Resolving the Antarctic contribution to sea-level rise: a hierarchical modelling framework†

Andrew Zammit-Mangion\textsuperscript{a,b,*}, Jonathan Rougier\textsuperscript{b}, Jonathan Bamber\textsuperscript{a} and Nana Schöna

Determining the Antarctic contribution to sea-level rise from observational data is a complex problem. The number of physical processes involved (such as ice dynamics and surface climate) exceeds the number of observables, some of which have very poor spatial definition. This has led, in general, to solutions that utilise strong prior assumptions or physically based deterministic models to simplify the problem. Here, we present a new approach for estimating the Antarctic contribution, which only incorporates descriptive aspects of the physically based models in the analysis and in a statistical manner. By combining physical insights with modern spatial statistical modelling techniques, we are able to provide probability distributions on all processes deemed to play a role in both the observed data and the contribution to sea-level rise. Specifically, we use stochastic partial differential equations and their relation to geostatistical fields to capture our physical understanding and employ a Gaussian Markov random field approach for efficient computation. The method, an instantiation of Bayesian hierarchical modelling, naturally incorporates uncertainty in order to reveal credible intervals on all estimated quantities. The estimated sea-level rise contribution using this approach corroborates those found using a statistically independent method. © 2013 The Authors. Environmetrics Published by John Wiley & Sons, Ltd.

Keywords: spatial statistics; stochastic partial differential equations; sea-level rise; source separation; remote sensing

1. INTRODUCTION

Larger than the conterminous USA, Antarctica is the biggest fresh water reservoir on Earth by an order of magnitude. It is estimated to contain 70% of the world’s fresh water and if it were to melt, sea level would rise by about 60 m. As a consequence, even a relatively small change in the mass balance (the sum of ice gains and losses) could have profound implications for low-lying land. Sea-level rise is considered to be one of the most serious consequences of future climate warming, with over 185 million people living in coastal areas vulnerable to a 1 m increase (Nicholls, 2011). In the last two decades, the rate of sea level has risen to nearly twice the mean value for the 20th century, and part of that increase is believed to be due to contributions from Greenland and Antarctica, with the latter having the potential for the largest future contribution and the greatest uncertainty surrounding its predicted behaviour (Bamber and Aspinall, 2013).

The Antarctic contribution to sea-level rise within a time interval is proportional to its change in ice sheet mass within that same interval. This is a straightforward relationship and is the basis of all (including the present) analyses carried out so far. However, estimating the ice sheet mass loss/gain from available data is extremely challenging. First, we need to consider several physical processes with markedly different spatio-temporal characteristics, such as flow of the Earth’s mantle, ice dynamics and precipitation. These are not directly observed as the observational data are a short recording of a linear combination of changes in height and mass caused by each of these processes. Thus, although only a subset of the processes are of direct relevance to sea-level rise, all must be considered. The problem of extracting individual components from observed linear combinations bears similarity to audio source separation (Cardoso, 1998), where sources composing a signal are separated from observed mixtures—we return to this discussion in Section 7.1. The second challenge regarding estimation is the presence of more latent components than observational data sets. Hence, the problem is, from a deterministic perspective, under-determined: even if we had noiseless data available at every desirable location, a unique solution for the latent variables does not exist.

In addition to challenges with regard to estimation, modelling the changes of the individual components is itself not straightforward. First, most processes are highly heterogeneous in nature. Precipitation, for instance, is considerably more variable at the coast than in the ice sheet interior. Second, some of the processes are highly correlated, that is, there is a linear dependence between a subset of these, which needs to be modelled. Third, the effect of these on the observed data is spatially variant. For example, the Antarctic ice sheet is surrounded by floating ice shelves that can exceed 1 km in thickness and 800 km in length and which are mainly in hydrostatic equilibrium with the ocean.

---

* Correspondence to: Andrew Zammit-Mangion, School of Geographical Sciences, University of Bristol, BS8 1SS, UK. E-mail: A.ZammitMangion@bristol.ac.uk

\textsuperscript{a} School of Geographical Sciences, University of Bristol, Bristol, BS8 1SS, UK

\textsuperscript{b} Department of Mathematics, University of Bristol, Bristol, BS8 1TW, UK

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

†This article is published in Environmetrics as part of a Special on Issue Physical-statistical modelling; guest-edited by Petra Kuhnert, CSIRO Computational Informatics, PMB 2, Glen Osmond, SA 5064 Australia.

Environmetrics 2014; 25: 245–264 © 2013 The Authors. Environmetrics Published by John Wiley & Sons, Ltd.
Environmetrics

2. THE ANTARCTIC ICE SHEET: PHYSICAL PROCESSES AND CURRENT APPROACHES

The two key processes affecting the Antarctic ice sheet mass and thickness are (i) surface processes, such as precipitation and ablation (sublimation and melt-water runoff), and (ii) ice dynamics due to internal deformation of the ice due to gravity and sliding over a lubricated bed. The latter is influenced by changes in the force balance at the ice-ocean interface, termed the grounding line, and longer-term climate variations. A shift from the balance state (i.e. the state required for zero net mass loss) in either of these processes alters the ice sheet thickness and mass. The changes are detected by satellite altimetry (height change) and gravimetry (mass change). However, there are at least three other mechanisms influencing altimetry and gravimetry in addition to surface and ice dynamics, which need to be accounted for, despite not directly affecting ice sheet mass balance. The first, glacio-isostatic adjustment (GIA), is the solid-earth vertical response to deglaciation since the last glacial maximum (20k years ago). GIA results in the flow of heavy mantle material into the unloaded region, thus significantly affecting gravimetry readings. The second, firn compaction, is the mechanism by which snow at the surface compacts to become ice (generally over decades). Because this is a compaction process (i.e. a change in density but not in mass), it is not detected by gravimetry but by altimetry. Third, there is an elastic response of the bedrock to changes in ice loading. The elastic response and firn compaction are confounded, as both are not detected by gravimetry. However, firn compaction fluctuations are on the order of centimetre/year, whilst elastic responses on the order of a few millimetre/year. Because of this difference, we assume a zero elastic response: every change in height, which does not cause a mass change, is assumed to be by firn compaction.

There are thus four processes that need to be accounted for: (i) surface processes, (ii) firn densification, (iii) ice dynamics and (iv) GIA. Whilst the former three need to be modelled to abruptly terminate at the Antarctic coastline (they all relate to ice), GIA may be taking place out at sea. Further, for mass budget estimates, it is important to assess whether the former three are within the grounding line (the point at which ice begins to float on the ocean and loses contact with bedrock), or outside it. Any melt in parts of the ice sheet, which are not grounded (i.e. floating ice shelves), does not contribute to sea-level rise. Observation modelling, discussed in Section 4.2, needs to cater for this spatial dependence. The coastline and grounding line demarcations may be seen in Figure 1 (left) for the West Antarctic Ice Sheet (WAIS) and the Antarctic Peninsula (AP). The WAIS and AP constitute about a sixth of the entire ice sheet. Because this is the first study of its kind, we will restrict our attention to this area which, however, contains the strongest contributors to sea-level rise in Antarctica (Rignot et al., 2008).
The four processes take place in the physical layers idealised in Figure 1 (right). Even with the use of a third data type (Global Positioning System (GPS) measurements of vertical bedrock motion), a deterministic consideration of all four results in an ill-posed problem (three observations, four unknowns), compelling researchers to make various prior assumptions of variable validity. In the work of Riva et al. (2009), firm compaction is assumed to play no role in observed height trends. Further, at each point in space, mass loss is assumed to be a result of changes in either ice or surface processes, not both. Both assumptions are motivated by the fact that changes due to surface processes and firm compaction are expected to be dominated by those from ice in the critical areas. In the work of Luthcke et al. (2013), GIA is assumed to be known using numerical models. The mass balance is then obtained by deducting the contribution of GIA from the gravimetry observations. This is the most direct way of estimating ice sheet mass loss, but the modelled GIA correction is uncertain and various divergent estimates exist (Horwath and Dietrich, 2009). GIA has length scales on the order of hundreds to thousands of kilometres. Hence, errors (which would be spatially correlated) of even a few millimetre/year can induce considerable changes in mass budget estimates. Finally, Wingham et al. (2006) use only altimetry to estimate the ice mass loss. GIA and firm compaction are assumed to be of negligible relevance when compared to height change by ice loss and thus ignored. A regional climate model (downscaled global climate hindcast) able to reconstruct surface processes is then used to quantify the change in surface processes. Finally, the ice thickness anomalies are obtained after subtraction of the surface height change from the altimetry observations. This approach relies heavily on an accurate assessment of the surface processes by the regional climate model.

As discussed in the introduction, the hierarchical modelling approach presented here has several advantages over those discussed earlier. There is no requirement of applying numerical model outputs to physical processes and observation models may be constructed, which explicitly relate the measurements to the processes generating them. In addition, hierarchical modelling is a fully probabilistic Bayesian framework and is thus also ideally suited for quantifying uncertainty. Uncertainty in one of the latent processes (for example, the one that is explicitly relate the measurements to the processes generating them. In addition, hierarchical modelling is a fully probabilistic Bayesian framework and is thus also ideally suited for quantifying uncertainty. Uncertainty in one of the latent processes (for example, the one that is least resolved in the data) is propagated naturally through all other processes. All these advantages are possible because hierarchical modelling is a process-based approach: whilst in the previous methods, prior process information (through numerical models or otherwise) is used to correct the data until they reflect the quantity being estimated (e.g. height change due to ice dynamics), with hierarchical modelling we start off by defining the characteristics of processes and then use the data to correct our prior belief. The investigated quantity is then obtained from our updated (posterior) understanding.

### 3. PROCESS MODEL

The height trends (in metre/year) of ice dynamics, surface and firm compaction are highly heterogeneous in space, especially along the coastline. The differing degree of heterogeneity amongst processes is key to separating the processes in the inference stage and thus must be captured in our prior model. Both (i) geostatistical (Cressie, 1993) and (ii) Gaussian Markov random field (GMRF; Rue and Held, 2005) approaches to constructing such models are easily suited to this task. The former is known to suffer from the big n problem because of the matrix inversion required for inference with a Gaussian field formulation. On the other hand, GMRFs are prone to loss of interpretability in the construction of the conditional-dependence matrix. Unlike Gaussian fields, GMRFs are, however, amenable to linear algebraic operations for fast, computationally tractable inference.

#### 3.1. Stochastic partial differential equation formulation

As shown by Lindgren et al. (2011), the advantages of the two spatial modelling approaches may be reconciled by assuming that the random fields are Gaussian processes, which are solutions to heterogeneous stochastic partial differential equations (SPDEs) of the form

$$A_z(s) := (\kappa^2(s) - \Delta)^{\alpha/2}(\tau(s)z(s)) = \mathcal{W}(s)$$

(1)
where \( s \in \mathcal{O} \subset \mathbb{R}^2 \) is the spatial domain of interest, \( z(s) \) is the height trend, \( \Delta \) is the Laplacian, \( \kappa(s) > 0 \) is a length-scale parameter, \( \tau(s) > 0 \) is a field-scaling parameter and \( W(s) \) is Gaussian spatial white noise. The exponent \( \alpha = v + d/2 \) where \( v \) is a smoothness parameter and \( d = 2 \) is the spatial dimensionality. The length-scale parameter \( \kappa(s) \) is related empirically to a range parameter \( l(s) \) corresponding to correlations of 0.1 as \( l(s) = \sqrt{80}/\kappa(s) \).

The geostatistical interpretation of the random field is preserved in the SPDE formulation through Whittle’s result (Whittle, 1954), which shows that for a spatially invariant \( (\tau, \kappa) \), \( z(s) \) is a Gaussian field with a kernel of the Matérn-type. A local interpretation for heterogeneous fields follows naturally: for smoothly varying \( \kappa(s) \), \( \tau(s) > 0 \), locally at \( \kappa(s) = \kappa^* \) and \( \tau(s) = \tau^* \), \( z(s) \) is a Gaussian field with Matérn kernel

\[
k(s, r) = \frac{\sigma^2(\tau^*, \kappa^*)}{2^{\nu-1}\Gamma(\nu)}(\kappa^*||s - r||)K_\nu(\kappa^*||s - r||)
\]

where \( \sigma^2(\tau^*, \kappa^*) \) is the marginal variance of the spatial process and \( K_\nu \) is the modified Bessel function of the second kind. The field-scaling parameter \( \tau(s) \) is related to the marginal variance of the SPDE \( \sigma_1^2(s) \) through \( \tau(s) = \sqrt{\sigma_1^2(s)/\sigma^2} \). The computational advantages of a GMRF formulation are then taken advantage of through dimensionality reduction of the SPDE, see Section 3.2.

Each of the four processes, say \( z_i(s) \), can be modelled through the SPDE formulation (1). The resulting spatial process is hence multivariate, that is, at each location \( s \), \( z(s) \in \mathbb{R}^4 \) has four components, one for each process. Known dependencies between the processes introduce cross-covariance structures, which need to be included within the model (Gneiting et al., 2010).

The multivariate spatial process

Denote the height trends due to the four processes (bedrock, ice, firm and surface) as \( z_R(s), z_I(s), z_F(s) \) and \( z_S(s) \). The processes \( z_R(s) \) and \( z_I(s) \) are independent from each other and from \( z_F(s), z_S(s) \). The latter two are, however, correlated. Correlations arise from the fact that the rate of firm compaction is highly dependent on the precipitation rate: increased precipitation will cause increased compaction (Arthern et al., 2010).

Consider, for now, two dependent Gaussian processes \( z_1(s), z_2(s) \). Define the cross-covariance matrix function of \( z(s) = [z_1(s) \ z_2(s)]^T \) as \( R(s, s') = \text{Cov}(z_1(s), z_2(s')) : \mathbb{R}^{2 \times 1} \times \mathbb{R}^{2 \times 1} \rightarrow \mathbb{R}^{2 \times 2} \). This may be constructed by modelling the joint field as a linear transform of two independent fields of unit variance \( \tilde{z}(s) = [\tilde{z}_1(s) \ \tilde{z}_2(s)]^T \), where \( \text{Cov} (\tilde{z}_1(s), \tilde{z}_2(s')) \) is diagonal, through a (possibly) spatially varying matrix transform \( \tilde{L}(s) \)

\[
z(s) = \tilde{L}(s)\tilde{z}(s)
\]

where \( \tilde{L}(s) \in \mathbb{R}^{2 \times 2} \) is the Cholesky factor of the cross-covariance matrix function \( R(s, s') \) (Finley et al., 2007). In this application, we find no reason to model a spatially variant dependence between the two processes and thus enforce spatial invariance by setting \( R(s, s') = \tilde{R} \), giving \( \tilde{L}(s) = \tilde{L} \). This is the simplest possible construction of covarying processes in what is commonly termed the coregionalisation approach (Wackernagel, 2003).

As a further simplification, suppose that \( z_1(s) \) and \( z_2(s) \) are a linear combination of two fields \( \tilde{z}_1(s) \) and \( \tilde{z}_2(s) \), which are i.i.d. Then, without loss of generality, one may fix \( z_1 = \sigma\tilde{z}_1 \) where \( \sigma^2 \) is the marginal variance of \( z_1 \) and model \( z_2(s) = a\tilde{z}_1 + b\tilde{z}_2(s) \) so that

\[
\begin{bmatrix}
  z_1(s) \\
  z_2(s)
\end{bmatrix}
= \sigma
\begin{bmatrix}
  1 & 0 \\
  a & b
\end{bmatrix}
\begin{bmatrix}
  \tilde{z}_1(s) \\
  \tilde{z}_2(s)
\end{bmatrix}
\]

(4)

The (pointwise) cross-covariance matrix \( R = \tilde{L}\tilde{L}^T \) is then given by

\[
R = \sigma^2
\begin{bmatrix}
  1 & a \\
  a & a^2 + b^2
\end{bmatrix}
\]

(5)

This special case is of interest to us as visual inspection of projected height trends in numerical models due to firm densification and surface process reveal practically identical spatial frequency characteristics. We use the similarity to justify the ‘identically distributed’ (up to some variance \( \sigma^2 \)) assumption. Let \( z_1(s) = z_S(s) \) and \( z_2(s) = z_F(s) \). An estimate for \( R \) is the empirical covariance matrix \( \tilde{R} \), which we obtain using a regional climate model (Lenaerts et al., 2012) and a firm densification model (Ligtenberg et al., 2011):

\[
\tilde{R} = \sigma^2
\begin{bmatrix}
  1 & -0.41 \\
  -0.41 & 0.24
\end{bmatrix}
\]

(6)

Equation (6) shows the negative correlation between the two processes and the smaller amplitude changes associated with trends in firm densification. The estimate \(|b| = 0.27\) (using \( \hat{a}^2 + \hat{b}^2 = 0.24 \)) is smaller than \(|a| = 0.41\) indicating that the trends due to firm processes are in fact strongly dependent on those at the surface.

System of SPDEs

The previous discussion leads us to the consideration of a system of SPDEs (Aune and Simpson, 2012). Consider once again \( \tilde{z}_1(s), \tilde{z}_2(s) \). Because these are i.i.d., they may be represented as follows:

\[
\begin{bmatrix}
  A & 0 \\
  0 & A
\end{bmatrix}
\begin{bmatrix}
  \tilde{z}_1(s) \\
  \tilde{z}_2(s)
\end{bmatrix}
= \begin{bmatrix}
  \mathcal{W}_1(s) \\
  \mathcal{W}_2(s)
\end{bmatrix}
\]

(7)
where \( \mathcal{W}_1(s), \mathcal{W}_2(s) \) are i.i.d. Without loss of generality, we assume that \( \sigma = 1 \) and that the marginal variance is embedded in \( \mathcal{A} \). Then, by (4), we have that

\[
\begin{bmatrix}
\mathcal{A} & 0 \\
0 & \mathcal{A}
\end{bmatrix} \mathbf{L}^{-1} \mathbf{z}(s) = \mathbf{W}(s)
\]

(8)

where \( \mathbf{W}(s) = [\mathcal{W}_1(s) \mathcal{W}_2(s)]^T \). Equation (8) may also be given in Kronecker form by treating \( \mathcal{A} \) as a matrix of size \( 1 \times 1 \)

\[
(\mathbf{L}^{-1} \otimes \mathcal{A}) \mathbf{z}(s) = \mathbf{W}(s)
\]

(9)

This formulation using the Kronecker product results in considerable computational simplifications, which will be taken advantage of in Section 3.2. We now construct the system governed by the problem under study. Denote the operators associated with the rock, ice, firn and interior. The same procedure was used for configuring

\[
A_r, A_i, A_f, A_s,
\]

respectively. Then the modelled random fields are the solution to the system

\[
\begin{bmatrix}
A_r & A_i \\
A_s & A_f \\
A_f & A_s \\
A_s & A_f
\end{bmatrix}
\begin{bmatrix}
z_R(s) \\
z_I(s) \\
z_S(s) \\
z_F(s)
\end{bmatrix} = \begin{bmatrix}
\mathbf{W}_R(s) \\
\mathbf{W}_I(s) \\
\mathbf{W}_S(s) \\
\mathbf{W}_F(s)
\end{bmatrix}
\]

(10)

where the sub-block

\[
\begin{bmatrix}
A_s & A_{FS} \\
A_{SF} & A_F
\end{bmatrix} = (\mathbf{L}^{-1} \otimes \mathcal{A}_s)
\]

(11)

The solutions to (10) are

\[
z_R \sim \mathcal{GP}(0, k_R(s, r)),
\]

(12a)

\[
z_I \sim \mathcal{GP}(0, k_I(s, r)),
\]

(12b)

\[
(z_S, z_F) \sim \mathcal{GP}\left(0, \hat{\mathbf{R}} \otimes k_S(s, r)\right)
\]

(12c)

where \( \mathcal{GP}(\mu(s), k(s, r)) \) denotes a Gaussian process with mean \( \mu(s) \) and covariance function \( k(s, r) \). The random fields \( (z_S, z_F) \) are jointly Gaussian with dependence defined through the cross-covariance matrix function.

**Matérn configuration**

The natural question to ask is what to choose for \( \tau(s), \kappa(s) \) in (2), or, equivalently, \( \sigma(s) \) and \( l(s) \) (the marginal deviation and range parameter, respectively), which we assume are slow varying. The parameters are needed to configure \( \mathcal{A}_i, i \in \{R, I, S\} \)—note that because of the dependence in (9), we do not need to explicitly specify \( \mathcal{A}_F \). The marginal deviation and range parameters can be themselves decomposed into smooth basis functions whose weights are subsequently estimated in a parameter estimation step. However, we find that numerical models of ice sheet processes are ideally suited to set the local properties of the covariance function. The models incorporate knowledge of geophysical characteristics such as elevation and precipitation levels, which correctly capture the spatial variability (in terms of range and field-scaling). Further, the resolution of the numerical models is generally sufficiently high to extract the local characteristics.

The first numerical model outputs we consider are the surface process anomalies \( (\mathcal{A}_S) \), calculated as the difference between the study-period mean (2003–2009) and the long-term mean (1979–2002). The long-term mean varies from 4000 mm water equivalent (mmweq) at the coast to a few hundred mmweq further inland and is thus spatially heterogeneous. Note that a 10 mmweq anomaly results in a 10 mmweq/year trend in height change. We found \( \sigma(s) \) and \( l(s) \) using the model output in conjunction with fixed-subregion analysis (Haas, 1990). We divided the domain of interest into a \( 6 \times 6 \) grid of squares with sides roughly 600 km in length and fitted zero-mean Matérn fields \((\nu = 3/2)\) to provide local estimates for \( l^* \) and \( \sigma^* \) collapsed to point estimates at the grid cells’ centres (all distances and areas quoted in a polar stereographic projection at 71° latitude). We next interpolated the points using Gaussian radial basis functions with standard deviations of 400 km (see also Higdon, 1998). The procedure for obtaining \( l(s) \) and \( \sigma(s) \) for surface process anomalies is shown in Figure 2. The result reflects our prior belief that surface processes close to the coast have a shorter wavelength and are larger in magnitude than in the drier interior. The same procedure was used for configuring \( \mathcal{A}_R \). In this case, the GIA model we employed (Simon et al., 2010) generated a field, which is sufficiently regular and homogeneous to warrant spatially invariant \( l = 3000 \) km and \( \sigma = 4.2 \) mm. The parameters \( \sigma \) and \( l \) were found by fitting once again a zero-mean Matérn to the model output.

Finally, we need to configure \( \mathcal{A}_f \). In this work, we do not make use of any numerical models, which predict change in height due to ice dynamics. Instead, we assume *a priori* a Matérn kernel with range parameter \( l = 50 \) km. This parameter is sufficiently large to accommodate regularity in large glaciers (e.g. the active Pine Island Glacier system) and sufficiently small for adjacent glaciers in the AP to be considered independent. However, we have a good indication of where mass loss, due to ice dynamics, is likely to occur by considering the ice surface velocity field obtained from satellite observations. In most cases, ice loss or gain can only occur where ice velocities are significantly high \((\geq 10 \) m/year\) (e.g. Hurkmans et al., 2012). We reflect this prior knowledge by setting the marginal standard deviation of the ice-loss trend to a sigmoid function:

\[
\sigma(s) = \frac{15}{1 + \exp(-(\nu(s) - 10))}
\]

(13)
where $v(s)$ is the ice velocity as detected by synthetic aperture radar (Rignot et al., 2011) and the saturation limit 15 m/year is a soft maximum for the ice-loss trend. One exception to this is the Kamb ice stream in the southern part of the WAIS. Here, ice accumulation is occurring at low ice velocities after the shut-down of a previously fast-flowing glacier, which occurred 150 years ago (Ng and Conway, 2004). We cater for this by setting $\sigma(s) = 0.5$ m/year in this drainage basin.

Combining physical model output and statistical models in this way was carried out for the first time, to the best of the authors’ knowledge, by Calder et al. (2011). Similar to this work, they use the numerical model output solely to elicit the dynamic characteristics of a spatio-temporal model able to describe flow direction and dispersion. Another approach to physical-statistical modelling, not used here, is to use physical models to assist dimensionality reduction (e.g. Wikle et al., 2001). Constraining modes of variation in a subset of the physical processes would also help in reducing posterior uncertainty on the marginal process distributions.

### 3.2. GMRF representation

In order to maintain a feasible computational complexity, we translate the Gaussian field representations into approximate GMRF representations (Lindgren et al., 2011). To do so, we discretise the system of SPDEs described in (10) using a basis of compact support. We construct finite elements on a triangulated domain in order to easily fit boundaries of arbitrary shapes and to accommodate a mesh density, which varies according to the spatial characteristics of the represented processes. We stress that in this case, the spatial properties are fully determined by the SPDEs and not by the triangulation/discretisation employed, which serve only to construct approximate solutions. This approach is hence different from standard GMRF modelling where the spatial characteristics are usually dependent on the specific discretisation employed.

We generated triangulations from the algorithms of Persson and Strang (2004), which maximise the minimum angles of the triangles (Delaunay triangulations). Because no heterogeneity is apparent from the GIA numerical models, we generated a regular triangulation with a minimum triangle length of 100 km. Firn densification, surface and ice processes on the other hand, as shown previously, are heterogeneous in nature, with range parameters generally reducing with proximity to the coast. We hence used the following proxy for mesh density (passed as the scaled-edge length function) for the three processes

$$f(s) = \lambda r - (\lambda - 1)rI(s)$$  \hspace{1cm} (14)

where $\lambda > 1$ penalises for whether the vertex lies within the mainland grounding line or an island ($I(s) = 1$) or not ($I(s) = 0$), and $r$ is the distance from the region of interest (the WAIS + AP grounding line delineation and neighbouring islands with area larger than 1000 km$^2$).

In this case, we chose $\lambda = 8$ and a minimum edge length of 40 km, which is sufficient to capture the majority of processes at the coast and small enough to encourage spatial regularity between neighbouring cells in the inferred ice process, see Figure 3. For all meshes, we used extended domains in order to reduce boundary effects. The total number of nodes in the mesh was $n = 7972$: 1744 for reconstructing the GIA and 2076 for the ice, firn and surface processes, respectively.

A set of basis functions on the constructed triangulation is found by setting each individual node to one and the rest to zero, for all nodes. The resulting basis functions, commonly known as tent functions, equal the field at their nodal points. Hence, for any decomposition $z(s) \approx \sum_{i=1}^{n} \phi_i(s)x_i$, for each $\phi_i$ with central node $x_i$, $z(s_i) = x_i$. Moreover, the triangular graph is sparse: using the same graph to construct a GMRF prior will thus also result in sparse precision matrices.
We follow the approach of Lindgren et al. (2011) and employ the Galerkin method by decomposing each underlying height trend using the set of basis functions

\[ z_j(s) \approx \sum_{i=1}^{n} \phi_{j,i}(s)x_{j,i} = \phi_j(s)^T x_j \] (15)

where throughout, \( j \in \{R, I, S\} \) is used to denote the rock, ice and surface fields, respectively. Each of these fields, governed by (1), is then projected through the inner product \( \langle \cdot, \cdot \rangle \) with respect to \( \phi_j(s) \) to give

\[ (\phi_j(s), (\kappa_j^2(s) - \Delta)^{\alpha_j/2} \tau_j(s)\phi_j(s)^T) x_j = (\phi_j(s), \mathcal{W}(s)) \] (16)

For \( v_j = 1, \alpha_j = 2 \), after application of Green’s Theorem and zero Neumann boundary conditions, (16) reduces to

\[ (\tilde{M}_j + \tilde{K}_j)x_j = w_j, \quad w_j \sim \mathcal{N}(0, M_j) \] (17)

where

\[ \tilde{M}_j^{(k,l)} = \int_{\mathcal{O}} \phi_{j,k}(s)\phi_{j,l}(s)\kappa_j^2(s)\tau_j(s)ds \] (18)

\[ \tilde{K}_j^{(k,l)} = \int_{\mathcal{O}} \nabla \phi_{j,k}(s)\nabla \phi_{j,l}(s)\tau_j(s)ds \] (19)

\[ M_j^{(k,l)} = \int_{\mathcal{O}} \phi_{j,k}(s)\phi_{j,l}(s)ds \] (20)

\[ K_j^{(k,l)} = \int_{\mathcal{O}} \nabla \phi_{j,k}(s)\nabla \phi_{j,l}(s)ds \] (21)

and the superscript \((i, j)\) denotes the \(i, j^{th}\) matrix element. If \( \tau_j(s) \) and \( \kappa_j(s) \) are sufficiently regular to be piecewise constant over

\[ s \in \bigcup_{k,l} (\text{supp}(\phi_{j,k}(s)), \text{supp}(\phi_{j,l}(s))) \quad \forall k, l \] (22)

then \( \tilde{M}_j \) and \( \tilde{K}_j \) reduce to scaled versions of the mass and stiffness matrices \( M_j, K_j \), which can be obtained using standard finite element techniques. Because, from (17),

\[ \text{prec}(x_j) = Q_j = (\tilde{M}_j + \tilde{K}_j)M_j^{-1}(\tilde{M}_j + \tilde{K}_j) \] (23)

which is dense for non-diagonal \( M_j \), we approximate \( M_j \) as a diagonal using a row-sum lumping method (see also Bueche et al., 2000). Note that from Lindgren et al. (2011) Section 2.3, \( v \) (and hence \( \alpha \)) are restricted to be positive integers in \( \mathbb{R}_+^2 \). For GIA trends we employ
νG = 2 (i.e. aG = 3) fields. The discretisation of this smooth (homogeneous) field can be obtained using recursive Galerkin formulations, see Lindgren et al. (2011) for further details. The reduction procedure results in a set of states for each process trend: xR for GIA, xI for ice dynamics, xF for firn compaction and xS for surface processes. The final concatenated state vector is \( x = [x^T_R, x^T_I, x^T_F, x^T_S]^T \in \mathbb{R}^n \), which is Gaussian with prior expectation \( x^p = 0 \) and prior precision matrix \( Q \in \mathbb{R}^{n \times n} \), where Q is 4×4 in block form, with all four processes, except for \( x_S \) and \( x_F \), conditionally independent a priori. From (9), it follows that the finite-dimensional representation of the correlated fields is (recall that \( \sigma = 1 \) is assumed in (5))

\[
\begin{pmatrix}
L^{-1} \otimes (M_S + \tilde{K}_S)
\end{pmatrix}
\begin{pmatrix}
x_S \\
x_F
\end{pmatrix}
= \begin{pmatrix}
w_S \\
w_F
\end{pmatrix}
\tag{24}
\]

where \( w_S \) and \( w_F \) are i.i.d. with covariance matrix \( M_S \). By repeated application of the mixed-product property of the Kronecker product, the block precision matrix is found as

\[
\text{prec} \left( \begin{pmatrix}
x_S \\
x_F
\end{pmatrix} \right) = \left( L^{-1} \otimes (M_S + \tilde{K}_S) \right)^T \left( I_S \otimes M_S^{-1} \right) \left( L^{-1} \otimes (M_S + \tilde{K}_S) \right)
= \left( L^{-T} \otimes (M_S + \tilde{K}_S) M_S^{-1} (M_S + \tilde{K}_S) \right)
= \tilde{R}^{-1} \otimes Q_S = \begin{pmatrix}
Q_S & Q_{FS} \\
Q_{SF} & Q_F
\end{pmatrix}
\tag{25}
\]

where \( I_m \) is the \( m \times m \) identity matrix. The full precision matrix is then similar in form to (10)

\[
Q = \begin{pmatrix}
Q_R & Q_I \\
Q_S & Q_{FS} \\
Q_{SF} & Q_F
\end{pmatrix}
\tag{26}
\]

and is sparse. In the presence of known parameters, computation of the posterior distribution of \( x \) follows from the Cholesky factorisation; however, the nature of the observation model, described in the next section, hinders its direct computation for large (say \( n \gtrsim 10^4 \)) problems.

4. THE OBSERVATION MODEL

We obtain spatially referenced temporal trend data using conventional trend analysis with ordinary least-squares by considering the largest time-window in which all data sets overlap: March 2003 to November 2009. With observations, the residuals are used to estimate trend errors (see, for example, Yan and Su, 2009). We make use of three data sets (summarised as temporal trends): (i) satellite altimetry, (ii) gravimetry and (iii) GPS. The fourth data set, synthetic aperture radar interferometry (InSAR), is used solely for prior construction. These are discussed in turn. Figure 4 depicts their scope in relation to the hierarchical model.

**Figure 4.** High-level information flow schematic. The altimetry, gravimetry and Global Positioning System (GPS) observations are generated (solid arrow) by the four latent processes, which contain prior information extracted from (dashed arrow) numerical models and ice velocity readings. The left and centre columns depict dependencies within the observation model, whilst the right and centre columns depict the information flow used for process model construction.
4.1. Data

4.1.1. Altimetry

Elevation changes are key to estimating trends in mass loss. Two satellite altimetry data sets are used in this work: the Ice, Cloud and Land Elevation Satellite (ICESat), which is laser-based, and the Envisat, which is radar-based. For a comparison between the two, see Connor et al. (2009). Because the footprint on the ground (i.e. the area illuminated by the satellite instrument) of ICESat is 70 m and that of Envisat is 2–3 km, the two altimetry data sets are able to provide high-resolution detail in the resolved fields. ICESat was corrected for a bias of $-6.5 \pm 2.9$ mm/year, found through a comparison to (known) mean sea-level rise. ICESat and Envisat provide sufficient detail for temporal trends in height to be estimated at a 1 km resolution. However, it is likely that small-scale effects (Cressie and Wikle, 2011; Katzfuss and Cressie, 2012) and correlated errors (Wingham et al., 1998) highly corrupt height trends obtained using altimetry at this scale. Although these characteristics can be included into the observation model, for simplicity, we follow Riva et al. (2009) and assume that they can be ignored when averaging the 1 km trends on a 20 km grid. The standard deviation of the trends within every 20 km grid cell is then used as the aggregated error.

4.1.2. Gravimetry

The Gravity Recovery and Climate Experiment records changes in the gravity field due to mass redistribution at and below the Earth surface and generates data, which may explicitly be used to quantify mass change. The data are made available in mass concentrations (mascons), which lie roughly on a 100 km polar grid. These, however, are an upsampling of a lower-frequency signal (Luthcke et al., 2013; Wahr et al., 1998). To simplify the highly correlated observation process, we averaged the signal on a polar grid (arc 300 km, radius 250 km) and further assumed that the resulting mass trends $y_G$ are a diffused image of the localised mass $y^{\ast}_{ETX}$, that is, $y_G = P y^{\ast}_{ETX}$, where

$$P(i,j) = \begin{cases} \theta, & i = j \\ 1 - (\theta i) & i \sim j \end{cases}$$

and where $\theta_i$ denotes the number of neighbours of mass concentration $i$, see Figure 5. A rough estimate for $\theta, \theta = 0.1$, was obtained by minimising $\| y_G - P(\theta) y^{\ast}_{ETX} \|$ where $y^{\ast}_{ETX}$ was simulated by using GIA and surface models and ice mass loss was estimated solely from altimetry readings. In the full inference scheme, $\theta$ is treated as an unknown parameter. Mascons were assumed to be neighbours in $P$ if their geometric centres differed by no more than 450 km under the projection employed.

Signals obtained through gravimetry are compensated for atmospheric and ocean tide anomalies using forward models, which may contain systematic errors. In addition, there are leakage-in and leakage-out errors from unmodelled sources of mass variation surrounding the ice sheet (Luthcke et al., 2013). In this work, we thus assume that the true error of the gravimetry trend is a factor $\gamma_G$ of the regression error where $\gamma_G$ is treated as an unknown.

4.1.3. Global Positioning System

Because GPS stations are placed on bedrock, GPS time series are able to provide an indication of the rate of GIA. Unfortunately, the elastic signal, caused by instantaneous changes in loading, is superimposed on that from GIA and can be larger in magnitude (Thomas et al., 2011). Unlike altimetry, the number of GPS readings is limited to a few dozen at most.

Figure 5. Left: Original gravimetry mascons. Note how these original readings are spatially smooth. Right: Aggregated mascons together with assumed neighbourhood structure of $P$ (450 km local neighbourhood)
4.1.4. Synthetic aperture radar interferometry

The InSAR (Rignot et al., 2011) is used to measure ice surface velocity. Unlike the other observations, the ice velocity map obtained from InSAR is solely used to help construct a spatially varying prior marginal variance for the ice field as seen in (13). The velocities in regions of the ice sheet not covered by the satellites or suffering high signal-to-noise ratio were replaced by balance velocities (i.e. the ice velocity required for an ice sheet in dynamic equilibrium) obtained from Bamber et al. (2000).

4.2. Observation zones

As in all geophysical applications, it is important to understand how the physical processes generate the data. A complication specific to this study is the spatial dependence of the observation process due to the distinction between the grounding line and the coastline. These spatial demarcations split up the domain of interest into observation zones (recall Figure 1).

Changes in mass of layers upstream (inland) of the grounding line (Zone 1) are detected by gravimetry. Floating ice shelves (Zone 2), which are downstream of the grounding line, are already in hydrostatic equilibrium, and melting in this zone does not contribute to sea-level rise. This spatial demarcation is important as it highlights from which regions, through the observational model, mass loss is permitted. The second demarcation, the coastline, separates the ice shelves (Zone 2) from the ocean (Zone 3). As stated previously, everything in Zone 2 is undetectable by gravimetry; however, height changes do occur and need to be incorporated in order to ensure smoothness of the surface, firn densification and ice processes across the grounding line.

These boundaries occur not only on mainland Antarctica but also on surrounding islands, all of which potentially also contribute to sea-level rise. Because grounding line/coastline data are not available other than for the mainland, we assume that all height changes within an island occur within the grounding line (Zone 1).

4.3. Model definition

Trends recorded by altimetry (averaged on a 20 km grid) are a superposition of the bedrock, ice, firn and surface height trends. For each altimeter recording \( y_{Ai} \) at spatial location \( s_i \), the observation model is

\[
y_{Ai} = z_R(s_i) + z_I(s_i) + z_S(s_i) + z_F(s_i) + e_A(s_i), \quad i \in \{1, 2, \ldots, m_A\} \tag{28}
\]

where \( m_A \) is the number of altimetry trends (on a 20 km grid) and \( e_A = [e_A(s_1) \ldots e_A(s_{m_A})] \) is a spatially uncorrelated noise vector independent of \([z_j(s_i)]_{j \in \{R, I, S, F\}}\).

The observation model for GPS differs from that of altimetry because stations are situated on bedrock. For each GPS recorded trend \( y_{GPS_i} \) at spatial location \( s_i \), the observation model is

\[
y_{GPS_i} = z_R(s_i) + e_{GPS}(s_i), \quad i \in \{1, 2, \ldots, m_{GPS}\} \tag{29}
\]

where \( m_{GPS} \) is the number of GPS stations. Equation (29) should also include the elastic response; however, this is of local scope and averages out over the large distances at which GPS stations are separated (generally several hundred kilometres). The elastic signal is thus assumed to be absorbed within the error term.

The gravimetry observations are slightly more involved as they record averages over an irregular domain with a spatial footprint of several thousands of squared kilometres. Each mascon reading is given as

\[
y^*_{Gi} = \int_{O_{Gi} \cap O_{ER}} \rho_I z_I(s) + \rho_S z_S(s) ds + \int_{O_{Gi}} \rho_R z_R(s) ds, \quad i \in \{1, 2, \ldots, m_G\} \tag{30}
\]

where \( i \) refers to the \( i^{th} \) observation, \( y_G = [y_{G1}, \ldots, y_{Gm_G}]^T \) is the vector of gravimetry mass concentration trends, \( O_{Gi} \) is the spatial footprint of the gravimetry reading, \( O_{ER} \) is the spatial domain enclosed by the grounding line, \( \rho_I \) is the density of ice (917 kg/m\(^3\)), \( \rho_S(s) \) is the density at which surface process fluctuations occur (300–600 kg/m\(^3\)), \( \rho_R \) is the effective mantle density (3800 kg/m\(^3\)). Moreover, as noted earlier, the smoothing applied in post-processing results in the averages being a diffused version of the original reading themselves. Hence, from (27), the observation model becomes

\[
y_{Gi} = P(\theta)y^*_{Gi} + e_{Gi}(s) \tag{31}
\]

The error process \( e_{Gi}(s) = [e_{Gi,1}(s) \ldots e_{Gi,m_G}(s)]^T \) is assumed to be spatially uncorrelated because of the aggregation carried out (Section 4.1). To cater for temporal filtering, we set \( e_{Gi} = y_G e^*_G \), where \( e^*_G \) is the trend standard error. Note that (30) reflects gravimetry not detecting ice and surface mass change outside the grounding line. Also, because firn densification is a compaction process, it does not contribute to any change in mass and is hence not present in the observation model.

Finally, to enforce zero height trends over the sea due to surface, firn and ice processes outside of the coastline, we add pseudo-observations at all the relevant nodes. The model for pseudo-observations is

\[
y^*_P = e_P(s_i), \quad i \in \{1, 2, \ldots, m_P\} \tag{32}
\]

where \( \text{var}(e_P(s_i)) = 10^{-6} \) introduces the confidence required on zero height trends. These pseudo-observations in Zone 3 are important to reconstruct the abrupt discontinuity present at the coastline. However, these are not critical for mass budget estimates as the demarcation is catered for during their calculation and also in the gravimetry observation process.
Following the Galerkin method on each $z_i(s)$, one can re-write the observation model in vector-matrix notation. Let $y_A = [y_A^1, \ldots, y_A^m]^{T}$ be the vector of altimetry trends, then $y_A = C_A x + e_A$ where $C_A^j$, the $j^{th}$ row of $C_A$, is given by

$$C_A^j = \left[ \phi_F^T(s_j), \phi_T^T(s_j), \phi_S^T(s_j), \phi_P^T(s_j) \right]$$

(33)

Similarly, let $y_{GPS} = [y_{GPS}^1, \ldots, y_{GPS}^m]^{T}$ be the vector of GPS trends, then $y_{GPS} = C_{GPS} x + e_{GPS}$ where $C_{GPS}^j = \left[ \phi_R^T(s_j) \right]$. Pseudo-observations are treated similarly. Let $y_P = [y_P^1, \ldots, y_P^m]^{T}$, then $y_P = C_P x + e_P$ where $C_P^j$ is zero everywhere except for the element corresponding to the node that is being softly constrained to zero; this element is set equal to 1.

For gravimetry, we first re-write (31) as follows:

$$y_G = P(\theta) \left[ B_1(\rho_I z_I(s) + \rho_S z_S(s)) + B_2(\rho_R z_R(s)) \right] + e_G(s)$$

(34)

where $B_1, B_2 \in \mathbb{R}^{mG}$ are vectors of linear operators, which compute the integral (to be numerically approximated) over the observational windows of the gravimetry observations corresponding to (30). The gravimetry observation model then reduces to $y_G = C_G(\theta)x + e_G$ where $C_G = P(\theta)D$ and

$$D^j = \left[ B_2(\rho_R \phi_R^T(s)), B_1(\rho_I \phi_I^T(s)), B_1 (\rho_S \phi_S^T(s)), e^T \right]$$

(35)

Define $y = [y_A, y_{GPS}, y_G, y_P]^T$ as the observation vector and $C(\theta) = \left[ C_A^T, C_{GPS}^T, C_G(\theta)^T, C_P^T \right]^T$ the respective observation matrix. In this problem, we have $m_A = 8857, m_{GPS} = 22, m_G = 104$ and $m_P = 1483$ so that $y \in \mathbb{R}^{10466 \times 1}$. The full observation model can then be summarised as

$$y \sim \mathcal{N}(C(\theta)x, \Sigma_{obs}(\gamma_G))$$

(36)

where $\Sigma_{obs}(\gamma_G)$ is diagonal with all elements known (up to unknown $\gamma_G^2$) from the regression residuals. Note that because altimetry, GPS and pseudo-observations can be treated as point observations on a triangulated domain, each row of $C_A$ and $C_{GPS}$ contains at most three non-zero observations per process; these observation matrices are thus sparse. On the other hand, the sub-blocks corresponding to gravimetry, $C_G$, are dense, with the effect of correlating all the states in the update step, discussed next.

5. BAYESIAN ANALYSIS

The full model, graphically shown in Figure 6 (left), is as follows:

$$y|\theta, x, \gamma_G \sim \mathcal{N}(C(\theta)x, \Sigma_{obs}(\gamma_G))$$

(37a)

$$x \sim \mathcal{N}(x_P, Q^{-1})$$

(37b)

$$\gamma_G^2 \sim IG(\alpha_{\gamma_G}, \beta_{\gamma_G})$$

(37c)

$$\theta \sim \mathcal{N}(\mu_{\theta}, \sigma_{\theta}^2)$$

(37d)

A more suitable prior distribution for $\theta$ would be a beta distribution because $0 < \alpha \ll 1$. However, the use of a normal prior distribution results in a normal conditional distribution, which can be readily sampled from in a Gibbs sampling step. Hyper-priors can form a fourth layer within the hierarchical framework; however, there is sufficient prior knowledge to treat these as fixed parameters. Following the study outlined in Section 4.1, we configured $p(\theta)$ to reflect our prior belief on the diffusion parameter of the gravimetry observations by setting

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Figure 6. Left: Graphical model for Antarctic trend data. Right: Blocking (partitioning) the state-space to aid mixing. Ice, firn and surface processes are blocked together in small spatial groups of 600 km squares (Blocks 1–4). The glacio-isostatic adjustment field (Block 5) is grouped separately.}
\end{figure}
\( \mu_q = 0.1 \) and \( \sigma_q = 0.1 \). Finally, we reflect our uncertainty on the gravimetry variance-reduction by choosing \( p(y_G^2) \) such that the 95 percentile \( q_{0.95} \approx 400 \) and the 5 percentile \( q_{0.05} \approx 1 \). Suitable values are \( \alpha_{yG} = 0.64 \) and the scale parameter \( \beta_{yG} = 3.17 \).

For a given \( y_G \) and \( \theta \), standard results (e.g. Rasmussen and Williams, 2006, Section 2.1) give

\[
\begin{align*}
\hat{Q} &= C(\theta)^T Q_{obs}(y_G) C(\theta) + \hat{Q}, \\
\hat{x} &= C(\theta)^T Q_{obs}(y_G) + \hat{Q} x_p
\end{align*}
\]

(38a) (38b)

Computation of the conditional posterior follows by finding the Cholesky factor of \( \hat{Q} \). We then find \( \hat{x} \) using standard linear algebraic routines. Computational efficiency relies on having a sparse Cholesky factor. However, for this problem, \( \hat{Q} \) is relatively dense (7% full), which, by Rue and Held (2005) Corollary 2.2, implies a Cholesky factor of equal or higher density. One can reorder \( \hat{Q} \) such that the number of non-zeros in the Cholesky factor is as low as possible. However, even with a suitable reordering, computation of the factor at each sampling step was considered prohibitive. This high density is attributed to the diffusion model for the gravimetry, which correlates all the states \( x \). For example, when no diffusion was present (\( \theta = 0 \)), the density would have been less than 2%. For this reason, it is required to break-down the problem into smaller sub-problems. We thus partitioned the state-space into blocks from which Cholesky factors with moderate number of non-zeros could be computed. This allowed for quick and efficient sampling of the latent states.

When blocking, it is important to identify components, which, as much as possible, are loosely correlated in the stationary distribution. As the ice, firn and surface processes are considerably correlated

\[
\text{Irreducibility ensures convergence of a Markov chain within a Markov chain Monte Carlo framework to the stationary distribution. However,}
\]

\[
\text{for this reason, it is required to break-down the problem into smaller sub-problems. We thus partitioned the state-space into blocks from which Cholesky factors with moderate number of non-zeros could be computed. This allowed for quick and efficient sampling of the latent states.}
\]

5.1. Initialisation

Irreducibility ensures convergence of a Markov chain within a Markov chain Monte Carlo framework to the stationary distribution. However, good starting values are able to reduce the burn-in period. This is useful when each sample is computationally expensive, as in the present case. A reasonable starting value can be obtained by computing the Cholesky factorisation of the posterior precision matrix for plug-in values of \( \theta \) and \( y_G \) in (38); however, this is not scalable with the current setup. For this reason, we approximate the Gaussian posterior using an approximate inference method.

Variational Bayes (VB) has been frequently employed in spatio-temporal modelling for fast approximate inference in applications where the likelihood is intractable (e.g. Zammit-Mangion et al., 2012). However, VB can also be effectively used to approximate a tractable posterior by a simpler posterior, which is easier to compute, by taking advantage of the graphical model of the underlying latent field. Usually, this approximate posterior is one that assumes that subsets of the latent variables are independent conditioned on the data. Denote the indices of the \( d \)-dimensional subset of \( x \), \( x' \), as \( I_d \) and define \( I = \bigcup_{d=1}^{d} I_d \). Then, for \( d \) subsets, the VB approximation is

\[
p(x | y) \approx \prod_{d=1}^{d} q_i(x_i')
\]

The partitioning scheme suggested previously is important in this case too as highly correlated blocks can lead to slow convergence and inaccurate approximate posteriors.

The approximate distribution in this problem is given as

\[
q_i(x) = \prod_{i=1}^{d} q_i(x_i') = q_1(x_R) \prod_{j=2}^{d} q_j(x_j') \cdot x_j'^i, x_j'^j, x_j'^f
\]

(39)

where \( d-1 \) is the number of spatial partitions (in this case, \( d = 22 \)). By the general theory of VB (Šmídl and Quinn, 2005), each approximate posterior is given by

\[
q_i(x_i') \propto \exp \left( \mathbb{E}_{q_i(x_i') | q_{\theta}, q_{y_G}} \ln p(y | x) \right), \quad i = 1 \ldots d,
\]

(40)

which, through conditioning, yields

\[
q_i(x_i') = N \left( \hat{x}_i', \hat{Q}_i' \right)
\]

where

\[
\hat{Q}_i' = C_i'^T Q_{obs} C_i' + Q_i'^i,
\]

(41a)

\[
\hat{x}_i' = C_i'^T Q_{obs} \left( y - \sum_{j \in [1 \ldots d] / i} C_i'^j \hat{x}_j' \right) + Q_i'^i x_p - \sum_{j \in [1 \ldots d] / i} \left( Q_i'^j (\hat{x}_j' - x_p') \right)
\]

(41b)

where \( Q_i'^j, i, j \in [1 \ldots d] \) is the matrix composed of the rows in \( Q \) with indices in \( I_i \) and the columns of \( Q \) with indices in \( I_j \), and where \( C_i' \) is the matrix composed of the columns in \( C \) with column indices in \( I_i \). For a latent Gaussian process, (41a) and (41b) are identical to...
the Gibbs sampler conditionals with the expected means replaced with a sample value. Initially, we thought that sampling could be avoided and that VB could be used to obtain final estimates. However, careful inspection of the variational posterior estimates showed over- and under-confidence in some regions, prompting us to employ Gibbs sampling instead.

5.2. Gibbs sampler convergence

We ran the VB algorithm with plug-in values $\theta = 0.1$ and $\gamma_G = 10$. The algorithm was assumed to have converged when the maximum change in the expectation of any state was less than 1 mm/year. Convergence was reached after 69 iterations and only required a few seconds using R Software (R Core Team, 2012) on a high-spec desktop computer. After setting the VB expectation as the first sample, we ran the Gibbs sampler for 5000 iterations and used the first 500 iterations as burn-in. Convergence of the samples was assessed through visual inspection and a Geweke diagnostic (Geweke, 1992) on the two parameters and a random subset of latent states (Wikle et al., 2001). The sampling procedure took 3 h to complete.

6. SUMMARY OF RESULTS

Despite the high-dimensional state-space inference is fast and produces promising results. In all figures, we use the triangulation’s dual mesh, the Voronoi tessellation, in order to convey the resolution of the modelled fields.

6.1. Gravimetry parameter estimation

The results for the parameter estimates are given in Table 1. As expected, $\theta$ is centred at a reasonably high value (0.1) which, with the given neighbourhood structure, corresponds to an average signal diffusion of 60%. This diffusion suggests a correlation length scale within the observations of several hundreds of kilometres and supports the evident smoothness apparent in Figure 5. The error amplification of $\gamma_G = 11.2$ is a clear sign that trends estimated from the mascons are indeed over-confident. It is interesting to ask whether this high value renders the observations non-informative. Treating the mascon observation $\gamma_G$ as the signal, we define the signal-to-noise ratio of each reading as $\text{SNR} = 20 \log_{10} (\gamma_G^1 / (\gamma_G \sigma_G^2))$. The gravimetry mascons have a maximum SNR of 37 dB with $\gamma_G = 1$, which drops to 16 dB at $\gamma_G = 11.2$. The latter corresponds to a reciprocal absolute coefficient of variation of 6 (i.e. the observed trend is six times its standard deviation) and can thus be considered informative.

6.2. Process estimation

In Figure 7 (left), we show the inference on the ice trend, accompanied by uncertainty estimates. Prominent features include the well-known ice-loss in the Amundsen Sea Embayment (Flament and Rémy, 2012) and ice accumulation in the Kamb ice stream (see Figure 1, left, for a map). Several isolated regions in the uncertainty map display high confidence. These correspond to regions of low ice velocity as recorded by InSAR and thus minimal height change due to ice dynamics. Height changes in these areas are attributed to surface and/or firn compaction processes, which in turn are inferred at lower uncertainty (Figure 8). In all trend maps, we show the circular region within which no altimetry readings are available. This is reflected in the ice process uncertainty map where regions of high velocity are also reported with high uncertainty.

Interestingly, considerable positive height change in the Southern AP is attributed to ice build-up, despite the generally accepted view that this region is losing mass overall (e.g. Kunz et al., 2012). Trends in surface processes close to the coast cause mass loss at relatively high densities (close to that of ice), and as a consequence, there could be some confounding between ice and surface processes in this area. However, the regional climate model output shows a negative surface anomaly in this region (Figure 2, left). Moreover, there is some recent evidence from ice cores of a positive accumulation anomaly in the AP that predates the beginning of the regional climate model time series (1980). This corroborates our result (Nield et al., 2012; Kunz et al., 2012). Unlike in the Southern AP, ice loss and gain in the Northern AP display short wavelength variability. However, an overall negative trend in this region can be observed, which will be quantified explicitly in Section 6.3. The GIA trend in Figure 7 (right) is seen to be positive overall. This corroborates expert belief that there is an overall uplift of the bedrock in the WAIS. Peaks are evident over the Filchner-Ronne and Ross ice shelves, which is a common feature in forward numerical models. The peak uplift rate is 5 mm/year, which is about 2–3 mm/year less than that obtained from models (Sasgen et al., 2012; Whitehouse et al., 2012).

Most mass loss and height change is correctly attributed to ice dynamics, despite the ill-constrained nature of the problem. In the AP, where data are sparse, surface and firn process uncertainty is high. Firn densification is seen, by construction, to correlate negatively with

| Parameter | 5 percentile | 50 percentile | 95 percentile |
|-----------|--------------|---------------|---------------|
| $\theta$  | 0.101        | 0.107         | 0.11          |
| $\gamma_G$| 9.7          | 11.2          | 12.9          |
The prior model-data balance is visualised by plotting the Kullback–Leibler divergence (KLD) between the posterior and the prior distributions at each point in space for each process. The KLD at some node \( x \in \mathbb{R} \) is defined as
\[
KL(p(x)\|p(x|y)) = \int_{-\infty}^{\infty} p(x) \ln \left( \frac{p(x)}{p(x|y)} \right) dx
\]
(42)

For computational reasons, we approximate the marginal posterior \( p(x|y) \) as a normal density from the sample histogram. The KLD between two normals can be computed analytically, see for instance Roberts and Penny (2002), Appendix B.

The KLD shows where data are significantly affecting our posterior estimates and, for this reason, is sometimes referred to as a measure of surprise (Itti and Baldi, 2006). Two representative maps are shown for the surface processes and GIA in Figure 9. A high KLD can indicate a shift in the mean, uncertainty or both. For GIA, we note that most adjustment occurs where GPS stations are relatively plentiful. The peak...
on the peninsula coincides with a GPS station, which provided an accurate trend estimate. KLDs of surface and firm processes are low over the peninsula where data are scarce. Unsurprisingly, adjustment for the surface trends is mostly present in regions of low ice velocity where detected height changes cannot be attributed to ice dynamics. High KLDs in Zone 3 are due to pseudo-observations (Section 4.3) forcing the posterior to zero beyond the coastline.

### 6.3. Mass-balance validation

We validate our approach on a sector level against estimates obtained using the input–output method (IOM, Rignot et al., 2008) in 2006, half-way through our period of study. The IOM varies considerably to the approach taken here: whilst here, we estimate trends caused by a shift from a hypothetical balance state, the IOM simply subtracts an output (ice discharge) from an input (snow accumulation) in a given year. As a result, the role the data play in estimation is different. The IOM makes use of InSAR to directly estimate output flow, whilst in the present work, InSAR is only used to construct a prior on the magnitude of allowable height loss due to ice dynamics. One would thus expect the errors between the two to be largely uncorrelated, especially when considering the range of other data sources used here. Uncertainties in the IOM were estimated by comparison of the regional climate model and the firm densification model to in situ observations, and reported errors using InSAR (which, unlike here, play an important role in the estimated ice discharge).
The results of the comparison are shown in Figure 10 and Table 2. The mean estimates are seen to mostly match within uncertainty levels. The large deviation of the means in Sector 324 is a result of the observed positive height trends being attributed to ice dynamics in the Southern AP. It is still unclear what is causing the large discrepancy in Sector 323, also apparent in the comparison by Shepherd et al. (2012), Figure 1. Our method is seen to report trends at a lower uncertainty. This, however, could be due to a number of factors. First, we are quoting uncertainty in trends, which span over 7 years (2003–2009), whilst the uncertainty of the IOM is for a single year. Averaging results obtained with the IOM over a similar time span would give lower errors (although the largest errors in the IOM are temporally correlated). Second, the model we employ still omits some sources of uncertainty concerning (i) the elastic response, (ii) the density profiles and (iii) the correlated errors within altimetry readings. Inclusion of these sources would increase our posterior uncertainty on all fields and hence also on our mass budget estimates.

### 6.4 Imbalance contribution to sea-level rise

We estimate that between 2003 and 2009, the combined mass loss rate of the WAIS and the AP was $-70 \pm 15$ Gt/year, which equates to a global sea-level rise of $0.20 \pm 0.04$ mm/year. Due to the mismatch in Sector 323, this is considerably lower than the IOM estimate of $0.52 \pm 0.18$ mm in 2006. Still, although this is a relatively short time interval and modest contribution to sea level, it represents about 10% of the mean rate for the 20th century. Importantly, the rate also appears to be accelerating (Rignot et al., 2011). This is of critical interest and importance as a large part of the WAIS is believed to be inherently unstable (Bamber et al., 2009; Schoof, 2007).
7. DISCUSSION

The major challenge in this field is the source separation problem. We briefly describe why this is the case in Section 7.1. In Section 7.2, we show how the problem features in some of the enhancements to the methods, which we are developing. Section 7.3 concludes the work.

7.1. Source separation

In signal processing, the separation of source signals from observed mixtures is known as blind source separation (e.g. Comon and Jutten, 2010). In the present case, the sources are the four physical processes, whilst the mixtures are the observed signals. Blind refers to the fact that the mixture process (analogous to the observation model in this context) is structurally unknown. Here, the observation model is structurally defined a priori; however, as we have more signals (four) than mixtures (three), the problem is under-determined.

Under-determined source separation has been recently applied in a Gaussian process context (Liu et al., 2011) for audio signal separation. Cardoso (1998) states that source separation is facilitated by spectral diversity and mixture diversity. Placing this into a BHM context, spectral diversity is manifested in prior process models having different covariance structures. For spatial processes, this translates to varying degrees of isotropy and/or length scales across the processes. The mixture diversity is reflected in the observation model, which is required to differentiate amongst the sources. Consider for example $x_1, x_2$ jointly Gaussian with zero mean. For an observation model $y = ax_1 + bx_2 + v$ where $v$ is independent from both $x_1$ and $x_2$, it can be easily shown that the posterior means $\hat{x}_1 = \mathbb{E}[x_1|y]$ and $\hat{x}_2 = \mathbb{E}[x_2|y]$ are linearly related through

$$\hat{x}_1 = \frac{a\sigma_1^2 + b\sigma_{12}}{a\sigma_{12} + b\sigma_2} x_2$$

where $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2 \end{bmatrix}$ is the posterior covariance matrix. This implies zero mixture diversity—the conditional mean of $x_1$ is fully determined by that of $x_2$ (and vice-versa), independent of the data. Although the conditional means might be separated, we term this type of separation trivial separation, as any separation is fully determined by the model specification and is data-invariant. One need not have perfect trivial separation in practice. A linear relationship between the posterior conditional means is, however, a strong indication of this effect.

Trivial separation is also possible in the spatial context. With no spectral dissociation (e.g. identical prior Matern fields), the posterior means of two Gaussian processes are perfectly linearly related under the observation model

$$y(s) = af_1(s) + bf_2(s) + e(s)$$

where $e(s)$ is spatially uncorrelated. This is clearly seen in Figure 11 where the linear relationship diminishes in the case of prior spectral dissimilarity. However, neither differing prior marginal variances nor correlations between $f_1(s)$ and $f_2(s)$ remove this linear dependence.

This scenario is of interest to us, as the firm process and surface process exhibit remarkably similar spectral properties. In fact, we equip them with identical (up to some linear correlation) process priors. Altimetry records a simple superposition of the four processes, which is hence only able to trivially separate the two processes. The separation can only be carried out using density differences through gravimetry. The diffused scope of gravimetry, however, makes this process extremely hard, and in fact, a near perfectly linear correlation is observed a posteriori between the surface and the firm conditional means. Trivial separation is not necessarily detrimental to the analysis; however, it does mean that caution should be exercised when expressing inferential statements—in this case, the function distribution of the firm process follows nearly deterministically from that of the surface process due to our prior assumed relationships between the two fields and not due to observed data. Unfortunately, the consideration of more processes, such as the elastic response, will only aggravate the source separation problem.

### Table 2. Posterior estimates from 2003 to 2009 trend data compared with the IO method in 2006

| Sector | Estimated trend (Gt/year) | IO method (Gt, 2006) |
|--------|--------------------------|----------------------|
| 301    | 8.80 ± 1.4              | −1.1 ± 12.8          |
| 318    | 17.7 ± 13.1             | 25.9 ± 6.4           |
| 319    | −1.1 ± 1.2              | 2.3 ± 3.3            |
| 320    | −11.8 ± 1.9             | −14.5 ± 39.2         |
| 321    | −48.1 ± 3.3             | −54.6 ± 21.2         |
| 322    | −27.3 ± 2.0             | −35.4 ± 16.7         |
| 323    | 2.7 ± 2.6               | −51 ± 27.0           |
| 324    | 14.2 ± 2.2              | −11.1 ± 30.0         |
| 325    | −24.8 ± 3.3             | −47 ± 9.0            |

Errors are at the 1σ level.
7.2. Framework extension

The simplest way to alleviate the source separation problem is to introduce more mixtures, that is, diverse data sets (increasing spatial coverage or accuracy of the individual instruments is less critical). In this work, we intentionally set up a framework that is able to readily incorporate data from other sources including, for example, those from ice cores (Frezzotti et al., 2012). Ice cores are obtained by extracting firm/ice samples from depths of a few tens to thousands of metres into the ice sheet, and give indications of annual snow accumulation layers over decades to millennia. Surface anomalies are relatively easily obtained from ice core data and could be treated as point observations (as with GPS), which could constrain the surface process layers. Indeed, this might be the only way to remedy trivial separability between the surface and the firn processes. In addition, spatially more extensive data relating to surface and firn processes are being generated from airborne surveys, which include the use of radar systems that can detect and track subsurface layers (isochrones) from the air. Such data are limited to relatively small sectors of the whole continent, but they are generally of high resolution and could also help further constrain and separate the latent fields, at least at a regional scale.

An important avenue for future work is the treatment of the temporal component within the process models. Spatio-temporal process priors may provide additional spectral dissimilarity required to separate the underlying physical processes. GIA, for example, is spatially and temporally smooth, whilst surface and firn processes have a high frequency temporal component (one day, for example, in the case of a snowfall event). Height loss due to ice dynamics is spatially rough but temporally rather smooth and without sub-annual variability. On the other hand, ice-shelf break-ups, such as that of Larsen B in 2002, result in sudden altimetric and gravimetric changes. These sudden change points provide a clear signature for ice dynamics. The spatio-temporal framework will have to allow for a spatio-temporal observation model, which would be achieved through the use of temporally indexed \( C_t \). At a monthly resolution, computational complexity will have to be reduced by exploiting matrix sparsity. The use of dynamic spatio-temporal models in large-scale inference (as opposed to space-time Gaussian fields) is a complex problem and an active area of research (see, for example, Hartikainen et al., 2011).

7.3. Concluding remarks

The analysis in Section 6 suggests that a statistical approach for determining the behaviour of the Antarctic ice sheet has a number of advantages over other methods that have been used in the past. The specific methodology designed for this work makes use of advanced statistical techniques that have revolutionised large-scale statistics in the last two decades. We do not claim that it is superior to those making more extensive use of numerical models, nor that the solution presented here is the final one. However, we present it as an entirely new way of approaching this challenging problem and as one that may open new doors in the analysis of ice sheet processes.

There are improvements that can be made to the current model. As mentioned in Section 6.3, the elastic response to ice loading, uncertainty in the densities \( \rho_S(s) \) and \( \rho_R \), and the correlation errors within the altimetry may be modelled. All these modifications will invariably increase the posterior uncertainty. However, the optimal model was not meant to be the focus of the work, rather to introduce a new methodology. The tools we present here are versatile enough for the consideration of most stochastic models, which, in any case, may change as our understanding of processes in the cryosphere improves with time.

Acknowledgements

This work was made possible by the NERC grant NE/I027401/1. The authors would like to thank Finn Lindgren, Botond Cseke, Ruud Hurkmans, Bert Wouters and Volker Klemann for helpful suggestions and discussions, and also Scott Luthcke, Frédérique Remy, Thomas Flament, Matt King and Liz Petrie for support throughout the project.
REFERENCES

Arthern RJ, Vaughan DG, Rankin AM, Mulvaney R, Thomas ER. 2010. In situ measurements of Antarctic snow compaction compared with predictions of models. Journal of Geophysical Research 115(F3): DOI: 10.1029/2009JF001306.

Aune E, Simpson D. 2012. The use of systems of stochastic als for priers for multivariate models with discrete structures. Technical Report Statistics no. 7/2012, Norwegian University of Science and Technology, Trondheim, Norway.

Bamber JL, Aspinall WP. 2013. An expert judgement assessment of future sea level rise from the ice sheets. Nature Climate Change 3: 424–427.

Bamber JL, Riva RE, Vermeersen BL, LeBrocq AM. 2009. Reassessment of the potential sea-level rise from a collapse of the West Antarctic Ice Sheet. Science 324(5929): 901–903.

Bamber JL, Vaughan DG, Joughin I. 2000. Widespread complex flow in the interior of the Antarctic Ice Sheet. Science 287(5456): 1248–1250.

Berliner LM, Wikle CK, Cressie N. 2000. Long-lead prediction of Pacific SSTs via Bayesian dynamic modeling. Journal of Climate 13(22): 3953–3968.

Bueche D, Sukumar N, Moran B. 2000. Dispersive properties of the natural element method. Computational Mechanics 25(2): 207–219.

Calder CA, Berrett C, Shi T, Xiao N, Munroe DK. 2011. Modeling space—time dynamics of aerosols using satellite data and atmospheric transport model output. Journal of Agricultural, Biological, and Environmental Statistics 16(4): 495–512.

Cardoso JF. 1998. Blind signal separation: statistical principles. Proceedings of the IEEE 86(10): 2009–2025.

Comon P, Jutten C. 2010. Handbook of Blind Source Separation: Independent Component Analysis and Applications. Academic press: New York.

Connor LN, Laxon SW, Ridout AL, Krabill WB, McAdoo DC. 2009. Comparison of Envisat radar and airborne laser altimeter measurements over Arctic sea ice. Remote Sensing of Environment 113(3): 563–570.

Cressie N. 1993. Statistics for Spatial Data. Wiley; New York.

Cressie N, Wikle CK. 2011. Statistics for Spatio-Temporal Data. Wiley: Hoboken, NJ.

Finley AO, Banerjee S, Carlin BP. 2007. spBayes: an R package for univariate and multivariate hierarchical point-referenced spatial models. Journal of Statistical Software 19(4): 1–24.

Flament T, Rény F. 2012. Dynamic thinning of Antarctic glaciers from along-track repeat radar altimetry. Journal of Glaciology 58(211): 830–840.

Frezzotti M, Scarchilli C, Becagli S, Proposito M, Urbani S. 2012. A synthesis of the Antarctic surface mass balance during the last eight centuries. The Cryosphere 6: 821–848.

Gelfand AE, Zhu L, Carlin BP. 2001. On the change of support problem for spatio-temporal data. Biostatistics 2(1): 31–45.

Geweke J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In Proceedings of the Fourth International Conference on Bayesian Statistics. Bernardo JM, Smith AFM, Dawid AP, Berger JO (eds). Oxford University Press: New York; 169–193.

Gneiting T, Kleiber W, Schlather M. 2010. Matérn cross-covariance functions for multivariate random fields. Journal of the American Statistical Association 105(491): 1167–1177.

Haas TC. 1990. Kriging and automated variogram modeling within a moving window. Proceedings of the Fourth International Riva REM, Gunter BC, Urban TJ, Vermeersen BLA, Lindenbergh RC, Bamber JL, van de Wal RSW, van den Broeke MR, Schutz BE. 2009. The ANTARCTIC CONTRIBUTION TO SEA-LEVEL RISE

Higdon D. 1998. A process-convolution approach to modelling temperatures in the North Atlantic Ocean. Environmental and Ecological Statistics 5(2): 173–190.

Horwath M, Dietrich R. 2009. Signal and error in mass change inferences from GRACE: the case of Antarctica. Geophysical Journal International 177(3): 849–864.

Hurkmans RTWL, Bamber JL, Sørensen LS, Joughin IR, Davis CH, Krabill WB. 2012. Spatiotemporal interpolation of elevation changes derived from satellite altimetry for Jakobshavn Isb., Greenland. Journal of Geophysical Research 117(F3): DOI: 10.1029/2011JF002072.

Itti L, Baldi P. 2006. Bayesian surprise attracts human attention. Advances in Neural Information Processing Systems (NIPS 2005). Vol. 18. Cambridge University Press: Cambridge, MA: 547–554.

Katzfuss M, Cressie N. 2012. Bayesian hierarchical spatio-temporal smoothing for very large datasets. Environmetrics 23(1): 94–107.

Kunz M, King MA, Mills P, Miller PE, Fox AJ, Vaughan DG, Marsh SH. 2012. Multi-decadal glacier surface lowering in the Antarctic Peninsula. Geophysical Research Letters 39(19): DOI: 10.1029/2012GL052823.

Lenaerts JTM, van den Broeke MR, van de Berg WJ, van Meijgaard E, Munneke PK. 2012. A new, high-resolution surface mass balance map of Antarctica (1979–2010) based on regional atmospheric climate modeling. Geophysical Research Letters 39(4): L04501.

Ligtenberg SRM, Helsen MM, van der Broeke MR. 2011. An improved semi-empirical model for the densification of Antarctic firn. The Cryosphere 5(4): 809–819.

Lindgren F, Rue H, Lindström J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society B 73(4): 423–498.

Liuksu A, Badeau R, Richard G. 2011. Gaussian processes for underdetermined source separation. IEEE Transactions on Signal Processing 59(7): 3155–3167.

Lutkemeier SB, Saba TJ, Loomis BD, Arendt AA, McCarthy JJ, Camp J. 2013. Antarctica, Greenland and Gulf of Alaska land ice evolution from an iterated GRACE global mascon solution. Journal of Glaciology 59(216): 613–631.

Ng F, Conway H. 2004. Fast-flow signature in the stagnant Kamb Ice Stream, West Antarctica. Geology 32(6): 481–484.

Nichols AJ. 2011. Planning for the impacts of sea level rise. Geology 39(2): 144–157.

Nield GA, Whitehouse PL, King MA, Clarke PJ, Bentley MJ. 2012. Increased ice loading in the Antarctic Peninsula since the 1850s and its effect on glacial isostatic adjustment. Geophysical Research Letters 39(19): L17504.

Persson P, Strang G. 2004. A simple mesh generator in MATLAB. SIAM Review 46: 329–345.

R Core Team. 2012. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria.

Rasmussen CE, Williams CKI. 2006. Gaussian Processes for Machine Learning. The MIT Press: Cambridge, MA.

Rignot E, Bamber JL, van den Broeke MR, Davis CH, Li Y, van de Berg WJ, van Meijgaard E. 2008. Recent Antarctic ice mass loss from radar interferometry and regional climate modeling. Nature Geoscience 1(2): 106–110.

Rignot E, Mouginot J, Scheuchl B. 2011. Ice flow of the Antarctic ice sheet. Science 333(6048): 1427–1430.

Riva REM, Gunter BC, Urban TJ, Vermeersen B.A.L., Lindenbergh RC, Helsen MM, Bamber JL, van de Wal RSW, van den Broeke MR, Schutz BE. 2009. Glacial isostatic adjustment over Antarctica from combined ICESat and GRACE satellite data. Earth and Planetary Science Letters 288(3–4): 516–523.

Roberts GO, Sahu SK. 2002. Updating schemes, correlation structure, blocking and parameterization for the Gibbs sampler. Journal of the Royal Statistical Society B 65(2): 291–317.

Roberts SJ, Penny WD. 2002. Variational Bayes for generalised autoregressive models. IEEE Transactions on Signal Processing 50(9): 2245–2257.

Rue H, Held L. 2005. Gaussian Markov Random Fields: Theory and Applications. Chapman Hall/CRC Press: Boca Raton, FL.

Sasgen I, Konrad H, Ivins ER, van den Broeke MR, Bamber JL, Mayewski, Z, Kleman J. 2012. Coastal Antarctic ice mass balance 2002 to 2011: regional re-analysis of GRACE satellite gravimetry measurements with improved estimate of glacial-isostatic adjustment. The Cryosphere 6: 3703–3732.

Schoof C. 2007. Ice sheet grounding line dynamics: steady states, stability, and hysteresis. Journal of Geophysical Research 112. F03S28.
Shepherd A, Shepherd A, Ivins ER, Genua A, Barletta VR, Bentley MJ, Bettadpur S, Briggs KH, Bromwich DH, Forsberg R, Galin N, Horwath M, Jacobs S, Joughin I, King MA, Lenaerts JTM, Li J, Ligtenberg SRM, Luckman A, Luthcke SB, McMillan M, Meister R, Milne G, Mouginot J, Muir A, Nicolas JP, Padin J, Payne AJ, Pritchard H, Rignot E, Rott H, Sandberg Sørensen L, Scambos TA, Scheuchl B, Schrama EJO, Smith B, Sundal AV, van Angelen JH, van de Berg WJ, van den Broeke MR, Vaughan DG, Velicogna I, Wahr J, Whitehouse PL, Wingham DJ, Yi D, Young D, Zwally HJ. 2012. A reconciled estimate of ice-sheet mass balance. Science 338(6111): 1183–1189.

Simon KM, James TS, Ivins ER. 2010. Ocean loading effects on the prediction of Antarctic glacial isostatic uplift and gravity rates. Journal of Geodesy 84(5): 305–317.

Šmídl V, Quinn A. 2005. The Variational Bayes Method in Signal Processing. Springer-Verlag: New York.

Thomas ID, King MA, Bentley MJ, Whitehouse PL, Penna NT, Williams SDP, Riva REM, Lavallee DA, Clarke PJ, King EC, Hindmarsh RCA, Koivula H. 2011. Widespread low rates of Antarctic glacial isostatic adjustment revealed by GPS observations. Geophysical Research Letters 38: L22302.

Wackernagel H. 2003. Multivariate Geostatistics. Springer-Verlag: New York.

Wahr J, Molenaar M, Bryan F. 1998. Time variability of the Earth’s gravity field: hydrological and oceanic effects and their possible detection using GRACE. Journal of Geophysical Research 103(B12): 30205–30229.

Whitehouse PL, Bentley MJ, Milne GA, King MA, Thomas ID. 2012. A new glacial isostatic adjustment model for Antarctica: calibrated and tested using observations of relative sea-level change and present-day uplift rates. Geophysical Journal International 190: 1464–1482.

Whittle P. 1954. On stationary processes in the plane. Biometrika 41: 434–449.

Wikle CK, Berliner LM. 2005. Combining information across spatial scales. Technometrics 47(1): 80–91.

Wikle CK, Milliff RF, Nychka D, Berliner LM. 2001. Spatiotemporal hierarchical bayesian modeling: tropical ocean surface winds. Journal of the American Statistical Association 96(454): 382–397.

Wingham DJ, Ridout AJ, Scharroo R, Arthern RJ, Shum CK. 1998. Antarctic elevation change from 1992 to 1996. Science 282(5388): 456–458.

Wingham DJ, Shepherd A, Muir A, Marshall GI. 2006. Mass balance of the Antarctic Ice Sheet. Philosophical Transactions of the Royal Society A 364(1844): 1627–1635.

Yan X, Su X. 2009. Linear Regression Analysis: Theory and Computing. World Scientific Publishing: Singapore.

Zammit-Mangion A, Sanguinetti G, Kadirkamanathan V. 2012. Variational estimation in spatiotemporal systems from continuous and point-process observations. IEEE Transactions on Signal Processing 60(7): 3449–3459.