Supplement of Adv. Stat. Clim. Meteorol. Oceanogr., 11, 107–121, 2025
https://doi.org/10.5194/ascmo-11-107-2025-supplement
© Author(s) 2025. CC BY 4.0 License.

Supplement of

Comparison of sea surface temperatures and marine air temperatures in the tropical Pacific

Peter F. Craigmile and Peter Guttorp

Correspondence to: Peter F. Craigmile (peter.craigmile@hunter.cuny.edu)

The copyright of individual parts of the supplement might differ from the article licence.

S1 Supplemental tables and figures

Table S1 shows the percentage of quality codes for the observed buoy-based measurements of sea surface temperature (SST) and marine air temperature (MAT), summarized over all locations over the years 1992–2020.

Table S1: Rounded to one decimal place, the percentage of quality codes for the observed buoy-based measurements of sea surface temperature (SST) and marine air temperature (MAT), summarized over all locations over the years 1992–2020.

Quality code	SST	MAT
1 (Good data)	15.0%	13.3%
2 (Probably good data)	80.1%	80.8%
3 (Questionable data)	4.8%	5.9%
5 (Adjusted data)	0.1%	$<$0.1%

Figure S1 displays the percentage of missingness of the daily buoy measurements of sea surface temperature (SST), marine atmospheric temperature (MAT), and their difference (SST−MAT) over the years 1992–2020. Figure S2 displays the corresponding percentage of missingness of the monthly buoy measurements, but now over the years 1996–2018.

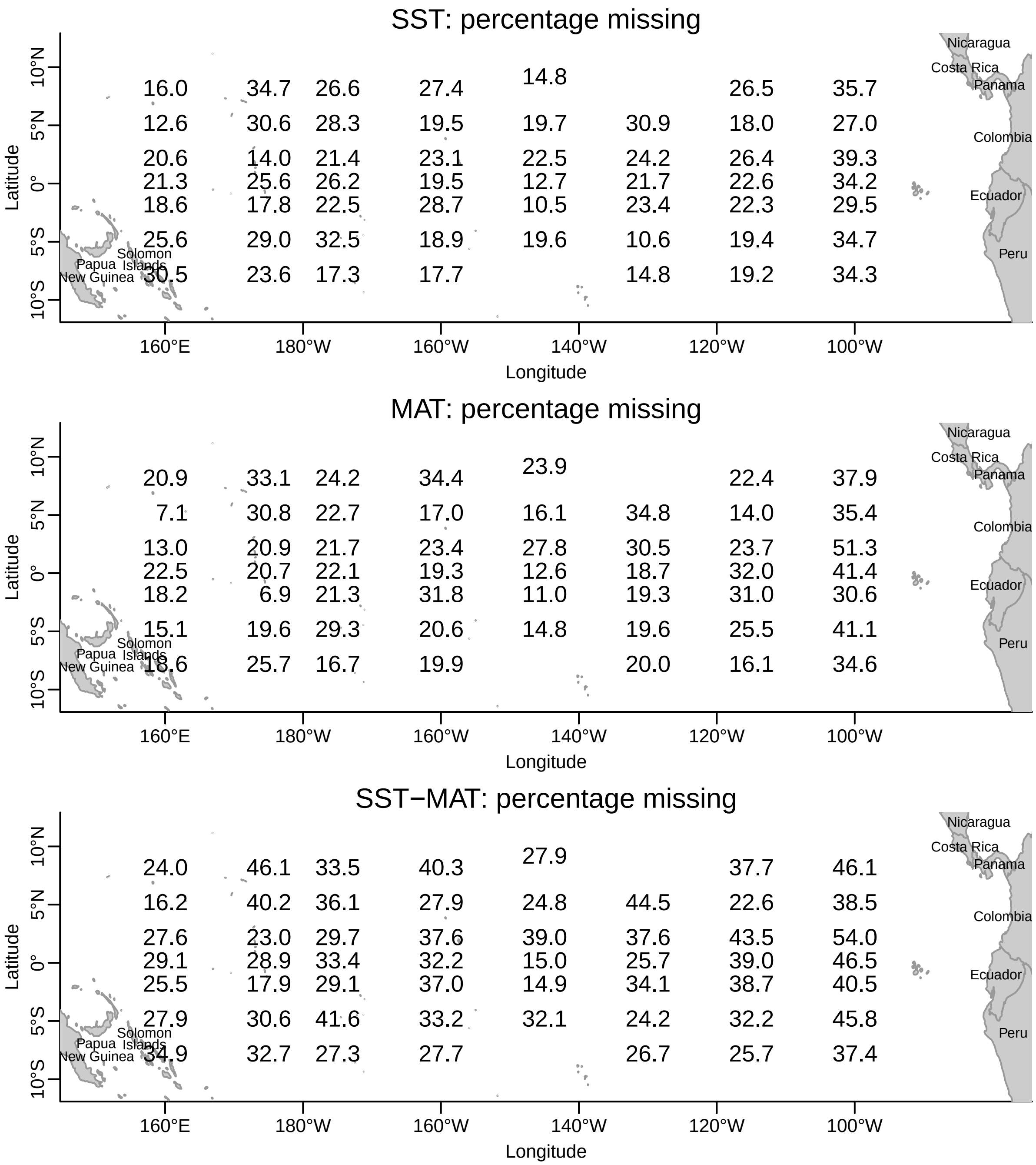

Figure S1: Top panel displays the percentage of days of buoy-based measurements of sea surface temperature (SST) that are missing by location over the years 1992–2020. The middle panel displays the percentage of days of buoy-based measurements of marine atmospheric temperature (MAT) that are missing by location over the same period. The bottom panel displays the percentage of days that the differences (SST−MAT) are missing by location.

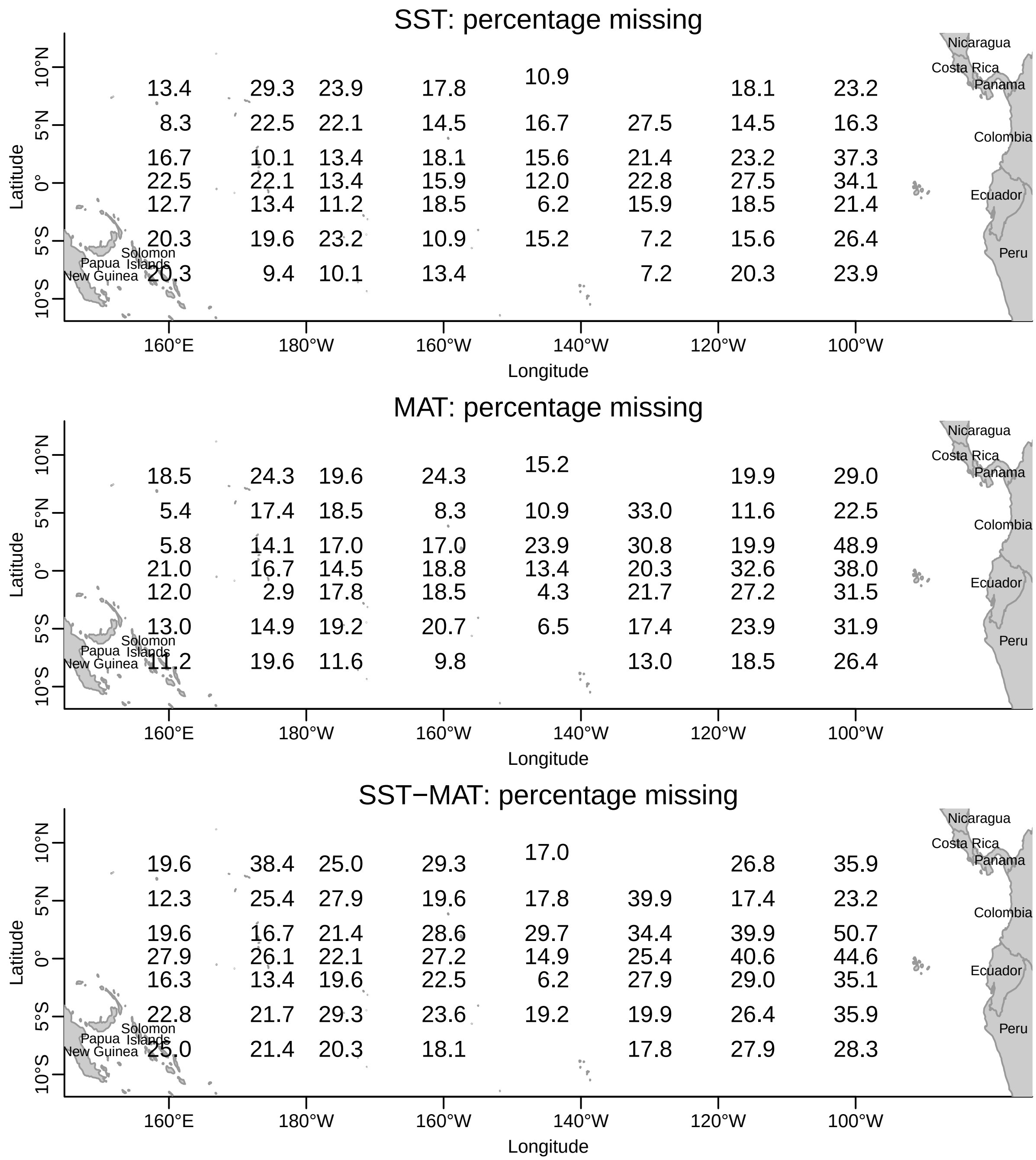

Figure S2: The percentage of missingness of the monthly buoy measurements, over the years 1996–2018 (top panel: SST; middle panel: MAT; bottom panel: SST−MAT).

S2 The MCMC algorithm

Before we present the MCMC algorithm to sample from the posterior distribution of the parameters given the data, we write the model in vector-matrix notation.

We have m spatial locations, $\{\boldsymbol{s}_j : j = 1, \ldots, m\}$ and the entire time period is indexed from $t = 1, \ldots, N$. Let $o(\boldsymbol{s}_j, t) = 1$ if we observe differences $Z(\boldsymbol{s}_j, t)$ at location $\boldsymbol{s}_j$ and time t; otherwise let $o(\boldsymbol{s}_j, t) = 0$ if there is no observed difference.

Let $\boldsymbol{Z}_t = (Z(\boldsymbol{s}_j, t) : j = 1, \ldots, m;\ o(\boldsymbol{s}_j, t) = 1)^T$ denote the vector of observations available at time point t and let $l_t = |\boldsymbol{Z}_t|$ denote how many observations at time point t. We let $\boldsymbol{Y}_t = (Y(\boldsymbol{s}_j, t) : j = 1, \ldots, m)^T$ denote the latent spatio-temporal difference process at time t over all spatial locations and let $\boldsymbol{A}_t$ be a $n_j \times m$ matrix of ones and zeros that maps the spatial locations from $\boldsymbol{Y}_t$ to $\boldsymbol{Z}_t$ for each time point t.

Conditionally independent over t, we have that

$$\boldsymbol{Z}_t | \boldsymbol{Y}_t \sim N_{l_t}(\boldsymbol{A}_t \boldsymbol{Y}_t, \sigma^2 \boldsymbol{I}_{l_t}), \quad t = 1, \ldots, N.$$

For each t, the jth element of the m-vector $\boldsymbol{\mu}_t$ is

$$[\mu_t]_j = \sum_{k=1}^{K} \beta_k(\boldsymbol{s}_j) x_{k,t}.$$

Also let $\boldsymbol{\Phi} = \mathrm{diag}(h(\eta(\boldsymbol{s}_j)) : j = 1, \ldots, m)$. Next, let $\boldsymbol{\Sigma}_1$ denote the $m \times m$ matrix with (j, j') element $\mathrm{cov}(\zeta(\boldsymbol{s}_j, 1), \mathrm{cov}(\boldsymbol{s}_{j'}, 1))$, and $\boldsymbol{\Sigma}$ denote the $m \times m$ matrix with (j, j') element $\mathrm{cov}(\zeta(\boldsymbol{s}_j, 2), \mathrm{cov}(\boldsymbol{s}_{j'}, 2))$. Then

$$\boldsymbol{Y}_1 \sim N_m(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1),$$

independent of the sequence of conditionally independent random vectors:

$$\boldsymbol{Y}_t | \boldsymbol{Y}_{t-1} \sim N_m(\boldsymbol{\mu}_t + \boldsymbol{\Phi}[\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}_{t-1}], \boldsymbol{\Sigma}), \quad t = 2, \ldots, N.$$

S2.1 Updating $\boldsymbol{Y}_t$, $t = 1, \ldots, N$:

S2.1.1 Updating $\boldsymbol{Y}_1$:

We have

$$\begin{aligned}
\boldsymbol{Z}_1 &\sim N_{l_1}(\boldsymbol{A}_1 \boldsymbol{Y}_1, \sigma^2 \boldsymbol{I}_{l_1}); \\
\boldsymbol{Y}_1 &\sim N_m(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1); \\
\boldsymbol{Y}_2 | \boldsymbol{Y}_1 &\sim N_m(\boldsymbol{\mu}_2 + \boldsymbol{\Phi}(\boldsymbol{Y}_1 - \boldsymbol{\mu}_1), \boldsymbol{\Sigma}).
\end{aligned}$$

Thus

$$\boldsymbol{Y}_2 - \boldsymbol{\mu}_2 + \boldsymbol{\Phi}\boldsymbol{\mu}_1 = N_m(\boldsymbol{\Phi}\boldsymbol{Y}_1, \boldsymbol{\Sigma}).$$

Hence, to sample from $\boldsymbol{Y}_1$ given everything else we draw from $N_m(\boldsymbol{P}^{-1}\boldsymbol{q}, \boldsymbol{P}^{-1})$ where

$$\begin{aligned}
\boldsymbol{P} &= \frac{\boldsymbol{A}_1^T \boldsymbol{A}_1}{\sigma^2} + \boldsymbol{\Sigma}^{-1} + \boldsymbol{\Phi}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Phi}, \text{ and} \\
\boldsymbol{q} &= \frac{\boldsymbol{A}_1^T \boldsymbol{Z}_1}{\sigma^2} + \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1 + \boldsymbol{\Phi}^T \boldsymbol{\Sigma}^{-1} \left[\boldsymbol{Y}_2 - \boldsymbol{\mu}_2 + \boldsymbol{\Phi}\boldsymbol{\mu}_1\right].
\end{aligned}$$

S2.1.2 Updating $\boldsymbol{Y}_t$, $t = 2, \ldots, N-1$:

We have

$$\begin{aligned}
\boldsymbol{Z}_t &\sim N_{l_t}(\boldsymbol{A}_t \boldsymbol{Y}_t, \sigma^2 \boldsymbol{I}_{l_t}); \\
\boldsymbol{Y}_t | \boldsymbol{Y}_{t-1} &\sim N_m(\boldsymbol{\mu}_t + \boldsymbol{\Phi}(\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}_{t-1}), \boldsymbol{\Sigma}); \\
\boldsymbol{Y}_{t+1} | \boldsymbol{Y}_t &\sim N_m(\boldsymbol{\mu}_{t+1} + \boldsymbol{\Phi}(\boldsymbol{Y}_t - \boldsymbol{\mu}_t), \boldsymbol{\Sigma}).
\end{aligned}$$

Thus

$$\boldsymbol{Y}_{t+1} - \boldsymbol{\mu}_{t+1} + \boldsymbol{\Phi}\boldsymbol{\mu}_t = N_m(\boldsymbol{\Phi}\boldsymbol{Y}_t, \boldsymbol{\Sigma}).$$

Hence, to sample from $\boldsymbol{Y}_t$ given everything else we draw from $N_m(\boldsymbol{P}^{-1}\boldsymbol{q}, \boldsymbol{P}^{-1})$ where

$$\begin{aligned}
\boldsymbol{P} &= \frac{\boldsymbol{A}_t^T \boldsymbol{A}_t}{\sigma^2} + \boldsymbol{\Sigma}^{-1} + \boldsymbol{\Phi}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Phi}, \text{ and} \\
\boldsymbol{q} &= \frac{\boldsymbol{A}_t^T \boldsymbol{Z}_t}{\sigma^2} + \boldsymbol{\Sigma}^{-1} \left[\boldsymbol{\mu}_t + \boldsymbol{\Phi}(\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}_{t-1})\right] + \boldsymbol{\Phi}^T \boldsymbol{\Sigma}^{-1} \left[\boldsymbol{Y}_{t+1} - \boldsymbol{\mu}_{t+1} + \boldsymbol{\Phi}\boldsymbol{\mu}_t\right].
\end{aligned}$$

S2.1.3 Updating $\boldsymbol{Y}_N$:

In this case

$$\begin{aligned} \boldsymbol{Z}_N &\sim N_{l_N}(\boldsymbol{A}_N \boldsymbol{Y}_N, \sigma^2 \boldsymbol{I}_{l_N}); \\ \boldsymbol{Y}_N | \boldsymbol{Y}_{N-1} &\sim N_m(\boldsymbol{\mu}_N + \boldsymbol{\Phi}(\boldsymbol{Y}_{N-1} - \boldsymbol{\mu}_{N-1}), \boldsymbol{\Sigma}). \end{aligned}$$

Hence, to sample from $\boldsymbol{Y}_N$ given everything else we draw from $N_m(\boldsymbol{P}^{-1}\boldsymbol{q}, \boldsymbol{P}^{-1})$ where

$$\begin{aligned} \boldsymbol{P} &= \frac{\boldsymbol{A}_N^T \boldsymbol{A}_N}{\sigma^2} + \boldsymbol{\Sigma}^{-1}, \text{ and} \\ \boldsymbol{q} &= \frac{\boldsymbol{A}_N^T \boldsymbol{Z}_N}{\sigma^2} + \boldsymbol{\Sigma}^{-1} \left[\boldsymbol{\mu}_N + \boldsymbol{\Phi}(\boldsymbol{Y}_{N-1} - \boldsymbol{\mu}_{N-1}) \right]. \end{aligned}$$

S2.2 Updating σ^2:

Let

$$S = \sum_{j=1}^{m} \sum_{t \in O_j} [Z(\boldsymbol{s}_j, t) - Y(\boldsymbol{s}_j, t)]^2$$

with $L = \sum_{j=1}^m |O_j|$. Then if the prior for σ^2 is $\text{InvGamma}(\alpha_{\sigma^2},\ \beta_{\sigma^2})$ to sample from σ^2 given all the other parameters, we sample from

$$\text{InvGamma}(\alpha_{\sigma^2} + L/2, \beta_{\sigma^2} + S/2).$$

S2.3 Updating $\boldsymbol{\eta}$:

We have

$$\boldsymbol{Y}_1 \sim N_m(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1),$$

and

$$\boldsymbol{Y}_t | \boldsymbol{Y}_{t-1} \sim N_m(\boldsymbol{\mu}_t + \boldsymbol{\Phi}(\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}_{t-1}), \boldsymbol{\Sigma}), \quad t = 2, \ldots, N,$$

with $\boldsymbol{\Phi} = \mathrm{diag}(h(\eta(\boldsymbol{s}_j)) : j = 1, \ldots, m)$, where we assume that

$$\boldsymbol{\eta} = (\eta(\boldsymbol{s}_j) : j = 1, \ldots, m)^T \quad \sim \quad N_m(\boldsymbol{\mu}_\eta, \boldsymbol{V}_\eta).$$

We update $\boldsymbol{\eta}$ given the other parameters using a random walk Metropolis-Hastings algorithm. Let $\boldsymbol{\eta}^c$ be the current value of $\boldsymbol{\eta}$ and we propose a new value via

$$\boldsymbol{\eta}^n \sim N_m(\boldsymbol{\eta}^c, \boldsymbol{\Lambda}_\eta),$$

for some proposal covariance matrix $\boldsymbol{\Lambda}_\eta$. Let $\boldsymbol{\Phi}^n = \mathrm{diag}(h(\eta^n(\boldsymbol{s}_j)) : j = 1, \ldots, m)$ and $\boldsymbol{\Phi}^c = \mathrm{diag}(h(\eta^c(\boldsymbol{s}_j)) : j = 1, \ldots, m)$. Let $\boldsymbol{\Sigma}^n$ be the covariance matrix $\boldsymbol{\Sigma}$ calculated above using the new values of $\boldsymbol{\eta}$, and $\boldsymbol{\Sigma}_1^c$ be the covariance matrix calculated with the current values of $\boldsymbol{\eta}$. (Similarly we define the new and current covariance matrices at time $t = 1$ by $\boldsymbol{\Sigma}_1^n$ and $\boldsymbol{\Sigma}_1^c$, respectively.) Then we accept the proposal $\boldsymbol{\eta}^n$ with probability $\min(1, \exp(a - b))$, where

$$a = \log n_m(\boldsymbol{Y}_1; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1^n) + \sum_{t=2}^{N} \log n_m(\boldsymbol{Y}_t; \boldsymbol{\mu}_t + \boldsymbol{\Phi}^n(\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}_{t-1}), \boldsymbol{\Sigma}) + \log n_m(\boldsymbol{\eta}^n; \boldsymbol{\mu}_\eta, \boldsymbol{V}_\eta)$$

and

$$b = \log n_m(\boldsymbol{Y}_1; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1^c) + \sum_{t=2}^{N} \log n_m(\boldsymbol{Y}_t; \boldsymbol{\mu}_t + \boldsymbol{\Phi}^c(\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}_{t-1}), \boldsymbol{\Sigma}) + \log n_m(\boldsymbol{\eta}^c; \boldsymbol{\mu}_\eta, \boldsymbol{V}_\eta).$$

If we do not accept the proposal, we stay at $\boldsymbol{\eta}^c$. (Here $n_m(\boldsymbol{z}; \boldsymbol{m}, \boldsymbol{V})$ denotes the density for a m-variate normal with mean $\boldsymbol{m}$ and covariance $\boldsymbol{V}$, evaluated at $\boldsymbol{z}$.)

In practice we update $\boldsymbol{\eta}$ in blocks of length 9 (one sixth of the total number of spatial locations), tuning the acceptance probability to be around 0.2 over the blocks.

S2.4 Updating τ_ζ^2 and λ_ζ in $\boldsymbol{\Sigma}$:

We have that $\boldsymbol{\Sigma}_1 = \tau_\zeta^2 \boldsymbol{R}_{1,\zeta}(\lambda_\zeta)$ where

$$[\boldsymbol{R}_{1,\zeta}(\lambda_\zeta)]_{j,k} = \frac{\exp(-d(\boldsymbol{s}_j, \boldsymbol{s}_k)/\lambda_\zeta)}{1 - \phi(\boldsymbol{s}_j)\phi(\boldsymbol{s}_k)}, \quad j, k = 1, \ldots, m.$$

and $\boldsymbol{\Sigma} = \tau_\zeta^2 \boldsymbol{R}_\zeta(\lambda_\zeta)$ where

$$[\boldsymbol{R}_\zeta(\lambda_\zeta)]_{j,k} = \exp(-d(\boldsymbol{s}_j, \boldsymbol{s}_k)/\lambda_\zeta)\,, \quad j, k = 1, \ldots, m.$$

Assuming an $\text{InvGamma}(\alpha_{\tau_\zeta^2},\, \beta_{\tau_\zeta^2})$ prior distribution for τ_ζ^2, our update for τ_ζ^2 given all the other parameters is

$$\text{InvGamma}(\alpha_{\tau_\zeta^2} + Nm/2, \beta_{\tau_\zeta^2} + S/2).$$

In the above equation let

$$S \;=\; \boldsymbol{\zeta}_1^T \boldsymbol{R}_{1,\zeta}^{-1}(\lambda_\zeta)\boldsymbol{\zeta}_1 + \sum_{t=2}^{N} \boldsymbol{\zeta}_t^T \boldsymbol{R}_\zeta^{-1}(\lambda_\zeta)\boldsymbol{\zeta}_t.$$

where

$$\boldsymbol{\zeta}_t \;=\; \begin{cases} \boldsymbol{Y}_1 - \boldsymbol{\mu}_1, & t = 1; \\ \boldsymbol{Y}_t - \boldsymbol{\mu}_t - \boldsymbol{\Phi}(\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}_{t-1}), & t = 2, \ldots, N, \end{cases}$$

denotes the vector-valued time series innovations.

We update λ_ζ via Metropolis-Hastings. Assume the current value is λ_ζ^c we propose a new one on the log scale:

$$\log \lambda_\zeta^n \sim N(\log \lambda_\zeta^c, v_{\lambda_\zeta}^2),$$

where $v_{\lambda_\zeta}^2$ is some proposal variance. Let $\boldsymbol{\Sigma}^n = \tau_\zeta^2 \boldsymbol{R}_\zeta(\lambda_\zeta^n)$ and $\boldsymbol{\Sigma}^c = \tau_\zeta^2 \boldsymbol{R}_\zeta(\lambda_\zeta^c)$, with $\boldsymbol{\Sigma}_1^n$ and $\boldsymbol{\Sigma}_1^c$ defined similarly. Then we accept the proposal λ_ζ^n with probability $\min(1, \exp(a - b))$, where

$$a = \log n_m(\boldsymbol{Y}_1; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1^n) + \sum_{t=2}^{N} \log n_m(\boldsymbol{Y}_t; \boldsymbol{\mu}_t + \boldsymbol{\Phi}(\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}_{t-1}), \boldsymbol{\Sigma}^n) + \log \pi(\lambda_\zeta^n) + \log \lambda_\zeta^n$$

and

$$b = \log n_m(\boldsymbol{Y}_1; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1^c) + \sum_{t=2}^{N} \log n_m(\boldsymbol{Y}_t; \boldsymbol{\mu}_t + \boldsymbol{\Phi}(\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}_{t-1}), \boldsymbol{\Sigma}^c) + \log \pi(\lambda_\zeta^c) + \log \lambda_\zeta^c;$$

otherwise we stay at λ_ζ^c. In the above equations $\pi(\lambda_\zeta)$ is the prior density for λ_ζ.

S2.5 Updating $\boldsymbol{\beta}_\eta$, τ_η^2, λ_η:

(This allows for spatial trend in the transformed spatially varying AR(1) parameters through the introduction of a design matrix $\boldsymbol{X}_\eta$.)

We have

$$\boldsymbol{\eta} = (\eta(\boldsymbol{s}_j) : j = 1, \ldots, m)^T \quad \sim \quad N_m(\boldsymbol{\mu}_\eta, \boldsymbol{V}_\eta).$$

where $\boldsymbol{\mu}_\eta = \boldsymbol{X}_\eta \boldsymbol{\beta}_\eta$ and $\boldsymbol{V}_\eta = \tau_\eta^2 \, R(\lambda_\eta)$ with

$$[\boldsymbol{R}(\lambda_\eta)]_{j,k} = \exp(-d(\boldsymbol{s}_j, \boldsymbol{s}_k)/\lambda_\eta)\,, \quad j, k = 1, \ldots, m.$$

We use a multivariate normal update for $\boldsymbol{\beta}_\eta$, inverse gamma update for τ_η^2, and random walk Metropolis-Hastings update for $\log \lambda_\eta$. This is implemented in the R function `update.GSP.pars` in the file `GSP_Bayes.R`.

S2.6 Updating the parameters characterizing $\boldsymbol{\mu}_t$:

We have that

$$\boldsymbol{\mu}_t = \sum_{k=1}^{K} x_{k,t} \, \boldsymbol{\beta}_k,$$

where $x_{k,t}$ is the kth covariate value at time t and

$$\boldsymbol{\beta}_k = (\beta_k(\boldsymbol{s}_j) : j = 1, \ldots, m)^T.$$

Suppose we want to update $\boldsymbol{\beta}_k$ for some $k = 1, \ldots, K$. Then $\boldsymbol{\mu}_t = x_{k,t} \, \boldsymbol{\beta}_k + \boldsymbol{\nu}_{k,t}$ where $\boldsymbol{\nu}_{k,t} = \sum_{l \neq k} x_{l,t} \, \boldsymbol{\beta}_l$.

Independently over t we have

$$\boldsymbol{Y}_1 \sim N(x_{k,1} \, \boldsymbol{\beta}_k + \boldsymbol{\nu}_{k,1}, \boldsymbol{\Sigma}_1),$$

with

$$\boldsymbol{Y}_t | \boldsymbol{Y}_{t-1} \sim N_m(x_{k,t} \, \boldsymbol{\beta}_k + \boldsymbol{\nu}_{k,t} + \boldsymbol{\Phi}(\boldsymbol{Y}_{t-1} - x_{k,t-1} \, \boldsymbol{\beta}_k - \boldsymbol{\nu}_{k,t-1}), \boldsymbol{\Sigma}), \quad t = 2, \ldots, N.$$

Thus

$$\boldsymbol{C}_{k,1} \;=\; \boldsymbol{Y}_1 - \boldsymbol{\nu}_{k,1} \sim N_m(\boldsymbol{D}_{k,1}\,\boldsymbol{\beta}_k, \boldsymbol{\Sigma}_1),$$

which is independent of

$$\boldsymbol{C}_{k,t} \;=\; \boldsymbol{Y}_t - \boldsymbol{\nu}_{k,t} - \boldsymbol{\Phi}[\boldsymbol{Y}_{t-1} - \boldsymbol{\nu}_{k,t-1}] \;\sim\; N_m(\boldsymbol{D}_{k,t}\,\boldsymbol{\beta}_k, \boldsymbol{\Sigma}),$$

over $t = 2, \ldots, N$. In the above equations

$$\boldsymbol{D}_{k,t} \;=\; \begin{cases} x_{k,1}\,\boldsymbol{I}_m, & t = 1; \\ x_{k,t}\,\boldsymbol{I}_m - x_{k,t-1}\,\boldsymbol{\Phi}, & t = 2, \ldots, N. \end{cases}$$

Thus our update for $\boldsymbol{\beta}_k$ given all the other parameters is $N_m(\boldsymbol{P}^{-1}\boldsymbol{q}, \boldsymbol{P}^{-1})$ where

$$\begin{aligned} \boldsymbol{P} \;&=\; \boldsymbol{D}_{k,1}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{D}_{k,1} + \sum_{t=2}^{N}\boldsymbol{D}_{k,t}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{D}_{k,t} + \boldsymbol{V}_{\beta_k}^{-1}, \text{ and} \\ \boldsymbol{q} \;&=\; \boldsymbol{D}_{k,1}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{C}_{k,1} + \sum_{t=2}^{N}\boldsymbol{D}_{k,t}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{C}_{k,t} + \boldsymbol{V}_{\beta_k}^{-1}\boldsymbol{\mu}_{\beta_k}. \end{aligned}$$

S2.7 Updating the hyperparameters for $\boldsymbol{\beta}_k$, $k = 1, \ldots, K$:

(This allows for spatial trend in the spatially varying temporal trend parameters through the introduction of a design matrix $\boldsymbol{X}_{\beta_k}$.)

Independently over $k = 1, \ldots, K$ we have

$$\boldsymbol{\beta}_k = (\beta_k(\boldsymbol{s}_j) : j = 1, \ldots, m)^T \;\sim\; N_m(\boldsymbol{\mu}_{\beta_k}, \boldsymbol{V}_{\beta_k}),$$

where $\boldsymbol{\mu}_{\beta_k} = \boldsymbol{X}_{\beta_k}\boldsymbol{\beta}_{\beta_k}$ and $\boldsymbol{V}_{\beta_k} = \tau_{\beta_k}^2\,\boldsymbol{R}(\lambda_{\beta_k})$ with

$$[\boldsymbol{R}(\lambda_{\beta_k})]_{j,l} = \exp(-d(\boldsymbol{s}_j, \boldsymbol{s}_l)/\lambda_{\beta_k}), \quad j, l = 1, \ldots, m.$$

For each k, we use a multivariate normal update for $\boldsymbol{\beta}_{\beta_k}$, inverse gamma update for $\tau_{\beta_k}^2$, and random walk Metropolis-Hastings update for $\log \lambda_{\beta_k}$. This is implemented in the R function `update.GSP.pars` in the file `GSP_Bayes.R`.

S3 Analyzing differences of seasonally standardized values

We now investigate whether there is evidence of spatial trends over time when comparing the monthly SST and MAT values, after seasonal standardization. By seasonal standardization we mean that at each spatial location, after removing the longest possible monthly temporal average, we then divide by the longest possible monthly temporal standard deviation. Figure S3 compares the monthly temporal standard deviations for the TAO buoy SST, TAO buoy MAT, ERA5 reanalysis product SST, and reanalysis product MAT. We can see that the standard deviations by month are more different near the equator, and more similar further away from the equator.

In Figure S4 we display heat maps of the seasonally standardized SST, seasonally standardized MAT, and their difference (seasonally standardized SST minus seasonally standardized MAT) for the TAO buoys in the left panels and the ERA5 reanalysis product in the right panels. There is evidence of strong seasonal patterns in the seasonally standardized SST and seasonally standardized MAT for both the TAO buoy data and ERA5 reanalysis product across the study region. There is little evidence of seasonal patterns in the difference of the seasonally standardized values for the TAO buoys, but some evidence of seasonality in the differences for the ERA5 reanalysis product. Figure S5 compares the differences of the seasonally standardized values for the TAO buoys and ERA5 reanalysis product, and indicates evidence of increased variability for both series in the west and north of the study region in the tropical Pacific. There is also a suggestion that the distribution of the differences of the seasonally standardized TAO buoys measurements and ERA5 reanalysis values may be different in the west.

We fit our hierarchical spatio-temporal Bayesian model to the monthly seasonally standardized differences for the TAO buoys and our spatio-temporal Bayesian model to the monthly seasonally standardized differences for ERA5. For both models we do not include a seasonal terms that vary over spatial locations, as we did for the models fit to the seasonally adjusted

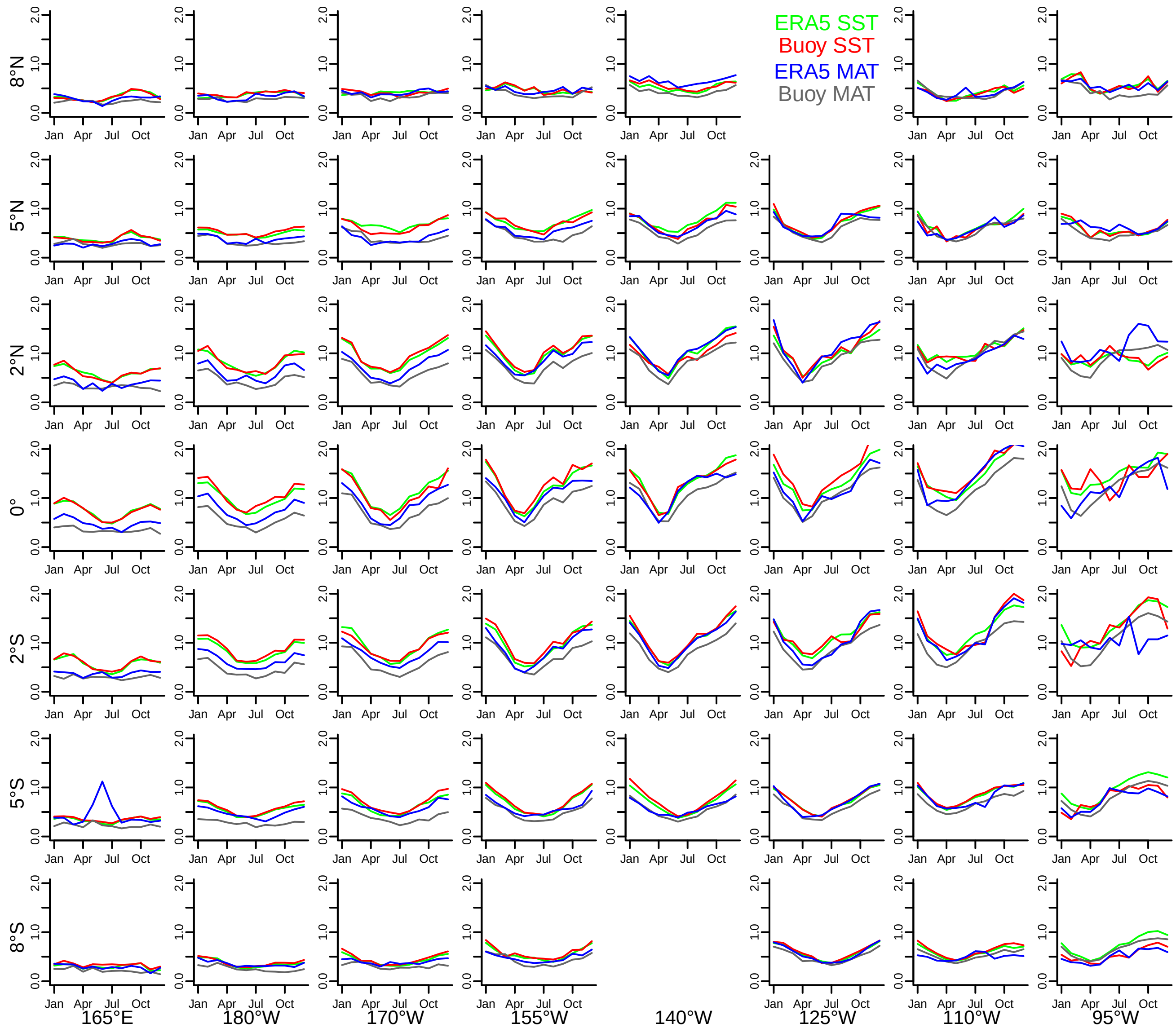

Figure S3: Monthly standard deviations for Buoy SST (red), Buoy MAT (gray), ERA5 SST (green), ERA5 MAT (blue) for the $0.25° \times 0.25°$ grid square containing the corresponding buoy location. This pictorial representation does not accurately represent the distances between different buoy locations, and the series presented at 8°N and 140°W is actually at at location 9°N and 140°W.

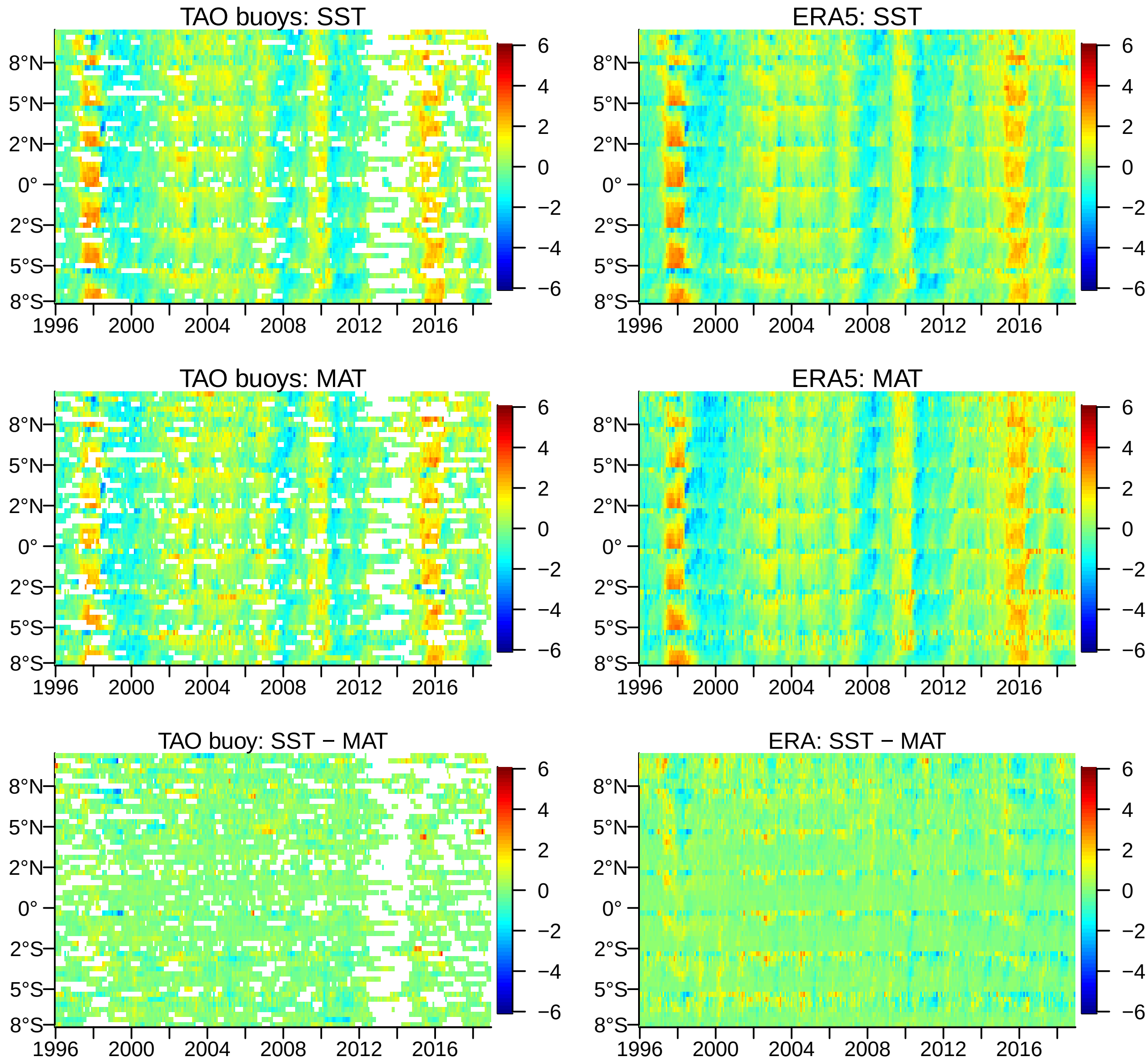

Figure S4: Heat maps of the seasonally standardized monthly SST, MAT, and their difference for the TAO buoy data (left panels) and ERA5 reanalysis product (right panels). In each panel the monthly time series by year (horizontal axis) are ordered from north to south, and within each latitude from west to east (vertical axis)

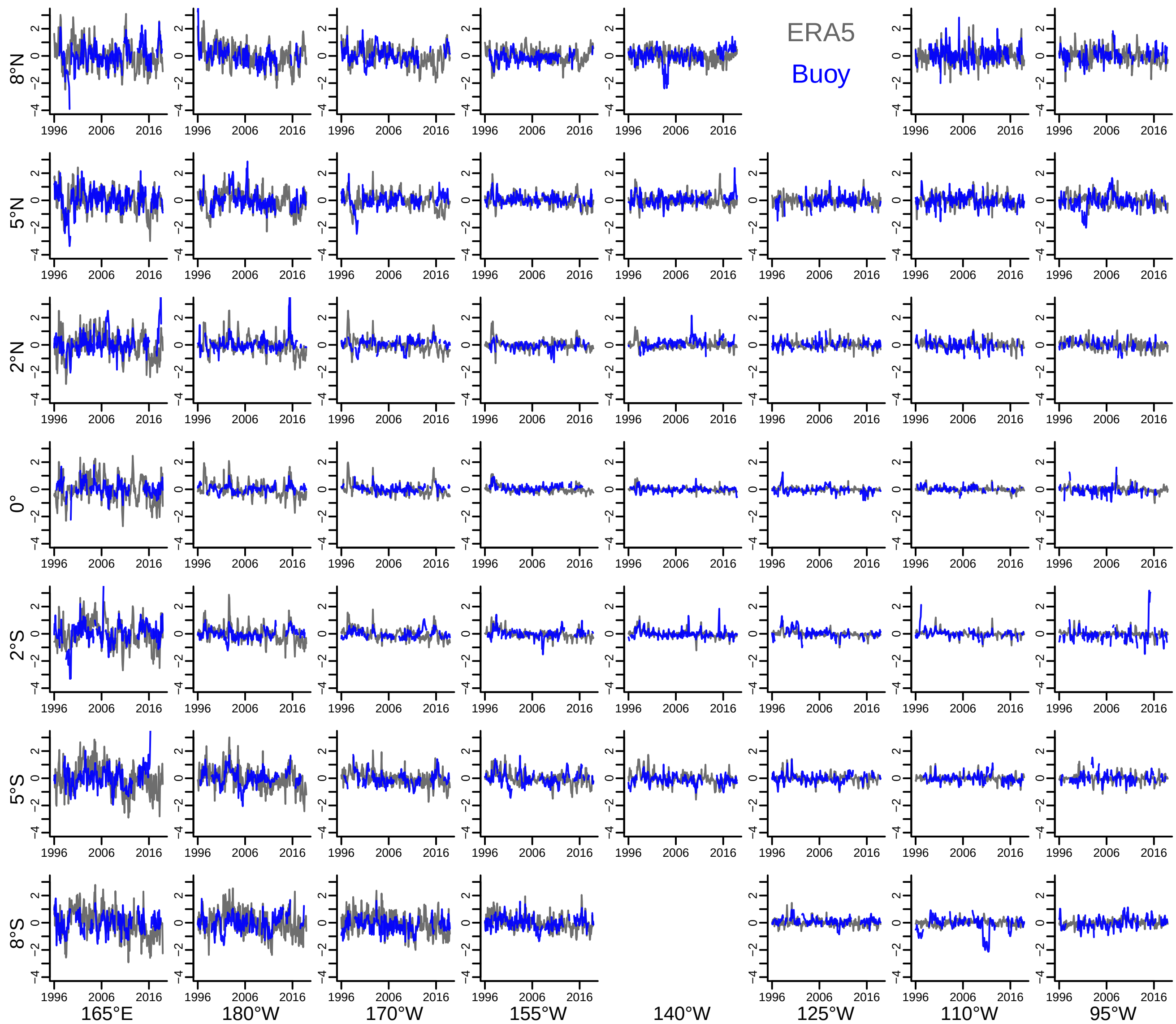

Figure S5: Buoy (blue) and ERA5 (gray) seasonally standardized monthly differences for the 0.25° × 0.25° grid square containing the corresponding buoy location. This pictorial representation does not accurately represent the distances between different buoy locations, and the series presented at 8°N and 140°W is actually at at location 9°N and 140°W.

differences. Again we still include a constant and linear trend term at each location, that can vary spatially over the tropical Pacific.

Column (a) of Table S2 present summaries of the posterior distributions of hyperparameters in the model fit to the monthly seasonally standardized differences for the buoys, while column (b) of the same table summarizes the posteriors for the model fit to the monthly seasonally standardized differences for ERA5. The key finding is that there is evidence that while the intercept and slope terms (as measured by μ_{β_1} and μ_{β_2}, respectively) are not different from zero for the TAO buoys, there is evidence of a decreasing trend over time when comparing the seasonally standardized SST with the seasonally standardized MAT for the ERA5 reanalysis product: the posterior mean for μ_{β_2} is estimated to decrease by 0.018 units per year, with a 95% credible interval of $(0.011, 0.046)$ for this decrease.

Figure S6 displays posterior summaries of the spatially varying parameters in each of the two hierarchical spatio-temporal Bayesian models fit to the seasonally standardized monthly differences. This plot is arranged similarly to Figures 6 and 7 in the main article. Once we seasonally standardize we learn that there is no evidence of linear trends over the tropical Pacific comparing the seasonally standardized SST to the seasonally standardized MAT for the TAO buoys, but there is evidence of significant negative linear trends in the west of the tropical Pacific when we compare the seasonally standardized SST to the seasonally standardized MAT for the ERA5 reanalysis product. We should caution that while we find differences between the standardized values over time for the ERA5 reanalysis product, it is harder to interpret the implication of these results on the original Celsius scale. For that we reason, we prefer to interpret the results presented in the main article.

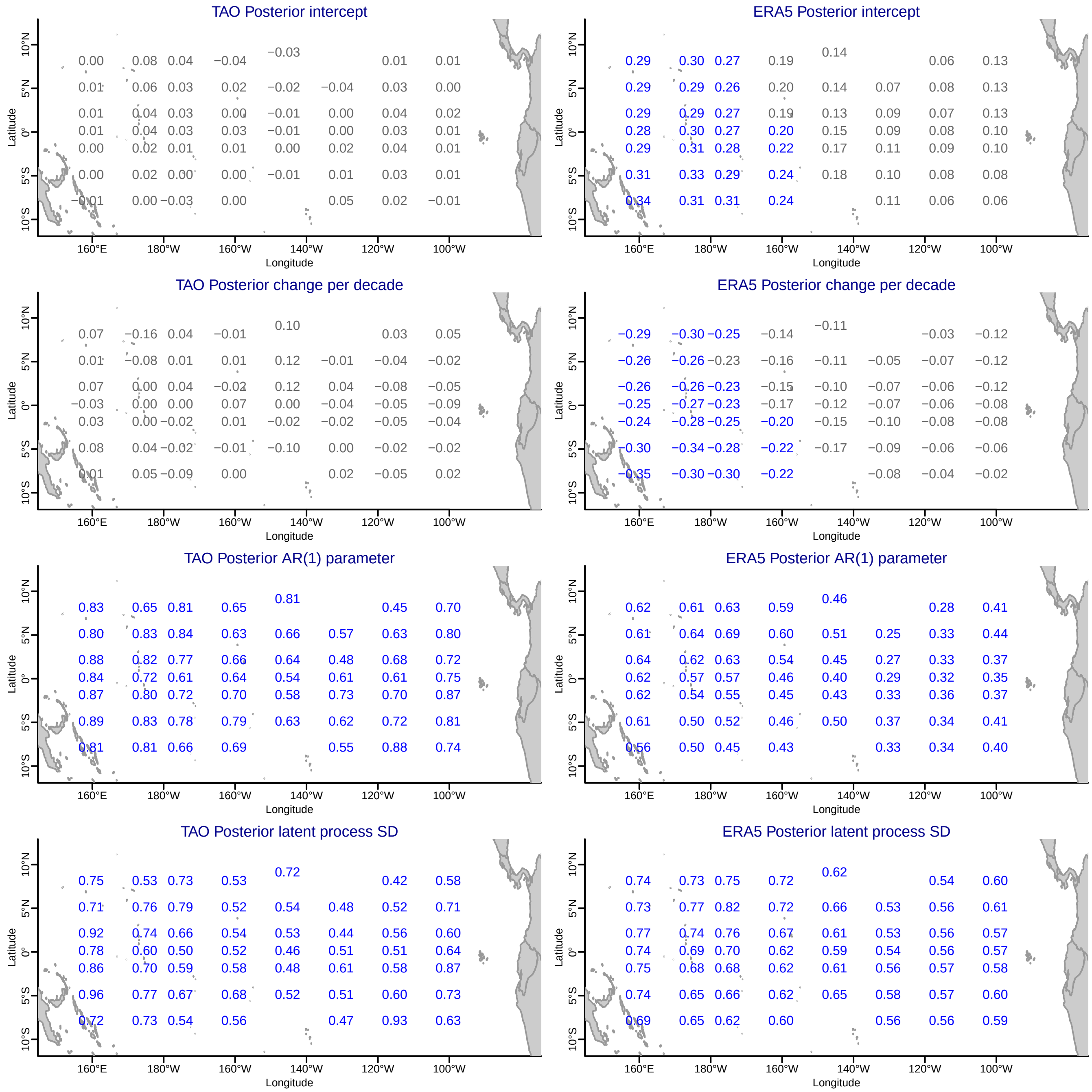

Figure S6: A summary of spatially varying posteriors from the hierarchical spatio-temporal Bayesian models fit to the seasonally standardized monthly differences, SST-MAT, for TAO buoys (first column) and ERA5 (second column). The values are the posterior means for the spatially varying intercepts (first row), spatially varying decadal trend (second row), the spatially varying AR(1) parameters (third row), and spatially varying process standard deviations (fourth row). Gray values indicates that simultaneous 95% credible intervals for the parameter at that location includes zero, whereas blue values indicate that the interval does not contain zero.

Table S2: Posterior means with 95% credible intervals in parentheses for hierarchical spatio-temporal Bayesian models fit to four different datasets. Each column shows the results for each dataset: (a) the seasonally standardized monthly differences, SST-MAT, for TAO buoys; (b) the seasonally standardized monthly differences, SST-MAT, for ERA5.

	Seasonally standardized differences	
Parameter	(a) TAO buoy	(b) ERA5
σ	0.214 (0.201, 0.225)	0.160 (0.130, 0.185)
μ_η	1.890 (1.137, 2.659)	1.031 (0.704, 1.365)
τ_η	0.817 (0.565, 1.145)	0.316 (0.215, 0.448)
λ_η	1.768 (1.029, 2.699)	2.189 (1.432, 3.133)
τ_ζ	0.312 (0.298, 0.325)	0.456 (0.438, 0.476)
λ_ζ	0.574 (0.480, 0.690)	0.804 (0.686, 0.946)
μ_{β_1}	0.002 (-0.116, 0.114)	0.191 (0.080, 0.308)
τ_{β_1}	0.092 (0.051, 0.155)	0.089 (0.056, 0.132)
λ_{β_1}	2.062 (1.279, 3.043)	2.474 (1.644, 3.477)
μ_{β_2}	0.001 (-0.030, 0.034)	-0.018 (-0.046, 0.011)
τ_{β_2}	0.027 (0.022, 0.034)	0.023 (0.019, 0.028)
λ_{β_2}	3.198 (2.213, 4.367)	3.830 (2.737, 5.114)