understanding.

Beyond surface temperature, the idea of constructing an SG with SHT can be extended to various climate variables such as wind speed and atmospheric gas concentrations. After the removal of the deterministic component, the stochastic component can be efficiently expanded using spherical harmonics. The model in the spectral domain is contingent on the specific nature of the variable, which may ask for including other influential factors. Moreover, for variables that may exhibit interdependence, such as temperature and precipitation, exploring methods for their joint modeling in the spectral domain is also of interest. Investigating regional climate change, especially on land, is crucial for several reasons. On the one hand, climate change does not affect all regions equally. On the other hand, different regions have varying levels of vulnerability to climate change, for example, extreme weather events, sea-level rise, or precipitation patterns. Local communities, governments, and businesses may be interested in how specific areas are being impacted so that they can adapt and plan for the future. Therefore, another possible extension is to build an SG for high-resolution climate simulations on a specific region.

Supplementary Materials

Additional information about the CESM2-LENS2 data, inference process, derivation, validation, and results for the case studies are contained in the online supplementary materials. (.pdf file)

The R code and instructions for downloading, processing the data, and reproducing the results in the article are available at the GitHub repository: https://github.com/SpatialTemporalStats/LENS2_Emulator_Reproducibility_Materials

Disclosure Statement

The authors report there are no competing interests to declare.

References

Baker, A. H., Xu, H., Hammerling, D. M., Li, S., and Clyne, J. P. (2017), "Toward a Multi-Method Approach: Lossy Data Compression for Climate Simulation Data," in *International Conference on High Performance Computing*, pp. 30–42, Springer. [2494]

Banerjee, S., Gelfand, A. E., Finley, A. O., and Sang, H. (2008), "Gaussian Predictive Process Models for Large Spatial Data Sets," *Journal of the Royal Statistical Society*, Series B, 70, 825–848. [2494]

Bolin, D., Simas, A. B., and Xiong, Z. (2023), "Covariance-based Rational Approximations of Fractional SPDEs for Computationally Efficient Bayesian Inference," *Journal of Computational and Graphical Statistics*, 33, 64–74. [2494]

Branstator, G., and Teng, H. (2010), "Two Limits of Initial-Value Decadal Predictability in a CGCM," *Journal of Climate*, 23, 6292–6311. [2496]

Castruccio, S., and Genton, M. G. (2016), "Compressing an Ensemble with Statistical Models: An Algorithm for Global 3D Spatio-Temporal Temperature," *Technometrics*, 58, 319–328. [2505]

——— (2018), "Principles for Statistical Inference on Big Spatio-Temporal Data from Climate Models," *Statistics & Probability Letters*, 136, 92–96. [2496,2497]

Castruccio, S., and Guinness, J. (2017), "An Evolutionary Spectrum Approach to Incorporate Large-Scale Geographical Descriptors on Global Processes," *Journal of the Royal Statistical Society*, Series C, 66, 329–344. [2494]

Castruccio, S., Hu, Z., Sanderson, B., Karspeck, A., and Hammerling, D. (2019), "Reproducing Internal Variability with Few Ensemble Runs," *Journal of Climate*, 32, 8511–8522. [2494]

Castruccio, S., McInerney, D. J., Stein, M. L., Crouch, F. L., Jacob, R. L., and Moyer, E. J. (2014), "Statistical Emulation of Climate Model Projections based on Precomputed GCM Runs," *Journal of Climate*, 27, 1829–1844. [2494,2497,2500]

Castruccio, S., and Stein, M. L. (2013), "Global Space-Time Models for Climate Ensembles," *The Annals of Applied Statistics*, 7, 1593–1611. [2494]

Chang, W., Haran, M., Olson, R., and Keller, K. (2014), "Fast Dimension-Reduced Climate Model Calibration and the Effect of Data Aggregation," *The Annals of Applied Statistics*, 8, 649–673. [2494]

Cressie, N., and Johannesson, G. (2008), "Fixed Rank Kriging for Very Large Spatial Data Sets," *Journal of the Royal Statistical Society*, Series B, 70, 209–226. [2494]

Datta, A., Banerjee, S., Finley, A. O., and Gelfand, A. E. (2016), "Hierarchical Nearest-Neighbor Gaussian Process Models for Large Geostatistical Datasets," *Journal of the American Statistical Association*, 111, 800–812. [2494]

Dong, C., Peings, Y., and Magnusdottir, G. (2023), "Regulation of Southwestern United States precipitation by non-ENSO Teleconnections and the Impact of the Background Flow," *Journal of Climate*, 36, 7415–7433. [2495]

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016), "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization," *Geoscientific Model Development*, 9, 1937–1958. [2493]

Furrer, R., and Sain, S. R. (2010), "spam: A Sparse Matrix R Package with Emphasis on MCMC Methods for Gaussian Markov Random Fields," *Journal of Statistical Software*, 36, 1–25. [2502]

Guinness, J. (2021), "Gaussian Process Learning via Fisher Scoring of Vecchia's Approximation," *Statistics and Computing*, 31, 25. [2494]

Guinness, J., and Hammerling, D. (2018), "Compression and Conditional Emulation of Climate Model Output," *Journal of the American Statistical Association*, 113, 56–67. [2494]

Hu, W., and Castruccio, S. (2021), "Approximating the Internal Variability of Bias-Corrected Global Temperature Projections with Spatial Stochastic Generators," *Journal of Climate*, 34, 8409–8418. [2494]

Huang, H., Castruccio, S., Baker, A. H., and Genton, M. G. (2023), "Saving Storage in Climate Ensembles: A Model-based Sstochastic Approach" (with Discussion), *Journal of Agricultural, Biological and Environmental Statistics*, 28, 324–344. [2493,2495,2496,2497,2498,2500,2503,2505]

Jeong, J., Castruccio, S., Crippa, P., and Genton, M. G. (2018), "Reducing Storage of Global Wind Ensembles with Stochastic Generators," *The Annals of Applied Statistics*, 12, 490–509. [2494,2495,2498]

Jeong, J., Jun, M., and Genton, M. (2017), "Spherical Process Models for Global Spatial Statistics," *Statistical Science*, 32, 501–513. [2494,2497]

Jeong, J., Yan, Y., Castruccio, S., and Genton, M. G. (2019), "A Stochastic Generator of Global Monthly Wind Energy with Tukey g-and-h Autoregressive Processes," *Statistica Sinica*, 29, 1105–1126. [2494,2495,2496]

Jones, R. H. (1963), "Stochastic Processes on a Sphere," *The Annals of Mathematical Statistics*, 34, 213–218. [2494,2497,2498]

Jun, M., and Stein, M. (2008), "Nonstationary Covariance Models for Global Data," *The Annals of Applied Statistics*, 2, 1271–1289. [2494]

Kaufman, C. G., Schervish, M. J., and Nychka, D. W. (2008), "Covariance Tapering for Likelihood-based Estimation in Large Spatial Data Sets," *Journal of the American Statistical Association*, 103, 1545–1555. [2494]

Kay, J. E., DeRepentigny, P., Holland, M. M., Bailey, D. A., DuVivier, A. K., Blanchard-Wrigglesworth, E., Deser, C., Jahn, A., Singh, H., Smith, M. M., et al. (2022), "Less Surface Sea Ice Melt in the CESM2 Improves Arctic Sea Ice Simulation with Minimal Non-Polar Climate Impacts," *Journal of Advances in Modeling Earth Systems*, 14, e2021MS002679. [2495]

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S., Danabasoglu, G., Edwards, J., et al. (2015), "The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability," *Bulletin of the American Meteorological Society*, 96, 1333–1349. [2493,2500]

Kennedy, M. C., and O'Hagan, A. (2001), "Bayesian Calibration of Computer Models," *Journal of the Royal Statistical Society*, Series B, 63, 425–464. [2494]

López-Pintado, S., and Romo, J. (2009), "On the Concept of Depth for Functional Data," *Journal of the American Statistical Association*, 104, 718–734. [2500]

Lorenz, E. N. (1963), "Deterministic Nonperiodic Flow," *Journal of Atmospheric Sciences*, 20, 130–141. [2496]

Muñoz, S. E., Dee, S. G., Luo, X., Haider, M. R., O'Donnell, M., Parazin, B., and Remo, J. W. F. (2023), "Mississippi River Low-Flows: Context, Causes, and Future Projections," *Environmental Research: Climate*, 2, 031001. [2495]

National Oceanic and Atmospheric Administration (NOAA) (2022), "NOAA Celebrates 200 Years of Science, Service, and Stewardship," Revised October 24, 2022. [2493]

Nychka, D., Hammerling, D., Sain, S., and Lenssen, N. (2016), "LatticeKrig: Multiresolution Kriging Based on Markov Random Fields," R package version 8.4. [2498]

Oakley, J. E., and O'Hagan, A. (2002), "Bayesian Inference for the Uncertainty Distribution of Computer Model Outputs," *Biometrika*, 89, 769–784. [2494]

——— (2004), "Probabilistic Sensitivity Analysis of Complex Models: A Bayesian Approach," *Journal of the Royal Statistical Society*, Series B, 66, 751–769. [2494]

O'Neill, B., Kriegler, E., Riahi, K., Ebi, K., Hallegatte, S., Carter, T., Mathur, R., and Vuuren, D. (2014), "A New Scenario Framework for Climate

Change Research: The Concept of Shared Socioeconomic Pathways," *Climatic Change*, 122, 401–414. [2495]

Riahi, K., van Vuuren, D. P., Kriegler, E., et al. (2017), "The Shared Socioeconomic Pathways and their Energy, Land Use, and Greenhouse Gas Emissions Implications: An Overview," *Global Environmental Change*, 42, 153–168. [2495]

Rodgers, K. B., Lee, S.-S., Rosenbloom, N., et al. (2021), "Ubiquity of human-induced changes in climate variability," *Earth System Dynamics*, 12, 1393–1411. [2495]

Rue, H., Martino, S., and Chopin, N. (2009), "Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations," *Journal of the Royal Statistical Society*, Series B, 71, 319–392. [2494]

Sacks, J., Schiller, S., and Welch, W. (1989), "Designs for Computer Experiments," *Technometrics*, 31, 41–47. [2494]

Santambrogio, F. (2015), *Optimal Transport for Applied Mathematicians* (Vol. 55), pp. 58–63, New York: Birkäuser, 55. [2501]

Stein, M. L. (2007), "Spatial Variation of Total Column Ozone on a Global Scale," *The Annals of Applied Statistics*, 1, 191–210. [2494,2497,2498]

——— (2008), "A Modeling Approach for Large Spatial Datasets," *Journal of the Korean Statistical Society*, 37, 3–10. [2494,2497]

Sun, Y., and Genton, M. G. (2011), "Functional Boxplots," *Journal of Computational and Graphical Statistics*, 20, 316–334. [2500]

Tagle, F., Genton, M. G., Yip, A., Mostamandi, S., Stenchikov, G., and Castruccio, S. (2020), "A High-Resolution Bilevel Skew-t Stochastic Generator for Assessing Saudi Arabia's Wind Energy Resources" (with Discussion), *Environmetrics*, 31, e2628. [2494]

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012), "An Overview of CMIP5 and the Experiment Design," *Bulletin of the American Meteorological Society*, 93, 485–498. [2493]

Underwood, R., Bessac, J., Di, S., and Cappello, F. (2022), "Understanding the Effects of Modern Compressors on the Community Earth Science Model," in *2022 IEEE/ACM 8th International Workshop on Data Analysis and Reduction for Big Scientific Data (DRBSD)*, IEEE, pp. 1–10. [2494]

Vecchia, A. V. (1988), "Estimation and Model Identification for Continuous Spatial Processes," *Journal of the Royal Statistical Society*, Series B, 50, 297–312. [2494]

Yan, Y., and Genton, M. G. (2019), "Non-Gaussian Autoregressive Processes with Tukey g-and-h Transformations," *Environmetrics*, 30, e2503. [2494]