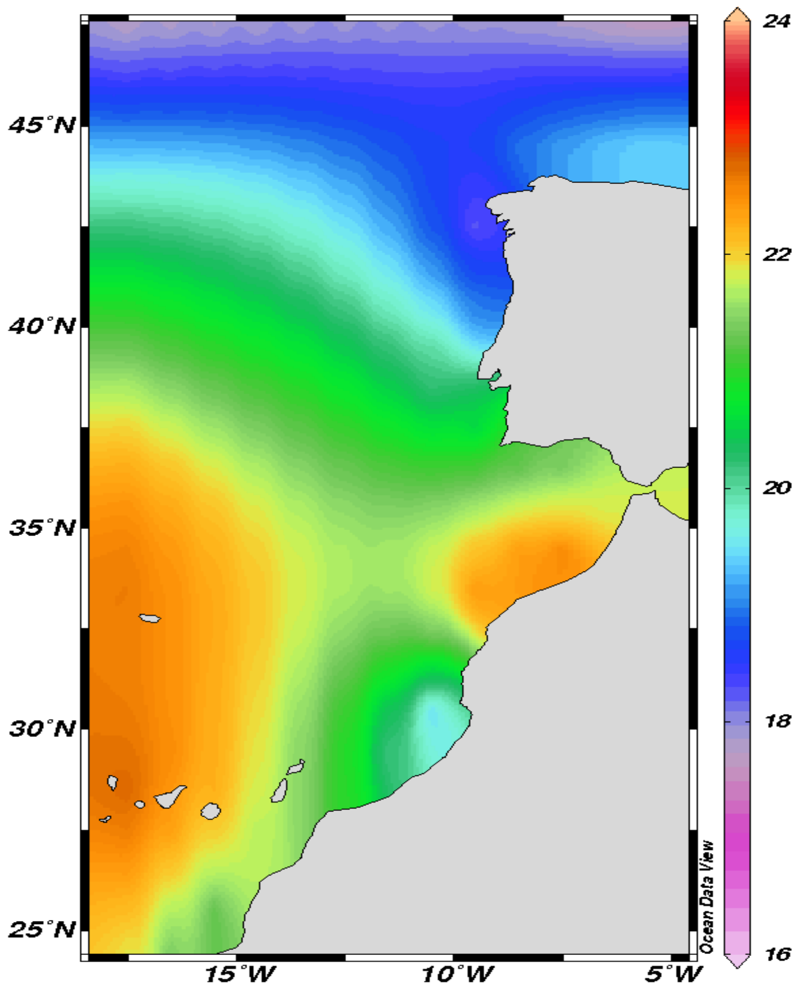
A Hierarchical Bayesian Model for Ocean Properties Reconstructions

Bruno Sansó

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

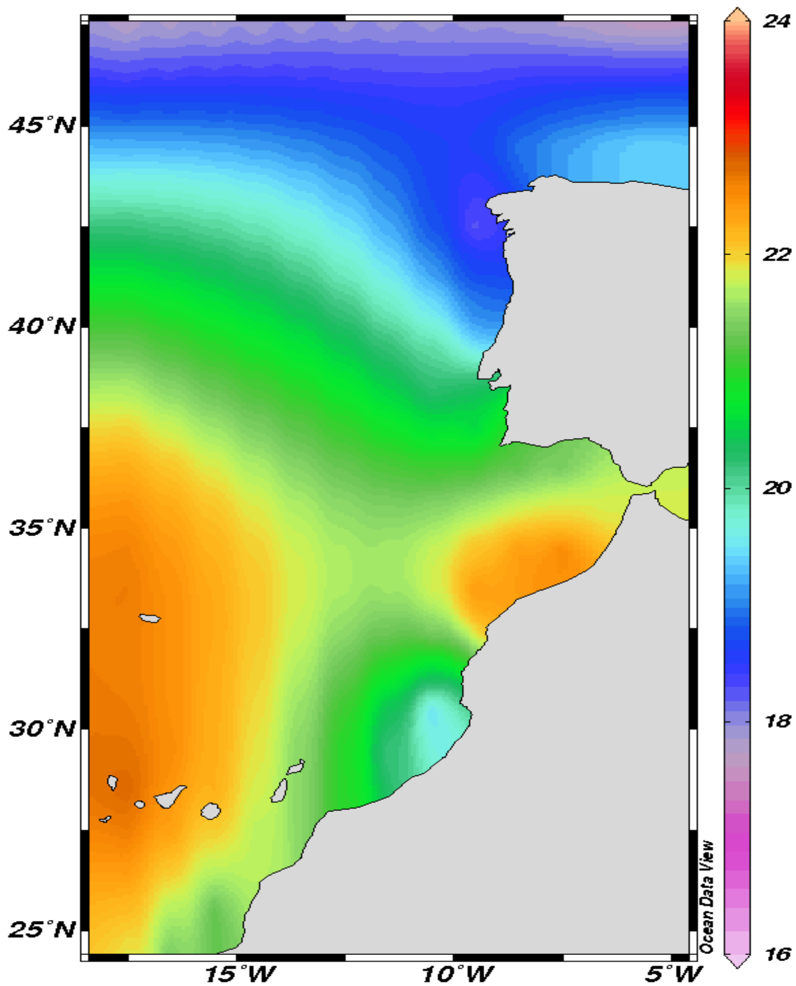
In collaboration with Ricardo Lemos, NOAA-NMFS, Pacific Grove, CA.

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.

Goals

To produce a realistic climatology we decided to obtain a reconstruction of SSTs with the following properties:

Goals

To produce a realistic climatology we decided to obtain a reconstruction of SSTs with the following properties:

- The reconstruction must be smooth and realistic, when compared to other existing products.

Goals

To produce a realistic climatology we decided to obtain a reconstruction of SSTs with the following properties:

- The reconstruction must be smooth and realistic, when compared to other existing products.
- Anomaly fields must capture medium- to large-scale features and average zero everywhere

Goals

To produce a realistic climatology we decided to obtain a reconstruction of SSTs with the following properties:

- The reconstruction must be smooth and realistic, when compared to other existing products.
- Anomaly fields must capture medium- to large-scale features and average zero everywhere
- Smooth trend fields.

Goals

To produce a realistic climatology we decided to obtain a reconstruction of SSTs with the following properties:

- The reconstruction must be smooth and realistic, when compared to other existing products.
- Anomaly fields must capture medium- to large-scale features and average zero everywhere
- Smooth trend fields.
- All fields must be accompanied by measures of uncertainty.

Goals

To produce a realistic climatology we decided to obtain a reconstruction of SSTs with the following properties:

- The reconstruction must be smooth and realistic, when compared to other existing products.
- Anomaly fields must capture medium- to large-scale features and average zero everywhere
- Smooth trend fields.
- All fields must be accompanied by measures of uncertainty.
- The method must be useful in large geographical domains and long time frames.

Goals

To produce a realistic climatology we decided to obtain a reconstruction of SSTs with the following properties:

- The reconstruction must be smooth and realistic, when compared to other existing products.
- Anomaly fields must capture medium- to large-scale features and average zero everywhere
- Smooth trend fields.
- All fields must be accompanied by measures of uncertainty.
- The method must be useful in large geographical domains and long time frames.
- Observational errors should be accounted for.

Bayesian Hierarchical Models

We propose a statistical model based on the Bayesian hierarchical structure. This consists of three layers:

• Observation Equation: relationship between observations and the true state of nature.

• Process equation: probabilistic description of the space and time variability of the state vector.

• Prior distributions: prior knowledge on the parameters that define the previous two layers.

Bayesian Hierarchical Models

- The hierarchal approach offers a coherent probabilistic framework to combine information from different sources.

Bayesian Hierarchical Models

- The hierarchal approach offers a coherent probabilistic framework to combine information from different sources.
- All estimation uncertainty is propagated through the different levels of the model and fully accounted for.

Bayesian Hierarchical Models

- The hierarchal approach offers a coherent probabilistic framework to combine information from different sources.
- All estimation uncertainty is propagated through the different levels of the model and fully accounted for.
- These models lend themselves to the use of Monte Carlo methods. This provides for easy probabilistic inference for quantities derived from the model parameters.

A non-homogeneous spatially-varying process can be written as

$$X(\boldsymbol{s}) = \sum_{i=1}^{m} b(\boldsymbol{s} - \boldsymbol{s}_i^*; \boldsymbol{w}(s))\gamma_i, \quad \boldsymbol{\gamma} \sim N(0, \boldsymbol{K})$$

where, $\boldsymbol{s}_1^*, \ldots, \boldsymbol{s}_m^*$ is a grid, and

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

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

A non-homogeneous spatially-varying process can be written as

$$X(\boldsymbol{s}) = \sum_{i=1}^{m} b(\boldsymbol{s} - \boldsymbol{s}_i^*; \boldsymbol{w}(s))\gamma_i, \quad \boldsymbol{\gamma} \sim N(0, \boldsymbol{K})$$

where, $\boldsymbol{s}_1^*, \ldots, \boldsymbol{s}_m^*$ is a grid, and

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

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

The distance is given as

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

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).

The ellipsoidal shape is controlled by the parameters in

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

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

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

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

The spatial variation of $\boldsymbol{\omega}$ is obtained, with a normalized b, as

$$\boldsymbol{\omega}(\boldsymbol{s}) = \sum_{i=1}^{m} b(\boldsymbol{s} - \boldsymbol{s}_i^*; \boldsymbol{u})\boldsymbol{\rho}(\boldsymbol{s}_i^*) \quad \boldsymbol{u} = (2, 1, 0, 0)$$

with appropriate uniform priors on each $\rho_k(\boldsymbol{s}_i^*), k = 1, \ldots, 4$.

The SST observation $x_{i,m,y}(\boldsymbol{s})$ corresponding to data set $i = 1, \ldots, 4$ (OSD, CTD, XBT, MBT), in month m, year y and location $\boldsymbol{s}$ follows

$$x_{i,m,y}(\boldsymbol{s}) = \theta_{m,y}(\boldsymbol{s}) + \varepsilon_{i,m,y}(\boldsymbol{s}), \;\; \varepsilon_{i,m,y}(\boldsymbol{s}) \sim N\left(0, \tau_i^2\right).$$

Different τ_i allow for different observational error variances depending on the instrument.

Climatology:

$$\Xi_m(\boldsymbol{s}) = \frac{1}{30} \sum_{y=1961}^{1990} \theta_{m,y}(\boldsymbol{s}),$$

Climatology:

$$\Xi_m(\boldsymbol{s}) = \frac{1}{30} \sum_{y=1961}^{1990} \theta_{m,y}(\boldsymbol{s}),$$

Anomaly:

$$\Delta_{m,y}(\boldsymbol{s}) = \theta_{m,y}(\boldsymbol{s}) - \Xi_m(\boldsymbol{s}).$$

The true SST is

$$\theta_{m,y}(\boldsymbol{s}) \sim N\left(\sum_{\boldsymbol{j}} b(\boldsymbol{s}-\boldsymbol{j};\boldsymbol{\Lambda}(\boldsymbol{s}))\gamma(\boldsymbol{j}), \Phi(\boldsymbol{s})^2\right)$$

where

$$\gamma(\boldsymbol{j}) = \alpha(\boldsymbol{j}) + \boldsymbol{\beta}_t(\boldsymbol{j})\boldsymbol{w}_t^T + \eta(\boldsymbol{j})(t-180),$$

$\boldsymbol{j}$ denotes a points in a 4° resolution grid, and

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

Here $t = m + 12(y-1961)$ denotes time in months since December 1960. Note that $\theta_{m,y}(\boldsymbol{s})$ is continuous in space.

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.

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.

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.

Implementation

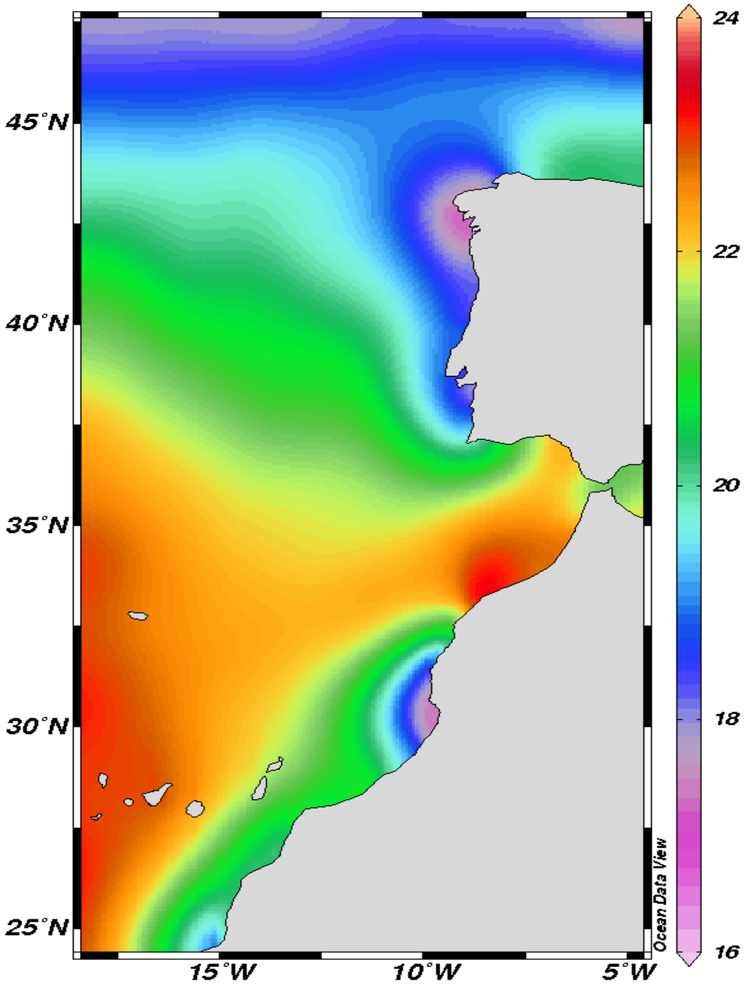
The use of a compactly supported kernel allows for an efficient parallel implementation. We use 13 processors, each one working with two columns of J.

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

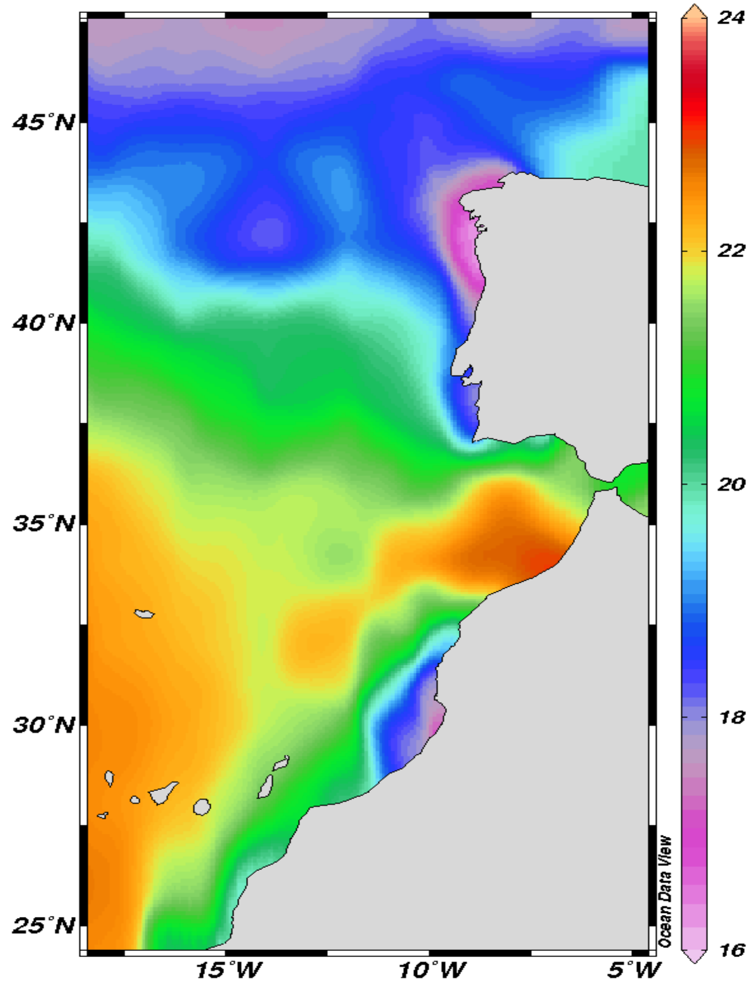
We run an MCMC. To determine convergence we use the diagnostics available in `BOA` to set the burn-in (1,200 iterations), the thinning (1/3) and the sample size (6,000 from the thinned chain). We also performed two separate runs, a warm start configuration (30°) and a cold one (15°).

Results: August Climatology

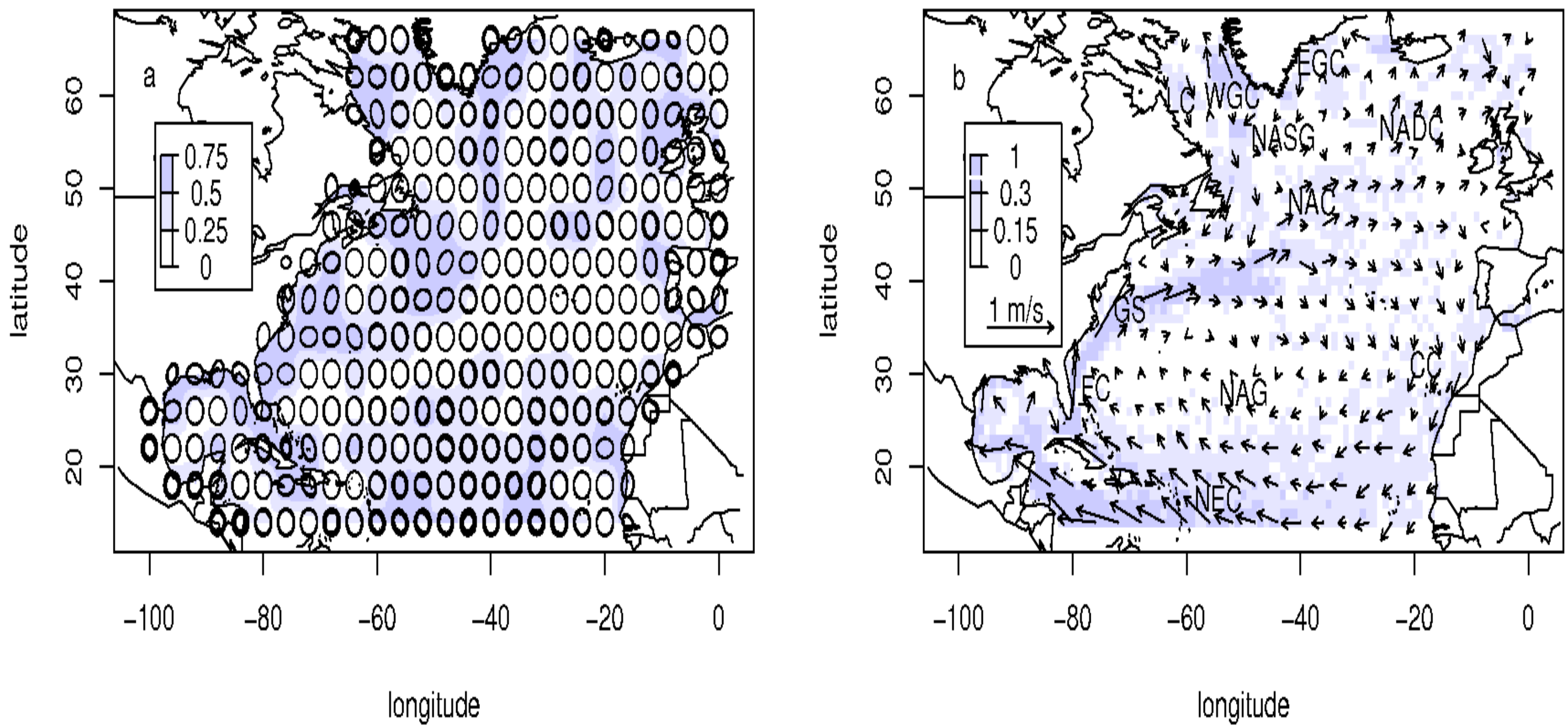
August SST from WOA01

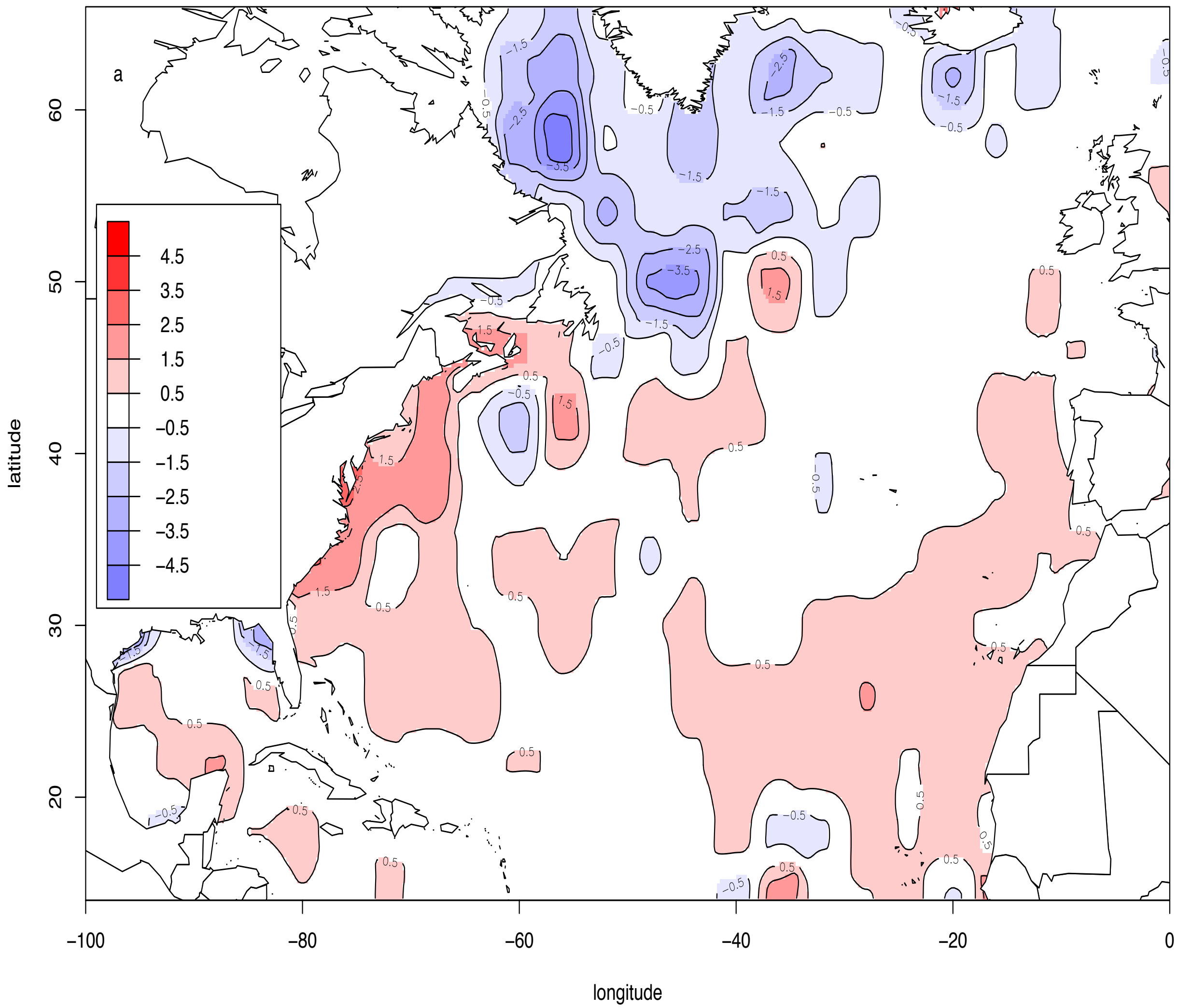


August SST from LS09



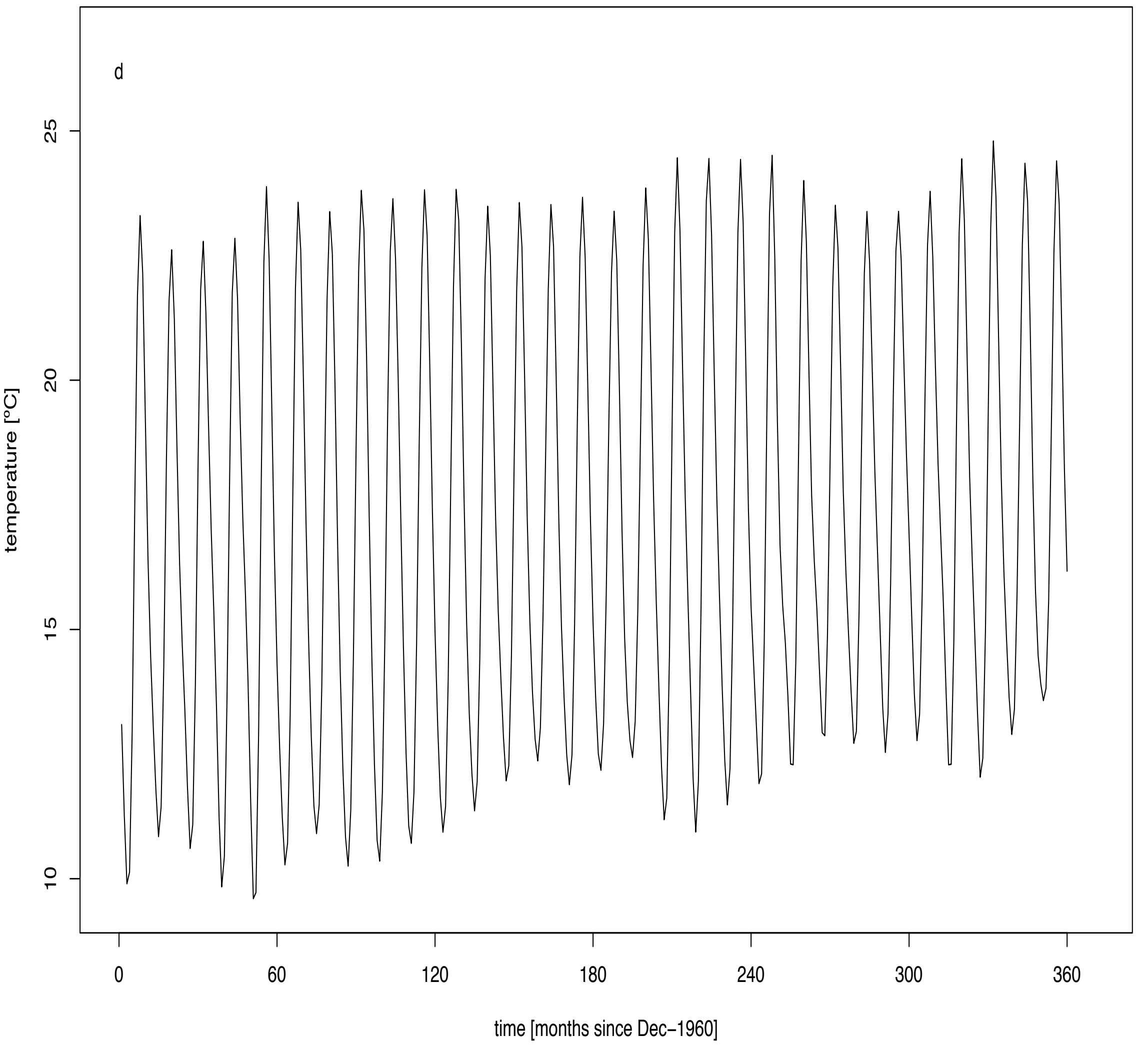
Results: Kernels Vs Currents





Trends (°C/30 years)

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



Model for Salinity and Temperature

The development of a model that considers temperature and salinity jointly over a 3D domain needs to account for the following issues:

Model for Salinity and Temperature

The development of a model that considers temperature and salinity jointly over a 3D domain needs to account for the following issues:

- Data sparsity. There are substantially more observations for the surface of the ocean than for the deep water. Furthermore, salinity is sampled about 10 times less than temperature.

Model for Salinity and Temperature

The development of a model that considers temperature and salinity jointly over a 3D domain needs to account for the following issues:

• Data sparsity. There are substantially more observations for the surface of the ocean than for the deep water. Furthermore, salinity is sampled about 10 times less than temperature.

• The column of water needs to satisfy a density stability constraint. That is, density must increase with depth. Density is related to temperature and salinity via the equation of state.

Model for Salinity and Temperature

The development of a model that considers temperature and salinity jointly over a 3D domain needs to account for the following issues:

- Data sparsity. There are substantially more observations for the surface of the ocean than for the deep water. Furthermore, salinity is sampled about 10 times less than temperature.

- The column of water needs to satisfy a density stability constraint. That is, density must increase with depth. Density is related to temperature and salinity via the equation of state.

- The previous two points imply that we need to consider a hierarchical structure for our statistical model that incorporates some physical constraints.

An Ocean Model

A full description of the ocean dynamics requires the estimation of the following properties:

An Ocean Model

A full description of the ocean dynamics requires the estimation of the following properties:

- Velocities: $\boldsymbol{u} = (u^x, u^y, u^z)$

An Ocean Model

A full description of the ocean dynamics requires the estimation of the following properties:

- Velocities: $\boldsymbol{u} = (u^x, u^y, u^z)$
- Temperature T and salinity S

A full description of the ocean dynamics requires the estimation of the following properties:

- Velocities: $\boldsymbol{u} = (u^x, u^y, u^z)$
- Temperature T and salinity S
- Pressure p and density ρ.

Based on observations that we denote as $\boldsymbol{Y}$ we want to build a statistical model to explore the posterior distribution

$$\pi(\boldsymbol{u}, T, S, p, \rho | \boldsymbol{Y})$$

We write

$$\pi(\boldsymbol{u}, T, S, p, \rho | \boldsymbol{Y}) = \pi(T, S | \boldsymbol{u}, \boldsymbol{Y}) \pi(\boldsymbol{u} | \rho, p, \boldsymbol{Y}) \pi(p | \rho, \boldsymbol{Y}) \pi(\rho | \boldsymbol{Y})$$

For the remaining of the talk I will focus on the model for density, $\pi(\rho|\boldsymbol{Y})$.

Observational Model

We let $\boldsymbol{Y}_t = (Y_t(\boldsymbol{s}_1), \ldots, Y_t(\boldsymbol{s}_{n_t}))^T$, be the vector of densities at time t for all locations where records are available at that time. Then

$$\boldsymbol{Y}_t \sim N(\boldsymbol{F}^o \boldsymbol{\rho}_t, \mathrm{diag}(\boldsymbol{u}_t))$$

where $\boldsymbol{F}^o$ is an incidence matrix and $\boldsymbol{\rho}_t$ is a vector of stacked true, unobserved, density profiles on a $.25°$ resolution grid.

Process Model

Following the same framework of the SST model, we let

$$\boldsymbol{\rho}_t \sim N(\boldsymbol{F}(\boldsymbol{\mu}_{m(t)} + \boldsymbol{W}\boldsymbol{\gamma}_t), \tau \boldsymbol{I})$$

$m(t)$ is the month corresponding to time t. Thus, there are 12 vectors $\boldsymbol{\mu}_t$ that determine the regular seasonal cycle.

Process Model

Following the same framework of the SST model, we let

$$\boldsymbol{\rho}_t \sim N(\boldsymbol{F}(\boldsymbol{\mu}_{m(t)} + \boldsymbol{W}\boldsymbol{\gamma}_t), \tau \boldsymbol{I})$$

$m(t)$ is the month corresponding to time t. Thus, there are 12 vectors $\boldsymbol{\mu}_t$ that determine the regular seasonal cycle.

The elements of $\boldsymbol{\mu}_t$ are sampled from a multivariate normal truncated to impose that its components are in increasing order. $\boldsymbol{F}$ is obtained from Bezièr kernels centered on a 1° resolution grid. Whence, the product $\boldsymbol{F}\boldsymbol{\mu}_t$ has its components in increasing order.

$\boldsymbol{W}$ is a matrix of spatial factors and $\boldsymbol{\gamma}_t$ is a low dimensional vector capturing time-varying trends and low frequency cycles. Thus

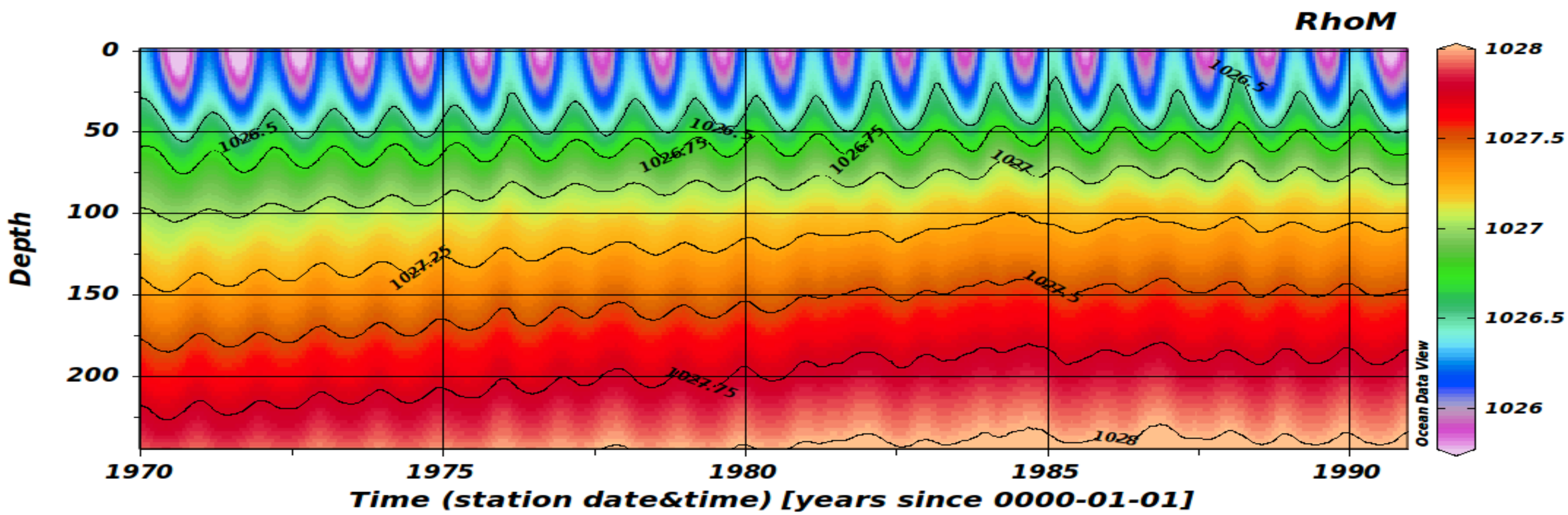
$$\boldsymbol{\gamma}_t \sim N(\boldsymbol{G}\boldsymbol{\gamma}_{t-1}, \boldsymbol{H})$$

for a rotation matrix $\boldsymbol{G}$ and an appropriately chosen matrix $\boldsymbol{H}$.

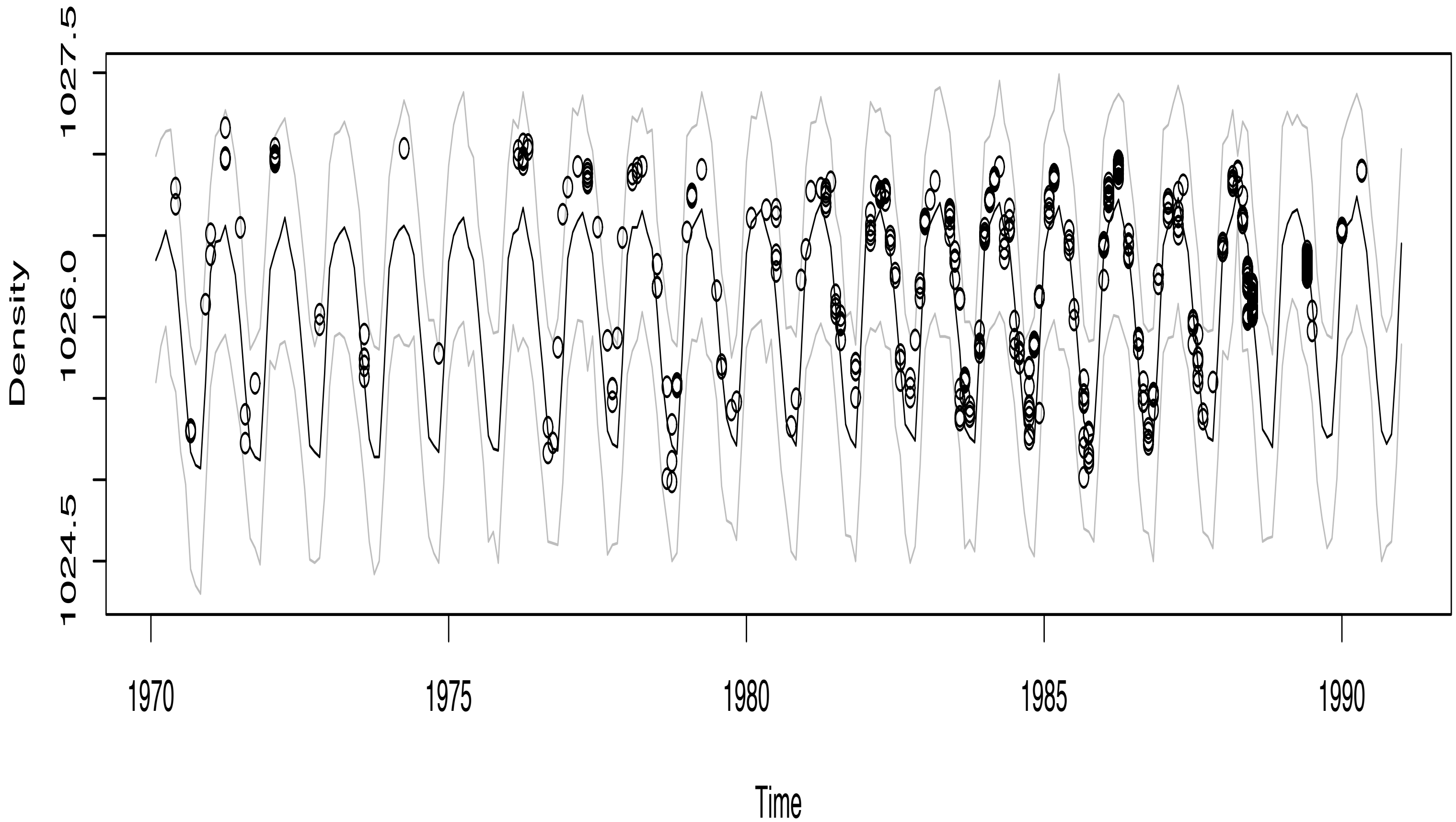
Preliminary Results

We considered a $2^\circ \times 2^\circ$ grid with 33 layers in the vertical centered at 39N 12W (off Lisbon). We used NODC's OSD and CTD data from 1970 to 1991

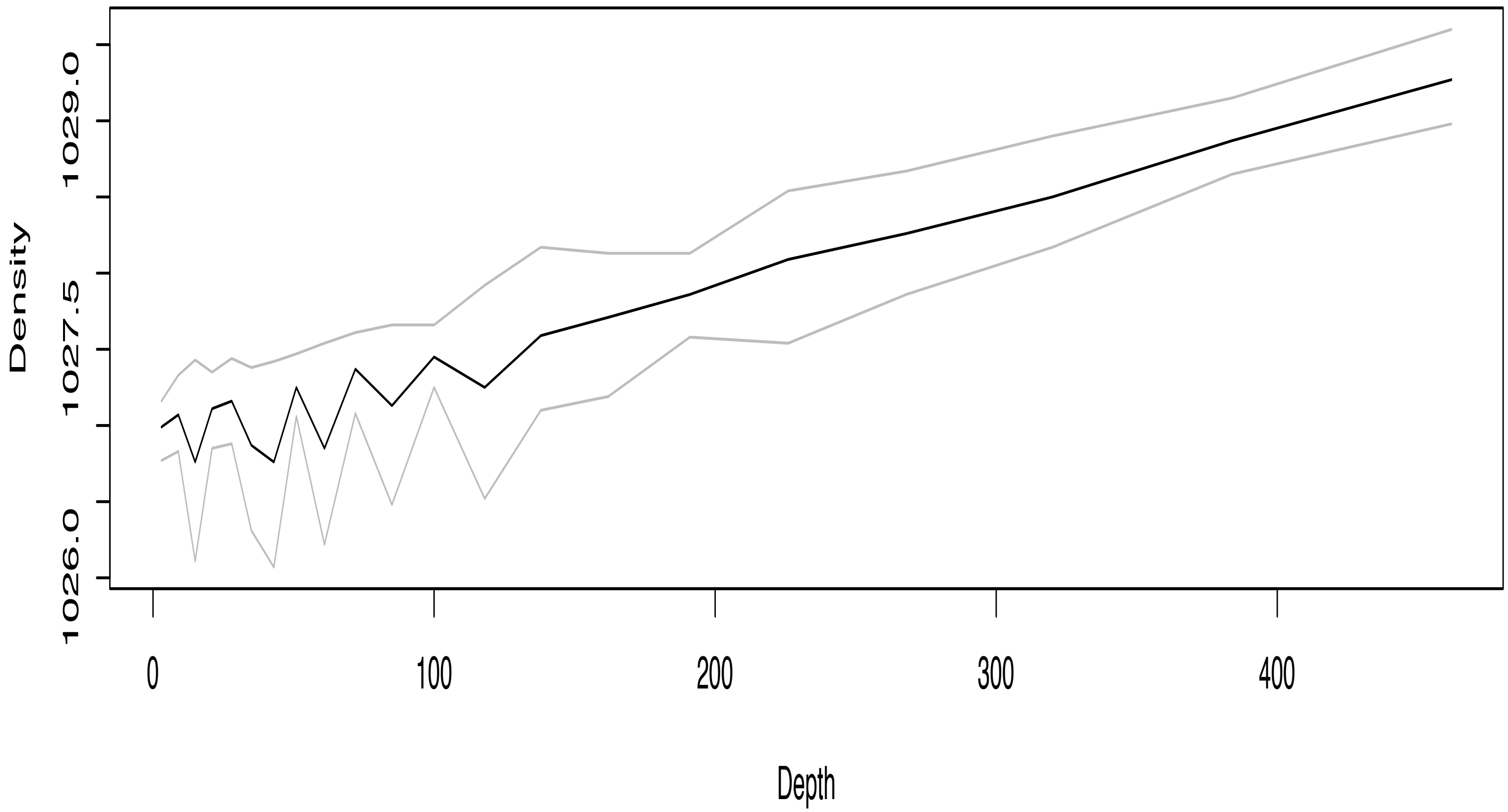
Density Profile



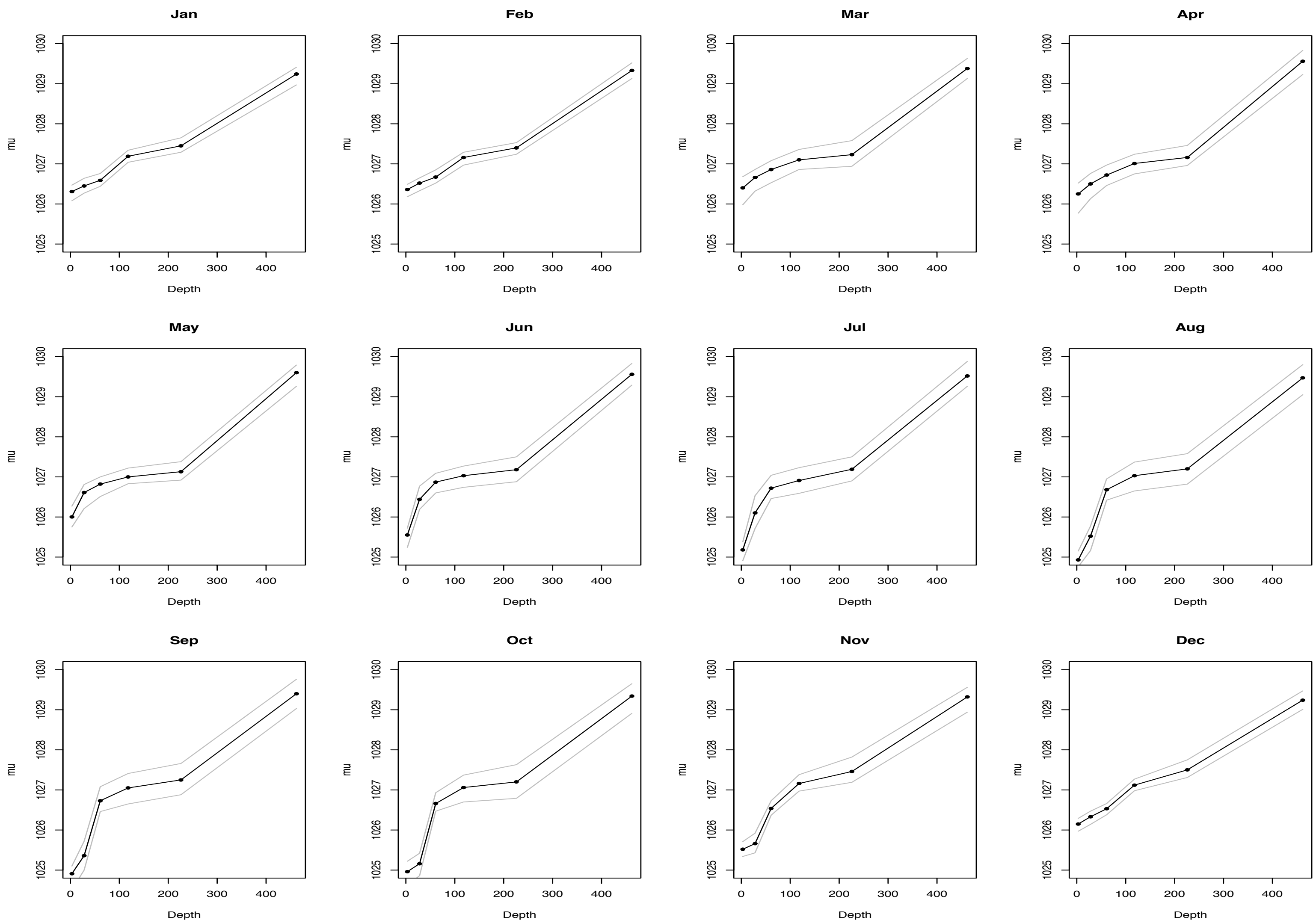
Surface Density



Single Location Profile



Mean Density Profile



Discussion

- We have developed spatio-temporal models for ocean variables that are non-stationary in either space or time. We are able to capture important space- and time-varying features of the ocean.

Discussion

• We have developed spatio-temporal models for ocean variables that are non-stationary in either space or time. We are able to capture important space- and time-varying features of the ocean.

• Our model includes observational errors and realistic descriptions of the latent processes governing the evolution of ocean variables.

Discussion

• We have developed spatio-temporal models for ocean variables that are non-stationary in either space or time. We are able to capture important space- and time-varying features of the ocean.

• Our model includes observational errors and realistic descriptions of the latent processes governing the evolution of ocean variables.

• The model is able to provide probabilistic assessments of the variabilities included in the estimated quantities. All estimation variabilities are accounted for in the final product.

Discussion

- We have developed spatio-temporal models for ocean variables that are non-stationary in either space or time. We are able to capture important space- and time-varying features of the ocean.

- Our model includes observational errors and realistic descriptions of the latent processes governing the evolution of ocean variables.

- The model is able to provide probabilistic assessments of the variabilities included in the estimated quantities. All estimation variabilities are accounted for in the final product.

- Our model is able to handle large data sets. By using kernels with compact support and making use of the structure of the CDLM we are able to parallelize the estimation algorithms.

- Reference: Ricardo T. Lemos, Bruno Sansó (2006)

"Spatio-temporal Variability of Ocean Temperature in the Portugal Current System". *Journal of Geophysical Research Oceans*, 111, C04010, doi:10.1029/2005JC003051.

Discussion

- Reference: Ricardo T. Lemos, Bruno Sansó (2006) "Spatio-temporal Variability of Ocean Temperature in the Portugal Current System". *Journal of Geophysical Research Oceans*, 111, C04010, doi:10.1029/2005JC003051.

- Reference: 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. **Winner of the 2010 Mitchell Prize to the best applied Bayesian paper!**