

ABSTRACT

Over time scales between 10 days and 10-20 years – the macroweather regime – atmospheric fields, including the temperature, respect statistical scale symmetries, such as power-law correlations, that imply the existence of a huge memory in the system that can be exploited for long-term forecasts. The Stochastic Seasonal to Interannual Prediction System (StocSIPS) is a stochastic model that exploits these symmetries to perform long-term forecasts. It models the temperature as the high-frequency limit of the (fractional) energy balance equation (fractional Gaussian noise) which governs radiative equilibrium processes when the relevant equilibrium relaxation processes are power law, rather than exponential. They are obtained when the order of the relaxation equation is fractional rather than integer and they are solved as past value problems rather than initial value problems. Long-range weather prediction is conventionally an initial value problem that uses the current state of the atmosphere to produce ensemble forecasts. In contrast, StocSIPS predictions for long memory processes are “past value” problems that use historical data to provide conditional forecasts. Cross-correlations can be used to define teleconnection patterns, and for identifying possible dynamical interactions, but they do not necessarily imply any causation. Using the precise notion of Granger causality, we show that for long-range stochastic temperature forecasts, the cross-correlations are only relevant at the level of the innovations – not temperatures. Extended here to the multivariate case, (m-StocSIPS) produces realistic space-time temperature simulations. Although it has no Granger causality, we are able to reproduce emergent properties including realistic teleconnection networks and El Niño events and indices.

STOCSIPS

In the macroweather regime, the temperature anomalies are nearly Gaussian in time and their statistics satisfy a scaling symmetry. They can be modeled for every pixel as a fractional integration order $1/2 + H$ of a Gaussian white noise process, $\gamma(t)$ (fractional Gaussian noise):

$$T(t) = \frac{c_H}{\Gamma(H + 1/2)} \int_{-\infty}^t (t - t')^{-(1/2-H)} \gamma(t') dt'$$

For the discrete case:

$$\mathbf{T}_t = \mathbf{I}_{t,t}^H \gamma_t \quad \text{and} \quad \gamma_t = \text{inv}(\mathbf{I}_{t,t}^H) \mathbf{T}_t$$

where $\mathbf{I}_{t,t}^H$ is given by the Choleski decomposition of the correlation matrix. The forecast at time $t + k$ is:

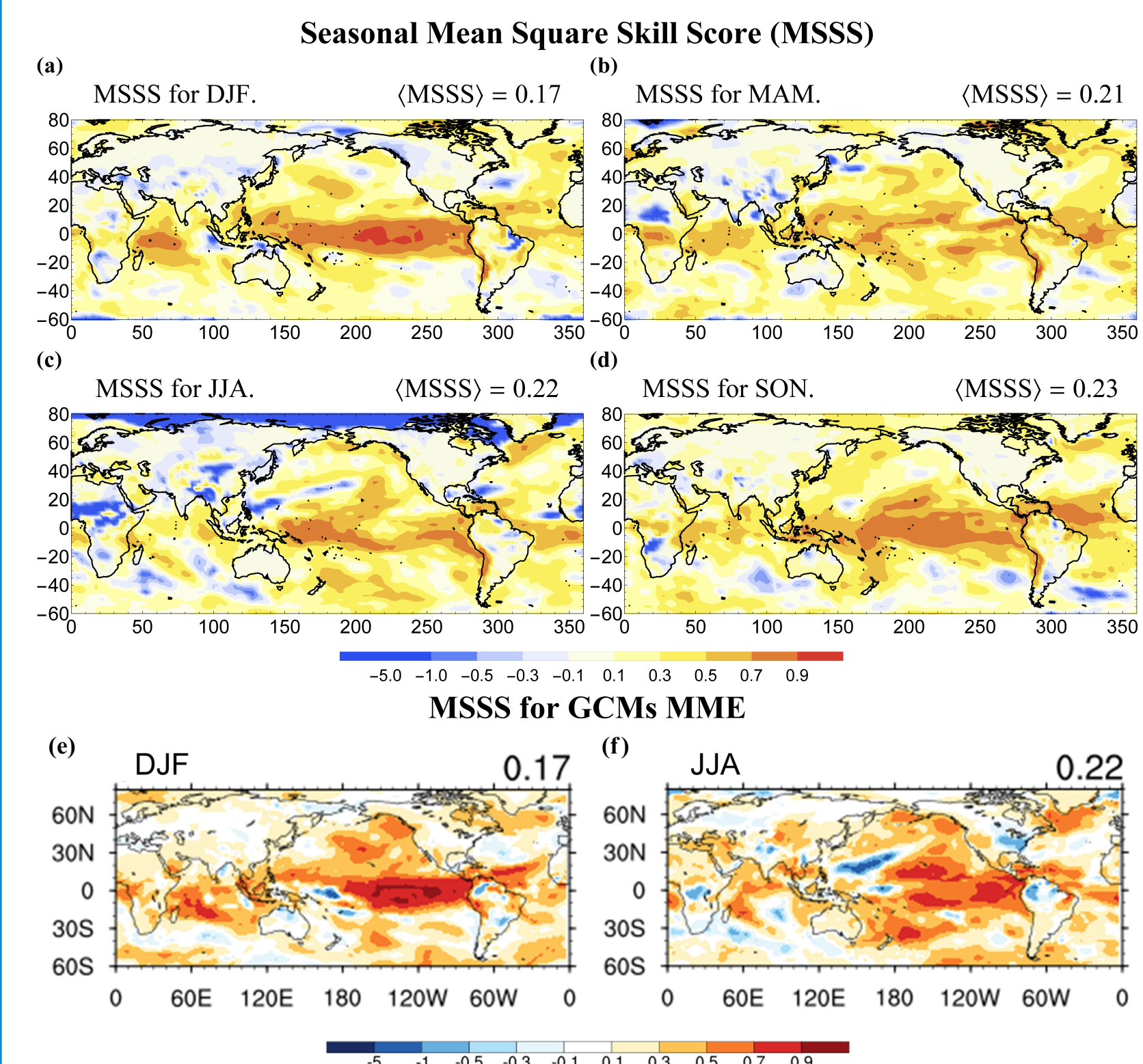
$$\hat{\mathbf{T}}_{t+k} = \mathbf{I}_{t+k,t}^H \gamma_t$$

and based on pass data:

$$\hat{\mathbf{T}}_{t+k} = \mathbf{G}_{t+k,t}^H \mathbf{T}_t$$

The predictor maximizes the mean square skill score.

Compared to an optimal MME combination of 6 GCMs, StocSIPS shows similar or better Skill.



M-STOCSIPS

StocSIPS predicts every pixel independently using its own past. To investigate the possibility of improving StocSIPS forecasts using spatial cross-correlations, we extended it to the multivariate version (m-StocSIPS):

$$T_i(t) = \sum_j \int_{-\infty}^t \kappa_{ij}(t - t') \gamma_j(t') dt'$$

where:

$$\kappa_{ij}(t) = \delta_{ij} \frac{c_{H_i}}{\Gamma(H_i + 1/2)} t^{-(1/2-H_i)}$$

and

$$\rho_{ij}(t - t') = \langle \gamma_i(t) \gamma_j(t') \rangle = a_{ij} \delta(t - t'),$$

with $-1 \leq a_{ij} \leq 1$.

The temperature cross-covariance for lag $\Delta t > 0$ is:

$$C_{ij}(\Delta t) = \langle T_i(t) T_j(t + \Delta t) \rangle = \sum_{mn} \int_0^\infty \kappa_{im}(t') \kappa_{jn}(t' + \Delta t) a_{mn} dt'$$

hence the cross-correlation is:

$$R_{ij}(\Delta t) = \frac{C_{ij}(\Delta t)}{\sqrt{C_{ii}(0) C_{jj}(0)}}$$

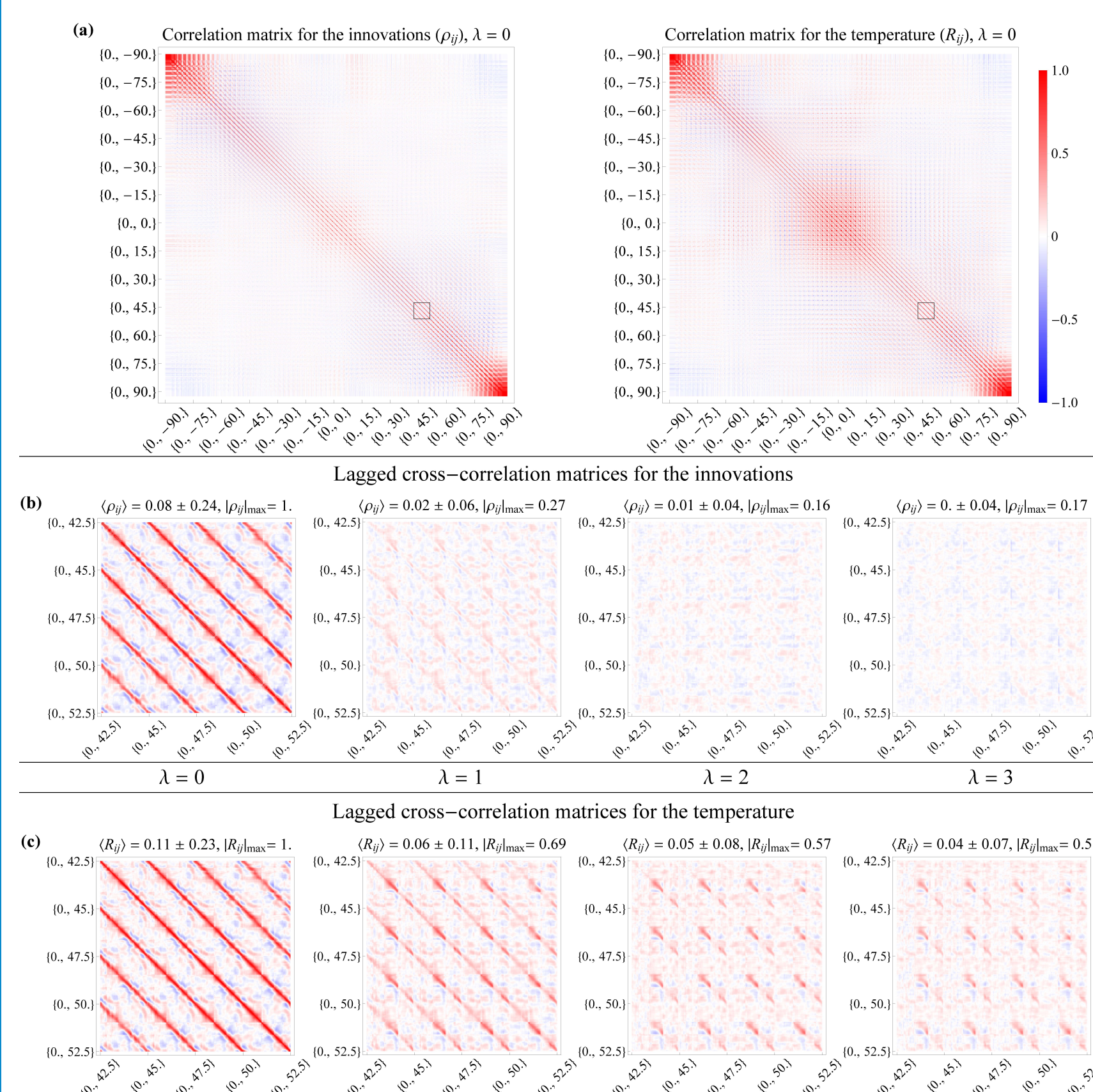
If the model is valid ($\rho_{ij}(t - t') = a_{ij} \delta(t - t')$), then any predictor that is a linear combination of past temperature values from any position j , is orthogonal to the error of the predictor obtained from the past at location i :

$$\langle T_j(s) E_i(t) \rangle = 0; \quad s < 0, \quad t > 0$$

Hence, the StocSIPS predictor based on a single pixel is optimal given the full field $T(x, t)$ for $t \leq 0$. This is a precise statement of Granger causality for the m-StocSIPS model. Although there are large cross-correlations inherited from the innovation matrix a_{ij} , the information of past temperatures from other locations does not help improve the forecast.

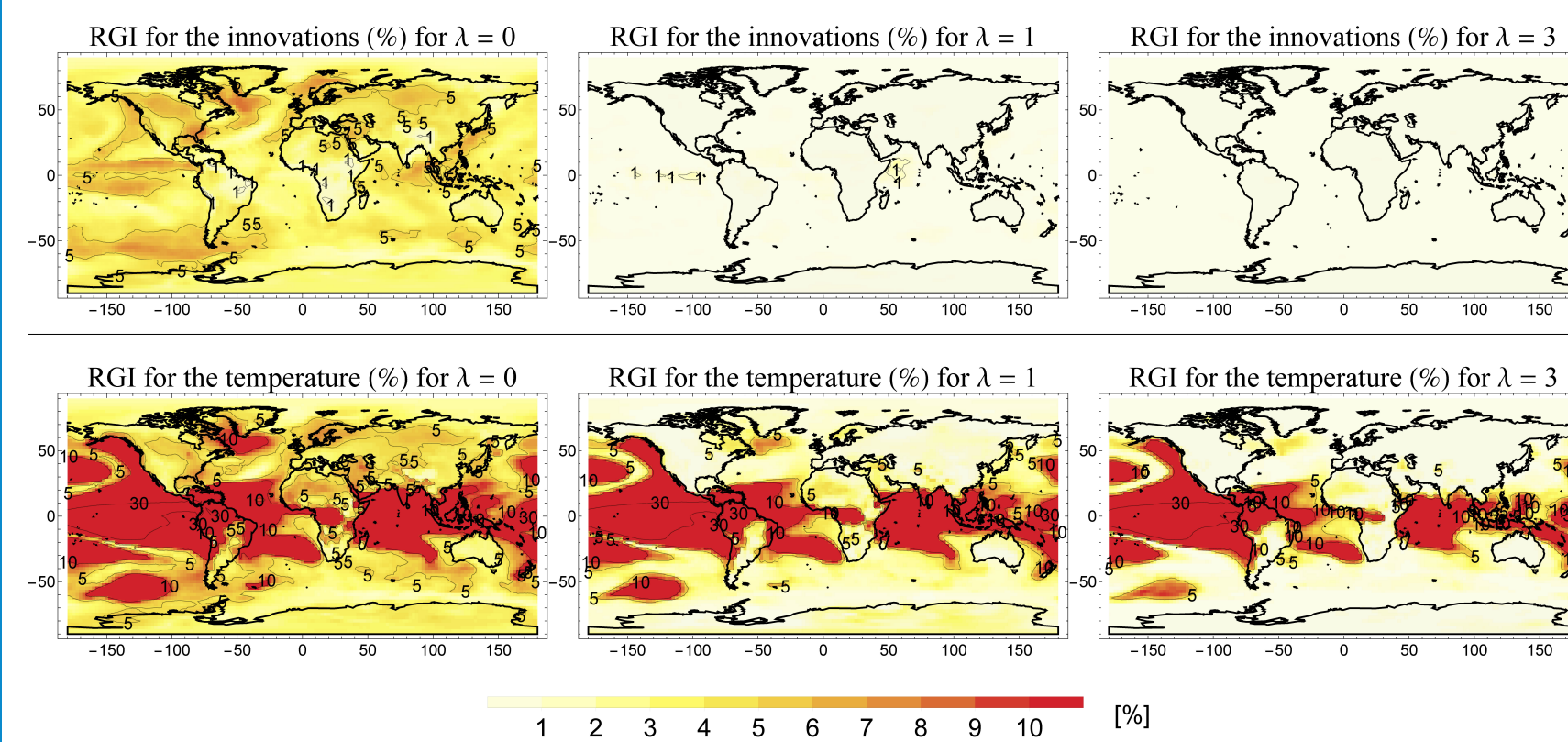
MODEL VALIDITY

Empirical cross-correlations



Ratio of global influence (RGI)

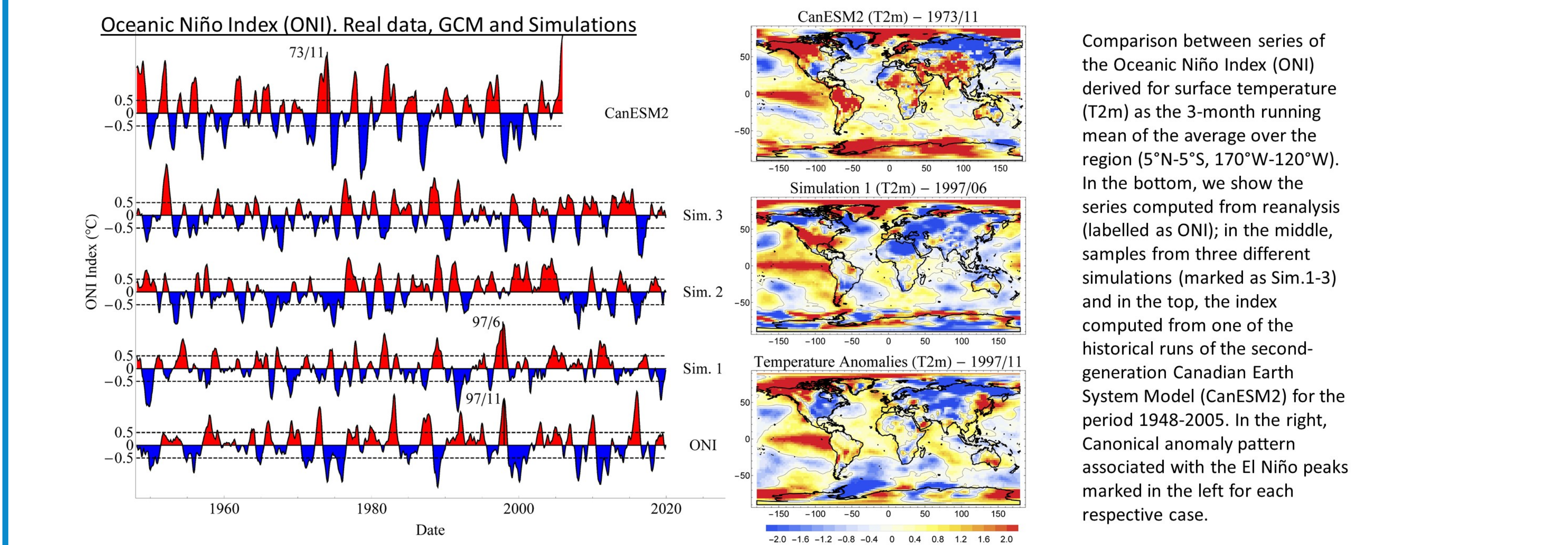
The RGI for pixel i is the fraction of the area of the planet for which $|R_{ij}(\lambda)| > 0.2$, averaged over all j (for innovations $|\rho_{ij}(\lambda)| > 0.2$).



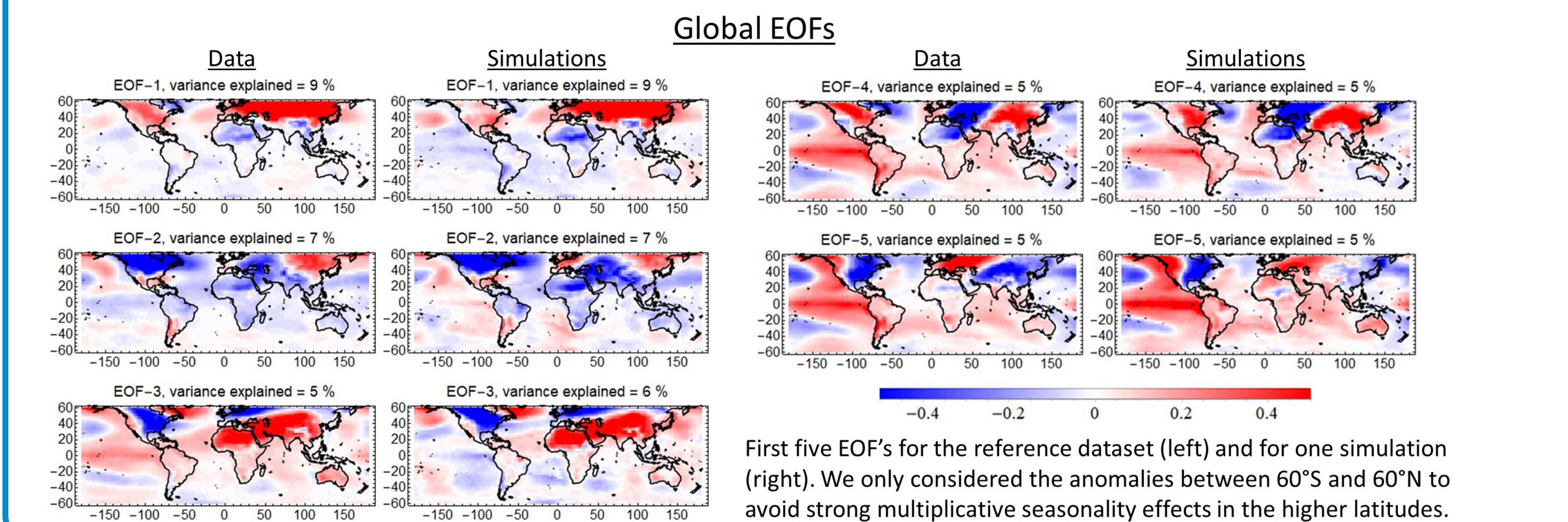
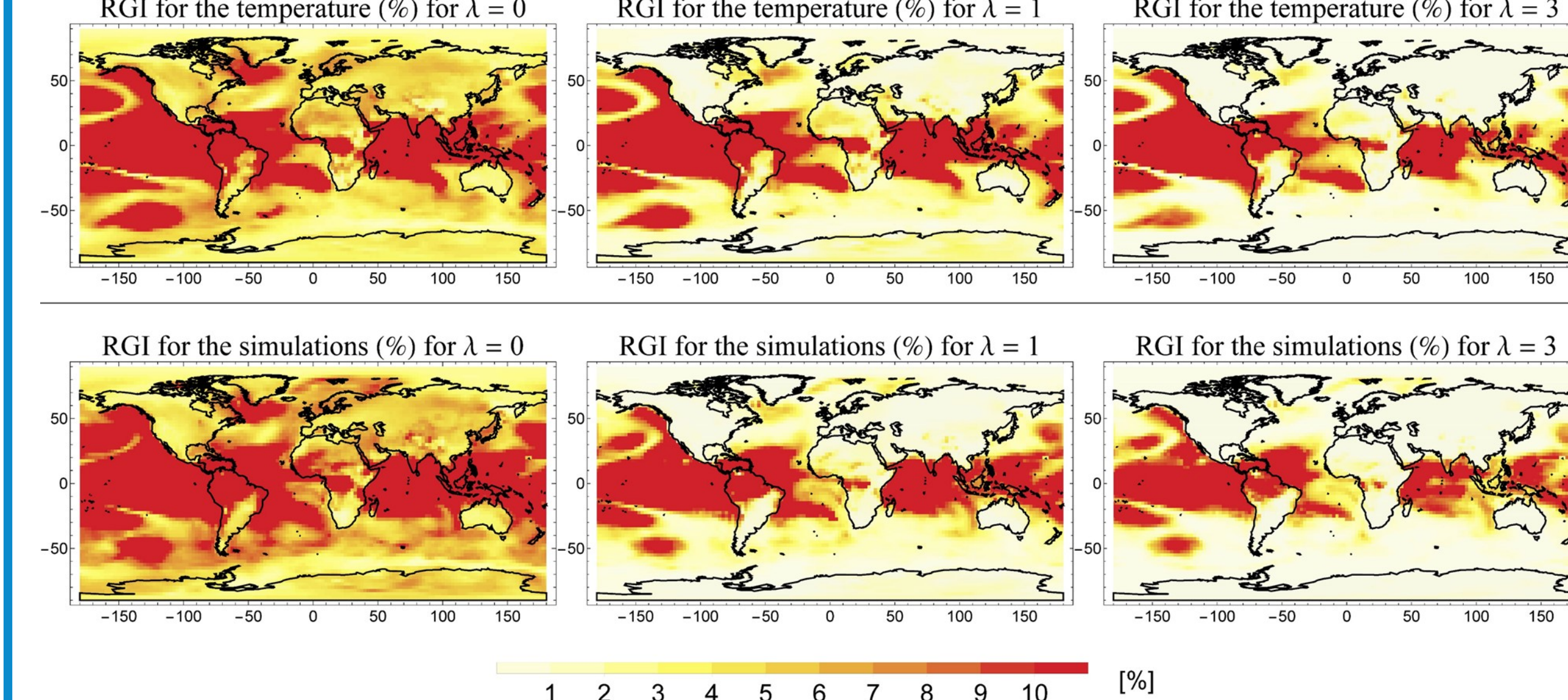
For the innovations, almost all the correlation is lost for $\lambda > 0$: there is no significant influence on future values for any pixel. For $\lambda > 0$, the region of largest innovation influence is the tropical Pacific where $\text{RGI} \approx 5\%$. For temperature anomalies (panel (d)), much larger correlations and RGIs are obtained. For $\lambda > 0$, almost all the influence from land disappears, but the ocean's influence is preserved up to around 1 year (not shown).

SIMULATIONS AND EMERGENT PROPERTIES

To compare StocSIPS space-time statistical structures to reanalysis and to GCM outputs, we produced simulations with the same resolutions and overall length as our reference NCEP/NCAR Reanalysis 1 dataset (864 months, 2.5° resolution). Full movies of the model outputs are available at <https://youtu.be/1MvRcE2-LKc>. Here we focus on El Niño events that are particularly difficult to simulate.



Ratio of Global Influence (RGI) for different lags. Real data and Simulations



First five EOF's for the reference dataset (left) and for one simulation (right). We only considered the anomalies between 60°S and 60°N to avoid strong multiplicative seasonality effects in the higher latitudes.

SUMMARY

- Scaling-based long-range stochastic forecasting is a past value problem not an initial value problem.
- Granger causality shows that spatial correlations in temperature cannot be used to improve on memory-based predictions of infinitely long series.
- Real-world statistics and teleconnection patterns can be reproduced with stochastic simulations with a total lack of causal relationships.

References

- 2021.- Del Rio Amador, L., and Lovejoy, S., (2021) Long-range Forecasting as a Past Value Problem: Untangling Correlations and Causality with Scaling, Geophys. Res. Lett., 48, e2020GL092147.
<https://doi.org/10.1029/2020GL092147>
- 2021.- Del Rio Amador, L., and Lovejoy, S., (2021) Using regional scaling for temperature forecasts with the Stochastic Seasonal to Interannual Prediction System (StocSIPS), Clim. Dyn.,
<https://doi.org/10.1007/s00382-021-05737-5>
- 2019.- Del Rio Amador, L., and Lovejoy, S., (2019) Predicting the global temperature with the Stochastic Seasonal to Interannual Prediction System (StocSIPS), Clim. Dyn.,
<https://doi.org/10.1007/s00382-019-04791-4>
- For more information about StocSIPS, visit: <http://www.physics.mcgill.ca/StocSIPS/index.html>