

UNIVERSITY OF CALIFORNIA  
SANTA CRUZ

**MEASURING BIAS OF SEA SURFACE TEMPERATURE  
MEASUREMENT DEVICES IN THE MEDITERRANEAN SEA**

A project submitted in partial satisfaction of the  
requirements for the degree of

Master of Science

in

STATISTICS AND APPLIED MATHEMATICS

by

**Celeste Tretto**

August 2014

The project of Celeste Tretto  
is approved by:

---

Professor Bruno Sanso

---

Professor Tatiana Xifara

# Table of Contents

<b>List of Figures</b>	<b>iv</b>
<b>List of Tables</b>	<b>v</b>
<b>Acknowledgments</b>	<b>vi</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Review of measurement methods</b>	<b>4</b>
2.1 Ships . . . . .	4
2.1.1 Buckets . . . . .	5
2.1.2 ERIs . . . . .	6
2.2 Buoys . . . . .	6
2.3 Preliminary data analysis . . . . .	7
<b>3 Model</b>	<b>12</b>
3.1 Sea Surface Temperature Process . . . . .	13
3.1.1 Convolution-based Gaussian processes . . . . .	13
3.2 Device bias . . . . .	17
3.2.1 Hierarchical models . . . . .	17
3.2.2 Mixture models for missing metadata . . . . .	20
3.3 Hierarchical model for observational variance . . . . .	21
3.4 Markov Chain Monte Carlo . . . . .	23
3.4.1 Description of the algorithm . . . . .	24
<b>4 Model results</b>	<b>26</b>
4.1 Estimation of device bias . . . . .	26
4.1.1 Influence of rounded data . . . . .	29
4.1.2 Robustness of device type prior . . . . .	31
4.2 Temperature process . . . . .	32
4.2.1 Comparison with NOAA adjusted data . . . . .	32
4.3 Residual analysis . . . . .	35

<b>5 Conclusion</b>	<b>37</b>
<b>A Full Conditionals</b>	<b>39</b>

# List of Figures

2.1	Locations and device types of data points . . . . .	9
2.2	Summary of observations w.r.t. types of devices . . . . .	9
2.3	Number of observations available for each single device . . . . .	10
3.1	Hierarchical structure for device bias parameters . . . . .	18
4.1	Posterior densities of device type parameters $\mu$ and $\tau$ . . . . .	29
4.2	Posterior mean of densities of bias parameters $\beta$ and $\sigma$ . . . . .	30
4.3	Posterior mean device bias per location . . . . .	30
4.4	Simulated unbiased temperature process . . . . .	35
4.5	Daily estimated SST for December 2003 from NOAA adjusted temperatures	36
4.6	Map for model residual values . . . . .	36

# List of Tables

4.1	Posterior mean and confidence intervals for device type mean bias $\mu_j$ . . .	28
4.2	Posterior standard deviation confidence intervals for distribution of device type bias $\tau_j$ . . . . .	28
4.3	Posterior mean and confidence intervals for device type mean bias $\mu_j$ from different MCMC simulations . . . . .	33
4.4	Posterior standard deviation and confidence intervals for distribution of standard deviation of device type bias $\tau_j$ from different MCMC simulations	34

## **Acknowledgments**

I would like to thank my advisor Professor Bruno Sansò for his continuous support not just on this projects, but through the entire year. I would also like to thank my co-reader Professor Tatiana Xifara. In addition, I'd like to thank John Kennedy (Met Office) for the useful comments and providing the data used in this project.

# Chapter 1

## Introduction

Sea Surface Temperatures (SST) are data records on water temperatures close to the surface of the sea. SST measurements can be taken at three depth layers: SST-skin and SSTsubskin are temperatures measured at the superficial layer of water that absorbs solar radiations, within and immediately below the conductive laminar sublayer of sea surfaces. Such measurements are sensitive to solar radiation and therefore subject to large diurnal cycles. The temperature at those layers is usually recorded with remote-sensing techniques such as infrared radiometers or microwave radiometers in satellites. The surface temperatures below these superficial layers are defined as SST-depth and are normally recorded by ships or buoys.

Historical records of Sea Surface Temperatures are pivotal in understanding the dynamics of the Earth's climate, since they capture information about the changes in solar radiation and air masses from the atmosphere. Sea surface temperatures are used to define boundary conditions in general circulation models that numerically simulate slow

changes in climate.

Because of their fundamental role in climate change studies, accuracy in estimating sea surface temperatures is crucial. Previous studies (J. J. Kennedy 2011) have worked on measuring the impact of device biases due to changes in instrumentation and data availability and have found that “Uncertainties of estimated trends in global and regional-average sea surface temperature due to bias adjustments since the Second World War are found to be larger than uncertainties arising from the choice of analysis techniques, indicating that this is an important source of uncertainty in analyses of historical sea-surface temperatures.” (J. J. Kennedy 2011)

Unfortunately, the problem of estimating measurement device biases can not be considered resolved, as previous studies rarely found common agreement points, and previous literature mentions how “it is worth reflecting on the fact that the vast majority of work being done to investigate climate change is directed at sorting out the considerable confusion that arises from the numerous errors that each instrument and techniques introduces, and then applying a correction factor”. (Strangeways 2010, p. 79)

In addition to measurement errors due to the single measurement devices used, environmental factors and technology changes in measuring techniques are the main source of uncertainty in modeling SST measurement bias. In general, the studies produced so far agree that the quality of measurements has improved with time, thanks to the creation and the widespread adoption of guidelines by meteorological institutions such as the Intergovernmental Oceanography Commission (IOC). Technology improvements

are a second pivotal factor in the progressive decrease in measurement errors.

In this project, we propose a Bayesian hierarchical model for bias of devices that let the single device bias depend on a common distribution for each type of device used. A hierarchical approach has two main advantages: first, it provides a robust structure to the model when information for a single device comes from few data points. Secondly, with a hierarchical approach, we attempt to discern between the presence of a measurement bias that is attributable to characteristics of the type of device used (e.g. device material) and bias that depends on conditions attributable to the single devices.

# Chapter 2

## Review of measurement methods

### 2.1 Ships

Until the 1980s, ships have been the main resource for data on sea surface temperatures. Today, there are around 4000 ships that belong to the Voluntary Observing Fleet (VOF), and they represent around 30% of the data available worldwide (J. J. Kennedy 2011).

Each National Meteorological Service is in charge of recruiting their own ships, but the exchange of information, supply materials and instruments between foreign fleets is very active. A detailed description of the type of ship and the type of data that is collected can be found in the Marine Observer's Handbook Archive (1995).

The two main categories of devices used by ships are buckets and ERIs (Engine Room Inlet) measurements. In general, there has been a decrease over the years in the bias of these devices due to spreading of better measurement guidelines and the shift from

manually operated thermometers to digital sensors.

### 2.1.1 Buckets

The sample of ocean water is obtained by throwing a bucket attached to a rope into the sea and pulling it back when full. The measurement is mostly taken manually with a thermometer.

Buckets are among the first measurement methods used, dating back to the 17th century. At first, buckets were made with wood, and resembled little barrels. One of the first examples of standardization of the observations was made at the Brussels Marine Conference in 1853, where canvas buckets were introduced. The latest guidelines from the 1950s recommend the use of a heavier bucket made of rubber, to reach a deeper layer of water and for better insulation of the sample. However, it is estimated that the switch to rubber buckets was not completed until 1980, and the measurements preceding that year come from a mixture of canvas and rubber, depending on each ship (Strangeways 2010).

Bucket samples are subject to the most varying external conditions: from the initial state of the bucket (e.g. whether it was left in the sun or in the shade) and the consequent temperature adjustment to the water sampled, to the location of the sample (e.g. depth of water sample). One of the main factors that influence the quality of the water samples is the speed of the ship, and therefore the pressure of sea water inside the bucket (E. C. Kent 2006).

### **2.1.2 ERIs**

Engine-room intakes measure the temperature of the water that is pumped to cool the engines. ERI measurements were used since the early 1900s, but started to become the predominant method used by ships during the Second World War, because of the dangers of using lights at night (Strangeways 2010).

The point at which the measurement is taken can influence the temperature: water that travelled in pipes close to the engines will likely be heated up. On the other hand, the samples of water come from a deeper sea level than buckets and other measurements, and that factor might result in cooler samples (J. J. Kennedy 2011).

Originally, the measurements were mostly taken manually with a thermometer. More recently, electronic sensors are being increasingly used. According to Kent and Taylor (E. C. Kent 2006), there might be significant rounding of SST measurements taken manually. In addition, parallax errors might erroneously estimate the depth of the measurement.

## **2.2 Buoys**

Buoys are floating devices that consist of a plastic and steel sphere attached to a drogue. The measurements are taken by a sensor embedded at the bottom of the buoy, at a depth of around 25cm. Moored buoys provide repeated temperature measurements at a fixed location, while drifting buoys can move following ocean currents and cover a larger area.

Weather buoys were introduced in the early 1980s, progressively taking over ship measurements as main source of data. Measurements taken by drifting buoys are considered to be more consistent than moored buoys, following a standardization of their design in the 1990s (Strangeways 2010).

Larger moored buoys are used near shore, while smaller drifting buoys are used farther out at sea. Most buoys sample temperatures at shallower depths than ships, and are therefore more subject to diurnal changes. However, temperature deviations are easily detectable and adjustments can be easily proposed, since buoys take hourly measurements (E. C. Kent 2006).

Drifting buoys became more standardized in 1993, following the adoption of new guidelines by the IOC Global Ocean Observing System. Currently, drifting buoys collect the greatest proportion of SST data (J. J. Kennedy 2011).

## 2.3 Preliminary data analysis

The model presented in this project uses data on temperatures in the Mediterranean Sea region for December 2003. The data used is part of the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) provided by NOAA's National Climatic Data Center. ICOADS is a digital database that collects weather observations from ships and buoys, from 1662 to 2007. Fig 2.1 shows the locations of the data points for each type of device. Temperature measurements from ships (buckets and ERIs) cover a larger area compared to buoys.

As shown in fig 2.2a and fig 2.2b, drifting and moored buoys provide most of the data points while the most common type between all single devices is ERI. Fig 2.3 shows a breakdown of the number of data points provided by each single device.

There is a large variability in the number of observations available for each single device, with buoys being the more robust source, registering repeated observations at the same locations. Buckets and ERIs measurements are rarely able to provide repeated measurements at one location. For some devices we might register only one observation. The dataset contains some missing data. In some cases, the ID of the device is known, but its type (whether ship or buoy) is unknown. In other cases, the device ID is unknown.

From a first exploration of the data, temperatures do not seem to be influenced by the diurnal cycle, except for a few observations of unknown type.

The precision of both temperatures and spatial coordinates (latitude/longitude) is of one decimal place. Many of the ERIs measurements are subject to rounding errors to a whole degree. While it is not possible to give a precise indication on how many measurements were rounded, we estimate a minimum of 80 devices for a total of nearly 2,000 data points (out of a total of 12,210). Since we expect the order of magnitude of the estimates of device bias to be smaller than a whole degree, rounding errors might cause difficulties in detecting the presence of biases. The impact of rounding on the estimation of measurement biases is discussed in section 4.1.1.

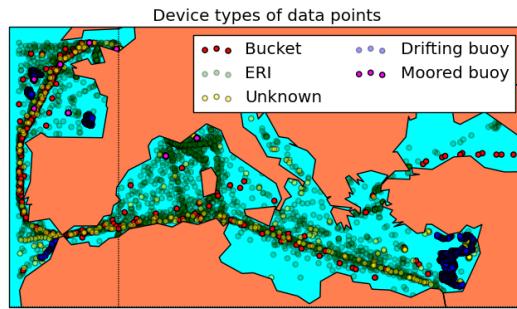


Figure 2.1: Locations and device types of data points

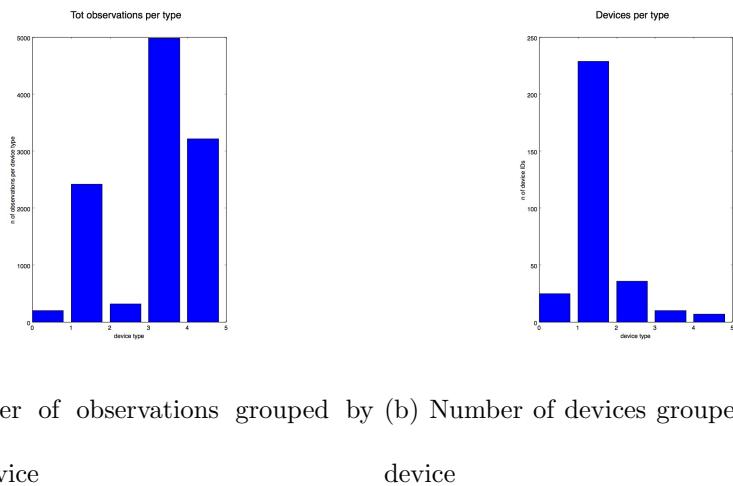


Figure 2.2: Summary of observations w.r.t. types of devices

Number of observations for each device

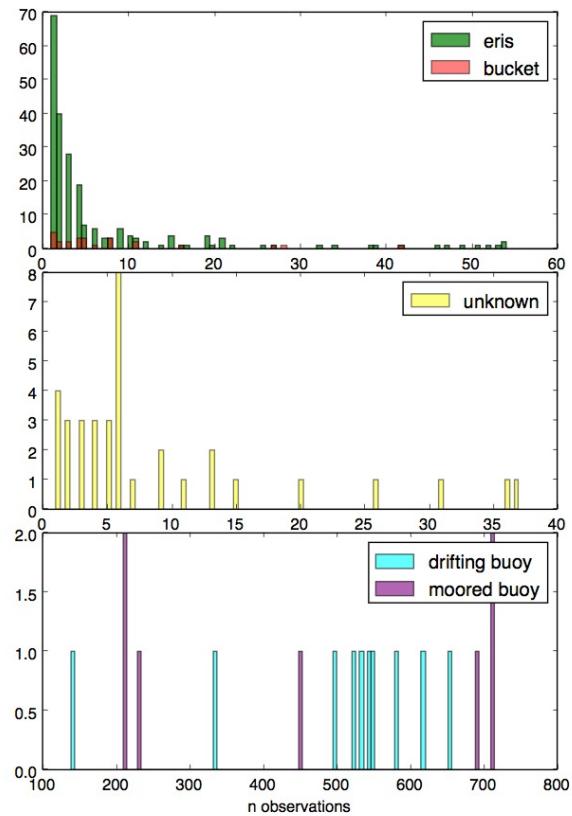


Figure 2.3: Number of observations available for each single device

While this section briefly reviewed the evolution of SST measurement methods and devices and their main characteristics, this project focuses on a precise month (December 2003) for a small geographical area (Mediterranean sea). Temperature measurements in the Mediterranean cover the majority of the sea surface area, especially because of the presence of numerous commercial ships. Thanks to a large availability of data, the Mediterranean region is therefore a good candidate for a first attempt at estimating bias of devices with the model presented in the next section.

# Chapter 3

## Model

Sea Surface Temperature data points are modeled using a linear combination of an underlying “unbiased” temperature process, and a “bias” coefficient. The temperature process is denoted with  $\theta(s_k) : [\theta(s_1), \dots, \theta(s_K)]$  and is uniquely dependent on the 2 dimensional geographical location  $s_k$  (latitude and longitude). The bias component is denoted with  $\beta_i : [\beta_1, \dots, \beta_I]$  and depends only on the device used to sample the temperature datapoint.  $\epsilon(s_k)_i : [\epsilon(s_1)_1, \dots, \epsilon(s_K)_1, \dots, \epsilon(s_K)_I]$  represents the model residuals. The observational variance  $\sigma_i^2 : [\sigma_1^2, \dots, \sigma_I^2]$  is device-dependent, but not location-dependent.

Specification:

$$\mathbf{y} = \boldsymbol{\theta}(s) + \boldsymbol{\beta} + \boldsymbol{\epsilon}, \text{ where:}$$

$$\mathbf{y} = [y_1, \dots, y_N] \quad \boldsymbol{\theta}(s) = [\theta(s_1), \dots, \theta(s_K)] \quad \boldsymbol{\beta} = [\beta_1, \dots, \beta_I]$$

$$\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$$

$n = 1, \dots, N$  ; where n represent the observation ( $N = 12,210$ )

$i = 1, \dots, I$  ; where i represents the device ( $I = 371$ )

$k = 1, \dots, K$  ; where k represents the location ( $K = 3,072$ )

Because latitude and longitude coordinates for temperature data points are rounded to one decimal point, locations are determined with one decimal degree precision. The rounding often results in identifying nearby locations with the same latitude and longitude coordinates. Because of this, slowly moving devices like drifting buoys often present repeated measurements for the same location.

More generally, in the case of buckets, ERIs, and drifting buoys, we can have measurements for the same device at numerous locations, while moored buoys generally present repeated measurements at the same location.

## 3.1 Sea Surface Temperature Process

### 3.1.1 Convolution-based Gaussian processes

A spatial Gaussian process is a stochastic process where, for any arbitrary collection of locations  $z_1, \dots, z_l$ , the joint distribution is multivariate Normal. Any

Gaussian process is completely specified by its first two moments.

A Gaussian process is stationary when:

- $\mu(x) = x$ , it has constant mean for all  $x$
- $Cov(x_i, x_j) = Cov(x_i - x_j)$ , the covariance depends only on the distance between  $x_i$  and  $x_j$ .

Spatial Gaussian processes have widespread use in geostatistics, when the available data constitutes a finite sample (or is statistically related to) of an underlying continuous phenomenon, such as sea surface temperatures. A more complete description of spatial Gaussian processes can be found in P. J. Diggle (2007).

In order to model continuous phenomena, Gaussian processes can be represented as a process convolution, by specifying a latent process  $x(s)$  and a smoothing kernel  $k(s)$  (Higdon).

In this project, we start by defining a discrete Gaussian process used to model a latent discrete temperature process  $\mathbf{z}(\mathbf{u})$ .  $\mathbf{z}(\mathbf{u})$  is a vector of normal distributions, each associated with a point  $u$  in a regularly spaced spatial grid over the Mediterranean Sea. The prior distribution for  $\mathbf{z}(\mathbf{u})$  is:

$$\mathbf{z}(\mathbf{u}) \sim MVN(\mathbf{0}, (\phi \mathbf{W})^{-1})$$

This latent non-stationary spatial process is defined with Gaussian Markov Random Fields specifications, by specifying a neighborhood system for the precision matrix  $\mathbf{W}$ .

$\mathcal{F}$  is said to be a Markov Random Field on a grid  $\mathcal{S}$  with neighbors  $\mathcal{N}$  when it satisfies the following conditions:

- positivity:  $Pr(f > 0) = 1, f \in \mathcal{F}$
- Markovianity:  $Pr(f_i|f_{\mathcal{S}[-i]}) = Pr(f_i|f_{\mathcal{N}_i})$

The precision matrix of  $\mathbf{z}$  follows a first-order, locally linear neighboring system:

$$\mathbf{W} = \begin{cases} n_{lm} & \text{if } l = m \text{ (} n_{lm} \text{ is the number of neighbors.)} \\ -1 & \text{if } l \sim m \text{ (} l \text{ is a neighbor of } m \text{)} \\ 0 & \text{otherwise} \end{cases}$$

The regularity of the process is controlled by the scale parameter  $\phi$ .

Modeling sea surfaces with MRF models allows to move from working with irregularly spaced data (the single temperature measurements taken at given latitude and longitude coordinates) to working with a regularly spaced grid of latitude/longitude locations. In addition, calculations can be further simplified by defining an optimal grid width that permits a reduction in the locations to be considered, while still maintaining a desired level of granularity of the information. On the other hand, since the locations  $k$  are irregularly spaced, we should represent the spatial process as a convolution based (kernel) model, in order to work with a continuous process.

A convolution process is determined by a latent process  $\mathbf{z}(\mathbf{u})$  and a smoothing kernel function  $Bez(k, l)$ . The chosen kernel is the Bezier kernel:

$$Bez(k, u_l, u_m, \psi, \alpha) = B(k, l) = \begin{cases} \left(1 - \left(\frac{\|k-u_l\|}{\psi\|u_l-u_m\|}\right)^2\right)^\alpha & \text{if } \frac{\|k-u_l\|}{\psi\|u_l-u_m\|} < 1 \\ 0 & \text{otherwise} \end{cases}$$

where  $u_l, u_m$  are adjacent centers in a latitude/longitude degrees grid.  $k$  is the location.  $\alpha$  and  $\psi$  are fixed.

By adjusting the kernel width  $\psi\|u_l - u_m\|$  we can obtain a smoother or more rugged representation of sea surface temperatures. Ideally, for large and heterogeneous spatial data sets, we could let  $\psi$  vary with location.

The Bezier kernel has a compact support and its computations are not very demanding. On the other hand, the kernel is isotropic and therefore might not be able to capture some of sea surface temperatures dynamics such as currents or vicinity to land. A more general kernel for DPC in a spatial grid is defined in Lemos and Sanso (R. T. Lemos 2012).

The resulting process is an intrinsic random function, where both the parameter  $\phi$  and the kernel width  $\psi\|u_l - u_m\|$  control the regularity of the process  $\mathbf{z}(\mathbf{u})$ :

$$\mathbf{x}(s) = \mathbf{B}'\mathbf{z}(\mathbf{u}) \sim MVN(\mathbf{0}, \frac{\mathbf{1}}{\phi}\mathbf{B}'\mathbf{W}^{-1}\mathbf{B})$$

An intrinsic random process is a non-stationary stochastic process with stationary increments, such that  $\Delta x(s) = x(s) - x(s - u)$  is stationary.

### 3.1.1.1 SST process

The underlying SST process for  $\theta(s)$  depends on the location of the temperature measurements and is given by the following specifications:

Let  $\theta(s_k)^* = \theta(s_k) - \bar{\theta}$ , then:

$$\theta(s_k)^* = \sum_{l=1}^L \text{Bez } (k, l) z(u_l) + v_k = \mathbf{B}_k \mathbf{z}(\mathbf{u}) + v_k$$

$v_k \sim N(0, \nu_k^2)$ ,  $s_k$  is a specific location,

$\mathbf{z}(\mathbf{u})$  is a Markov Random Field:  $\mathbf{z} = [z(u_1), \dots, z(u_L)] \sim N(\mathbf{0}, (\phi \mathbf{W})^{-1})$ , where

$$\phi \sim Gamma(a_\phi, b_\phi)$$

$l : \{1, \dots, L\}$  are centers in a lat/lng degrees grid. The centers are 1-degree distant.

## 3.2 Device bias

### 3.2.1 Hierarchical models

Modeling the bias of the single devices requires a hierarchical approach. A hierarchical model is a way to represent multiple parameters that are related, by having

$$\begin{array}{c}
\tau_j^2 \sim InvGamma(a_\tau, b_\tau) \\
\downarrow \\
\mu_j \sim N(m_j, k_0 \tau_j^2) \\
\downarrow \\
\beta_{ij} \sim (\mu_j, \tau_j^2)
\end{array}$$

Figure 3.1: Hierarchical structure for device bias parameters

them depend on the same population distribution. In models with many parameters, such structure avoids overfitting issues, especially in cases in which the ratio of number of data points over number of parameters is small. In particular, hierarchical models are a common solution in cases in which there are more parameters than data points. Hierarchical models are a straightforward expression of a Bayesian setting. The dependency structure for the device bias component in the model is given by fig 3.1.

Let the device type be denoted by  $j$ , where  $\mathbf{j} : [\text{bucket}, \text{eri}, \text{drifting buoy}, \text{moored buoy}]$ .

Then:

$$\boldsymbol{\mu} = \{\mu_1, \dots, \mu_4\}, \text{ and } \boldsymbol{\tau}^2 = \{\tau_1^2, \dots, \tau_4^2\}.$$

The bias for single devices are indicated by  $\beta_{ij}$ , and:  $\boldsymbol{\beta} = \{\beta_{11}, \dots, \beta_{I_1 1}, \dots, \beta_{I_4 4}\}$ , where  $I_1 + \dots + I_4 = I$  represent the total number of single devices that belong to  $j = 1, \dots, j = 4$  respectively.

The joint prior distribution will be as follows:

$$p(\boldsymbol{\beta}, \boldsymbol{\mu}, \tau^2) \propto p(\boldsymbol{\beta}|\boldsymbol{\mu}, \tau^2)p(\boldsymbol{\mu}|\tau^2)p(\tau^2) \quad (3.2.1)$$

The joint posterior:

$$p(\boldsymbol{\beta}, \boldsymbol{\mu}, \tau^2, \boldsymbol{\Sigma} | \mathbf{y}) \propto p(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\mu}, \tau^2, \boldsymbol{\Sigma}^2)p(\boldsymbol{\beta}, \boldsymbol{\mu}, \tau^2)p(\boldsymbol{\Sigma}^2) \quad (3.2.2)$$

where  $\boldsymbol{\Sigma}^2$  is the observational variance of  $\mathbf{y} \sim N(\boldsymbol{\beta} + \boldsymbol{\theta}, \boldsymbol{\Sigma})$ .

Hierarchical models tend to be flexible in adaptation to the data: by looking at the posterior samples, it is possible to determine whether the parameters that are lower in the hierarchy really have a dependency on their common distribution. The closer the posterior samples of the bias of the single devices are concentrated around the samples from the common distribution for the bias of the type of device, the more this dependency structure is justified.

In addition, hierarchical models are useful in incorporating pieces of information from previous literature. In particular, we let the mean of the bias for the single devices  $\beta_{ij}$  depend on a common distribution for each type  $j$  ( $\mu_1, \dots, \mu_4$ ) they belong to. The hyperparameters for the common type distribution are chosen according to previous papers (?).

A way to account for the difference between studies is to increase the prior hyperparameters for variance. The variance  $\tau_j^2$  has an inverse gamma prior distribution. We fix the expected values of the prior distribution of  $\tau_j^2$  in order to both reflect previous literature and account for extra uncertainty.

### 3.2.1.1 Hierarchical models for device bias

In the model, we let the single device biases be normally distributed with a common mean and variance according to the device type they belong to by applying a one-way normal random-effects model, with unknown variance and a normal population distribution for the group means.

The model describing a prior for device bias is:

$$p(\beta_{ij} | \mu_j, \tau_j^2) \sim N(\mu_j; \tau_j^2)$$

$$p(\mu_j | \tau_j^2) \sim N(m_j, k_0 \tau_j^2)$$

$$p(\tau_j^2) \sim IG(a_{\tau_j}, b_{\tau_j})$$

$m_j, a_{\tau_j}, b_{\tau_j}, k_0$  are fixed hyperparameters.

### 3.2.2 Mixture models for missing metadata

As mentioned in section 2.3, for some data points the type of device is not known due to missing metadata. In other words, the specific distributions of the mean  $\mu_j$  for some  $\beta_{ij}$  are unknown.

Assuming independency of the mixture components, we let the unknown-type  $\beta_{ij}$  be distributed according to a finite convex mixture of four normal distributions:

$$p(\beta_{ij} | \mu_j, \tau_j^2) = \sum_{j=1}^4 \gamma_{ij} N(\mu_j; \tau_j^2)$$

The weights of the distributions are given a Dirichlet prior:

$$p(\boldsymbol{\gamma}_i) \sim Dir(\boldsymbol{\alpha}_{\gamma})$$

We introduce a vector of latent variables  $\boldsymbol{\lambda}_i$  to allow for more efficient sampling.

Given  $\boldsymbol{\lambda}_i$ , the distribution of the single device bias  $\beta_{ij}$  is the following:

$$p(\beta_{ij} | \boldsymbol{\lambda}_i) = \prod_{j=1}^4 N(\beta_{ij} | \mu_j; \tau_j^2)^{\lambda_{ij}}$$

$\boldsymbol{\lambda}_i$  has a Multinomial distribution. Notice that only one term of the above product needs to be sampled at each iteration, because for  $\lambda_i = [\lambda_{i1}, \lambda_{i2}, \lambda_{i3}, \lambda_{i4}]$  only one component will be = 1 while the others will be 0.

The prior for  $\boldsymbol{\lambda}_i$  is given by:

$$\boldsymbol{\lambda}_i = (\lambda_{i1}, \lambda_{i2}, \lambda_{i3}, \lambda_{i4}) \sim Multiv(1; \boldsymbol{\gamma}_i)$$

### 3.3 Hierarchical model for observational variance

We let the variance  $\tau_j^2$  of the device bias  $\beta_{ij}$  depend on the device type only, while the observational variance  $\sigma_i^2$  depends on the single measurement device i in order

to account for variability due to the single devices. In some cases, only few observations are available for a single device, providing few degrees of freedom for the parameters of  $\sigma_i^2$ , causing problems in the estimability of the variance. As explained earlier, introducing a hierarchy in the parameters helps with giving a more robust structure to the posterior distributions and avoids overfitting, especially in cases in which few data points are available. We therefore let the observational variances  $\sigma_1^2, \dots, \sigma_I^2$  depend on a common parameter  $\sigma^2$  according to the following hierarchical structure:

$$\sigma^2 \sim \text{Gamma}(a_\sigma, b_\sigma)$$

$$\sigma_i^2 \sim \text{InGamma}(\alpha_\sigma, (\alpha_\sigma - 1)\sigma^2)$$

The parameters of the conditional posterior distribution for  $\sigma_i^2$  depend on both  $\sigma^2$  and the data. The role of the hyperparameter  $\alpha_\sigma$  can be inferred by looking at the marginal posterior distribution of  $(\sigma_1^2, \dots, \sigma_I^2)$ :

$$\begin{aligned} p(\sigma_1^2, \dots, \sigma_I^2) &= \int_{\sigma^2} p(\sigma^2) \prod_1^I p(\sigma_i^2 | \sigma^2) p(y_i | \sigma_i^2) d\sigma^2 \propto \\ &\int_{\sigma^2} (\sigma^2)^{a_\sigma-1} \exp(-b_\sigma \sigma^2) \\ &\prod_{i=1}^I \frac{1}{((\alpha_\sigma - 1)\sigma^2)^{\alpha_\sigma}} \frac{1}{(\sigma_i^2)^{\alpha_\sigma+1}} \exp\left(-\frac{\sigma^2(\alpha_\sigma - 1)}{\sigma_i^2}\right) \frac{1}{(\sigma_i^2)^{\frac{1}{2}}} \exp\left(-\frac{(y(s_k) - \theta(s_k) - \beta_{ij})^2}{2\sigma_i^2}\right) d\sigma^2 \propto \\ &\dots \\ &\sum_{i=1}^I \left(\frac{\alpha_\sigma - 1}{\sigma_i^2} + b\right)^{-(a+\alpha_\sigma I)} \prod_{i=1}^I \frac{1}{(\sigma_i^2)^{\alpha_\sigma + \frac{1}{2} + 1}} \exp\left(-\frac{(y(s_k) - \theta(s_k) - \beta_{ij})^2}{2\sigma_i^2}\right) \end{aligned}$$

So that for  $\alpha = 1$  the posterior marginal distribution of  $\sigma_1, \dots, \sigma_I$  is an inverse gamma with parameters:

$$\sigma_1, \dots, \sigma_I \sim \prod_{i=1}^I InvGamma \left( 1 + \frac{1}{2}; \frac{(y(s_k)_i - \beta_{ij} - \theta(s_k))^2}{2} \right)$$

### 3.4 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) is a method to draw simulations from a posterior distribution  $p(\theta|y)$ , when direct sampling is not possible (A. Gelman 2003). MCMC samples from a Markov Chain that, under fairly general conditions, is guaranteed to have an equilibrium distribution that is equal to the posterior distribution.

A sequence of  $X_1, X_2, \dots$  random variables is a Markov Chain if  $f(x_{n+1}|x_1, x_2, \dots, x_n) = f(x_{n+1}|x_n)$ . In other words, if the distribution of a random variable in a sequence, conditional on the random variables that preceded in the sequence, depends only on the last. The most common MCMC methods use Markov chains with stationary transition probabilities, where the conditional distribution of  $(x_{n+1}|x_n)$  does not depend on n. (S. Brooks 2011)

The Gibbs sampler is a particular class of MCMC algorithms and is used when the posterior distribution  $p(\theta|y)$  is defined by sampling directly one component or one sub-vector of components at a time. Suppose that, given the joint density  $f(x_1, x_2, x_3)$  we are interested in obtaining information on the marginal density:  $f(x_1) = \int \int f(x_1, x_2, x_3) dx_2, dx_3$ . In cases in which the marginal density  $f(x_1)$  is hard to obtain, the Gibbs Sampler algorithm suggests that a large enough sampling from the conditional distribution

$f(x_1|x_2, x_3)$  will provide accurate information on the characteristics of  $f(x_1)$  (G. Casella 1992).

### 3.4.1 Description of the algorithm

Let  $\Theta = \{\beta, \mu, \tau^2, \sigma^2, \theta, \nu, z, \phi\}$

Let the target distribution be:  $p(\Theta|y)$ .

The algorithm that describes the Gibbs Sampling works as follows:

```

for m in range(1, M) do
    sample  $\beta_1^{(m)} \sim p(\beta_1^{(m)} | \beta_2^{(m-1)}, \dots, \phi_n^{(m-1)})$ 
    sample  $\beta_2^{(m)} \sim p(\beta_2^{(m)} | \beta_1^{(m)}, \dots, \phi_n^{(m-1)})$ 
    ...
    sample  $\phi^{(m)} \sim p(\phi^{(m)} | \beta_1^{(m)}, \dots, z_l^{(m)})$ 
end for

```

Where M indicates the maximum number of iterations. Initial values for  $\Theta^0 = [\beta_1^0, \dots, \phi^{(0)}]$  have to be set.

Full conditionals for the parameters of the model are specified in Appendix A. The next chapter will explore the samples from the posterior distributions obtained for the model quantities. The main challenges with the implementation of this Gibbs sampler can be related to presence of few or missing data points to inform the posterior

distributions of some bias parameters. In addition, the described MCMC aims to collect samples from around 4,000 parameters, which consists in a third of the available data points.

# Chapter 4

## Model results

The Gibbs Sampler was run for 7,000 iterations, with a burn in of 400 samples and a thinning of 3, resulting in 2,200 samples for each parameter's posterior distribution. Below we analyze the samples obtained from the posterior distributions of the parameters of the model. In general, the Gibbs sampler showed almost immediate convergence of the posterior samples, except in some cases for which the posterior sample depended on few data points, or when a large slice of the data for a type of device was rounded. We discuss these cases more in detail in section 4.1.1. Furthermore, section 4.1 analyzes possible identifiability issues for the device bias component of the model.

### 4.1 Estimation of device bias

Posterior probability intervals for device type bias  $\mu$  and device type standard deviation  $\tau$  are described in table 4.1 and 4.2. Fig 4.1a and fig 4.1b show the posterior densities for  $\mu$  and  $\tau$ , Fig 4.2a and fig 4.2b show a histogram for the distribution of the

posterior means of  $\beta$  (device bias) and  $\sigma$  (observational standard deviation).

The posterior means for device type bias  $\mu_1, \dots, \mu_4$  are negative and close to 0. Buckets and ERIs show the same posterior mean bias (-0.016). Drifting buoys have the smallest posterior means, in line with previous research that states that drifting buoys are in general the most accurate measurement method. It is interesting to point out that, while all device type biases are concentrated around 0, their standard deviation  $\tau$  changes. Drifting buoys, for example, have a much smaller variability than static buoys, as shown in 4.2a and fig 4.2b. A similar comparison can be made for buckets and ERIs. From a first look at these results, we could hypothesize that the device bias is more subject to conditions dependent on the individual device (such as sensor calibration for buoys and human mistakes for ships) rather than the type of device they belong to. The next section considers another possible interpretation due to presence of rounded data.

Fig 4.3 shows the mean device bias by location of the single devices. The biases seem to be mainly influenced by the type of device prevalent in the single locations and do not seem to show any other geographical pattern, as expected.

The mixture model of normal priors  $\mu$  for  $\beta_{ij}$ , used in the case the device type  $j$  is unknown, might present identifiability issues. The posterior distributions for  $\beta_{ij}$  with  $j$  unknown present multimodality. In addition, we would expect the posterior samples

of the latent variables  $\lambda_i$  to converge and indicate what are the posterior probabilities of the device  $\beta_{ij}$  belonging to a type  $j = 1, 2, 3, 4$ . On the contrary, the chain of posterior samples presents numerous jumps.

Device type	Mean	5% probability	95% probability
Bucket	-0.016	-0.103	0.077
ERI	-0.016	-0.075	0.046
Drifting buoy	-0.004	-0.103	0.101
Moored buoy	-0.043	-0.172	0.067

Table 4.1: Posterior mean and probability intervals for device type mean bias  $\mu_j$

Device type	Mean	5% probability	95% probability
Bucket	0.024	0.016	0.037
ERI	0.004	0.003	0.004
Drifting buoy	0.030	0.018	0.050
Moored buoy	0.038	0.020	0.073

Table 4.2: Posterior standard deviation probability intervals for distribution of device type bias  $\tau_j$

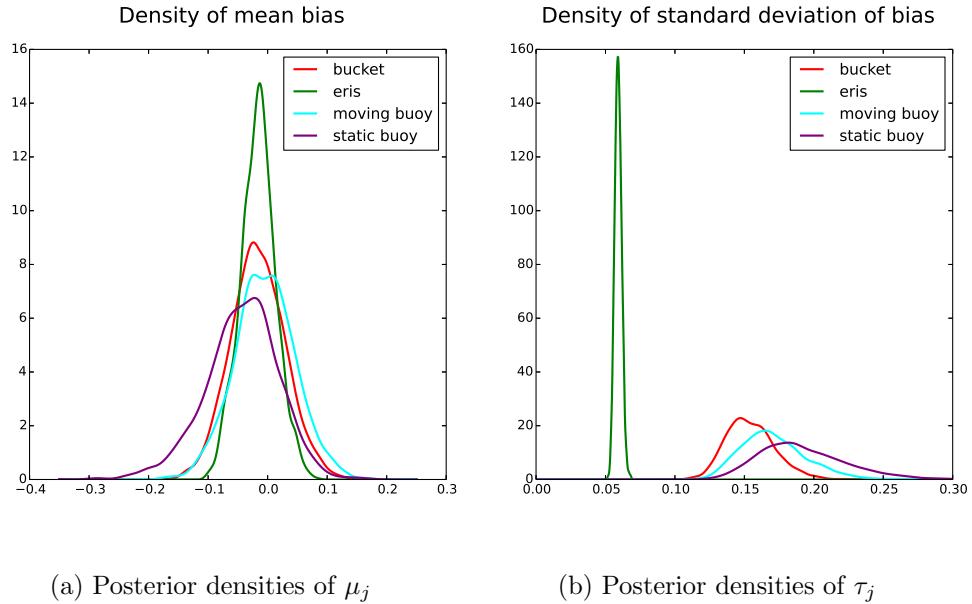


Figure 4.1: Posterior densities of device type parameters  $\mu$  and  $\tau$

#### 4.1.1 Influence of rounded data

The results for the bias of ERIs (and, less predominantly, buckets) might have been skewed towards 0 because of the presence of significant rounding of the reported temperature measurements, as mentioned in the introduction. More than half of the ERIs devices show reports of measurements that are very likely to have been rounded. The presence of data points rounded to the nearest integer might lead to two main effects that cause erroneous estimation of bias of ships:

- for a single device  $i$  in a single location  $s_k$ , repeated measurements rounded to the nearest integers cause the variance of the sample to be zero and might seriously affect the estimation of the average bias; because of a very small variance,

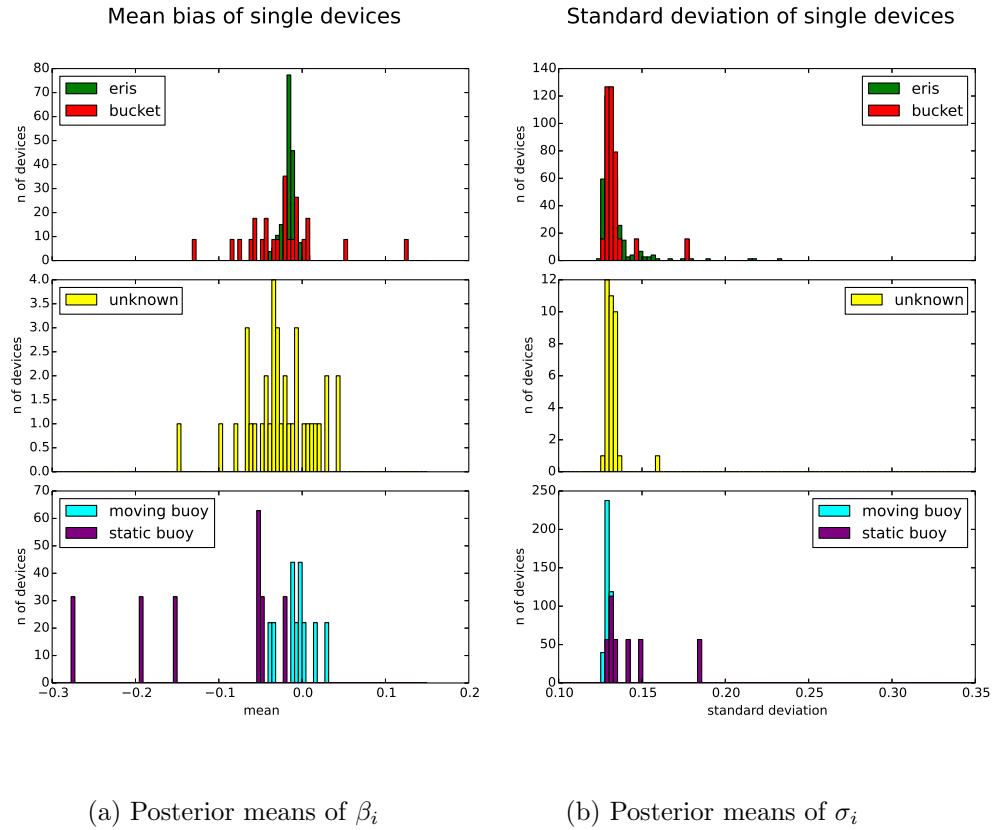


Figure 4.2: Posterior mean of densities of bias parameters  $\boldsymbol{\beta}$  and  $\boldsymbol{\sigma}$

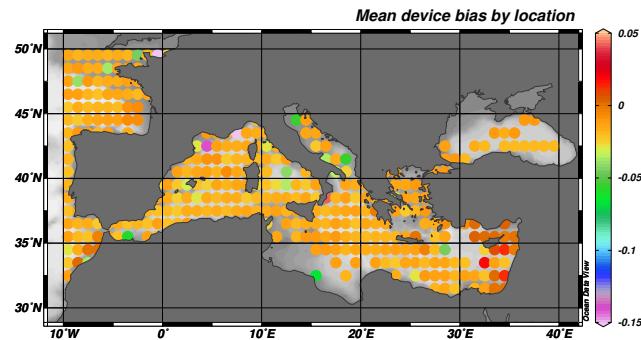


Figure 4.3: Posterior mean device bias per location

autocorrelation of the posterior samples is high;

- the rounding error has a mean of 0 and a variance larger than the variance of the device bias, therefore the rounding error outweighs the measuring bias error.

#### 4.1.2 Robustness of device type prior

We tested the Gibbs sampler for robustness with respect to device prior parameters by running the following MCMC simulations:

- $M_1$ : uninformative prior for  $\mu_1, \dots, \mu_4$ :

$$p(\mu_j | \tau_j^2) \propto 1$$

$$p(\tau_j^2) \sim IG(a_{\tau_j}, b_{\tau_j})$$

- $M_2$ : extreme values for hyper parameters:

$$p(\mu_j | \tau_j^2) \sim N(m_j, k_0 \tau_j^2)$$

$$\text{where } \mathbf{m} = [\mathbf{m}_1, \dots, \mathbf{m}_4] = [\mathbf{3}, \mathbf{3}, -\mathbf{3}, -\mathbf{3}]$$

$$p(\tau_j^2) \sim IG(a_{\tau_j}, b_{\tau_j})$$

- $M_3$ : fix the prior mean for all drifting buoys to zero  $\mu_{j=3} = 0$  and the biases of the single drifting buoys to zero  $\beta_{1(j=3)}, \dots, \beta_{i(j=3)} = 0$

Posterior probability intervals are described in Table 4.3 and 4.4.

The posterior results are not significantly influenced by modifications in the prior parameters. Choosing an uninformative prior ( $M_1$ ) leads the posterior means for buckets, ERIs

and drifting buoys  $\mu_1, \mu_2, \mu_3$  to concentrate more around 0, while  $\mu_4$  (moored buoy) is slightly more negative. The distance between buoys and ship biases is therefore slightly higher. Choosing extreme prior values ( $M_2$ ) does not lead to relevant changes in the posterior means of  $\boldsymbol{\mu}$ , but shows an increase in the confidence intervals of the posterior distributions. Wider probability intervals are expected, since extreme prior values lead to a larger exploration of the posterior space in the Gibbs sampler. Lastly, estimating the bias of the device types for bucket, ERI and moored buoy ( $M_3$ ) relatively to the drifting buoys does not provide different results, except in the case of the moored buoy type, which presents a more negative posterior mean.

## 4.2 Temperature process

Figure 4.4 shows a series of 9 simulations of the temperature process  $\boldsymbol{\theta}$  at a grid of 1000 k-locations in the Mediterranean Sea.

The values for the temperature process are obtained by sampling  $\mathbf{z}(\mathbf{u})$  from its posterior distribution. The unbiased temperature patterns appear similar among the simulations. The geographical areas with most variability (the Lybian coasts and the Black Sea) correspond to the areas for which few or no data points were available.

### 4.2.1 Comparison with NOAA adjusted data

In order to evaluate the external consistency of the SST process, we compare the results to data on SSTs curated by NOAA. The source of the data and a description of the method used to estimate SSTs can be found in et. al (2007). The data are

Model	Device type	Mean	5% probability	95% probability
$M_1$	Bucket	-0.009	-0.134	0.119
	ERI	-0.007	-0.041	0.029
	Drifting buoy	-0.001	-0.183	0.172
	Moored buoy	-0.076	-0.293	0.114
$M_2$	Bucket	-0.146	-0.422	0.108
	ERI	-0.187	-0.405	0.046
	Drifting buoy	-0.048	-0.322	0.230
	Moored buoy	-0.280	-0.797	0.134
$M_3$	Bucket	-0.007	-0.112	0.096
	ERI	-0.007	-0.042	0.032
	Drifting buoy (fixed at 0)	0	0	0
	Moored buoy	-0.055	-0.213	0.104

Table 4.3: Posterior mean and probability intervals for device type mean bias  $\mu_j$  from different MCMC simulations

Model	Device type	Mean	5% probability	95% probability
$M_1$	Bucket	0.038	0.022	0.065
	ERI	0.004	0.003	0.005
	Drifting buoy	0.072	0.035	0.136
	Moored buoy	0.075	0.030	0.164
$M_2$	Bucket	0.305	0.168	0.532
	ERI	0.288	0.151	0.477
	Drifting buoy	0.265	0.143	0.470
	Moored buoy	0.486	0.198	1.047
$M_3$	Bucket	0.034	0.020	0.055
	ERI	0.004	0.003	0.005
	Drifting buoy (fixed at 0)	0	0	0
	Moored buoy	0.065	0.026	0.131

Table 4.4: Posterior standard deviation and probability intervals for distribution of standard deviation of device type bias  $\tau_j$  from different MCMC simulations

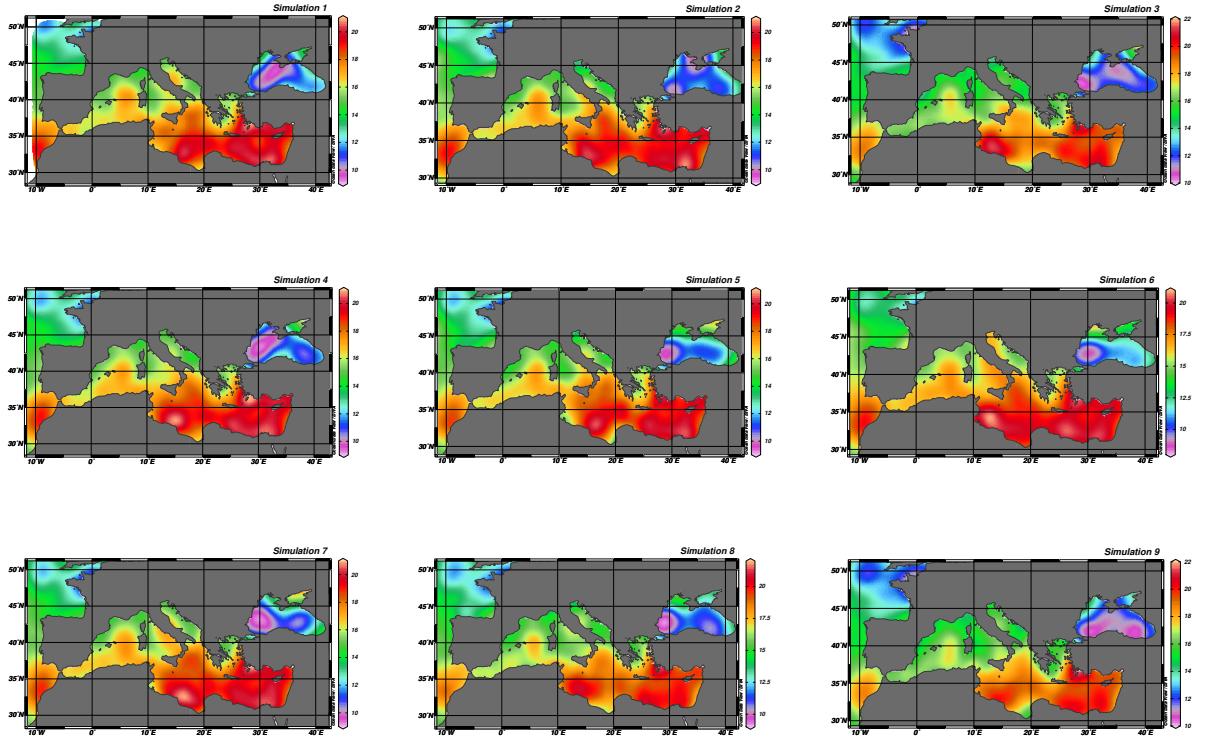


Figure 4.4: Simulated unbiased temperature process

collected from ships, buoys and satellites.

The main differences in estimated temperatures can be found on the Lybian coast and the Black Sea. As previously mentioned, there are few data points available for the Black Sea, and none for the Lybian coast.

### 4.3 Residual analysis

Model residuals  $\hat{\epsilon} = \mathbf{y} - \hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\beta}}$  are shown in fig 4.6.  $\hat{\boldsymbol{\theta}}$  and  $\hat{\boldsymbol{\beta}}$  are the posterior means of the Gibbs samples. The residuals do not show any evident pattern and are all

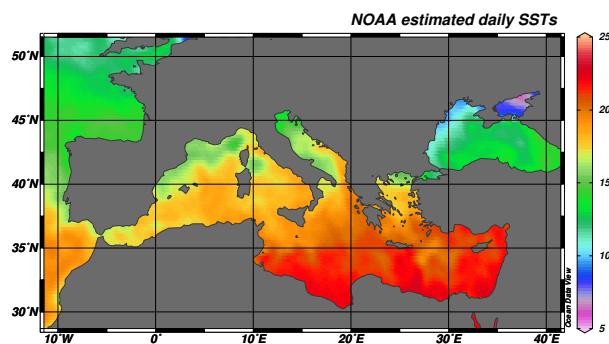


Figure 4.5: Daily estimated SST for December 2003 from NOAA adjusted temperatures

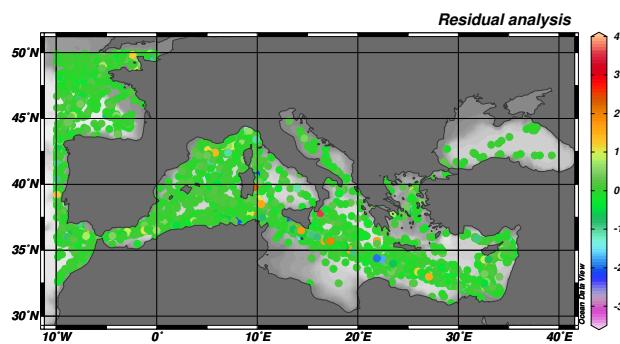


Figure 4.6: Map for model residual values

symmetrically concentrated around 0.

# Chapter 5

## Conclusion

In this project, we presented a statistical model for the determination of the biases of the devices used to measure sea surface temperatures. The model decomposes temperature data points in a spatial temperature process and a hierarchical model for bias. The hierarchical model lets devices of the same type (buckets, ERIs, moored and drifting buoys) depend on the same population mean parameter, in a one-way normal random effects for group mean model.

The distribution for the mean bias of each measurement type is negative and concentrated around 0. On the other hand, there are evident differences in the variability of the biases for the single measurement devices, with buckets and moored buoys being more variable than ERIs and drifting buoys. The results indicated that the measurement bias might be more dependent on characteristics peculiar to the single devices rather than the type of devices used.

The model bias parameters were tested for robustness by repeating the MCMC sim-

ulations changing hyperparameters and initial values. A residual analysis shows no significant patterns. The autocorrelation plots of the posterior samples do not show autocorrelation in the samples, except in the cases explained by section on rounded data 4.1.1. The results on bias from ships are not completely lined up with previous literature, that expects the bias for such typed of devices to be larger compared to buoys. Such difference in results might be due to the presence of rounded data, the size of the dataset considered, and the narrow geographical scope of the analysis.

# Appendix A

## Full Conditionals

### Device bias priors

Prior values for  $\mu_j$ :

- $p(\mu_{\text{buoy}}) \sim N(0.08; k_0 \tau_{\text{buoy}}^2)$  same prior for static and moving buoy
- $p(\mu_{\text{ship}}) \sim N(0.2; k_0 \tau_{\text{ship}}^2)$
- $p(\mu_{\text{bucket}}) \sim N(0.4; k_0 \tau_{\text{bucket}}^2)$

## Joint posterior for device bias component

$$y(s_k)_{in} = \theta(s_k) + \beta_{ij} + \epsilon(s_k)_{in}$$

$$\epsilon(s_k)_{in} \sim N(0, \sigma_i^2)$$

$i = 1, \dots, I_1, \dots, I_4$ ; where i represents the single device

$k = 1, \dots, K$ ; where  $s_k$  represents the location

$n = 1, \dots, N$ ; where n represents the single observation for device i at location k

We map the model from  $N * I * K$  dimensions to I dimensions:

$$\tilde{Y}_i = \beta_{ij} + \tilde{\epsilon}_i$$

Where  $\tilde{Y}_i$  represents the average bias for each device with variance  $\tilde{\sigma}_i^2$ .

Take:

$$D_i(k) = \{\text{observations from device } i \text{ at location } k\}$$

$$n_{ik} = \text{number of observations in } D_i(k)$$

$$K_i = \text{total number of locations touched by device } i$$

Then:

$$\tilde{Y}_i = \frac{1}{K_i} \sum_{k=1}^{K_i} \left( \frac{1}{n_{ik}} \sum_{n \in D_i(k)} (y(s_k)_i - \theta(s_k)) \right)$$

Joint posterior:

$$p(\beta_{ij}, \mu_j, \tau_j^2, \lambda_i, \tilde{\sigma}_i^2 | \tilde{Y}_i) \propto p(\tilde{Y}_i | \beta_{ij}, \tilde{\sigma}_i^2) p(\beta_{ij} | \mu_j, \tau_j, \lambda_i) p(\mu_j | \tau_j^2) p(\tau_j^2) p(\lambda_i | \gamma_i) p(\gamma_i) p(\tilde{\sigma}_i^2)$$

Full conditionals:

$$p(\beta_{ij} | \mu_1, \dots, \mu_4, \tau_1^2, \dots, \tau_4^2, \lambda_i, \gamma_i, \tilde{\sigma}_i^2) \sim \prod_{j=1}^4 N \left( \left[ \frac{\tilde{Y}_i}{\tilde{\sigma}_i^2} + \frac{\mu_j}{\tau_j^2} \right] \left( \frac{1}{\tilde{\sigma}_i^2} + \frac{1}{\tau_j^2} \right)^{-1}; \left( \frac{1}{\tilde{\sigma}_i^2} + \frac{1}{\tau_j^2} \right)^{-1} \right)^{\lambda_{ij}}$$

$$p(\lambda_{ij} = 1 | \gamma_i, \beta_{ij}, \tilde{Y}_i, \tilde{\sigma}_i^2, \mu_j^2, \tau_j^2) = \frac{\gamma_{ij} N(\beta_{ij} | \mu_j^2, \tau_j^2, \tilde{Y}_i, \tilde{\sigma}_i^2)}{\sum_{l=1}^4 \gamma_{il} N(\beta_{il} | \mu_l^2, \tau_l^2, \tilde{Y}_i, \tilde{\sigma}_i^2)}$$

$$p(\gamma_i | \lambda_i, \beta_{ij}, \tilde{Y}_i, \tilde{\sigma}_i^2, \mu_j^2, \tau_j^2) = \prod_{j=1}^4 \gamma_{ij}^{\alpha_\gamma + \lambda_{ij} - 1} \sim Dir(\alpha_\gamma + \lambda_i)$$

$$\mu_j \sim N \left( \left( \frac{\sum_{i \in j} \beta_{ij}}{\tau_j^2} + \frac{m_j}{k_0 \tau_j^2} \right) \left( \frac{n_i}{\tau_j^2} + \frac{1}{k_0 \tau_j^2} \right)^{-1}; \left( \frac{n_i}{\tau_j^2} + \frac{1}{k_0 \tau_j^2} \right)^{-1} \right)$$

$$\tau_j \sim IG \left( a_\tau + \frac{n_{ij} + 1}{2}; b_\tau + \frac{\sum_{i \in j} (\beta_{ij} - \mu_j)^2}{2} + \frac{(\mu_j - m_0)^2}{2} \right)$$

Where  $n_{ij}$  is number of devices of type j.

## Observational variance

We let  $\sigma_i^2$  depend on a common  $\sigma^2$ :

$$(\sigma^2 | a_\sigma; b_\sigma) \sim G(a_\sigma; b_\sigma)$$

$$(\sigma_i^2 | \sigma^2) \sim IG(\alpha_\sigma; (\alpha_\sigma - 1)\sigma^2)$$

Full conditionals:

$$(\sigma^2 | \tilde{\sigma}_1^2, \dots, \tilde{\sigma}_I^2) \sim G\left(a_\sigma + I\alpha_\sigma; b_\sigma + (\alpha_\sigma - 1) \sum_{i=1}^I \frac{1}{\tilde{\sigma}_i^2}\right)$$

$$(\tilde{\sigma}_i^2 | \sigma^2) \sim IG\left(\frac{1}{2} + \alpha_\sigma; \sigma(\alpha_\sigma - 1) + \frac{(\tilde{y}_i - \beta_{ij})^2}{2}\right)$$

## SST process

We map the model from  $N * I * K$  dimensions to K dimensions:

$$\tilde{Y}(s_k) = \boldsymbol{\theta}(s_k) + \boldsymbol{\varepsilon}(s_k)$$

Where  $\tilde{Y}(s_k)$  represents the average “unbiased” temperature at each location with variance  $\text{var}(\epsilon(s_k))$ .

Let:

$$D_k(i) = \{\text{observations at location } s_k \text{ for device } i\}$$

$$n_{ki} = \text{number of observations in } D_k(i)$$

$$I_k = \text{number of devices that touched location } s_k$$

Then:

$$\tilde{Y}(s_k) = \frac{1}{I_k} \sum_{i=1}^{I_k} \left( \frac{1}{n_{ki}} \sum_{j \in D_k(i)} (y_{ikn} - \beta_{ij}) \right)$$

### Model for $\theta(s_k)$

The prior for  $\theta(s_k)$  is written as:

$$\theta(s_k)^* = \theta(s_k) - \bar{\theta}$$

$$\theta(s_k)^* = \mathbf{B}_k \mathbf{z}(\mathbf{u}) + \mathbf{v}(\mathbf{s}_k)$$

$$v \sim N(0, \mathbf{V})$$

$$\text{where } V = \nu^2 I$$

where  $\mathbf{B}_k$  is a Bezier kernel vector for location  $s_k$  for the spatial grid of  $\mathbf{u}$  centers  $1 : 1, \dots, L$

Prior for  $\mathbf{z}(\mathbf{u})$ :

$$\mathbf{z} \sim N(\mathbf{0}, (\phi \mathbf{W})^{-1})$$

Structure of  $\phi \mathbf{W}$  (first order neighborhood system):

$$\mathbf{W} = \begin{cases} n_{lm} & \text{if } l = m \\ -1 & \text{if } l \sim m \\ 0 & \text{otherwise} \end{cases}$$

$n_{nm}$  is the number of neighbors.

Prior for  $\phi$ :

$$\phi \sim G(a_z; b_z)$$

Prior for  $\nu^2$ :

$$\nu^2 \sim IG(a_\nu; b_\nu)$$

Joint Posterior:

$$p(\boldsymbol{\theta}, \nu^2, \mathbf{z}, \psi | \tilde{Y}) = p(\tilde{Y} | \boldsymbol{\theta}, \nu^2) p(\boldsymbol{\theta} | \mathbf{z}, \nu^2) p(\mathbf{z} | \psi) p(\psi) p(\nu^2) =$$

Full conditionals:

$$p(\nu^2 | \boldsymbol{\theta}, \mathbf{z}, \psi, \tilde{Y}) \sim IG\left(a_\nu + \frac{K}{2}; b_\nu + \frac{(\boldsymbol{\theta} - \mathbf{B}\mathbf{z})'(\boldsymbol{\theta} - \mathbf{B}\mathbf{z})}{2}\right)$$

$$\theta(s_k) \sim N\left(\left(\frac{Y(\tilde{s}_k)}{\sigma(\tilde{s}_k)^2} + \frac{\mathbf{B}'_k \mathbf{z}}{\nu(s_k)^2}\right) \left(\frac{1}{\sigma(\tilde{s}_k)^2} + \frac{1}{\nu(s_k)^2}\right)^{-1}; \left(\frac{1}{\sigma(\tilde{s}_k)^2} + \frac{1}{\nu(s_k)^2}\right)^{-1}\right)$$

Take  $\mathbf{Z}^* = (\mathbf{B}' \mathbf{V}^{-1} \mathbf{B} + \phi \mathbf{W})^{-1} \mathbf{B}' \mathbf{V}^{-1} \tilde{\boldsymbol{\theta}}$ , then

$$\mathbf{Z} \sim N(\mathbf{Z}^*, (\mathbf{B}' \mathbf{V}^{-1} \mathbf{K} + \phi \mathbf{W})^{-1})$$

Full conditional for  $\phi$ :

$$\phi \sim G\left(a_z + \frac{L}{2}; b_z + \frac{\mathbf{z}' \mathbf{W} \mathbf{z}}{2}\right)$$

# Bibliography

H. S. Stern D. B. Rubin A. Gelman, J.B. Carlin. *Bayesian data analysis*. Chapman et Hall, 2003.

MET Office Publications Archive. Marine observer's handbook, 1995.

A. Kaplan E. C. Kent. Toward estimating climatic trends in sst. part 3: systematic biases. *Journal of Atmospheric and Oceanic Technology*, 2006.

Reynolds et. al. Daily high-resolution blended analyses. Dataset:  
(<ftp://eclipse.ncdc.noaa.gov/pub/OI-daily-v2/daily-sst.pdf>), 2007.

E. I. George G. Casella. Explaining the gibbs sampler. *Journal of the American Statistical Association*, 1992.

D. Higdon. A primer on space-time modeling from a bayesian perspective. Statistical Sciences Group, Los Alamos Laboratory.

R. O. Smith D. E. Parker M. Saunby J. J. Kennedy, N. A. Rayner. Reassessing biases and other uncertainties in sea-surface temperature observations measured in situ since 1850, part 2: biases and homogenization. *Journal of Geophysical Research*, 2011.

- P. J. Ribeiro Jr. P. J. Diggle. *Model based geostatistics*. Springer, 2007.
- B. Sanso R. T. Lemos. A spatio-temporal model for mean, anomaly, and trend fields of north atlantic sea surface temperature. *Journal of the American Statistical Association*, 2012.
- G. L. Jones X. Meng S. Brooks, A. Gelman. *Handbook of Markov Chain Monte Carlo*. Chapman et Hall, 2011.
- I. Strangeways. *Measuring Global Temperatures. Their analysis and interpretation*. Cambridge University Press, 2010.