

Supplementary information to “Evaluating climate field reconstruction techniques using improved emulations of real-world conditions”

Jianghao Wang¹, Julien Emile-Geay¹,
Dominique Guillot², and Jason E. Smerdon³

¹University of Southern California, Los Angeles, California

²Stanford University, Stanford, California

³Lamont-Doherty Earth Observatory of Columbia University, Palisades, New
York

1 Methods

Assume a linear relationship exists between temperature T and proxy P :

$$\mathbf{T} = \mathbf{B}\mathbf{P} + \varepsilon, \quad (1)$$

where ε denotes the error term, following a standard multivariate normal distribution. One can show that when $\mathbf{P}^t\mathbf{P}$ is non-singular, i.e. \mathbf{P} is full rank, the estimate of \mathbf{B} is given by the ordinary least square estimates:

$$\hat{\mathbf{B}} = (\mathbf{P}^t\mathbf{P})^{-1}\mathbf{P}^t\mathbf{T} \quad (2)$$

In paleoclimate reconstructions, however, it is often the case $\mathbf{P}^t\mathbf{P}$ is rank-deficient, making traditional regression methods no longer applicable. Four CFR techniques are used here to solve the problem.

1.1 RegEM-TTLS [*Schneider, 2001*]

Reconstructions with RegEM-TTLS were performed using the standard algorithm described in *Schneider* [2001]. In particular, the implementation of TTLS follows the description in *Fierro et al.* [1997]. First, SVD is calculated on the augmented data matrix $A \equiv (\mathbf{T}_{\text{calib}}, \mathbf{P}_{\text{calib}})$, in which small singular values are treated as zeros, and $\mathbf{T}_{\text{calib}}, \mathbf{P}_{\text{calib}}$ refer to \mathbf{T}, \mathbf{P} over the calibration period. Regularization is achieved by retaining only the first k large eigenvalues of the SVD of A . The choice of k determines to what extent the SVD of A is regularized, and can be user-specified or chosen adaptively by an algorithm. In this study, k is selected by detecting the first break in the log eigenvalue spectrum of A prior to input into feeding to the RegEM algorithm.

1.2 M09 [*Mann et al., 2009*]

The M09 implementation of TTLS uses a hybrid version of RegEM-TTLS that treats low-frequency and high-frequency signals separately (a 20-year frequency split). The reconstruction is then performed in a forward ‘stepwise’ approach century by century. For each century, only the first M leading modes of surface temperature are retained, determined by the number of degrees of freedom in the proxy network. M09 also uses a semi-adaptive choice for both low-frequency and high-frequency truncation parameters, k_l and k_h respectively. The method selects: 1) k_l to retain 33% of the low-frequency multivariate data variance, and 2) k_h by detecting the first break in the log eigenvalue spectrum of high-frequency multivariate data variance, which are also calculated for each century of reconstruction.

1.3 CCA [*Smerdon et al., 2010*]

Canonical correlation analysis (CCA) is applied as described by *Smerdon et al.* [2010]. The primary goal of CCA is to perform dimensional reductions on \mathbf{T}, \mathbf{P} and \mathbf{B} in Eq 1. CCA first decomposes \mathbf{T} and \mathbf{P} through singular value decomposition (SVD), and only

uses the first few leading modes (d_p, d_t) to estimate the covariance matrix $\hat{\Sigma}$, and then further truncates the eigen decomposition of $\hat{\Sigma}$ (via the selection of d_{cca}) to estimate $\hat{\mathbf{B}}$. The truncation parameters are selected based on "leave-half-out" cross-validation (CV), as described by *Smerdon et al.* [2010]. See Fig S2 for an illustration of the implementation of CCA. The procedure is straightforward, and only depends on the choice of three discrete parameters. Since the set of optimal (d_p, d_t, d_{cca}) is chosen through CV, model parameters for CCA are not set a priori, making the method fully data-adaptive.

Below are some technical notes on speeding up CCA and employing the method in a temporally variant network. Originally, in the study by *Smerdon et al.* [2010], CCA is applied on a temporally invariant network, which means that (d_p, d_t, d_{cca}) only needs to be chosen once over the entire reconstruction period. In reality (and so in our study), proxy availability declines back in time, so the truncation parameters should also be chosen adaptively. In this study, we define patterns of missing values and group proxies with the same pattern together. Here we provide a simple example to illustrate the process. Assume we know that 35 proxies are available until A.D. 1000, when estimating the set of truncation parameters, only temperatures and proxies over the calibration period are used. In this example, only these 35 proxies will be used for the estimation.

To speed up the computation, we also modified the original code (`predccabp.m`) provided in the supplementary info on *Smerdon et al.* [2010]. Instead of calculating SVD for \mathbf{T} and \mathbf{P} every time when performing a CCA reconstruction in the CV process, we apply SVD only once before using CV to choose the optimal set of parameters.

The code is available at the website (<https://code.google.com/p/common-climate/>).

1.4 GraphEM [*Guillot et al.*, submitted]

The GraphEM algorithm benefits from recent advances in high-dimensional statistics by using a Gaussian graphical model [GGMs, a.k.a. Markov random fields, *Whittaker*, 1990; *Lauritzen*, 1996] to estimate the covariance structure of climate fields (Σ , a critical ingredient of all CFR methods). GGMs provide a natural and flexible framework for modeling the inherent spatial heterogeneities of high-dimensional spatial fields, which would in general be more difficult with standard parametric covariance models. At the same time, they provide the necessary parameter reduction for obtaining precise and well-conditioned estimates of Σ , even when the sample size is much smaller than the number of variables (as is typically the case in paleoclimate applications).

In the case of multivariate normal data, two variables are conditionally independent given the rest of the variables if and only if the corresponding element of the inverse covariance matrix $\Omega_{ij} = (\Sigma^{-1})_{ij}$ is equal to zero. Thus, when many variables are conditionally independent (as is typically the case in climate fields), the inverse covariance matrix Ω is sparse, and the number of parameters to estimate is greatly reduced.

In GraphEM, the conditional independence structure of the climate field is first estimated using an ℓ_1 -penalized maximum likelihood method (graphical lasso). Once the conditional independence structure is known, the covariance matrix Σ can be estimated in accordance with these conditional independence relations within the EM algorithm (see Fig S1 for an illustration of the GraphEM algorithm).

As described in *Guillot et al.* [submitted], when estimating the graphical structure of

the field, three parameters need to be specified to determine the target sparsity of the graph for the temperature-temperature part (**TT**), proxy-proxy part (**PP**), and the temperature-proxy part (**TP**). For all three parameters, a value that is large enough has to be specified so that the true graph is contained in the estimated one. On the other hand, the sparsity has to be small enough for the covariance matrix to be well-conditioned. For this study, the parameters have been set to $(\mathbf{TT}, \mathbf{PP}, \mathbf{TP}) = (5\%, 5\%, 5\%)$, following the choices in [Guillot *et al.*, submitted]. Future studies will explore objective criteria for choosing these parameters.

2 Results

2.1 Spatial metrics

Figs. S3 - S12 below plot the spatiotemporal variations of CE. The figures are based on each ensemble median, using the flat network.

2.2 Global average

Table S1 and S2 summarize the weighted global mean ($\hat{\mu}$) verification statistics (CE, RE, R^2) for each century, with the estimated standard deviation ($\hat{\sigma}$) displayed in parentheses.

By definition, CE is the most stringent diagnostic. In Table S2 and S1, it is also evident that reconstruction skill based on CE indicates the worst results. The RE statistic, which is less stringent than CE, indicates that reconstructions are skillful (if $RE > 0$) for periods after A.D 1300 in the local SNR case, and remain skillful during the entire reconstruction interval for all methods in the max SNR network. R^2 is the least strict statistic because it only accounts for similarity between two time series by trend. As in Table S2, $R^2 > 0$ for all methods in all cases. With R^2 as the judging criterion, CCA stands out in the max SNR network, which is consistent with the mechanisms of the CCA algorithm, i.e. maximize the cross-correlation between proxies and temperature time series.

2.3 Additional diagnostics

We also examined the relationship between:

1. CE vs. number of proxies per grid box (Figs. S13 and S14);
2. CE vs. average SNR per grid box (Figs. S15 and S16);
3. CE vs. sum of SNR per grid box (Figs. S17 and S18).

The results indicate that there are no simple apparent relationships between these variables of interest.

2.4 Error cancellation in index reconstructions

Following the approach of *Guillot et al.* [submitted], we consider the positive and negative contributions to the error in the reconstructed global mean temperature index (Figs. S19 and S20). For every spatial point s in a given year t , we compute the deviation from the true field $\hat{T}(t, s) - T(t, s)$, where \hat{T} is the reconstructed temperature and T is the target (here the GCM-simulated temperature). Such deviations are then partitioned into positive and negative contributions, and the area-weighted average of each is computed to form the timeseries displayed in Figs. S19 and S20.

This allows decomposition of the sum of reconstruction errors (black curves) into contributions stemming from regions of warm or cool biases. In an ideal reconstruction, the deviations should be small everywhere. In practice, however, reconstructions are often biased warm ($\hat{T}(t, s) > T(t, s)$), but in some locations are biased cold ($\hat{T}(t, s) < T(t, s)$). As long as the positive values are comparable with negative values, the value of the global mean deviation $\langle \hat{T}(t, s) - T(t, s) \rangle$, which is a simple linear combination of the positive and negative values, will be small, and so will the RMSE obtained from it.

As shown on Figs. S19 and S20, similar positive errors arise from all methods, but negative errors arising from RegEM-TTLS and M09 are much higher than those from CCA and GraphEM. Because the absolute values of RegEM-TTLS and M09's negative deviations are comparable with their positive deviations, their overall sum is smaller than those arising from CCA and GraphEM. Thus we have the counterintuitive circumstance that larger spatial deviations result in smaller overall errors when composite indices (e.g. global or hemispheric means) are computed. This finding has important implications for interpreting CFR results, showing in particular that the global mean temperature reconstruction is a poor indicator of spatial skill.

References

- Fierro, R. D., G. H. Golub, P. C. Hansen, and D. P. O' Leary, Regularization by truncated total least squares, *SIAM J. Sci. Comput.*, 18, 1223–1241, 1997.
- Guillot, D., B. Rajaratnam, and J. Emile-Geay, Statistical paleoclimate reconstructions via markov random fields, *Ann. Appl. Stat.*, submitted.
- Lauritzen, S. L., *Graphical Models*, Oxford: Clarendon Press., 1996.
- Mann, M. E., Z. Zhang, S. Rutherford, R. S. Bradley, M. K. Hughes, D. Shindell, C. Ammann, G. Faluvegi, and F. Ni, Global signatures and dynamical origins of the little ice age and medieval climate anomaly, *Science*, 326(5957), 1256–1260, 2009.
- Schneider, T., Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values., *J. Clim.*, 14, 853–871, 2001.
- Smerdon, J. E., A. Kaplan, D. Chang, and M. N. Evans, A pseudoproxy evaluation of the cca and regem methods for reconstructing climate fields of the last millennium*, *Journal of Climate*, 23(18), 4856–4880, 2010.

Whittaker, J., *Graphical Models in Applied Multivariate Statistics*, John Wiley and Sons, Chichester, United Kingdom, 1990.

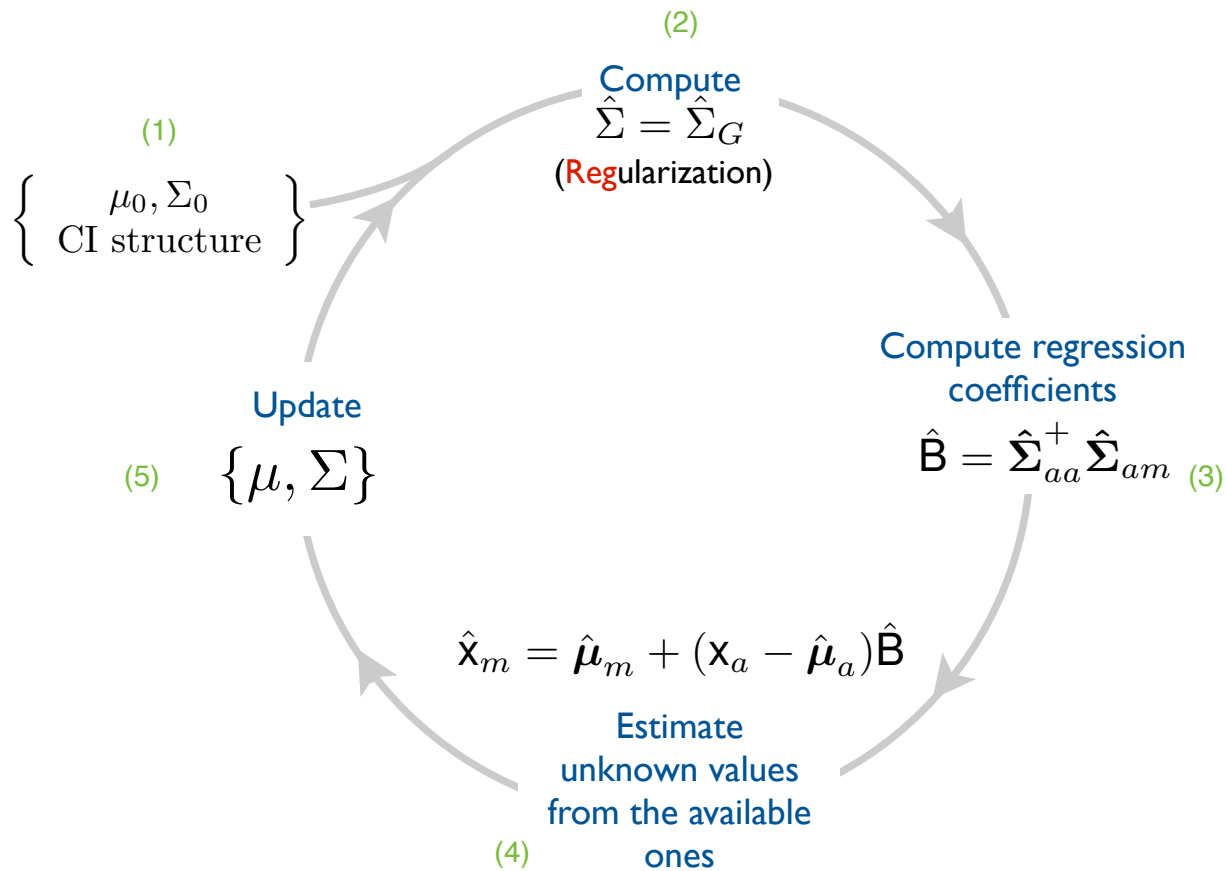


Figure S1: A summary of the GraphEM algorithm. Iterates until convergence. From *Guillot et al.* [submitted]

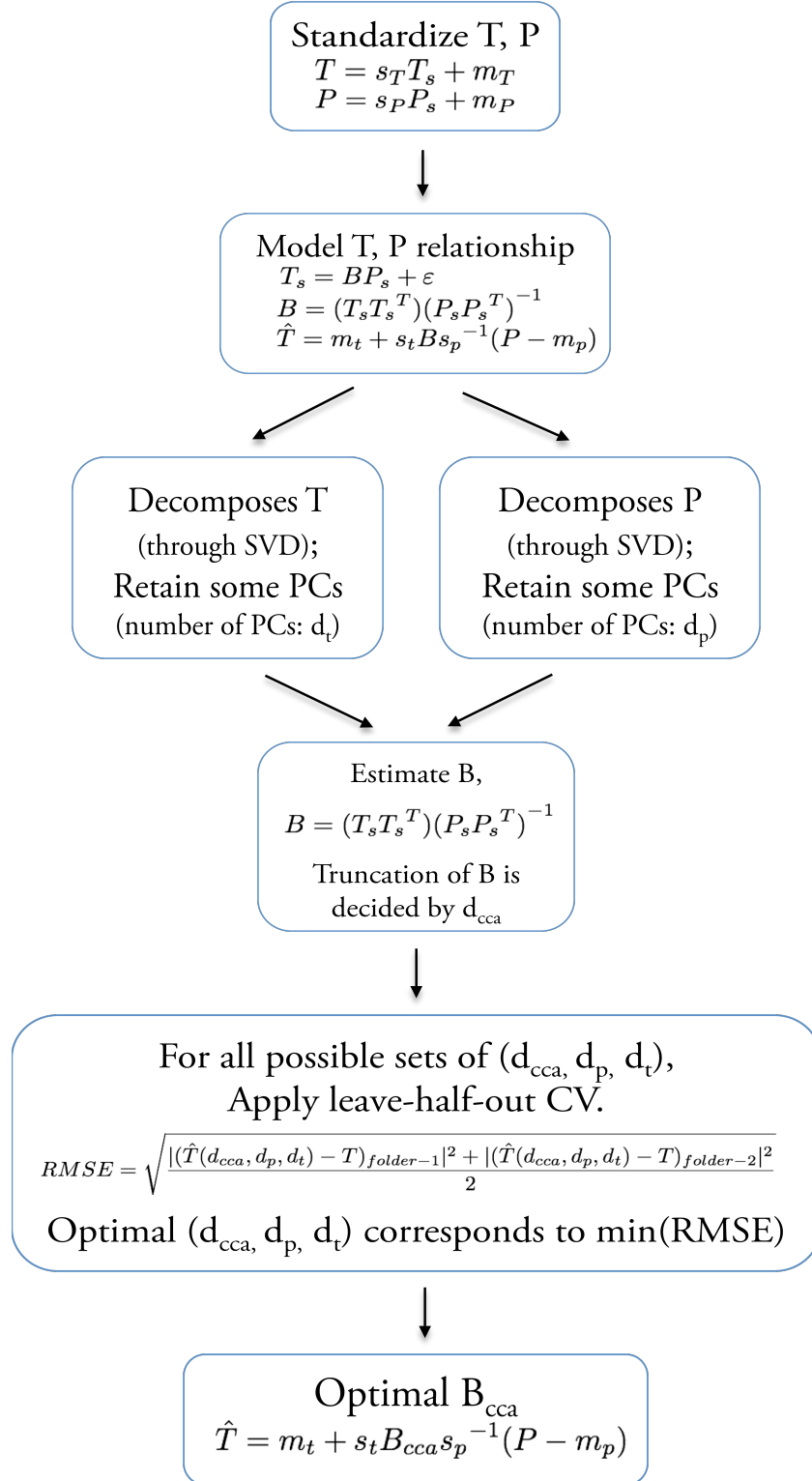


Figure S2: A summary of the CCA algorithm. Reproduced from *Smerdon et al.* [2010]. T_s, P_s are temperature T and proxy P standardized over the temporal dimension spanning the calibration interval, with mean m_T, m_P and standard deviation s_T, s_P , respectively. B is the regression coefficient matrix.

RegEM-TTLS, flat network, local SNR

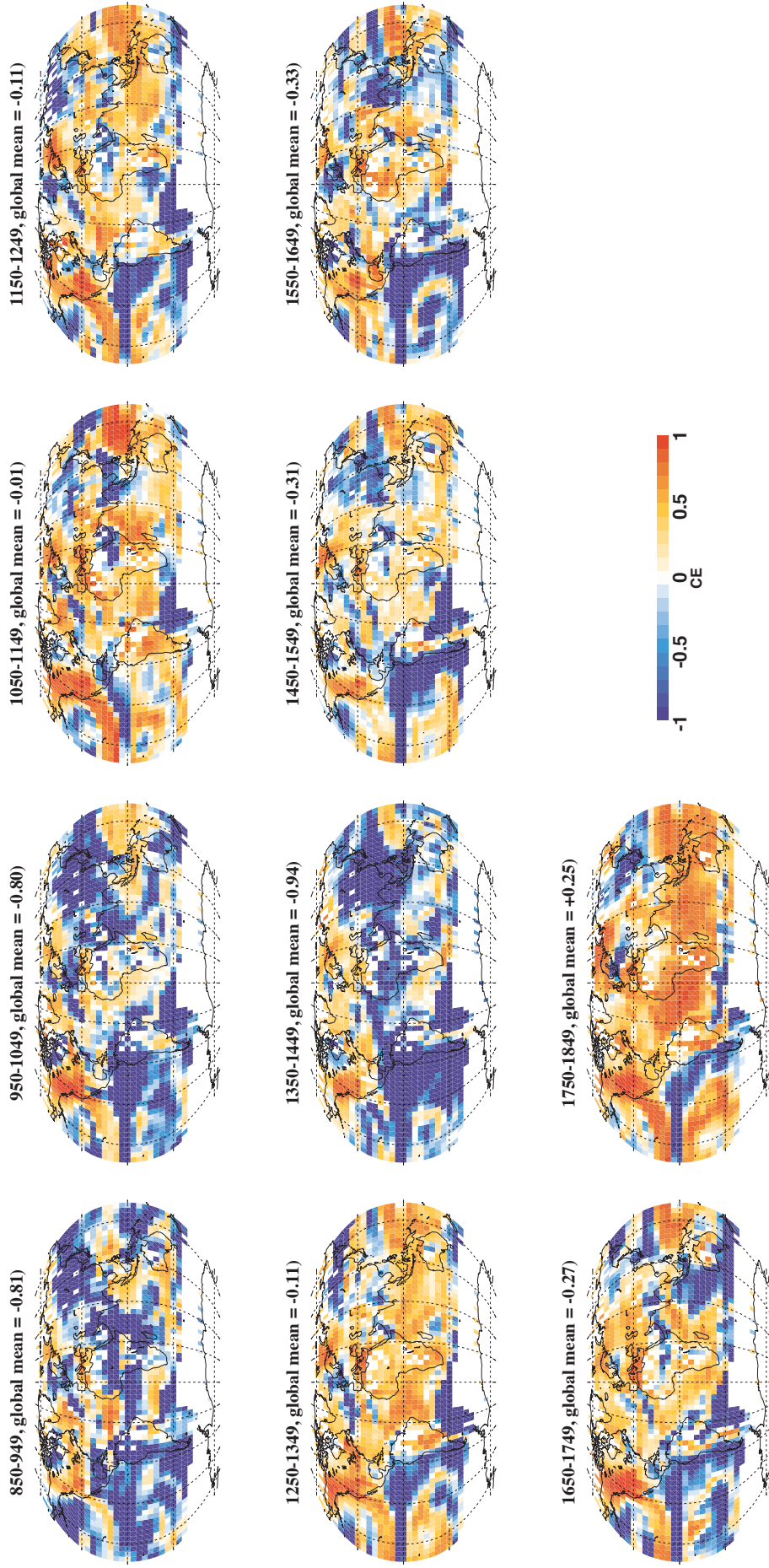


Figure S3: Spatiotemporal variation of RegEM-TTLS reconstructions, using the local SNR, flat network, based on the median of 100-member ensemble.

RegEM-TTLS, flat network, max SNR

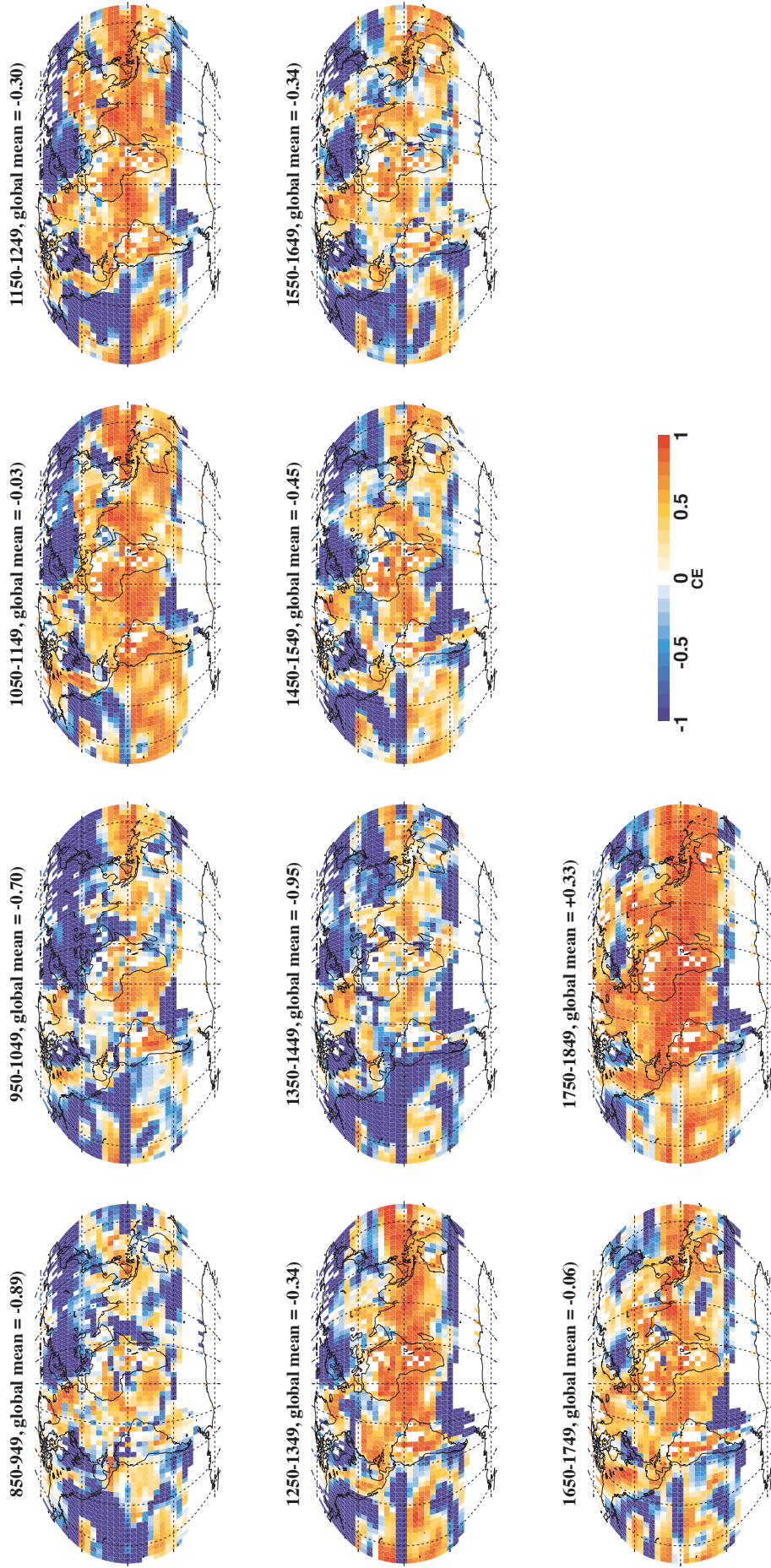


Figure S4: Spatiotemporal variation of RegEM-TTLS reconstructions, using the max SNR, flat network, based on the median of 100-member ensemble.

M09, flat network, local SNR

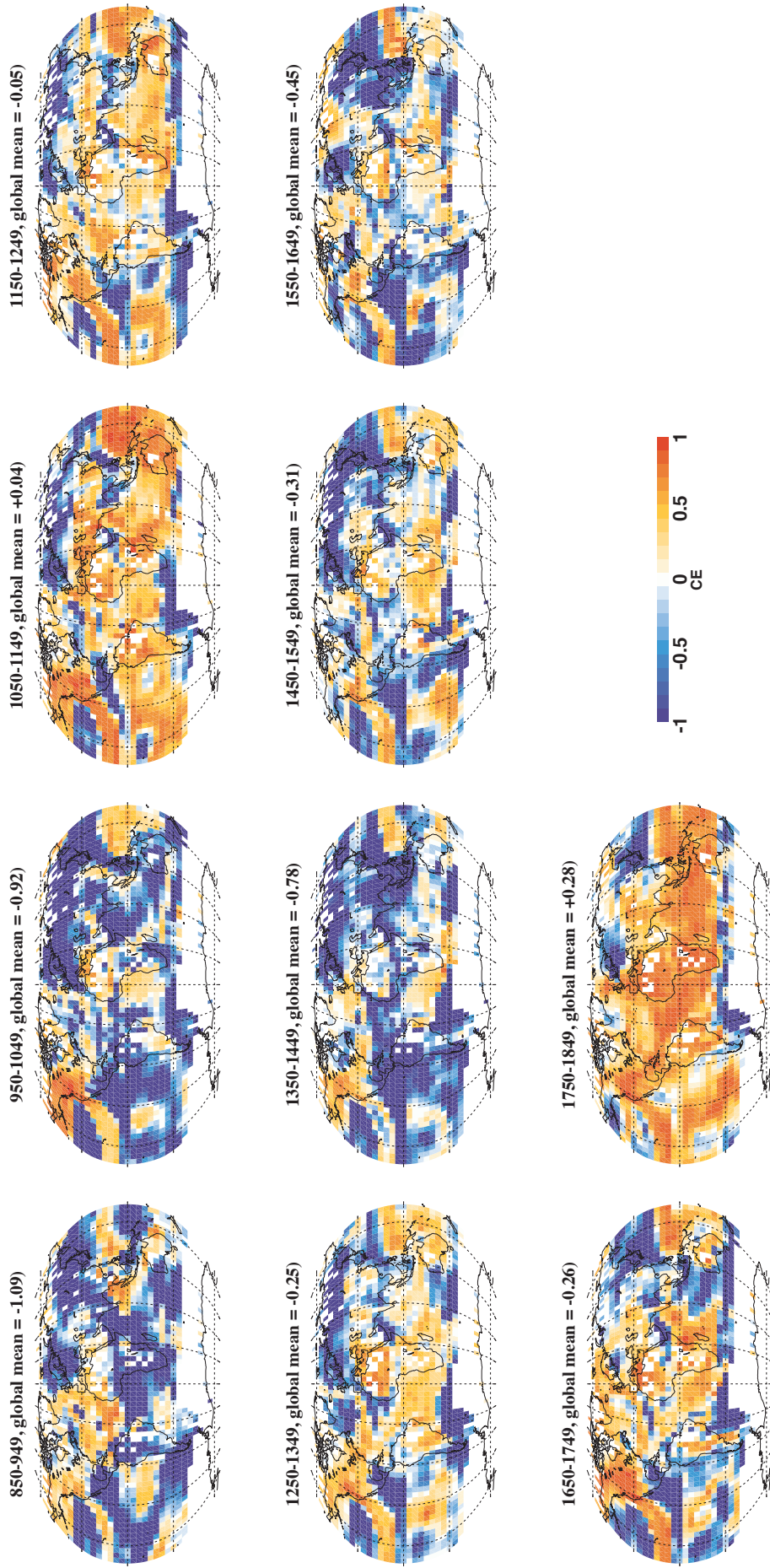


Figure S5: Spatiotemporal variation of M09 reconstructions, using the local SNR, flat network, based on the median of 100-member ensemble.

M09, flat network, max SNR

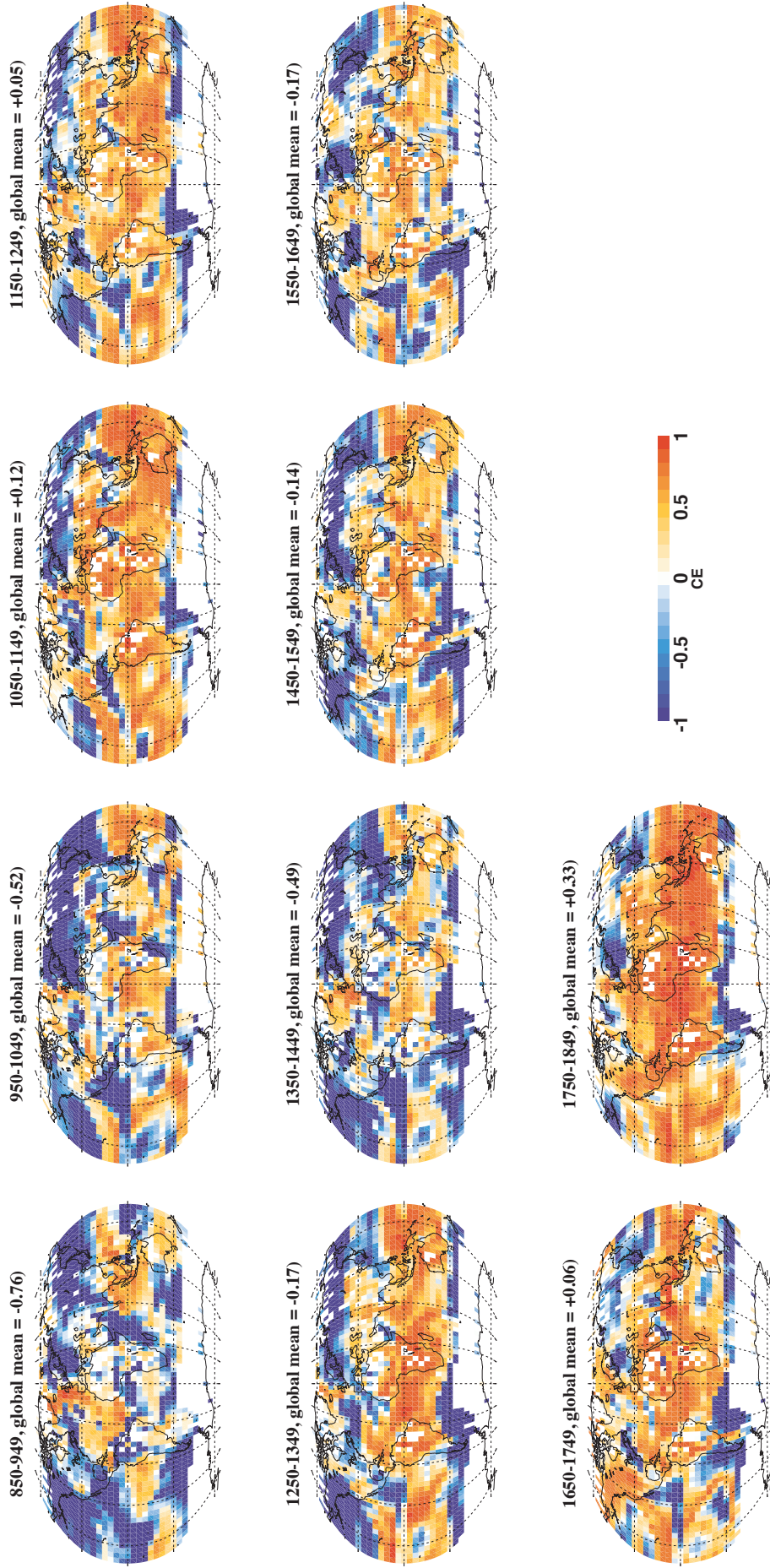


Figure S6: Spatiotemporal variation of M09 reconstructions, using the max SNR, flat network, based on the median of 100-member ensemble.

CCA, flat network, local SNR

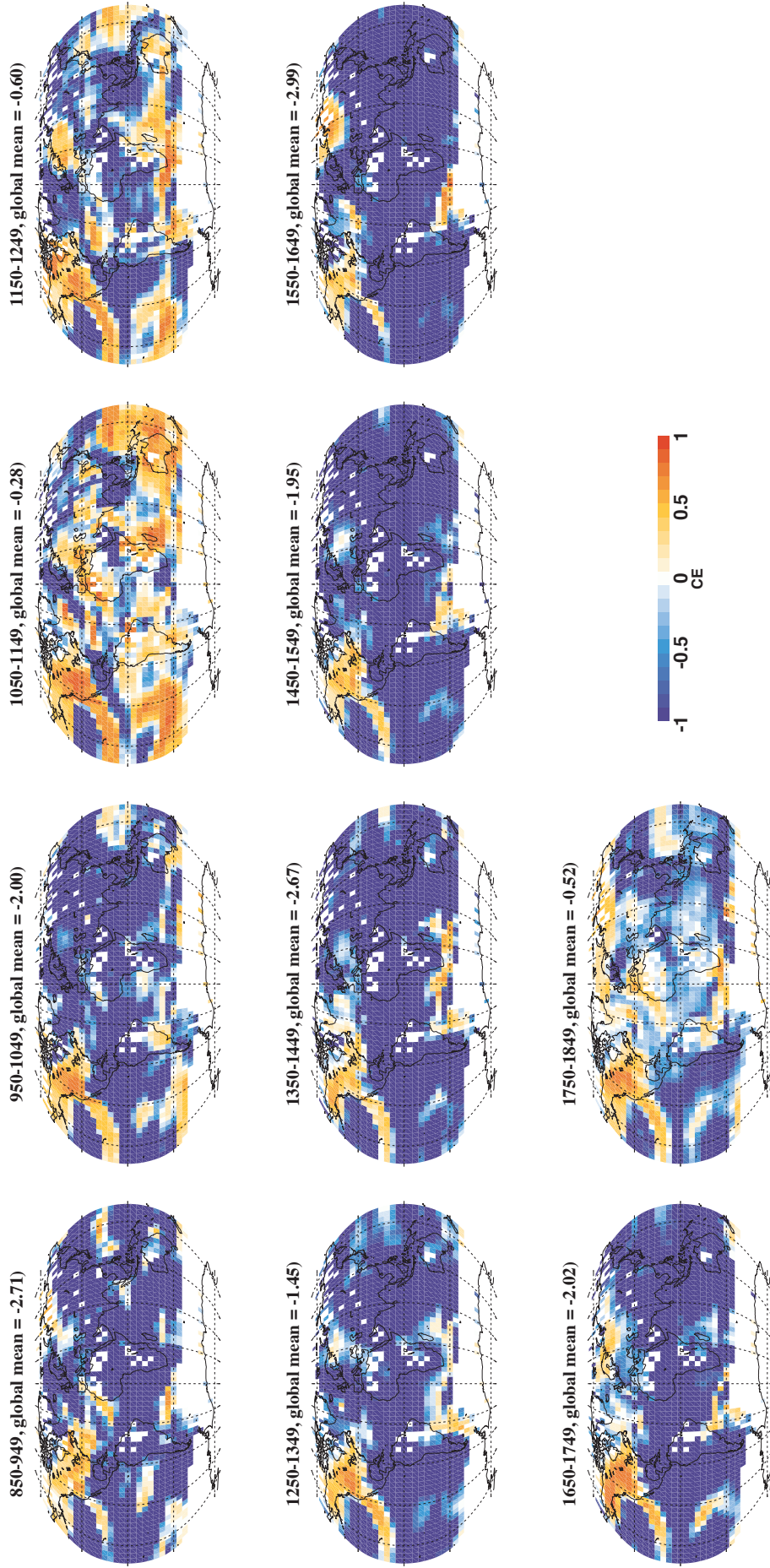


Figure S7: Spatiotemporal variation of CCA reconstructions, using the local SNR, flat network, based on the median of 100-member ensemble.

CCA, flat network, max SNR

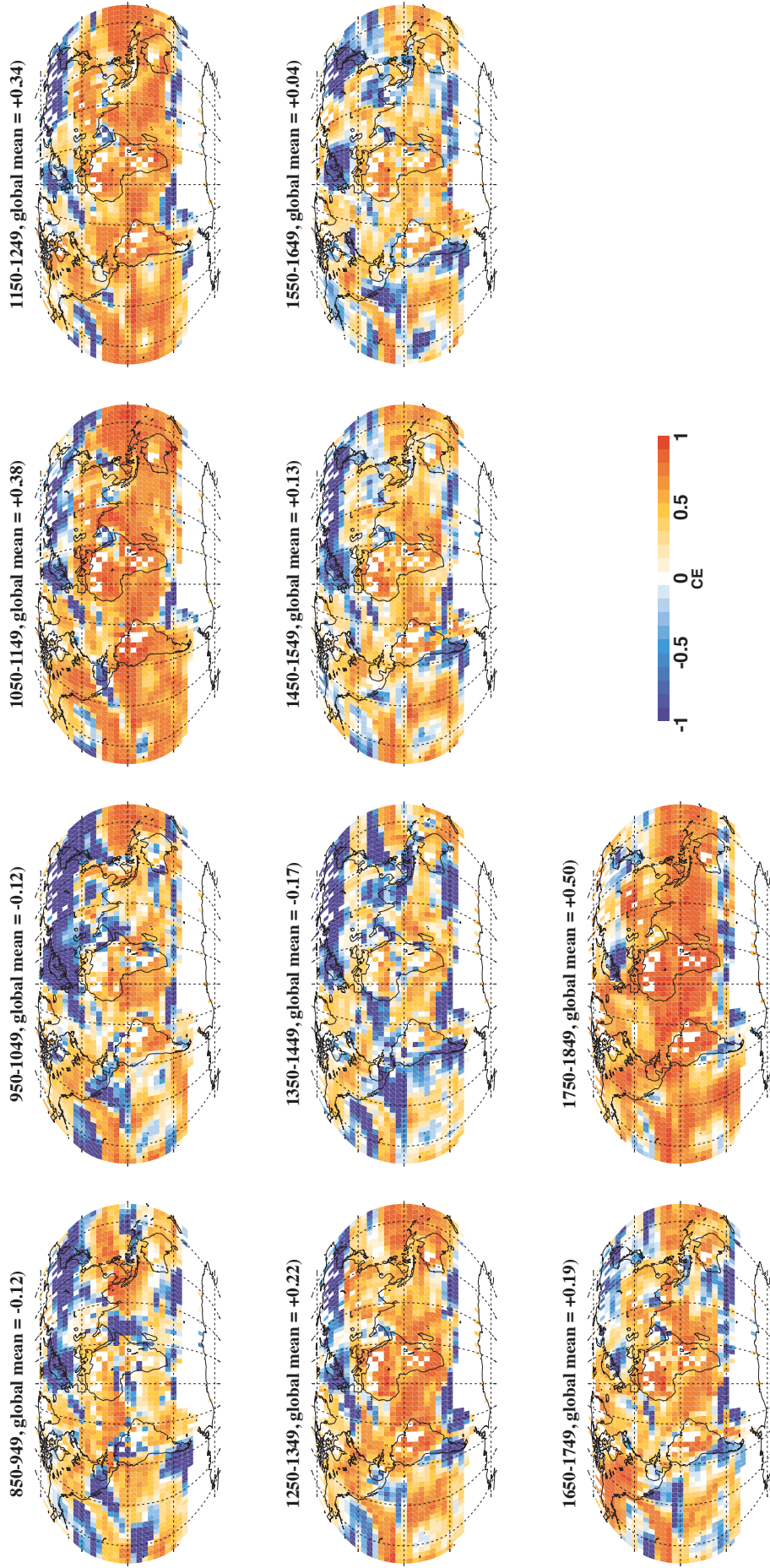


Figure S8: Spatiotemporal variation of CCA reconstructions, using the max SNR, flat network, based on the median of 100-member ensemble.

GraphEM, flat network, local SNR

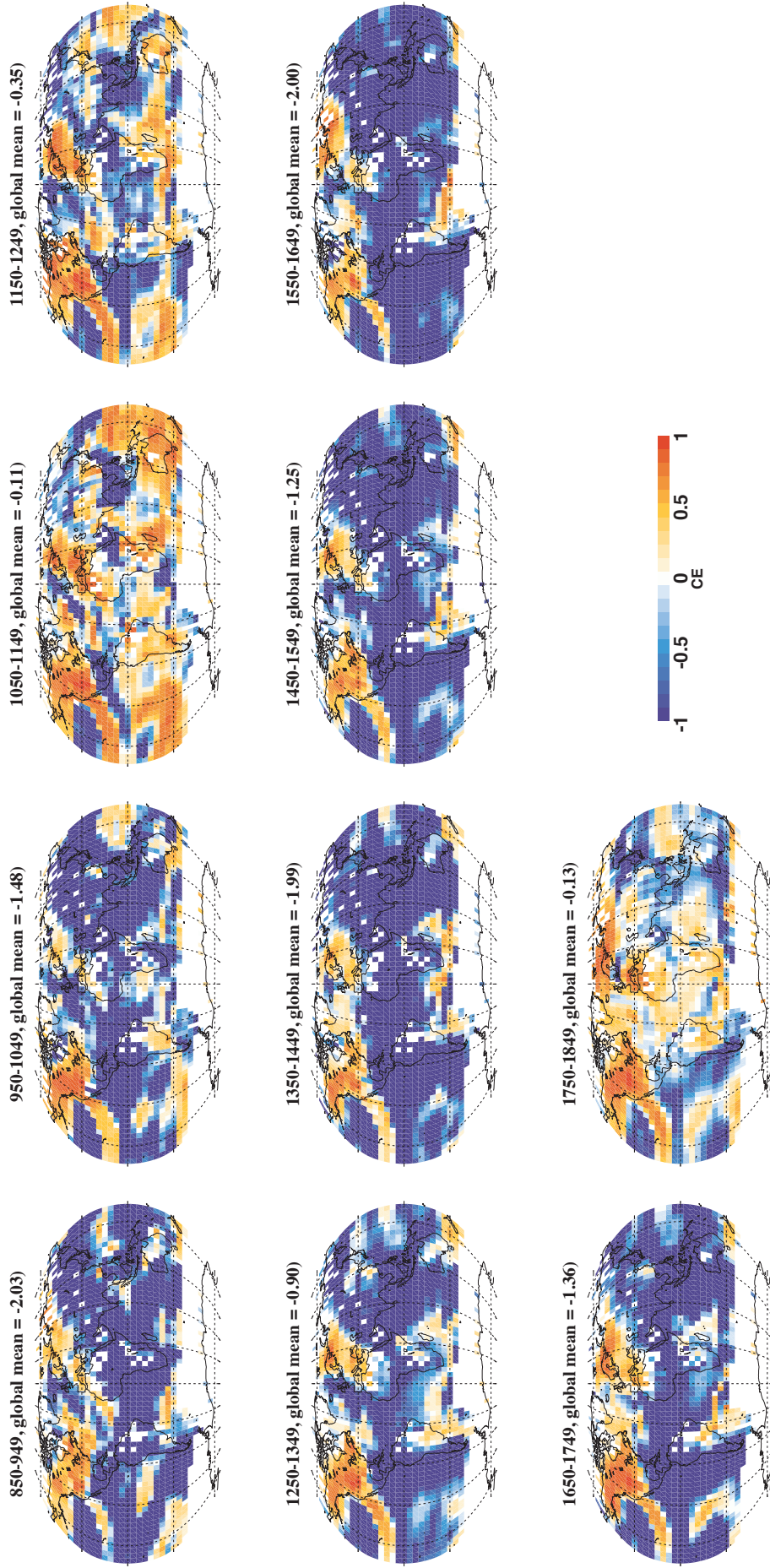


Figure S9: Spatiotemporal variation of GraphEM reconstructions, using the local SNR, flat network, based on the median of 100-member ensemble.

GraphEM, flat network, max SNR

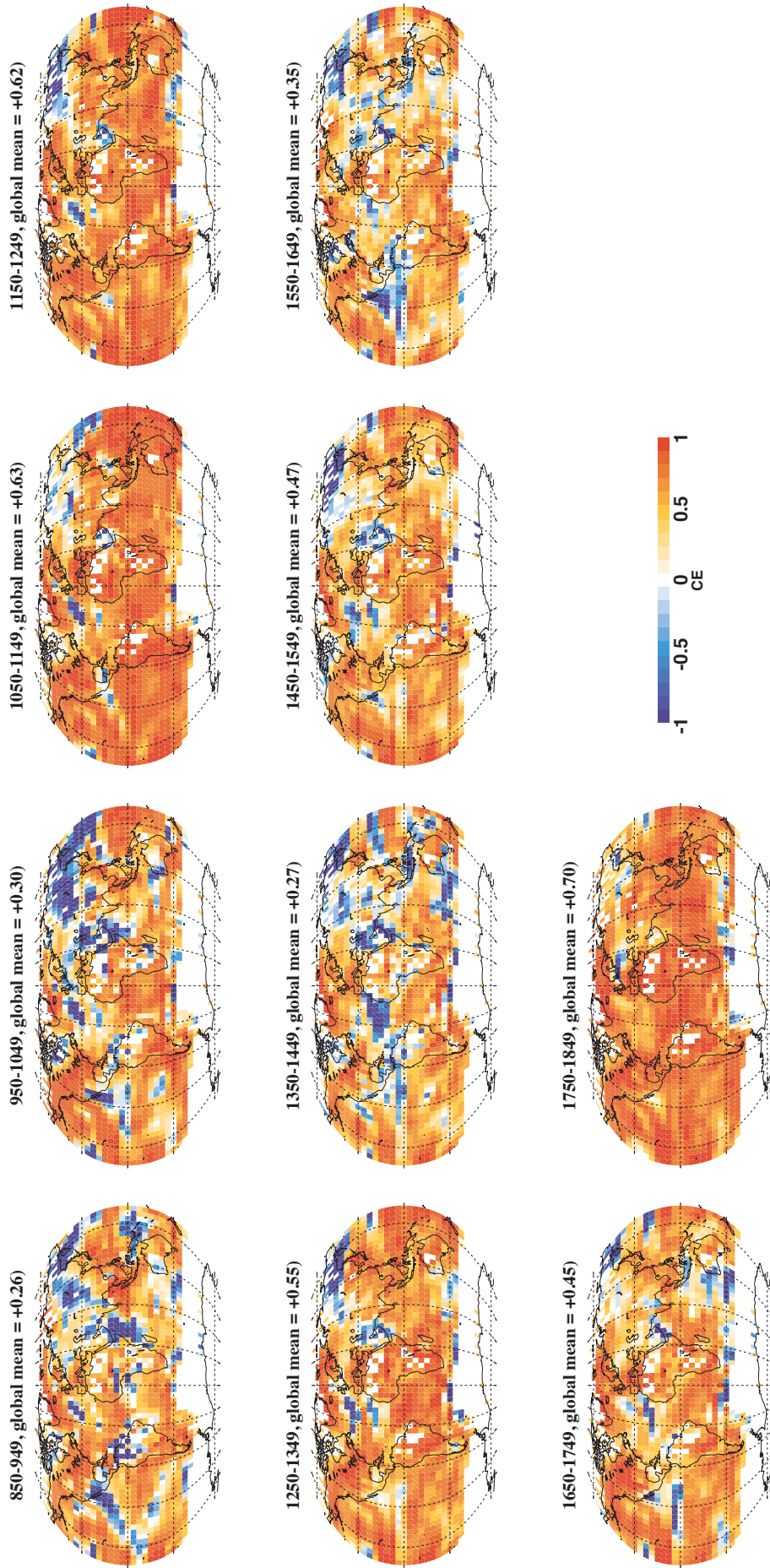


Figure S10: Spatiotemporal variation of GraphEM reconstructions, using the max SNR, flat network, based on the median of 100-member ensemble.

GraphEM, flat network, SNR = 0.25

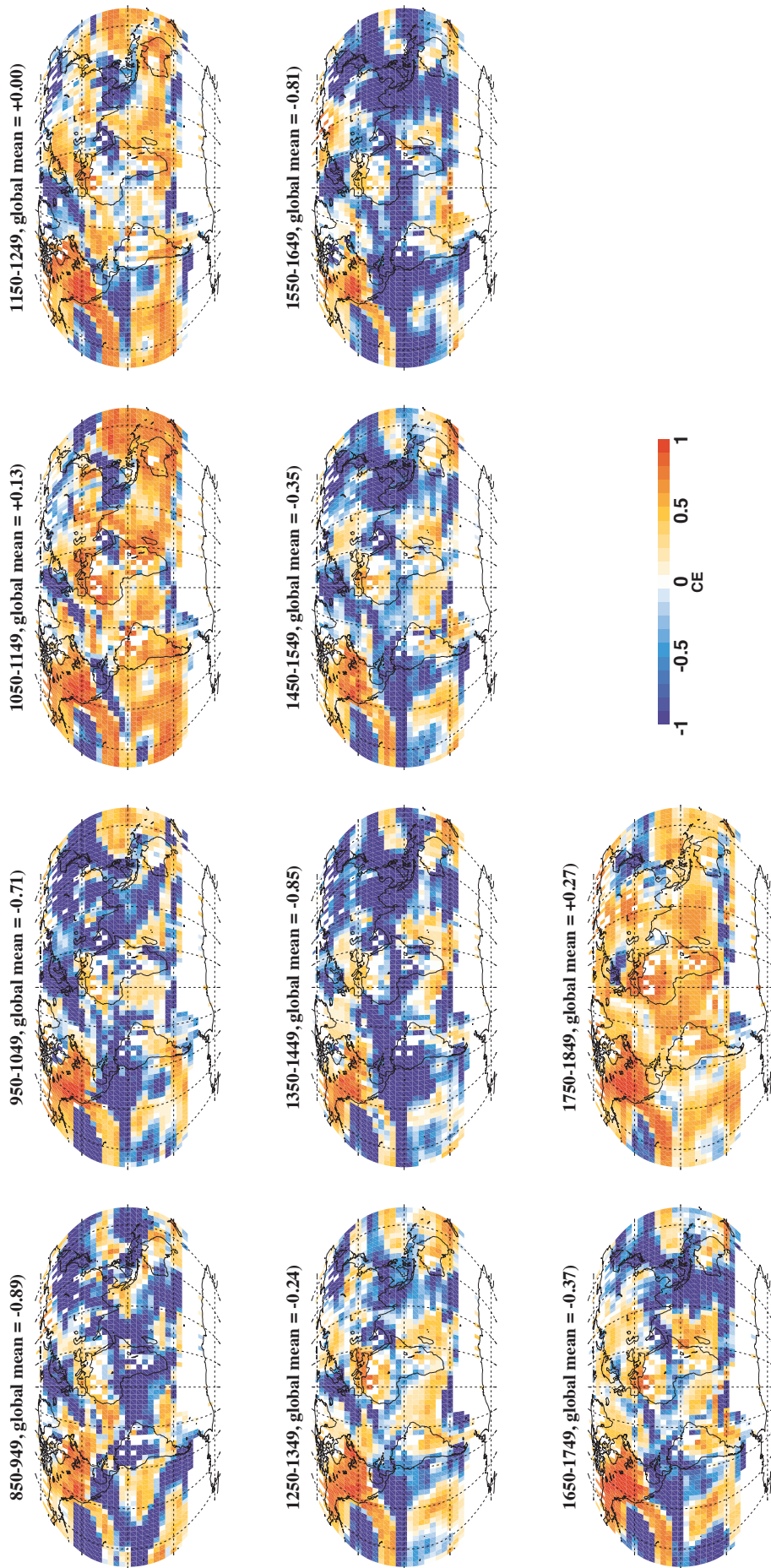


Figure S11: Spatiotemporal variation of GraphEM reconstructions, using the SNR = 0.25, flat network, based on the median of 100-member ensemble.

GraphEM, flat network, SNR = 0.5

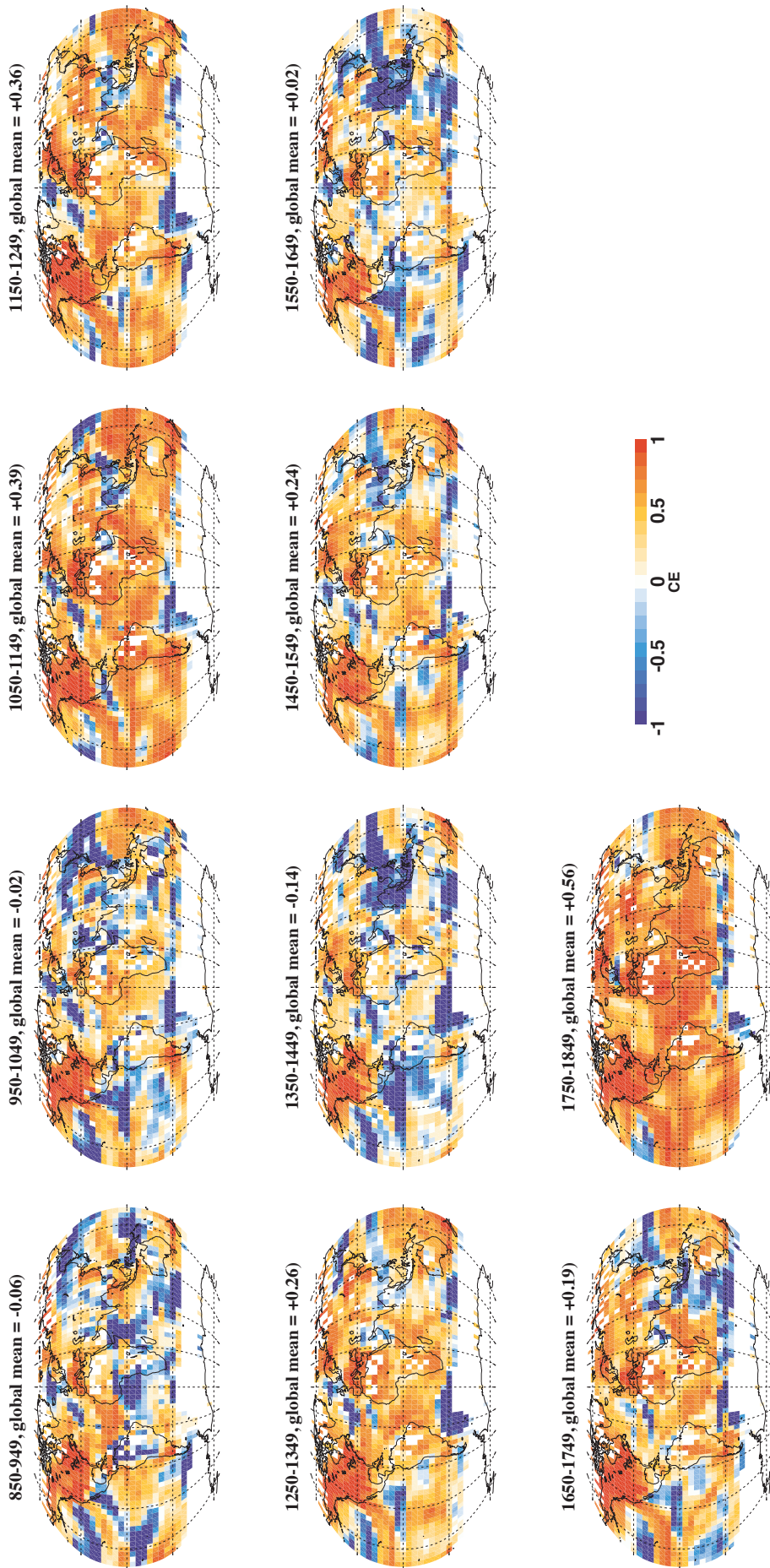


Figure S12: Spatiotemporal variation of GraphEM reconstructions, using the SNR = 0.5, flat network, based on the median of 100-member ensemble.

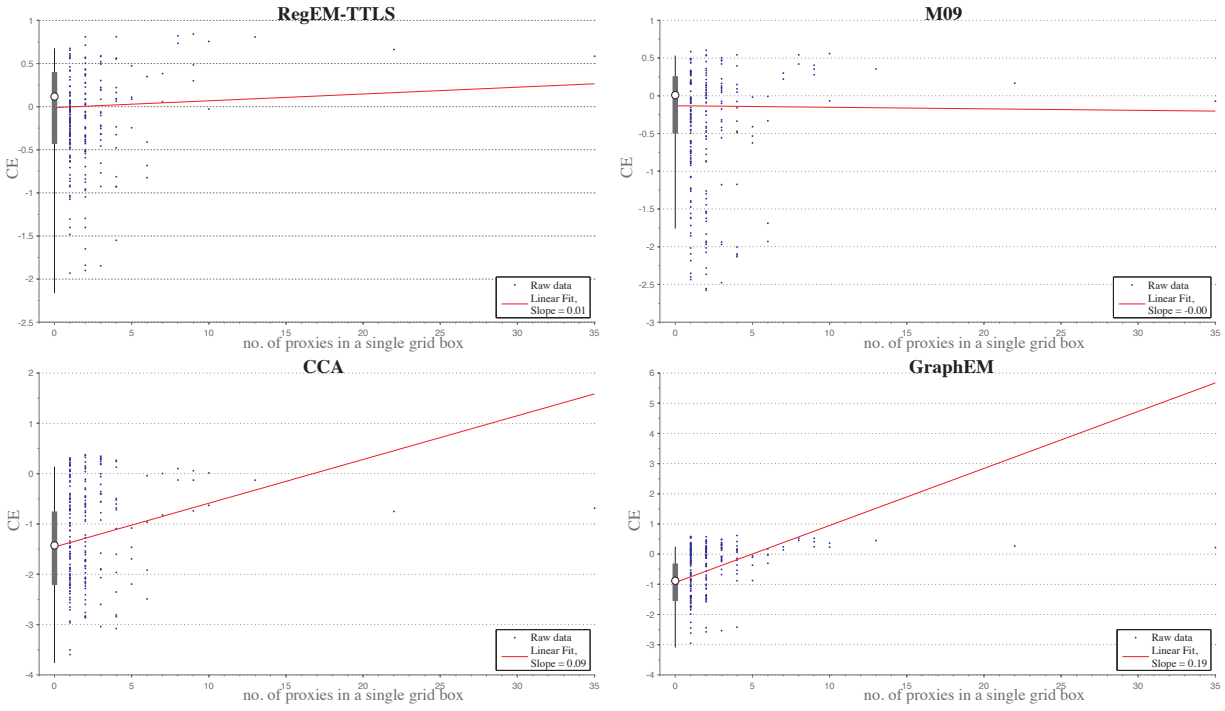


Figure S13: Relationship between CE and number of proxies per grid box, using the staircase network with local SNR.

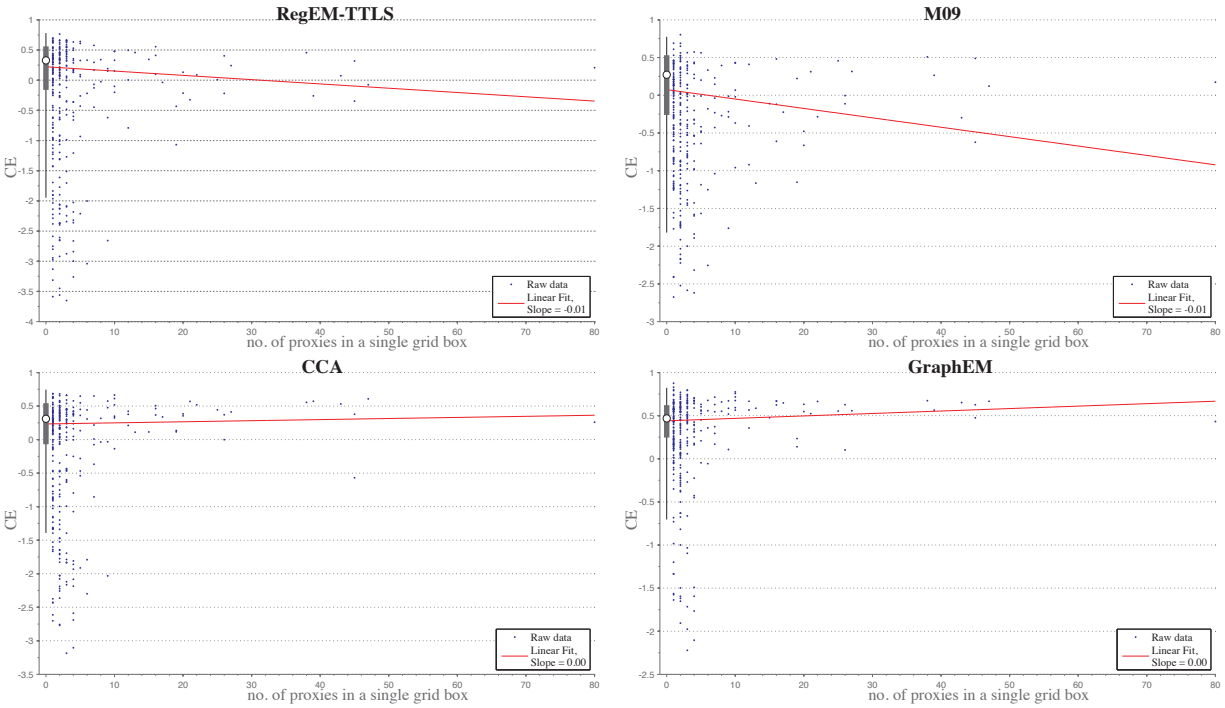


Figure S14: Relationship between CE and number of proxies per grid box, using the staircase network with max SNR.

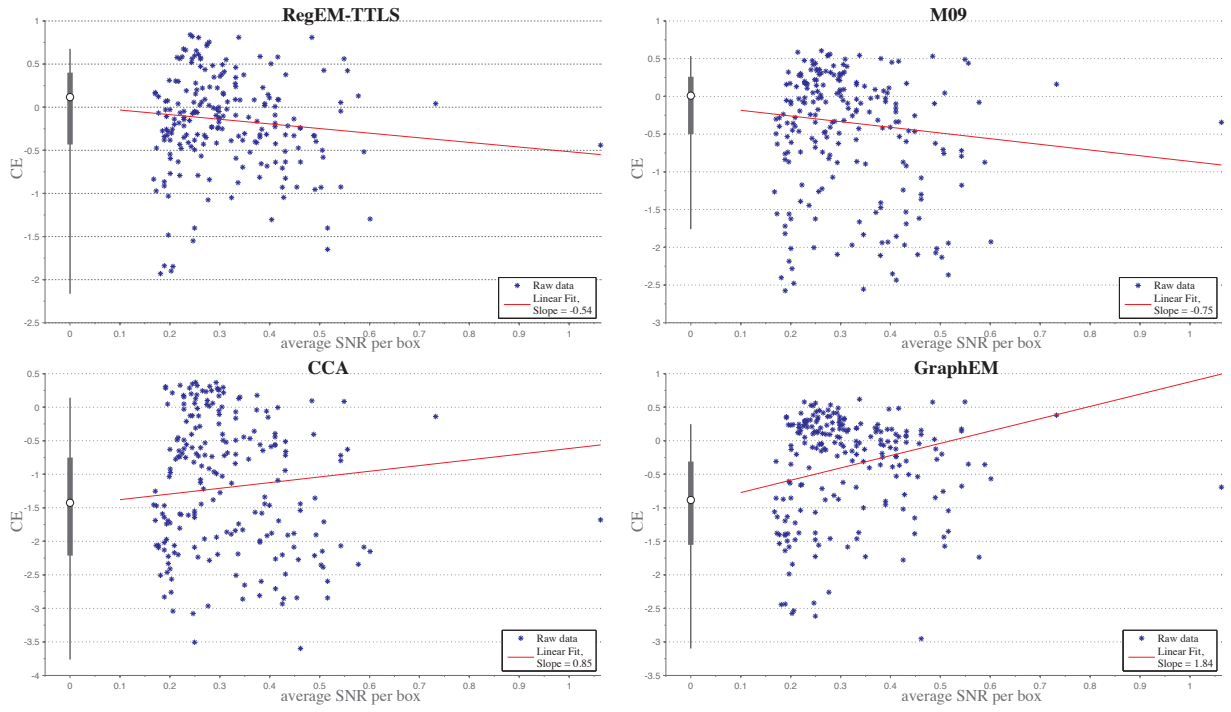


Figure S15: Relationship between CE and average SNR per grid box, using the staircase network with local SNR.

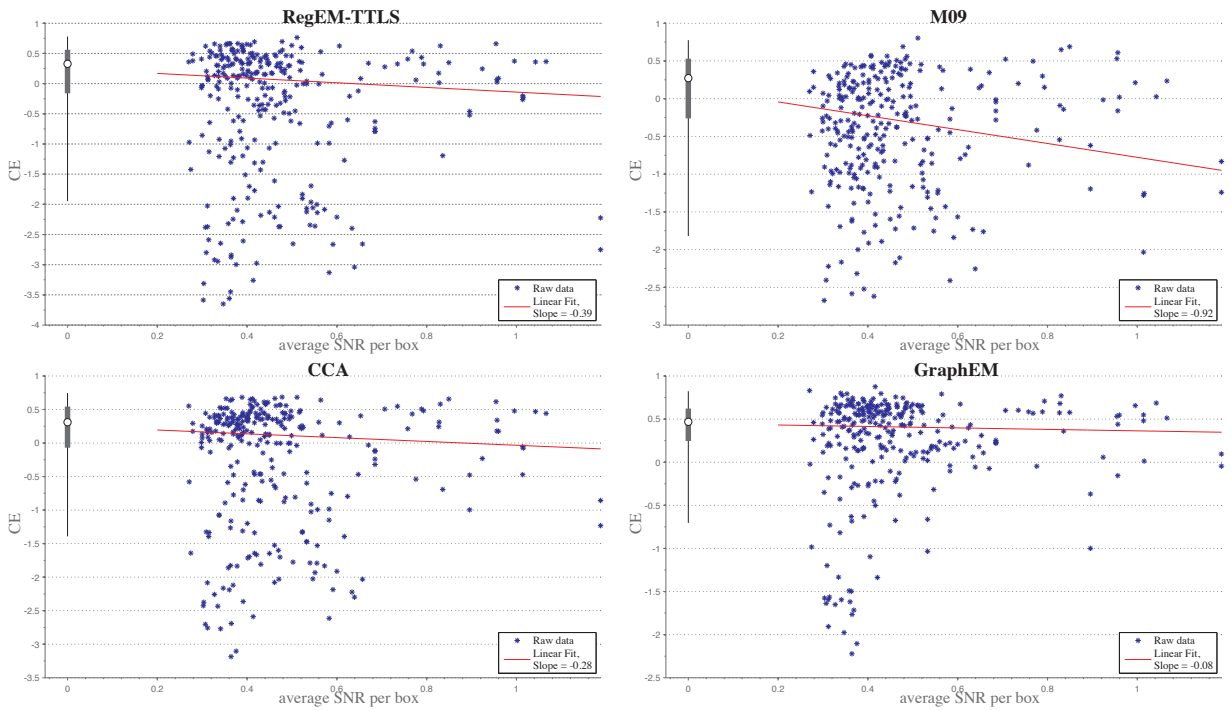


Figure S16: Relationship between CE and average SNR per grid box, using the staircase network with max SNR.

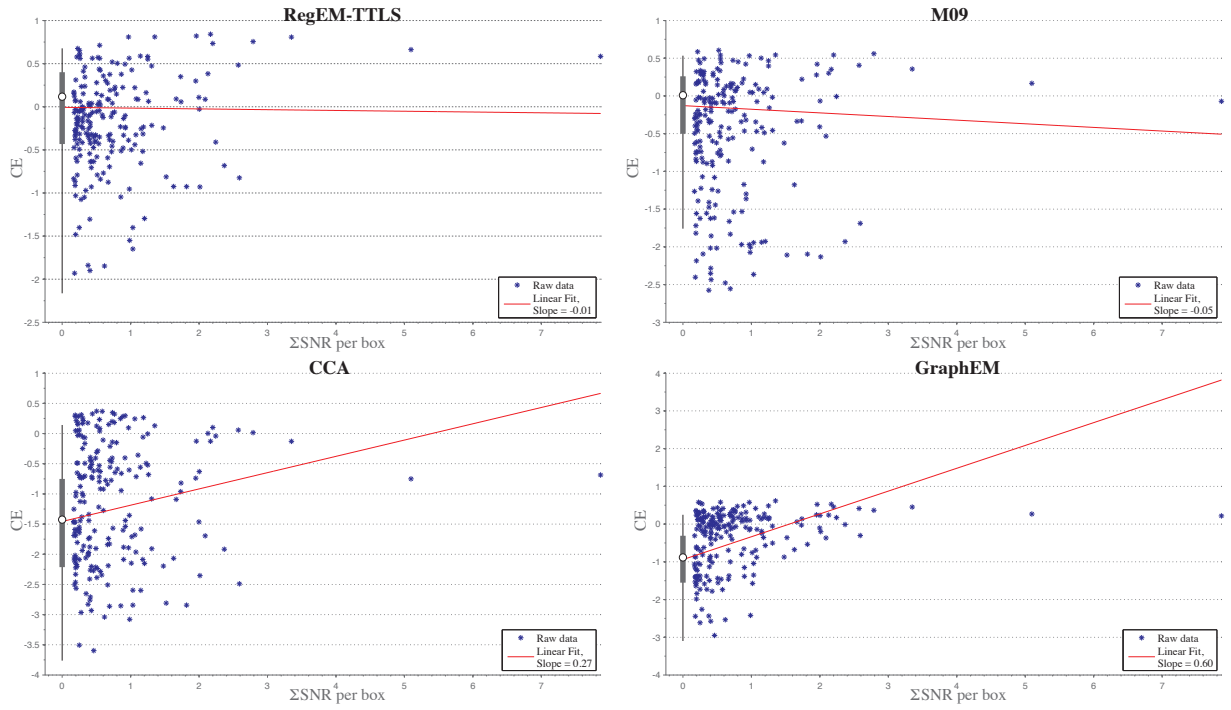


Figure S17: Relationship between CE and sum of SNR per grid box, using the staircase network with local SNR.

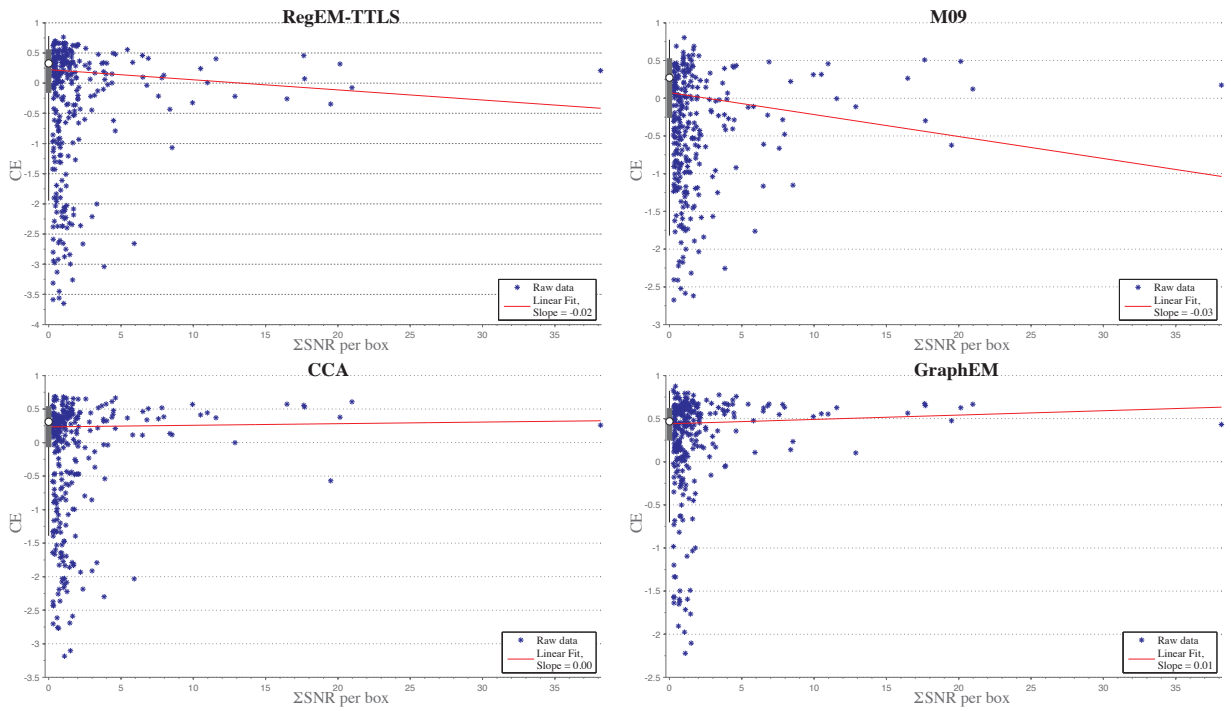


Figure S18: Relationship between CE and sum of SNR per grid box, using the staircase network with max SNR.

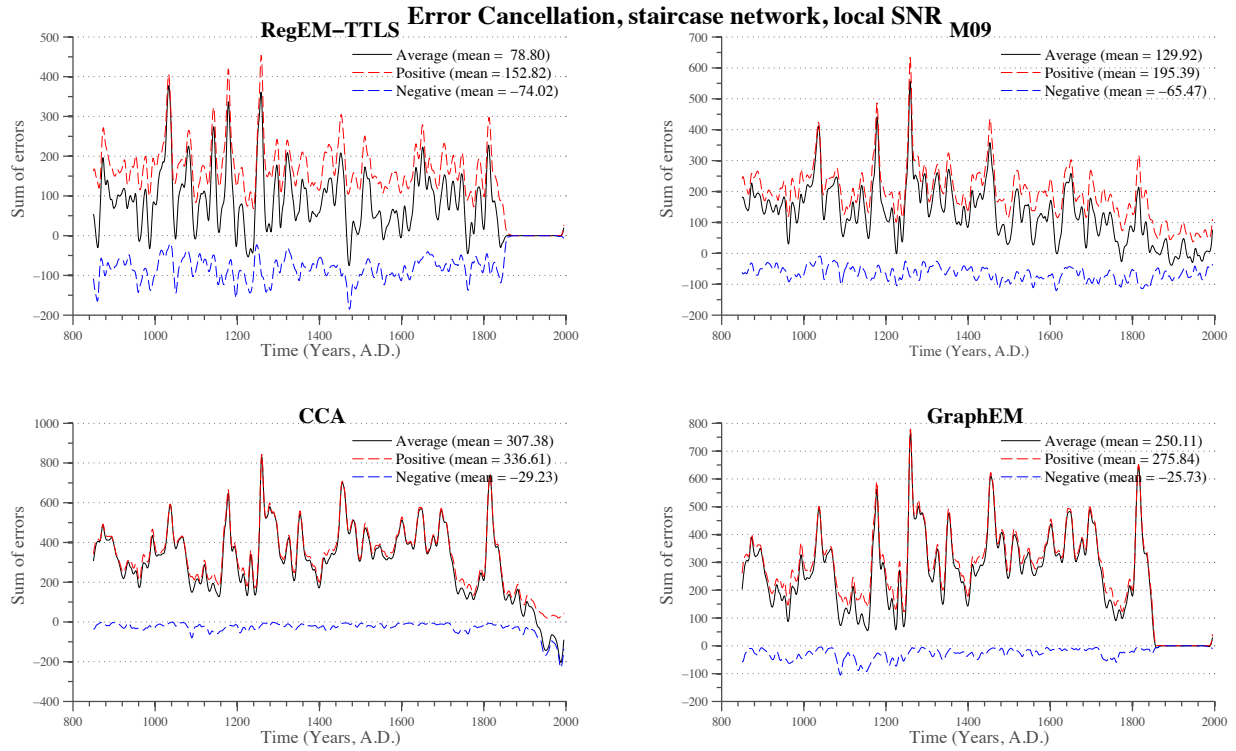


Figure S19: Error cancellation analysis, using the staircase network with local SNR.

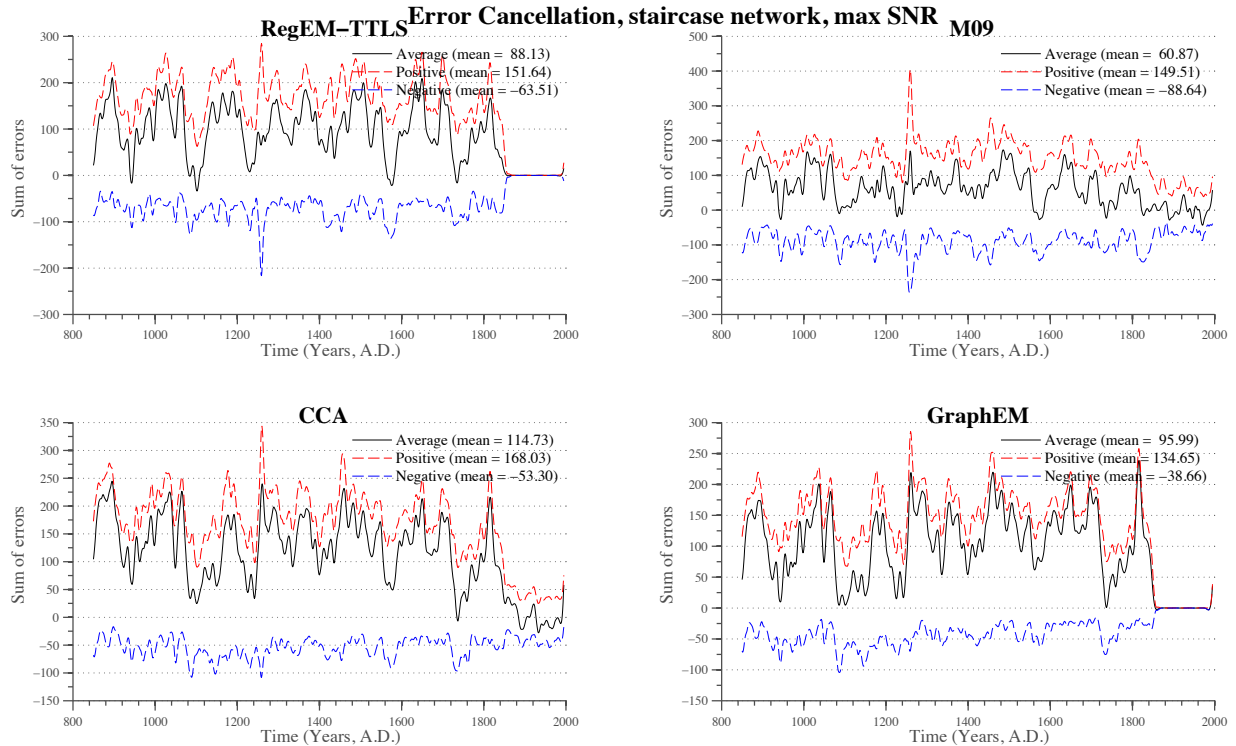


Figure S20: Error cancellation analysis, using the staircase network with max SNR.

SNR	Metric	900	1000	1100	1200	1300	1400	1500	1600	1700	1800
RegEM - TTLS											
local	CE	-1.51 _(0.49)	-1.31 _(0.34)	-0.31 _(0.19)	-0.46 _(0.23)	-0.42 _(0.18)	-1.48 _(0.35)	-0.67 _(0.23)	-0.82 _(0.28)	-0.66 _(0.26)	+0.02 _(0.14)
	RE	+0.42 _(0.10)	+0.35 _(0.08)	+0.09 _(0.16)	+0.13 _(0.14)	+0.54 _(0.06)	+0.38 _(0.07)	+0.53 _(0.06)	+0.65 _(0.04)	+0.61 _(0.05)	+0.59 _(0.05)
	R ²	+0.20 _(0.03)	+0.19 _(0.03)	+0.28 _(0.04)	+0.31 _(0.04)	+0.39 _(0.06)	+0.19 _(0.03)	+0.21 _(0.04)	+0.25 _(0.03)	+0.47 _(0.04)	+0.53 _(0.05)
max	CE	-1.34 _(0.48)	-1.01 _(0.22)	-0.33 _(0.18)	-0.65 _(0.30)	-0.65 _(0.21)	-1.60 _(0.38)	-0.80 _(0.26)	-0.60 _(0.18)	-0.27 _(0.13)	+0.16 _(0.10)
	RE	+0.35 _(0.12)	+0.39 _(0.08)	+0.07 _(0.13)	-0.00 _(0.17)	+0.45 _(0.07)	+0.27 _(0.10)	+0.46 _(0.07)	+0.64 _(0.04)	+0.63 _(0.04)	+0.63 _(0.05)
	R ²	+0.24 _(0.03)	+0.29 _(0.02)	+0.36 _(0.02)	+0.43 _(0.03)	+0.49 _(0.02)	+0.30 _(0.02)	+0.36 _(0.02)	+0.29 _(0.02)	+0.47 _(0.02)	+0.63 _(0.02)
M09											
local	CE	-1.76 _(1.08)	-1.47 _(0.94)	-0.29 _(0.37)	-0.49 _(0.45)	-0.56 _(0.37)	-1.30 _(0.60)	-0.65 _(0.40)	-0.77 _(0.41)	-0.47 _(0.21)	+0.09 _(0.15)
	RE	+0.45 _(0.16)	+0.37 _(0.17)	+0.24 _(0.19)	+0.19 _(0.20)	+0.53 _(0.08)	+0.44 _(0.08)	+0.58 _(0.07)	+0.65 _(0.06)	+0.63 _(0.04)	+0.61 _(0.06)
	R ²	+0.16 _(0.02)	+0.20 _(0.04)	+0.29 _(0.05)	+0.30 _(0.07)	+0.33 _(0.08)	+0.15 _(0.04)	+0.22 _(0.06)	+0.19 _(0.04)	+0.48 _(0.05)	+0.57 _(0.07)
max	CE	-0.80 _(0.13)	-0.55 _(0.07)	+0.09 _(0.03)	+0.02 _(0.06)	-0.21 _(0.06)	-0.55 _(0.07)	-0.19 _(0.06)	-0.21 _(0.05)	+0.03 _(0.03)	+0.30 _(0.04)
	RE	+0.55 _(0.03)	+0.59 _(0.02)	+0.43 _(0.02)	+0.41 _(0.02)	+0.58 _(0.02)	+0.53 _(0.02)	+0.64 _(0.02)	+0.71 _(0.01)	+0.71 _(0.01)	+0.68 _(0.02)
	R ²	+0.20 _(0.01)	+0.34 _(0.01)	+0.40 _(0.01)	+0.48 _(0.01)	+0.56 _(0.01)	+0.37 _(0.01)	+0.42 _(0.01)	+0.31 _(0.01)	+0.58 _(0.01)	+0.68 _(0.01)
CCA											
local	CE	-2.76 _(0.40)	-2.05 _(0.28)	-0.32 _(0.10)	-0.63 _(0.15)	-1.47 _(0.21)	-2.71 _(0.35)	-1.98 _(0.30)	-3.02 _(0.45)	-2.04 _(0.30)	-0.54 _(0.14)
	RE	+0.38 _(0.06)	+0.36 _(0.05)	+0.34 _(0.04)	+0.32 _(0.05)	+0.41 _(0.04)	+0.38 _(0.04)	+0.42 _(0.05)	+0.43 _(0.05)	+0.43 _(0.05)	+0.42 _(0.05)
	R ²	+0.18 _(0.03)	+0.20 _(0.03)	+0.29 _(0.04)	+0.33 _(0.04)	+0.38 _(0.06)	+0.17 _(0.04)	+0.22 _(0.05)	+0.19 _(0.04)	+0.46 _(0.04)	+0.54 _(0.05)
max	CE	-0.18 _(0.05)	-0.17 _(0.05)	+0.35 _(0.02)	+0.31 _(0.02)	+0.19 _(0.03)	-0.22 _(0.06)	+0.09 _(0.04)	-0.01 _(0.05)	+0.15 _(0.04)	+0.47 _(0.02)
	RE	+0.72 _(0.01)	+0.70 _(0.01)	+0.61 _(0.01)	+0.61 _(0.01)	+0.75 _(0.01)	+0.69 _(0.01)	+0.75 _(0.01)	+0.79 _(0.01)	+0.78 _(0.01)	+0.78 _(0.01)
	R ²	+0.41 _(0.01)	+0.47 _(0.01)	+0.53 _(0.01)	+0.60 _(0.01)	+0.68 _(0.01)	+0.49 _(0.01)	+0.56 _(0.01)	+0.43 _(0.02)	+0.64 _(0.01)	+0.76 _(0.01)
GraphEM											
local	CE	-2.29 _(0.37)	-1.70 _(0.27)	-0.24 _(0.10)	-0.49 _(0.13)	-1.04 _(0.19)	-2.23 _(0.30)	-1.41 _(0.22)	-2.23 _(0.34)	-1.54 _(0.24)	-0.24 _(0.10)
	RE	+0.44 _(0.05)	+0.41 _(0.05)	+0.36 _(0.04)	+0.35 _(0.04)	+0.50 _(0.04)	+0.44 _(0.04)	+0.51 _(0.04)	+0.53 _(0.04)	+0.52 _(0.04)	+0.52 _(0.04)
	R ²	+0.18 _(0.02)	+0.21 _(0.03)	+0.29 _(0.03)	+0.33 _(0.04)	+0.37 _(0.05)	+0.18 _(0.03)	+0.24 _(0.03)	+0.20 _(0.03)	+0.43 _(0.03)	+0.52 _(0.04)
max	CE	-0.02 _(0.06)	+0.06 _(0.06)	+0.49 _(0.02)	+0.46 _(0.03)	+0.38 _(0.05)	-0.00 _(0.07)	+0.28 _(0.05)	+0.10 _(0.07)	+0.25 _(0.06)	+0.57 _(0.03)
	RE	+0.77 _(0.01)	+0.76 _(0.01)	+0.68 _(0.02)	+0.71 _(0.02)	+0.83 _(0.01)	+0.78 _(0.01)	+0.82 _(0.01)	+0.83 _(0.01)	+0.82 _(0.01)	+0.82 _(0.01)
	R ²	+0.47 _(0.02)	+0.54 _(0.02)	+0.58 _(0.02)	+0.65 _(0.01)	+0.72 _(0.01)	+0.55 _(0.02)	+0.62 _(0.02)	+0.50 _(0.02)	+0.65 _(0.01)	+0.78 _(0.01)

Table S1: Verification statistics summary for the weighted global average, with the flat network. CE , RE and R^2 are computed for each 100-year time slice of interest. All numbers given outside of parentheses are the mean of the 100-member ensemble; numbers in parentheses are the corresponding standard deviation.

SNR	Metric	900	1000	1100	1200	1300	1400	1500	1600	1700	1800
RegEM - TTLS											
local	CE	-10.98 _(6.52)	-6.31 _(2.35)	-3.34 _(1.33)	-2.68 _(1.12)	-1.58 _(0.56)	-2.56 _(0.75)	-0.84 _(0.29)	-0.75 _(0.29)	-0.42 _(0.19)	0.21 _(0.08)
	RE	-1.85 _(1.57)	-0.87 _(0.58)	-1.61 _(0.82)	-0.93 _(0.57)	0.24 _(0.16)	0.15 _(0.16)	0.47 _(0.08)	0.65 _(0.04)	0.66 _(0.03)	0.66 _(0.03)
	R ²	0.12 _(0.02)	0.12 _(0.03)	0.17 _(0.05)	0.18 _(0.05)	0.25 _(0.08)	0.14 _(0.04)	0.21 _(0.05)	0.20 _(0.04)	0.48 _(0.03)	0.59 _(0.04)
max	CE	-1.11 _(0.32)	-1.03 _(0.18)	-0.18 _(0.13)	-0.27 _(0.19)	-0.23 _(0.10)	-0.84 _(0.14)	-0.23 _(0.08)	-0.30 _(0.08)	-0.00 _(0.04)	0.39 _(0.02)
	RE	0.48 _(0.08)	0.46 _(0.05)	0.25 _(0.08)	0.29 _(0.10)	0.62 _(0.03)	0.53 _(0.03)	0.65 _(0.02)	0.73 _(0.01)	0.72 _(0.01)	0.75 _(0.01)
	R ²	0.22 _(0.02)	0.29 _(0.02)	0.33 _(0.03)	0.41 _(0.02)	0.51 _(0.02)	0.30 _(0.02)	0.39 _(0.02)	0.32 _(0.01)	0.54 _(0.01)	0.70 _(0.01)
M09											
local	CE	-1.67 _(0.61)	-1.42 _(0.54)	-0.35 _(0.20)	-0.35 _(0.20)	-0.76 _(0.31)	-1.44 _(0.50)	-0.79 _(0.36)	-0.83 _(0.40)	-0.51 _(0.35)	0.09 _(0.15)
	RE	0.47 _(0.08)	0.42 _(0.11)	0.28 _(0.11)	0.30 _(0.08)	0.50 _(0.07)	0.44 _(0.08)	0.51 _(0.09)	0.63 _(0.06)	0.62 _(0.07)	0.61 _(0.06)
	R ²	0.13 _(0.04)	0.14 _(0.05)	0.23 _(0.06)	0.22 _(0.10)	0.31 _(0.09)	0.14 _(0.05)	0.16 _(0.05)	0.17 _(0.04)	0.46 _(0.05)	0.57 _(0.07)
max	CE	-0.68 _(0.22)	-0.54 _(0.08)	0.04 _(0.07)	0.03 _(0.10)	-0.45 _(0.16)	-0.65 _(0.13)	-0.24 _(0.09)	-0.23 _(0.06)	0.01 _(0.04)	0.29 _(0.04)
	RE	0.56 _(0.07)	0.55 _(0.02)	0.39 _(0.06)	0.34 _(0.06)	0.48 _(0.05)	0.46 _(0.05)	0.63 _(0.02)	0.71 _(0.01)	0.71 _(0.01)	0.68 _(0.02)
	R ²	0.19 _(0.02)	0.31 _(0.02)	0.35 _(0.03)	0.46 _(0.02)	0.52 _(0.03)	0.35 _(0.02)	0.40 _(0.01)	0.30 _(0.01)	0.58 _(0.01)	0.68 _(0.01)
CCA											
local	CE	-3.72 _(0.43)	-2.88 _(0.27)	-0.65 _(0.10)	-0.95 _(0.12)	-1.94 _(0.21)	-3.04 _(0.35)	-2.32 _(0.29)	-3.39 _(0.39)	-2.16 _(0.24)	-0.54 _(0.11)
	RE	0.23 _(0.06)	0.20 _(0.05)	0.20 _(0.05)	0.21 _(0.04)	0.30 _(0.05)	0.32 _(0.05)	0.35 _(0.05)	0.39 _(0.05)	0.42 _(0.04)	0.42 _(0.04)
	R ²	0.12 _(0.03)	0.12 _(0.04)	0.22 _(0.06)	0.17 _(0.09)	0.32 _(0.09)	0.12 _(0.03)	0.17 _(0.06)	0.18 _(0.04)	0.45 _(0.04)	0.56 _(0.04)
max	CE	-0.79 _(0.15)	-0.62 _(0.10)	0.15 _(0.04)	0.12 _(0.04)	-0.00 _(0.06)	-0.45 _(0.09)	-0.02 _(0.05)	-0.07 _(0.05)	0.15 _(0.04)	0.48 _(0.02)
	RE	0.63 _(0.02)	0.61 _(0.02)	0.53 _(0.02)	0.51 _(0.02)	0.70 _(0.01)	0.65 _(0.01)	0.73 _(0.01)	0.78 _(0.01)	0.78 _(0.01)	0.78 _(0.01)
	R ²	0.22 _(0.03)	0.35 _(0.02)	0.39 _(0.03)	0.50 _(0.02)	0.61 _(0.02)	0.40 _(0.02)	0.50 _(0.02)	0.39 _(0.02)	0.64 _(0.01)	0.76 _(0.01)
GraphEM											
local	CE	-2.40 _(0.32)	-1.87 _(0.22)	-0.28 _(0.10)	-0.59 _(0.11)	-1.48 _(0.17)	-2.54 _(0.33)	-1.67 _(0.24)	-2.57 _(0.36)	-1.78 _(0.28)	-0.35 _(0.11)
	RE	0.42 _(0.04)	0.38 _(0.04)	0.36 _(0.04)	0.32 _(0.03)	0.40 _(0.04)	0.40 _(0.04)	0.47 _(0.04)	0.49 _(0.04)	0.48 _(0.04)	0.48 _(0.04)
	R ²	0.14 _(0.03)	0.14 _(0.05)	0.23 _(0.05)	0.22 _(0.07)	0.31 _(0.07)	0.15 _(0.04)	0.19 _(0.04)	0.19 _(0.03)	0.40 _(0.04)	0.48 _(0.05)
max	CE	-0.28 _(0.10)	-0.20 _(0.07)	0.35 _(0.03)	0.30 _(0.04)	0.19 _(0.06)	-0.24 _(0.09)	0.10 _(0.06)	-0.08 _(0.09)	0.12 _(0.08)	0.50 _(0.03)
	RE	0.71 _(0.01)	0.70 _(0.01)	0.61 _(0.02)	0.62 _(0.02)	0.77 _(0.01)	0.72 _(0.01)	0.78 _(0.01)	0.80 _(0.01)	0.79 _(0.01)	0.79 _(0.01)
	R ²	0.29 _(0.02)	0.41 _(0.02)	0.45 _(0.03)	0.54 _(0.02)	0.64 _(0.01)	0.44 _(0.02)	0.53 _(0.02)	0.42 _(0.02)	0.61 _(0.02)	0.74 _(0.01)

Table S2: Same as Table S1, but with the staircase network.