Sea Surface Temperature Estimation

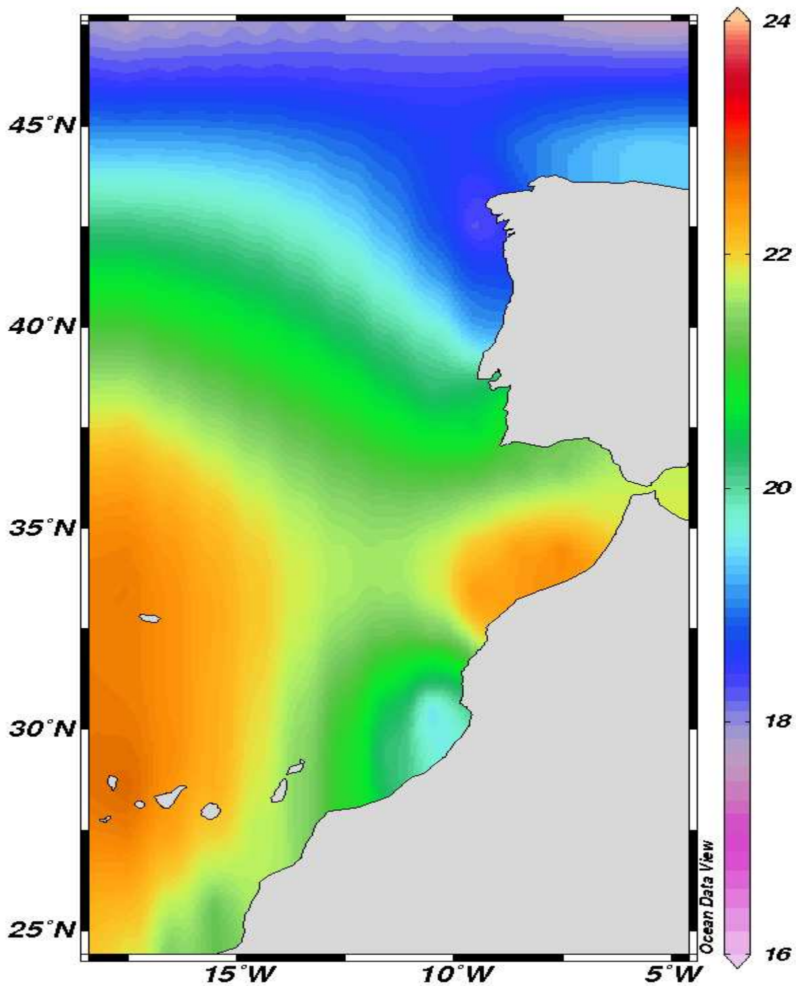
Bruno Sansó

Department of Applied Mathematics and Statistics
University of California Santa Cruz
http://www.ams.ucsc.edu/∼bruno

In collaboration with Ricardo Lemos (U. Queensland), and Celeste Tretto (UCSC). We acknowledge the advise of John Kennedy (Met Office, UK).

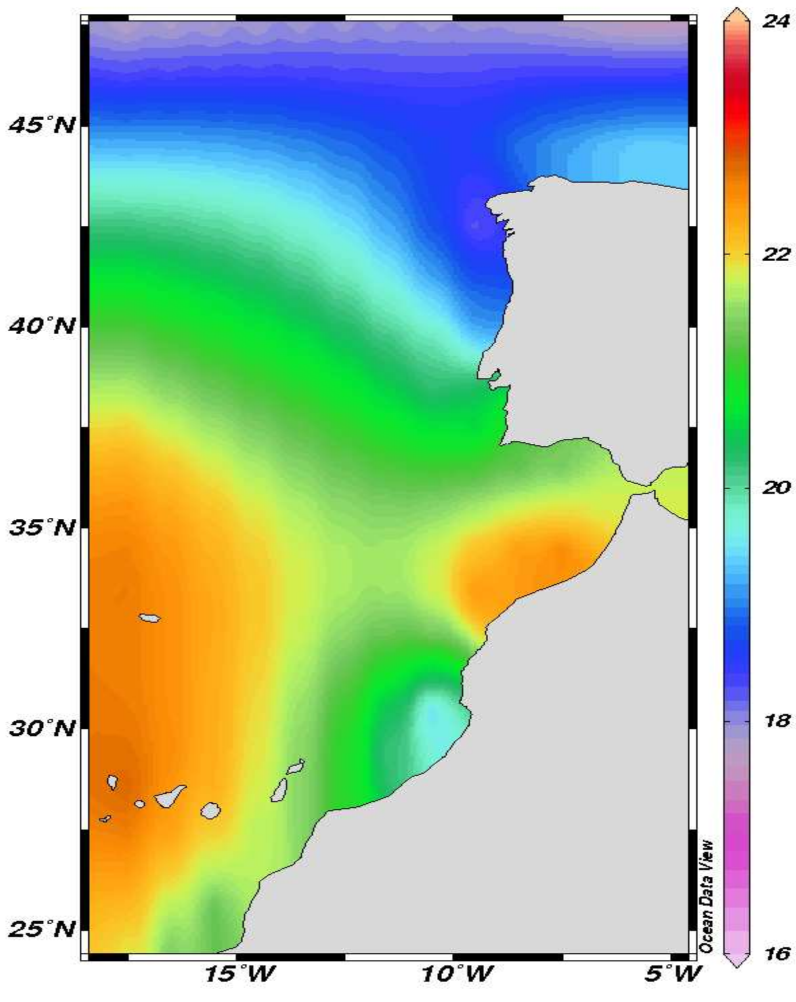
Motivating Issue

August average SST from WOA98



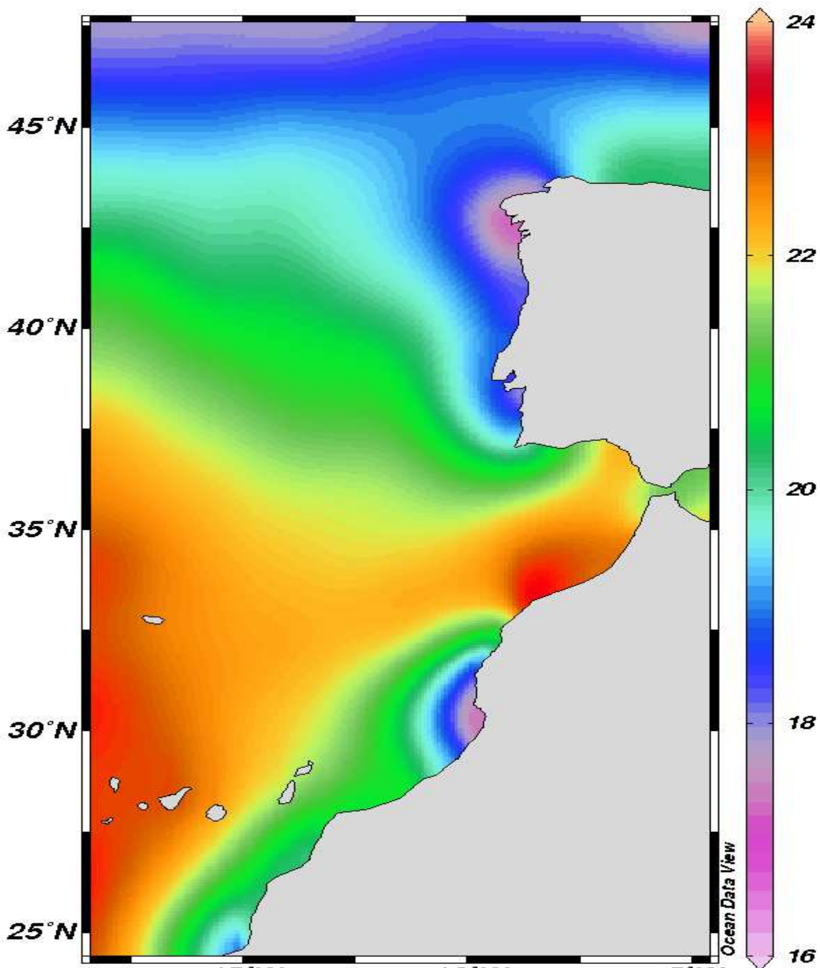
- The '98 WOA climatology has a resolution of 1°. It is very smooth with a number of circular features. These are probably due to a large spherical kernel.

August average SST from WOA98



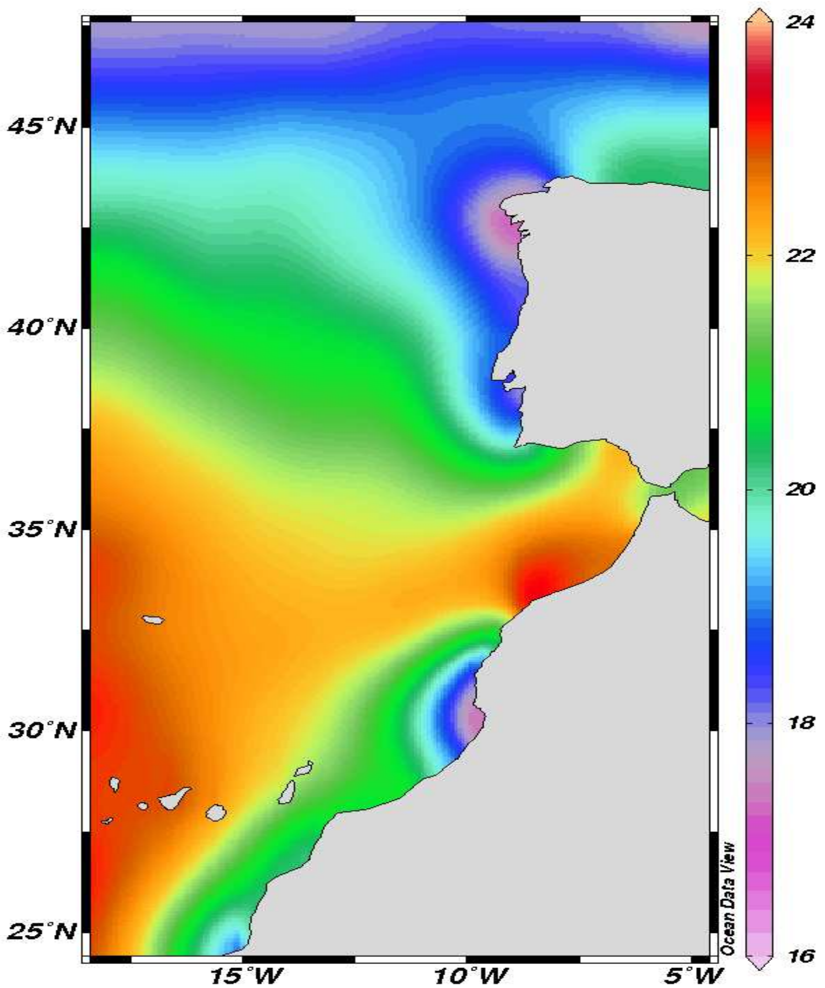
- The '98 WOA climatology has a resolution of 1°. It is very smooth with a number of circular features. These are probably due to a large spherical kernel.
- There is no upwelling off the Iberian peninsula and a very weak one in Western Africa.

August average SST from WOA01



- The WOA01v2 climatology has 1/4°resolution. It has been available since 2006 and it includes data sources that entered the NODC after 1998.

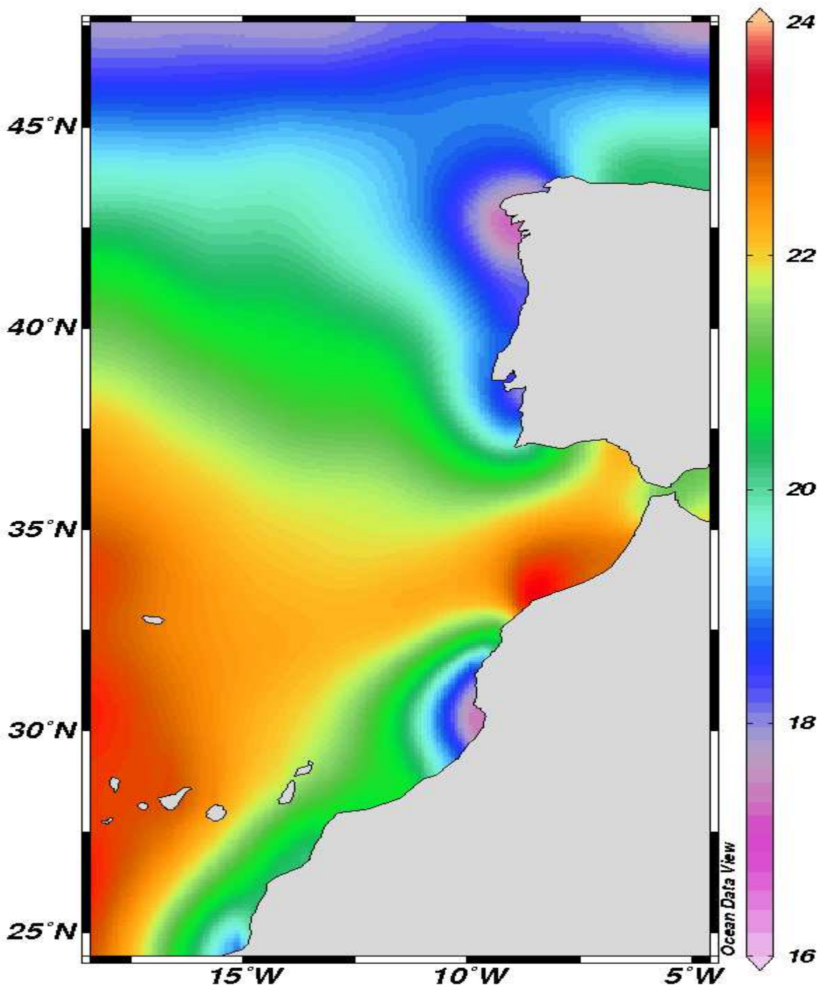
August average SST from WOA01



- The WOA01v2 climatology has 1/4°resolution. It has been available since 2006 and it includes data sources that entered the NODC after 1998.
- Circular kernels have a visible effect, e.g. in NW Iberia.

World Ocean Atlas

August average SST from WOA01



- The WOA01v2 climatology has 1/4°resolution. It has been available since 2006 and it includes data sources that entered the NODC after 1998.
- Circular kernels have a visible effect, e.g. in NW Iberia.
- Not considering the year of observation for computations mixes cold decades with warm ones, thereby producing that strong front in NW Africa (32.5°N)

- The current version of the World Ocean Atlas is WOA13, and has 1/4°resolution. It uses only in situ data.

Worl Ocean Atlas

- The current version of the World Ocean Atlas is WOA13, and has $1/4°$resolution. It uses only in situ data.
- The WOA13 provides monthly climatologies for the period 1955–2012, not only for SST, but also for 120 depths down to 5,500 meters. It also includes salinity and four other variables.

Worl Ocean Atlas

- The current version of the World Ocean Atlas is WOA13, and has 1/4°resolution. It uses only in situ data.
- The WOA13 provides monthly climatologies for the period 1955–2012, not only for SST, but also for 120 depths down to 5,500 meters. It also includes salinity and four other variables.
- It is obtained by smoothing gridded averages using Optimal Interpolation (Kaplan et al. 1997). This is based on assuming Gaussian errors, projecting the state variable on the basis of EOFs and finding the MAP induced by a Gaussian prior.

OSTIA

The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) produces a daily analysis of SST blending in situ and satellite observations (L4).

OSTIA

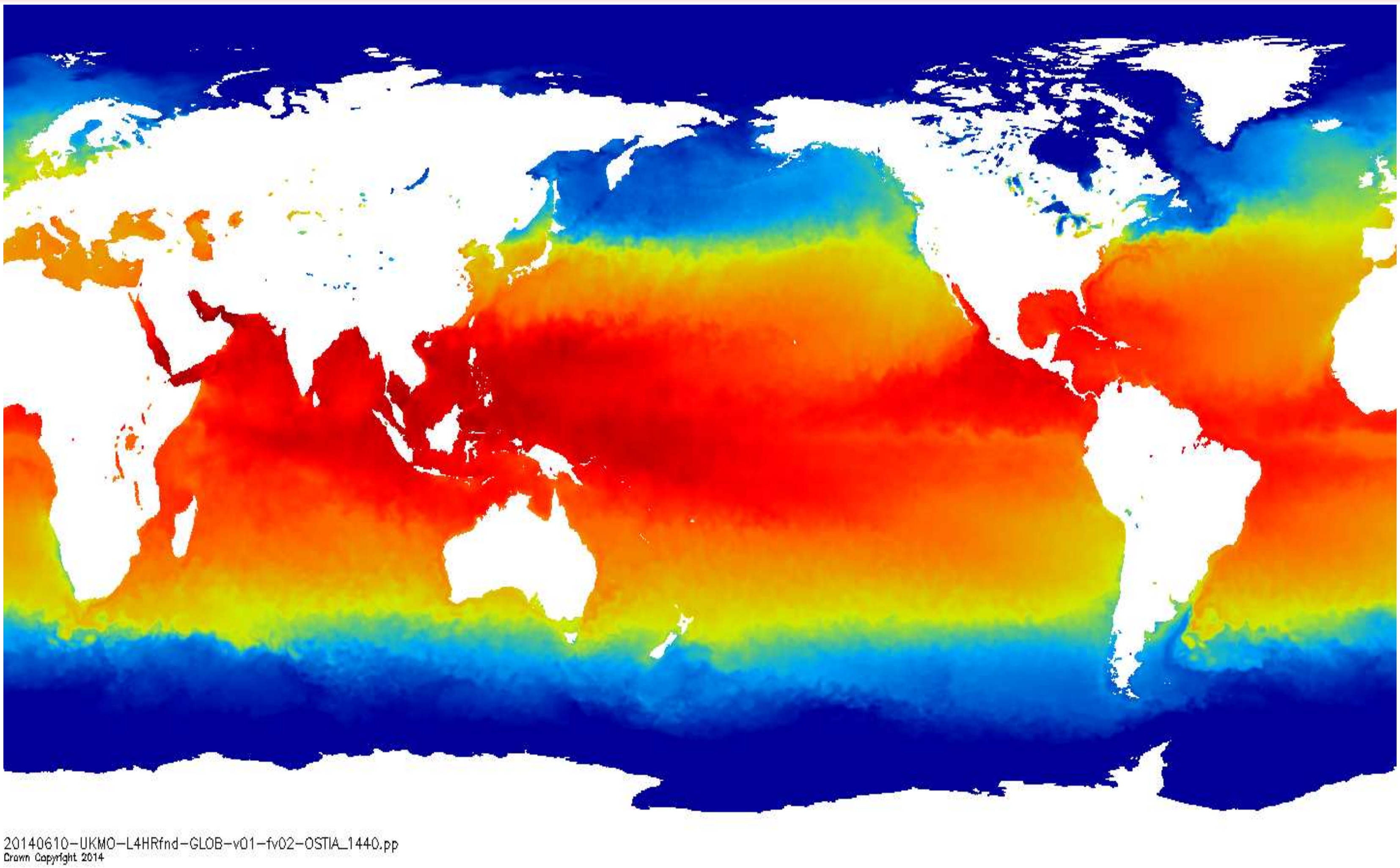
The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) produces a daily analysis of SST blending in situ and satellite observations (L4).

The analysis $\boldsymbol{x}_k$ consists, essentially, on the estimation of the posterior mode of a multivariate posterior, assuming Gaussian errors and using the estimate $\boldsymbol{x}_{k-1}$ from the previous day, as a prior.

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} + \boldsymbol{A}(y_k - h(\boldsymbol{x}_{k-1}))$$

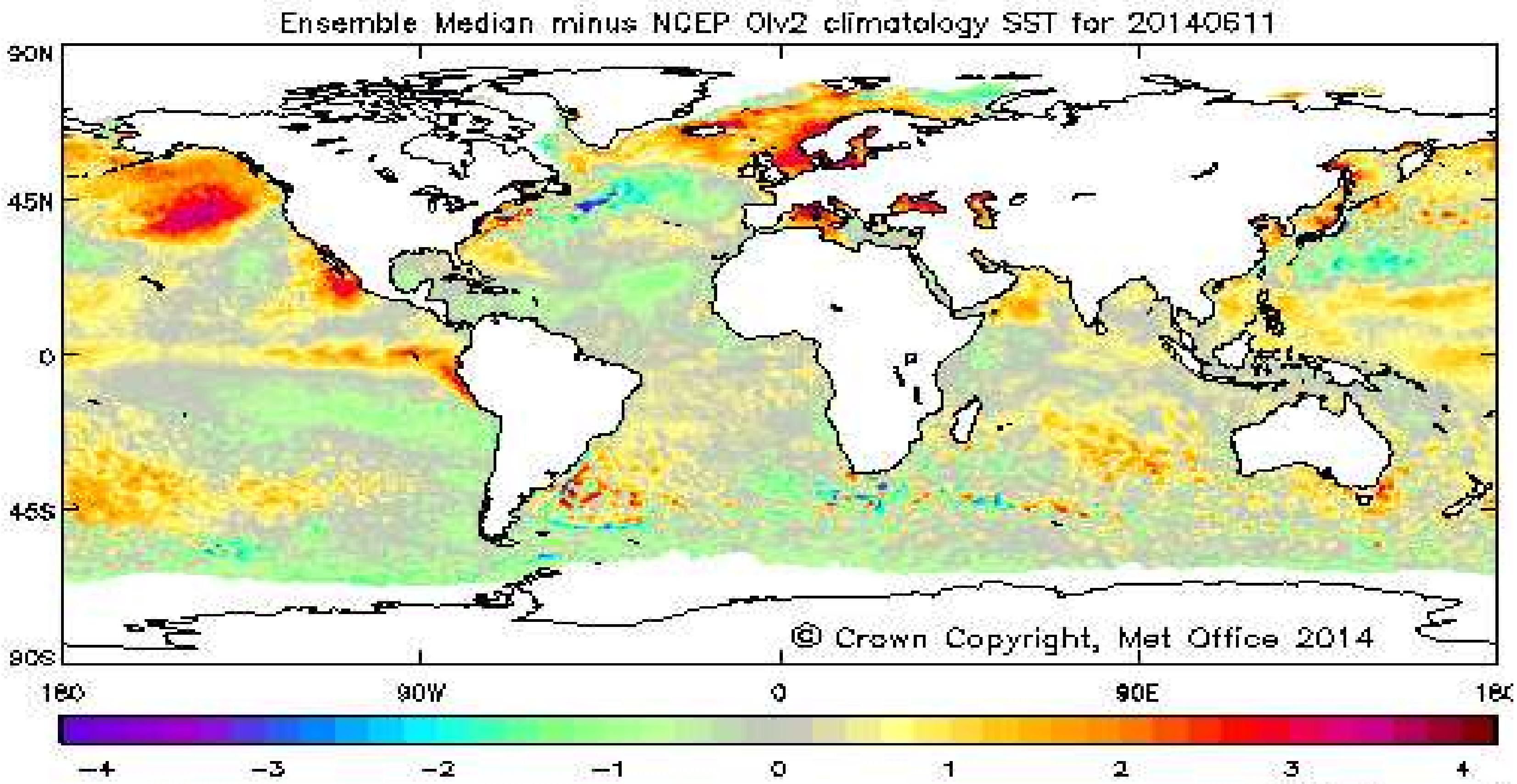
The function h maps the observations to the grid. The matrix $\boldsymbol{A}$ is obtained from such mapping, the prior covariance matrix and the observational variance.

OSTIA



SST for June 10, 2014

GHRSST



OSTIA is one of the analyses of the Group for High Resolution SST. Here we show the anomalies of the median of the 11 member ensemble, WRT the climatology for June 11, 2014.

Why a Space-Time Model?

Why use a sophisticated spatio-temporal model for climatologies when an average is all that is needed?

Why a Space-Time Model?

Why use a sophisticated spatio-temporal model for climatologies when an average is all that is needed?

• Data distribution is uneven in space and time. A model can borrow strength from data-dense periods and locations.

Why a Space-Time Model?

Why use a sophisticated spatio-temporal model for climatologies when an average is all that is needed?

• Data distribution is uneven in space and time. A model can borrow strength from data-dense periods and locations.

• To account for location- and time-varying seasonal cycles, long term trends and high frequency variability.

Why a Space-Time Model?

Why use a sophisticated spatio-temporal model for climatologies when an average is all that is needed?

• Data distribution is uneven in space and time. A model can borrow strength from data-dense periods and locations.

• To account for location- and time-varying seasonal cycles, long term trends and high frequency variability.

• To account for observational error.

Why a Space-Time Model?

Why use a sophisticated spatio-temporal model for climatologies when an average is all that is needed?

- Data distribution is uneven in space and time. A model can borrow strength from data-dense periods and locations.

- To account for location- and time-varying seasonal cycles, long term trends and high frequency variability.

- To account for observational error.

- To incorporate different sources information, including structural knowledge.

Why a Space-Time Model?

Why use a sophisticated spatio-temporal model for climatologies when an average is all that is needed?

• Data distribution is uneven in space and time. A model can borrow strength from data-dense periods and locations.

• To account for location- and time-varying seasonal cycles, long term trends and high frequency variability.

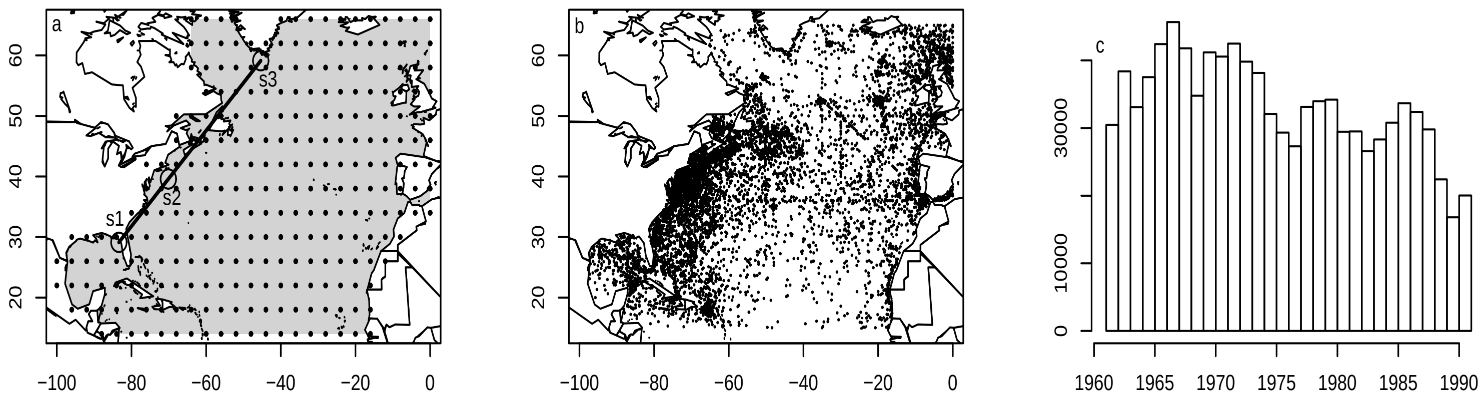
• To account for observational error.

• To incorporate different sources information, including structural knowledge.

• To produce probabilistic measures of uncertainty.

North Atlantic Climatology: Domain

The domain S (gray area) and the grid J (bullets; $r_J = 4°$) used for the convolving process. A transect with three "case study" points. A random sample of 1% of the data. Temporal distribution of the data.



Observation Equation

We use data from the NODC World Ocean Database 2005, collected with four types of instruments between 1961 and 1990.

We use data from the NODC World Ocean Database 2005, collected with four types of instruments between 1961 and 1990.

The SST observation $x_{i,m,y}(\boldsymbol{s})$ collected with instrument $i = 1, \ldots, 4$, in month m, year y and location $\boldsymbol{s}$ follows

$$x_{i,m,y}(\boldsymbol{s}) \sim N\left(\theta_{m,y}(\boldsymbol{s}), \tau_i^2\right).$$

Represent a Gaussian process X as

$$\theta(\boldsymbol{s}) = \sum_{j=1}^{m} B_j(\boldsymbol{s})\gamma_j = \boldsymbol{B}(\boldsymbol{s})^T\boldsymbol{\gamma}\ , \quad \boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_m)^T \sim N_m(0, \boldsymbol{K})$$

Low Rank Models

Represent a Gaussian process X as

$$\theta(\boldsymbol{s}) = \sum_{j=1}^{m} B_j(\boldsymbol{s})\gamma_j = \boldsymbol{B}(\boldsymbol{s})^T \boldsymbol{\gamma}\ , \quad \boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_m)^T \sim N_m(0, \boldsymbol{K})$$

- Predictive Gaussian processes (Banerjee et al. '08) that take γ as a Gaussian process over a grid. $B(s)$ is obtained by calculating the predictive expectation for points outside the grid.

Low Rank Models

Represent a Gaussian process X as

$$\theta(\boldsymbol{s}) = \sum_{j=1}^{m} B_j(\boldsymbol{s})\gamma_j = \boldsymbol{B}(\boldsymbol{s})^T\boldsymbol{\gamma}\ , \quad \boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_m)^T \sim N_m(0, \boldsymbol{K})$$

• Predictive Gaussian processes (Banerjee et al. '08) that take γ as a Gaussian process over a grid. $B(s)$ is obtained by calculating the predictive expectation for points outside the grid.

• Reduced rank kriging (Cressie and Johanesson '08) where $B(s)$ corresponds to a multi-resolution, non-orthogonal, basis.

Low Rank Models

Represent a Gaussian process X as

$$\theta(\boldsymbol{s}) = \sum_{j=1}^{m} B_j(\boldsymbol{s})\gamma_j = \boldsymbol{B}(\boldsymbol{s})^T\boldsymbol{\gamma} \; , \quad \boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_m)^T \sim N_m(0, \boldsymbol{K})$$

- Predictive Gaussian processes (Banerjee et al. '08) that take γ as a Gaussian process over a grid. $B(s)$ is obtained by calculating the predictive expectation for points outside the grid.
- Reduced rank kriging (Cressie and Johanesson '08) where $B(s)$ corresponds to a multi-resolution, non-orthogonal, basis.
- Process convolutions (Higdon '07) where $\gamma_j = \gamma(s_j^*)$ is a process over a grid $s_1^*, \dots, s_J^*$ and $B_j(s) = b(s - s_j^*)$.

To model the SSTs we use process convolutions:

$$\theta_{m,y}(\boldsymbol{s}) \sim N\left(\sum_{\boldsymbol{j}} b(\boldsymbol{s}-\boldsymbol{j};\boldsymbol{\Lambda}(\boldsymbol{s}))\left(\alpha(\boldsymbol{j})+\boldsymbol{\beta}_t(\boldsymbol{j})\boldsymbol{w}_t^T+\right.\right.$$
$$\left.\left.\eta(\boldsymbol{j})(t-180)\right), \Phi(\boldsymbol{s})^2\right),$$

$$\Phi(\boldsymbol{s})^2=\sum_{\boldsymbol{j}} b(\boldsymbol{s}-\boldsymbol{j};\boldsymbol{\Omega}(\boldsymbol{s}))\exp\left(\sigma(\boldsymbol{j})\right).$$

Here $t = m + 12(y - 1961)$ denotes time in months since December 1960.

Spatially-Varying Kernels

The space-varying kernels are

$$b(\boldsymbol{s}-\boldsymbol{j};\boldsymbol{\omega}) \equiv \begin{cases} \left(1-||\boldsymbol{s}-\boldsymbol{j}||_{\boldsymbol{\Sigma}}^{2}\right)^{\omega_1} & \text{if } ||\boldsymbol{s}-\boldsymbol{j}||_{\boldsymbol{\Sigma}} < 1 \\ 0 & \text{otherwise.} \end{cases}$$

and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_4)$.

Spatially-Varying Kernels

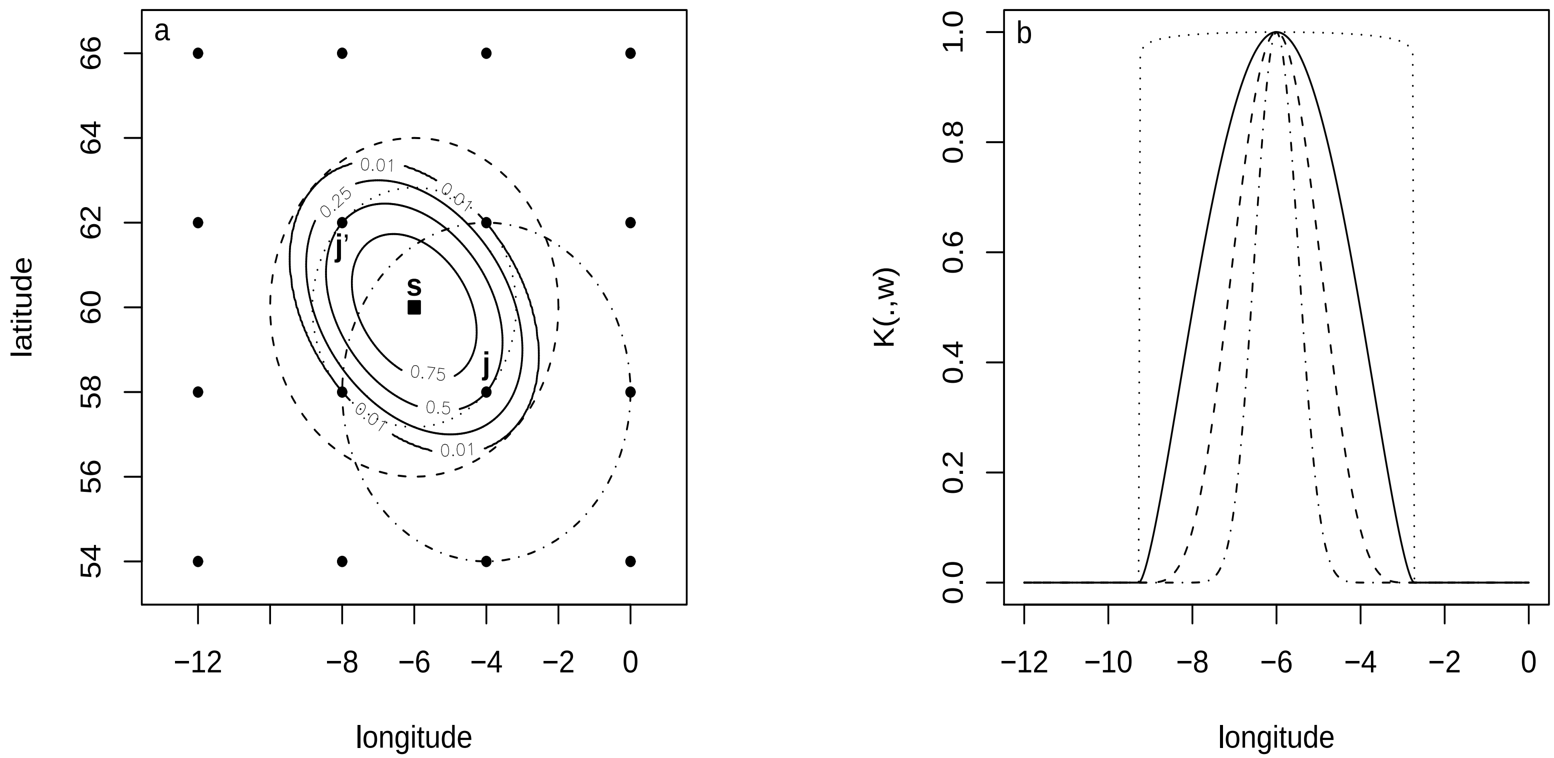
The space-varying kernels are

$$b(\boldsymbol{s}-\boldsymbol{j};\boldsymbol{\omega}) \equiv \begin{cases} \left(1-||\boldsymbol{s}-\boldsymbol{j}||^2_{\boldsymbol{\Sigma}}\right)^{\omega_1} & \text{if } ||\boldsymbol{s}-\boldsymbol{j}||_{\boldsymbol{\Sigma}} < 1 \\ 0 & \text{otherwise.} \end{cases}$$

and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_4)$.

The distance is given as

$$||\boldsymbol{s}-\boldsymbol{j}||_{\boldsymbol{\Sigma}} \equiv \sqrt{((x_s - x_j),(y_s - y_j))\,\boldsymbol{\Sigma}^{-1}\,((x_s - x_j),(y_s - y_j))^T}.$$


a
latitude
longitude
s
j
j'
0.01
0.25
0.5
0.75
b
K(.,w)
longitude

Spatially-Varying Kernels

The ellipsoidal shape is controlled by the parameters in

$$\mathbf{\Sigma}^{-1} \equiv \begin{pmatrix} \Psi_1 + \Psi_2 \cos 2\pi\omega_4 & \Psi_2 \sin 2\pi\omega_4 \\ \Psi_2 \sin 2\pi\omega_4 & \Psi_1 - \Psi_2 \cos 2\pi\omega_4 \end{pmatrix}$$

$$\mathbf{\Psi} = \frac{1}{2}\left(\frac{1}{a^2} + \frac{1}{A^2}, \frac{1}{a^2} - \frac{1}{A^2}\right)$$

$$a = L + \omega_2(U - L), \quad A = a + \omega_3(U - a), \;\; \omega_2, \omega_3 \in (0, 1)$$

So the semi-minor and semi-major axes a and A belong to (L, U).

Anisotropy of the kernels is achieved by letting $\boldsymbol{\Lambda}(\boldsymbol{s}) = (\Lambda_1(\boldsymbol{s}), \ldots, \Lambda_4(\boldsymbol{s}))$ and $\boldsymbol{\Omega}(\boldsymbol{s}) = (\Omega_1(\boldsymbol{s}), \ldots, \Omega_4(\boldsymbol{s}))$, with

$$\boldsymbol{\Lambda}(\boldsymbol{s}) = \sum_{\boldsymbol{j}} b(\boldsymbol{s} - \boldsymbol{j}, \boldsymbol{u}) \boldsymbol{\kappa}(\boldsymbol{j}),$$

$$\boldsymbol{\Omega}(\boldsymbol{s}) = \sum_{\boldsymbol{j}} b(\boldsymbol{s} - \boldsymbol{j}, \boldsymbol{u}) \boldsymbol{\rho}(\boldsymbol{j}),$$

with $\boldsymbol{u} = (2, r_j, r_j, 0)$. Which implies that the kernels are spherical. κ_1 and κ_4 have Uniform priors in $(1.5, 5)$ and $(-\pi/2, \pi/2)$, respectively; the joint prior for κ_2 and κ_3 is proportional to $I_{r_J/\sqrt{2} < \kappa_3 \leq \kappa_2 < r_J}$. We assign analogous priors to $\boldsymbol{\rho}$.

Lack of stationarity in time is handled by letting

$$\boldsymbol{\beta}_t \sim N\left(\boldsymbol{\beta}_{t-1}, \boldsymbol{W}_t\right)$$

and

$$\boldsymbol{w}_t = \left(\sin\left(\frac{2\pi t}{12}\right), \cos\left(\frac{2\pi t}{12}\right), \sin\left(\frac{2\pi t}{6}\right), \cos\left(\frac{2\pi t}{6}\right)\right).$$

Where $\boldsymbol{W}_t$ is modeled using a space-dependent discount factor.

Baseline, α, long-term trend, η, initial seasonality, $\boldsymbol{\beta}_1$, and variance, σ, coefficients are modeled with a MRF as in

$$\alpha(\boldsymbol{j}) \sim N\left(\frac{\alpha(\mathrm{N}(\boldsymbol{j})) + \alpha(\mathrm{S}(\boldsymbol{j})) + \alpha(\mathrm{E}(\boldsymbol{j})) + \alpha(\mathrm{W}(\boldsymbol{j}))}{4}, \tau_\alpha^2\right).$$

The use of a compactly supported kernel allows for an efficient parallel implementation. We use 13 processors, each one working with two columns of J.

Implementation

The use of a compactly supported kernel allows for an efficient parallel implementation. We use 13 processors, each one working with two columns of J.

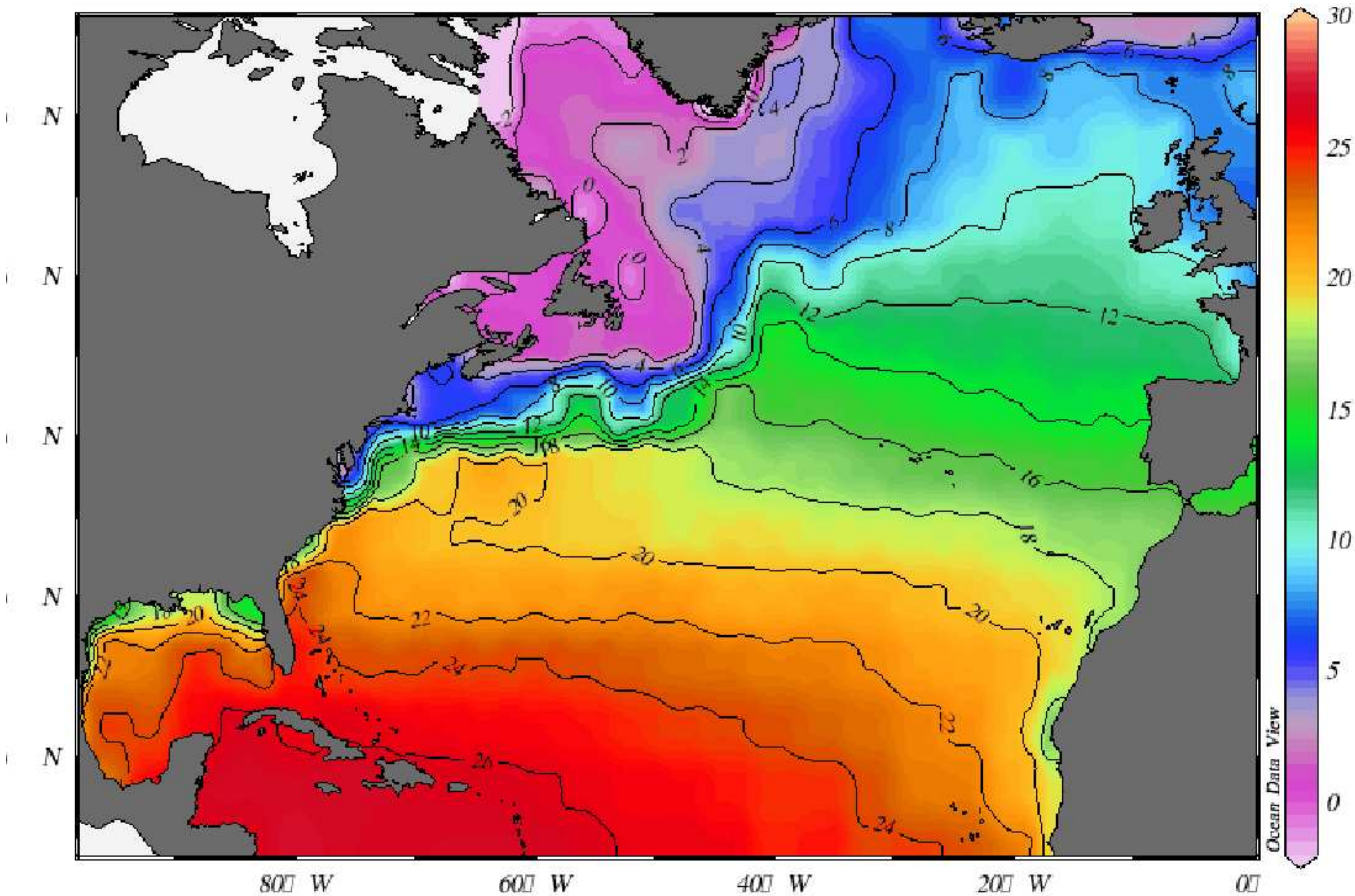
We use reasonably vague inverse gamma priors for all variance parameters. Posterior inference shows that the data do provide information about those parameters.

The use of a compactly supported kernel allows for an efficient parallel implementation. We use 13 processors, each one working with two columns of J.

We use reasonably vague inverse gamma priors for all variance parameters. Posterior inference shows that the data do provide information about those parameters.

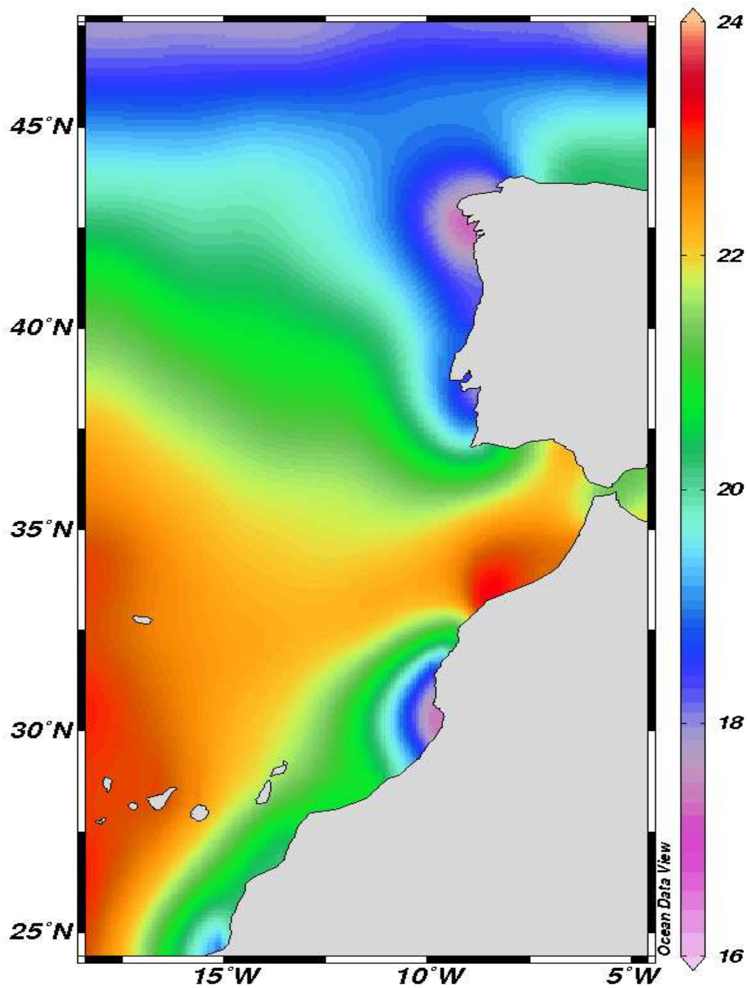
We run an MCMC. To determine convergence we use the diagnostics available in `BOA` to set the burn-in (1,200 iterations), the thinning (1/3) and the sample size (6,000 from the thinned chain). We also performed two separate runs, a warm start configuration (30°) and a cold one (15°).

Results: January Climatology

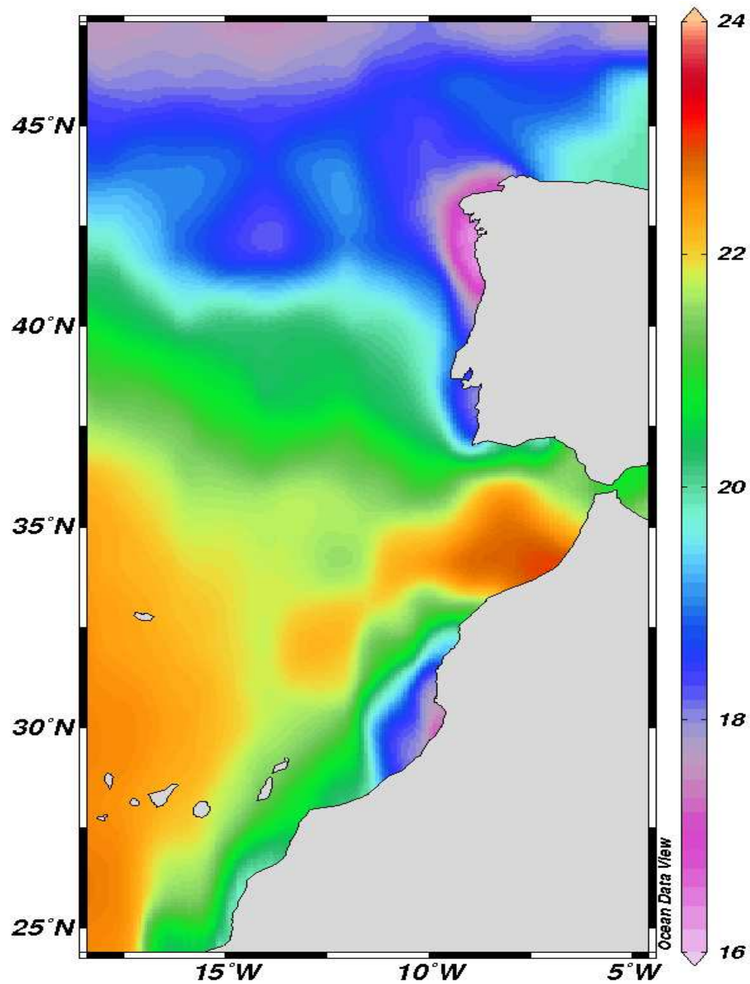


Results: August Climatology

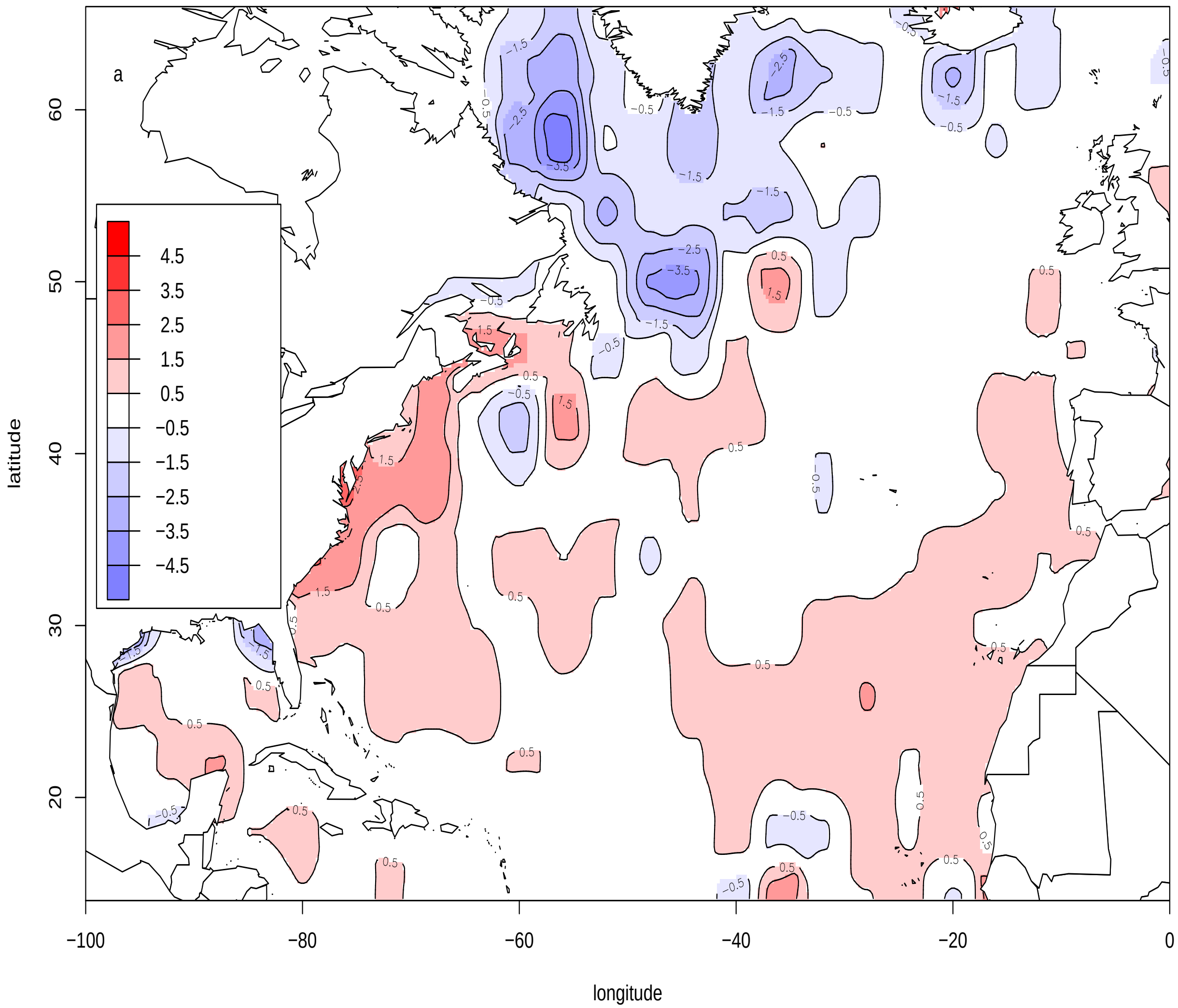
August SST from WOA01



August SST from LS09

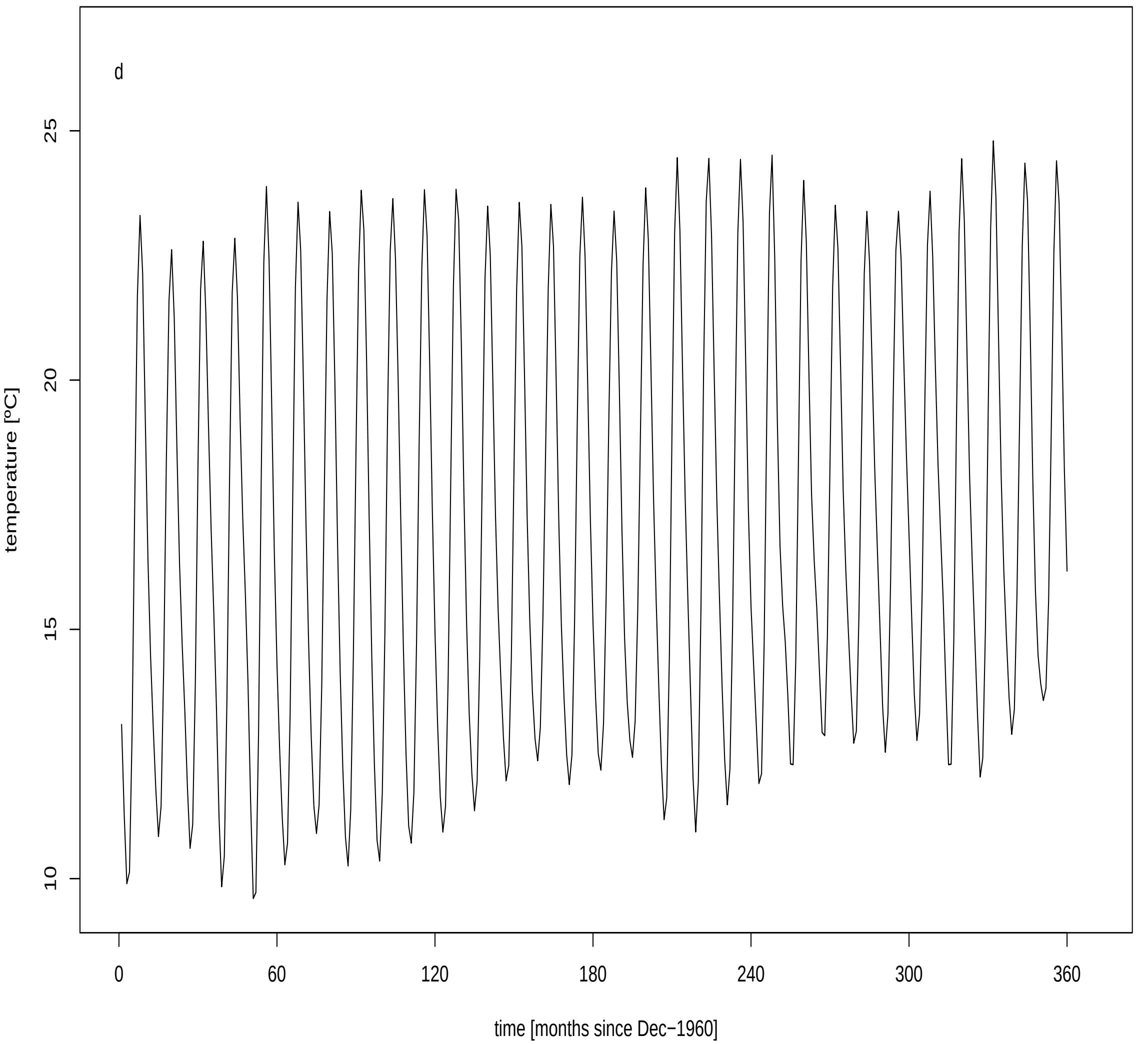


Results: Trends



Trends (°C/30 years)

Posterior mean for monthly SST at $\boldsymbol{s}_2$ (°C).



SST Biases

A collection of SST data from 1662 to 2007 is available from International Comprehensive Ocean-Atmosphere Data Set (ICOADS).

SST Biases

A collection of SST data from 1662 to 2007 is available from International Comprehensive Ocean-Atmosphere Data Set (ICOADS).

ICOADS reports 12 different measurement methods that correspond to observations collected by ships, research vessels, moored buoys, drifting buoys, and oceanographic stations.

SST Biases

A collection of SST data from 1662 to 2007 is available from International Comprehensive Ocean-Atmosphere Data Set (ICOADS).

ICOADS reports 12 different measurement methods that correspond to observations collected by ships, research vessels, moored buoys, drifting buoys, and oceanographic stations.

Each measurement method has a different associated bias. Moreover, measurement errors corresponding to the same device are likely to be correlated.

SST Biases: Observation Equation

The SST metadata often does not report the measurement method used for a given observation. This uncertainty needs to be accounted for in the analysis. We achieve this by proposing a mixture model for the biases.

SST Biases: Observation Equation

The SST metadata often does not report the measurement method used for a given observation. This uncertainty needs to be accounted for in the analysis. We achieve this by proposing a mixture model for the biases.

Ignore the time index and let $x_{ik}(\boldsymbol{s})$ be the k-th observation at location $\boldsymbol{s}$ from device i. Then

$$x_{ik}(\boldsymbol{s}) = \theta(\boldsymbol{s}) + \beta_i + \varepsilon_{ik}(\boldsymbol{s}), \quad \varepsilon_{ik}(\boldsymbol{s}) \sim N(0, \sigma_i^2)$$

SST Biases: Process Equation

We use a reduced rank model for the random field of SSTs, thus

$$\theta(\boldsymbol{s}) = \mu(\boldsymbol{s}) + \boldsymbol{B}'(\boldsymbol{s})\boldsymbol{\gamma} + v(\boldsymbol{s}) \ \ , \ \ v \sim N(0, \Phi^2(\boldsymbol{s}))$$

Here $\mu(\boldsymbol{s})$ is a linear function of the location. $\boldsymbol{B}(\boldsymbol{s})$ is a vector obtained from the evaluations of basis functions, that may depend on a set of parameters $\boldsymbol{\phi}$. Finally $\boldsymbol{\gamma} \sim N(\boldsymbol{0}, \boldsymbol{K})$.

SST Biases: Process Equation

We use a reduced rank model for the random field of SSTs, thus

$$\theta(\boldsymbol{s}) = \mu(\boldsymbol{s}) + \boldsymbol{B}'(\boldsymbol{s})\boldsymbol{\gamma} + v(\boldsymbol{s}) \ \ , \ \ v \sim N(0, \Phi^2(\boldsymbol{s}))$$

Here $\mu(\boldsymbol{s})$ is a linear function of the location. $\boldsymbol{B}(\boldsymbol{s})$ is a vector obtained from the evaluations of basis functions, that may depend on a set of parameters $\boldsymbol{\phi}$. Finally $\boldsymbol{\gamma} \sim N(\boldsymbol{0}, \boldsymbol{K})$.

In our application $\boldsymbol{K}(\boldsymbol{s})$ is obtained from kernel convolutions and $\boldsymbol{\gamma}$ is a Markov random field.

SST Biases: Priors

As there is sometimes uncertainty about the type of measurement, we consider the prior for β_i

$$\beta_i \sim \sum_{j=1}^{p} \lambda_{ij} N(\mu_j, \tau_j^2) \ \ , \ \ p(\mu_j, \tau_j^2) = N(\mu_j | m_j, \kappa\tau_j^2) IG(\tau_j^2 | a_j, b_j).$$

Information about the parameters of these priors can be obtained from the published literature.

SST Biases: Priors

As there is sometimes uncertainty about the type of measurement, we consider the prior for β_i

$$\beta_i \sim \sum_{j=1}^{p} \lambda_{ij} N(\mu_j, \tau_j^2) \ \ , \ \ p(\mu_j, \tau_j^2) = N(\mu_j | m_j, \kappa \tau_j^2) IG(\tau_j^2 | a_j, b_j).$$

Information about the parameters of these priors can be obtained from the published literature.

The variance of each device follows the prior

$$p(\sigma_i^2 | \sigma^2) = IG(\sigma_i^2 | \alpha + 1, \alpha \sigma^2) \ \ , \ \ p(\sigma^2) = Ga(\sigma^2 | a, b).$$

Thus, a priori, we assume that $E(\sigma_i^2) = \sigma^2$.

SST Biases: Priors

As there is sometimes uncertainty about the type of measurement, we consider the prior for β_i

$$\beta_i \sim \sum_{j=1}^{p} \lambda_{ij} N(\mu_j, \tau_j^2) \;\;,\;\; p(\mu_j, \tau_j^2) = N(\mu_j | m_j, \kappa\tau_j^2) IG(\tau_j^2 | a_j, b_j).$$

Information about the parameters of these priors can be obtained from the published literature.

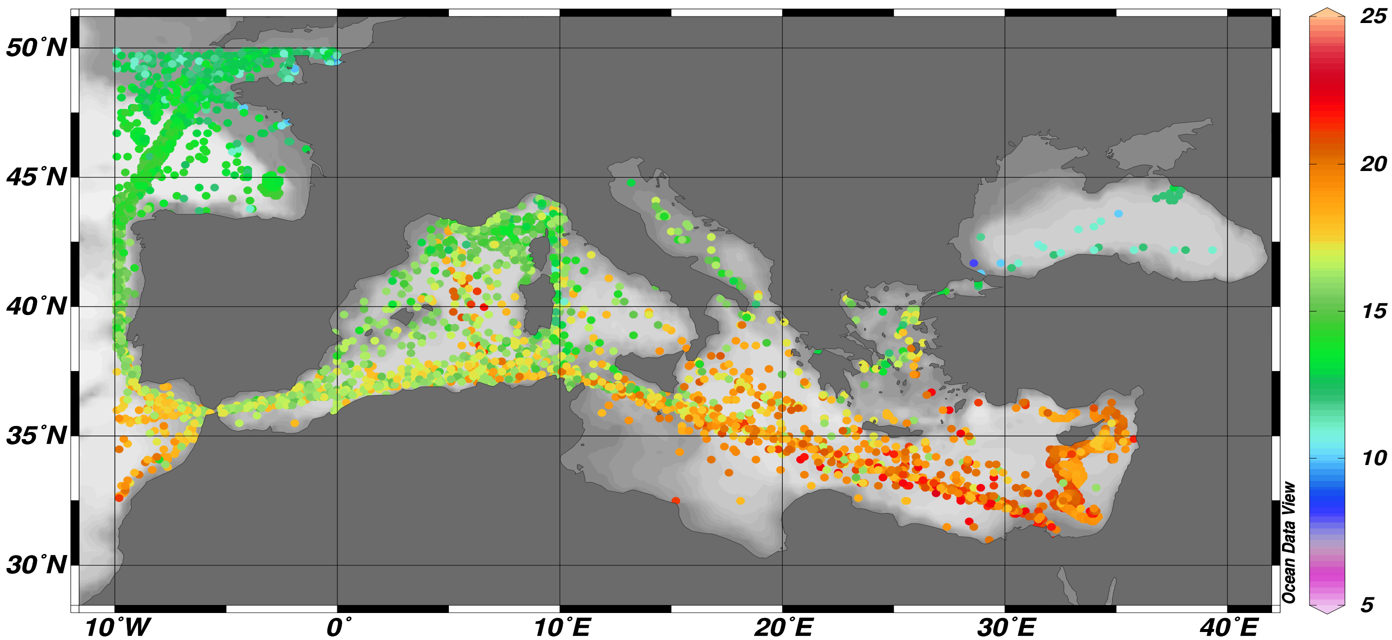
The variance of each device follows the prior

$$p(\sigma_i^2 | \sigma^2) = IG(\sigma_i^2 | \alpha + 1, \alpha\sigma^2) \;\;,\;\; p(\sigma^2) = Ga(\sigma^2 | a, b).$$

Thus, a priori, we assume that $E(\sigma_i^2) = \sigma^2$.

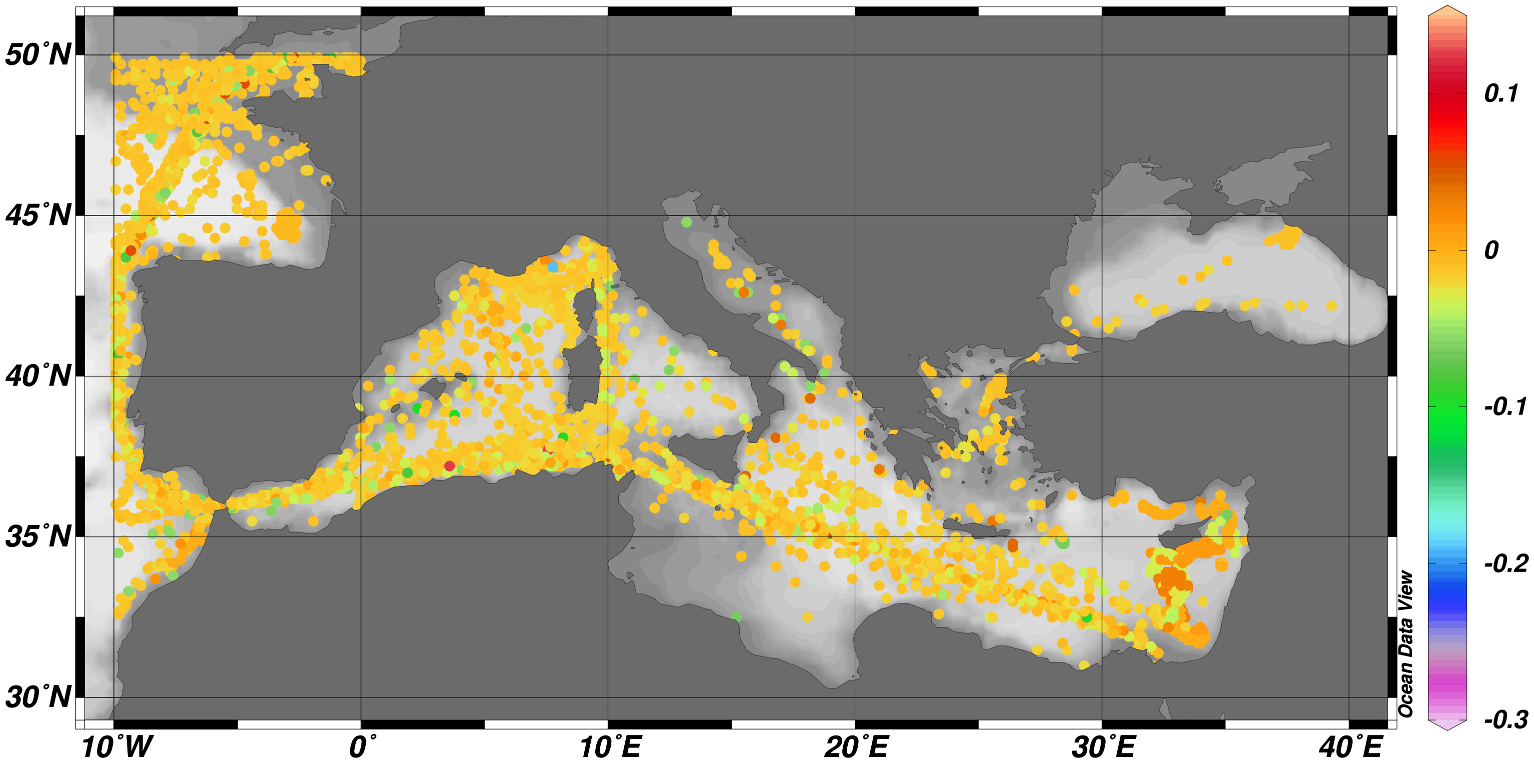
Finally, $\boldsymbol{\lambda}_i = (\lambda_{i1}, \ldots, \lambda_{ip}) \sim Mult(1; \pi_{i1}, \ldots, \pi_{ip})$.

SST Mediterranean Data



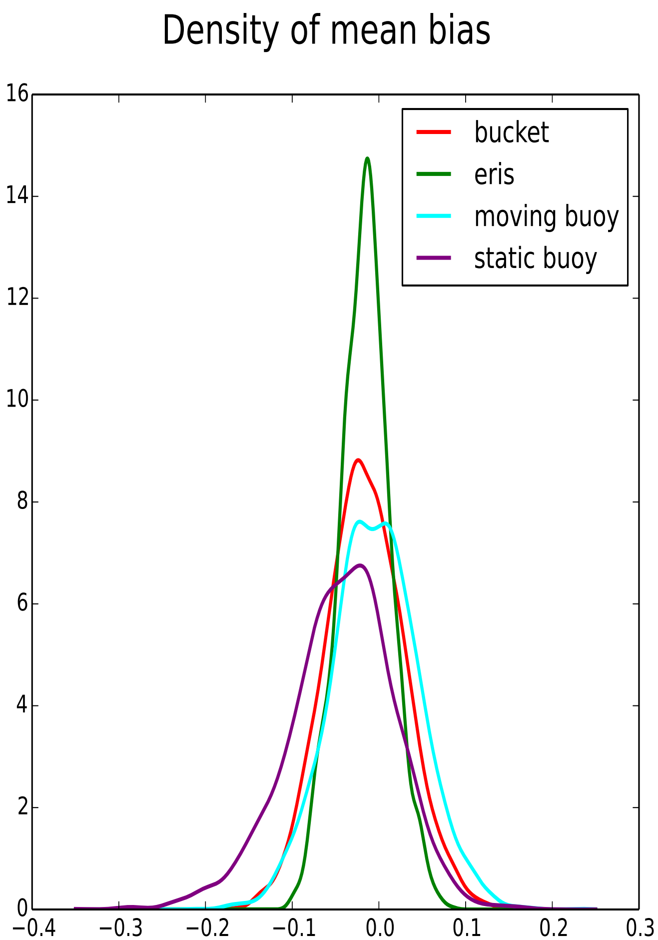
- We consider the data for the Mediterranean Sea in December 2003.
- There are 12,210 SST observations in this datatset.
- We consider four types of measurements: moored buoys, moving buoys, engine room intake, and buckets.

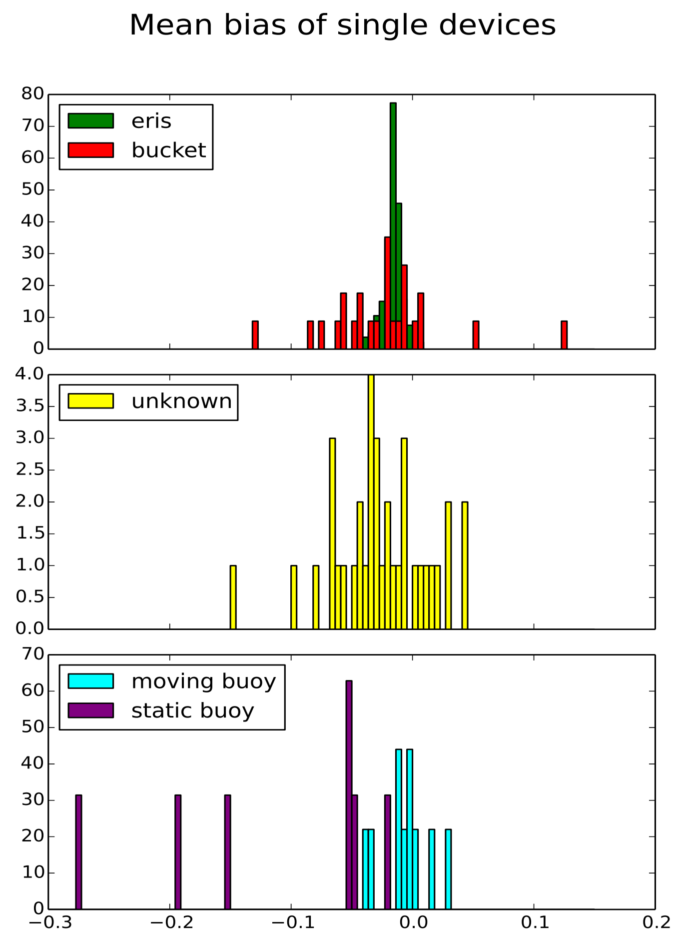
Estimated Biases



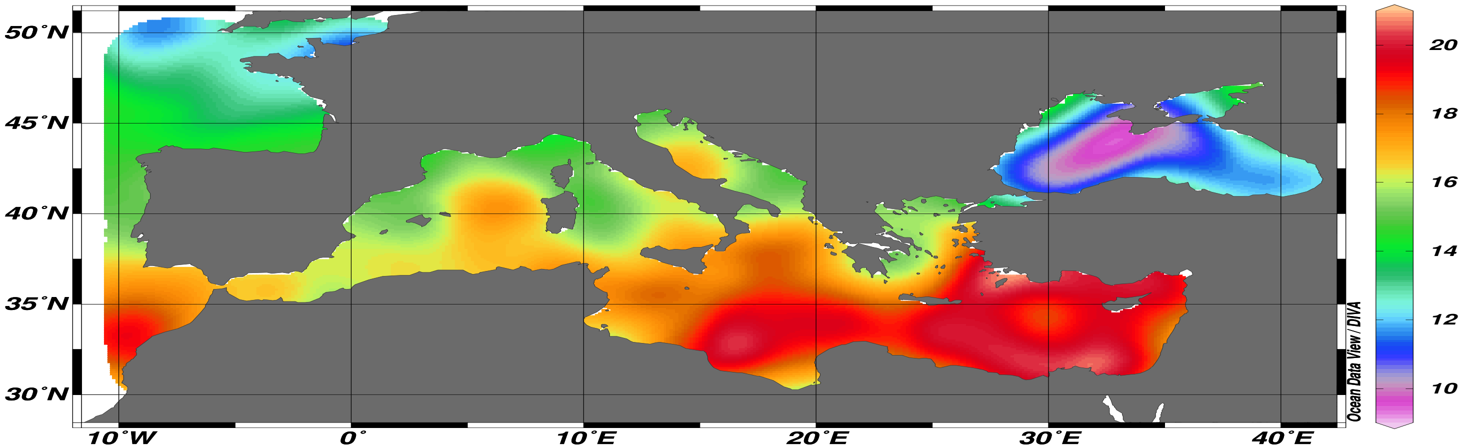
Mean estimated bias by location

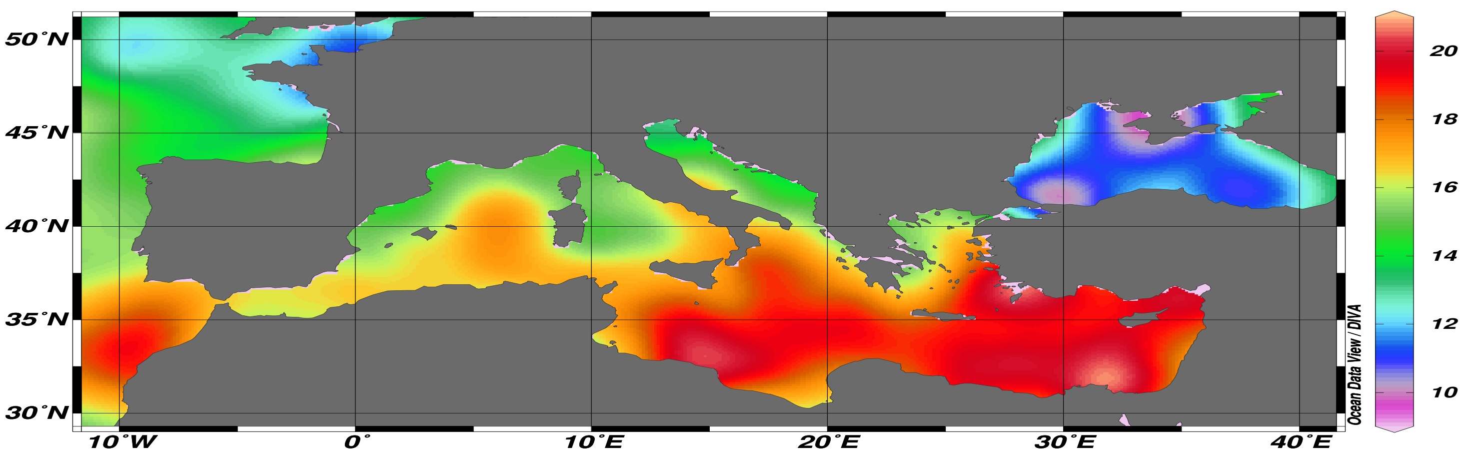
Distributions of the Biases





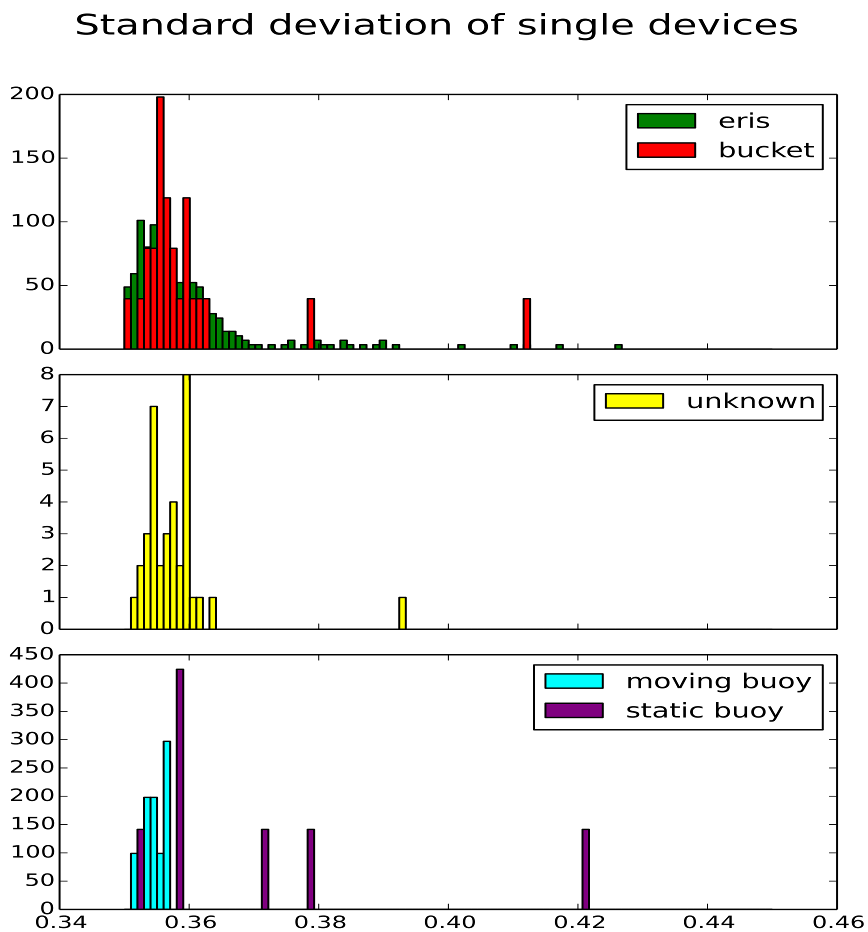
Estimated SST





Two samples of the posterior predictive SST

Observational Error Variance



Histograms of the estimated standar deviations of the observational errors for the different devices, by type.

Model Extensions

- Extend the model to larger domains.
- Consider a temporal component that includes trends and time-varying cycles.
- Consider space-varying biases.
- Include satellite data. This requires dealing with a change of support problem.

References

- Ricardo T. Lemos, Bruno Sansó (2009) “A Spatio-Temporal Model for Mean, Anomaly and Trend Fields of North Atlantic Sea Surface Temperature (with discussion)”. *Journal of the American Statistical Association*, 104, pp. 5–25.