Supplementary Information to A causality-based perspective on Atlantic and tropical Pacific teleconnections to Northeast Brazil in seasonal forecast systems

Corresponding author: Aleksander Lacima-Nadolnik (aleksander.lacima@bsc.es)

Contents of Supplementary Information:

Methods

- **S1 Detrending**
- **S2 EOF projection**
- **S3 MSSS decomposition**

Figures & Tables

- **Figure S1**: Correspondence between PCA-based modes and conventional box-averaged climate indices in ERA5.
- **Figure S2**: ACC scorecards for tropical Atlantic SST.
- **Figure S3**: ACC scorecards for Niño3.4 SST.
- **Figure S4**: ACC scorecards for South Atlantic SST.
- **Figure S5**: ACC scorecards for South Atlantic SLP.
- **Figure S6**: MSSS scorecards for surface wind over NEB.
- **Figure S7**: MSSS scorecards for precipitation over NEB.
- **Figure S8**: MSSS scorecards for tropical Atlantic SST.
- **Figure S9**: MSSS scorecards for Niño3.4 SST.
- **Figure S10**: MSSS scorecards for South Atlantic SST.
- **Figure S11**: MSSS scorecards for South Atlantic SLP.
- **Figure S12**: Information transfer from the second mode of tropical Atlantic SST (i.e., AMM) to surface wind at global scale.
- **Figure S13**: Information transfer from the second mode of tropical Atlantic SST (i.e., AMM) to precipitation at global scale.
- **Figure S14**: Information transfer from the first mode of tropical Pacific SST (i.e., ENSO) to surface wind at global scale.
- **Figure S15**: Information transfer from the first mode of tropical Pacific SST (i.e., ENSO) to precipitation at global scale.
- **Figure S16**: Information transfer from the third mode of South Atlantic SLP (i.e., SASH meridional displacements) to surface wind at global scale.
- **Figure S17**: Information transfer from the first mode of South Atlantic SST (i.e., SASD/SAOD) to surface wind at global scale.

- **Figure S18**: Information transfer from the first mode of South Atlantic SST (i.e., SASD/SAOD) to precipitation at global scale.
- **Figure S19**: Information transfer from the second mode of South Atlantic SST to surface wind at global scale.
- **Figure S20**: Information transfer from the second mode of South Atlantic SST to precipitation at global scale.
- **Figure S21**: Information transfer and correlation between the dominant large-scale drivers and surface wind over NEB for each initialisation month across the C3S multi-system.
- **Figure S22**: Information transfer and correlation between the dominant large-scale drivers and precipitation over NEB for each initialisation month across the C3S multi-system.
- **Figure S23**: Relationship between causal fidelity and average forecast skill for precipitation over NEB.
- **Table S1**: Correlation between causal fidelity and each term of the MSSS decomposition for surface wind and precipitation over NEB.
- **Table S2**: Causal fidelity and MSSS amplitude error of each C3S system for surface wind over NEB in the extended 1980–2025 reference period.
- **Table S3**: Correlation between causal fidelity and MSSS for surface wind and precipitation over NEB, without control and controlling for driver skill and ensemble size.
- **Figure S24**: One-point correlation between surface wind anomalies over NEB and global precipitation anomalies.

S1. Detrending

At each grid point, the anomaly field, $y(x,t)$, is decomposed as:

$$y(x,t) = F(x,t) + \varepsilon(x,t) \quad \text{(S1)}$$

where F is the forced component, and ε is the residual, which contains the internal variability, any forced signal not captured by scaling with the nonlinear reference trend and noise. The forced component is then expressed as:

$$F(x,t) = a(x) + b(x) \cdot L_f[GMST(t)] \;\; ; \quad f = \delta/T \quad \text{(S2)}$$

where a is the intercept, b is the local sensitivity to the nonlinear reference trend, L_f is the LOWESS operator, f is the parameter controlling the fraction of reference time series used locally by the LOWESS operator at each step, $GMST$ is the global mean surface temperature time series used as reference, δ is the selected smoothing window, here $\delta = 120$ months (i.e., 10 years), and T is the length of the reference time series. The detrended anomaly field, $y'(x,t)$, is then the regression residual:

$$y'(x,t) \equiv \varepsilon(x,t) = y(x,t) - a(x) - b(x) \cdot L_f[GMST(t)] \quad \text{(S3)}$$

As the regression includes the intercept $a(x)$, the residual has zero temporal mean over the fitting period by construction (i.e., $\overline{\varepsilon(x,t)} = 0$). In other words, $a(x)$ accounts for the nonzero mean of the anomaly over the fitting period. If the anomaly field $y(x,t)$ is not locally sensitive to the nonlinear reference trend, then $b(x) \simeq 0$ and the detrended anomaly field is $y'(x,t) \simeq y(x,t) - a(x)$ (i.e., the original anomaly with its average over the fitting period removed). Each variable is detrended using its respective LOWESS-smoothed GMST time series, derived from within each dataset, ensuring that the forced signal estimate is internally consistent. The reanalysis is detrended over the longer 1950–2025 period to avoid LOWESS edge effects in 1980, whereas the hindcasts are detrended over the common hindcast period (1993–2016). Over this shorter 24-year window, the LOWESS-smoothed reference is effectively linear, so the nonlinear character of the reference trend is most relevant for the reanalysis. We verified that the detrending is insensitive to the fitting period. Over the 1980–2025 and 1993–2016 overlaps, the GMST reference computed from the longer record differs from the shorter references by only ~2% of its total range on average, and the detrended regional anomalies correlate at r ≥ 0.999. The largest discrepancies occur near the start of the shorter windows, which is consistent with LOWESS edge effects.

S2. EOF projection

To ensure that each driver is expressed with respect to the same spatial pattern, we define a common basis from the ERA5 EOFs and project each hindcast anomaly field onto it. For a given mode, the hindcast PC at each time step is obtained as:

$$PC_h(i,l,m) = \frac{1}{\sigma_{ref}} \cdot \sum_{x,y} \sqrt{cos(\phi_y)} \cdot X_h(i,l,m,x,y) \cdot EOF_{ref}(x,y) \quad \text{(S4)}$$

where i, l, and m correspond to initialisation, lead time, and member, respectively, ϕ_y is the latitude of grid row y, $X_h(i,l,m,x,y)$ is the hindcast anomaly field, and the sum is taken over the ocean grid cells

for SST-based modes or over the full spatial domain for SLP-based modes. The rescaling factor σ_{ref} is the standard deviation of ERA5 PC over the period 1980–2025, placing the projected hindcast PCs on a scale consistent with the ERA5-based ones, while preserving relative amplitude differences across models.

S3. MSSS decomposition

Following Murphy (1988), the MSSS is decomposed into its phase agreement, amplitude (i.e., conditional bias), and mean bias terms as follows:

$$\boldsymbol{MSSS} = \frac{\boldsymbol{r^2}}{\boldsymbol{1+q}} - \frac{\boldsymbol{(r-s)^2}}{\boldsymbol{1+q}} - \frac{\boldsymbol{b^2}}{\boldsymbol{1+q}} + \frac{\boldsymbol{q}}{\boldsymbol{1+q}} = \boldsymbol{P - A - B + Q} \qquad \text{(S5)}$$

where r is the Pearson correlation between the ensemble-mean hindcast and the reference anomalies, $s = s_{h'}/s_{o'}$ is the ratio of their standard deviations, $b = (\bar{h}' - \bar{o}')/s_{o'}$ is the standardised mean bias, and$q = \bar{o}'^2/s_{o'}^2$ is the squared mean of the reference anomalies relative to their variance. P, A, and B are the phase agreement, amplitude, and mean bias terms, respectively. The term $Q = q/(1+q)$ arises because the reference anomalies have a small non-zero mean over the verification period (1993–2016). For $\bar{o}' = 0$ the term vanishes and the expression reduces to that of Murphy (1988). The amplitude term vanishes for $s = r$, when the ensemble-mean amplitude matches the level justified by its correlation with the reference.

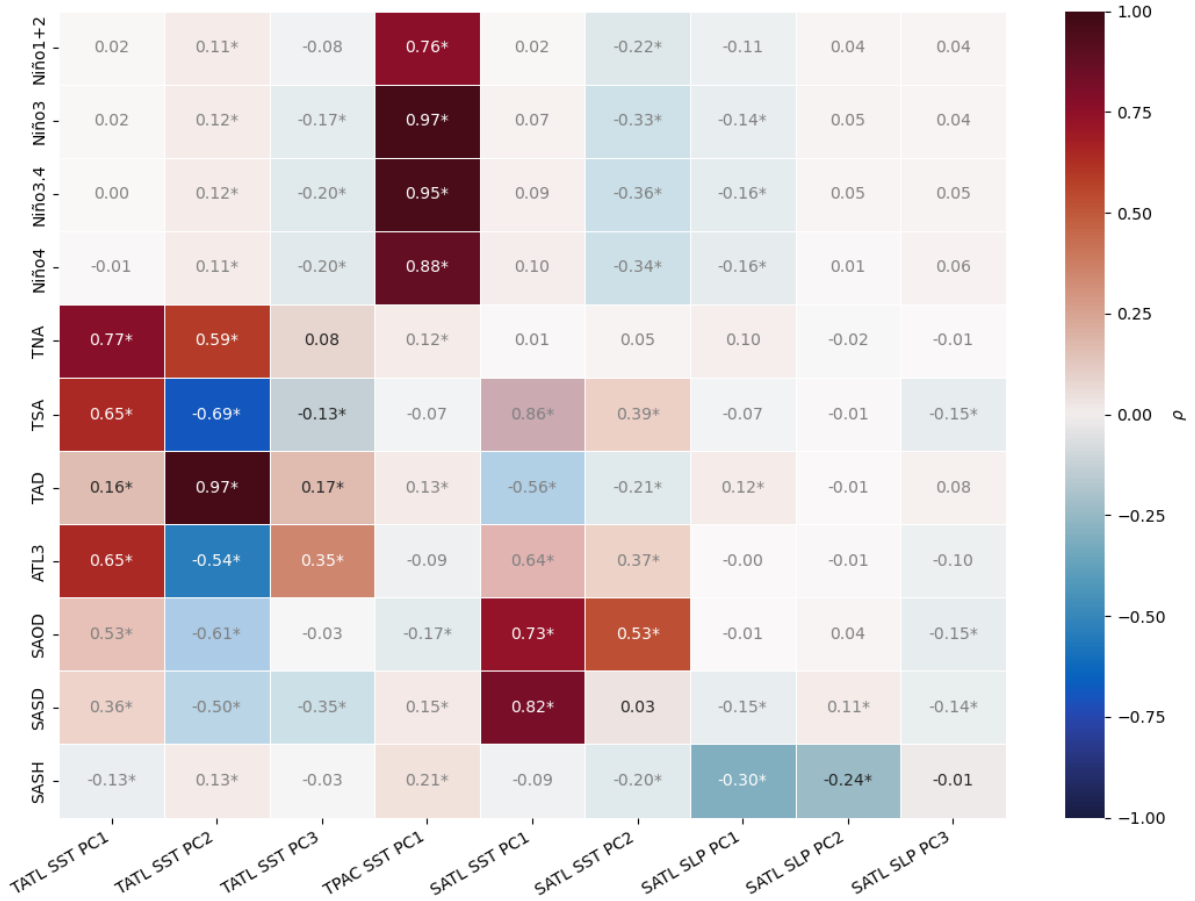

Figure S1. Spearman correlations between the PC time series from Fig. 2 and the box-averaged indices commonly used to characterise the large-scale modes of variability used in the study in the ERA5 reanalysis over 1980-2025. From top to bottom these are the Niño1+2, Niño3, Niño3.4 and Niño4 indices, the tropical North Atlantic (TNA), tropical South Atlantic (TSA) and tropical Atlantic dipole (TAD; i.e., TNA-TSA) indices, the Atlantic Niño (ATL3) index, the South Atlantic subtropical dipole (SASD) and South Atlantic Ocean dipole (SAOD) indices, and the South Atlantic subtropical high (SASH) index. Asterisks indicate when correlation values are statistically significant at the 99 % confidence level (α = 0.01), obtained through a standard two-tailed Student's t-test. Non-faded cells indicate the correspondence between box-averaged indices and PCA-based indices.

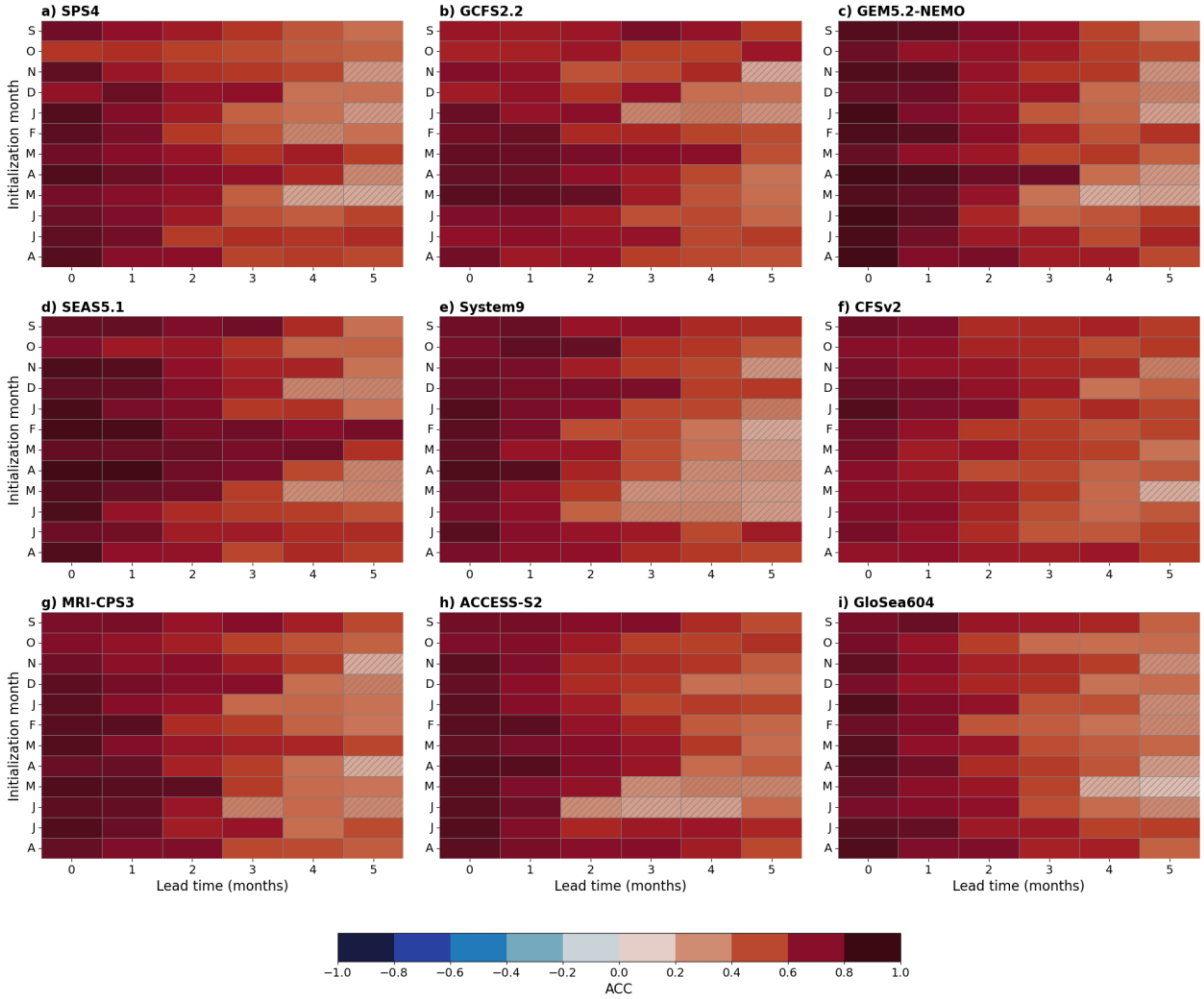

Figure S2. Same as Fig. 3 but for SST in the tropical Atlantic region shown in Fig. 2.

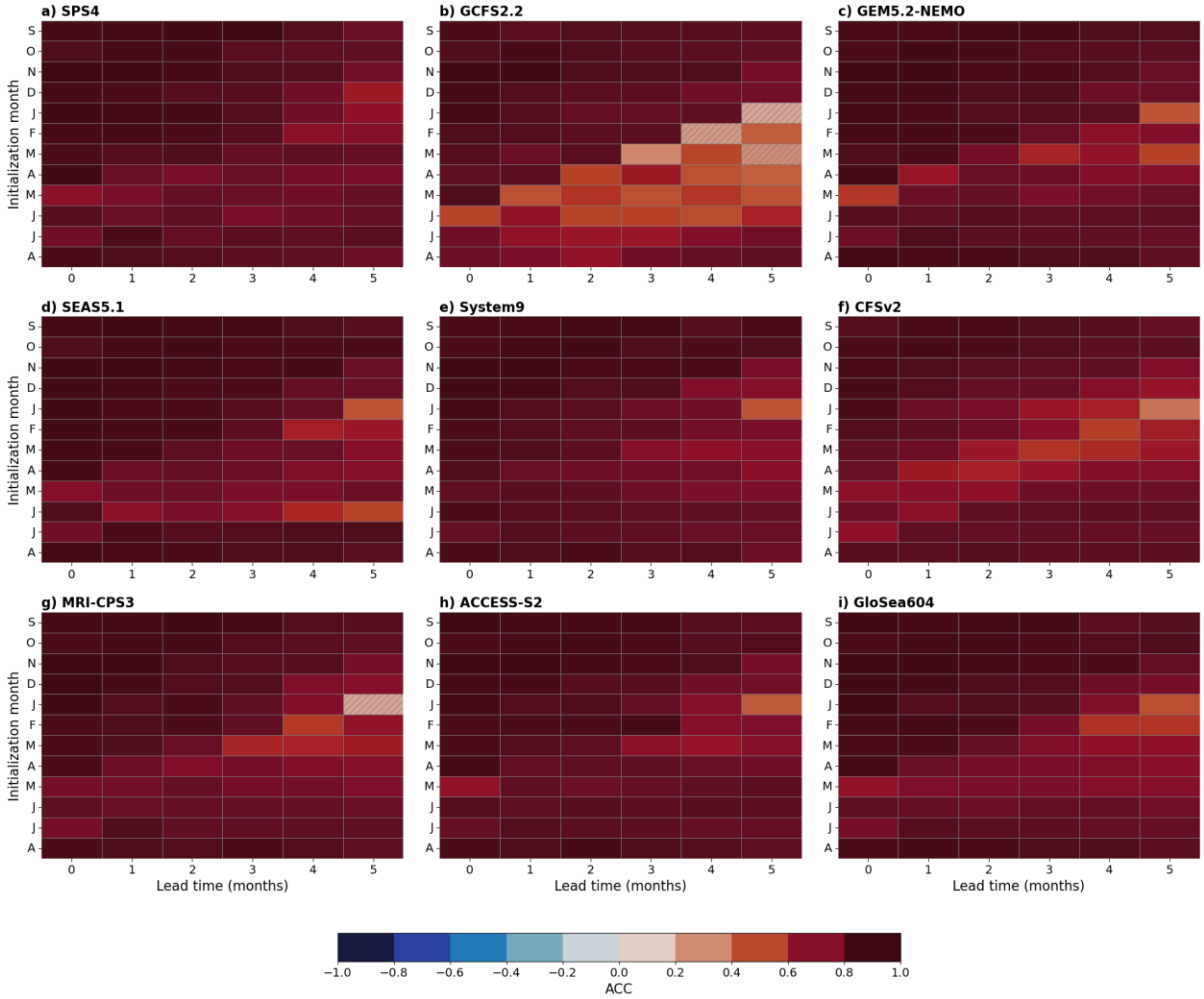

Figure S3. Same as Fig. 3, but for SST in the Niño 3.4 region.

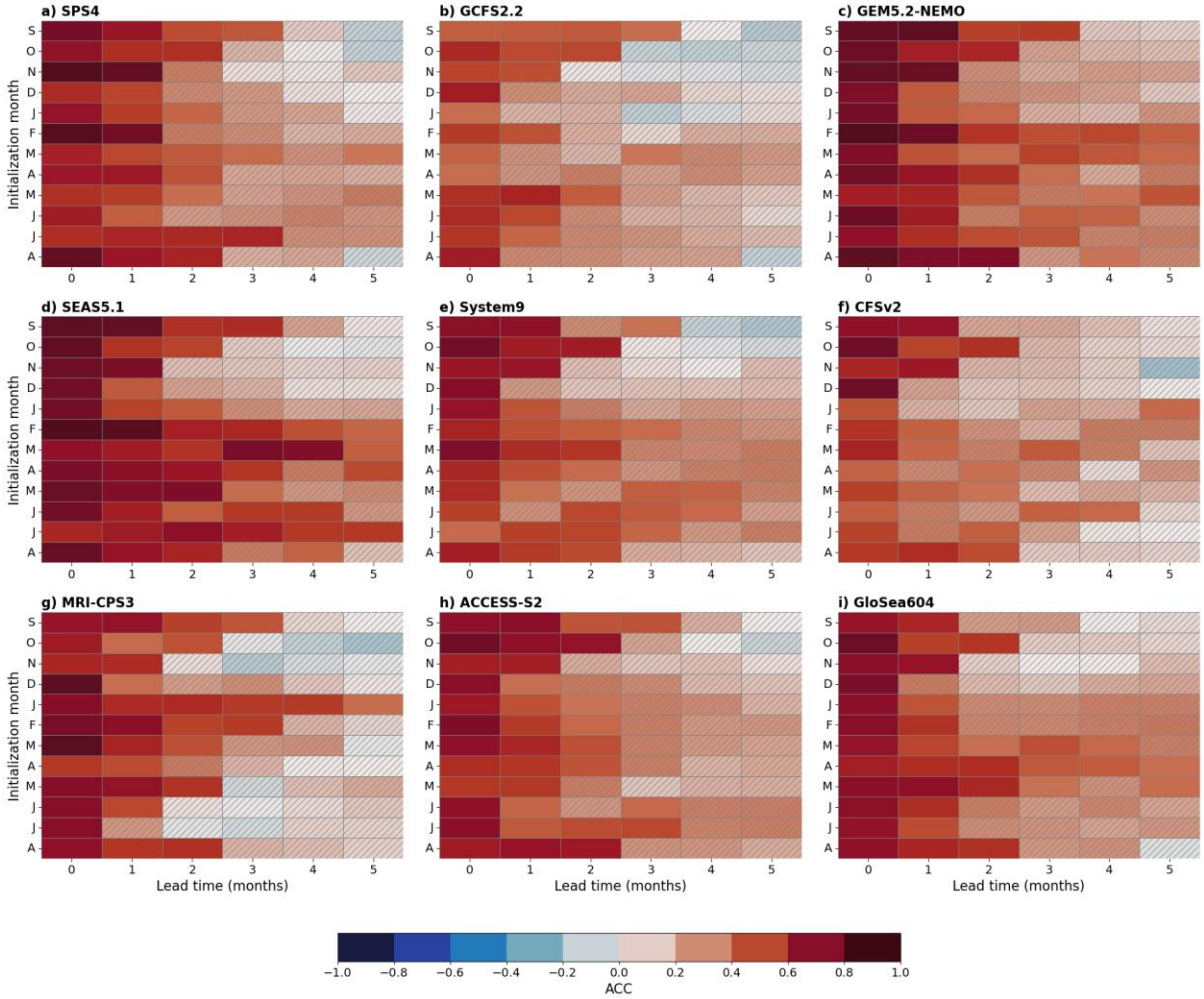

Figure S4. Same as Fig. 3 but for SST in the South Atlantic region shown in Fig. 2.

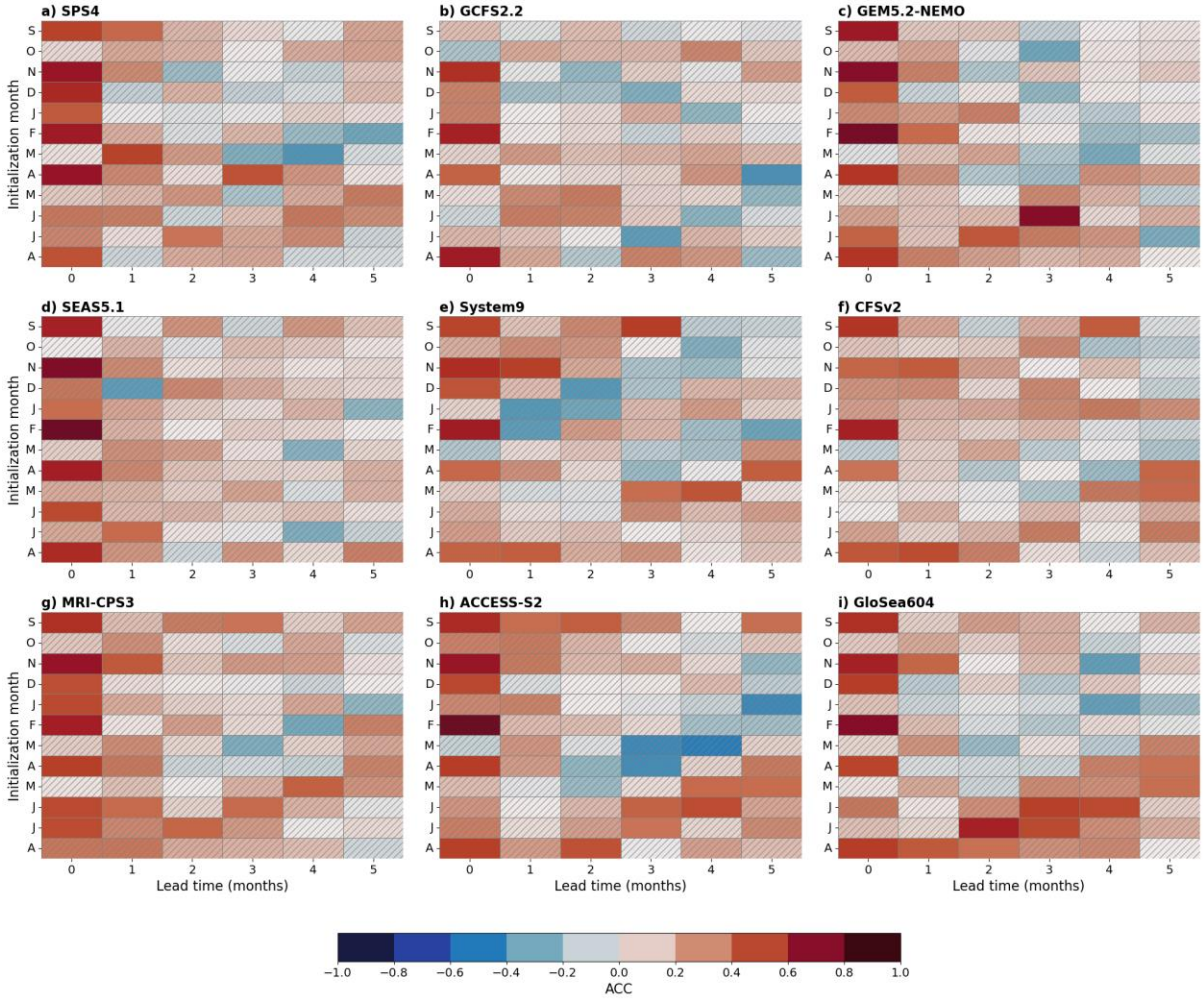

Figure S5. Same as Fig. 3 but for SLP in the South Atlantic region shown in Fig. 2.

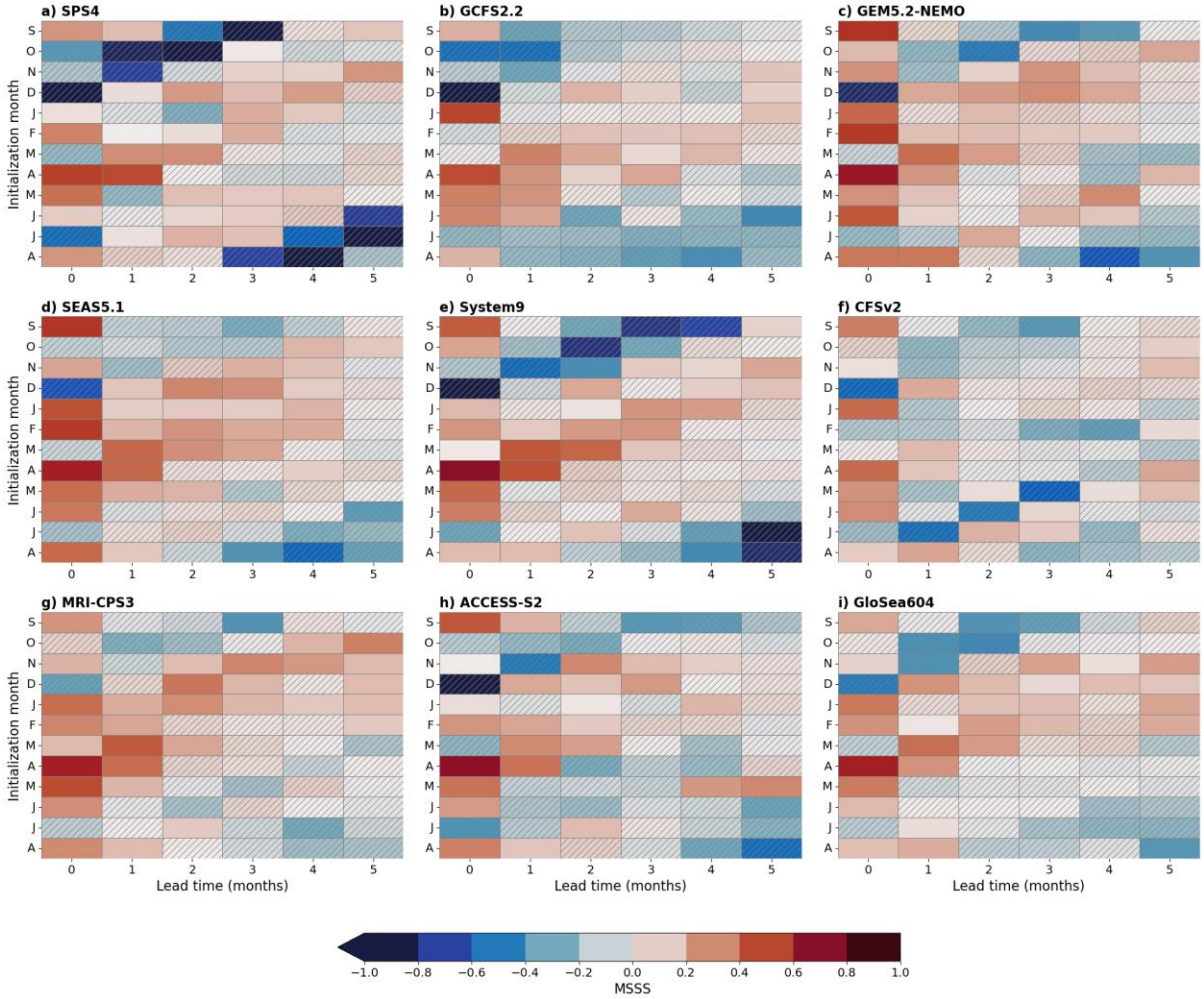

Figure S6. MSSS scorecards of surface wind in NEB for all initialisation months and lead times for a) SPS4, b) GCFS2.2, c) GEMS5.2-NEMO, d) SEAS5.1, e) System9, f) CFSv2, g) MRI-CPS3, h) ACCESS-S2, and i) GloSea604. Non-hatched cells indicate statistically significant MSSS at the 95 % confidence level (α = 0.05), obtained through a one-tailed permutation test with 1000 random permutations. Note that the initialisation months go from September to August.

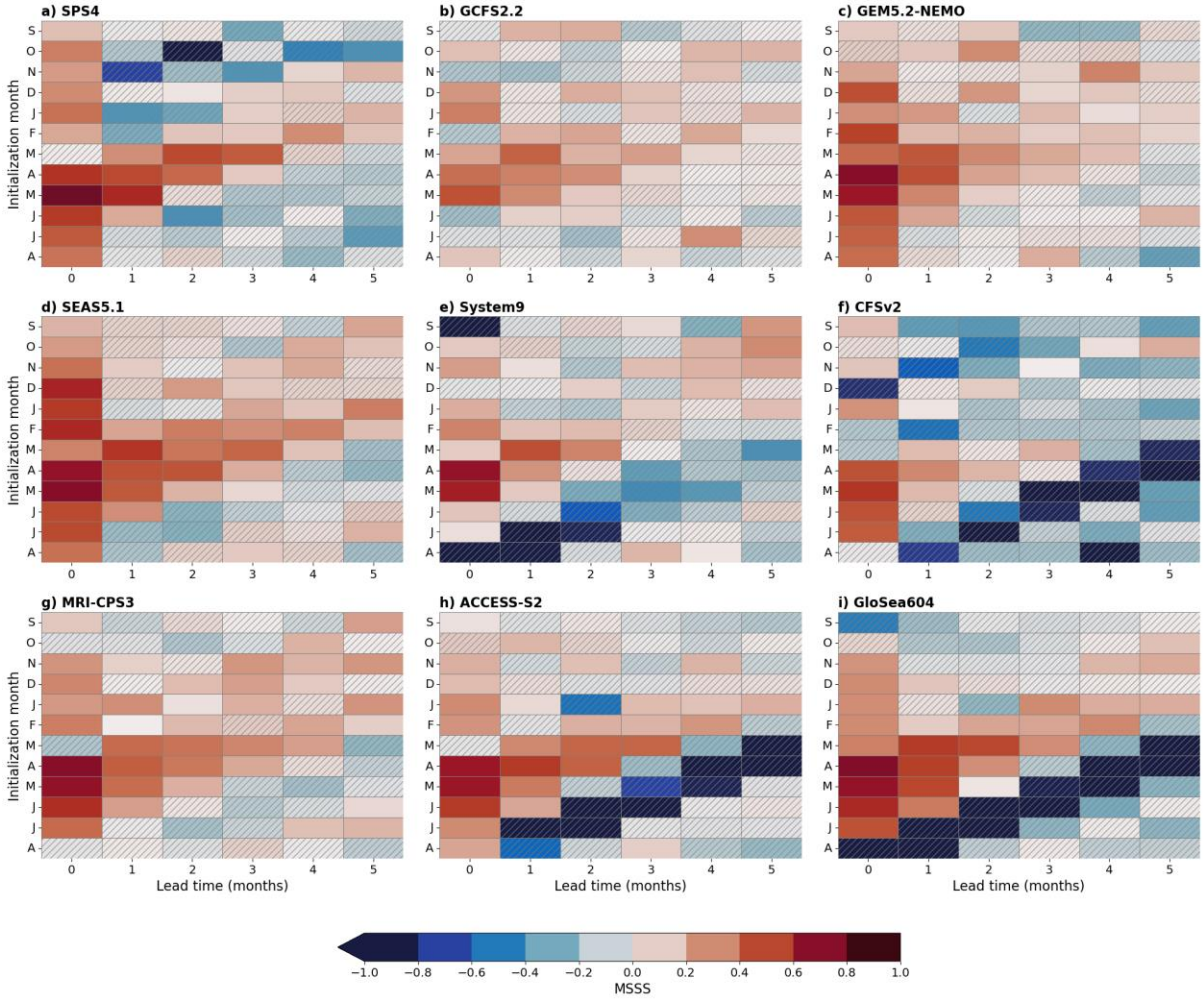

Figure S7. Same as Fig. S6 but for precipitation in NEB.

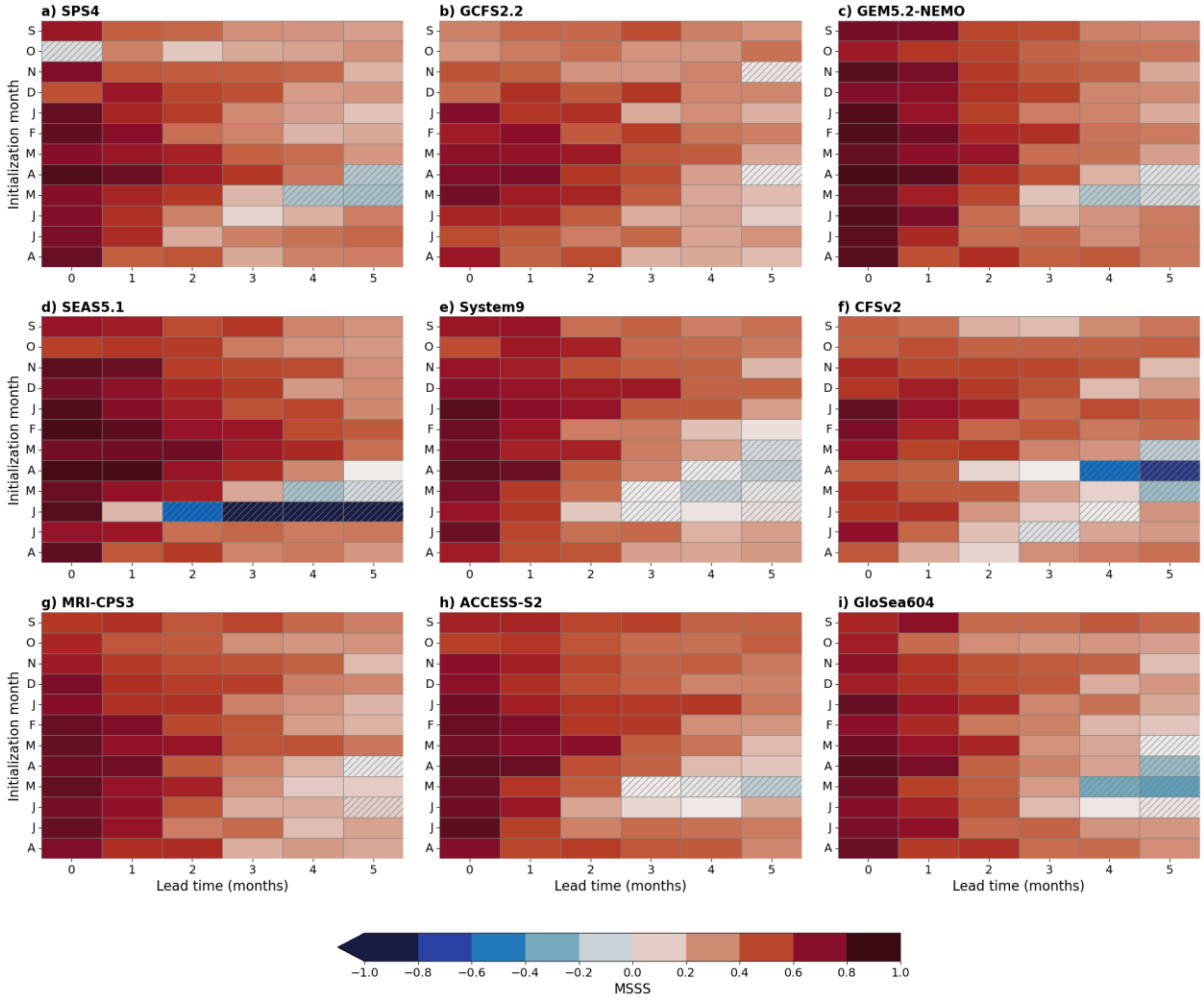

Figure S8. Same as Fig. S6 but for SST in the tropical Atlantic region shown in Fig. 2.

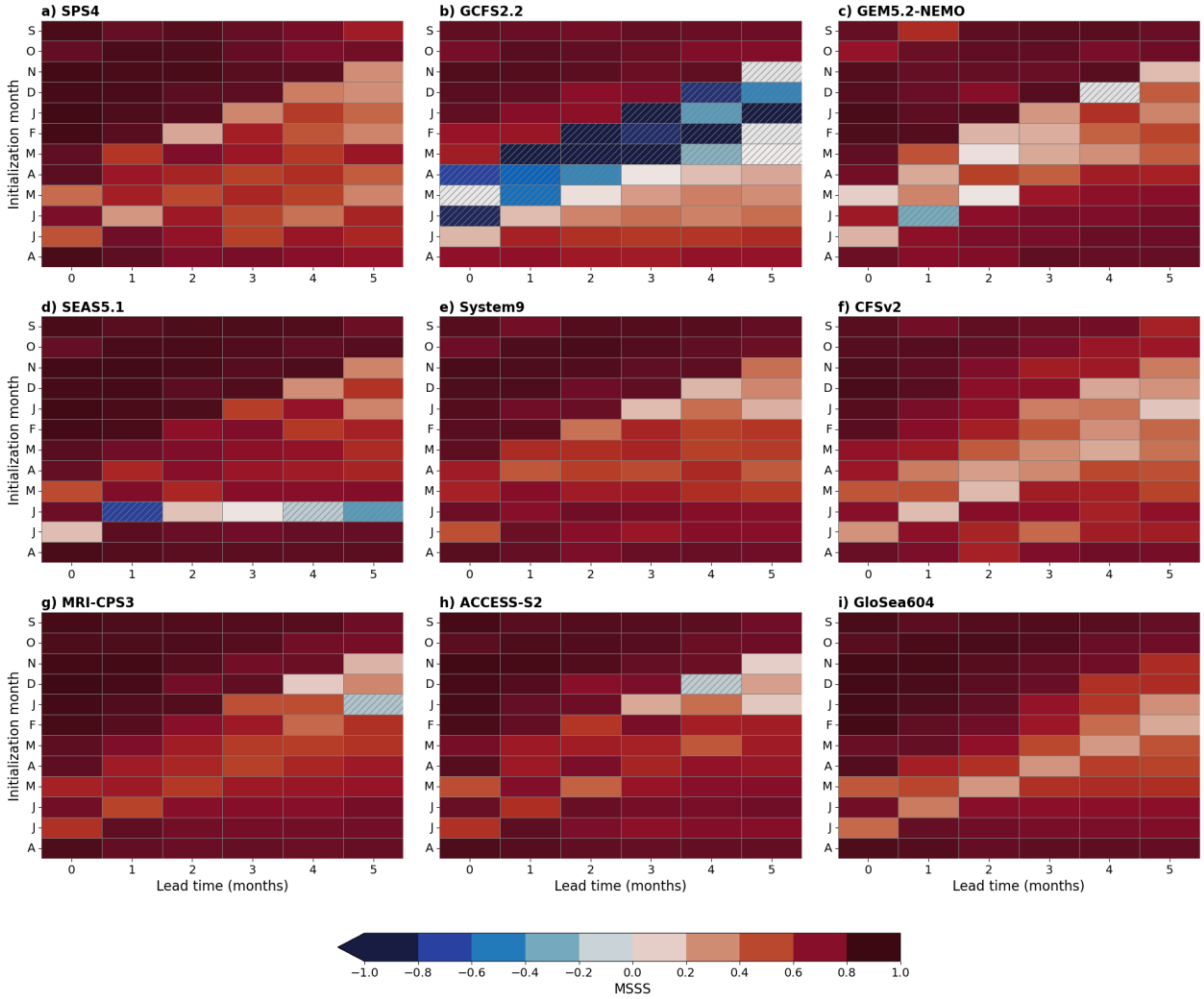

Figure S9. Same as Fig. S6 but for SST in the Niño3.4 region.

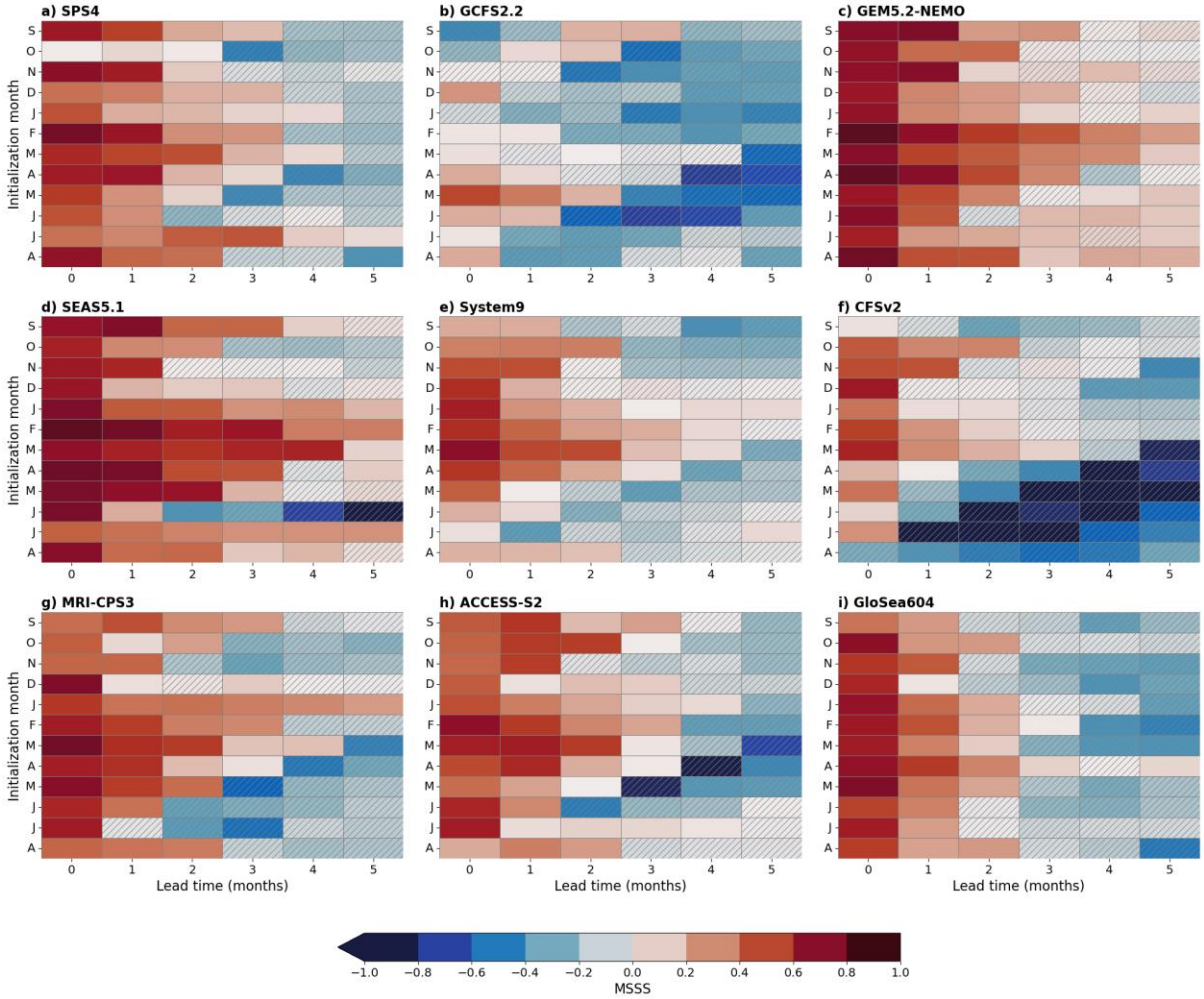

Figure S10. Same as Fig. S6 but for SST in the South Atlantic region shown in Fig. 2.

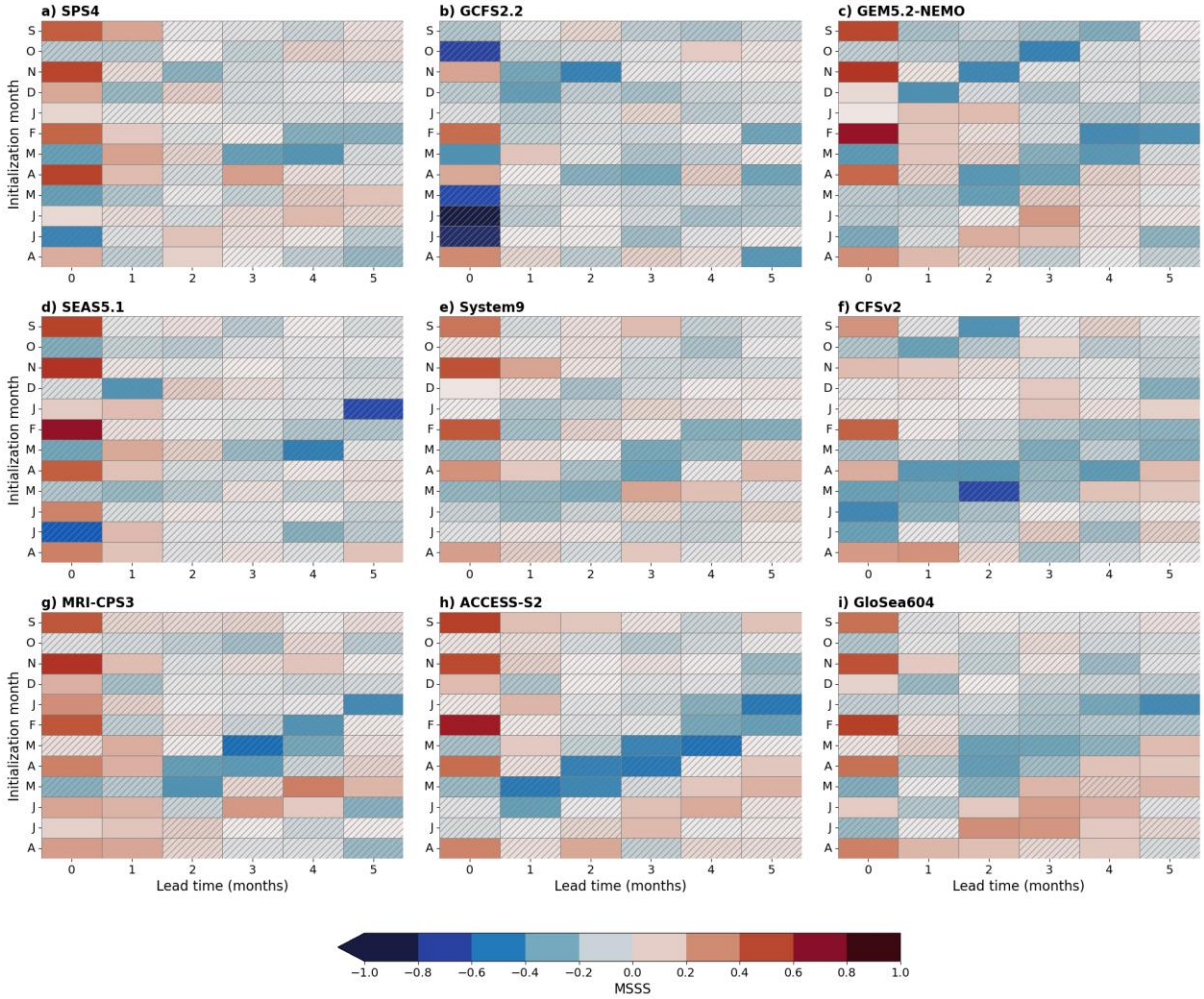

Figure S11. Same as Fig. S6 but for SLP in the South Atlantic region shown in Fig. 2.

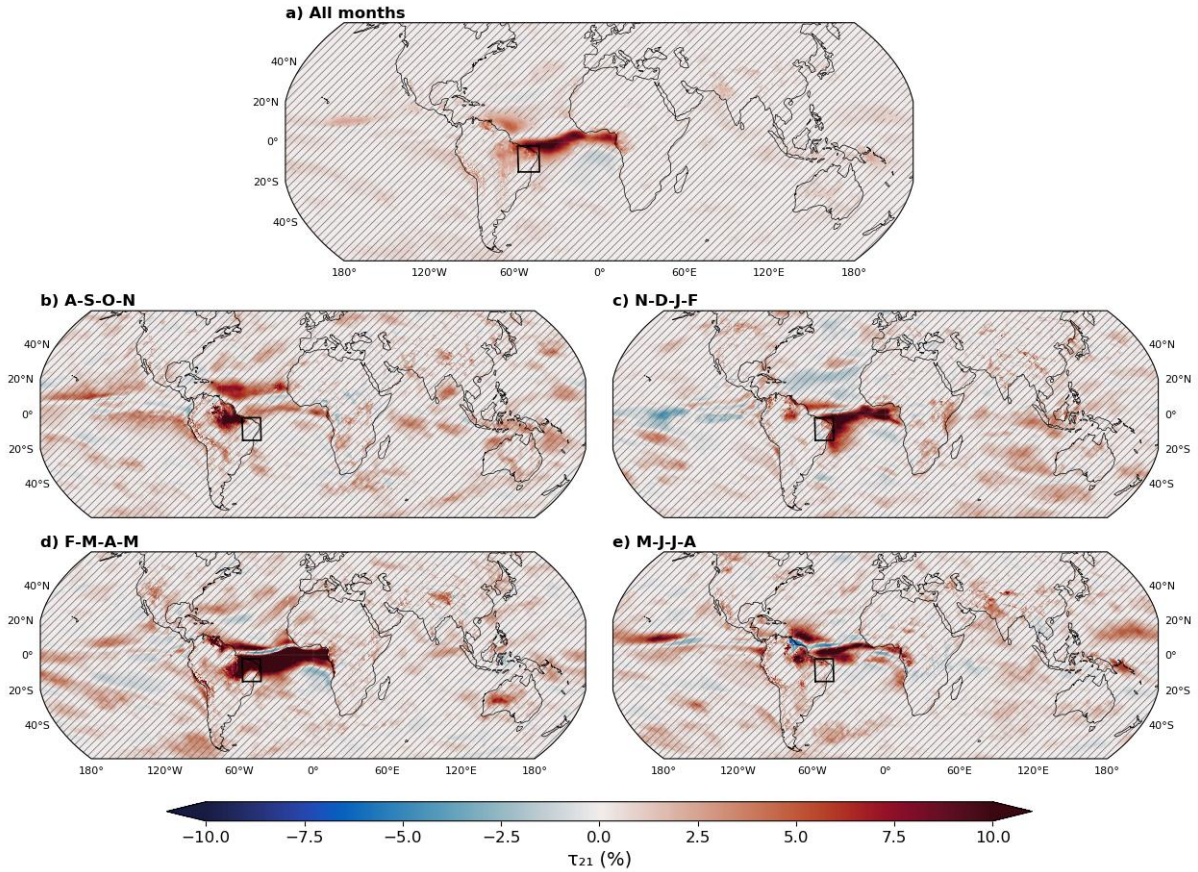

Figure S12. Normalised univariate information transfer, τ_{21}, between the second mode of tropical Atlantic SST (i.e., TATL SST EOF2) and surface wind at global scale for all months, a), and the periods N-D-J-F, b), F-M-A-M, c), M-J-J-A, d), and A-S-O-N, e), over 1980-2025 in the ERA5 reanalysis. Hatching indicates where τ_{21} is not statistically significant at the 95 % confidence level (α = 0.05), obtained through a bootstrap-based two-tailed Wald z-test.

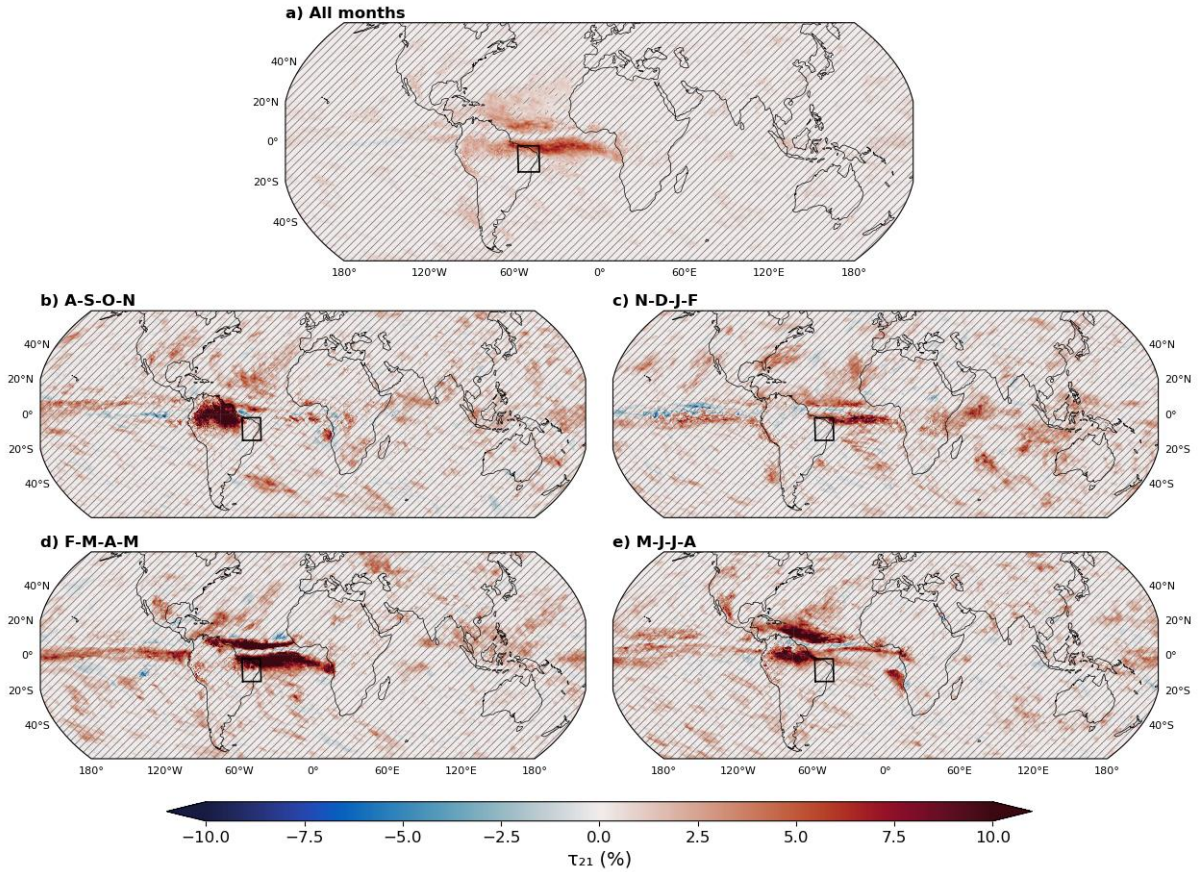

Figure S13. Same as Fig. S12 but for precipitation.

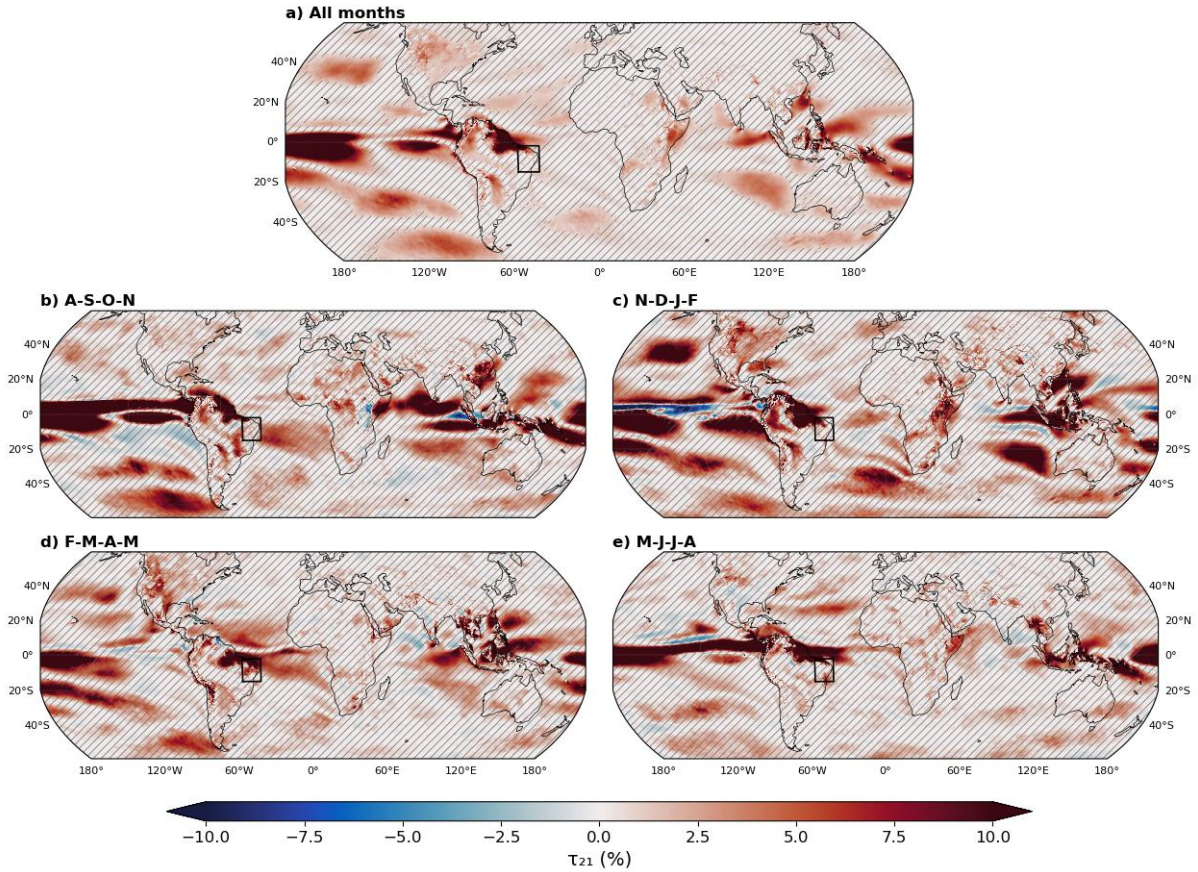

Figure S14. Same as Fig. S12 but for the first mode of tropical Pacific SST (i.e., TPAC SST EOF1) and surface wind.

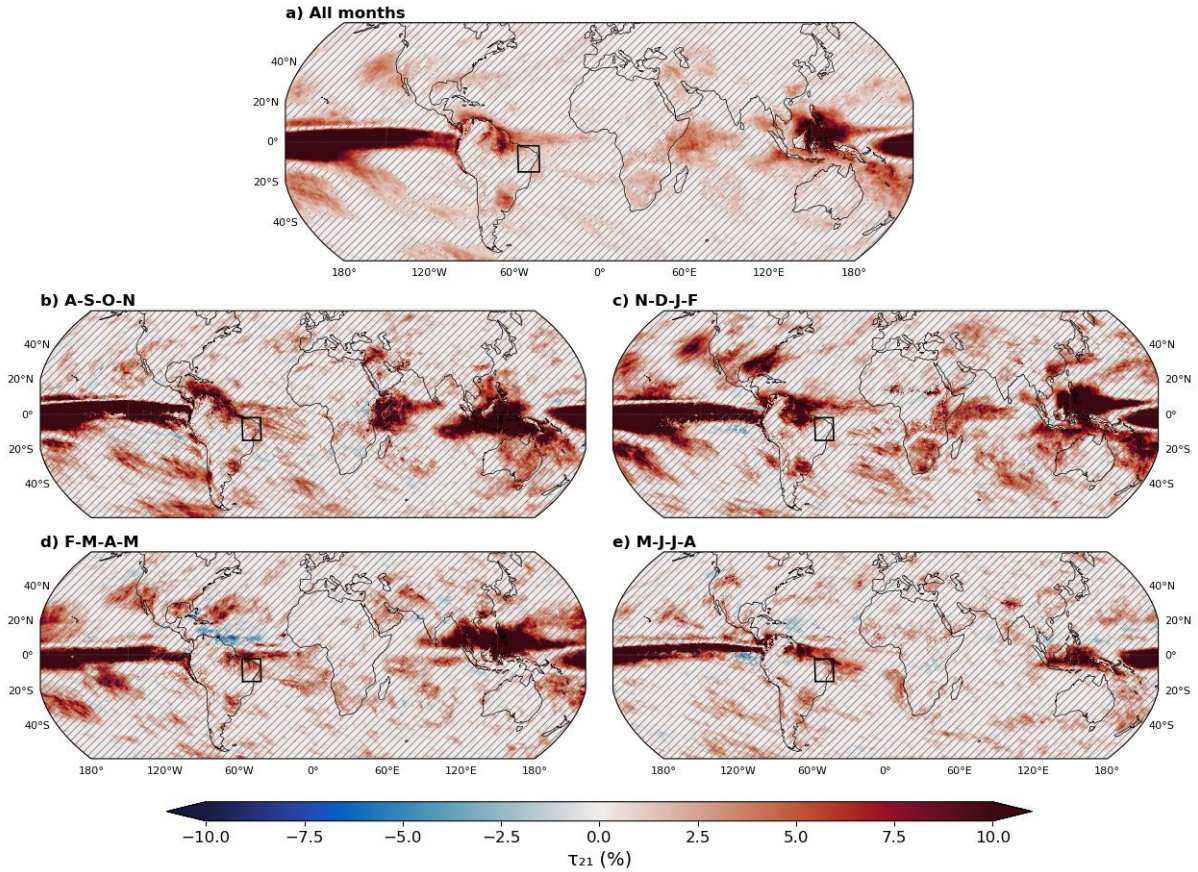

Figure S15. Same as Fig. S12 but for the first mode of tropical Pacific SST (i.e., TPAC SST EOF1) and precipitation.

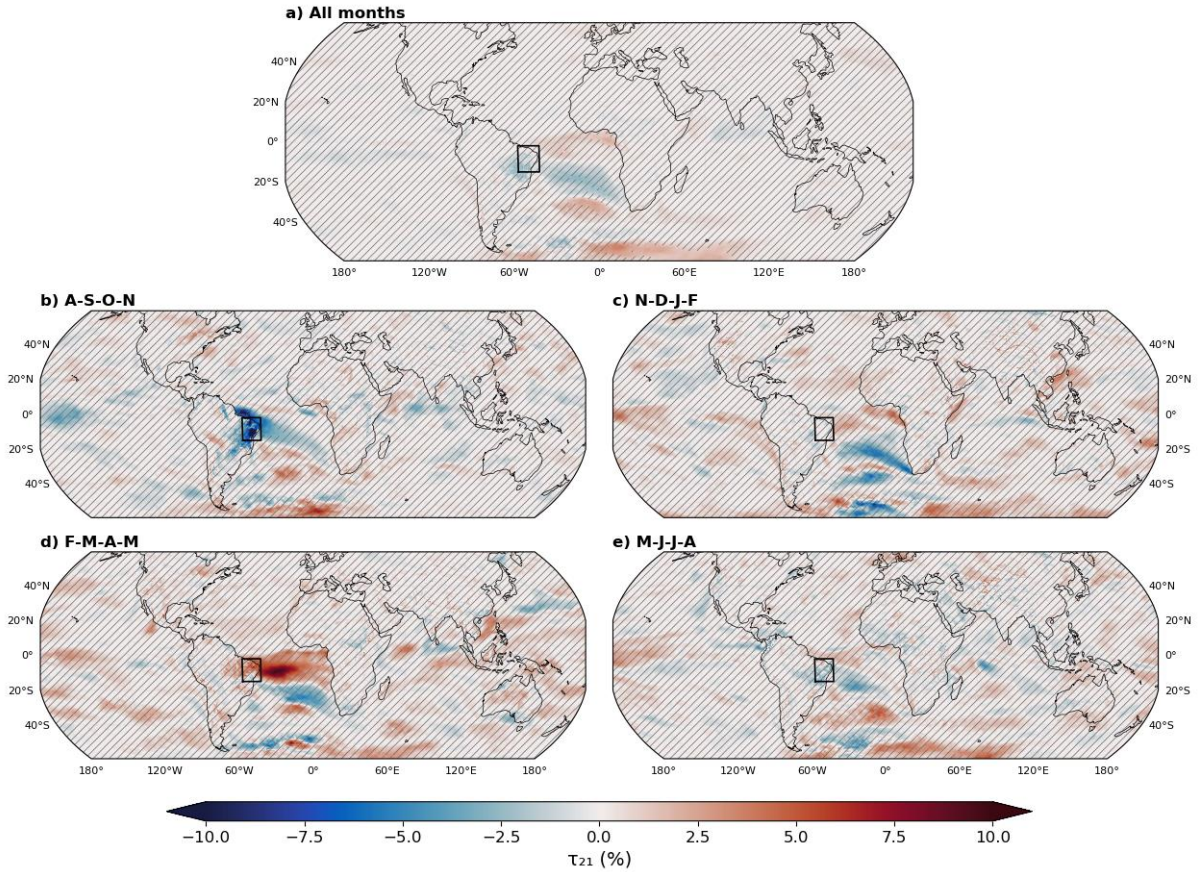

Figure S16. Same as Fig. S12 but for the third mode of South Atlantic SLP (i.e., SATL SLP EOF3) and surface wind.

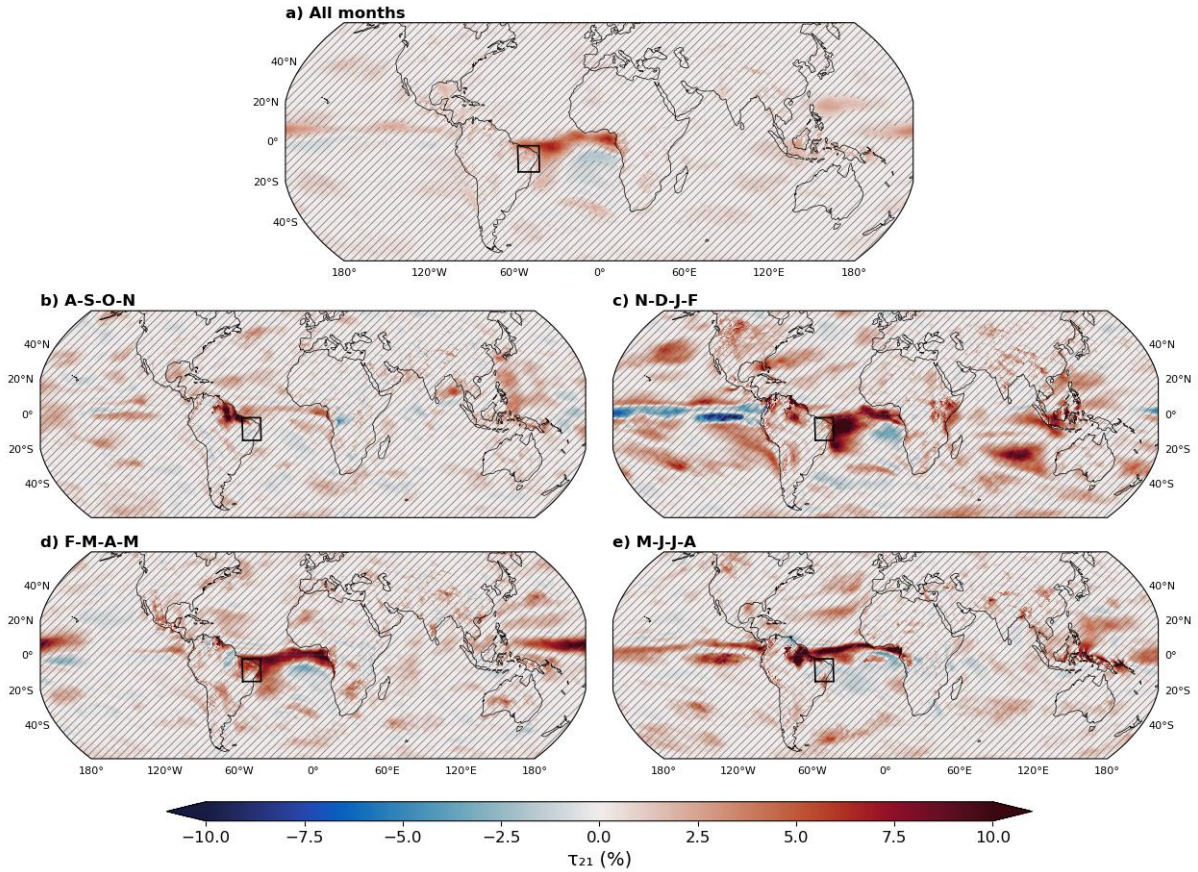

Figure S17. Same as Fig. S12 but for the first mode of South Atlantic SST (i.e., SATL SST EOF1) and surface wind.

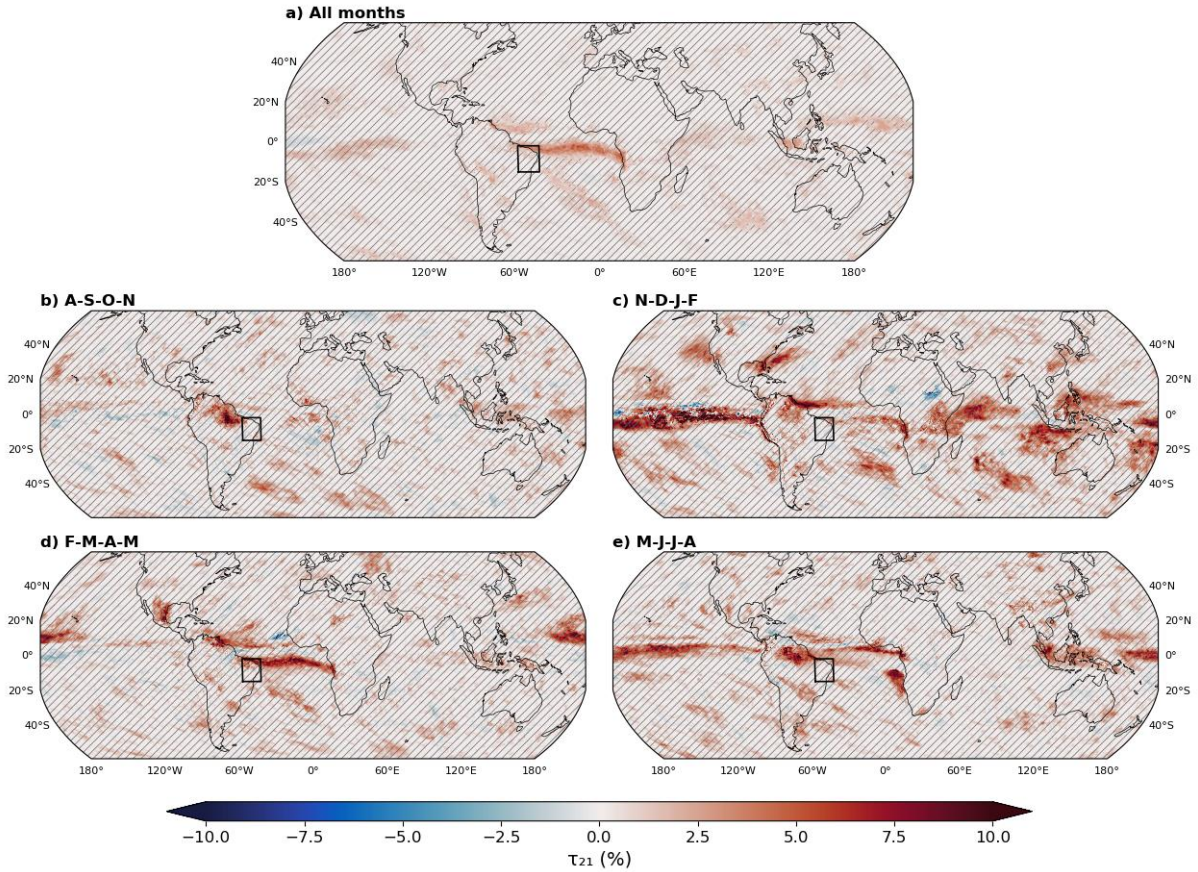

Figure S18. Same as Fig. S12 but for the first mode of South Atlantic SST (i.e., SATL SST EOF1) and precipitation.

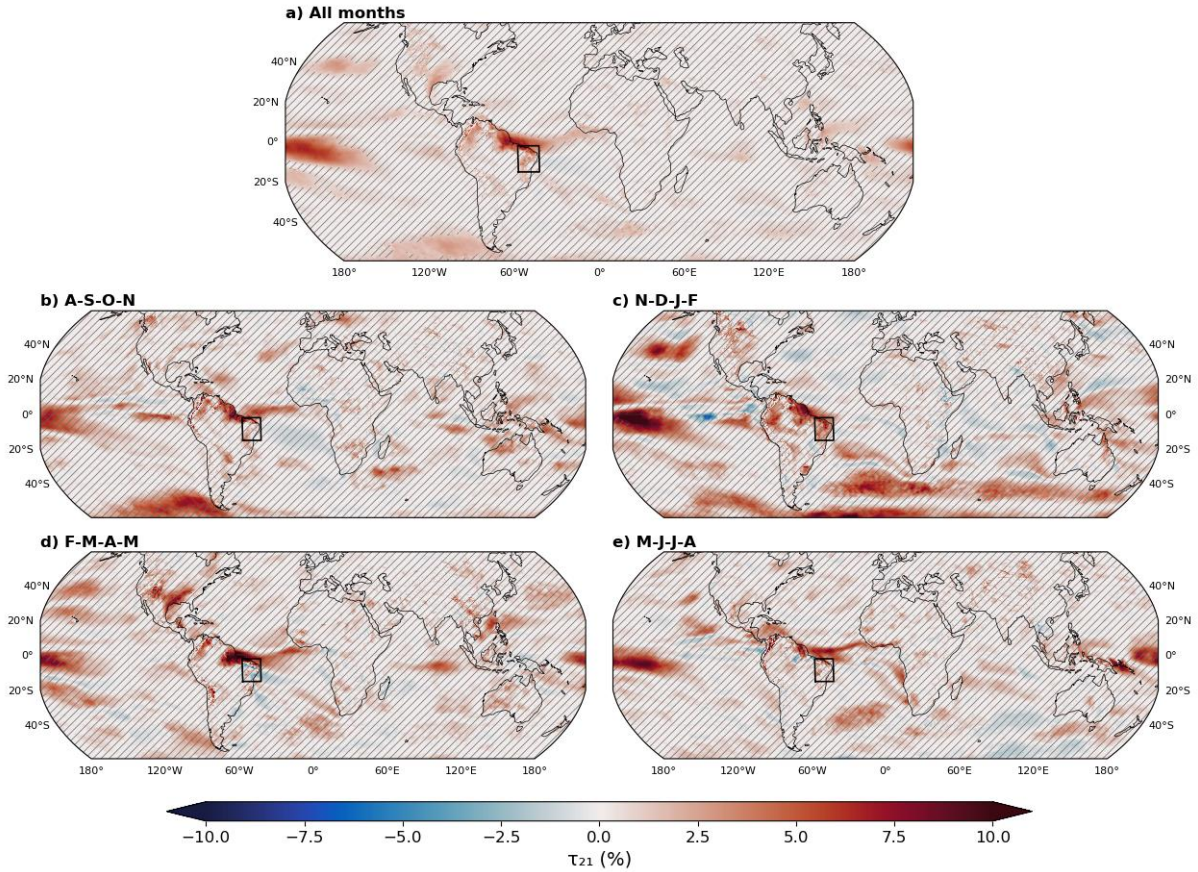

Figure S19. Same as Fig. S12 but for the second mode of South Atlantic SST (i.e., SATL SST EOF2) and surface wind.

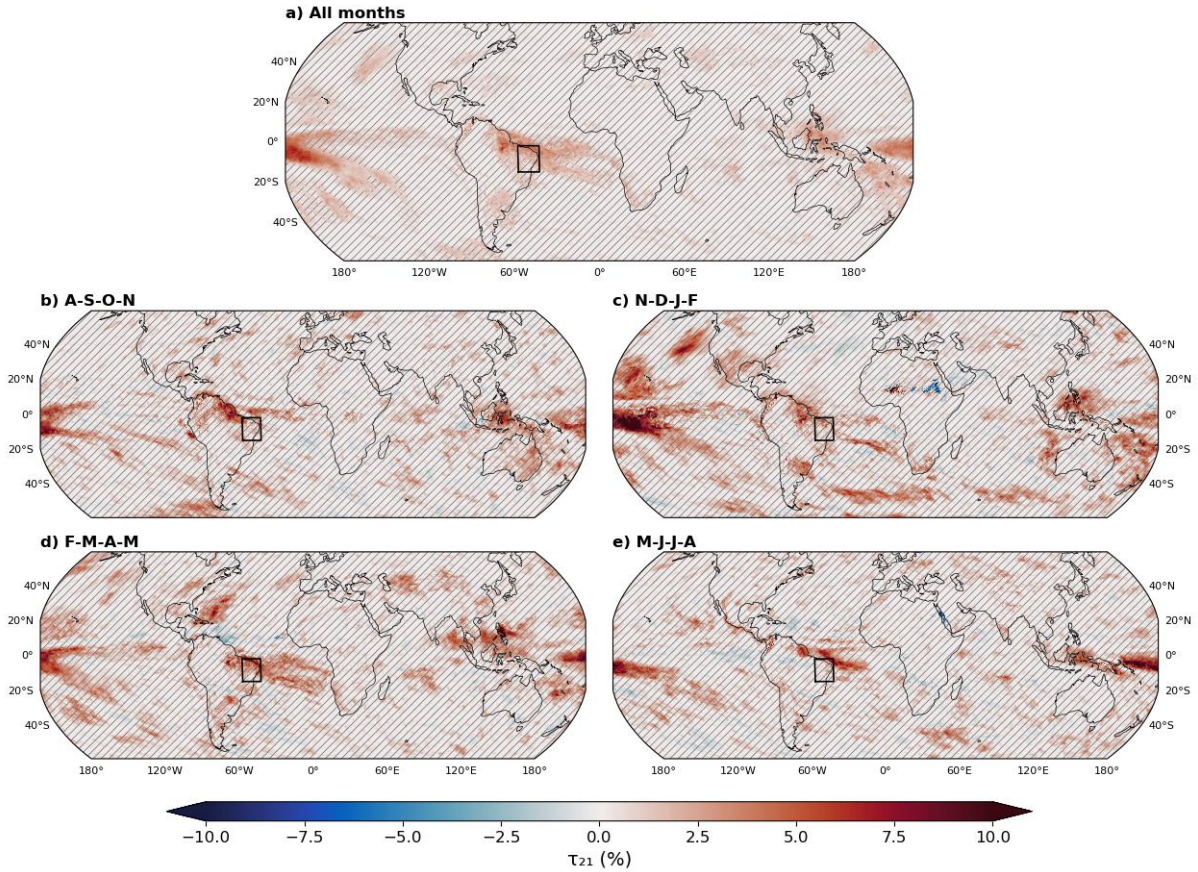

Figure S20. Same as Fig. S12 but for the second mode of South Atlantic SST (i.e., SATL SST EOF2) and precipitation.

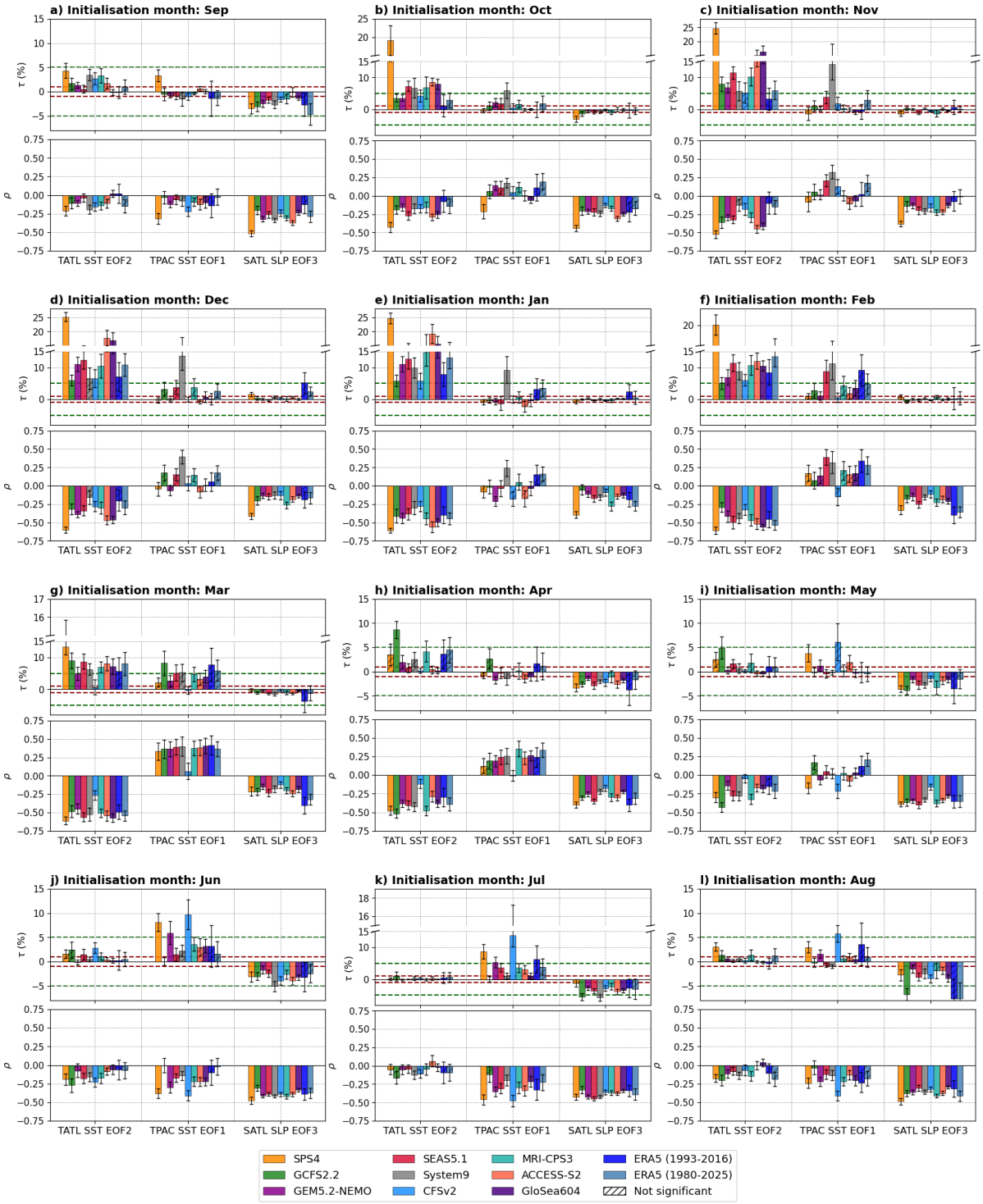

Figure S21. Normalised multivariate information transfer, τ (top rows), and Spearman correlation coefficient, ρ (bottom rows), between the dominant large-scale modes of NEB's climate variability (i.e., drivers) and surface wind in NEB (i.e., target) for all initialisation months, a)-l), over the common hindcast period (1993-2016) up to lead time 3 in the C3S multi-system, compared to ERA5 reanalysis over the periods 1993-2016 and 1980-2025. The drivers are the second mode of tropical Atlantic SST (left column), the first mode of tropical Pacific SST (middle column), and the third mode of South Atlantic SLP (right column). Hatching indicates where τ and ρ are not statistically significant at the 95 % confidence level ($\alpha = 0.05$), obtained through a bootstrap-based two-tailed Wald z-test. Error bars indicate bootstrap standard errors, calculated as the standard deviation across 1000 bootstrap resamples over

independent years. The red and green dashed horizontal lines indicate where $|\tau| = 1\,\%$ and $|\tau| = 5\,\%$, respectively. Note that the initialisation months go from September to August.

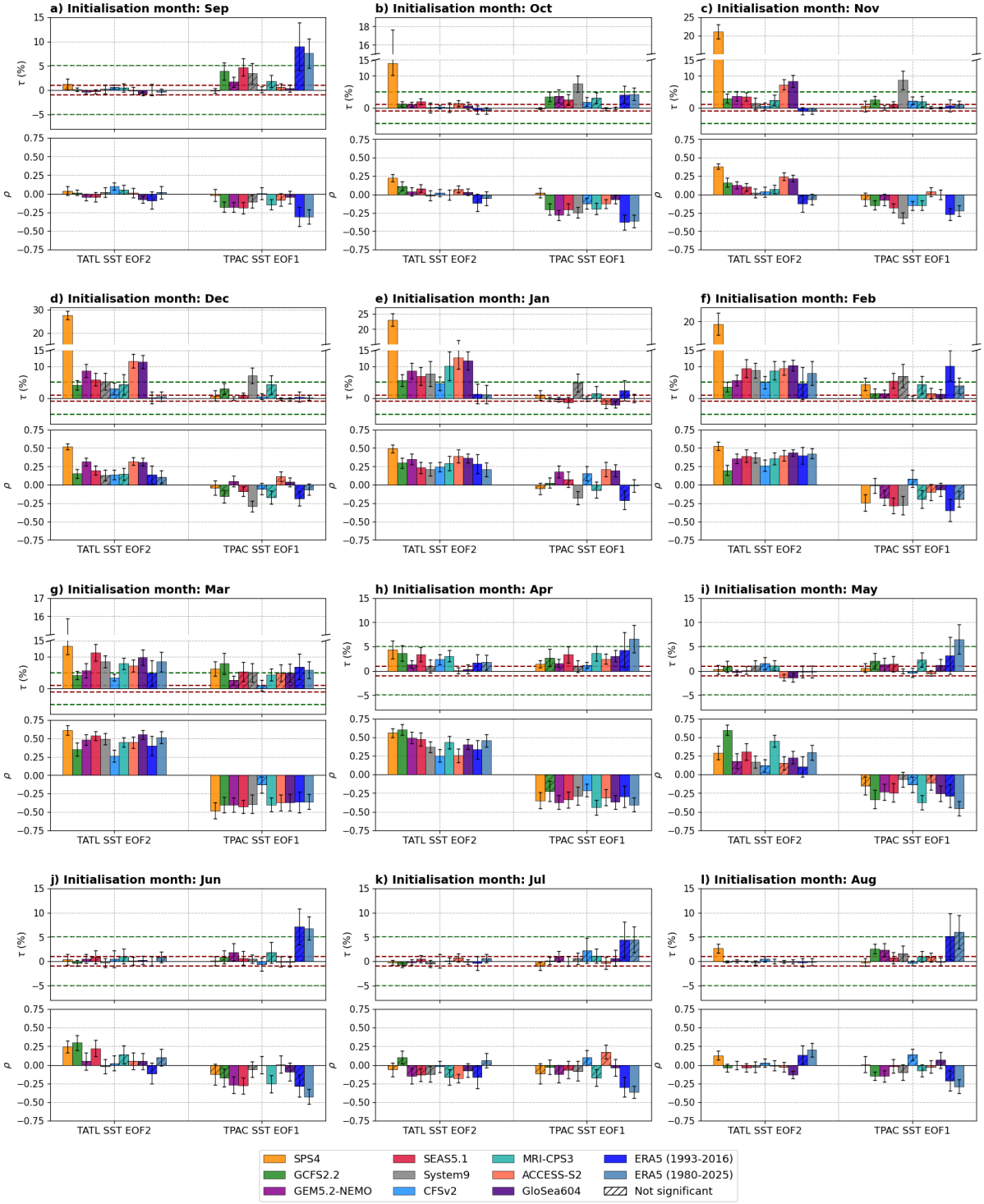

Figure S22. Normalised multivariate information transfer, τ (top rows), and Spearman correlation coefficient, ρ (bottom rows), between the dominant large-scale modes of NEB's climate variability (i.e., drivers) and precipitation in NEB (i.e., target) for all initialisation months, a)-l), over the common hindcast period (1993-2016) up to lead time 3 in the C3S multi-system, compared to ERA5 reanalysis over the periods 1993-2016 and 1980-2025. The drivers are the second mode of tropical Atlantic SST (left column) and the first mode of tropical Pacific SST (right column). Hatching indicates where τ or ρ are not statistically significant at the 95 % confidence level (α = 0.05), obtained

through a bootstrap-based two-tailed Wald z-test. Error bars indicate bootstrap standard errors, calculated as the standard deviation across 1000 bootstrap resamples over independent years. The red and green dashed horizontal lines indicate where $|\tau| = 1$ % and $|\tau| = 5$ %, respectively. Note that the initialisation months go from September to August.

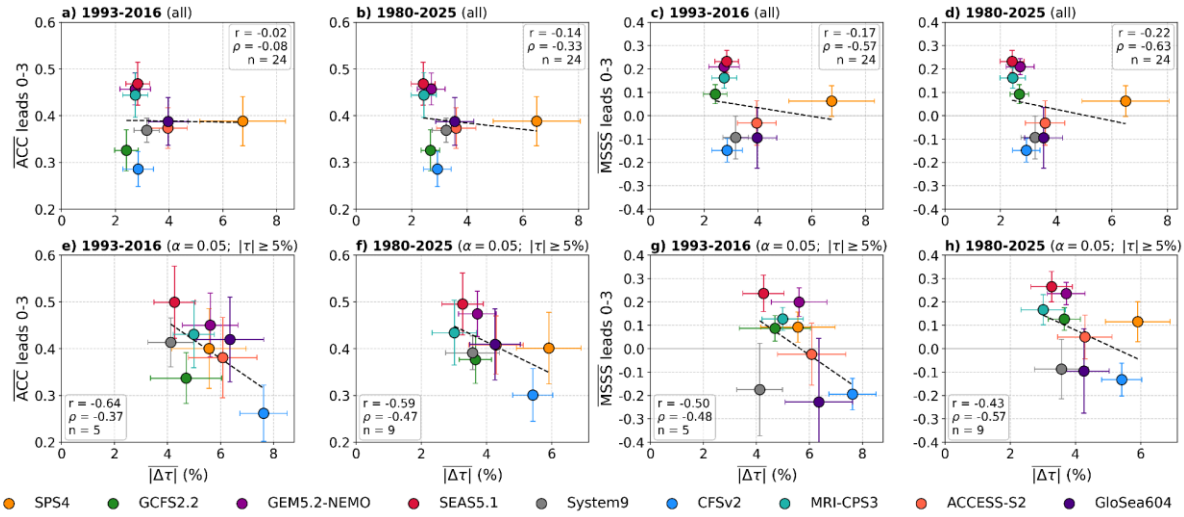

Figure S23. Relationship between causal fidelity, $\overline{|\Delta\tau|}$, and average forecast skill over lead times 0-3, in terms of ACC, a)-b), e)-f), and MSSS, c)-d), g)-h), for precipitation in NEB across the C3S multi-system. For each metric, the left column uses ERA5 over 1993-2016 as the causal reference and the right column ERA5 over 1980-2025. For each system, $|\Delta\tau|$ is averaged over all initialisation months and drivers (i.e., AMM, ENSO, SASH meridional displacements). The top panels, a)-d), use all initialisation months, whereas the bottom panels, e)-h), retain only those for which the ERA5 causal link is either significant at the 95 % confidence level ($\alpha = 0.05$), obtained through a bootstrap-based two-tailed Wald z-test, or for which $|\tau| \geq 5$ %. The black dashed line is the ordinary least-squares fit across systems. The legend boxes include the Pearson, r, and Spearman, ρ, correlation coefficients, and the number of init-month × driver samples, n, entering the $\overline{|\Delta\tau|}$ average. Error bars show the standard error of the mean, taken over n for $\overline{|\Delta\tau|}$ and over the retained initialisation months for the corresponding skill metric.

Variable	Reference	Links	$r(\overline{\lvert\Delta\tau\rvert}, P)$	$r(\overline{\lvert\Delta\tau\rvert}, A)$	$r(\overline{\lvert\Delta\tau\rvert}, B)$	$\bar{P}$	$\bar{A}$	$\bar{B}$	$\bar{Q}$
Precipitation	1993–2016	All	0.16	0.31	−0.01	0.212	0.169	0.014	0.015
		Retained	−0.43	0.43	0.07	0.210	0.197	0.019	0.020
	1980–2025	All	0.05	0.31	−0.00	0.212	0.169	0.014	0.015
		Retained	−0.25	0.45	0.02	0.239	0.168	0.016	0.017
Surface wind	1993–2016	All	0.20	**0.88**	0.04	0.165	0.142	0.028	0.028
		Retained	0.06	**0.67**	0.04	0.172	0.129	0.023	0.024
	1980–2025	All	−0.10	**0.94**	0.13	0.165	0.142	0.028	0.028
		Retained	−0.26	**0.83**	−0.03	0.176	0.146	0.027	0.027

Table S1. Pearson correlations between causal fidelity, $\overline{\lvert\Delta\tau\rvert}$, and the terms of the MSSS decomposition (Sect. S3) across the C3S seasonal forecast systems for phase, P, amplitude, A, and mean bias, B. The last four columns give the multi-system means of each term, where MSSS = P - A - B + Q. In the links column, "All" includes all initialisations, whereas "Retained" includes only those where the ERA5 causal link is either significant at the 95 % confidence level (α = 0.05), obtained through a bootstrap-based two-tailed Wald z-test, or for which ≥ 5 %. Bold values indicate when r is significant at the 95 % confidence level (α = 0.05), obtained through a standard two-tailed Student's t-test.

System	$\overline{\lvert\Delta\tau\rvert}$ (%)		A		s		r		$s-r$		Members
	Ret.	All	Ret.	All	Ret.	All	Ret.	All	Ret.	All	
SEAS5.1	2.30	1.49	0.10	0.09	0.60	0.56	0.41	0.37	0.19	0.19	25
MRI-CPS3	2.55	1.40	0.06	0.07	0.46	0.44	0.38	0.33	0.08	0.11	10
System9	2.75	2.44	0.18	0.15	0.65	0.58	0.34	0.33	0.31	0.25	31
GEM5.2-NEMO	3.07	1.71	0.10	0.10	0.53	0.53	0.39	0.37	0.13	0.16	20
GCFS2.2	3.55	1.94	0.16	0.16	0.54	0.54	0.26	0.26	0.28	0.28	30
GloSea6	4.06	2.05	0.11	0.11	0.55	0.52	0.33	0.28	0.23	0.24	28
ACCESS-S2	4.61	2.31	0.19	0.17	0.64	0.61	0.37	0.31	0.27	0.30	27
CFSv2	5.08	2.88	0.14	0.16	0.58	0.62	0.25	0.28	0.33	0.34	20
SPS4	8.30	4.09	0.29	0.28	0.76	0.77	0.34	0.35	0.42	0.42	30

Table S2. System values for surface wind against the 1980–2025 reference, for retained initialisations (Ret.) and all initialisations (All), as defined in Table S1. From left to right: causal fidelity, $\overline{\lvert\Delta\tau\rvert}$, amplitude term of the MSSS, A, ratio of hindcast to reference standard deviation, s, hindcast–reference Pearson correlation, r, amplitude excess, $s-r$, and number of ensemble members. Values are averaged over the corresponding initialisation months and lead times 0–3. Systems are sorted by increasing causal fidelity for retained initialisations.

Variable	Reference	Links	No control		Driver skill (EOF)		Driver skill (index)		Ensemble size	
			$r(x,y)$	$\rho(x,y)$	$r(x,y\|z)$	$\rho(x,y\|z)$	$r(x,y\|z)$	$\rho(x,y\|z)$	$r(x,y\|n)$	$\rho(x,y\|n)$
Precipitation	1993–2016	All	−0.17	−0.57	−0.12	−0.37	−0.22	−0.69	−0.04	−0.51
		Retained	−0.50	−0.48	0.33	−0.11	−0.46	−0.47	−0.59	**−0.73**
	1980–2025	All	−0.22	−0.63	−0.19	−0.41	−0.25	−0.69	−0.07	−0.58
		Retained	−0.43	−0.57	−0.05	−0.05	−0.42	−0.56	−0.37	−0.54
Surface wind	1993–2016	All	**−0.68**	−0.50	**−0.76**	−0.46	**−0.92**	−0.68	−0.59	−0.35
		Retained	−0.47	−0.18	−0.30	−0.00	**−0.80**	−0.52	−0.45	−0.35
	1980–2025	All	**−0.85**	**−0.87**	**−0.85**	**−0.84**	**−0.94**	**−0.88**	**−0.81**	**−0.83**
		Retained	**−0.77**	**−0.70**	**−0.70**	−0.53	**−0.85**	−0.63	**−0.76**	**−0.76**

Table S3. Pearson, r, and Spearman, ρ, correlations between $\overline{|\Delta\tau|}$, x, and average MSSS, y, across the C3S forecast systems, without control and as partial correlations controlling for driver skill, z, or ensemble size, n. Driver skill is estimated from the MSSS of area-weighted means over each EOF domain (EOF) and from the MSSS of the Niño3.4 and TAD indices (index) averaged over the same initialisation–driver cells as $\overline{|\Delta\tau|}$. The standard index control excludes the SASH, as no standard index exists for this mode. For precipitation, x and z estimated from EOF-domain averages are strongly correlated for the retained initialisations, so their contributions cannot be separated. "All" and "Retained" follow the same definition as in Table S1. Bold values indicate when the correlation is significant at the 95 % confidence level (α = 0.05), obtained through a standard two-tailed Student's t-test.

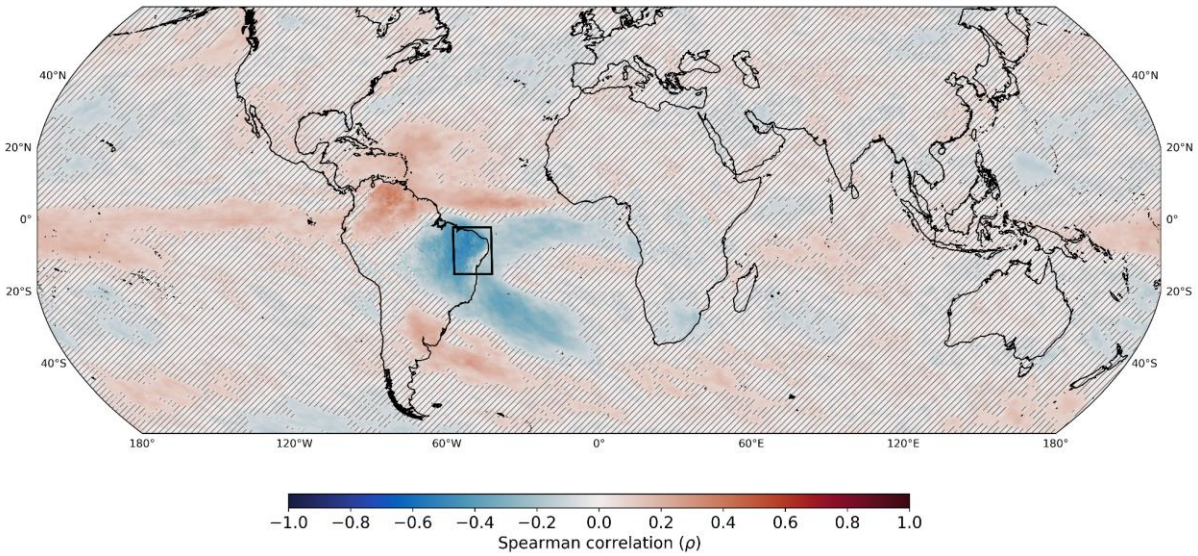

Figure S24. One-point Spearman correlation of surface wind monthly anomalies in NEB against global precipitation monthly anomalies over the period 1980-2025 in the ERA5 reanalysis. The anomaly fields have been detrended as described in Sect. 2.1. The red rectangle highlights the location of NEB. Hatching indicates where ρ is not statistically significant at the 95 % confidence level (α = 0.05), obtained through a standard two-tailed Student's t-test.

References

Murphy, A. H. (1988). Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient. Monthly Weather Review, 116(12), 2417-2424. https://doi.org/10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2