https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

Climate models lack decadal variability in precipitation because of a weak atmospheric response to sea surface temperature anomalies

Robin Noyelle[1], Robert Jnglin Wills[1], Erich Fischer[1], Heini Wernli[1], and Reto Knutti[1]
[1]Institute for Atmospheric and Climate Science, ETH Zürich, Zürich, Switzerland
Correspondence: Robin Noyelle (robin.noyelle@env.ethz.ch)

Abstract.

A reliable representation of the mechanisms leading to decadal variability in climate models is crucial for projections of multi-year droughts and other long-lasting extremes. Here we show that CMIP6 climate models have too low autocorrelation at one-year time lag and beyond, and as a result too weak decadal variability in both precipitation and sea level pressure (SLP). We employ a statistical learning technique to systematically quantify the extent to which SSTs condition the anomalies in these surface variables across the globe. In the extratropics, SSTs explain only a small portion of the variability of precipitation and SLP in models, even at multi-year time scales. ERA5 and 20CRv3 reanalyses suggest distinctly higher autocorrelation and decadal variability because of a stronger conditioning by SSTs in the real world. We attribute the weak decadal variability of precipitation and SLP in models to a weak dynamical atmospheric response to decadal variability of SSTs. Our results suggest that the models considered underestimate the dynamical ocean-atmosphere coupling that is important for decadal variability in extratropical precipitation and SLP, especially outside the regions impacted by the sub-decadal tropical Pacific ocean variability.

1 Introduction

Long-lasting climate extremes present major challenges to societies. Examples include megadroughts — exceptional multi-year dry events (Cook et al., 2022), such as the Dust Bowl in the 1930s in the US (Schubert et al., 2004), the Sahel droughts in the 1980s (Folland et al., 1986) or the ongoing megadrought in Chile since 2010 (Garreaud et al., 2020). The occurrence of these events is primarily a manifestation of the decadal and multi-decadal variability of the climate system (Ghil, 2020), which has been well-documented by paleo-proxy studies (Cook et al., 2022; Laepple et al., 2023). In addition to long-term trends from natural and anthropogenic forcings, regional and local climates can thus exhibit large and persistent anomalies at multi-year time scales. Representing these correctly in climate models is essential to make reliable predictions for the next decades (O'Reilly et al., 2021).

Although climate models reproduce the decadal and multi-decadal variability at large scales, such as for the global mean surface temperature, they underestimate the variability at regional scale (Laepple and Huybers, 2014; Dee et al., 2017; Kravtsov, 2017; Cheung et al., 2017; Laepple et al., 2023). For example, the North Atlantic Oscillation (NAO) behaves mainly as white noise in models (Simpson et al., 2018; Eade et al., 2022), and therefore its decadal and multi-decadal variability is underestimated (Woollings et al., 2014; Kim et al., 2018; O'Reilly et al., 2019; Schurer et al., 2023; Bonnet et al., 2024). Any

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

such underestimation of the variability has important consequences on projections for future regional climates (Gastineau and Frankignoul, 2015; Simpson et al., 2019; Al-Yaari et al., 2026). Given the intrinsic short time scales of the atmospheric circulation, it is expected that its decadal variability beyond random sampling is forced by the decadal variability of the oceans (Barsugli and Battisti, 1998; Meehl et al., 2009), and this indicates that atmosphere-ocean interactions may play an important role in this underestimated variability.

The atmosphere and oceans form a coupled system which presents distinct modes of variability from those that would arise in either of the two separately (Hirst, 1986). In general the tropical oceans have the strongest impacts on the atmosphere at a global scale (Lau and Nath, 1994; Timmermann et al., 2018). The decadal and multi-decadal variability in the extratropical oceans has been classically conceptualized as the integration of random, short-term variability in the atmosphere (Hasselmann, 1976; Frankignoul and Hasselmann, 1977). The impact of SST anomalies on the dynamics of the atmosphere at large scales was thus thought to be weak at sub-annual time scales (Kushnir et al., 2002), although increasing at longer time scales (Hoskins and Karoly, 1981; Frankignoul, 1985; Battisti et al., 2019) — a "small oceanic forcing" regime (Czaja et al., 2019). However, isolating the atmospheric response to SST anomalies is challenging both in models and observations. Atmosphere-only simulations show that the response to SST anomalies is sensitive to the base state of the atmosphere, even for similar model configurations (Thomson and Vallis, 2018b, a). Coupled and uncoupled models show a very different response to anomalous SSTs, due to nonlinear feedbacks of atmospheric eddies on SSTs (Liu and Wu, 2004). In models, there is for example no consistent feedback of the Atlantic Multi-decadal Variability (AMV) onto the atmospheric circulation (Peings et al., 2016) and its impact on the European climate is too weak (Qasmi et al., 2017) compared to the impacts inferred from observations (Davini et al., 2015; O'Reilly et al., 2017; Simpson et al., 2019). Models have been shown to have limited prediction skills at decadal time scales due to a poor representation of ocean-atmosphere feedbacks, especially in the extratropics (Kolstad, 2026). The decadal predictability is thus also underestimated, a potential consequence of the signal-to-noise paradox (Scaife and Smith, 2018; Smith et al., 2020). Even in the tropical Pacific, the response of the Walker circulation to ENSO-related SST anomalies in models is underestimated (Wang et al., 2021).

Thus, it remains elusive whether fully-coupled climate models correctly represent the link between SST anomalies and atmospheric variables at decadal time scales. In this study, we examine the decadal variability of three atmospheric surface variables — sea level pressure (SLP), total precipitation and 2-m temperature (T2M) — in six large ensemble climate simulations from the Coupled Model Intercomparison Project Phase 6 (CMIP6). We address the question of whether these climate models correctly represent the decadal variability of these variables and their link with SSTs at multi-year time scales. In particular, we employ a statistical learning technique with a rigorous regularization procedure to quantify systematically and globally how much the variability in SST anomalies explains the variability of the surface variables from monthly to decadal time scales. We compare the results obtained for the CMIP6 models with those for two reanalyses and three observational datasets.

The paper is structured as follows. In Sect. 2 we present the reanalysis, observational datasets and climate model data used, the computation of the metrics to diagnose decadal variability and the ridge regression employed to quantify the link between SSTs and surface variables. In Sect. 3 we present the results obtained. The conclusions reached are presented and discussed in Sect. 4.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

2 Data and methods

2.1 Climate models, reanalyses and observational data sets

In this study we use six large ensembles of climate models over the historical period (1850-2014) from the CMIP6 ensemble: MPI-ESM1-2-LR (50 members), MIROC6 (50 members), IPSL-CM6A-LR (33 members), CanESM5 (40 members), ACCESS-ESM1-5 (40 members) and CESM2-LE (100 members). These models were chosen to sample the climate model families diversity (Knutti et al., 2013) as much as possible while keeping the analyses presented computationally feasible. All models are regridded on a regular 2.5° × 2.5° grid for all variables. For all models, we consider their monthly mean SST, 2-m temperature (T2M) and sea-level pressure (SLP) and their monthly sum total precipitation. Over the quasi-preindustrial period (1850-1900) we assume that the climate fluctuations are small enough (including the forcing by volcanic aerosols) for the climate to be mostly stationary and thus the output of models are not detrended (except for the estimation of spectral slopes, see below). Over the ERA5-like period (1930-2014), all members are detrended in a similar way as for the ERA5 reanalysis (see below).

In addition to these CMIP6 models, we also compare our results to 11 runs of the CESM2 model forced according to the CMIP6 protocol over the historical period and 11 additional AMIP-like runs of the same model for which the SSTs are taken from the 11 coupled runs. These latter runs thus present an atmospheric variability compatible with the evolution of the coupled runs while not having a feedback of the atmospheric variability onto the oceanic variability.

To compare our results to the real-world behavior of the climate system, we use two reanalyses: the ERA5 reanalysis (Soci et al., 2024) over the period 1940-2024 and the 20CRv3 reanalysis (Slivinski et al., 2019) over the period 1806-2015. The SST forcing of the ERA5 reanalysis comes from the Operational Sea Surface Temperature and Ice Analysis (OSTIA) produced by the UK Met Office from September 2007 onward (Good et al., 2020). SST data from the HadISST2.1.0.0 dataset produced by the Met Office Hadley Centre are used before September 2007 (Titchner and Rayner, 2014). For the 20CRv3 reanalysis, SST fields are prescribed from two eight-member ensembles: version 3 of the Simple Ocean Data Assimilation with sparse input (SODAsi.3), reanalysis for 1836–1980 and the Hadley Centre Sea Ice and Sea Surface Temperature dataset, version 2.2 (HadISST2.2), for 1981–2015. Both reanalyses present substantial temporal trends due to external natural and anthropogenic forcings. In order to correct for these trends, we regress at each grid point the monthly surface variables and SSTs on the global mean surface temperature (GMST) derived from each dataset's own T2M field, smoothed with a running mean of 30 years. In addition, we use three long-term observation products: HadSLP2 (Allan and Ansell, 2006) over the period 1900-2004 for SLP, gridded Global Precipitation Climatology Centre (GPCC) precipitation dataset (Schneider et al., 2016) over the period 1900-2020 for land precipitation and the ensemble mean of the DCENT-I dataset (Chan et al., 2026) over the period 1900-2025 for T2M. All observation products are detrended with respect to the GMST derived from DCENT-I over their respective time period. In the following, we refer to both reanalyses and observations as "observational datasets".

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

2.2 Autocorrelation and spectral slopes

We compute the one-year time lag autocorrelation using the Spearman, i.e. rank-based, correlation. We chose the Spearman autocorrelation to avoid being influenced by the non-Gaussian distribution of some surface variables, especially precipitation. For models over the quasi-industrial period we compute the autocorrelation by pooling all members together. For the reanalyses and for individual members of climate models over the ERA5-like period (1930-2014), all variables are detrended with respect to the GMST of each individual member before computing the autocorrelation (similar to ERA5 and 20CRv3).

The spectral slope β of a variable is defined as the scaling factor between its power spectral density $E(f)$ and the frequencies f: $E(f) \propto f^{-\beta}$. A variable that shows more variability on long rather than short time scales has $\beta > 0$. In contrast, a variable with no amplification of low-frequency variability has $\beta = 0$ (white noise). Power spectral densities are computed with the multi-tapering method. The spectral slopes are computed using linear regression after uniform logarithmic frequency binning (Huybers and Curry, 2006; Ault et al., 2013). We use a median regression rather than an ordinary least-square regression to decrease the influence of outliers data points in the slope estimation. We target the decadal time scale and thus compute spectral slopes over the range of frequencies covering time scales between 5 and 30 years. Choosing slightly different starting and ending time scales does not strongly influence the values of the spectral slopes (not shown). For this analysis only, the models over the quasi pre-industrial period are detrended by removing a linear trend with respect to time fitted on the ensemble mean, in order to remove spurious increase of the power spectral density at long time scales. Over the quasi pre-industrial period we compute the power spectral density for each member independently and then average over members to get an ensemble mean power density. The spectral slopes are then computed on this ensemble mean as explained above.

2.3 Ridge regression

In order to quantify the capacity of SSTs to explain the variance of surface variables, we adapt the method of Noyelle et al. (2026). We make two assumptions based on physical knowledge of the atmosphere typical time scales: (i) the evolution time scale of the atmosphere is short enough so that it reacts to the oceanic boundary conditions in less than one month, thus the conditional distribution of all surface variables only depends on the SSTs of the preceding month and not the months before ; (ii) the mixing time scale in the atmosphere is short enough so that the state of the atmosphere at month m does not influence the state of the atmosphere at month $m+1$ conditional on the oceanic boundary conditions. We use a one-month time lag for SSTs to avoid a potential atmospheric confounder that would influence both SSTs and surface variables and thus bias the estimation of the link between the two. Nonetheless, because of the slow time scales of the ocean, preliminary tests showed that using the SSTs of the same month gives similar results to those with a one-month time lag. Thus, we are confident that the results presented do not depend strongly on the two assumptions stated here. Under these two assumptions, there is a conditional probability distribution function of each surface variable X^m at month m at each grid point that depends only on the SSTs of the previous month: $\rho(X^m \mid \mathbf{SST}^{m-1})$, where $\mathbf{SST}^{m-1}$ is the *global* field of SSTs at a preceding month $m-1$. This conditional distribution quantifies how strongly SSTs condition the values of surface variables. We want to measure the

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.



average of this conditional distribution and assume that it is linear with respect to $\mathbf{SST}^{m-1}$:

$$\mathbb{E}(X^m \mid \mathbf{SST}^{m-1}) := \int X^m \rho(X^m \mid \mathbf{SST}^{m-1}) dX^m = X_0^m + \boldsymbol{\beta^m} \cdot \mathbf{SST}^{m-1} \quad (1)$$

where X_0^m is the climatological average of variable X at month m and $\boldsymbol{\beta^m}$ is a vector of coefficients to be estimated. The dot represents the usual scalar product.

In our climate data sets, the number of predictors $\mathbf{SST}^{m-1}$ is typically one order of magnitude larger than the number of predictands X^m. The latter indeed represents grid point time series over all ensemble members while the former represents all ocean grid points. Moreover, the predictors in $\mathbf{SST}^{m-1}$ are highly spatially correlated. To take into account these two limitations, we estimate the coefficients $\boldsymbol{\beta^m}$ via a linear regularized model with an L^2 penalty, namely ridge regression (Hastie et al., 2009):

$$\hat{\boldsymbol{\beta}}^m = \underset{\boldsymbol{\beta}^m}{\operatorname{argmin}} \left(\sum_{l=1}^{n} (X^{m,l} - X_0^m - \boldsymbol{\beta}^m \cdot \mathbf{SST}^{m-1,l})^2 + \lambda \|\boldsymbol{\beta}^m\|^2 \right) \quad (2)$$

where $l = 1, ..., n$ is the number of samples encompassing different years and ensemble members, $\mathbf{SST}^{m-1,l}$ is the global anomalous SSTs for sample l, $\lambda \geq 0$ is the regularization parameter and $\|\cdot\|^2$ is the Euclidean L^2 norm. SST anomalies are computed by removing the monthly climatological mean and dividing by the monthly climatological standard deviation. Regularization shrinks the elements of the vector $\boldsymbol{\beta}^m$, and the magnitude of the shrinkage depends on the hyperparameter λ. The value set for the regularization parameter λ is crucial and needs to be carefully evaluated. A value too high will shrink the coefficients too much and lead to artificially low predictive power. A value too small risks producing overfitted coefficients, with artificially high predictive power on training data but low extrapolation capability on unseen data. We determine an optimal regularization parameter independently for each grid point, climate model, surface variable and month by cross-validation. The cross-validation procedure is detailed in appendix A, with validation figures.

Importantly, the proportion of grid points where the model underfits is very low: always below 0.5% and even equal to 0 for most models and months (Fig. A1). This implies that *under the linear hypothesis* the ridge regression does not underestimate the proportion of variance explained by SSTs. On the contrary, the proportion of grid points where the model overfits (i.e. overestimates the proportion of variance explained) is higher. While it is typically below 5% for both SLP and T2M and between 5 to 10% for precipitation for all models and for 20CRv3, it is between 15 to 25% for precipitation in ERA5. This illustrates the need for a large amount of data to perform this regression analysis and the fact that the results obtained with ERA5 must be critically considered. Note that even if the proportion of grid points where the model overfits for precipitation is not negligible, those grid points are located at high latitudes (not shown) and the analysis suggests anyway that SSTs explain a very low percentage of the variance of precipitation variability there (Fig. 5).

We evaluate the quality of the ridge regression fit by measuring its capacity to generalize to unseen data. We compare the percentage of variance explained at yearly time scales between the training data and the unseen test data for models for all surface variables (Fig. A2 for SLP, Fig. A3 for precipitation, and Fig. A4 for T2M). Models generalize well, with similar R^2 scores for training and test data. The difference in corresponding R^2 scores are much higher for reanalyses than for the models and very geographically varying. In addition to the overfitting problem due to the limited amount of data, this also partly comes

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

from the fact that R^2 are evaluated over very few data points (8 years for ERA5, 20 years for 20CRv3 for test data). Thus, our results for the reanalyses are more problematic and rest on the hypothesis that the cross-validation procedure, which works well for models, can also be trusted for reanalyses (as evidenced by Fig. A1 for 20CRv3).

2.4 Spectral coherence

Once the coefficients $\hat{\boldsymbol{\beta}}^m$ are determined, we can generate the SST-driven anomalies of surface variables:

$$\tilde{X}^m_{SST} = \hat{\boldsymbol{\beta}}^m \cdot \mathbf{SST}^{m-1}. \tag{3}$$

We then evaluate a time scale-dependent metric of the explanatory power of SSTs on the surface variables using the spectral coherence between the SST-driven anomalies $\tilde{X}^m_{SST}$ and the actual anomalies $\tilde{X}^m := X^m - X^m_0$ of surface variables:

$$C(f) := \frac{|\, E_{\tilde{X}_{SST},\tilde{X}}(f)\,|^2}{E_{\tilde{X}_{SST}}(f) E_{\tilde{X}}(f)} \tag{4}$$

where $E_{\tilde{X}_{SST},\tilde{X}}(f)$ is the cross-spectral density between the SST-driven anomalies and the actual anomalies, and $E_{\tilde{X}_{SST}}(f)$ and $E_{\tilde{X}}(f)$ are the spectral densities of the SST-driven anomalies and the actual anomalies respectively. All spectral and cross-spectral densities are evaluated using the multi-tapering method similarly to spectral slopes.

The spectral coherence gives values between 0 and 1 that can be interpreted as a percentage of variance explained, similar to an R^2. However, an increase in the explained variance of 0.1 between two time scales has a different meaning if the initial explained variance was 0 (small gain in information) or 0.85 (large gain in information). We thus transform the spectral coherence using an analog of the change in mutual information from information theory (MacKay, 2003):

$$\Delta C(f_1 \to f_2) := K \log\left(\frac{1 - C(f_1)}{1 - C(f_2)}\right), \tag{5}$$

where we choose the proportionality constant K such that the gain of information of an increase of the spectral coherence from 0 to 0.1 is equal to 1, i.e. $K = -1/\log(0.9)$. Note that using this metric, when the spectral coherences are close to 1, a small change in the absolute spectral coherence can lead to a large change in ΔC.

3 Results

3.1 Climate models underestimate the persistence of precipitation and SLP at multi-year time scales

We begin by considering the 1-year lag autocorrelation as a simple metric to show that both SLP and precipitation have low persistence at multi-year time scales in models (Fig. 1). We aggregate the autocorrelations over two climatological zones, the tropics (30°S-30°N) and the extratropics (80°S-30°S and 30°N-80°N). Climate models have low or very low autocorrelation in precipitation, typically below 0.1 in absolute value and even below 0.05 over mid-latitude land (Fig. 1b, see also Fig. S2). SLP has slightly higher autocorrelation values, especially in the tropics (Fig. 1a and Fig. S1), but overall models consistently

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

estimate autocorrelation values very close to zero. For these two variables, models underestimate the autocorrelation compared to observational datasets, especially in the extratropics. The underestimation of the autocorrelation is a systematic and robust result across observational datasets and models (Fig. 1g,h). In contrast, the autocorrelation in T2M is comparable (Fig. 1c,f and Fig. S3), and the differences are more regional and model dependent (Fig. 1i). This is especially the case over oceans where T2M is a very good proxy for SSTs at monthly and longer time scales (Pearson correlation typically above 0.99, not shown).

Computing autocorrelation member-by-member independently after self-detrending over the period 1930-2014 (similarly to how we treat observational datasets, see Sect. 2) gives similar results (Fig. S4 and Fig. S6-S7). Compared to the autocorrelation obtained by grouping all members together as in Fig. 1, this even suggests that the autocorrelation estimated from the detrended observational datasets may be slightly underestimated (compare Fig. 1e and Fig. S4e for the simulations, for example), which could arise because of centennial-scale variability that is removed by detrending member-by-member. On a regional basis, the SLP autocorrelation is consistently underestimated by models over most of Africa, the Indian Ocean, Eastern Asia, Northern and Southern America (Fig. 1g and Fig. S5). The signal is more mixed for precipitation (Fig. 1h and Fig. S6), which also shows a global underestimation of the autocorrelation but with more regional variability. This is likely a result of the high variance of precipitation at yearly time scales, which makes the estimation of its autocorrelation less stable than for SLP. In contrast, for T2M there are regions of under- and overestimation of the autocorrelation that are mainly model-dependent and not always robust over observational datasets (Fig. 1i and Fig. S7).

The persistence of a time series, as measured by the autocorrelation at every time lag, is linked to its long-term variability via the Wiener-Khinchin theorem (Chatfield and Xing, 2019), which states that the Fourier transform of the (Pearson) autocorrelation function is equal to the power spectral density of the time series. Positive autocorrelation at a given time scale typically translates into positive spectral slopes. Positive spectral slopes indicate that variability increases at longer time scales. In contrast, a spectral slope close to zero indicates that the time series is behaving as white noise, i.e. as a process without memory, at these time scales. We therefore compute spectral slopes β for atmospheric surface variables for time scales from 5 to 30 years to measure decadal variability in the models and reanalyses. Similar to the results for autocorrelation, models underestimate the spectral slopes for SLP (Fig. 2a,d) and precipitation (Fig. 2b,e). For most models, precipitation is for example displaying spectral slopes close to zero in the mid-latitudes (Fig. 2e) while reanalyses suggest spectral slopes around 0.2-0.3 (see also Fig. S9 and Fig. S10 for the spatial distribution of spectral slopes). This stands in contrast with the behavior of T2M, which has similar spectral slopes in models and reanalyses over both land and ocean (Fig. 2c,f and Fig. S11). These results are robust when computing spectral slopes independently member-by-member over the transient period (Fig. S8). Similarly to the autocorrelation, the observational datasets give spectral slopes outside the 90% range of ensemble members for SLP (Fig. S12) while the results are more unstable for precipitation (Fig. S13), and both regionally and model dependent for T2M (Fig. S14).

3.2 SST anomalies explain little of the yearly variability of precipitation and SLP outside the tropics and the Pacific in models

Given the prominent 1-15-day time scale of internal atmospheric variability, it is physically expected that, apart from random sampling, most of the internal decadal variability in atmospheric surface variables is forced by the decadal variability of the

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

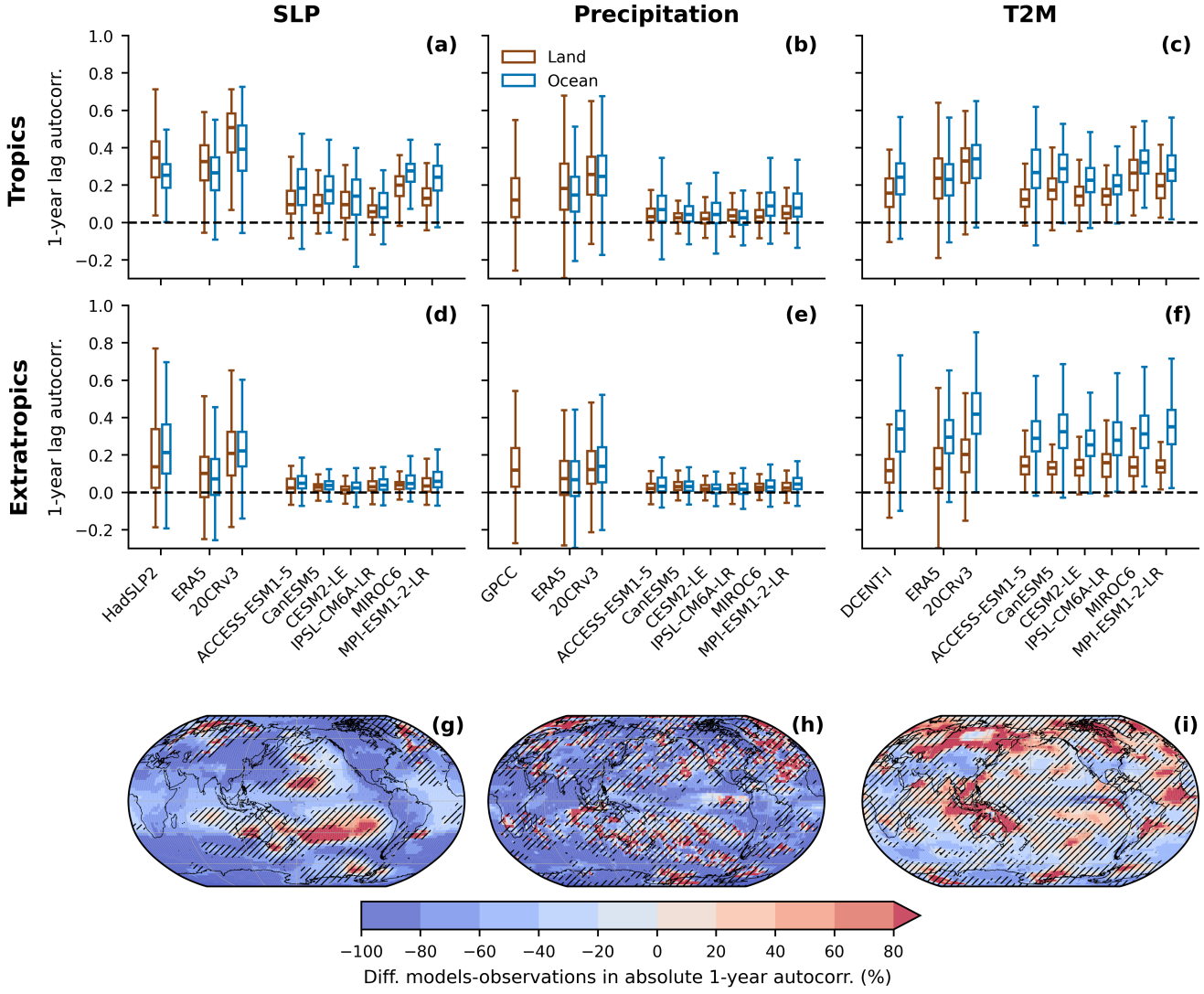


Figure 1. Autocorrelation of surface variables. Spearman 1-year lag autocorrelation for (**first column**) SLP, (**second column**) yearly accumulated precipitation and (**third column**) T2M. The boxplots show the distribution of autocorrelation for grid points in (**first row**) the tropics (30°S-30°N) and (**second row**) the extratropics (80°S-30°N and 30°N-80°N). The land mask of CESM2 is used to separate land (brown) and ocean (blue) grid points in all datasets. (**Third row**) Difference between multi-model mean and multi-observation mean absolute 1-year autocorrelation relative to the multi-observation mean value. The stippling shows grid points where less than 75% of all model-observation differences have the same sign. HadSLP2 and DCENT-I autocorrelation maps have been regridded bilinearly to the same 2.5° × 2.5° map as the other datasets.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

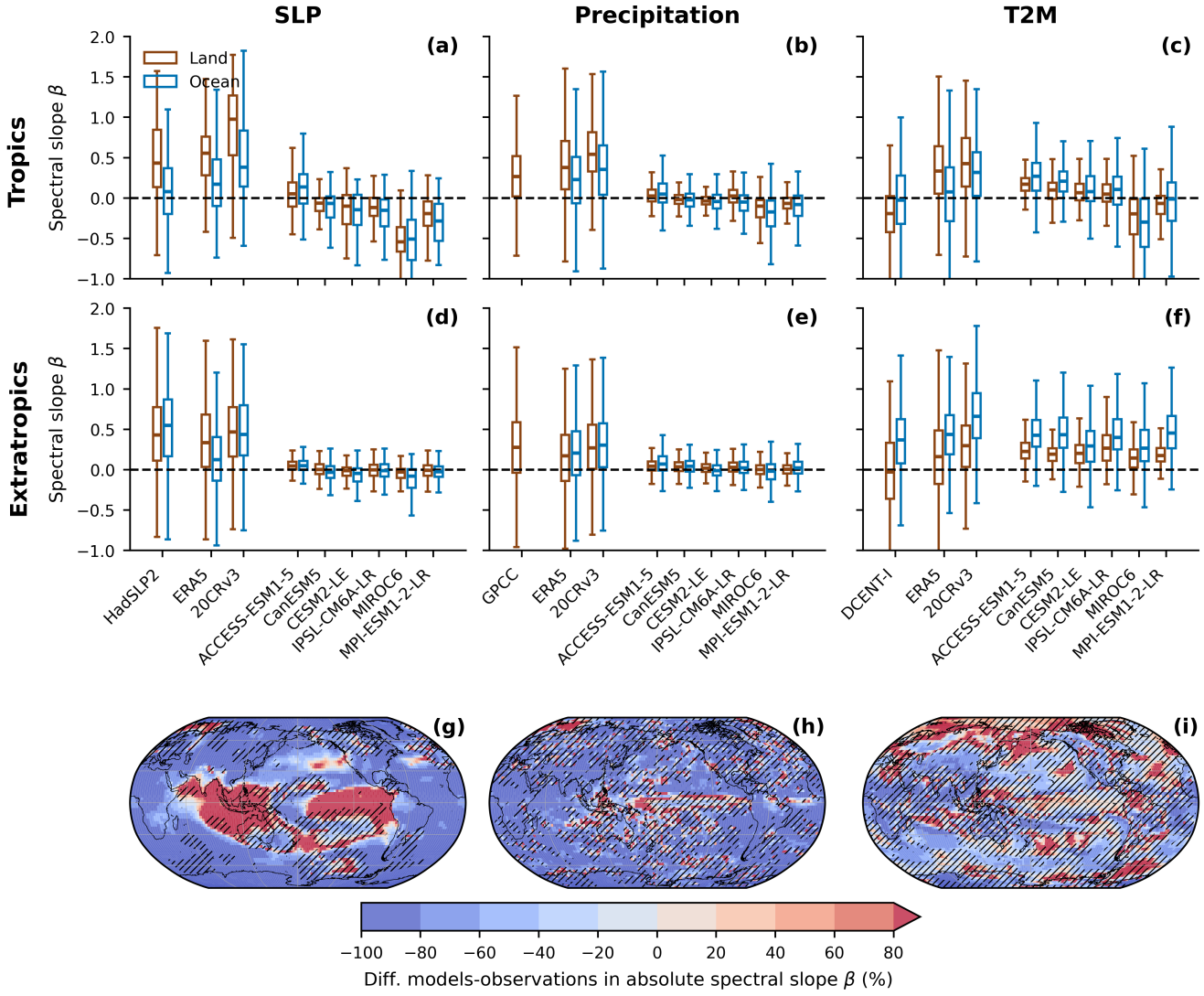


Figure 2. Spectral slopes estimated between time scales of 5 to 30 years for surface variables. Same as Figure 1 for spectral slopes between 5 and 30 years.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

oceans (Frankignoul, 1985; Barsugli and Battisti, 1998). The latter is mostly similar in models, reanalyses and observations, as demonstrated by our analysis of the spectral slopes of T2M over oceans in the previous section. However, the decadal variability of SLP and precipitation differs between models and observational datasets. We investigate this discrepancy by leveraging a statistical learning technique, namely ridge regression as detailed in Sect. 2. The strength of the SSTs constraint on the surface variables can then be evaluated via the spectral coherence between their actual anomalies and their SST-driven part given by the ridge regression (see Sect. 2). For a given time scale, the spectral coherence can be interpreted as a proportion of variance explained by the SSTs. In the following, we thus use interchangeably 'variance explained' or 'coherence'. In this section, we focus on the yearly time scales and evaluate longer time scales in the next section.

At yearly time scales, the proportion of variance of SLP and precipitation explained by SSTs is low in climate models over both land and ocean in the extratropics (Fig. 3d,e), typically around 20% for SLP and 10% for precipitation. There are geographical and model differences (Fig. 4 and Fig. 5), but these values are overall robust across models and this represents a large underestimation compared to reanalyses (Fig. 3d,e). The latter rather suggest coherence at yearly time scales between 40 to 80%, although the limited sample size of reanalyses likely makes them prone to overfitting (see tests in Sect. 2). While the exact magnitude of the reanalyses results needs to be interpreted with caution, they are consistently higher and this is coherent with the higher autocorrelation values than in models found in the previous section.

In the tropics, and especially over the central Pacific, the models' coherence is larger (Fig. 3a,b), with generally more agreement with the reanalyses despite still an underestimation of the variance explained. This higher coherence can be interpreted physically as a moderate to strong conditioning of SLP patterns and associated precipitation in the tropics by SSTs. This is expected insofar as the multi-year ENSO dynamics in both models and reanalyses strongly conditions the variability of all surface variables in the region (Timmermann et al., 2018) — although the behavior of ENSO can be very different in models compared to observations (Planton et al., 2021).

T2M generally shows larger coherence values than SLP and precipitation in models (Fig. 3c,f and Fig. 6), especially over oceans. This is no surprise given that T2M is closely related to SSTs even at one-month lag, and as a consequence models and reanalyses are not substantially different with respect to how much global SSTs condition T2M over tropical and subtropical oceans (Fig. 3i). However, the T2M coherence over land, especially in the extratropics, is lower than over oceans in models, contrary to what is estimated for reanalyses.

3.3 Precipitation and SLP variance explained by SSTs does not increase at long time scales in models

The low proportion of variance explained by SSTs at yearly time scales may not be surprising given the large atmospheric natural variability from year to year. However, at longer time scales, the atmospheric natural variability of surface variables decreases as yearly variability is smoothed out, which might lead to an increase in the proportion of variance explained by the SSTs. We refer the reader to Fig. 7 for a visual example of this mechanism using two grid points of the CESM2-LE. We illustrate in Fig. 8 the dependency on time scale of the spectral coherence between the surface variables and their SST-driven component as measured by the ridge regression, for six illustrative grid points, for the CESM2-LE (100 members) and the 20CRv3 reanalysis. The results are geographically dependent, but there is a general pattern of increasing coherence

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

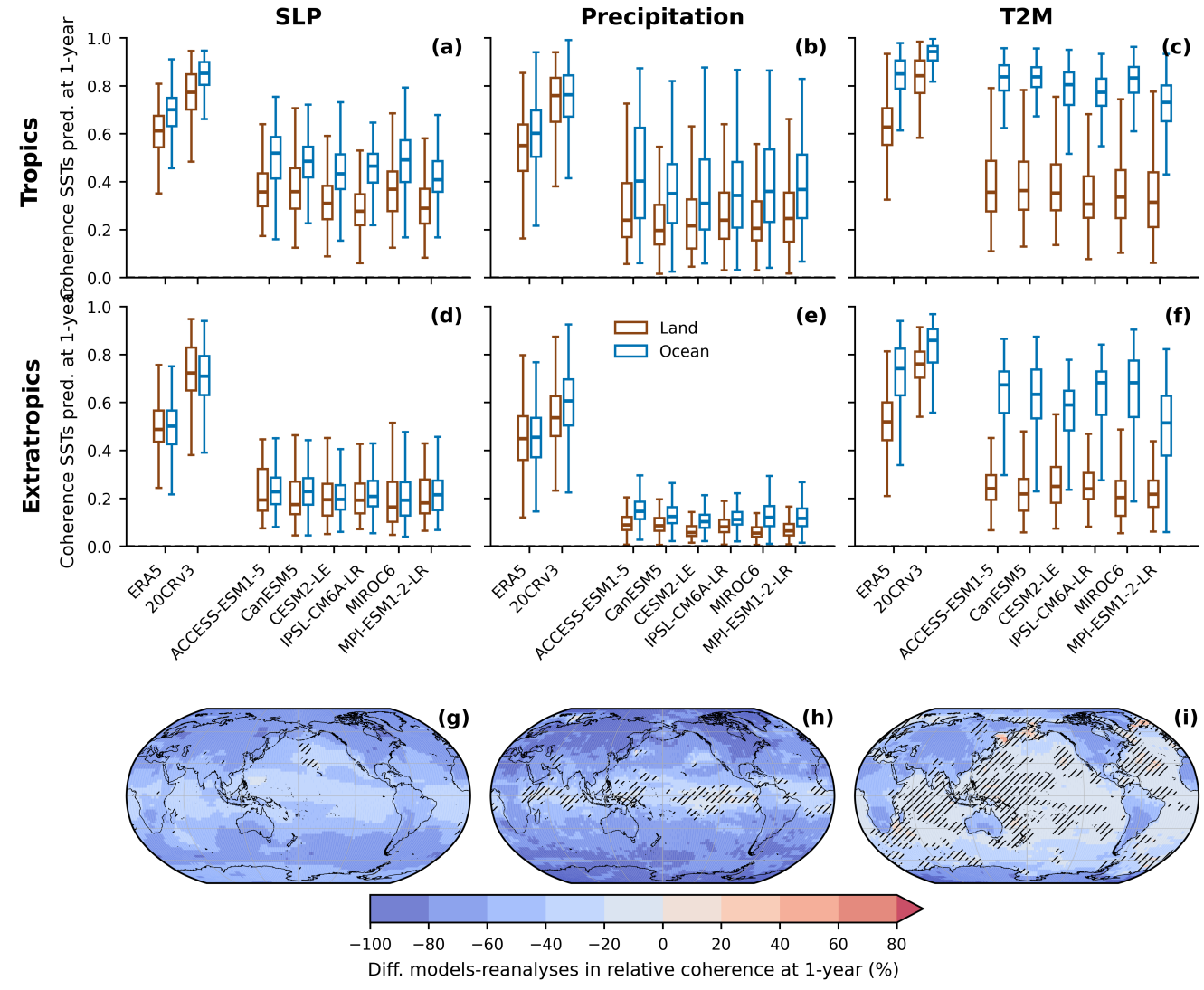


Figure 3. Coherence between surface variables and their SST-driven component. Spectral coherence at 1-year time scale between monthly anomalies and SST-driven monthly anomalies as predicted by the ridge regression (see Methods) for (**first column**) SLP, (**second column**) yearly accumulated precipitation and (**third column**) T2M. The coherence at 1-year time scale is computed by averaging the spectral coherence between time scales of 0.8 to 1.2 year. For the 20CRv3 reanalysis the coherence is computed using only the period 1900-2015. The boxplots show the distribution of autocorrelation for grid points in (**first row**) the tropics (30°S-30°N) and (**second row**) the extratropics (80°S-30°N and 30°N-80°N). The landmask of CESM2 is used for every reanalysis and model datasets to separate land and ocean grid points. (**Third row**) Difference between multi-model mean and multi-reanalysis mean coherence relative to the multi-reanalysis mean value. The stipplings show grid points where less than 75% of all model-reanalysis differences have the same sign.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

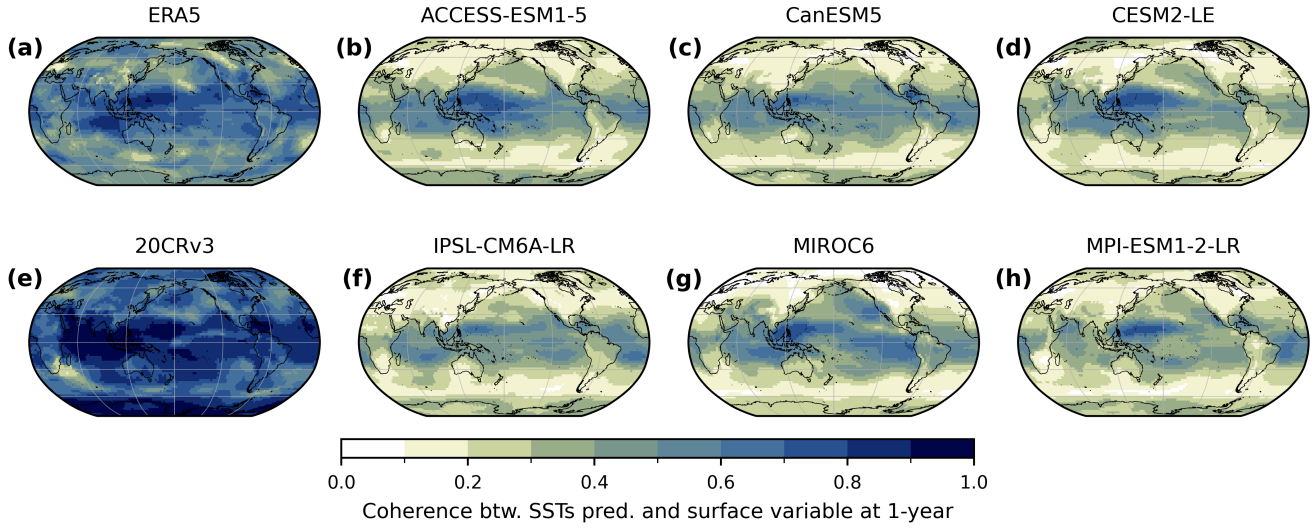


Figure 4. Spectral coherence at 1-year time scale between SLP and SST-driven SLP. The spectral coherence is computed between the monthly precipitation anomalies and the SST-driven monthly precipitation anomalies as predicted by the ridge regression. The coherence at 1-year time scale is computed by averaging the spectral coherence between time scales of 0.8 to 1.2 year. For the 20CRv3 reanalysis the coherence is computed using only the period 1900-2015.

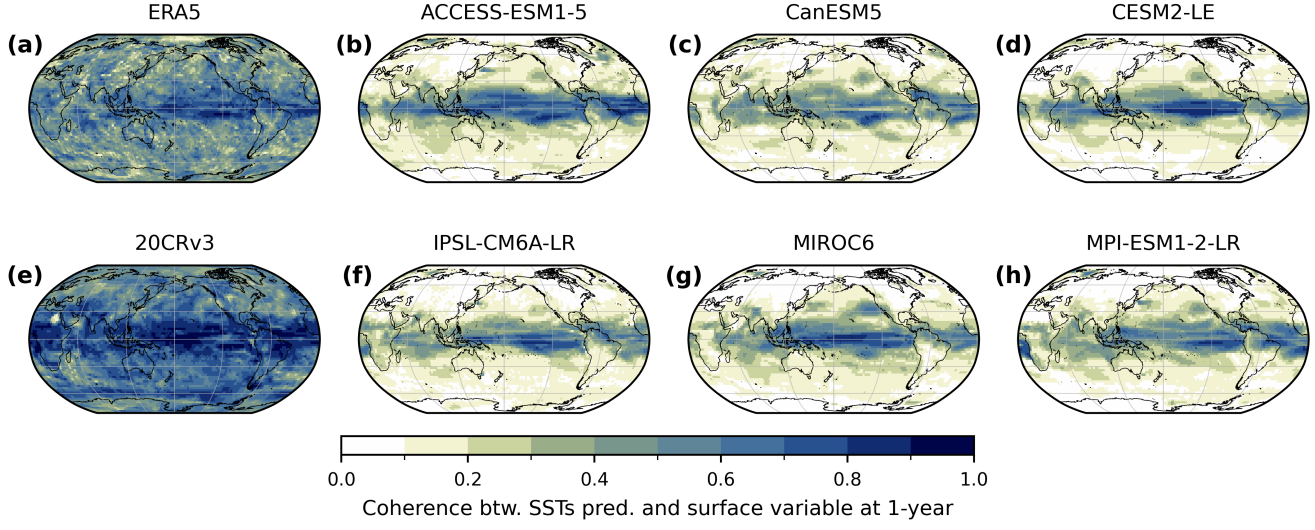


Figure 5. Spectral coherence at 1-year time scale between precipitation and SST-driven precipitation. Same as Fig. 4 for precipitation.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

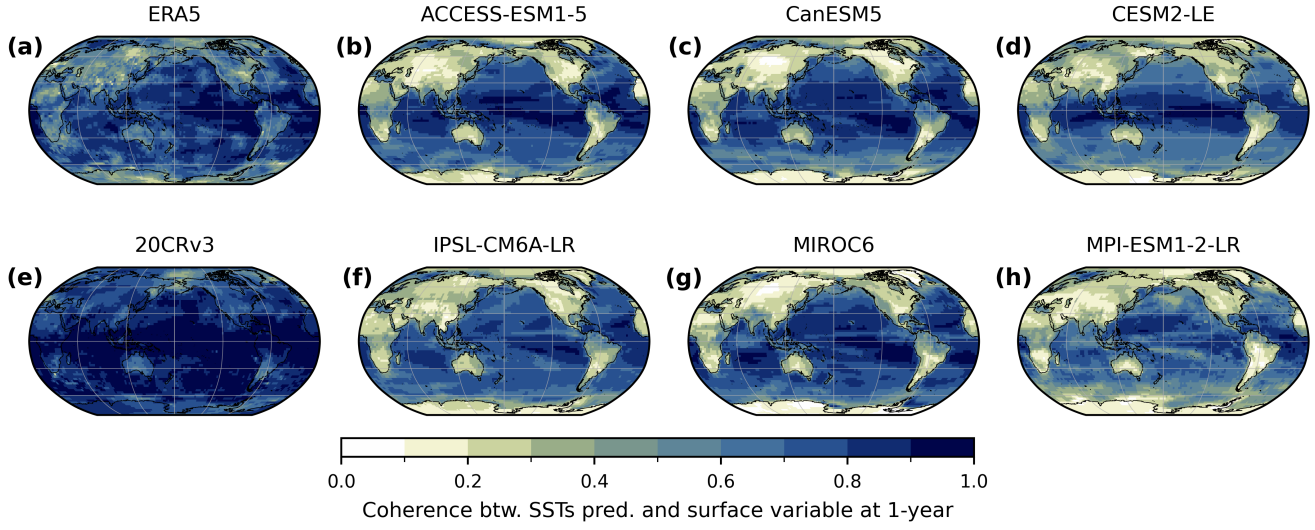


Figure 6. Spectral coherence at 1-year time scale between T2M and SST-driven T2M. Same as Fig. 5 for T2M.

at long time scales for the reanalysis for every grid point and variable. In contrast, the model does not show systematically increasing coherence at long time scales. For example, at the central Europe grid point (Fig. 8b), the coherence for both SLP and precipitation remains very low (around 10%) even at multi-decadal time scales. T2M presents a very different behavior, with increasing coherence from 25% at 1-year, up to 50% at 20-year time scales. Several tropical and sub-tropical grid points show a peak of coherence around 5 years for all variables in the model (Fig. 8g,l,e,o), which likely reflects the influence of ENSO that is then vanishing at longer time scales. In contrast, at the central Pacific grid point (Fig. 8j) the coherence increases and remains high at long time scales, similar to the reanalysis, which points towards a very strong constraint of the long-term variability of surface atmospheric variables by the SSTs in this region that is well captured by the model.

The other models show similar behaviors (Fig. 9). The proportion of variance of surface variables conditioned by the SSTs increases at longer time scales for all surface variables in the tropics. However, it increases at different paces and typically more for T2M than for precipitation and SLP. For the latter, subtropical highs over oceans show a large increase in their variance explained. On the contrary, the extratropics show little increase in variance explained at long time scales for precipitation and SLP. For example for SLP in Eurasia, the spectral coherence at 20 years in models is almost the same as at 1 year (Fig. 9a). Similar results stand for precipitation (Fig. 9b), with more regional variability — for example a small but robust increase in northern Siberia that is likely linked to the increased spectral coherence over the Barents-Kara sea. North America also displays an increased variance explained at long time scales for SLP, but the results are more mixed for precipitation. The behavior of SLP and precipitation stands in stark contrast to that of T2M: for almost all tropical and extratropical land, SSTs explain the variability of T2M better at long rather than short timescales.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

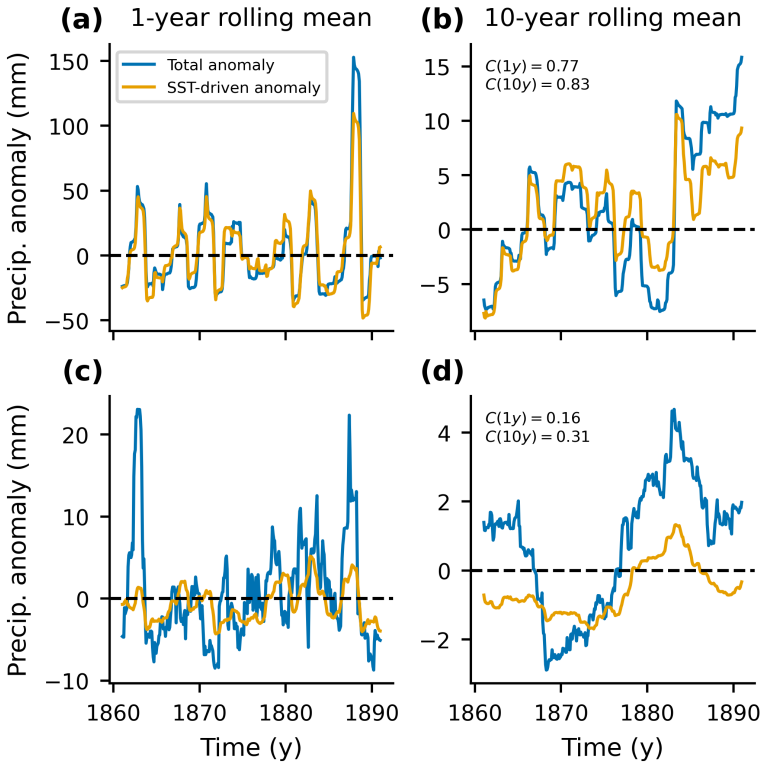


Figure 7. Illustration of the increasing conditioning of SSTs on surface variables at long time scales. Example time series of precipitation anomaly and SST-driven precipitation anomaly from the ridge regression for two grid points in one member of the CESM2-LE: (**a,b**) central Pacific grid point (lat=0°, lon=245°E) and (**c,d**) extratropical Pacific grid point (lat=30°N, lon=245°E). The text in (**b,d**) gives the spectral coherence between the anomalies and the SST-driven anomalies at time scales of 1 and 10 years, respectively.

4 Conclusions and discussion

The use of several large ensembles allows us to reach robust conclusions for CMIP-class climate models. For all of the six models investigated, we attribute the weak autocorrelation in precipitation and SLP from one year to another, and thus the white noise behavior at decadal time scales (Parsons et al., 2017; Simpson et al., 2018; Eade et al., 2022), to a weak forcing of these surface variables by decadal SST variability. This is especially the case in the extratropics, while the deep tropics behave differently because of the multi-year variability of ENSO and the strong conditioning it imposes on the variability of atmospheric surface variables in the region (see examples in Fig. 8). We have shown that T2M does behave differently than SLP and precipitation, with more variance explained by the SSTs, especially at long time scales. A plausible interpretation of the diverging results for the decadal variability of T2M and SLP is that the atmosphere responds mainly thermodynamically but not dynamically to oceanic forcings in these models. In other words, in the models, the decadal variability of T2M over

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

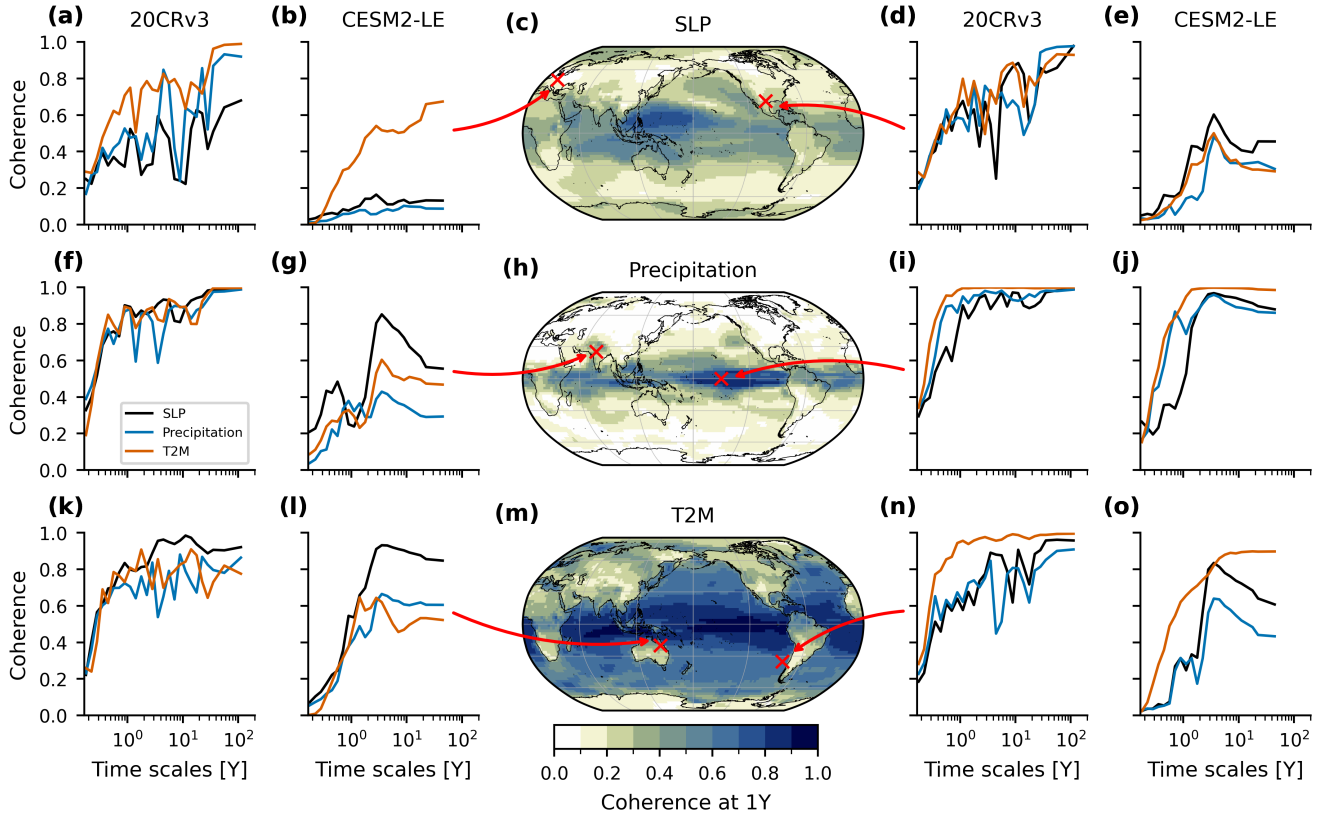


Figure 8. Illustration of the evolution of the spectral coherence with time scales for surface variables for CESM2 and comparison with 20CRv3. Maps show the spectral coherence between surface variables and their SST-driven component given by the ridge regression at the 1-year time scale for (**c**) SLP, (**h**) precipitation and (**m**) T2M. Panels show the dependency of spectral coherence with time scale for the grid points shown by the crosses on the maps for (**first column**) the 20CRv3 reanalysis (1900-2015) and (**second column**) the CESM2-LE (100 members, 1850-1900). For visualization, the spectral coherence was smoothed by averaging over bins of size 0.1 in log-space.

land is likely mainly the result of the advection of anomalous ocean-atmosphere heat fluxes by the mean circulation, while the atmospheric circulation itself, and thus SLP and precipitation, remains weakly affected by these fluxes.

One may wonder whether at long time scales our interpretation of the SST anomalies explaining the variability in surface variables can still be maintained, considering that the atmosphere can also be seen as the source of oceanic variability (Hasselmann, 1976). However, we consider the link between global SST anomalies — and thus implicitly the dominant, large-scale modes of variability — and anomalies of local surface variables. Thus, the risk of an atmospheric variable confounding the proportion of variance explained is mitigated. More fundamentally, the source of oceanic variability does not matter for our conclusions because we should also find a larger portion of variance explained if the oceanic variability is of atmospheric origin and the dynamical coupling is correctly represented — cf. also our time scales discussion in Sect. 2. We also tested

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

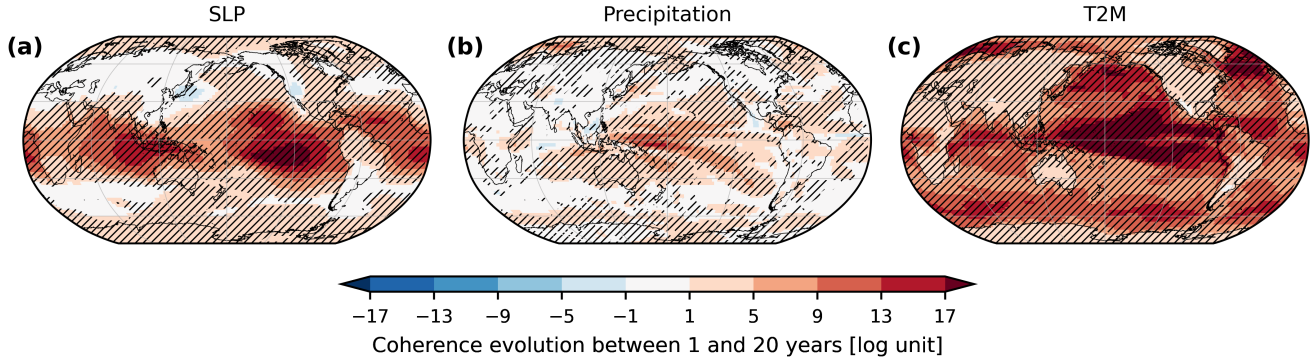


Figure 9. Variation of the spectral coherence of surface variables with time scales beyond 1 year. Multi-model mean of the log-difference (see Methods) of the spectral coherence between short (1 year) and long (20 years) time scales for (**a**) SLP, (**b**) precipitation and (**c**) T2M. The spectral coherence is averaged between 0.8 and 1.2 years for short time scales, and between 15 and 25 years for long timescales. A log-unit of 1 is equivalent to an increase of the spectral coherence from 0 to 0.1. Red and blue values indicate that SSTs explain a larger and smaller portion of the variability of surface variables, respectively, at long rather than short time scales. Hatching shows where all models agree on the sign of the change in coherence.

our analysis on two sets of 11 runs of the CESM2 model over the quasi pre-industrial period: 11 fully coupled runs and 11 atmospheric-only simulations forced by the SSTs of the coupled 11 runs. The latter thus have the same atmospheric variability conditional on the SSTs imposed as the coupled runs, but no direct feedback from the atmospheric variability to the oceanic variability. We find similar autocorrelation (not shown), spectral slopes (not shown) and spectral coherence (Fig. 10) for the coupled and atmospheric-only simulations, which confirms our interpretation.

Our results demonstrate that models show a weak conditioning of the decadal variability of surface variables by global SST anomalies in the extratropics. Physically, it could be expected that persistent SST anomalies should have an influence beyond the local boundary layer, impact the free tropospheric temperature gradients and thus the mean position of mid-latitude jets via the thermal wind relationship. As the atmospheric circulation is impacted, this should then imprint on SLP and precipitation variability, contrary to what we have shown for the models. It is challenging to validate that the real world is behaving differently than these climate models with the exact same ridge regression methodology. This is because of the lack of long and global observational products for these three surface variables. With some additional assumptions (see Sect. 2), we tried such a comparison with three observational products (HadSLP2, GPCC and DCENT-I) and the ERA5 and 20CRv3 reanalyses. We showed that the observational products suggest a robust underestimation of the decadal variability of SLP and precipitation by climate models, at least based on autocorrelation and spectral slopes, which is consistent with previous studies (Simpson et al., 2018; Laepple et al., 2023). We acknowledge that the results of the ridge regression procedure for the two reanalyses are less robust than for climate models, but the results obtained are nevertheless consistent with larger decadal variability coming from a stronger conditioning of surface variables by the SSTs in the real world. This is also consistent with previous observational studies (O'Reilly et al., 2017; Simpson et al., 2019; Smith et al., 2020).

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

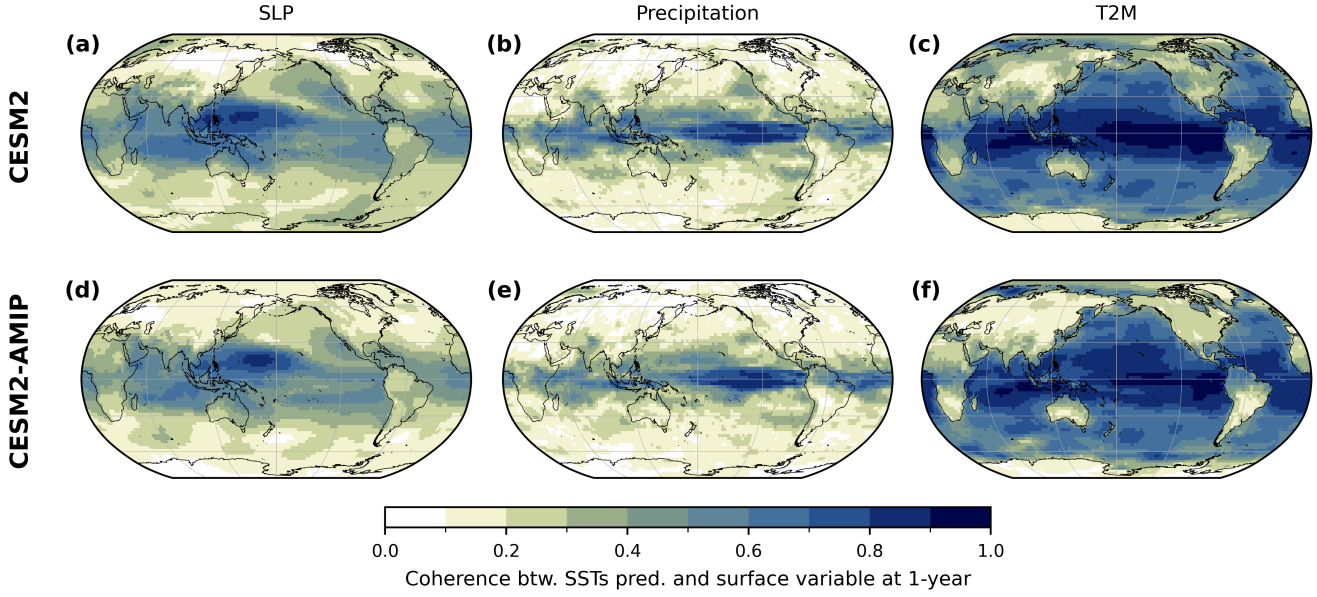


Figure 10. Comparison between spectral coherence at 1-year time scale between coupled and uncoupled simulations. 1-year spectral coherence using two 11-member ensembles from the CESM2 model over the quasi-industrial period (1850-1900) with (**first row**) fully-coupled simulations and (**second row**) atmosphere-only simulations using the SSTs of the fully-coupled simulations. The coherence at 1-year time scale is computed by averaging the spectral coherence between time scales of 0.8 to 1.2 year.

Our global and systematic investigation of the decadal variability in models gives consistent results with studies that showed a likely underestimation of the variability of the climate system at these time scales by CMIP-class models (Laepple and Huybers, 2014; Dee et al., 2017; Kravtsov, 2017; Cheung et al., 2017; Laepple et al., 2023). Our results may help explain the signal-to-noise paradox at decadal time scales (Scaife and Smith, 2018; Smith et al., 2020; Weisheimer et al., 2024): if the atmosphere is not responding strongly enough to SST anomalies in models, as we show here, then the spread in atmospheric surface variables conditional on a particular SST pattern will be too high, and observations will correlate better with the ensemble mean than each individual ensemble member. More generally, our results imply that the decadal variability of precipitation as projected by climate models should be considered carefully, especially when it comes to decadal predictions and the projection of multi-year extremes (Ault et al., 2013; O'Reilly et al., 2021). Why the information of the SST variability is not transmitted dynamically to the atmosphere in the extratropics remains a conundrum at this point. Several modeling studies showed that the relatively low resolution of these models could be the culprit. There has indeed been suggestions for non-linear responses of synoptic-scale dynamical processes in the atmosphere with respect to the model resolution for example (Smirnov et al., 2015; Czaja et al., 2019; Wills et al., 2024). Mesoscale (10-100 km) eddies in the ocean, especially along western boundary currents, have also been shown to be able to force the variability of the jet and could provide a mechanism to explain the decadal variability in

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

the real atmosphere (Wills and Thompson, 2018; Scaife et al., 2019; Joyce et al., 2019; Jüling et al., 2021; Zhang et al., 2021; Paolini et al., 2022; Seo et al., 2023).

Code availability. [The code to compute the quantities shown and the data produced will be made available after revisions.]

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix A: Cross-validation of the ridge regression model

This appendix details the cross-validation procedure of the ridge regression. We first isolate 10% of ensemble members of each model (test data set). On the remaining members (training data set) we use a 5-fold cross-validation. The training data set is first partitioned into 5 folds that contain 20% of the training data each. As the SST data, i.e. the predictors, are temporally very correlated we use a block cross-validation strategy (Roberts et al., 2017) where each block represents a continuous portion of years and members. This ensures that the SSTs of each block are as different as possible. For each fold, a fit over candidate λ values is estimated based on the remaining 80% of the data. These fits are then evaluated on the selected fold using the percentage of variance explained (R^2 score). We then compute the average R^2 score over the 5 folds and select the regularization parameter which maximizes it. The candidate values for λ are defined a priori using a semi-logarithmic scale: $\lambda = a \times 10^b$ for $a = 1, 2, ..., 9$ and $b = 1, 2, ..., b_{max}$. Based on preliminary tests we use a different b_{max} for each surface variable: $b_{max} = 12$ for SLP, $b_{max} = 14$ for precipitation and $b_{max} = 10$ for T2M. This range of values was chosen based on preliminary tests for different months, grid points and models in order to balance computational feasibility and robustness of the ridge regression fit. We evaluate whether the ridge regression is under- or overfitting based on whether the optimal regularization parameter λ_{opt} is equal to $\lambda_{min} = 10^1$ (underfitting) or $\lambda_{max} = 9 \times 10^{b_{max}}$ (overfitting). After determining the optimal regularization parameter, the ridge regression is fitted on the whole data set (training+test) and we show the results for this fit.

Note that for the reanalyses the amount of data is much more limited than for the large ensembles of climate models, only 85 months for each monthly regression for ERA5 for example. To mitigate this limited sampling issue, for both reanalysis, we train each monthly regression model by also using the data of the two adjacent months. For example, for the model to predict precipitation of the month of March using SSTs of the month of February, we also train the model with the (January SST, February precipitation) and (March SST, April precipitation) couples. For the 20CRv3 reanalysis, we use the entire period of the reanalysis, from 1806 to 2015. This is done only during the cross-validation and the training phases, the evaluation and the results shown are computed on the 1900-2015 period. To define the test dataset, we isolate one year every 10 years for the two reanalyses in order to sample a diversity of SST background.

Figure A1 shows for each reanalysis/model and each month the proportion of grid points that have a regularization parameter equals to λ_{max} (overfitting) and λ_{min} (underfitting). Figures A2-A4 show the difference in the R^2 between test and training data.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

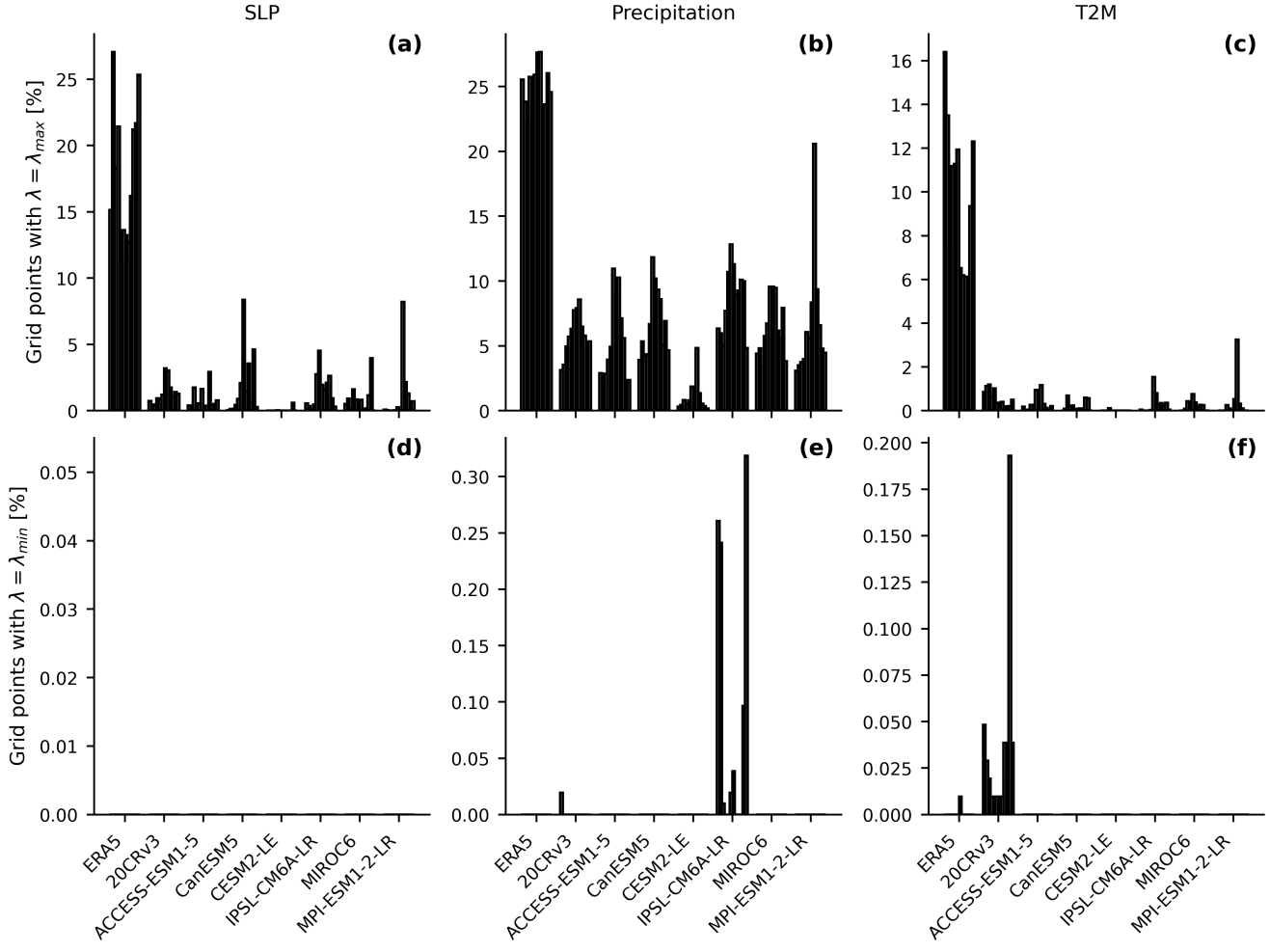


Figure A1. Percentage of grid points where the ridge regression on SSTs over- or underfit. (**First row**) Percentage of grid points where the regularization parameter λ is equal to the maximum regularization parameter allowed during the cross-validation procedure (overfitting) for (**a**) SLP, (**b**) precipitation and (**c**) T2M. Each bar shows the percentage of grid points for one month, starting in January. (**Second row**) Same for grid points where the regularization parameter is equal to the minimum regularization parameter allowed (underfitting).

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

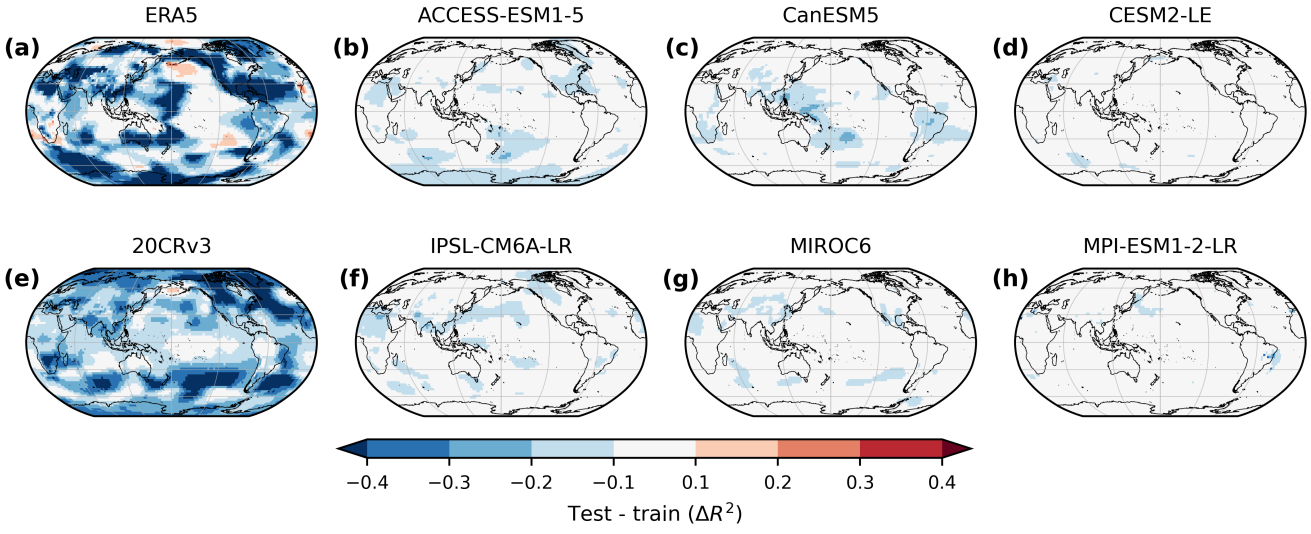


Figure A2. Evaluation of the generalization capability of the ridge regression for SLP. Difference in the percentage of yearly variance of SLP explained (R^2) by SSTs with the ridge regression model between training and test data. Test data represents the last 10% of ensemble members for the models. For the reanalyses, test data consist of one year taken every 10 years (8 years for ERA5, 20 years for 20CRv3).

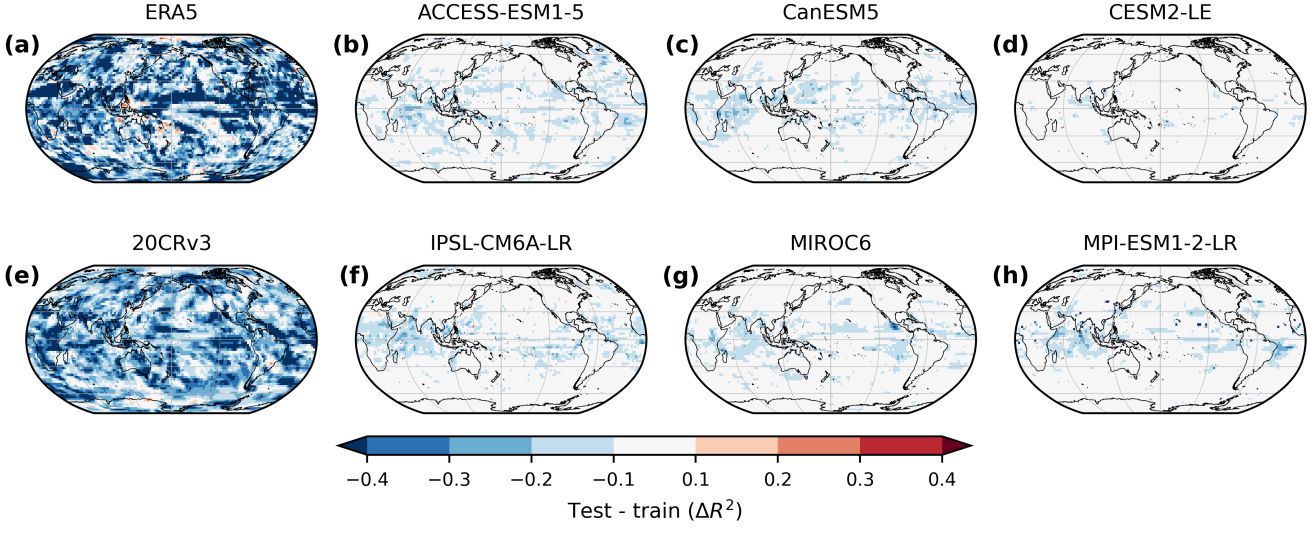


Figure A3. Evaluation of the generalization capability of the ridge regression for precipitation. Same as Fig. A2 for precipitation.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

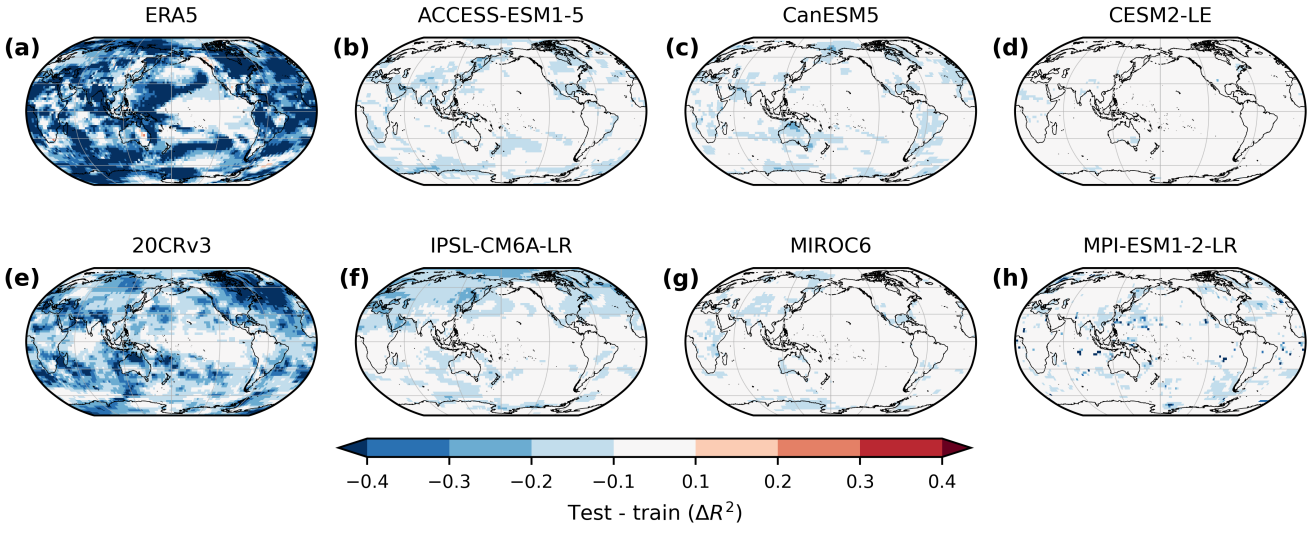


Figure A4. Evaluation of the generalization capability of the ridge regression for T2M. Same as Fig. A3 for T2M.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

Author contributions. Conceptualization: all authors. Methodology: R.N. Investigation: R.N. Visualization: R.N. Writing—original draft: R.N. Writing—review and editing: all authors.

Competing interests. The authors declare no competing interest.

Acknowledgements. RN has received funding from the Swiss National Science Foundation (SNSF) via grant no. 216710. RW is supported by the SNSF (grant no. 20337). RN would like to thank G. Gyuleva and P. Helpap for discussions on the analysis presented in this manuscript.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Al-Yaari, A., Swingedouw, D., Braconnot, P., Boyall, L., Lincoln, P., Marti, O., Caley, T., Extier, T., and Martin-Puertas, C.: Multi-centennial internal variability in the North Atlantic could drive additional warming over Europe, Nature Communications, 17, 2614, https://doi.org/10.1038/s41467-026-69209-2, 2026.

Allan, R. and Ansell, T.: A New Globally Complete Monthly Historical Gridded Mean Sea Level Pressure Dataset (HadSLP2): 1850–2004, Journal of Climate, 19, 5816–5842, https://doi.org/10.1175/JCLI3937.1, 2006.

Ault, T. R., Cole, J. E., Overpeck, J. T., Pederson, G. T., George, S. S., Otto-Bliesner, B., Woodhouse, C. A., and Deser, C.: The Continuum of Hydroclimate Variability in Western North America during the Last Millennium, Journal of Climate, 26, 5863–5878, https://doi.org/10.1175/JCLI-D-11-00732.1, 2013.

Barsugli, J. J. and Battisti, D. S.: The Basic Effects of Atmosphere–Ocean Thermal Coupling on Midlatitude Variability, Journal of the Atmospheric Sciences, 55, 477–493, https://doi.org/10.1175/1520-0469(1998)055<0477:TBEOAO>2.0.CO;2, 1998.

Battisti, D. S., Vimont, D. J., and Kirtman, B. P.: 100 Years of Progress in Understanding the Dynamics of Coupled Atmosphere–Ocean Variability, Meteorological Monographs, 59, 8.1–8.57, https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0025.1, 2019.

Bonnet, R., McKenna, C. M., and Maycock, A. C.: Model spread in multidecadal North Atlantic Oscillation variability connected to stratosphere–troposphere coupling, Weather and Climate Dynamics, 5, 913–926, https://doi.org/10.5194/wcd-5-913-2024, 2024.

Chan, D., Chan, S. C., Siddons, J. T., Cable, A., Faulkner, A., Cornes, R. C., Kent, E. C., Gebbie, G., and Huybers, P.: DCENT-I: A Globally Infilled Extension of the Dynamically Consistent ENsemble of Temperature Dataset, Geoscience Data Journal, 13, e70054, https://doi.org/10.1002/gdj3.70054, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/gdj3.70054, 2026.

Chatfield, C. and Xing, H.: The analysis of time series: an introduction with R, Chapman and hall/CRC, 2019.

Cheung, A. H., Mann, M. E., Steinman, B. A., Frankcombe, L. M., England, M. H., and Miller, S. K.: Comparison of Low-Frequency Internal Climate Variability in CMIP5 Models and Observations, Journal of Climate, 30, 4763–4776, https://doi.org/10.1175/JCLI-D-16-0712.1, 2017.

Cook, B. I., Smerdon, J. E., Cook, E. R., Williams, A. P., Anchukaitis, K. J., Mankin, J. S., Allen, K., Andreu-Hayles, L., Ault, T. R., Belmecheri, S., Coats, S., Coulthard, B., Fosu, B., Grierson, P., Griffin, D., Herrera, D. A., Ionita, M., Lehner, F., Leland, C., Marvel, K., Morales, M. S., Mishra, V., Ngoma, J., Nguyen, H. T. T., O'Donnell, A., Palmer, J., Rao, M. P., Rodriguez-Caton, M., Seager, R., Stahle, D. W., Stevenson, S., Thapa, U. K., Varuolo-Clarke, A. M., and Wise, E. K.: Megadroughts in the Common Era and the Anthropocene, Nature Reviews Earth & Environment, 3, 741–757, https://doi.org/10.1038/s43017-022-00329-1, 2022.

Czaja, A., Frankignoul, C., Minobe, S., and Vannière, B.: Simulating the Midlatitude Atmospheric Circulation: What Might We Gain From High-Resolution Modeling of Air-Sea Interactions?, Current Climate Change Reports, 5, 390–406, https://doi.org/10.1007/s40641-019-00148-5, 2019.

Davini, P., Hardenberg, J. v., and Corti, S.: Tropical origin for the impacts of the Atlantic Multidecadal Variability on the Euro-Atlantic climate, Environmental Research Letters, 10, 094010, https://doi.org/10.1088/1748-9326/10/9/094010, 2015.

Dee, S. G., Parsons, L. A., Loope, G. R., Overpeck, J. T., Ault, T. R., and Emile-Geay, J.: Improved spectral comparisons of paleoclimate models and observations via proxy system modeling: Implications for multi-decadal variability, Earth and Planetary Science Letters, 476, 34–46, https://doi.org/10.1016/j.epsl.2017.07.036, 2017.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

Eade, R., Stephenson, D. B., Scaife, A. A., and Smith, D. M.: Quantifying the rarity of extreme multi-decadal trends: how unusual was the late twentieth century trend in the North Atlantic Oscillation?, Climate Dynamics, 58, 1555–1568, https://doi.org/10.1007/s00382-021-05978-4, 2022.

Folland, C. K., Palmer, T. N., and Parker, D. E.: Sahel rainfall and worldwide sea temperatures, 1901–85, Nature, 320, 602–607, https://doi.org/10.1038/320602a0, 1986.

Frankignoul, C.: Sea surface temperature anomalies, planetary waves, and air-sea feedback in the middle latitudes, Reviews of Geophysics, 23, 357–390, https://doi.org/10.1029/RG023i004p00357, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/RG023i004p00357, 1985.

Frankignoul, C. and Hasselmann, K.: Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability, Tellus, 29, https://doi.org/10.3402/tellusa.v29i4.11362, 1977.

Garreaud, R. D., Boisier, J. P., Rondanelli, R., Montecinos, A., Sepúlveda, H. H., and Veloso-Aguila, D.: The Central Chile Mega Drought (2010–2018): A climate dynamics perspective, International Journal of Climatology, 40, 421–439, https://doi.org/10.1002/joc.6219, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.6219, 2020.

Gastineau, G. and Frankignoul, C.: Influence of the North Atlantic SST Variability on the Atmospheric Circulation during the Twentieth Century, Journal of Climate, 28, 1396–1416, https://doi.org/10.1175/JCLI-D-14-00424.1, 2015.

Ghil, M.: The physics of climate variability and climate change, Reviews of Modern Physics, 92, https://doi.org/10.1103/RevModPhys.92.035002, 2020.

Good, S., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Reid, R., Roberts-Jones, J., Searle, T., Waters, J., While, J., and Worsfold, M.: The Current Configuration of the OSTIA System for Operational Production of Foundation Sea Surface Temperature and Ice Concentration Analyses, Remote Sensing, 12, 720, https://doi.org/10.3390/rs12040720, 2020.

Hasselmann, K.: Stochastic climate models Part I. Theory, Tellus, 28, 473–485, https://doi.org/10.3402/tellusa.v28i6.11316, _eprint: https://doi.org/10.3402/tellusa.v28i6.11316, 1976.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning, Springer Series in Statistics, Springer, New York, NY, ISBN 978-0-387-84857-0 978-0-387-84858-7, https://doi.org/10.1007/978-0-387-84858-7, 2009.

Hirst, A. C.: Unstable and Damped Equatorial Modes in Simple Coupled Ocean-Atmosphere Models, Journal of the Atmospheric Sciences, 43, 606–632, https://doi.org/10.1175/1520-0469(1986)043<0606:UADEMI>2.0.CO;2, 1986.

Hoskins, B. J. and Karoly, D. J.: The Steady Linear Response of a Spherical Atmosphere to Thermal and Orographic Forcing, Journal of the Atmospheric Sciences, 38, 1179–1196, https://doi.org/10.1175/1520-0469(1981)038<1179:TSLROA>2.0.CO;2, 1981.

Huybers, P. and Curry, W.: Links between annual, Milankovitch and continuum temperature variability, Nature, 441, 329–332, https://doi.org/10.1038/nature04745, 2006.

Joyce, T. M., Kwon, Y.-O., Seo, H., and Ummenhofer, C. C.: Meridional Gulf Stream Shifts Can Influence Wintertime Variability in the North Atlantic Storm Track and Greenland Blocking, Geophysical Research Letters, 46, 1702–1708, https://doi.org/10.1029/2018GL081087, 2019.

Jüling, A., von der Heydt, A., and Dijkstra, H. A.: Effects of strongly eddying oceans on multidecadal climate variability in the Community Earth System Model, Ocean Science, 17, 1251–1271, https://doi.org/10.5194/os-17-1251-2021, 2021.

Kim, W. M., Yeager, S., Chang, P., and Danabasoglu, G.: Low-Frequency North Atlantic Climate Variability in the Community Earth System Model Large Ensemble, Journal of Climate, 31, 787–813, https://doi.org/10.1175/JCLI-D-17-0193.1, 2018.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.



Knutti, R., Masson, D., and Gettelman, A.: Climate model genealogy: Generation CMIP5 and how we got there, Geophysical Research Letters, 40, 1194–1199, https://doi.org/10.1002/grl.50256, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/grl.50256, 2013.

Kolstad, E. W.: Deficient ocean–atmosphere feedbacks constrain seasonal NAO prediction, Weather and Climate Dynamics, 7, 507–522, https://doi.org/10.5194/wcd-7-507-2026, 2026.

Kravtsov, S.: Pronounced differences between observed and CMIP5-simulated multidecadal climate variability in the twentieth century, Geophysical Research Letters, 44, 5749–5757, https://doi.org/10.1002/2017GL074016, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2017GL074016, 2017.

Kushnir, Y., Robinson, W. A., Bladé, I., Hall, N. M. J., Peng, S., and Sutton, R.: Atmospheric GCM Response to Extratropical SST Anomalies: Synthesis and Evaluation, Journal of Climate, 15, 2233–2256, https://doi.org/10.1175/1520-0442(2002)015<2233:AGRTES>2.0.CO;2, 2002.

Laepple, T. and Huybers, P.: Ocean surface temperature variability: Large model–data differences at decadal and longer periods, Proceedings of the National Academy of Sciences, 111, 16 682–16 687, https://doi.org/10.1073/pnas.1412077111, 2014.

Laepple, T., Ziegler, E., Weitzel, N., Hébert, R., Ellerhoff, B., Schoch, P., Martrat, B., Bothe, O., Moreno-Chamarro, E., Chevalier, M., Herbert, A., and Rehfeld, K.: Regional but not global temperature variability underestimated by climate models at supradecadal timescales, Nature Geoscience, 16, 958–966, https://doi.org/10.1038/s41561-023-01299-9, 2023.

Lau, N.-C. and Nath, M. J.: A Modeling Study of the Relative Roles of Tropical and Extratropical SST Anomalies in the Variability of the Global Atmosphere-Ocean System, Journal of Climate, 7, 1184–1207, https://doi.org/10.1175/1520-0442(1994)007<1184:AMSOTR>2.0.CO;2, 1994.

Liu, Z. and Wu, L.: Atmospheric Response to North Pacific SST: The Role ofOcean–Atmosphere Coupling, Journal of Climate, 17, 1859–1882, https://doi.org/10.1175/1520-0442(2004)017<1859:ARTNPS>2.0.CO;2, 2004.

MacKay, D. J.: Information theory, inference and learning algorithms, Cambridge university press, 2003.

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M. A., Greene, A. M., Hawkins, E., Hegerl, G., Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D., Stammer, D., and Stockdale, T.: Decadal Prediction: Can It Be Skillful?, Bulletin of the American Meteorological Society, 90, 1467–1486, https://doi.org/10.1175/2009BAMS2778.1, 2009.

Noyelle, R., Fischer, E., Knutti, R., and Wernli, H.: Does Oceanic Variability Explain the Intensity of Megadroughts and Megapluvials in the CESM2 Climate Model?, Journal of Climate, 2026.

O'Reilly, C. H., Woollings, T., and Zanna, L.: The Dynamical Influence of the Atlantic Multidecadal Oscillation on Continental Climate, Journal of Climate, 30, 7213–7230, https://doi.org/10.1175/JCLI-D-16-0345.1, 2017.

O'Reilly, C. H., Zanna, L., and Woollings, T.: Assessing External and Internal Sources of Atlantic Multidecadal Variability Using Models, Proxy Data, and Early Instrumental Indices, Journal of Climate, 32, 7727–7745, https://doi.org/10.1175/JCLI-D-19-0177.1, 2019.

O'Reilly, C. H., Befort, D. J., Weisheimer, A., Woollings, T., Ballinger, A., and Hegerl, G.: Projections of northern hemisphere extratropical climate underestimate internal variability and associated uncertainty, Communications Earth & Environment, 2, 194, https://doi.org/10.1038/s43247-021-00268-7, 2021.

Paolini, L. F., Athanasiadis, P. J., Ruggieri, P., and Bellucci, A.: The Atmospheric Response to Meridional Shifts of the Gulf Stream SST Front and Its Dependence on Model Resolution, Journal of Climate, 35, 6007–6030, https://doi.org/10.1175/JCLI-D-21-0530.1, 2022.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

Parsons, L. A., Loope, G. R., Overpeck, J. T., Ault, T. R., Stouffer, R., and Cole, J. E.: Temperature and Precipitation Variance in CMIP5 Simulations and Paleoclimate Records of the Last Millennium, Journal of Climate, 30, 8885–8912, https://doi.org/10.1175/JCLI-D-16-0863.1, 2017.

Peings, Y., Simpkins, G., and Magnusdottir, G.: Multidecadal fluctuations of the North Atlantic Ocean and feedback on the winter climate in CMIP5 control simulations, Journal of Geophysical Research: Atmospheres, 121, 2571–2592, https://doi.org/10.1002/2015JD024107, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015JD024107, 2016.

Planton, Y. Y., Guilyardi, E., Wittenberg, A. T., Lee, J., Gleckler, P. J., Bayr, T., McGregor, S., McPhaden, M. J., Power, S., Roehrig, R., Vialard, J., and Voldoire, A.: Evaluating Climate Models with the CLIVAR 2020 ENSO Metrics Package, Bulletin of the American Meteorological Society, 102, E193–E217, https://doi.org/10.1175/BAMS-D-19-0337.1, 2021.

Qasmi, S., Cassou, C., and Boé, J.: Teleconnection Between Atlantic Multidecadal Variability and European Temperature: Diversity and Evaluation of the Coupled Model Intercomparison Project Phase 5 Models, Geophysical Research Letters, 44, 11,140–11,149, https://doi.org/10.1002/2017GL074886, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2017GL074886, 2017.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., and Dormann, C. F.: Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure, Ecography, 40, 913–929, https://doi.org/10.1111/ecog.02881, _eprint: https://nsojournals.onlinelibrary.wiley.com/doi/pdf/10.1111/ecog.02881, 2017.

Scaife, A. A. and Smith, D.: A signal-to-noise paradox in climate science, npj Climate and Atmospheric Science, 1, 28, https://doi.org/10.1038/s41612-018-0038-4, 2018.

Scaife, A. A., Camp, J., Comer, R., Davis, P., Dunstone, N., Gordon, M., MacLachlan, C., Martin, N., Nie, Y., Ren, H.-L., Roberts, M., Robinson, W., Smith, D., and Vidale, P. L.: Does increased atmospheric resolution improve seasonal climate predictions?, Atmospheric Science Letters, 20, e922, https://doi.org/10.1002/asl.922, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/asl.922, 2019.

Schneider, U., Ziese, M., Meyer-Christoffer, A., Finger, P., Rustemeier, E., and Becker, A.: The new portfolio of global precipitation data products of the Global Precipitation Climatology Centre suitable to assess and quantify the global water cycle and resources, in: Proceedings of IAHS, vol. 374, pp. 29–34, Copernicus GmbH, ISSN 2199-8981, https://doi.org/10.5194/piahs-374-29-2016, 2016.

Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D., and Bacmeister, J. T.: On the Cause of the 1930s Dust Bowl, Science, 303, 1855–1859, https://doi.org/10.1126/science.1095048, 2004.

Schurer, A. P., Hegerl, G. C., Goosse, H., Bollasina, M. A., England, M. H., Smith, D. M., and Tett, S. F. B.: Role of multi-decadal variability of the winter North Atlantic Oscillation on Northern Hemisphere climate, Environmental Research Letters, 18, 044 046, https://doi.org/10.1088/1748-9326/acc477, 2023.

Seo, H., O'Neill, L. W., Bourassa, M. A., Czaja, A., Drushka, K., Edson, J. B., Fox-Kemper, B., Frenger, I., Gille, S. T., Kirtman, B. P., Minobe, S., Pendergrass, A. G., Renault, L., Roberts, M. J., Schneider, N., Small, R. J., Stoffelen, A., and Wang, Q.: Ocean Mesoscale and Frontal-Scale Ocean–Atmosphere Interactions and Influence on Large-Scale Climate: A Review, Journal of Climate, 36, 1981–2013, https://doi.org/10.1175/JCLI-D-21-0982.1, 2023.

Simpson, I. R., Deser, C., McKinnon, K. A., and Barnes, E. A.: Modeled and Observed Multidecadal Variability in the North Atlantic Jet Stream and Its Connection to Sea Surface Temperatures, Journal of Climate, 31, 8313–8338, https://doi.org/10.1175/JCLI-D-18-0168.1, 2018.

Simpson, I. R., Yeager, S. G., McKinnon, K. A., and Deser, C.: Decadal predictability of late winter precipitation in western Europe through an ocean–jet stream connection, Nature Geoscience, 12, 613–619, https://doi.org/10.1038/s41561-019-0391-x, 2019.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., Allan, R., Yin, X., Vose, R., Titchner, H., Kennedy, J., Spencer, L. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Crouthamel, R., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Jourdain, S., Kaplan, A., Kubota, H., Blancq, F. L., Lee, T.-C., Lorrey, A., Luterbacher, J., Maugeri, M., Mock, C. J., Moore, G. W. K., Przybylak, R., Pudmenzky, C., Reason, C., Slonosky, V. C., Smith, C. A., Tinz, B., Trewin, B., Valente, M. A., Wang, X. L., Wilkinson, C., Wood, K., and Wyszyński, P.: Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system, Quarterly Journal of the Royal Meteorological Society, 145, 2876–2908, https://doi.org/10.1002/qj.3598, 2019.

Smirnov, D., Newman, M., Alexander, M. A., Kwon, Y.-O., and Frankignoul, C.: Investigating the Local Atmospheric Response to a Realistic Shift in the Oyashio Sea Surface Temperature Front, Journal of Climate, 28, 1126–1147, https://doi.org/10.1175/JCLI-D-14-00285.1, 2015.

Smith, D. M., Scaife, A. A., Eade, R., Athanasiadis, P., Bellucci, A., Bethke, I., Bilbao, R., Borchert, L. F., Caron, L.-P., Counillon, F., Danabasoglu, G., Delworth, T., Doblas-Reyes, F. J., Dunstone, N. J., Estella-Perez, V., Flavoni, S., Hermanson, L., Keenlyside, N., Kharin, V., Kimoto, M., Merryfield, W. J., Mignot, J., Mochizuki, T., Modali, K., Monerie, P.-A., Müller, W. A., Nicolí, D., Ortega, P., Pankatz, K., Pohlmann, H., Robson, J., Ruggieri, P., Sospedra-Alfonso, R., Swingedouw, D., Wang, Y., Wild, S., Yeager, S., Yang, X., and Zhang, L.: North Atlantic climate far more predictable than models imply, Nature, 583, 796–800, https://doi.org/10.1038/s41586-020-2525-0, 2020.

Soci, C., Hersbach, H., Simmons, A., Poli, P., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Villaume, S., Haimberger, L., Woollen, J., Buontempo, C., and Thépaut, J.-N.: The ERA5 global reanalysis from 1940 to 2022, Quarterly Journal of the Royal Meteorological Society, 150, 4014–4048, https://doi.org/10.1002/qj.4803, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.4803, 2024.

Thomson, S. I. and Vallis, G. K.: Atmospheric Response to SST Anomalies. Part II: Background-State Dependence, Teleconnections, and Local Effects in Summer, Journal of the Atmospheric Sciences, 75, 4125–4138, https://doi.org/10.1175/JAS-D-17-0298.1, 2018a.

Thomson, S. I. and Vallis, G. K.: Atmospheric Response to SST Anomalies. Part I: Background-State Dependence, Teleconnections, and Local Effects in Winter, Journal of the Atmospheric Sciences, 75, 4107–4124, https://doi.org/10.1175/JAS-D-17-0297.1, 2018b.

Timmermann, A., An, S.-I., Kug, J.-S., Jin, F.-F., Cai, W., Capotondi, A., Cobb, K. M., Lengaigne, M., McPhaden, M. J., Stuecker, M. F., Stein, K., Wittenberg, A. T., Yun, K.-S., Bayr, T., Chen, H.-C., Chikamoto, Y., Dewitte, B., Dommenget, D., Grothe, P., Guilyardi, E., Ham, Y.-G., Hayashi, M., Ineson, S., Kang, D., Kim, S., Kim, W., Lee, J.-Y., Li, T., Luo, J.-J., McGregor, S., Planton, Y., Power, S., Rashid, H., Ren, H.-L., Santoso, A., Takahashi, K., Todd, A., Wang, G., Wang, G., Xie, R., Yang, W.-H., Yeh, S.-W., Yoon, J., Zeller, E., and Zhang, X.: El Niño–Southern Oscillation complexity, Nature, 559, 535–545, https://doi.org/10.1038/s41586-018-0252-6, 2018.

Titchner, H. A. and Rayner, N. A.: The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations, Journal of Geophysical Research: Atmospheres, 119, 2864–2889, https://doi.org/10.1002/2013JD020316, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013JD020316, 2014.

Wang, X.-Y., Zhu, J., Chang, C.-H., Johnson, N. C., Liu, H., Li, Y., Song, C., Xin, M., Zhou, Y., and Li, X.: Underestimated responses of Walker circulation to ENSO-related SST anomaly in atmospheric and coupled models, Geoscience Letters, 8, 17, https://doi.org/10.1186/s40562-021-00186-8, 2021.

Weisheimer, A., Baker, L. H., Bröcker, J., Garfinkel, C. I., Hardiman, S. C., Hodson, D. L. R., Palmer, T. N., Robson, J. I., Scaife, A. A., Screen, J. A., Shepherd, T. G., Smith, D. M., and Sutton, R. T.: The Signal-to-Noise Paradox in Climate Forecasts: Re-visiting Our Understanding and Identifying Future Priorities, Bulletin of the American Meteorological Society, 105, E651–E659, https://doi.org/10.1175/BAMS-D-24-0019.1, 2024.

https://doi.org/10.5194/egusphere-2026-6524
Preprint. Discussion started: 4 November 2026
© Author(s) 2026. CC BY 4.0 License.



Wills, R. C. J., Herrington, A. R., Simpson, I. R., and Battisti, D. S.: Resolving Weather Fronts Increases the Large-Scale Circulation Response to Gulf Stream SST Anomalies in Variable-Resolution CESM2 Simulations, Journal of Advances in Modeling Earth Systems, 16, e2023MS004 123, https://doi.org/10.1029/2023MS004123, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2023MS004123, 2024.

Wills, S. M. and Thompson, D. W. J.: On the Observed Relationships between Wintertime Variability in Kuroshio–Oyashio Extension Sea Surface Temperatures and the Atmospheric Circulation over the North Pacific, Journal of Climate, 31, 4669–4681, https://doi.org/10.1175/JCLI-D-17-0343.1, 2018.

Woollings, T., Czuchnicki, C., and Franzke, C.: Twentieth century North Atlantic jet variability, Quarterly Journal of the Royal Meteorological Society, 140, 783–791, https://doi.org/10.1002/qj.2197, _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2197, 2014.

Zhang, W., Kirtman, B., Siqueira, L., Clement, A., and Xia, J.: Understanding the signal-to-noise paradox in decadal climate predictability from CMIP5 and an eddying global coupled model, Climate Dynamics, 56, 2895–2913, https://doi.org/10.1007/s00382-020-05621-8, 2021.