The data set

We focus on the Community Climate System Model version (CCSM4) model (Gent et al., 2011) from the CMIP5 ensemble (Taylor et al., 2012). The annual temperature at surface is on a reference height of 2 meters above ground level, and is resolved on a regular spatial grid over the sphere with $M = 142$ latitudinal bands and $N = 288$ longitudinal bands. We focus on the RCP4.5 scenario (Van Vuuren et al., 2011), which formulates a moderate increase of Carbon Dioxide and other greenhouse gases from 2006 to 2100, i.e. $K = 95$ years. We consider two realizations, e.g. runs with different initial conditions, so the total data set consists of $142 \times 288 \times 95 \times 2 \approx 7.7$ million points. Here $L_m, m = 1, \ldots, M$ will denote the latitude, $\ell_n, n = 1, \ldots, N$ the longitude, and $t_k = 1, \ldots, K$ the time.

Denote by $\mathbf{T}_r$ the vector of space-time temperature at surface for realization r. We operate under the assumption that the two runs are independent, conditional to the climate mean, i.e.

$$\mathbf{T}_r = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_r \tag{S1}$$

where $\boldsymbol{\varepsilon}_r \overset{\text{iid}}{\sim} \mathcal{F}$ for some distribution $\mathcal{F}$, for which we assume the first and second moments exist. The assumption of independence of $\boldsymbol{\varepsilon}_r$ is associated with the deterministic chaotic nature of the primitive equations of the climate model (Lorenz, 1963; Collins and Allen, 2002; Collins, 2002; Branstator and Teng, 2010), and it has been proven valid on multiple instances for rapidly mixing atmospheric processes such as temperature at surface (Castruccio and Stein, 2013; Castruccio and Guinness, 2017).

Local behavior

To study the local behavior, we consider the normalized difference between the climate model runs, i.e. $\mathbf{e} = \frac{\mathbf{T}_1 - \mathbf{T}_2}{\sqrt{2}}$ so that the process has mean zero according to (S1), and the local behavior can be studied with no confounding factor from the climate.

There is a wide availability of diagnostics tool for understanding high frequency spatial behavior, and to detect its change across space. In Figure S1(a) and (b) we consider a single latitudinal band at approximately $L = 41°$ north, and we consider the set of data comprising of only land or ocean points:

$$\mathbf{e}_m^j(n,k) = \{\mathbf{e}(L_m, \ell_n, t_k) I(\ell_n \in \text{j})\},$$

where $j \in \{\text{land, ocean}\}$. In Figure S1(a) we plot the log estimated spectra (i.e. the periodogram) averaged over time for this quantity:

$$|\hat{f}_m^j(c)|^2 = \frac{1}{K}\sum_{k=1}^{K}\left|\sum_{n=0}^{N}\mathbf{e}_m^j(n,k)e^{-i\ell_n c}\right|^2, \quad c = 0,\ldots,N/2,$$

where $j \in \{\text{land}, \text{ocean}\}$. The land periodogram is considerably more flat than the one for ocean, hence highlighting how its corresponding spatial behavior is less regular than over land. An equivalent diagnostic of local behavior can be obtained by considering the variogram, i.e.

$$\hat{\gamma}_m^j(n) = \frac{1}{B}\sum_{|n_1-n_2|=n}\{\mathbf{e}_m^j(n_1,k) - \mathbf{e}_m^j(n_2,k)\}^2,$$

where B is the cardinality of the set. The variogram shows a similar behavior for large lags, which corresponds to low frequencies in the periodogram. To detail the local dependence across all latitudes and understand the patterns of nonstationarity, the contrasts can be reported in a map without the averaging across longitude that is performed in the variogram. In other words, we compute

$$\Delta_{\text{ew};m,n} = \frac{1}{K}\sum_{k=1}^{K}\{\mathbf{e}_m(n,k) - \mathbf{e}_m(n+1,k)\}^2 \tag{S2}$$

and plot this as a function of longitude n and latitude m, the results are in Figure S1(c). The plot highlights how land and ocean have different structure, so a model with a changing local behaviour across these two domains is necessary (Castruccio and Guinness, 2017). Also, high mountain regions display a changing behaviour as well, as highlighed by the high value of the contrasts in regions such as the Himalayas. A model that allows for a changing spatial smoothness with altitude is thus preferable (Jeong et al., 2018).

References

Branstator, G. and Teng, H. (2010) Two limits of initial-value decadal predictability in a CGCM. *Journal of Climate*, **23**, 6292–6311.

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.

Castruccio, S. and Stein, M. L. (2013) Global space-time models for climate ensembles. *Annals of Applied Statistics*, **7**, 1593–1611.

Collins, M. (2002) Climate predictability on interannual to decadal time scales: the initial value problem. *Climate Dynamics*, **19**, 671–692.

Collins, M. and Allen, M. R. (2002) Assessing the relative roles of initial and boundary conditions in interannual to decadal climate predictability. *Journal of Climate*, **15**, 3104–3109.

Gent, P. R. et al. (2011) The community climate system model version 4. *J. Climate*, **24**, 4973–4991.

Jeong, J., Castruccio, S., Crippa, P. and Genton, M. G. (2018) Reducing storage of global wind ensembles with stochastic generators. *Annals of Applied Statistics*. In press.

Lorenz, E. (1963) Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, **20**, 130–141.

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

Van Vuuren, D. et al. (2011) The representative concentration pathways: an overview. *Climatic Change*, **109**, 5–31.

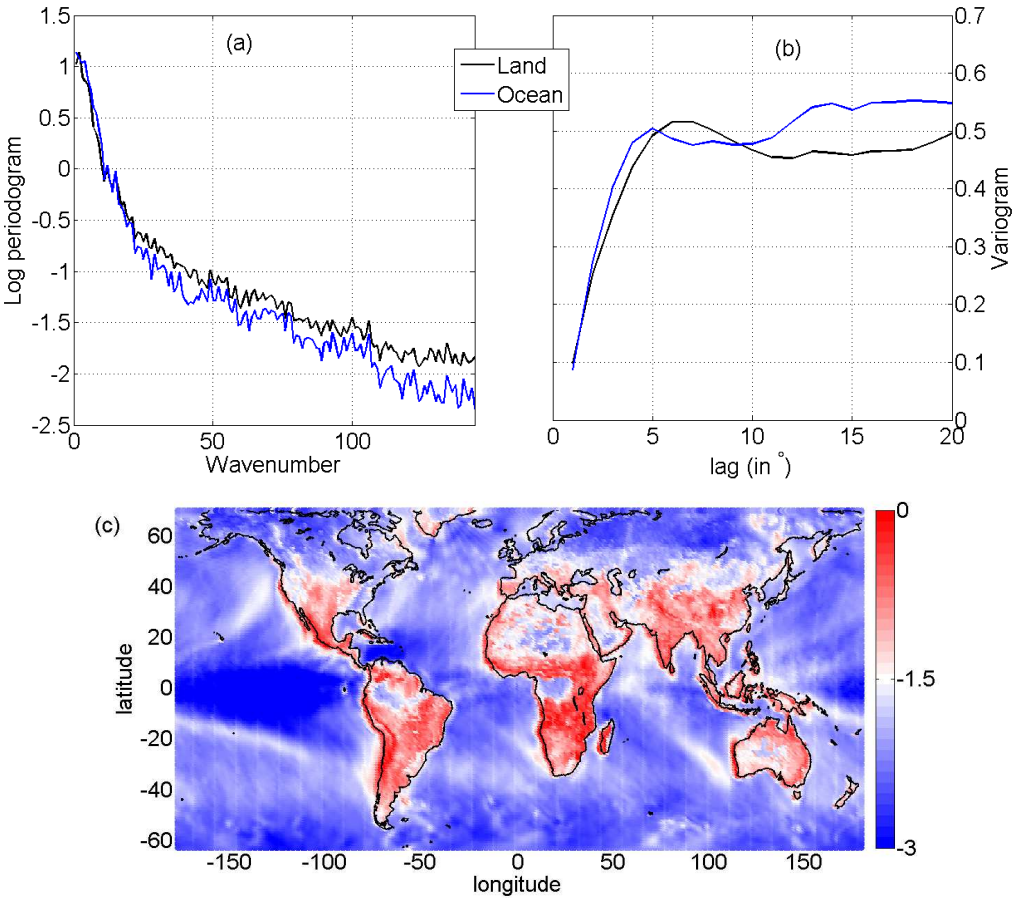

Figure S1: Static diagnostic tools for studying spatial behavior. (a) Average log periodogram (in time) at approximately 41 degrees north for all points in land (black) and over ocean (blue). (b) Average variogram (in time) at the same latitude as (a), with the same conventions for land and ocean. (c) A spatial map of the local contrasts as defined in (S2).