# Estimating the sea-level highstand during the Last Interglacial: a probabilistic massive ensemble approach

André Düsterhus[1,*], Mark E. Tamisiea[1], Svetlana Jevrejeva[1]

[1]*National Oceanography Centre, Liverpool, United Kingdom*

[*]*Current Affiliation: Institute of Oceanography, Center for Earth*

*System Research and Sustainability (CEN), University of Hamburg, Germany*

## SUMMARY

Essential to understanding sea-level change and its causes during the last interglacial is the quantification of uncertainties. In order to estimate the uncertainties, we develop a statistical framework for the comparison of paleao-climatic sea-level index points and GIA model predictions. For the investigation of uncertainties, as well as to generate better model predictions, we implement a massive ensemble approach by applying a data assimilation scheme based on particle filter methods. The different runs are distinguished through varying ice sheet reconstructions based on oxygen-isotope curves and different parameter selections within the GIA model. This framework has several advantages over earlier work, such as the ability to examine either the contribution of individual observations to the results or the probability of specific input parameters. This exploration of input parameters and data leads to a larger range of estimates than previously published work. We illustrate how the assumptions that enter into the statistical analysis, such as the existence of outliers in the observational database or the initial ice volume history, can introduce large vari-

ations to the estimate of the maximum highstand. Thus, caution is required to avoid over-interpreting results. We conclude that there are reasonable doubts whether the datasets previously used in statistical analyses are able to tightly constrain the value of maximum highstand during the last interglacial (LIG).

**Key words:** Sea-level change – Probability Distributions – Dynamics of lithosphere and mantle

# 1    INTRODUCTION

The latest IPCC report gives a likely range (66%) for global mean sea-level rise of 0.29-0.98 m by 2100 (Church et al. 2013a). This implies a 33% probability that sea-level change may lie outside this range, which is mainly due to difficulties assessing ice loss from Greenland and Antarctic ice sheets (Church et al. 2013b; Jevrejeva et al. 2014).

To learn more about the influence of ice sheets on sea level and the associated uncertainties, analogues from the past are investigated. By better understanding sea level under changing paleao-climatic conditions, conclusions on the possible development of sea level for the future are drawn (Siddall et al. 2009; Rohling et al. 2009; Bowen 2010). While the forcings of the past differ from those today and with them possibly also the physical reasons for sea-level change (Tzedakis et al. 2009; Ganopolski & Robinson 2011), understanding the response of the system to these changes can build confidence in physical models used in future projections.

In the past decade, investigations of past sea-level change have experienced further interest due to the increased availability of proxy data and sea-level indicators. For the Holocene, where salt marshes and mangroves are key indicators of sea-level change, methods based on age-depth models and transfer functions are widely applied (Kemp et al. 2009; Barlow et al. 2014; Horton et al. 2003). In other recent work, Stanford et al. (2011) and Lambeck et al. (2014) applied Markov Chain Monte Carlo (MCMC) methods to sediment and coral records after the last glacial maximum (LGM). Using records covering longer time periods, Medina-Elizalde (2012) used MCMC methods and sampling strategies to evaluate sea-level variations since 400 kyr BP.

Applying ensemble methods to determine palaeo ice sheets was also recently performed by Briggs et al. (2014). In this case, they focused on the Antarctic ice sheet and compared GIA model predictions with indicators, such as relative sea-level indicators or grounding line retreat data. The varied parameters within this approach control the physical parametrisations

of the ice-sheet model. The analysis is performed by a model-observation misfit, which was developed in Briggs & Tarasov (2013). In another instance, a Bayesian calibration was applied to the North American ice complex (Tarasov et al. 2012).

The analysis of the relatively more plentiful data of the interglacials, particularly the last interglacial, allows us to learn more about the development of global sea level during other periods where it was close to present-day values. Interpreting these data is not as simple as assessing the average over all sites, as the Earth's deformation and gravity changes will introduce geographic and temporal differences to the relative sea level experienced at the observation locations (Lambeck et al. 2012). These regional variations can be understood by modelling glacial isostatic adjustment (GIA), which accounts for deformational and gravitational changes caused by the changing ice sheets and oceans. The GIA models provide the connection between a given ice sheet history and estimates of local sea-level change. Dutton & Lambeck (2012) completed an intensive analysis of coral data while accounting for GIA and found a global sea-level highstand of 5.5 to 9 m during the LIG compared to today.

Kopp et al. (2009) introduced a statistical approach, coupling the outputs from GIA models with sea-level indicators, to obtain a probabilistic assessment of maximum sea level during the LIG. In this analysis, they used simple GIA models to connect a set of sea-level indicators to a consistent analysis and considered the evolution of the ice sheets and the Earth's and ocean's response to these changes. To estimate the global eustatic sea level, combinations of ice sheets derived from random variations along with a randomly sampled set of Earth model parameters were used to determine the global covariance of sea-level observations. The comparison of GIA model results was done using a krigging interpolation of the observational data onto a surface. Thus, the comparison is effectively done in model space. Afterwards, a Gibbs sampler is used to estimate the values for sea level and its uncertainties. The 33$^{\text{rd}}$ and 66$^{\text{th}}$ percentiles of global mean sea level (GMSL) were estimated as 8.0 and 9.4 m.

In this paper we introduce a different statistical approach, which particularly differs with regard to the generation of uncertainty estimates. In our comparison of the model results and the data, the stated uncertainties of the data are the primary controlling factor, compared to the model covariance of the Kopp et al. (2009) study. The differences between the approaches will be addressed throughout this paper. Ideally, if the observations are free of outliers, the uncertainties are not underestimated, and the GIA model predictions, which include the ice sheet history, accurately describe all of the spatial and temporal variation in sea level, then both approaches should deliver similar answers. However, we will demonstrate that this is not the case. To be able to compare the methodologies, we will use the same datasets as the

foundation of our analysis and explain which statistical assumptions have to be made to get to a specific result.

One of the largest drivers in the comparison of the data to model results was the timing and magnitude of the eustatic sea-level change. In order to allow systematic exploration of changes to this quantity that would best describe the data, we developed a massive ensemble approach. It includes a data assimilation scheme, which will be introduced in Section 2. This scheme is driven by a comparison of the model and data based on Bayesian statistics and further modifications thereof.

Section 3 examines the impact of different approaches to the observations, statistics, and GIA modelling on the resulting estimate of the GMSL highstand during the LIG. The discussions in Section 4 will focus on the advantages and disadvantages of the proposed methodology, leading to the conclusions in Section 5.

## 2   THEORETICAL FRAMEWORK AND DATA

Ideally, the comparison of sea-level indicators to predictions of sea-level change would be straight forward if GIA models adequately represented the response of the Earth and the spatial and temporal variation of the ice sheets was well known. Unfortunately, data to constrain the detailed Earth response and ice sheet history is very limited, and the sea-level indicators themselves are a large fraction of these constraints. Thus, one approach to understanding the general behaviour of the system is to use an ensemble approach, where a very large number of "simple" models sample the possible Earth response, ice sheet history and both of their uncertainties. A simple model approach allows a much simpler access to the basic processes of the physical system than complex models. This has led in climate science to a hierarchy of models, with different tasks for different types of models (Petersen 2000; Shackley et al. 1998). The statistical evaluation of simple models plays a very important role, as it is necessary to interpret the representativeness of the simplified model for the physical system (Zwiers & Von Storch 2004). As a consequence, approaches based on simple models are good candidates for the application of massive ensembles. The analysis of these simple models relies heavily on the chosen statistical framework and the ability to generate a large number of predictions of the sea-level variations, sampling a very wide range of model parameters.

The majority of this section details the statistical framework, as its approach is new to this problem (Section 2.2). This framework allows for comparison of the model results to the sea-level indicators and, through extension, data assimilation to estimate the probability distribution for global mean sea level. To facilitate comparisons with earlier work, we use the

sea-level indicator data set of Kopp et al. (2009) (see Section 2.1). The sea-level predictions will require a sampling of the ice sheet history (Section 2.3) and a description of the Earth and oceans response to these changes (i.e. the GIA model, Section 2.4). A simplified illustration of the framework tying the various components together is given in Section 2.5.

## 2.1    Observations

We use the dataset of Kopp et al. (2009). It includes 108 sea-level indicators at 47 sites that can be divided into two groups. The first is 30 observations from the Red Sea, dated via oxygen isotope correlation, the second is a mix of different indicators, e.g. corals, reef terraces and shoreline deposits. For each observation, the relative sea level and age is given with a Gaussian uncertainty. The database also has some additional information available, which we use in the analysis. We correct each observation with the given linear tectonic uplift or subsidence rates and their associated Gaussian uncertainty. For some observations stratigraphic sequence information is available, which we use in a sampling procedure of the observations, described in Section 2.2.2. We also use the information on limiting sea-level observations, which are given for 12 indicators. These indicate that sea level should be higher or lower than a given height.

## 2.2    Data assimilation

### 2.2.1    *General principle*

Our data assimilation scheme creates a large ensemble of GIA predictions by sequentially modifying the input parameters of the ice volume history to create a series of runs with good agreement between the sea-level indicators and the GIA model predictions. Our approach is based on a modified version of the sequential importance resampling filter (SIRF, van Leeuwen 2003). Each ensemble member, or run, is a GIA prediction for a specific ice volume history and Earth model and is evaluated by its fit to the observations (Section 2.2.2). An illustration of the applied scheme is shown in Fig. 1 and will be explained stepwise in the following. For a better understanding, Section 2.5 demonstrates a simplified example of the data assimilation scheme.

To start the process, we choose an Earth model and create a set of ice histories (size $N_s = 100$, see Section 2.3.3) that spans from some time in the past to present day. Given that the Earth's response to past and future (relative to the LIG, see Section 2.4) ice sheet changes affect observations of sea level in the LIG, the time period ($t_e$) must be larger than our period

of interest. Therefore, we start our modelling at the prior interglacial, 214 kyr BP. The GIA calculations are performed at discrete time points, indicated by the black stars in Fig. 2, with resolution varying between 0.5 kyr during the LIG and 2 kyr in the glaciation phases. The GIA model predictions are then statistically evaluated on the basis of the difference between the observations and the linearly-interpolated (in time and space) model results. Thus, a probability is assigned to each GIA model prediction, and therefore to each ice history. From this set, we identify the $N_k = 10$ best performing ensemble members.

To allow the ice sheets to adapt to the observations in the next round (see the box labelled "round" in Fig. 1), a new set of ice sheets is created (see Section 2.3.3). In contrast to the classical SIRF by van Leeuwen (2003), a fixed number of runs $N_r = 10$ is created from each of the $N_k$ ensemble members. With the choice of $N_s = N_k \cdot N_r$, the number of ensemble members calculated for each round is constant. Each of these new ensemble members shares the same ice history with its parent up until a given temporal analysis point. The time span $t_e$ is divided into $N_a = 64$ analysis points, illustrated by the vertical lines in Fig. 2, and for each of them a separate round is calculated in chronological order, starting with the earliest. Their spacing varies over time to allow a good temporal resolution of ice sheet changes during the LIG. For times between 170 kyr to 128 kyr BP and 118 kyr to 70 kyr BP, an analysis point occurs every 2 kyr. Between 127 and 119 kyr BP, a spacing of 0.5 kyr is used.

### 2.2.2 Statistical evaluation

The aim of this sub-section is to compare the results of a deterministic model ($\mathcal{M}$, the GIA model) with uncertain observations ($\vec{O}$, the sea-level indicators) to generate information on the model input parameters ($\vec{\eta}$). $\vec{\eta}$ can be the ice volume history, the Earth model parameters, such as lithospheric thickness or mantle viscosity, or any other specification defining the outcome of the GIA model prediction. To achieve this comparison, we use Bayesian statistics (Bayes 1763), which allows one to include prior knowledge into the analysis and can be expressed by the basic equation:

$$p\left(m|o\right) = \frac{p\left(o|m\right)p\left(m\right)}{p\left(o\right)}. \tag{1}$$

This equation states that the probability $p$ of the model $m$ given the observation $o$ is equal to the probability of the observation given the model, weighted by the model to observation ratio of prior probabilities.

The aim of the following analysis is to determine how probable each parameter set ($\vec{\eta_i}$, where the subscript $i$ indicates a particular parameter set from all those considered) is for a

given set of observations $\vec{O}$. Because the GIA model effectively connects both, this probability can be expanded as

$$p\left(\vec{\eta}_i|\vec{O}\right) = \int_{j\in\mathcal{J}} p\left(\vec{f}_j, \vec{\eta}_i|\vec{O}\right) d\vec{f}_j, \tag{2}$$

where $\vec{f}_j$ is one of possible solutions of the model. Allowing for the general case where the model may be non-deterministic, the index set $\mathcal{J}$ represents all possible solutions.

Under the assumption of a deterministic model (see Appendix A for full derivation), this leads to

$$p\left(\vec{\eta}_i|\vec{O}\right) = \frac{p\left(\vec{O}|\mathcal{M}\left(\vec{\eta}_i\right)\right) p\left(\vec{\eta}_i\right)}{\sum_{l\in\mathcal{L}} p\left(\vec{O}|\mathcal{M}\left(\vec{\eta}_l\right)\right) p\left(\vec{\eta}_l\right)}, \tag{3}$$

where the index set $\mathcal{L}$ represents all of the investigated parameter sets. Having rewritten eq. 2 to this simple form allows us to estimate the posterior probability of a parameter set from two factors. The first factor is the probability of the observations given the model results (or likelihood), and the second factor is the prior probability of each parameter set. Their product is normalised by the sum of the probabilities for all model runs under consideration. The main task now is to calculate the first term, which connects the observations to the model runs.

The following assumes that each observation has a Gaussian uncertainty, as these are provided by the database for most observational points (handling of non-Gaussian uncertainty associated with limiting observations will be explained later). Under this assumption, the probability of the observations given the results of the ensemble runs can be estimated by a multivariate normal distribution. This can be expressed by (Wilks 2011)

$$p\left(\vec{O}|\mathcal{M}\left(\vec{\eta}_i\right)\right) = \left(\sqrt{(2\pi)^q \det\left(\Sigma_O\right)}\right)^{-1} \cdot \exp\left(-\frac{1}{2}\left(\vec{O} - \mathcal{M}\left(\vec{\eta}_i\right)\right)^T \Sigma_O^{-1} \left(\vec{O} - \mathcal{M}\left(\vec{\eta}_i\right)\right)\right), \tag{4}$$

where $\Sigma_O$ is the covariance matrix of the observations, and $q$ is the dimension of the observational vector.

The dependence between the different observations is generally not known. The only relational information given within the database is stratigraphic (i.e. relative temporal differences) for some individual observational locations, which we exploit in Section 2.2.2. As a consequence of this lack of additional information, the only known dependence between the different observational sites is provided by the model $\mathcal{M}$ itself. Intrinsically, we are assuming that the model will account for the spatial correlations between the observations. Therefore we assume that all of the observations are independent samples of the sea-level change with respect to the modelled change. This assumption simplifies the covariance matrix to a diagonal matrix with the observational height uncertainties as its elements. This leads to the equation

$$p\left(\vec{\eta}_i|\mathcal{O}\right) = \frac{\prod\limits_{j} \exp\left(-\frac{1}{2}\left(\frac{O_j - \mathcal{M}(\vec{\eta}_i)_j}{\sigma_{O,j}}\right)^2\right) p\left(\vec{\eta}_i\right)}{\sum\limits_{l\in\mathcal{L}}\prod\limits_{j} \exp\left(-\frac{1}{2}\left(\frac{O_j - \mathcal{M}(\vec{\eta}_l)_j}{\sigma_{O,j}}\right)^2\right) p\left(\vec{\eta}_l\right)}. \tag{5}$$

Under the given assumptions, this equation states the formally-correct probability.

For our particular dataset, this description of the probability poses problems. The database of LIG sea-level indicators is derived from many different sources, which can lead to inconsistencies, particularly with regard to the uncertainty estimates. Some of the observations may also be inaccurate or misinterpreted. In addition, spatial and temporal correlations due to unmodelled processes, such as changes to tidal ranges and the environmental habitats, may still exist. These can introduce structural errors into our analysis. Addressing these issues is a challenge, and all potential solutions include advantages and disadvantages. For example, it would be possible to model the outliers directly and alter the Gaussian probability curves for the uncertainties. Nevertheless, we do not have any valid information as to which data points should be treated as outliers or to what the character of the probability curves might be. As we use a data assimilation scheme, which depends on a large number of forward models, it is also not feasible to exclude specific data points by reiterating the procedure with different sub-samples of the observations. Thus, we have chosen another means of addressing the problem, which is based on an alternate formulation of the probability.

For this formulation, we first estimate the probability of each observation separately. Secondly, we calculate a weighted mean of all derived probabilities, which includes a weighting factor $w_j$ that allow us to give the different observations a specific weight within the analysis. This leads to

$$\overline{p\left(O_k|\mathcal{M}\left(\vec{\eta}_i\right)\right)}\Big|_{k=1:n_{obs}} = \frac{\sum\limits_{j} w_j p\left(O_j|\mathcal{M}\left(\vec{\eta}_i\right)\right)}{\sum\limits_{j} w_j}. \tag{6}$$

An advantage of this equation is that outliers have a limited weight in the calculation of the estimated probability. Expressing it in a similar form as eq. 5 leads to

$$p\left(\vec{\eta}_i|\mathcal{O}\right)_U = \frac{\sum\limits_{j} w_j \exp\left(-\frac{1}{2}\left(\frac{O_j - \mathcal{M}(\vec{\eta}_i)_j}{\sigma_{O,j}}\right)^2\right) p\left(\vec{\eta}_i\right)}{\sum\limits_{l\in\mathcal{L}}\sum\limits_{j} w_j \exp\left(-\frac{1}{2}\left(\frac{O_j - \mathcal{M}(\vec{\eta}_l)_j}{\sigma_{O,j}}\right)^2\right) p\left(\vec{\eta}_l\right)}. \tag{7}$$

Note that the two eqs. 5 and 7 are not equivalent, even when the weighting factor is ignored.

In the following ensemble experiments, we will replace the multivariate likelihood $p\left(\vec{\eta}_i|\mathcal{O}\right)$, expressed in eq. 5 and named the multivariate approach in the following discussion, with the

mean of the probabilities leading to $p\left(\vec{\eta}_i|\mathcal{O}\right)_U$, eq. 7 and call this the univariate approach. It has to be stressed that this replacement represents a significant modification. We will show the effects of this modification on a given set of runs in Section 3.2, which will illustrate the problems of the multivariate approach with the inclusion of potential outliers.

For a better understanding of the influence of the age uncertainties of the observations, we introduce an additional sampling strategy. Therein, to generate the probability for a model run, the mean over $N_o = 100$ sampled estimates of all observations is taken. For each observation, a new estimate is generated by independently drawing the age and the relative sea level from the Gaussian uncertainty estimates around the original values. For the calculation of the new relative sea level, the appropriate tectonic uplift or subsidence for the newly-sampled time is taken into account. Where the relative age between two observations is available, such as data from the Red Sea or the Netherlands, the sampling is performed in an iterative process. The oldest observation is sampled from the Gaussian distribution of the age and the relative sea level, while each succeeding observation uses the relative age and its uncertainty with respect to the preceding sample. In addition, the dataset contains limiting observations. For these observations, the probability is set to 1 when the model predicts a value on the correct side of the limit. Otherwise it is handled the same as the other observations.

To prevent a large influence of observational points that are a large temporal distance to the current time analysis point, the weighting factor $w_j$ in eq. 6 is used. The selection and weighting of observations is done in three steps. In the first, all observations prior to the current analysis point get a weighting of 1. The second step consists of two parts. In the first part, the observations with an age of maximally 2 kyrs in the future get a full weight. For the observations between 2 to 6 kyr in the future of the analysis point, the weight is decreased linearly from 1 to 0. In the last step, we ensure that enough future observations have a weight of 1 within the analysis. Therefore, independent from the temporal distance to the analysis point, the first 20 future data points get a full weight. Apart for temporal sections with a high density of observations, such as those around 125 kyr BP, the consequence of this procedure is that all past and the next 20 observations have a full weight, while later observations are not included into the analysis. Furthermore, this technique allows for the focus on the LIG without neglecting the influence of the near-future observations on the assimilation procedure. With this current dataset, this weighting choice effectively limits the influence of the MIS 5a observations on the analysis until the analysis point reaches about 122 kyr BP.

### 2.2.3   Dependence of runs

A consequence of the shared ice history in the development of the ensemble members is the statistical dependence of the runs. This dependence must be taken into account in the estimation of the uncertainties. Because the weighting function $w_j$ effectively changes the observations utilised in each round (i.e., calculation of the ice sheets at the next time analysis point), the prior must be modified at each analysis point. For this, an iterative algorithm is used, which estimates the posteriors in a round-based approach. Our particular choice of algorithm for the recalculation of the prior has two goals. The first is to distribute sensibly the posterior from the previous analysis point of a branched run between itself and its descendants. The second is to ensure that a run that was not branched at some past analysis point still has some probability and may yet be eligible for branching should the changing weighting of the observations at each analysis point increase the run's likelihood.

In the first round the probability of the prior is evenly distributed to each ensemble member and is given by the following equation:

$$p\left(\vec{\eta}\right) = \frac{1}{N_s}. \tag{8}$$

For each subsequent temporal analysis point, we update the priors of all runs, both those from the former rounds and those newly-created at the current analysis point. In the case where a run is branched in the current round $(n)$, the new prior of the parent (par) is calculated by dividing its posterior from the previous round $(n-1)$ by the product of its likelihood and the likelihood of all its descendants (desc), both calculated from the former round (a derivation of this prior is shown in Appendix B):

$$p\left(\vec{\eta}_i\right)_n^{par} := \sqrt[N_{r+1}]{\frac{p\left(\vec{\eta}_i|\vec{O}\right)_{n-1}^{par}}{p\left(\vec{O}|\mathcal{M}\left(\vec{\eta}_i\right)\right)_{n-1}^{par} \cdot \prod_j p\left(\vec{O}|\mathcal{M}\left(\vec{\eta}_j\right)\right)_{n-1}^{desc}}}. \tag{9}$$

This newly-calculated prior of the parent is also given to each of the descendants, which gives none of the newly-branched runs or the parent an advantage. For runs that were not branched at some past analysis point, the same equation can be used, where the the product of the likelihoods of the decedents is set to one. The final step recalculates the posteriors for all ensemble members for the current round by applying eq. 7 before proceeding to the next round. Beside updating the posterior probabilities, this scheme also allows us to directly compare ensemble members over several rounds.

In Section 3, the results will be shown in two ways, either with this prior used in the

assimilation procedure or with the uniform prior defined in Section 3.1. Both priors are based on different assumptions and thus deliver different results.

## 2.3 Ice sheet generation

In the last section, $\vec{\eta}$ was defined as the input parameters that define the model run. These include the ice sheet history, the development of which is described in this section, and parameters that specify the Earth model response, described in Section 2.4. The initial ice histories are constructed in two parts: generation of the ice volume history and distribution of those ice volume estimates over specific regions. Finally, we describe the mechanisms for varying the resulting histories to allow for a better fit to the data during the assimilation procedure.

Similar to Kopp et al. (2009), the history of ice sheet volume is generated from scaling a $\delta^{18}O$ stack (Lisiecki & Raymo 2005, LR04). However, we use a more regional approach and develop two separate scaling mechanisms that allow us to explore the influence of the timing and variability of the initial ice sheet models (Section 2.3.1). To disperse the regional volumes, we use the ICE-5G (Peltier 2004) as a template (Section 2.3.2). This version has a both a glaciation and deglaciation phase (e.g. Peltier & Fairbanks 2006), though only the deglaciation phase is well constrained. Thus, in the scaling and distribution procedures described below, we only use the ICE-5G model values from present day to last glacial maximum (LGM).

### 2.3.1 Calculation of the amount of global and regional ice

In this analysis, we base our calculations on the two different initial estimates of the global temporal and spatial ice sheet history, which are then varied to allow for a better fit to the data. To get a better understanding on the influence of the variability and timing of the initial ice sheet histories, two different methods are used to calculate the variations of the ice sheet volumes. Both versions rely on oxygen-isotope curves, which are a well established prior for global sea-level change due to changing ice sheets (Shackleton 1987). Nevertheless, there are difficulties with this proxy, in that other effects, such as temperature, are present in the record, and thus the relationship between the oxygen-isotope ratio and ice volume is not constant over time (Shackleton 1987; Waelbroeck et al. 2002). It has been suggested that the depletion of the oxygen-isotope with the temperature related effects exist mainly during the interglacial (Cutler et al. 2003; Elderfield et al. 2012).

Kopp et al. (2009) allowed for this variation in this ratio by assigning a standard deviation of 10 m, estimated from Bintanja et al. (2005), as an uncertainty to the ice sheet volumes generated from the $\delta^{18}O$ stack of Lisiecki & Raymo (2005). The two versions of ice volume

histories that we create are also based on the LR04 stack. While one version is just a simple scaling of this curve to estimate the global ice volume, the other reduces the variability of this curve by a running mean and further measures.

The first version is more directly comparable to the approach of Kopp et al. (2009). It is a simple scaling of the LR04 curve so that the maximum difference in the LR04 values between LGM and present day produce a change in ice volume equivalent to that of ICE-5G (Peltier 2004) over the same period:

$$V_i^{ice} = \gamma_1 \cdot (\delta^{18}O_i - \min(\delta^{18}O_j)) + \min(v_k^{ice}), \tag{10}$$

where the scaling factor, $\gamma_1$, is given by

$$\gamma_1 = \frac{\max(v_k^{ice}) - \min(v_k^{ice})}{\max(\delta^{18}O_j) - \min(\delta^{18}O_j)}. \tag{11}$$

The subscripts $i$ and $j$ are time points of the LR04 curve, and $k$ are time points from the ICE-5G dataset. Both $j$ and $k$ are restricted to the period between 26 kyr BP and present day. Rather than applying eq. 10 to the global ice volume, we apply it separately to each region defined by Kopp et al. (2009), so that $v^{ice}$ is the regional time series of ICE-5G and $V^{ice}$ is the new time series.

As previous studies have suggested that the total volume of ice during the LIG can be smaller than its present-day value, the scaling described by eq. 10 would lead to some regions being assigned negative values of ice volume. For this first version of the ice volume history, we will preserve the global average. To do this, we redistribute the negative ice volumes from these regions proportionally to the regions that still have ice cover (e.g. Greenland or Antarctica) and then set the negative ice volumes to zero.

The results for this version are shown by the black lines in Fig. 3. For the sake of this plot, the estimates of GMSL are simply derived by assuming a constant area of the ocean (the only time this assumption is made). As can be seen, using LR04 directly leads to large, rapid changes in the total ice volume, particularly during the deglaciation period at 129 kyr BP and at the peak of the deglaciation at 124 kyr BP. These rapid variations are particularly evident in the Southern Hemisphere. In addition, the duration of the interglacial, in this case defined as values below present-day ice volume, is only 4.5 kyr. To understand the influence of this high variability and short interglacial duration on the estimates of GMSL, especially during the LIG, a second, smoother version of the initial ice sheet history is introduced.

The second version, in the following named modified-$\delta^{18}O$, represents an ice sheet with less high-frequency variability. It will be used to show the influence of a different base ice sheet within our analysis. To create this history, we first apply a 10-kyr moving average to the

LR04 stack, which will be represented by an over-bar in the following equations. This modified curve is then scaled to the regional values of the ICE-5G dataset by matching the standard deviation from 21 kyr BP to present day. In this case, we also assume the present-day values are known and fixed, which leads to

$$V_i^{ice} = \gamma_2 \left( \overline{\delta^{18}O_i} - \overline{\delta^{18}O_0} \right) + v_0^{ice}, \tag{12}$$

where $\gamma_2$ is given by

$$\gamma_2 = \frac{\sigma \left( v_j^{ice}|_{j=0...21} \right)}{\sigma \left( (\overline{\delta^{18}O_j} - \overline{\delta^{18}O_0})|_{j=0...21} \right)}. \tag{13}$$

We note that as the long-term mean has been applied to the time series, using the maxium and minimum between the LGM and present day as a basis for scaling is not appropriate. Thus, we use the standard deviation of each time series, which provides a natural scaling metric. Removing the long-term mean from the smoothed time series, and then scaling by the standard deviations, however, may produce very different ranges from maxima to minima in the time series when compared to the ICE-5G time series. Thus, we modify the resulting $V_i^{ice}$ further in order to obtain a reasonable time series. These changes are described further below.

To prevent very small ice sheets during the late Holocene, each regional ice sheet is only allowed to deviate from zero after its value of ice volume exceeds 1% of the LGM value. For each $v^{ice}$ time series, we identify the initial time point $t_l$ that meets this condition. The value of the modified ice sheet, $V^{ice}$, at the $t_l$ is then subtracted from the whole ice history:

$$V_i^{ice} = V_i^{ice} - V_{t_l}^{ice}. \tag{14}$$

For this version, negative values of ice volume are set to zero. Thus, the ice sheets up to the point $t_l$ do not deviate from zero.

The prescription described above can lead to the regional ice sheets differing considerably from the values ICE-5G, typically resulting in larger volumes. As we do not want to exceed the values of the ICE-5G ice sheets by too much for the base ice sheet, we apply a correction to ensure a maximum deviation of up to 2% compared to the ICE-5G ice sheet at the LGM. From the starting point $t_l$ to the maximum value up to 21 kyr BP, a linear anomaly time series is created, which starts with 0 and ends with the necessary positive or negative ice volume needed to meet the criteria. Afterwards, this anomaly time series is constant at the maximum value for the rest of the reconstruction. In a last step this time series is added to the modified ice history $V_{ice}$.

This modified-$\delta^{18}O$ ice volume history is also shown Fig. 3 (blue line). The most notable difference with the $\delta^{18}O$ ice volumes is the missing high frequency variations and a longer

interglacial of 8.5 kyrs. The procedure also results in some unusual behaviour for the small ice sheets in the southern hemisphere mid-latitudes (SML). Given their small volume compared to the other regions, we neglect this result. The purpose of the second base ice sheet model is to examine an alternative ice history with less short time scale variability. The subsequent adjustments to eq. 12 ensure a reasonable time series after the averaging.

### 2.3.2   Spatial distribution

The ice volume histories for each of the initial seven regions were generated in the last section. In order to generate the spatial distribution of the ice sheets, the ICE-5G data set is used. The construction is done for each region separately. In a first step, we identify the time step in ICE-5G data set from 26 kyr BP to present day that has the ice volume closest to our constructed value. Then this spatial distribution is scaled proportionally to fit the ice volume given by the time series. For volumes smaller/larger than those during this period of ICE-5G, the smallest/largest volume is scaled proportionally. It must be stated that using this procedure leads to the assumption that the spatial distributions of the ice sheets during the glaciation phase are the same as during the deglaciation phase. While this is a necessary simplification given that this version of ICE-5G is only well constrained during the deglaciation, it is clearly unlikely.

### 2.3.3   Variation of ice sheets for the data assimilation

For the data assimilation, variations of the two initial ice histories have to be generated. The variation is introduced by generating an anomaly time series, which is added to the global ice volume time series. These anomalies start at the analysis time and reach a maximum after a linear development of 0.5 kyrs. Afterwards this time series stays constant until 50 kyr BP. From then on the anomaly decreases linearly to zero at 10 kyr BP and stays at that value until present day.

The ensemble members are differentiated by the maximum of the anomaly. In the first round, with $N_s = 100$ deviations from the initial ice history, the anomaly of the members is equally distributed between -24.75 m and 24.75 m. This leads to parallel ice volume time series during the LIG with a difference of 0.5 m. For the subsequent rounds the maximum of the anomaly is equally distributed for the $N_r = 10$ ensemble members for each of the $N_k = 10$ seed runs between -6.75 m and 6.75 m. Thus, the parallel ice volume anomalies in this case have a difference of 1.5 m. These anomalies are then added to the seed run. Over the course of several analysis points, several anomalies can be added to the ice sheet time series

to form a given ensemble member. It has to be noted that in the procedure for generating a modified-$\delta^{18}O$ ice history, the setting of negative ice volumes to zero leads to a different ice volume than the input ice volume history.

With each of these new time series, we apply the procedure described in Section 2.3.2 to create the spatial distribution. With multiple analysis points and anomaly time series, it is possible to generate different ice sheet configurations.

## 2.4   GIA model

In addition to the spatial and temporal history of the ice sheets, GIA predictions also require a model of the Earth's response to the changing ice and water loads, as well as a formalism that describes the self-consistent redistribution of water in the oceans. In regard to the Earth models, we adopt the common assumption of a spherically-symmetric, self-gravitating, compressible Maxwell Earth model (Peltier 1974; Mitrovica & Peltier 1992). The newest generation of GIA models do not necessarily make these assumptions (for a review see Steffen & Wu 2011). However, our statistical approach requires tens of thousand of forward calculations, and thus we need to rapidly calculate the required GIA predictions of relative sea-level change at the observation locations. For this reason, we use the standard, and simpler, Earth models.

In this study, we use three different viscoelastic Earth models, distinguished on the basis of lithsopheric thickness and depth-varying mantle viscosity. We assume that the lithosphere behaves elastically, and that the mantle viscosity is constant within two layers that are separated by a boundary at 670 km depth. The elastic and density profile throughout Earth's interior is taken from the Preliminary Reference Earth Model (PREM) (Dziewonski & Anderson 1981). Our specific choices for the viscosity and lithospheric thickness are summarised in Table 1 and were adopted from previous modelling studies. The first Earth model, E1, is similar to the E1 model used in Lambeck et al. (2012). For E2, we use parameters similar to those in Raymo & Mitrovica (2012). The final model, E3, is the VM2 model (Peltier 1996). Typically, predictions of sea-level change in the far-field of the Late Pleistocene ice sheets are sensitive to the adopted value of lower mantle viscosity, and our choices for $\nu_{LM}$ span a range of the predicted sea-level behaviour at far-field locations. In general, the three models produce quite different predictions for a given ice sheet model, and in Section 3.4 we will examine the differences in the inferred maximum sea level during the LIG calculated with each.

The GIA modelling also requires a formalism that describes the redistribution of water in response to the exchange of mass between the oceans and the ice sheets. In this regard, we implement the pseudo-spectral sea-level algorithm detailed in Kendall et al. (2005). Their

treatment accurately accounts for grounded, marine-based ice, evolving shorelines, and rotation effects on sea level. With regard to this last effect, we compute perturbations in the rotation axis using the theory described in Mitrovica et al. (2005). In addition, the sea-level algorithm is constrained, via an iterative scheme, to yield present-day topography that converges to the observed.

For the statistical analysis it is important to acknowledge that the GIA model results are non-Markovian. This means that the results in any time step cannot be created by only knowing the evolution of parameters to this time point and the state of the model of the current time step. In particular, future variations in the ice sheet and Earth response also impact the values of relative sea level at any time step, because any relative sea-level curve is constrained to pass through zero at present-day. It is therefore necessary to evaluate the full ice histories and not only temporal subsets of them. Nevertheless, it is reasonable to assume that the sensitivity of the predictions at a given time progressively decreases as one considers loading in the distant past. Indeed, most of the variations, to first order, can be considered Markovian. For this reason, our analysis of peak relative sea level at the LIG will not consider the loading history prior to the penultimate glacial cycle (see Section 2.2). Furthermore, this first-order approximation allows us to apply the particle filter, as it must be based on a Markovian model.

## 2.5    Example of the data assimilation scheme

To illustrate the data assimilation scheme, we created a sketched example, shown in Fig. 4. In this example, we start with two runs with different anomalies to the global ice volume, which when combined with a base ice volume model would form the basis for the ice sheets used in the GIA model. These anomalies are shown in Fig. 4a as red and blue lines. Also shown are the observations, which are locally corrected for GIA. Note that if the local deviations of the observations from the global average were only due to GIA, and the GIA model perfectly estimated these effects, then the GIA-corrected observations should agree with the ice volume anomaly. As a different history of the ice sheets changes the local GIA effect, the observations have different values of relative sea level for the two runs. This is illustrated by using blue crosses for the observations corrected using the blue anomaly and red crosses for the observations corrected using the red anomaly. The black vertical line shows the current analysis time point. Under the assumption that both runs are the first and only runs in the data assimilation procedure, both have the same prior. As a consequence, the difference in the posterior probability is created by the distance to the observations only. The weighting

algorithm takes therein all observations before the line with a full weighting (represented by the thick black line above the two anomalies) and afterwards with a partial weighting described in Section 2.2.2 (represented by the thick grey line). To estimate the probability for each anomaly, 100 iteration of all of the observations within their two-dimensional uncertainty structures are drawn, and for each the probability is estimated with eq. 6. A mean over all separate 100 probabilities is taken and afterwards used in the estimation of the posterior. In this visual example, only the best run is used as a seed for the next round, which in this case is the red anomaly.

A branching is shown in Fig. 4b, which creates two new runs (red and blue) from the previously-selected best run. The previous anomalies are shown in grey, and the purple line indicates a common history shared by the red and blue anomalies. While the blue anomaly shares a value during part of the time history with the anomaly that was not branched, that does not imply that its distance to the observations, and with it the probability, is the same. Due to the different time history, the differences due to the GIA contribution can be quite large. Unlike in the former round, in this case the prior plays an important role. For this, the posterior for the runs of the former round will be estimated with the new weightings for the observations (the changed thick bar from Fig. 4a to 4b). In a next step, the prior for the branched run (the red line in Fig. 4a) is adapted as explained in Section 2.2.3, which simultaneously estimates the prior for the runs created at the second analysis point from the branched run. In a third step, posterior for the runs of created at the new analysis points are estimated. With this, the posteriors of all four runs are known, and in a next step the run with the highest posterior is selected as the new seed run. In the example shown in Fig. 4b, this would be the blue run. With this selection, the round is concluded, and the next analysis step is analysed (Fig. 4c). This round includes a new weighting and the determination of the best run. All in all the most likely run after all the analysis points is shown in Fig. 4d. It has adapted to the observations, which are shown for this run by the crosses.

## 3   RESULTS

In the following, we demonstrate the application of this statistical framework to determining the maximum GMSL during the LIG. After examining the behaviour of the most probable runs, we calculate the probability density functions (pdfs) and cumulative distribution functions (cdfs) of the maximum GMSL, illustrating the effect of the different initial ice histories and the input Earth models on the results. Most importantly, we are able to demonstrate how

each of the observations contribute to these probabilities to explain the characteristics of the pdfs.

To describe the probability accurately, we have to calculate the posterior and incorporate therewith a prior. However, the original prior used in the assimilation scheme (Section 2.2.3) does not allow us to vary the assumptions going into the analysis, such as excluding certain observations, without recalculating the runs. Thus, we introduce an alternate prior in Section 3.1 that allows us to evaluate a set of runs with respect to the observations without taking the run generating mechanism into account. This alternate prior also allows us to illustrate the impact of the choice of the prior on the results, with the alternate and original priors compared in Section 3.2.

### 3.1    GMSL estimate with uniform prior

For each of the three earth models, the data assimilation scheme is executed using both of the initial ice histories. The determination of the most probable runs at the analysis points is judged on the posterior of the univariate method described in Section 2.2.2, including the sampling strategy for the observations. All in all 39,000 (3 Earth models, 2 initial ice histories, 64 time analysis points plus the initial set of runs, 100 runs at each point) runs are calculated and analysed. Fig. 5 shows all the runs during the LIG. The shading of each line represents the relative probability, calculated with the application of a uniform prior for all runs at the last time analysis point on the basis of the original, unsampled observations. This prior gives every run the same probability, and the resulting posterior is equal to the likelihood. Thus the more probable runs, those with darker shading, are closer to all of the observations in terms of uncertainties than the runs with lighter shading.

The first thing to note in Fig. 5 is the dark band, which peaks near 8 m above present-day sea level. This band has a low variability during the interglacial. As the six experiments (an experiment being the 6500 runs associated with a given Earth model and base ice sheet) can be easily distinguished by their base ice history, we can identify that these high-probability runs are based upon the modified-$\delta^{18}O$ ice history. While the 19,500 runs generated by the three experiments based on the unmodified-$\delta^{18}O$ ice history are also shown, identifiable by the variation during the deglaciation near 128 kyr BP and the rapid fluctuations during the interglacial, their posterior using the uniform prior with the unsampled observations is noticeably lower.

Within the band of most-probable runs, several notable changes occur between 128 and 116 kyr BP. The first notable change happens at around 125 kyr BP, when the dark band splits into

several single recognisable runs. The primary reason for this is the number of observations. Given the weightings described in Section 2.2.2 and the requirement that at least twenty future observations are used at an analysis point when available, new observations come into the calculation at each time analysis point, changing the character of the most probable runs. With the large number of observations near 125 kyr BP, those future observations are closer to the current time analysis point than others, contributing to the splitting of the main band.

During the period from 125 to 120 kyr BP, the diverging runs generally indicate an increasing global mean sea level. In addition, the band of runs with much lower probability, illustrated by the lighter grey in the background of the main black runs, is widening. Near 122 kyr BP, the observations from MIS 5a (those at 74 and 85 kyr BP) enter into the analysis. The two observations from Barbados indicate a relatively high sea level at their respective times, while having a relatively small height uncertainty. As a consequence, they receive higher weighting and, when included, are a main driver for the data assimilation procedure. With a small number of sea-level observations, particularly in MIS 5d or MIS 5b, the constraints on the end of the interglacial are weak, leading to a much broader distribution of probable runs.

### 3.2 Maximum during the LIG

The primary goal of the framework is to be able to evaluate the probability of some prediction from the runs, such as the maximum highstand during the LIG. To do this, we use ensemble kernel dressing to calculate this probability and its uncertainty, applying the technique of Gaussian mixture modelling (GMM, Schölzel & Hense 2011). For each run, a normalised Gaussian curve, $\mathcal{N}(v_p, \sigma_{rep})$, is drawn around the predicted value, $v_p$. These curves have a mean of $v_p$, here the maximum sea level during the LIG, and a standard deviation, $\sigma_{rep}$, indicative of the range for which the run is representative, given the input parameters. This GIA model uncertainty should account for the effects on LIG predictions of unsampled uncertainty of the ice sheets after the LIG, as well as other possible sea-level effects, such as regional steric changes in sea level. We set $\sigma_{rep} = 1$ m, which is slightly larger than one-half of the 1.5 m of the branching spacing at the analysis points. These curves are summed for each value of a discretised range $v$, by weighting the curves with their probability $p_i$. This leads to

$$p(v) = \sum_i p_i \mathcal{N}_i (v_p, \sigma_{rep}) |_v. \tag{15}$$

As $p_i$ we use a posterior, calculated with one of the priors: either the prior used to create the runs (Section 2.2.3) or the uniform prior introduced in the last section.

The first example, shown in Fig. 6, uses all of the runs, the full observational dataset and the univariate estimation of the probability. We apply both the uniform prior with the sampling algorithm for the observations and the prior used to perform the data assimilation to obtain two pdfs, which are clearly non-Gaussian, with two distinct peaks. These peaks result from our use of the two base ice histories, which will be demonstrated and explored further in Section 3.3. From these pdfs, the cdfs and the percentiles can be calculated. We start with the cdf results using the prior of the data assimilation. The median value is 11.3 m, with the 33rd (likely) and the 66th percentiles (unlikely) given by 8.1 m and 14.2 m. It is notable that the 10th and 90th percentiles are 6.6 m and 15.6 m above present-day sea level. It is interesting to compare this with the results using the uniform prior (the cdf is not shown in Fig. 6, but the results are listed in Appendix C). In this case, the median is 11.2 m, with the percentiles for the likely and the unlikely values being 8.4 m and 14.1 m. The range between the 10th (6.1 m) and 90th (16.7 m) percentiles increases. The application of the prior of the data assimilation leads to more pronounced peaks and reduces the tail on the upper side. Primarily, we will show the results using the prior from the data assimilation, as it generally shows a smaller uncertainty range while having a similarly shaped pdf compared to the uniform prior. The percentiles for both priors are shown in Appendix C. We will return to these results in Section 3.3.

The second example illustrates the effect of excluding those observations that were also dropped by Kopp et al. (2009) during their analysis. It should be noted that through the assimilation procedure, these observations will still have had an impact on the results. However, here we recalculate the posterior without these observations, altering the runs' probability. Using the uniform prior, the comparison of the pdfs of the posterior calculated with and without these seven observations is shown in Fig. 7a and shows nearly identical results, as shown in the difference plot between the two analyses in Fig. 7c. The excluded observations have a very low probability for most runs and do not have any influence on the analysis when the posterior is calculated using the univariate analysis of eq. 6 including the sampling algorithm based on the full dataset. Thus, ignoring these observations has no effect.

As an illustration of why the altering of the probability from the multivariate to the univariate analysis reduces the impact of outliers, and to illustrate that these observations could have large impact on the results, we repeat the analysis of the runs with the multivariate approach. The results are shown in Figs. 7b and 7d, with those of the reduced dataset more notably different to those of the full dataset. However, the notable difference from Fig. 7a is the shift of the second peak associated with the $\delta^{18}O$ ice history to substantially higher values, which is common to both datasets. This is a result of the sampling procedure described in

Section 2.2.2. Some of the observations have very large temporal uncertainties, which effectively means that they could occur either during the interglacial, or during the deglaciation or glaciation on either side. This introduces a potentially large range of differences from the GIA model prediction. Given that multivariate analysis can assign more probability to runs matching these shifted data points, and results based on the $\delta^{18}O$ ice history are more susceptible to these changes (described further in the next section), the end result is a much higher, and unreasonable, estimate of the highstand.

In the end, the univariate approach for a given observational dataset distributes the probabilities over many runs. In contrast, the multivariate approach assigns most of the probability to a very few runs. This sensitivity and the low number of runs that have any probability lead us to the decision to use the univariate approach within the data assimilation procedure. Indeed, if we were to use the multivariate approach, it typically would not be appropriate to select 10 runs with the highest probability for the branching procedure.

### 3.3   Influence of the initial ice history

A third example demonstrates the influence of the initial ice history by applying the univariate approach with the sampled observations separately on the two groups of runs based upon the $\delta^{18}O$ and the modified-$\delta^{18}O$ ice histories. The comparison of the results of these two analyses is shown in Fig. 8. Both pdfs are nearly Gaussian, and the median highstand of the modified-$\delta^{18}O$ ice history is 7.5 m, around one-half as much as that of the $\delta^{18}O$ ice history. Both uncertainty ranges are similar, with an inter quartile range (IQR) of around 1.5 m. The distinct results for the two different base ice models demonstrates that the differences in the these base models leads to the distinct peaks in Section 3.2. Clearly, the variability and character of the initial ice model history can have a large impact on the results. As the probabilities related to the each base ice model separate, it is likely that the variability introduced in the assimilation scheme is insufficient to completely characterise the pdfs and cdfs associated with the maximum highstand during the LIG. For our purposes, we will refer to the results from each base ice sheet separately.

A number of reasons could lead to the higher median in the runs based upon the $\delta^{18}O$ ice history. With regard to the median values, the large spikes in GMSL coupled with the relatively short interglacial could lead to higher values, as the history is adjusted upwards to accommodate data at the ends of the interglacial. We discuss this further in Section 3.5. Additionally, small changes in an observation's age can cause a large drop or rise in the probability of a specific run, as this is only influenced by the difference between the observation's

value of relative sea level and the associated prediction from the GIA model run. In the case of the modified-$\delta^{18}O$ ice history, the smaller variability and longer interglacial results implies that changes the the sample's age will cause less variability of the probabilities within the sampling algorithm.

### 3.4    Influence of the Earth model

The Earth model parameters are another component of $\vec{\eta}$ that could introduce variation into the resulting pdf. To demonstrate the effect of this set of variables, the pdfs are recalculated separating the runs for each of the three different Earth models. As we have seen in the last section, changing the base ice history introduces the largest variability. Therefore, we will look at the variation between the Earth models separately for the two different base ice histories. The results are shown in Fig. 9. All of the individual curves are nearly Gaussian. For the $\delta^{18}O$ ice history results, the medians of the highstands are 15.0 m for the E3 model and 14.5 m and 14.6 m for the other two models. The other percentiles exhibit similar differences across the models. In the case of the modified-$\delta^{18}O$ ice history the median of the highstand for Earth model E1 is 6.7 m, for E2 7.7 m and for E3 8.4 m. The IQR for all of the results is near 1.4 m. There are relatively small variation of the pdfs that utilised the different Earth models. Compared with the much larger differences introduced by the initial ice sheet histories, this justifies the primary focus on adapting the ice sheet history to better fit the data.

### 3.5    Analysis of individual observations

One of the major advantages of our methodology compared with that of Kopp et al. (2009) is the ability to examine the contribution of individual observations to a GIA model prediction, such as the maximum highstand. This ability also allows us to better understand the behaviour of the pdfs in Fig. 8. To demonstrate this, we show the probabilities of the LIG highstand for each observation, using the two different initial ice histories and a fixed Earth model, E1. The results for this analysis are shown in Fig. 10 for the $\delta^{18}O$ ice history and in Fig. 11 for the modified-$\delta^{18}O$ history.

Both figures are constructed in the same way. On the x-axis the observations are sorted in chronological order, which leads to a non-linear temporal axis. Note that the large number of observations at 125 kyr BP leads to the number being repeated twice on the figure. The y-axis is non-linear as well and shows the maximum sea level during the LIG for each run. Thus, each row of the plot represents one run, sorted by the maximum highstand, and each column represents one observation. The colours indicate the probability of the run using

the uniform prior given the unsampled observations from the database. In the case of the univariate approach, the mean of one row is proportional to the probability of the specific run for the observations; in case of the multivariate approach, it is the product. It should be noted that these results are for the original, unsampled observations. Thus, taking the sum would not reproduce the corresponding result in Fig. 9, which does include the sampling.

In general, each observation has a band of high probability for several runs, with the probability decreasing towards runs with higher or lower maximum sea level. Nevertheless, this band of high probability is not always continuous in the direction of height. The reason for this is that each run takes a different path. For example, while two runs may have a similar value of the maximum highstand, that maximum may occur at different times. Thus, distance of an individual observation to the prediction of the two runs may be quite different. Other observations can be identified with a large band of a very high probability for all runs above or below a certain threshold. These are associated with the limiting observations in the database.

Both plots show that no runs have a high probability for all observations. Consequently, each observation contributes probability to different estimates of GMSL during the LIG in this framework. Both plots show a large number of runs (a wide band on the y-axis), with a small range of highstand values. This band exhibits discontinuities in probability at its borders in the columns for the individual observations. This set is the runs with the darker lines in Fig. 5. These runs share a common history that had the highstand prior to their eventual separation. Therefore, they are ordered near to each other. The results using the $\delta^{18}O$ ice history, Fig. 10, show three different periods of the behaviour of the observations. For the early observations, up to 125 kyr BP, there are large jumps between the most probable values of GMSL. While observations from the Red Sea indicate a low highstand, observations from the Netherlands and Bahamas indicate an extremely high value. This can be seen by the low highstand values associated with the Red Sea data in the bottom part of the plot and the higher values for the other observations at the top part. The next period covers the main timespan of the LIG between 117 and 125 kyr BP. During this time, most observations indicate a relatively stable maximum sea level. While outliers are still contributing to both high and low values, they are neither as prevalent nor systematic as before. Beginning around 123 kyr BP, we see that the band with nearly identical highstand values shows greater probability variability for some observation. In these cases, the history that these runs follow after their maximum alters their probability as additional observations enter the analysis. Only a small number of runs show larger values of probability for observations after 117 kyr BP. Due to the high sea-level values

from indicators in Barbados, the data assimilation scheme has to adapt the initial ice histories massively in order to obtain similar predictions. Thus, many of the runs that were generated in the rounds prior to the inclusion of this data exhibit no probability with these observations.

On average, the results using the modified-$\delta^{18}O$ ice history, Fig. 11, have a lower value for the maximum GMSL. In the first period, up to 125 kyr BP, the runs are consistently further away from the Red Sea data and therefore show high probabilities for runs with extreme low highstand. Unlike the last plot, the probability within the large band of similar highstands show a longer time span with nearly constant probabilities for these runs for the Red Sea data. This can be explained by the lower variability of the base ice history. If a run generates a higher sea level closer to the end of the interglacial, it is likely the higher value will be a new maximum value for the highstand. Thus, this run would likely be sorted to a higher position on the y-axis. Observations from Barbados after 117 kyr BP indicate an extremely high highstand. In this case the data assimilation scheme is not capable of bringing the main runs from the LIG period in line with these observations. It does try, however, and results can be seen in the drift towards higher values at the end of the LIG, as seen in Fig. 5.

This comparison of Figs. 10 and 11 is able to explain the differences in the pdfs of the high highstands seen in Fig. 8. The large band of similar highstands is systematically higher in Fig.10 than in Fig. 11, resulting from the different timing of the initial ice histories at the start and end of the interglacial. The modified-$\delta^{18}O$ ice history has an earlier end to the deglaciation and a later onset of glaciation than the $\delta^{18}O$ ice history. Given the shape of GMSL from the $\delta^{18}O$ ice history, on average the ice history must have less ice, i.e. a higher sea level, in order to extend the interglacial over this period.

### 3.6   Analysis of individual runs

To demonstrate a further capability of this data assimilation scheme, we examine two runs directly and compare them to the observations. We chose these two runs based upon them having the highest probability for each base ice history after applying the prior from the data assimilation scheme, with the result in Fig. 12 based upon the $\delta^{18}O$ ice history and that in Fig. 13 based upon the modified-$\delta^{18}O$ ice history. Therefore, they are representative of the runs associated with the two peaks within Fig. 8. Shown are the sea-level curve and relative to this the GIA- and uplift-corrected observations with their associated one standard deviation uncertainties in time and elevation.

Fig. 12, associated with the $\delta^{18}O$ ice history, shows that in the deglaciation phase of the global mean sea-level curve follows the Red Sea observations. It also demonstrates that

several observations indicate much higher sea levels at the very early stage of the interglacial. Between 118 and 126 kyr BP we see a large area of the plot with a very dark shade of grey where the uncertainties of the observations overlap. This area covers the a range of GMSL from 4 to 13 m. There are also some observations indicating much higher sea levels than is optimised by the data assimilation scheme. Those instances where the uncertainty range for the observations extend to the top plot are mainly due to observations that are limiting values. At the end of the interglacial the sea-level curve stops following the Red Sea data. This feature can be connected with the attempt of the assimilation scheme to generate GMSL results need to describe the observations at 74 and 85 kyr BP.

In Fig. 13, associated with the modified-$\delta^{18}O$ ice sheet, we see a much flatter sea-level curve. At the start of the interglacial, the sea-level curve is further away from the Red Sea observations than in the last plot but is closer to the other observations. From 122 kyr BP a second rise in GMSL can be seen. In this case, as with the previous result, the most probable run deviates from the Red Sea data at this time.

Much care must be taken in interpreting these plots, as changes to GMSL curve would result in changes to the GIA contribution. Thus, there is not a direct correspondence to between changes to the GMSL curve and the relative positions of the observations. The main outcome of these two plots is that the Red Sea data, which forms 30 of the 107 observations within the dataset, is obviously not in line with the other observations in timing of the LIG. It also shows that the uncertainties generated by the statistical methods might underestimate the real uncertainty, as systematic problems within the observational dataset are not adequately represented in the model.

## 4   DISCUSSION

The analysis shows that the values for the maximum highstand during the LIG and its uncertainty depend highly on the statistical approach used to compare the GIA models to the sea-level data, and on the related assumptions used in the analysis. That is especially true for inhomogeneous observational data with large uncertainties that are currently available.

Our results illustrate the sensitivity to the input models and statistical assumptions. In some cases, we obtain similar estimates of the maximum GMSL during the LIG to that of Kopp et al. (2009). For example, using the modified-$\delta^{18}O$ ice history with the univariate approach, we obtain a median estimate of 7.5 m, with a 1.1 m range between the 33rd and 66th percentiles. This compares to a range of 1.4 m (8.0-9.4 m) in Kopp et al. (2009). However, this result was highly dependent upon the base ice model. Using the $\delta^{18}O$ ice history result in

a much higher median (14.7 m). These results are pushed even higher using the more rigorous (and correct) multivariate approach, which is particularly affected by our assumption that there were no outliers or no structural uncertainties. The clear separation of the two base ice model results suggests that our assimilation procedure failed to completely characterise the pdfs and cdfs of the associated with the maximum highstand during the LIG. Nevertheless, it does highlight that these input models can have a large impact on the estimates, suggesting that the range of possible maximum highstand values during the LIG based on this dataset may be greater than suggested by Kopp et al. (2009).

The primary advantage of our approach is the possibility to connect the results with the input parameters of each ensemble run. With this knowledge it is possible to analyse the sources of uncertainty. This has been shown for varying observations (Section 3.2), different initial ice histories (Section 3.3) and different Earth models (Section 3.4). Furthermore, by having the opportunity to separately examine the contribution of each observation, as is shown in Sections 3.5 and 3.6, a very detailed analysis can be performed. Generalising, every input parameter used to create the ensemble runs can be analysed on the basis of any possible combination of the observations.

The statistical framework from this study in its simplest form allows one to compare two different ice histories and classify their performance given a set of observations and their associated uncertainties. In the case where initial set of input parameters generated runs with a consistently good fit to the observations, the framework could be used to determine the ice sheet model configuration during the LIG and Earth model parameters given the sea-level indicators. Nevertheless, this set of observations does not show a consistent picture of the ice sheet volume during the LIG, even allowing for GIA effects. This inconsistency in the observations is especially evident in the results of Section 3.5, where the framework allows us to illustrate the influence of each observation. In particular, there are obvious problems relating the early and late observations of this dataset with the timing of $\delta^{18}O$-derived ice volume history as it was demonstrated in Sections 3.6.

In general there are three options to explain the discrepancies with the observations. The first is that the GIA model predictions, in connection with the ice sheets driven by the LR04 curve, do not adequately represent the spatial and temporal variability of the physical system. This would not be a surprise, given that a simple model approach is used. (Note the further discussion below with regard to ice sheet models, both with regard to the simultaneous variation of the different ice sheets and to the simplified spatial distribution highlighted in Section 2.3.2.) For example, the limited number of Earth models utilised in the GIA modelling

may have contributed to upward drift in the GMSL observed in the runs towards the end of the LIG (Fig. 5). While the three models cover a range of behaviours possible given a variation of Earth model parameters, a very specific set of Earth models and ice sheet changes is needed to match near present-day sea level during MIS 5a in western Atlantic (Potter & Lambeck 2003). In addition, other spatial correlations may exist in the data, such a those introduced by changing tidal ranges and environmental habitats. However, there would need to be a further investigation into the actual causes and the consequences. Particular issues that would need to be explored are the timing of the LR04 curve, its relationship with global ice volume, and the discrepancies between the observations in West Australia and Barbados. Recently Lin et al. (2014) reported spatial variations in $\delta^{18}O$ time series with regard to the timing of the LIG, which may allow in future a different approach to connect them to ice sheet changes. It should be noted here that the Earth model parameters only had a small influence on the results. However, their importance may increase if the agreement between the initial GIA model predictions and the observations were closer. Problems relating the validity to use a benthic oxygen-isotope ratio at all as a basis for reconstruction of ice sheets were already mentioned in Section 2.3.1.

The second option would be that the observations, or their interpretation, are not sufficient to describe (i.e. sample) the system represented by the GIA model. In particular, the variety of observations in the database can lead to inhomogeneity in the uncertainties, which has a large influence on the results. When the uncertainties are underestimated, the associated observations receive a higher weight within the analysis. If the underestimated uncertainty is associated with an outlier, it can drive the assimilation scheme to an incorrect result. This issue is relevant for our results, in that since the dataset was published by Kopp et al. (2009), the timing of the most important subset of the dataset, the Red Sea data, was revised in Grant et al. (2012). The influence of this modified timing on the results, as well as additional observations compiled since then (e.g. Dutton & Lambeck 2012), needs further investigation.

The third possibility is driven by the statistical analysis method and explicitly by the assumptions made within it. We demonstrated that a simple assumption, whether outliers exist within the dataset or not (e.g. Fig. 7), has a considerable influence on the result. An outlier-free observational dataset can not be assumed given the results of the analysis on individual observations in Section 3.5. Identifying and reducing the influence of outliers is the aim of every statistical analysis of this kind. By analysing the potential outliers, the statistical framework can also be used in the sense of a quality assurance for the observational dataset. Furthermore, our analysis has not accounted for several potential structural uncertainties, which might

lead to different results. Among them are the correlation of observations over several sites, apart from those given by the GIA model, and non-linear tectonic movements. Non-Gaussian distribution of uncertainties were only partially included by the sampling algorithm and the special treatment of the limiting observations. Those and other structural uncertainties are generally not yet well defined by the underlying science. With additional information on the model and the observations, the statistical approach could be tailored to the problem and might lead to a better representation of the uncertainties.

There are several limitations of our analysis that could be improved in future studies. Currently, we assume that the different ice sheet volumes vary simultaneously, or at least to the extent allowed by the construction process in the case of the modified-$\delta^{18}O$ ice history. To address this limitation within the data assimilation scheme would require different scenarios for the ice sheets at every analysis point, substantially increasing the required computational time. In addition, we have used only one ice sheet history as the basis for the spatial reconstruction. It is possible that problems for certain regions might multiply within the data assimilation scheme. Such a scenario may arise if the observational dataset consists only of observations from a single region for a longer time span. Currently, a given ice sheet history from LGM to present day is used as the basis for the earlier reconstructions, causing a similar distribution during both the glaciation and deglaciation for a given ice volume. Ideally, one would use online coupled ice models or reconstructed ice sheets directly from that time span. An example for such an analysis is de Boer et al. (2013), who modelled $\delta^{18}O$ and eustatic sea level over 1 million years. Nevertheless, both the computational costs and the availability are a major problem in its realisation. Uncertainty about the validity of results would also exist due to the possibility that individual sea-level indicators are being used both to create the ice sheets and to constrain the statistical analysis, which would lead to a potential circular argumentation (Tebaldi & Knutti 2007). In general it has to be assumed that taking only a small subset of potential Earth models, using only LR04 and ICE-5G in the ice sheet reconstruction, and the synchronised changes on both hemisphere leads to an underestimation of the uncertainties. Even by using the data assimilation scheme and the different parameter settings, the real combination of ice sheets and Earth model parameters is not likely to have been tested. Furthermore, due to the sparseness of the data and the selection process at the analysis points in the data assimilation, it is hardly possible to perform a proper statistical verification of the results.

Our results imply that the observational dataset is too small and inhomogeneous to make a definitive statement on the sea level during the LIG. This conclusion is not necessarily in dis-

agreement with Kopp et al. (2009). Our median estimated with the modified-$\delta^{18}O$ ice history and their central value are similar, but we demonstrate a greater range of results is possible, depending upon the input models and statistical assumptions. Given that the datasets used by both analyses are the same, the difference in the uncertainty estimates originates from the different statistical methods and assumptions applied. We have demonstrated that the results can be very sensitive to the assumptions, e.g. the existence of outliers, as well as the initial input parameters, e.g. the initial ice volume history. The observational dataset is not large enough nor consistent enough to rectify the discrepancies between the different statistical analysis.

## 5   CONCLUSION

We start with a general Bayesian approach to calibrate simple models to observations. To address problems with outliers (large differences between the observations and the GIA model predictions), we deviate from this approach to a more adhoc data assimilation scheme based on particle filters. By applying the procedure to the datasets used by Kopp et al. (2009), we can obtain similar results for the median estimate of maximum GMSL during the LIG, but only with a particular set assumptions. Our results are very sensitive to the analysis assumptions, such as the initial ice history or the existence of outliers. This dataset of sea-level indicators and their uncertainties does not show a consistent picture for the sea-level highstand during the LIG, when it is assumed that the LR04-derived ice volume history and the GIA modelling produce reasonable predictions. These results suggest that additional data, as well as consistent and realistic assessments of the uncertainties, are necessary to generate better estimates of sea-level variations during the LIG.

The next step will be the usage of additional sea-level indicators, with a more consistent uncertainty assessments, and different ice sheet scenarios to explore the sources of the uncertainties. An aim will be to identify which component of the analysis, e.g. the observations, the statistical approach or the ice sheet and Earth model parameters, that is responsible for the large range of results. The statistical framework presented here is generally applicable for comparison of model predictions to observations. Thus, it could be widely applied in other studies.

## 6    ACKNOWLEDGEMENT

## REFERENCES

Barlow, N. L., Long, A. J., Saher, M. H., Gehrels, W. R., Garnett, M. H., & Scaife, R. G., 2014. Salt-marsh reconstructions of relative sea-level change in the north atlantic during the last 2000 years, *Quaternary Science Reviews*, **99**(0), 1–16.

Bayes, T., 1763. An essay towards solving a problem in the doctrine of chances, *Philosophical Transactions of the Royal Society of London*, **53**(370–418).

Bintanja, R., van de Wal, R. S. W., & Oerlemans, J., 2005. Modelled atmospheric temperatures and global sea levels over the past million years, *Nature*, **437**, 125–128.

Bowen, D. Q., 2010. Sea level ̃400 000 years ago (mis 11): analogue for present and future sea-level?, *Climate of the Past*, **6**, 19–29.

Briggs, R. D. & Tarasov, L., 2013. How to evaluate model-derived deglaciation chronologies: a case study using antarctica, *Quaternary Science Reviews*, **63**, 109–127.

Briggs, R. D., Pollard, D., & Tarasov, L., 2014. A data-constrained large ensemble analysis of antarctic evolution since the eemian, *Quaternary Science Reviews*, **103**, 91–115.

Campbell, E. P., 2005. Statistical modeling in nonlinear systems, *Journal of Climate*, **18**, 3388–3399.

Church, J., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., Merrifield, M. A., Milne, G. A., Nerem, R. S., Nunn, P. D., Payne, A. J., Pfeffer, W., Stammer, D., & Unnikrishnan, A., 2013a. Sea level change, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 1137–1216, eds Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., & Midgley, P., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., Merrifield, M. A., Milne, G. A., Nerem, R. S., Nunn, P. D., et al., 2013b. Sea-level rise by 2100, *Science*, **342**(6165), 1445–1445.

Cutler, K., Edwards, R., Taylor, F., Cheng, H., Adkins, J., Gallup, C., Cutler, P., Burr, G., & Bloom, A., 2003. Rapid sea-level fall and deep-ocean temperature change since the last interglacial period, *Earth and Planetary Science Letters*, **206**(3), 253–271.

de Boer, B., van de Wal, R., Lourens, L., Bintanja, R., & Reerink, T., 2013. A continuous simulation

of global ice volume over the past 1 million years with 3-d ice-sheet models, *Climate Dynamics*, **41**(5-6), 1365–1384.

Dutton, A. & Lambeck, K., 2012. Ice volume and sea level during the last interglacial, *Science*, **337**, 216–219.

Dziewonski, A. M. & Anderson, D. L., 1981. Preliminary reference earth model, *Phys. Earth Planet. Int.*, **25**, 297–356.

Elderfield, H., Ferretti, P., Greaves, M., Crowhurst, S., McCave, I., Hodell, D., & Piotrowski, A., 2012. Evolution of ocean temperature and ice volume through the mid-pleistocene climate transition, *Science*, **337**(6095), 704–709.

Ganopolski, A. & Robinson, A., 2011. Palaeoclimate: The past is not the future, *Nature Geoscience*, **4**, 661–663.

Grant, K. M., Rohling, E. J., Bar-Matthews, M., Ayalon, A., Medina-Elizalde, M., Bronk Ramsey, C., Satow, C., & Roberts, A. P., 2012. Rapid coupling between ice volume and polar temperature over the past 150,000 years, *Nature*, **491**, 744–747.

Horton, B., Larcombe, P., Woodroffe, S., Whittaker, J., Wright, M., & Wynn, C., 2003. Contemporary foraminiferal distributions of a mangrove environment, great barrier reef coastline, australia: implications for sea-level reconstructions, *Marine Geology*, **198**(3–4), 225 – 243.

Jevrejeva, S., Grinsted, A., & Moore, J., 2014. Upper limit for sea level projections by 2100, *Environmental Research Letters*, **9**(10), 104008.

Kemp, A. C., Horton, B. P., Culver, S. J., Corbett, D. R., van de Plassche, O., Gehrels, W. R., Douglas, B. C., & Parnell, A. C., 2009. Timing and magnitude of recent accelerated sea-level rise (north carolina, united states), *Geology*, **37**(11), 1035–1038.

Kendall, R. A., Mitrovica, J. X., & Milne, G. A., 2005. On post-glacial sea level – II. Numerical formulation and comparative results on spherically symmetric models, *Geophys. J. Int.*, **161**, 679–706.

Kopp, R. E., Simons, F. J., Mitrovica, J. X., Maloof, A. C., & Oppenheimer, M., 2009. Probabilistic assessment of sea level during the last interglacial stage, *Nature*, **462**, 863–867.

Lambeck, K., Purcell, A., & Dutton, A., 2012. The anatomy of interglacial sea levels: The relationship between sea levels and ice volumes during the last interglacial, *Earth and Planetary Science Letters*, **315–316**, 4–11.

Lambeck, K., Rouby, H., Purcell, A., Sun, Y., & Sambridge, M., 2014. Sea level and global ice volumes from the last glacial maximum to the holocene, *Proceedings of the National Academy of Sciences of the United States of America*, **111**(43), 15296–15303.

Lin, L., Khider, D., Lisiecki, L. E., & Lawrence, C. E., 2014. Probabilistic sequence alignment of stratigraphic records, *Paleoceanography*.

Lisiecki, L. E. & Raymo, M. E., 2005. A pliocene-pleistocene stack of 57 globally distributed benthic $\delta^{18}$O records, *Paleoceanography*, **20**(PA1003).

Medina-Elizalde, M., 2012. A global compilation of coral sea-level benchmarks: implications and new challenges, *Earth and Planetary Science Letters*, **362**, 310–318.

Mitrovica, J. & Peltier, W., 1992. A comparison of methods for the inversion of viscoelastic relaxation spectra, *Geophys. J. Int.*, **108**, 41–414.

Mitrovica, J. X., Wahr, J., Matsuyama, I., & Paulson, A., 2005. The rotational stability of an ice-age earth, *Geophys. J. Int.*, **161**, 491–506.

Peltier, W., 1974. The impulse response of a Maxwell Earth, *Rev. Geophys. Space Phys.*, **12**, 649–669.

Peltier, W. & Fairbanks, R., 2006. Global glacial ice volume and Last Glacial Maximum duration from an extended Barbados sea level record, *Quaternary Science Reviews*, **25**(23-24), 3322–3337, Critical Quaternary Stratigraphy.

Peltier, W. R., 1996. Mantle viscosity and ice-age ice sheet topography, *Science*, **273**(5280), 1359–1364.

Peltier, W. R., 2004. Global glacial isostacy and the surface of the ice-age earth: The ice-5g (vm2) model and grace, *Annual Review of Earth and Planetary Sciences*, **32**, 111–149.

Petersen, A. C., 2000. Philosophy of climate science, *Bulletin of the American Meteorological Society*, **81**(2), 265–271.

Potter, E.-K. & Lambeck, K., 2003. Reconciliation of sea-level observations in the Western North Atlantic during the last glacial cycle, *Earth Planet. Sci. Lett.*, **217**, 171–181.

Raymo, M. E. & Mitrovica, J. X., 2012. Collapse of polar ice sheets during the stage 11 interglacial, *Nature*, **483**, 453–455.

Rohling, E. J., Grant, K., Bolshaw, M., Roberts, A. P., Siddall, M., Hemleben, C., & Kucera, M., 2009. Antarctic temperature and global sea level closely coupled over the past five glacial cycles, *Nature Geoscience*, pp. 1–5.

Schölzel, C. & Hense, A., 2011. Probabilistic assessment of regional climate change in southwest germany by ensemble dressing, *Climate Dynamics*, **36**, 2003–2014.

Shackleton, N. J., 1987. Oxygen isotopes, ice volume and sea level, *Quaternary Science Reviews*, **6**, 183–190.

Shackley, S., Young, P., Parkinson, S., & Wynne, B., 1998. Uncertainty, complexity and concepts of good science in climate change modelling: Are gcms the best tools?, *Climatic Change*, **38**(2), 159–205.

Siddall, M., Stocker, T. F., & Clark, P. U., 2009. Constraints on future sea-level rise from past sea-level change, *Nature Geoscience*, **2**, 571 – 575.

Stanford, J. R., Hemingway, R., Rohling, E., Challenor, P., Medina-Elizalde, M., & Lester, J., 2011. Sea-level probability for the last deglaciation: a statistical analysis of far-field records, *Global and Planetary Change*, **79**, 193–203.

Steffen, H. & Wu, P., 2011. Glacial isostatic adjustment in Fennoscandia – A review of data and modeling, *Journal of Geodynamics*, **52**(3 – 4), 169 – 204.

Tarasov, L., Dyke, A. S., Neal, R. M., & Peltier, W. R., 2012. A data-calibrated distribution of deglacial chronologies for the north american ice complex from glaciological modeling, *Earth and Planetary Science Letters*, **315–316**, 30–40.

Tebaldi, C. & Knutti, R., 2007. The use of the multi-model ensemble in probabilistic climate projections, *Phil. Trans. R. Soc. A*, **365**(1857), 2053–2075.

Tzedakis, P. C., Raynaud, D., McManus, J. F., Berger, A., Brovkin, V., & Kiefer, T., 2009. Interglacial diversity, *Nature Geoscience*, **2**, 751–755.

van Leeuwen, P. J., 2003. A variance-minimizing filter for large-scale applications, *Monthly Weather Review*, **131**, 2071–2084.

Waelbroeck, C., Labeyrie, L., Michel, E., Duplessy, J. C., McManus, J. F., Lambeck, K., Balbon, E., & Labracherie, M., 2002. Sea-level and deep water temperature changes derived from benthic foraminifera isotopic records, *Quaternary Science Reviews*, **21**, 295–305.

Wilks, D. S., 2011. *Statistical methods in the atmospheric sciences*, Academic Press, 3rd edn.

Zwiers, F. W. & Von Storch, H., 2004. On the role of statistics in climate research, *International Journal of Climatology*, **24**(6), 665–680.
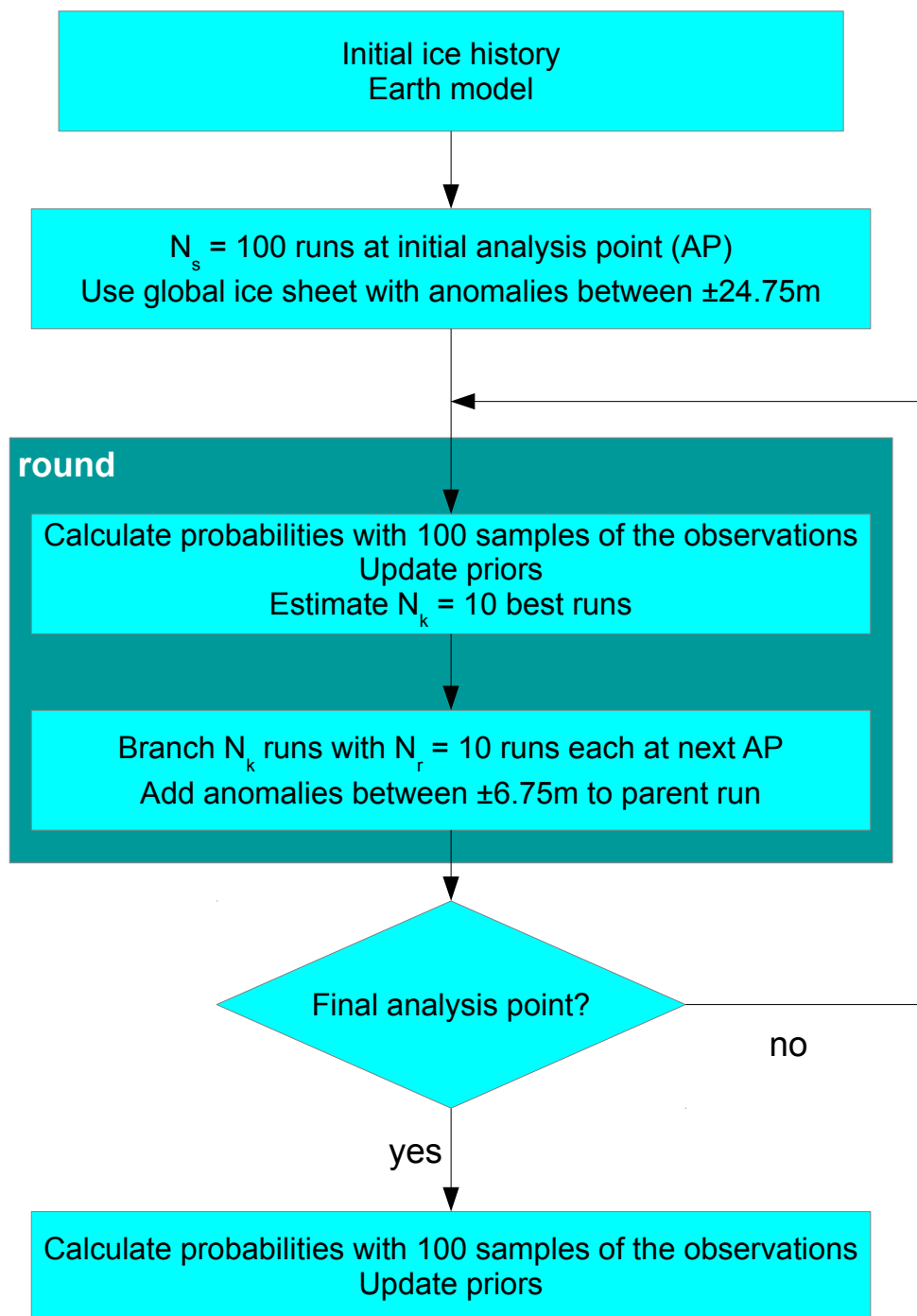
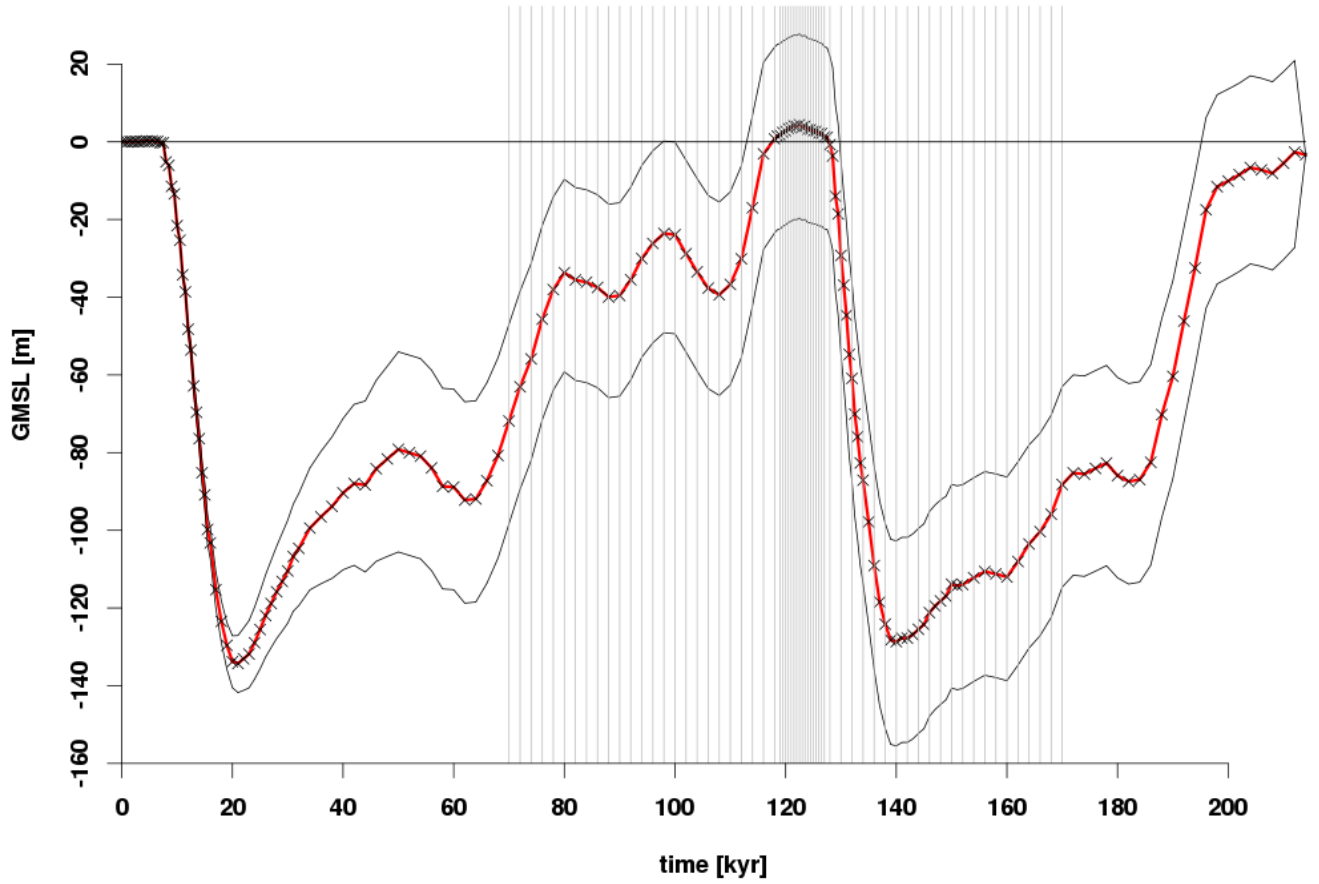**Figure 1.** Sketch of the data assimilation procedure.

**Figure 2.** Base modified-$\delta^{18}O$ ice volume history (red centre line) and the maximum and minimum ice volume history of the first branching (anomaly of -24.75 m and 24.75 m of eustatic sea level, black outer lines). Also shown are the analysis points (grey vertical lines) and calculations points (black crosses) over the analysis period.
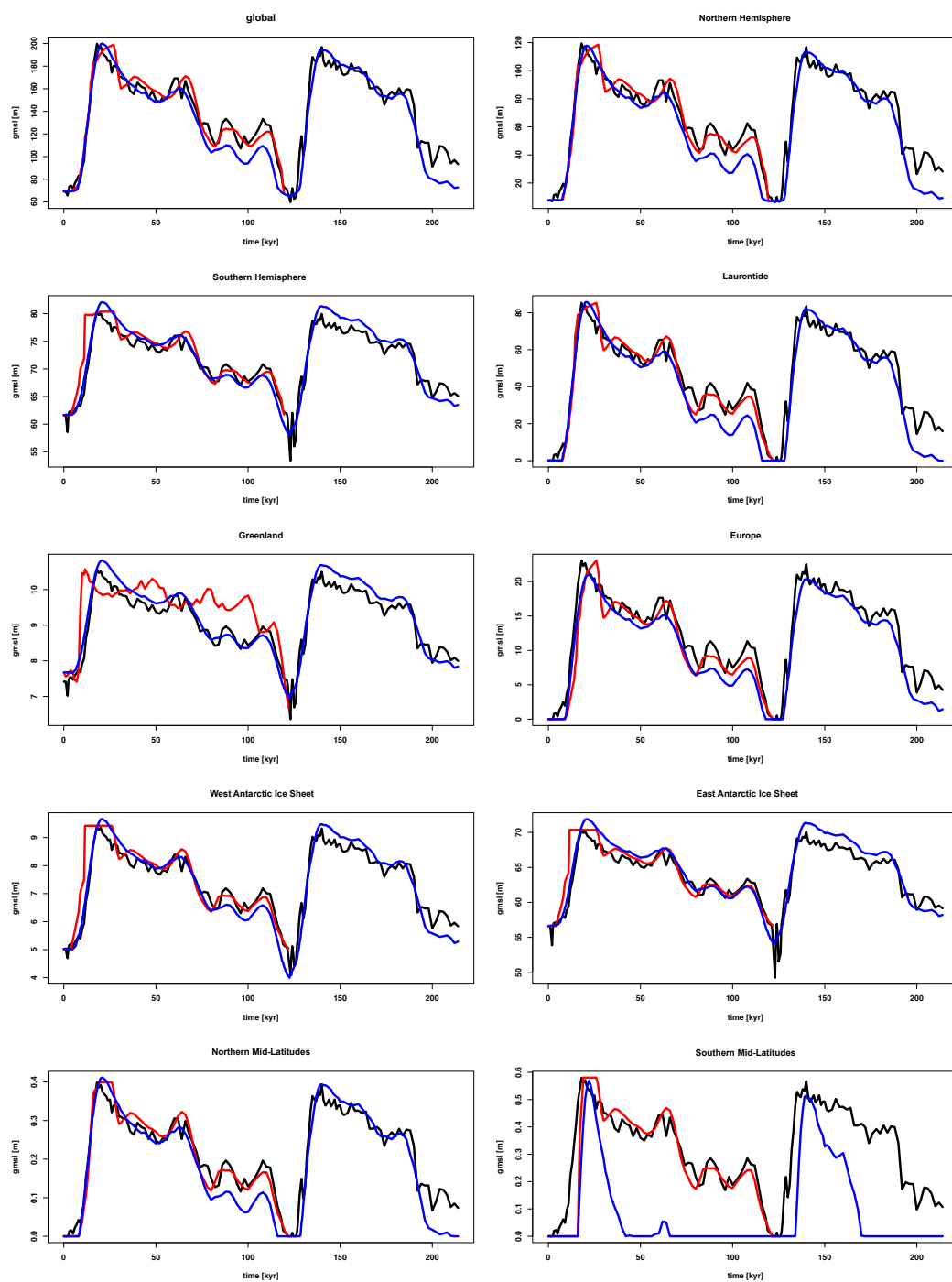
**Figure 3.** Initial regional ice sheet volumes. Shown are the $\delta^{18}O$ volumes in black, the modified-$\delta^{18}O$ volumes in blue and, for comparison, the ICE-5G ice sheet volumes in red. The first panel shows the sum (i.e. global total) of all of the regional ice sheet volumes.
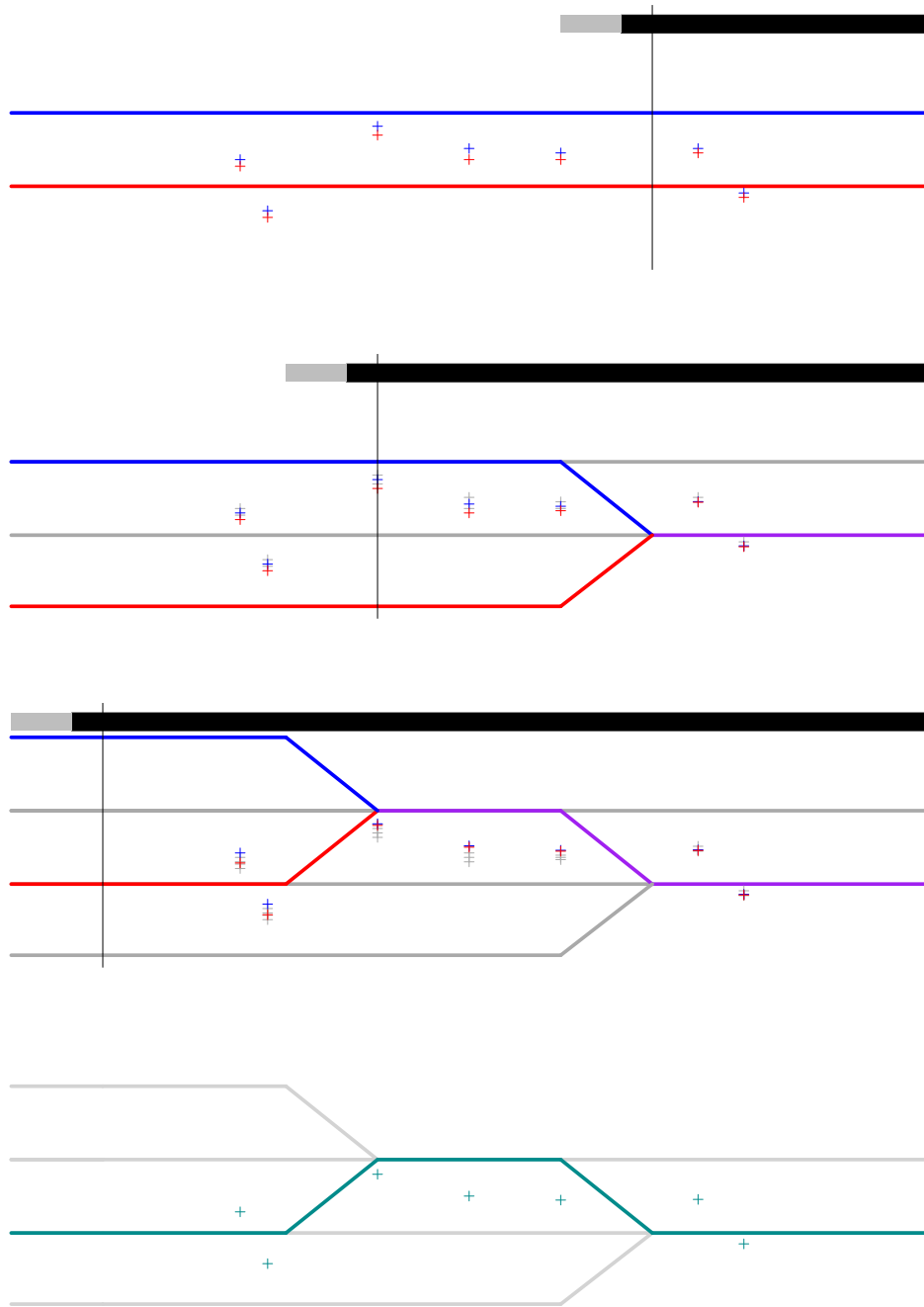
**Figure 4.** Simplified illustration of the data assimilation scheme, with each panel resenting a further step in the analysis. The red and blue lines represent different anomalies added to the base global ice volume, the grey lines are anomalies that were not branched, and the purple lines indicated a shared history between the red and blue anomalies. The green line in the last panel represents the best fitting anomaly. The plus signs (colours matching the respective anomaly) represent the GIA-corrected observations in position relative to corresponding anomaly. The vertical lines represent the analysis point. The thick line above each panel represents the weighing of the observations with black being a full weighting and grey being a partial weighing. The direction of time is the same as in proceeding figures, with present day on the left and further back in time on the right.
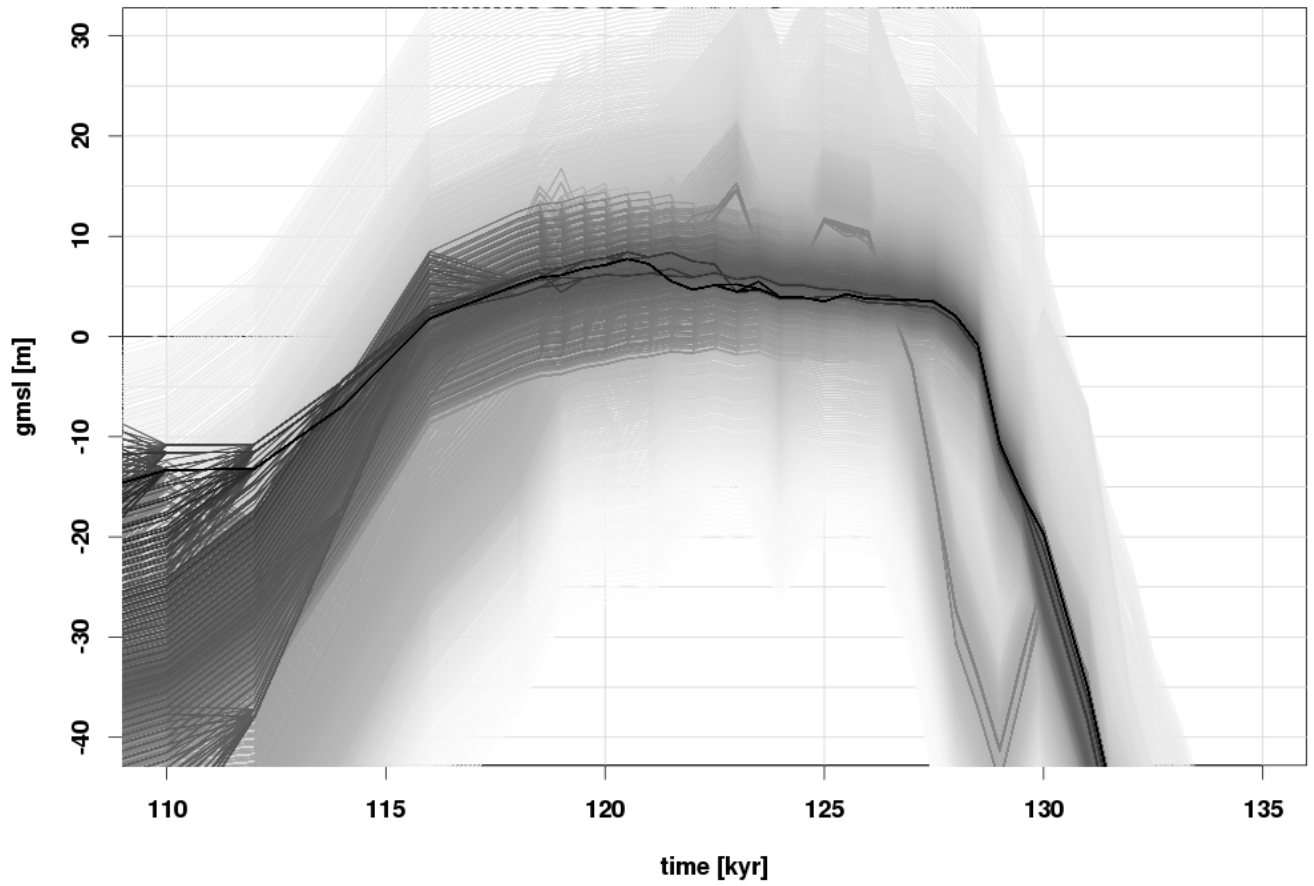
**Figure 5.** Global mean sea-level curves for all 39,000 runs during the LIG. Shading indicates the posterior of each curve calculated with the uniform prior with the original, unsampled observations. The darker greys are the most probable.
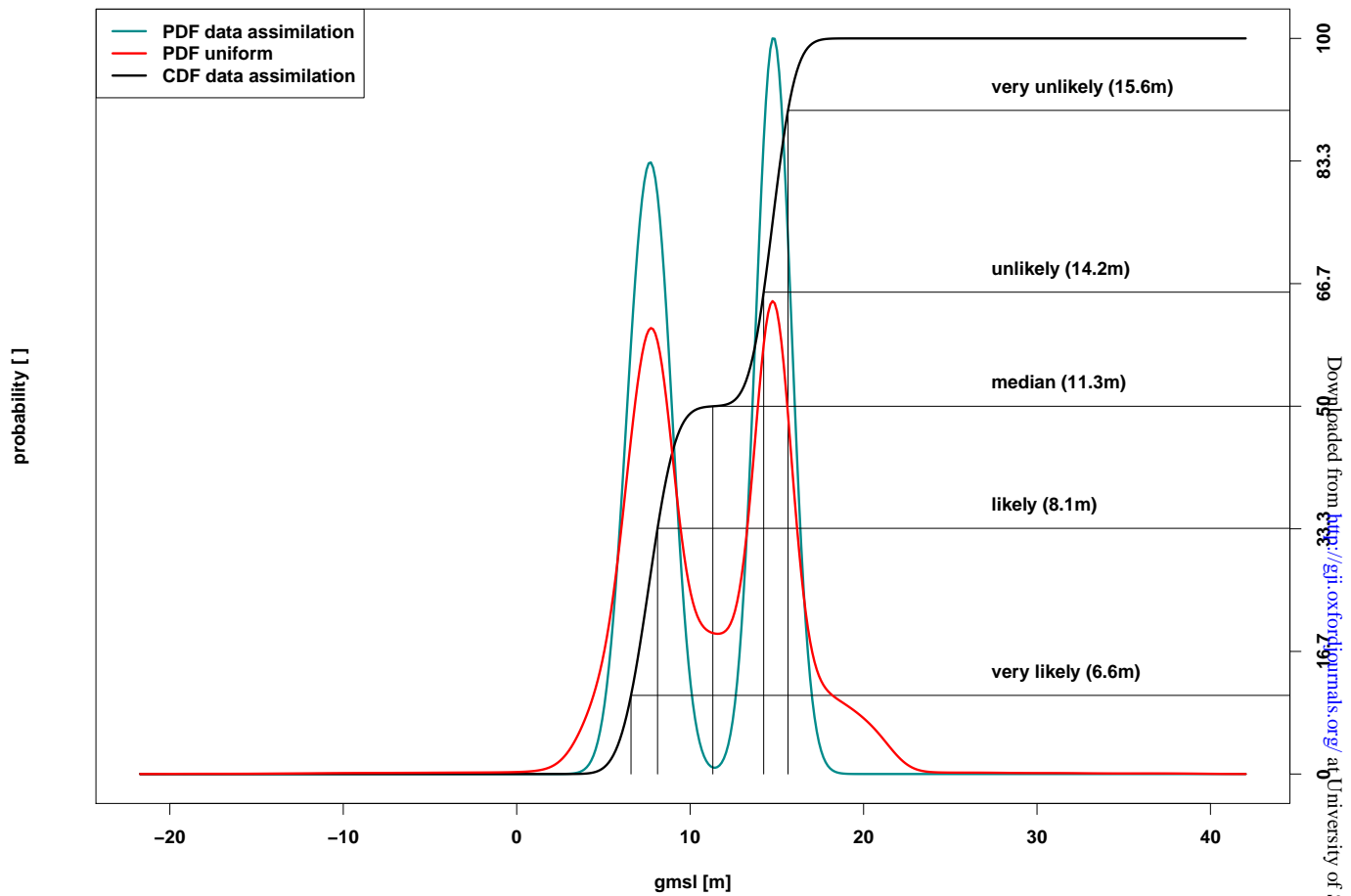
**Figure 6.** Estimation of the maximum highstand during the LIG. The results are calculated using Gaussian mixture modelling (GMM) based on the posterior with the application of either the uniform prior or the one used within the data assimilation. The green curve shows the probability density function (pdf) of the posterior of the prior used within the data assimilation and the black curve shows the corresponding cumulative distribution function (cdf). The red curve shows the pdf of the posterior of the uniform prior.
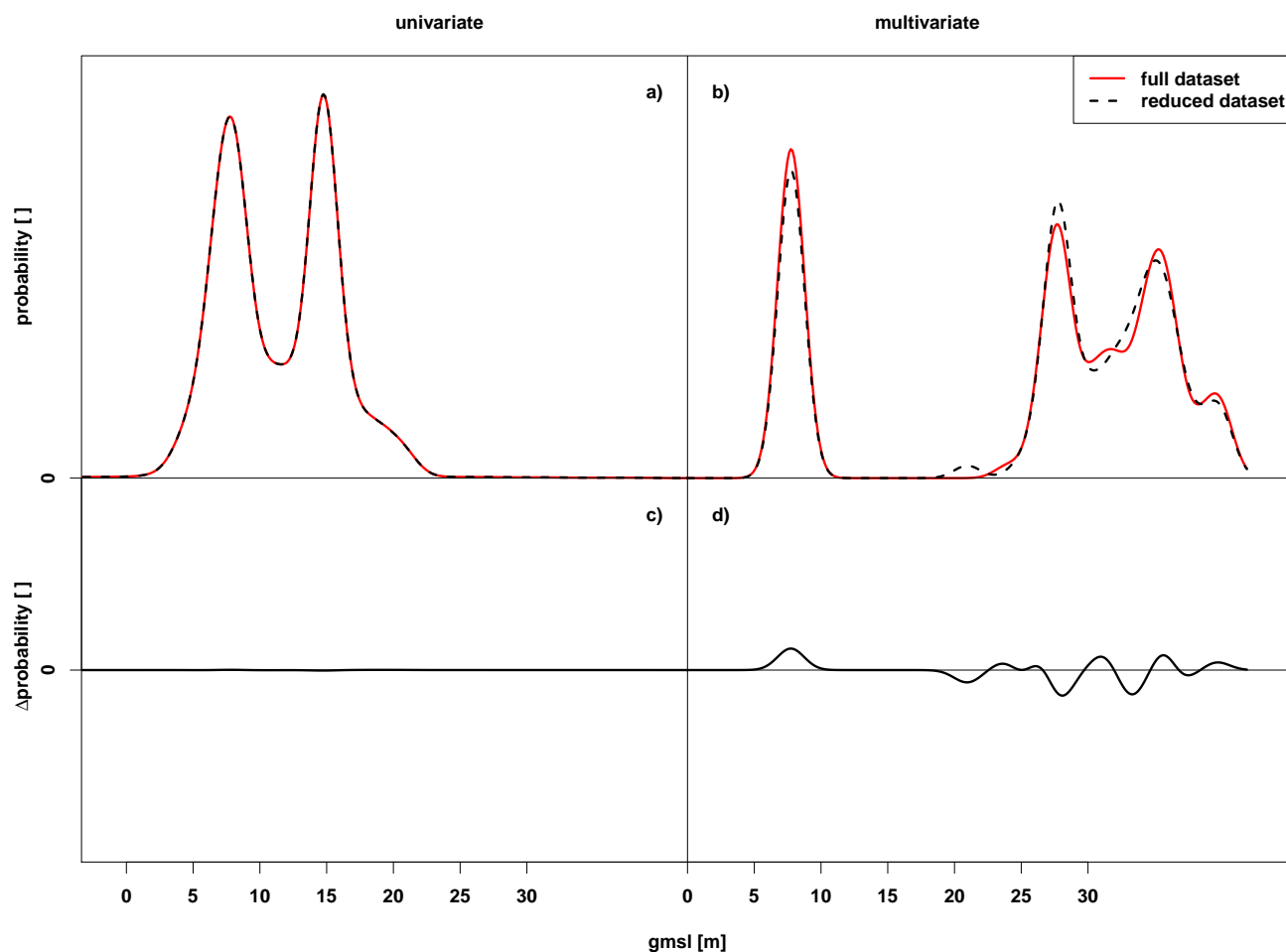
**Figure 7.** Estimation of the highstand during the LIG using the univariate and multivariate methods for calculating the probability showing the impact of outliers. (a) Comparison for the univariate method between the GMM results with the uniform prior for the full data set (red solid) and the reduced dataset (black dashed), which does not include the observations excluded in Kopp et al. (2009). (b) The same comparison for the multivariate method. (c) Difference of the full dataset minus reduced dataset for the univariate results in (a). (d) Difference of the multivariate results in (b).
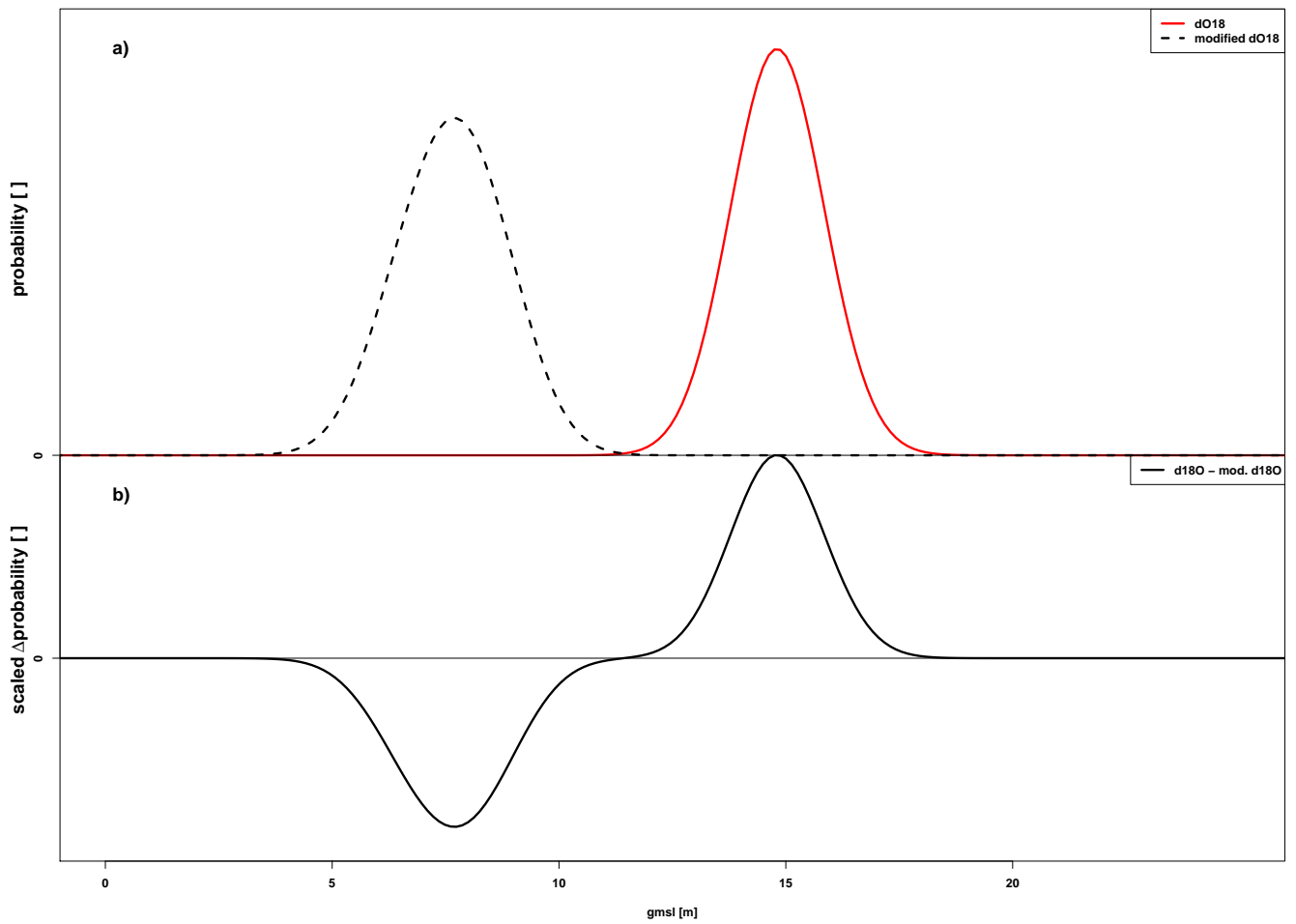
**Figure 8.** Estimation of the highstand for the LIG for the two different initial ice histories. (a) pdf of the runs using either the $\delta^{18}O$ (red solid line) or the modified-$\delta^{18}O$ (black dashed line) ice history. (b) The difference ($\delta^{18}O$ pdf minus modified-$\delta^{18}O$ pdf) is scaled down by 33% for graphical purposes.
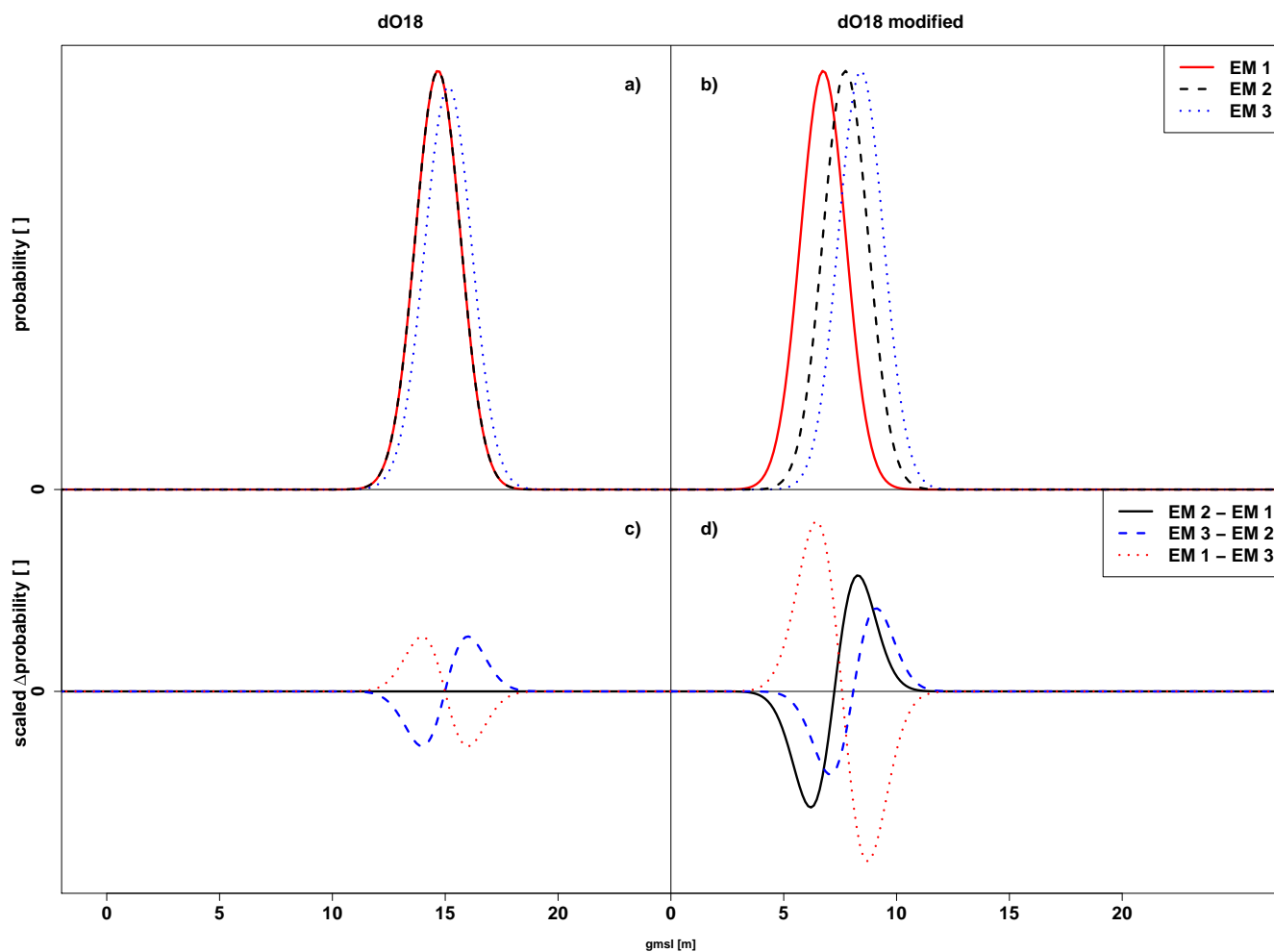
**Figure 9.** Estimation of the highstand for the LIG for the three different Earth models. (a) Maximum GMSL for E1 (red solid line), E2 (black dashed line) and E3 (blue dotted line) for the runs based on the $\delta^{18}O$ ice history. (b) Same as a, but based on the (b) Difference between the modified-$\delta^{18}O$. (c) and (d) show the scaled differences between the curves: E2-E3 (black solid line), E3-E1 (blue dashed line), E1-E2 (red doted line).
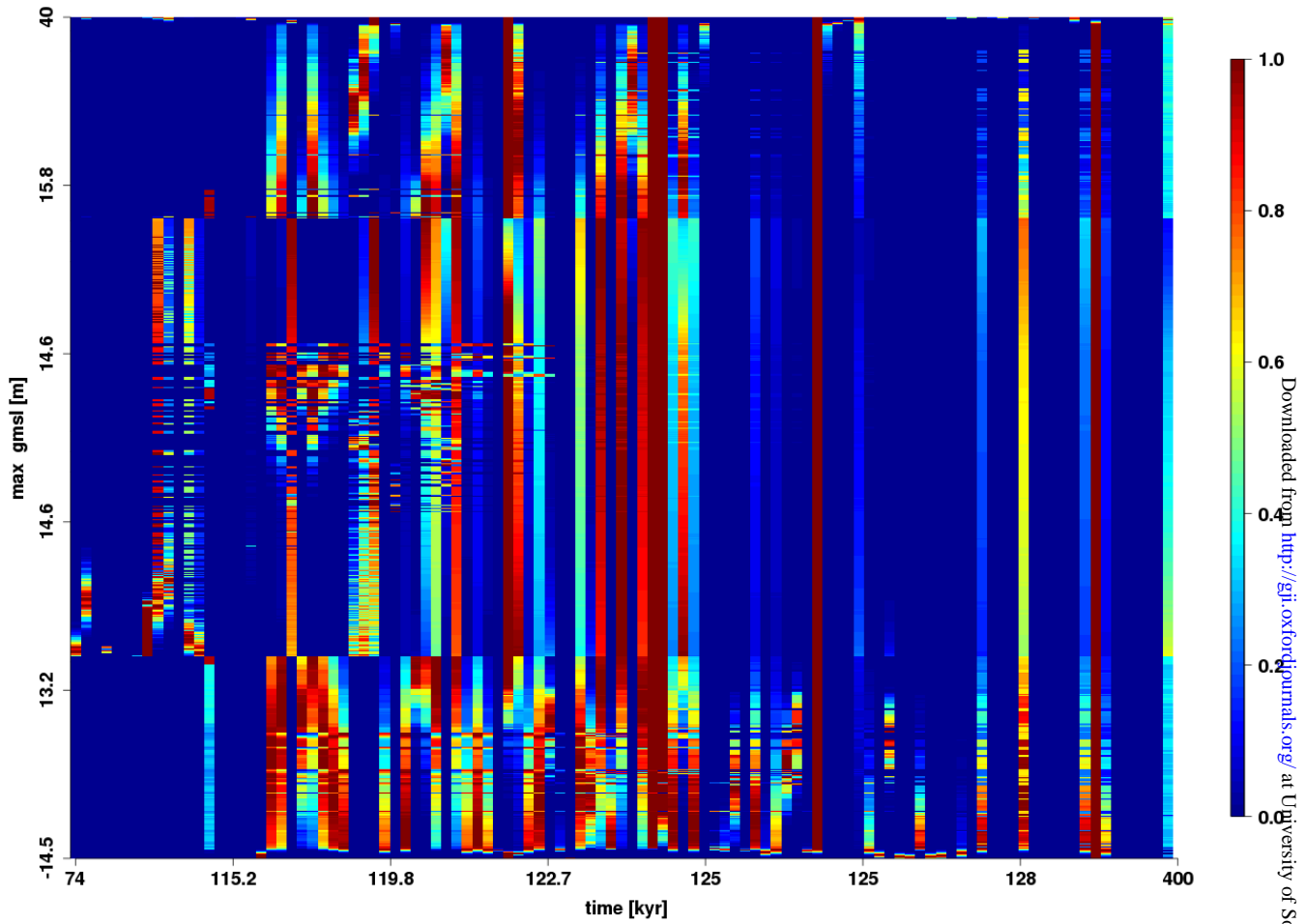
**Figure 10.** Analysis of observations for the $\delta^{18}O$ initial ice history and the E1 Earth model. Each run is represented by one row and each column represents one observation. The rows are sorted by the maximum highstand of each run, and the columns are sorted by the age estimate of the observation. Note that the resulting axes are nonlinear. The colours indicate the probability each run gets assigned by the associated observation.
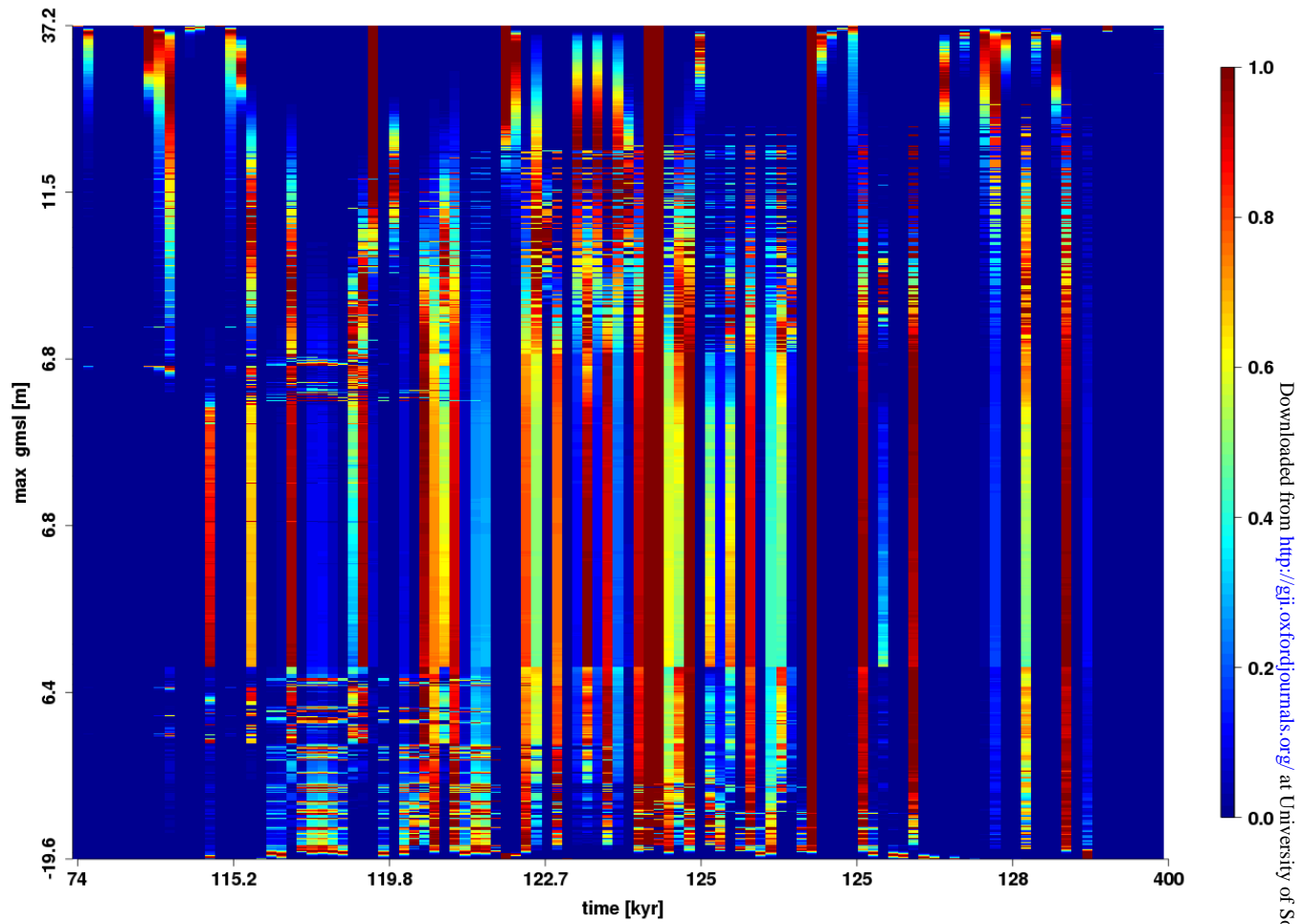
**Figure 11.** Analysis of observations for the modified-$\delta^{18}O$ initial ice history and the E1 Earth model. Description of the plot is the same as Fig. 10

**Figure 12.** Most probable run (red) of those based on the modified-$\delta^{18}O$ ice history using the data-assimilation prior. The boxes indicate the one standard deviation uncertainty, and the blue observations are the Red Sea data.
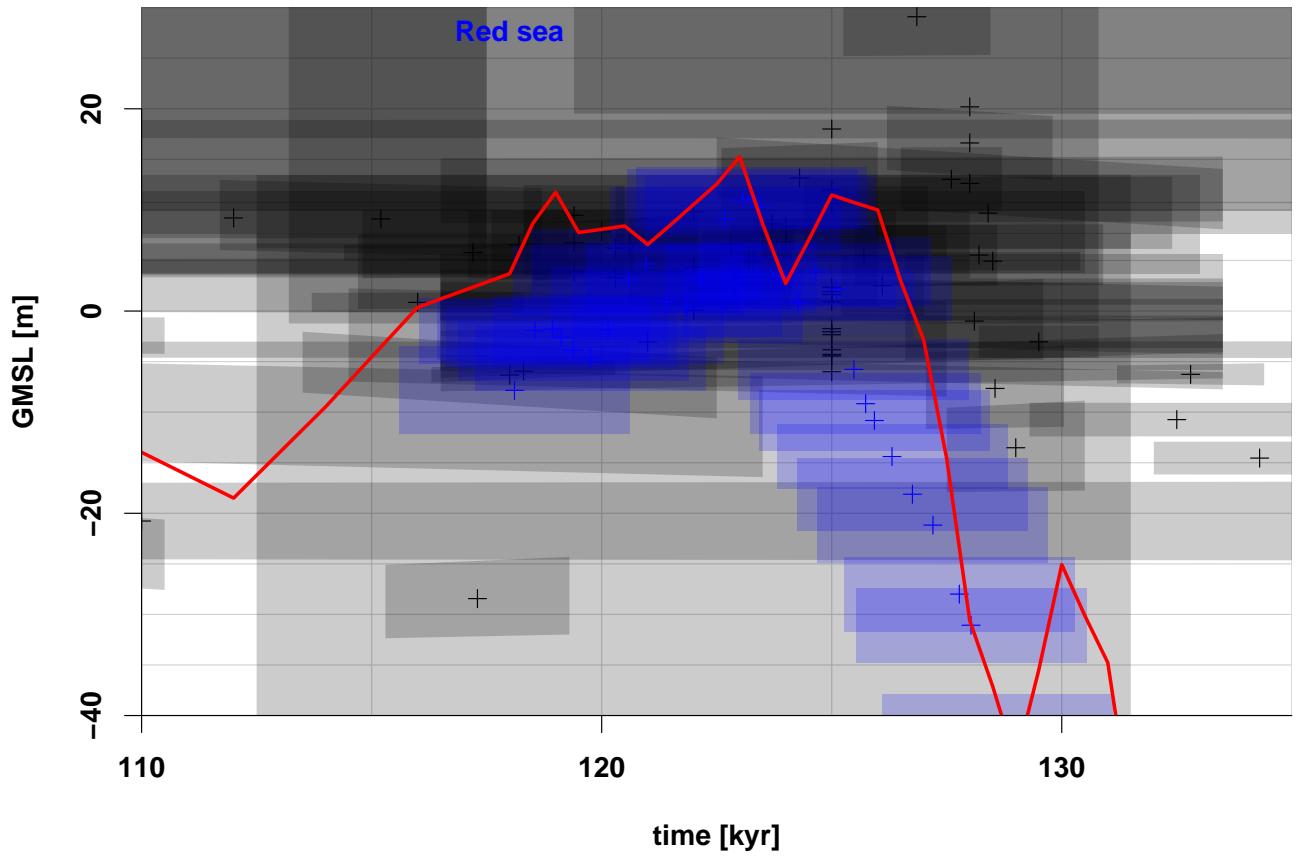
**Figure 13.** Most probable run (red) of those based on the modified-$\delta^{18}O$ ice history using the data-assimilation prior. The boxes indicate the one standard deviation uncertainty, and the blue observations are the Red Sea data.
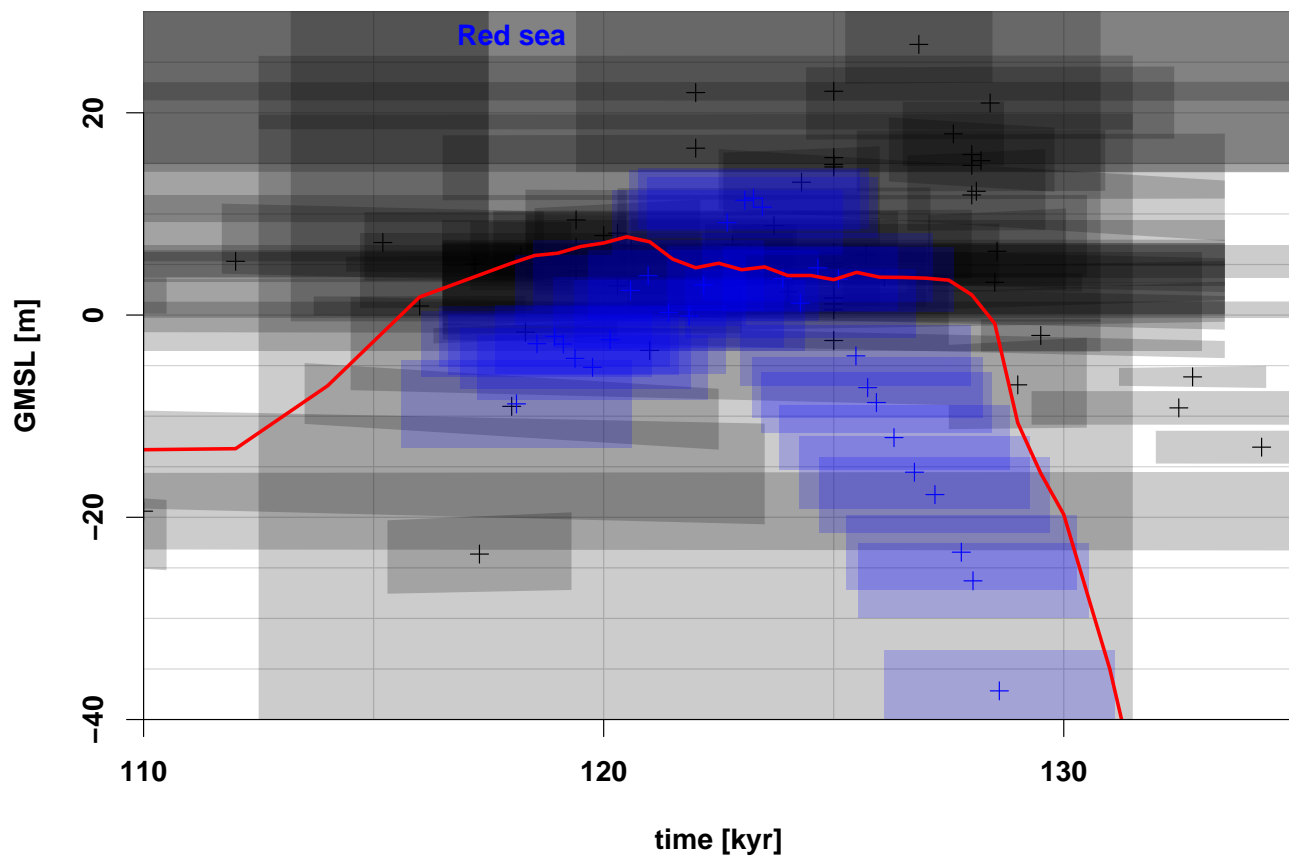
**Table 1.** Lithospheric thickness and viscosity values for the three Earth models.

| Earth Model | LT [km] | $\nu_{UM}$ [Pa s] | $\nu_{LM}$ [Pa s] |
|---|---|---|---|
| EM1 | 71 | $2 \cdot 10^{20}$ | $10 \cdot 10^{21}$ |
| EM2 | 96 | $5 \cdot 10^{20}$ | $5 \cdot 10^{21}$ |
| EM3 | 90 | VM2[1] | VM2[1] |

[1]For comparison, the an upper mantle viscosity of $5 \cdot 10^{20}$ Pa s and a lower mantle viscosity of $2.5 \cdot 10^{21}$ Pa s could be used as a two-layer approximation of VM2.

## APPENDIX A: DERIVATION OF BASIC EQUATION

The expanded probability on the right hand side of eq. 2 can be rewritten using the Bayesian eq. 1 (Campbell 2005):

$$p\left(\vec{f}_j, \vec{\eta}_i | \vec{O}\right) = \frac{p\left(\vec{O}|\vec{f}_j\right) p\left(\vec{f}_j|\vec{\eta}_i\right) p\left(\vec{\eta}_i\right)}{\sum\limits_{l\in\mathcal{L}} \sum\limits_{k\in\mathcal{J}_l} p\left(\vec{O}|\vec{f}_k\right) p\left(\vec{f}_k|\vec{\eta}_l\right) p\left(\vec{\eta}_l\right)}. \tag{A.1}$$

The index set $\mathcal{L}$ represents all of the investigated parameter sets, and $\mathcal{J}_l$ is the index set of the model solutions $\vec{f}_k$.

Because the GIA model is deterministic, its result, $\vec{f}$, is uniquely defined by the input parameters $\vec{\eta}$:

$$\mathcal{M}\left(\vec{\eta}_i\right) = \vec{f}_i. \tag{A.2}$$

Integrating over all possible model results, this leads to a Kronecker-Delta function, $\delta\left(\vec{f}_j - \mathcal{M}\left(\vec{\eta}_i\right)\right)$. Inserting it and eq. A.1 into eq. 2, we obtain

$$p\left(\vec{\eta}_i | \vec{O}\right) = \int\limits_{j\in\mathcal{J}} \left[p\left(\vec{O}|\vec{f}_j\right) \delta\left(\vec{f}_j - \mathcal{M}\left(\vec{\eta}_i\right)\right)\right] df_j \cdot \frac{p\left(\vec{\eta}_i\right)}{\sum\limits_{l\in\mathcal{L}} \sum\limits_{k\in\mathcal{J}_l} p\left(\vec{O}|\vec{f}_k\right) p\left(\vec{\eta}_l\right)}. \tag{A.3}$$

After execution of the $\delta$ function, this equation reduces to eq. 3.

## APPENDIX B: DERIVATION OF THE PRIOR

The difficulty in deriving the prior is connecting the former round $(n-1)$ with the new round $(n)$ when the seed run is branched with $N_r$ new runs. For both rounds Eq. 3 is valid for the seed run. Thus, we evaluate the equation for the former round and solve for the prior:

$$p\left(\vec{\eta}_i\right)^{n-1} = \frac{p\left(\vec{\eta}_i | \vec{O}\right)^{n-1} \left[\sum\limits_{l\in\mathcal{L}} p\left(\vec{O}|\mathcal{M}\left(\vec{\eta}_l\right)\right)^{n-1} p\left(\vec{\eta}_l\right)^{n-1}\right]}{p\left(\vec{O}|\mathcal{M}\left(\vec{\eta}_i\right)\right)^{n-1}}. \tag{B.1}$$

Next, we set the sum to 1, as this is simply a normalisation. We also define the prior of the new round as equal to the prior of the former round. Given that we want a likelihood that takes into account the changing fits to sea-level at past time points, we set the likelihood of the parent run in the former round equal with the likelihood of the parent run and its $N_r$ descendants. We assume that the prior is distibuted equal between the parent and the descendants, we take the $(N_r+1)^{\text{th}}$ root of the posterior divided by the product of the liklihoods. As a consequence we derive Eq. 9.

## APPENDIX C: MAXIMUM SEA LEVEL DURING THE LIG FOR THE DIFFERENT EXPERIMENTS

**Table A1.** Results of the different analyses.

| Experiment | Section | 10th perc. [m] very likely | 33rd perc. [m] likely | 50th perc. [m] median | 66th perc. [m] unlikely | 90th perc. [m] very unlikely | IQR[1] [m] |
|---|---|---|---|---|---|---|---|
| Uniform prior | | | | | | | |
| full dataset | 3.2 | 6.1 | 8.4 | 11.2 | 14.1 | 16.7 | 7.3 |
| reduced dataset | 3.2 | 6.1 | 8.4 | 11.2 | 14.1 | 16.7 | 7.3 |
| full dataset (mv)[2] | 3.2 | 7.4 | 26.6 | 29.2 | 33.1 | 37.2 | 25.3 |
| reduced dataset (mv)[2] | 3.2 | 7.5 | 26.9 | 29 | 33 | 37.1 | 23.6 |
| $\delta^{18}O$ | | 10.6 | 13.9 | 14.6 | 15.3 | 17.8 | 2.5 |
| modified-$\delta^{18}O$ | | 5.4 | 7.2 | 7.9 | 8.9 | 13.9 | 3.5 |
| E1, $\delta^{18}O$ | | 10 | 13.1 | 14.6 | 15.9 | 19.1 | 4.6 |
| E2, $\delta^{18}O$ | | 9.9 | 13.1 | 14.6 | 15.9 | 19.1 | 4.5 |
| E3, $\delta^{18}O$ | | 9.3 | 13.1 | 14.6 | 16 | 18.9 | 4.7 |
| E1, modified-$\delta^{18}O$ | | 3.5 | 6 | 7.5 | 8.9 | 13.6 | 4.8 |
| E2, modified-$\delta^{18}O$ | | 4.2 | 6.9 | 8.5 | 11.2 | 15.3 | 5.1 |
| E3, modified-$\delta^{18}O$ | | 4.6 | 7.3 | 8.8 | 10.3 | 14.7 | 4.7 |
| Original prior | | | | | | | |
| full dataset | 3.2 | 6.6 | 8.1 | 11.3 | 14.2 | 15.6 | 7.2 |
| $\delta^{18}O$ | 3.3 | 13.5 | 14.3 | 14.7 | 15.2 | 16.2 | 1.4 |
| modified-$\delta^{18}O$ | 3.3 | 6 | 7 | 7.5 | 8.1 | 9.1 | 1.7 |
| E1, $\delta^{18}O$ | 3.4 | 13.3 | 14.2 | 14.6 | 15 | 15.9 | 1.3 |
| E2, $\delta^{18}O$ | 3.4 | 13.3 | 14.2 | 14.5 | 15 | 15.9 | 1.3 |
| E3, $\delta^{18}O$ | 3.4 | 13.8 | 14.6 | 15 | 15.5 | 16.4 | 1.4 |
| E1, modified-$\delta^{18}O$ | 3.4 | 5.4 | 6.2 | 6.7 | 7.1 | 7.9 | 1.4 |
| E2, modified-$\delta^{18}O$ | 3.4 | 6.4 | 7.2 | 7.7 | 8.1 | 9 | 1.4 |
| E3, modified-$\delta^{18}O$ | 3.4 | 7.1 | 7.9 | 8.4 | 8.7 | 9.6 | 1.3 |

Notes: [1]inter quartile range. [2]multivariate analysis.