



RESEARCH ARTICLE

10.1029/2019JF005237

Dario Domingo and Irene Malmierca-Vallet are joint first authors of this article.

Using Ice Cores and Gaussian Process Emulation to Recover Changes in the Greenland Ice Sheet During the Last Interglacial

Dario Domingo^{1,2} , Irene Malmierca-Vallet³ , Louise Sime³ , Jochen Voss¹ , and Emilie Capron^{3,4} 

¹School of Mathematics, University of Leeds, Leeds, UK, ²Department of Mathematical Sciences, Durham University, Durham, UK, ³British Antarctic Survey, Cambridge, UK, ⁴Physics of Ice, Climate and Earth, Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

Key Points:

- Isotope data from ice cores allow inference about the Greenland ice sheet during the Last Interglacial
- A Gaussian process emulator reduces computational cost and makes data-model comparison possible
- A scenario-based approach allows a glimpse at the improvements possible with better constrained data

Supporting Information:

- Supporting Information S1

Correspondence to:

D. Domingo and I. Malmierca-Vallet, dario.domingo@durham.ac.uk; irealm37@bas.ac.uk

Citation:

Domingo, D., Malmierca-Vallet, I., Sime, L., Voss, J., & Capron, E. (2020). Using ice cores and Gaussian process emulation to recover changes in the Greenland Ice Sheet during the Last Interglacial. *Journal of Geophysical Research: Earth Surface*, 125, e2019JF005237. <https://doi.org/10.1029/2019JF005237>

Received 4 JUL 2019

Accepted 30 MAR 2020

Accepted article online 20 APR 2020

©2020. The Authors.

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

Abstract The shape and extent of the Greenland Ice Sheet (GIS) during the Last Interglacial (LIG) is a matter of controversy, with different studies proposing a wide range of reconstructions. Here, for the first time, we combine stable water isotopic information from ice cores with isotope-enabled climate model outputs to investigate the problem. Exploring the space of possible ice sheet geometries by simulation is prohibitively expensive. We address this problem by using a Gaussian process emulator as a statistical surrogate of the full climate model. The emulator is calibrated using the results of a small number of carefully chosen simulations and then permits fast, probabilistic predictions of the simulator outputs at untried inputs. The inputs are GIS morphologies, parameterized through a dimension-reduction technique adapted to the spherical geometry of the setting. Based on the emulator predictions, the characteristics of morphologies compatible with the available ice core measurements are explored, leading to a reduction in uncertainty on the LIG GIS morphology. Moreover, a scenario-based approach allows to assess the gains in uncertainty reduction which would result from the availability of better dated LIG measurements in Greenland ice cores.

1. Introduction

The extent and shape of the Greenland ice sheet (GIS) during the Last Interglacial (LIG; 129 to 116 ka) is a matter of some controversy. Mean global sea level is thought to have been increased by between 6 and 9 m compared to present-day level during this time period (e.g., Dutton et al., 2015; Kopp et al., 2009), with a likely contribution from both the Antarctic (e.g., DeConto & Pollard, 2016; Sutter et al., 2016) and the GISs (e.g., Fyke et al., 2011; Helsen et al., 2013; Lhomme et al., 2005; Tarasov & Peltier, 2003). The GIS contribution to the LIG highstand has however been difficult to pin down: Studies suggest that it may have contributed anywhere between +0.3 and +5.5 m to global mean sea level (Born & Nisancioglu, 2012; Cuffey & Marshall, 2000; Helsen et al., 2013; Otto-Bliesner et al., 2006; Quiquet et al., 2013; Robinson et al., 2011; Stone et al., 2013; Tarasov & Peltier, 2003; Yau et al., 2016; NEEM community members, 2013). Possible drivers for these changes include stronger summertime insolation at high northern latitudes, and Arctic land summer temperatures 4–5 °C higher than those that occurred during the preindustrial (e.g., CAPE Last Interglacial Project Members, 2006).

Evidence proximal to Greenland about the LIG comes from sediment discharged from Greenland and from deep ice cores (e.g., Colville et al., 2011; NGRIP Project Members, 2004; NEEM community members, 2013). From the sedimentary evidence from cores drilled off the Greenland coast it is difficult to produce an assessment of the extent and shape of the LIG GIS (Colville et al., 2011). Ice sections corresponding to the LIG have been found, near the bed, at up to seven Greenland ice sites: NEEM, NGRIP, GRIP, Renland, DYE3, GISP2, and Camp Century (CC) (Johnsen & Vinther, 2007). See Figure 1 for location of the sites. The suggestion of LIG ice in every one of these cores hinges on the occurrence of high stable water isotope ($\delta^{18}\text{O}$) measurements of the ice in the deepest parts of each core. At each of these sites, measurements from near the bottom show $\delta^{18}\text{O}$ values that are above the range that occurred during the better dated most recent 120 ka (e.g., Johnsen & Vinther, 2007; NEEM community members, 2013). Climate model simulations of the LIG, forced by appropriate greenhouse gas and orbital changes alone, have tended to fail to capture these observed $\delta^{18}\text{O}$

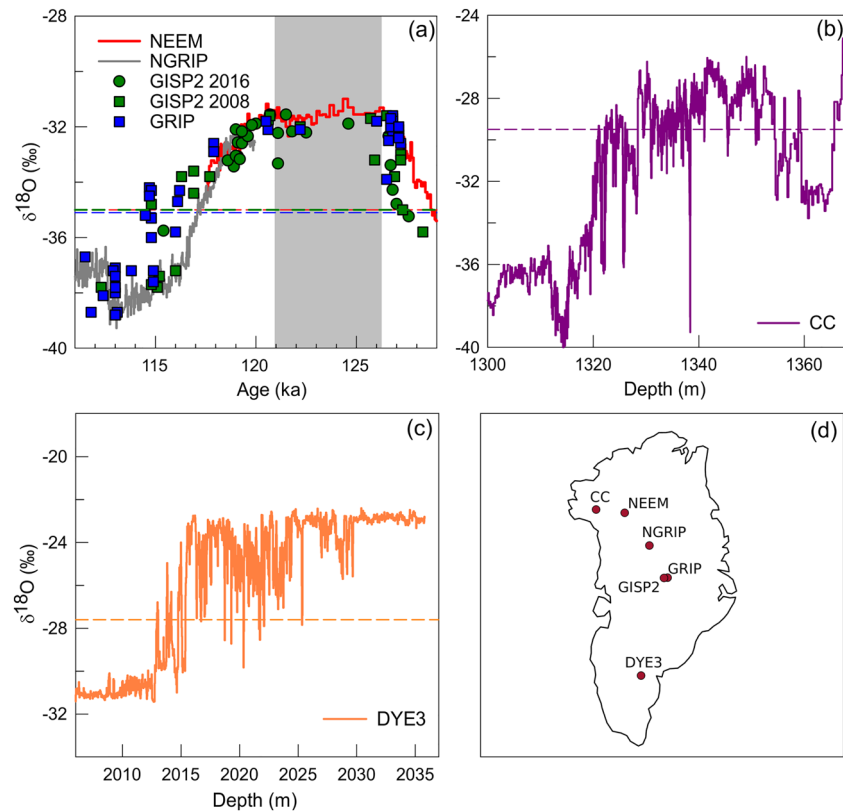


Figure 1. Summary of the data used. (a) $\delta^{18}\text{O}$ records from NEEM (red line, NEEM community members, 2013) and NGRIP (gray line, NGRIP Project Members, 2004), both displayed on the AICC2012 chronology (Govin et al., 2015; Bazin et al., 2013; Veres et al., 2013); GRIP (blue squares, dated by Landais et al., 2003); GISP2 (green squares dated by Suwa & Bender, 2008; and green circles dated by Yau et al., 2016). (b) Camp Century, $\delta^{18}\text{O}$ on a depth scale (Johnsen et al., 2001). (c) DYE3, $\delta^{18}\text{O}$ on a depth scale (Johnsen et al., 2001). (a–c) Present-day $\delta^{18}\text{O}$ values are indicated with horizontal dashed lines. (d) Locations of Greenland ice cores: NEEM (77.45°N, 51.06°W), NGRIP (75.1°N, 42.3°W), GRIP (72.6°N, 37.6°W), GISP2 (72.6°N, 38.5°W), Camp Century (77.2°N, 61.1°W), and DYE3 (65.2°N, 43.8°W).

(and temperature) changes (e.g., Gierz et al., 2017; Malmierca-Vallet et al., 2018; Masson-Delmotte et al., 2011; Sime et al., 2013; Sjolte et al., 2014).

If the LIG age interpretation of the near-bed ice at these ice core sites is correct, there must have been an extensive LIG GIS, with only a modest retreat during this period. However, this ice is difficult to date at several of the sites due to basal melting and complex flow interfering with age-depth modeling (NGRIP Project Members, 2004; NEEM community members, 2013; Landais et al., 2003). This LIG age uncertainty contributes to a wide range of possible LIG GIS scenarios. Some studies suggest that DYE3 was ice covered during the LIG (e.g., Colville et al., 2011; Willerslev et al., 2007), while others imply a complete absence of ice in southern Greenland (e.g., Koerner & Fischer, 2002; MacGregor et al., 2015). Some studies suggest strong melting in the south (e.g., Otto-Bliesner et al., 2006; Tarasov & Peltier, 2003), some in the north (e.g., Quiquet et al., 2013; Stone et al., 2013), and some in both (e.g., Born & Nisancioglu, 2012; Calov et al., 2015; Helsen et al., 2013; Robinson et al., 2011). Some studies even propose a two-dome structure with a small ice dome in south Greenland and a large ice dome in north Greenland (e.g., Calov et al., 2015; Langebroek & Nisancioglu, 2016).

Here we build on these studies with a new comprehensive investigation of the response of $\delta^{18}\text{O}$ to changes in the shape and extent of the GIS at 125 ka. We combine a synthesis of all available $\delta^{18}\text{O}$ data from deep Greenland ice cores, with simulation outputs from the Hadley Centre Coupled Model, Version 3 (HadCM3; Gordon et al., 2000; Tindall et al., 2009, referred to as simulator in the following). Since each simulation is computationally expensive and time consuming, the approach we follow is to use a Gaussian Process (GP) emulator: a fast and efficient statistical model providing probabilistic predictions of the simulator response to any GIS configuration, on the basis of the actual response on a selected set of GIS morphologies. Thanks

Table 1

Stable Water Isotopic ($\delta^{18}\text{O}$) Information From Six Greenland Deep Ice Cores: NEEM, NGRIP, GRIP, GISP2, Camp Century, and DYE3

	Anomaly values at each ice core site (LIG minus present day)					
	NEEM	NGRIP	GRIP	GISP2	Camp Century (CC)	DYE3
	$\delta^{18}\text{O}$ ‰	$\delta^{18}\text{O}$ ‰	$\delta^{18}\text{O}$ ‰	$\delta^{18}\text{O}$ ‰	$\delta^{18}\text{O}$ ‰	$\delta^{18}\text{O}$ ‰
Most likely ^a	+3.6	+3.1	+3.2	+2.7	+2.5	+4.7
Minimum ^b	+2.7	+2.5	+2.2	+1.7	+0.0	+0.0
Maximum ^b	+4.0	+3.8	+3.5	+3.4	+4.0	+5.2

Note. Most likely, maximum, and minimum values between 120 and 126 ka. All values are anomalies with respect to the present day. See section S1 in the supporting information for full details.

^aFor NEEM, we take the value at the age 126 ka. For the other ice cores, this is the value provided by Johnsen and Vinther (2007). ^bWe take the maximum/minimum value within 120–126 ka along the dated NEEM, GRIP, and GISP2 ice cores. For NGRIP, we take the maximum/minimum value from near the bed, as the ice core does not reach back to 126 ka. For CC and DYE3, we use the maximum value from near the bed as a maximum, and the present-day value as the minimum. Details are provided in section S1 of the supporting information.

to the emulator predictions, we inspect the properties of the GIS morphologies that are compatible with the available $\delta^{18}\text{O}$ records, reducing the uncertainty with respect to the previous studies.

The work is structured as follows. In section 2 we summarize the available information from the current LIG $\delta^{18}\text{O}$ measurements in Greenland. In section 3 the methodology is discussed. We describe the approach used to parameterize the range of possible GIS morphologies, section 3.1, and to select the ones used as input to the HadCM3 simulation runs, section 3.2. After a brief overview of the simulation setting, section 3.3, GP emulation is discussed in section 3.4: In section 3.4.1 a summary of the main steps and formulas needed to construct a GP emulator is provided, with the aim of providing a concise and self-contained reference of practical use; the specific choices and considerations to construct our GP emulators are detailed in sections 3.4.2 and 3.4.3. Section 3.5 illustrates the way in which the emulator predictions are compared to the ice core data, in order to identify record-compatible (RC) morphologies. In section 4, we discuss the emulator fit and show the results of the comparison between emulator predictions and ice core measurements, in terms of both mathematical properties and ice morphologies. Finally, implications for the GIS morphology during the LIG and for the ice core drilling community are discussed in section 5.

2. Data

We compile LIG stable water isotopic ($\delta^{18}\text{O}$) information from six Greenland deep ice cores: NEEM, NGRIP, GRIP, GISP2, CC, and DYE3; see Figure 1 for references and illustration of the data, and also Johnsen and Vinther (2007) and Suwa et al. (2006). While LIG ice layers have been found in numerous Greenland deep ice cores, it is not straightforward to infer the LIG peak $\delta^{18}\text{O}$ and the associated uncertainty range because of the difficulties attached to the dating of the LIG layers and the possibility of missing LIG layers.

NEEM is the only Greenland ice core depicting a continuous LIG climatic record; that is, it covers the time interval ~ 114.5 to ~ 128.5 ka (NEEM community members, 2013). Absolute dating uncertainties for NEEM are estimated to be of at least 2,000 years (Govin et al., 2015). Dating errors at the bottom for the other Greenland ice cores are likely larger. The ice stratigraphy at NGRIP is undisturbed but basal melting possibly removed the older LIG section. Hence, dating uncertainties are attached to the lack of data available to constrain the glaciological flow model used to establish the age model (NGRIP Project Members, 2004). While tentative reconstructions of the chronology of the bottom of the GRIP and GISP2 ice cores have been made using gas record synchronization with Antarctic ice cores (Landais et al., 2003; Suwa et al., 2006), the accuracy of dating the bottom of CC and DYE3 is limited by the poor preservation of deep ice samples. See section S1 in the supporting information for full details.

In Table 1, we summarize the previous information by providing a most likely, minimum, and maximum value at each site. All values are provided as anomalies with respect to the present day. At NEEM, GRIP, and GISP2, the LIG values are computed by considering the maximum and minimum $\delta^{18}\text{O}$ values measured across the time interval 120–126 ka (Figure 1). At NGRIP, we take the maximum/minimum value from near the bed, as the ice core does not reach back to 126 ka. We acknowledge that inferred uncertainties

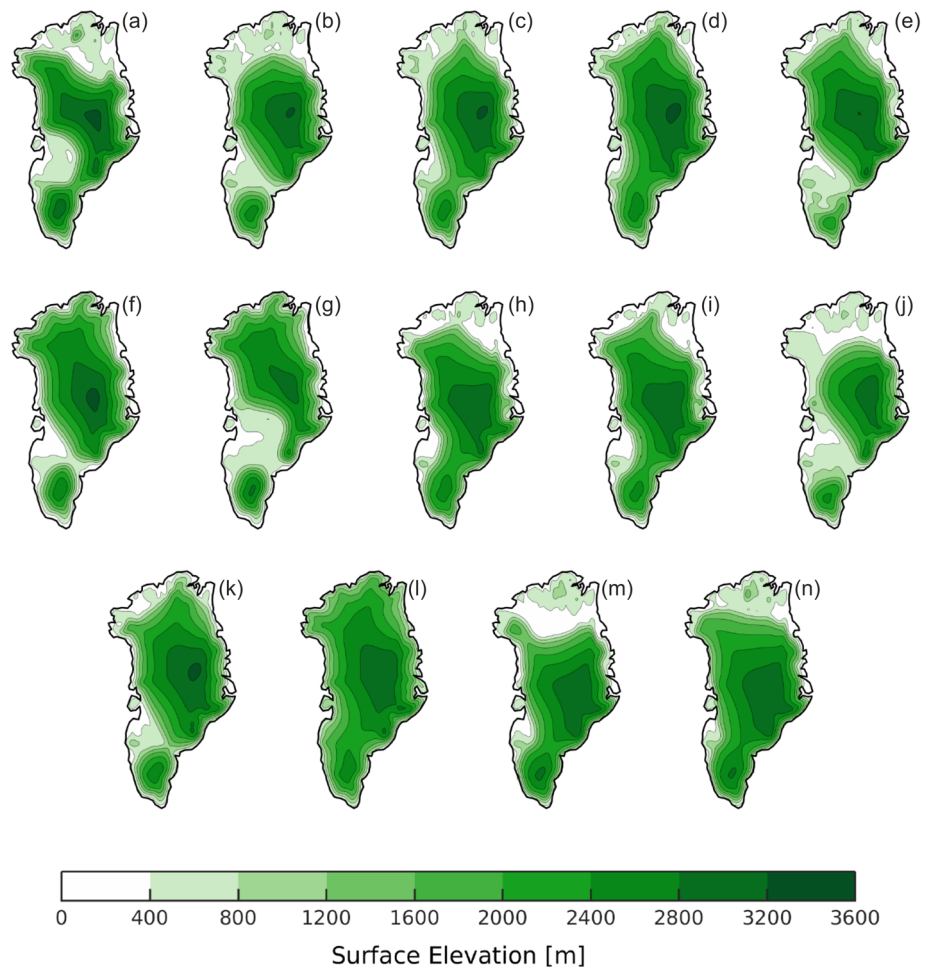


Figure 2. The surface height (meters) of LIG GIS morphologies used in this study. Panels contain data from (a) Born and Nisancioglu (2012); (b) Calov et al. (2015) at 121.3 ka; (c) Calov et al. (2015) at 122.1 ka; (d) Calov et al. (2015) at 123.3 ka; (e) Helsen et al. (2013); (f) Langebroek and Nisancioglu (2016) (minimum GIS for temperature lapse rate of 6.5 °C/km); (g) Langebroek and Nisancioglu (2016) (minimum GIS for temperature lapse rate of 8 °C/km); (h) Quiquet et al. (2013) (CNRM anomaly experiment); (i) Quiquet et al. (2013) (IPSL anomaly experiment); (j) Robinson et al. (2011) (strongly retreated GIS); (k) Robinson et al. (2011) (moderately retreated GIS); (l) Robinson et al. (2011) (weakly retreated GIS); (m) Stone et al. (2013) (GIS extent for the maximum contribution to sea level rise, +3.8 m); and (n) Stone et al. (2013) (GIS extent for the most likely contribution to sea level rise, +1.5 m).

attached to the NGRIP $\delta^{18}\text{O}$ minimum anomaly might be underestimated due to dating uncertainties as well as elevation-related assumptions (see supporting information for details). For CC and DYE3 sites, we use the maximum value from near the bed as the maximum, and the present-day value as the minimum (see supporting information for details). The last choice is due to the fact that, at both sites, it is not possible to confidently identify ice from the 120–126 ka time interval. In section 4, we investigate a series of scenarios that help to estimate how much information would be gained from recovering better dated LIG ice from the CC and DYE3 core locations.

3. Methods

To tackle the problem of using ice core data to help constrain the morphology of the GIS at 125 ka, we develop a methodology using climate simulations and Gaussian process emulation to assess the response of $\delta^{18}\text{O}$ in Greenland ice cores to changes in the shape and extent of the GIS. This requires a mathematical parameterization of the set of GIS morphologies (section 3.1), and the selection of a small number of design morphologies on which to run the $\delta^{18}\text{O}$ climate simulator (section 3.2). Based on the $\delta^{18}\text{O}$ -simulated

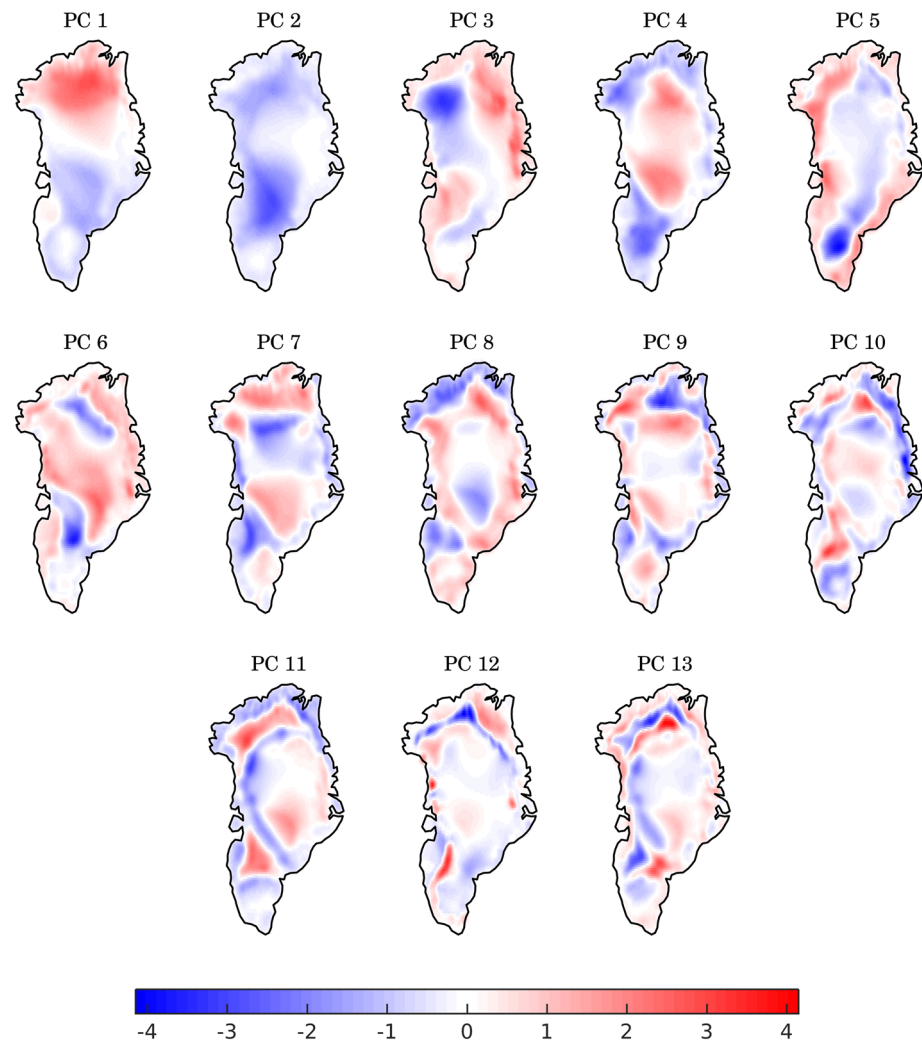


Figure 3. Plots of the PCs used to generate new synthetic morphologies in equation (1). Within each PC, areas of opposite sign yield opposite contributions to the surface height of the generated morphology. The values in the color bar do not have a physical meaning: The PCs have norm equal to one with respect to the scalar product used to carry out the PCA.

responses (section 3.3) on the design morphologies, an emulator is built (section 3.4) that allows almost instantaneous $\delta^{18}\text{O}$ predictions to any GIS morphology.

3.1. Parameterizing the Set of Ice Sheet Morphologies

We undertake the problem of parameterizing the set of ice sheet morphologies using a Principal Component Analysis (PCA) approach. This allows extrapolation of important morphology patterns from an initial set of LIG GIS reconstructions. A classical application of PCA to our ice sheet data set, however, would be inappropriate due to the spherical geometry of the Earth. We therefore apply the PCA ideas while weighting the components of a real vector, in our case representing the surface height of an ice sheet at different grid cells, by the area of the grid cells. This represents a particular case of generalized PCA, described, for example, in section 14.2.1 of Jolliffe (2002). The paper Salter et al. (2019) illustrates an application of the methodology within the uncertainty quantification literature. We provide full details of the approach applied to our case in section S4 of the supporting information.

To represent the wide range of GIS morphologies proposed for the LIG, we gather 14 GIS data sets from across several studies; see Figure 2 for references and illustration (and section S2 in the supporting information for full details). These ice sheet morphologies are used as a starting point to generate a much wider set of synthetic, potential morphologies. Prior to this, the 14 original morphologies are regridded onto a common,

Table 2

Standard Deviations σ_i Associated With the Principal Components Through PCA (First Row), and Corresponding Percentages of Explained Variance (Second and Third Rows)

	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10	PC 11	PC 12	PC 13
σ_i	271.12	228.56	145.45	141.41	91.37	86.16	75.40	65.21	56.62	47.80	44.45	36.72	25.71
% of Explained Variance	36.37	25.85	10.47	9.90	4.13	3.67	2.81	2.10	1.59	1.13	0.98	0.67	0.33
Cumulative % of Expl. Var.	36.37	62.22	72.69	82.59	86.72	90.40	93.21	95.31	96.90	98.03	99.01	99.67	100

rectangular grid with resolution of 0.2° in both directions, ranging over latitudes 59.7°N to 83.9°N and longitudes 73.6°W to 11°W . This results in a total of 122×314 grid cells for each morphology. The application of our PCA approach to this data set yields 13 Principal Components (PCs) V_1, \dots, V_{13} , each in the form of a 122×314 matrix of surface height values (Figure 3). We can generate new morphologies by taking linear combinations of the PCs, and adding the morphology \bar{M} obtained as the average of the original data set. That is, any synthetically generated morphology M in this work is of the following form:

$$M = \bar{M} + \sum_{i=1}^r x_i V_i \in \mathbb{R}^{122 \times 314} \quad (1)$$

for some real coefficients x_1, \dots, x_r . We choose $r = 8$ in light of the two following observations. Table 2 shows that the first eight PCs already account for more than 95% of the data set variance, hence capturing most of the variability displayed by the 14 original morphologies. Moreover, we observe that clear patterns can be seen in the plots of the first PCs: for example, PC1 in Figure 3 reveals that a positive coefficient x_1 in equation (1) will mainly increase the surface height of the generated morphology in the north, and decrease it in the south. Patterns can also be seen in higher PCs, but the omitted (9th to 13th) PCs seem more challenging to interpret in a physical sense.

Associating a morphology M to the vector $\mathbf{x} = (x_1, \dots, x_8)$ of its coefficients in equation (1), and vice versa, allows us to parameterize morphologies using the Euclidean space \mathbb{R}^8 . Within this parameterization, we can canonically associate the 14 original morphologies to vectors $\tilde{\mathbf{x}}^{(1)}, \dots, \tilde{\mathbf{x}}^{(14)} \in \mathbb{R}^8$, obtained by projecting each morphology onto the space generated by the first eight PCs. Notice that equation (1) with coefficients $\mathbf{x} = \mathbf{x}^{(j)}$ does not recover exactly the j th original morphology, due to the truncation error. However, the differences are small, since the neglected components account for less than 5% of the original data set variance. By denoting with $\tilde{x}_i^{(j)}$ the i th component of the vector $\tilde{\mathbf{x}}^{(j)}$, one can consider the sample variance

$$\sigma_i^2 = \text{Var} \left\{ \tilde{x}_i^{(1)}, \dots, \tilde{x}_i^{(14)} \right\}. \quad (2)$$

This provides a measure of the range covered by the original morphologies along the i th PC. The values of the standard deviation σ_i associated with the i th PC are shown in Table 2.

For our analysis we consider a prior distribution on the shape of all ice sheets or, equivalently, on the parameter vectors $\mathbf{x} \in \mathbb{R}^8$. Since the prior distribution should be concentrated in a region of the parameter space informed by the 14 original morphologies, we use

$$x_i \sim N(0, \sigma_i^2) \quad (3)$$

independently, that is, under the prior the parameters are independent, normally distributed with mean 0 and variance σ_i^2 . We will use this prior distribution, when designing the simulation runs in the next section, and when the compatibility with the ice core records is explored in section 4.2.

Finally, in order to avoid generating a GIS morphology with unrealistically low surface heights, at each grid point we consider either the bedrock height (using Stone et al., 2013 at 1.5 m of sea level rise) or the surface height from equation (1), whichever is greater. Once a morphology is generated, a corresponding land ice mask is required to carry out a $\delta^{18}\text{O}$ climate simulation: at each grid cell, this specifies whether the morphology is covered by ice or not. For the 14 original morphologies, the mask is provided. For a new morphology M , we generate the mask cell by cell, by comparing the cell height to the corresponding

heights of the original morphologies, further considering which of them have ice in the cell in question. Full details are provided in section S5 in the supporting information, where the mask generation method is also validated.

3.2. Choosing Morphologies for the $\delta^{18}\text{O}$ Simulations

The isotope-enabled coupled atmosphere-ocean general circulation model employed in this work is HadCM3. The model has been used extensively, with $\delta^{18}\text{O}$ specific code, to examine a variety of climate, sea ice, and ice sheet problems (e.g., Holloway et al., 2016, 2017; Malmierca-Vallet et al., 2018; Tindall et al., 2010). In the present study, HadCM3 is used to simulate $\delta^{18}\text{O}$ during the LIG, at the six Greenland locations introduced in Section 2.

In order to build an emulator of the climate simulator HadCM3, the simulated $\delta^{18}\text{O}$ response on some GIS morphologies is required. We call these design morphologies; the parameters $\mathbf{x}^{(j)} \in \mathbb{R}^8$ associated with them through equation (1) will be referred to as design points. Since each HadCM3 simulation is expensive, needing more than a week of real time to complete, only a few simulations can be performed. Careful selection of design morphologies is important to ensure a good calibration of the emulator (see Montgomery, 2017).

In this study, we perform simulations in two waves. For the first wave, our aim is to generate design morphologies which are well scattered within the region of space that the emulator is required to explore. For this, we use a quasi-random sample from the prior distribution (3): We generate a Halton quasi-random sequence (Kocis & Whiten, 1997) of well-scattered vectors $\hat{\mathbf{x}}^{(j)}$ in the eight-dimensional unit cube, and we apply the inverse of a cumulative distribution function of $N(0, \sigma_i^2)$ to the i th component of the vector $\hat{\mathbf{x}}^{(j)}$. This yields a sequence of design points $\mathbf{x}^{(j)}$ with the correct variance along each PC, and whose corresponding morphologies systematically cover different GIS scenarios for 125 ka. The GIS morphologies corresponding to the sequence of points $\mathbf{x}^{(j)}$ are generated, and the first 64 are selected that pass a basic plausibility test on their maximum height. The test consists in checking that the maximum surface height of a generated morphology is not greater than the mean plus four standard deviations of the set of maximum heights of the original morphologies. Two of the numerical simulations crashed for unknown reasons and were thus omitted from the analysis.

For the second wave, we use seven morphologies which are specially designed to assess how much small elevation changes in a morphology affect the simulator response. Each additional morphology is obtained upon perturbation of the parameter $\mathbf{x} \in \mathbb{R}^8$ associated with one of the first-wave morphologies (different in each case). In all seven cases, the induced difference on the surface elevation is order of only a few tens of meters. In two cases, differences of order of only few millimeters have been intentionally induced to test the chaoticity of the simulator response. This is also discussed in section 3.4.2, in relation to the use of a nugget term in equation (17). The total number of design morphologies that we use is therefore $n = 62 + 7 = 69$.

3.3. HadCM3 Simulation Runs

After the design morphologies are chosen, HadCM3 is used to reproduce the $\delta^{18}\text{O}$ response to these modified GIS morphologies at the ice core sites. All LIG simulations are forced with greenhouse gas values and orbital forcing which are appropriate for 125 ka. Further details on the model and the simulations are provided in supporting information section S3. In order to minimize the impact of climate biases of the simulator, anomalies with respect to a preindustrial simulation are considered: More precisely, we use the difference between the time-averaged $\delta^{18}\text{O}$ output from the last 50 years of the modified GIS LIG simulations and the time-averaged $\delta^{18}\text{O}$ output from the last 50 years of the preindustrial simulation.

A well-known limitation in comparing paleoclimate data with model results is the discrepancy in resolution. Model grid cells cover areas of hundreds of square kilometers, while the location where the ice core has been extracted corresponds to a single point on the globe. Here, we use bilinear interpolation over the different grid cells in order to compute the simulated $\delta^{18}\text{O}$ output at each of the ice core locations of interest. We note further that the coarse HadCM3 spatial resolution makes it difficult to reliably simulate $\delta^{18}\text{O}$ values at the coastal margins. Accordingly, the small Renland ice dome (where LIG ice has been recovered) is not analyzed in this study in that not adequately captured within the HadCM3 resolution.

In Figure 4, the outputs of the $n = 69$ simulation runs are shown for each ice core site, together with bands highlighting the minimum and maximum $\delta^{18}\text{O}$ anomalies reconstructed from the ice core records (Table 1). The figure shows that, for all pairs of locations, there are design morphologies whose simulated $\delta^{18}\text{O}$ anomalies are within the corresponding data ranges, illustrating that the model is able to reproduce the observed

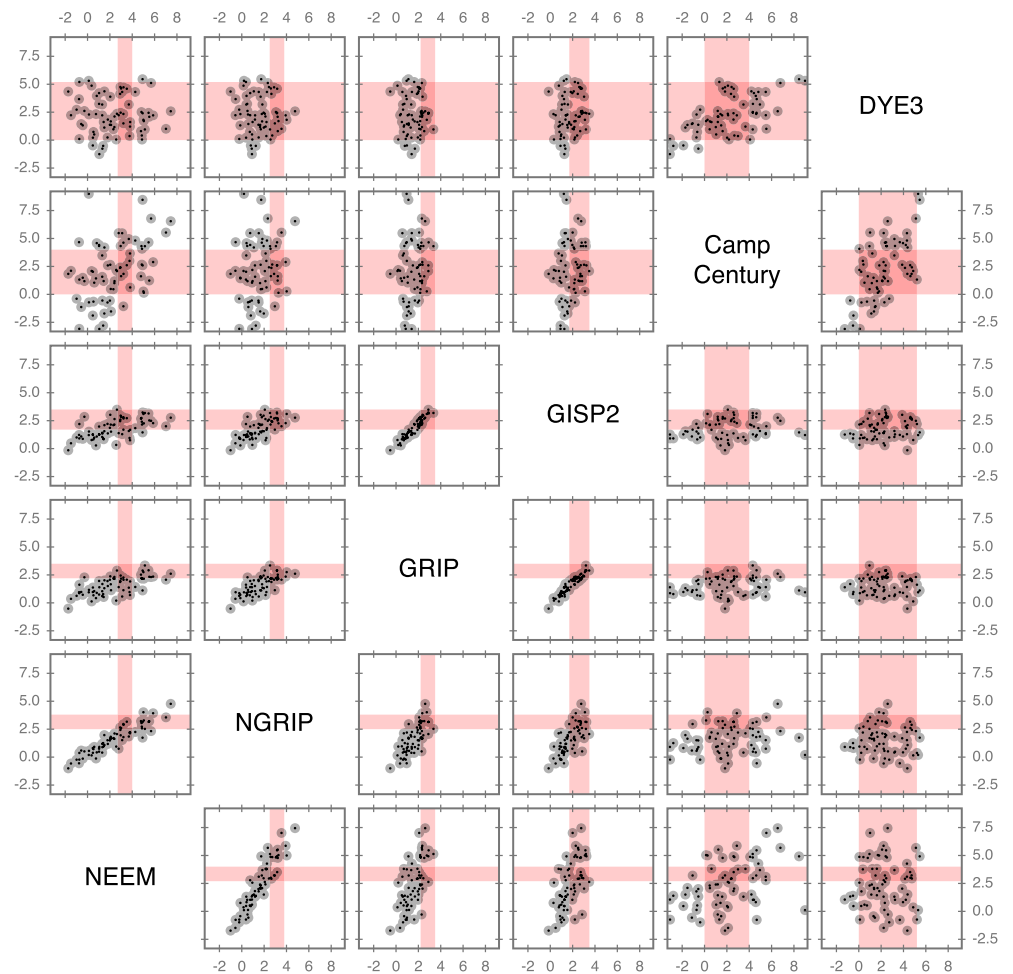


Figure 4. Simulated per-mille $\delta^{18}\text{O}$ anomalies for the design morphologies. Results are shown for pairs of locations. Shaded bands in red correspond to the ranges reconstructed using ice core records, using the minimum/maximum values shown in Table 1.

data in principle. Closer inspection shows that there are seven runs that match the data on five of the sites simultaneously. Due to the limited number of design morphologies, there is however no run which matches all six of the sites.

3.4. The Emulator

A GP emulator is a statistical model that effectively predicts the response of a complex simulator, on the basis of the actual simulator response at only a few inputs. In our case, we emulate the response of $\delta^{18}\text{O}$ anomalies to changes in the GIS morphology at 125 ka, starting from the climate simulator response on the design morphologies. For the convenience of the reader, in section 3.4.1 we summarize the formulas needed to build an emulator. In the subsequent sections, we explain and justify the particular choices undertaken in relation to the problem explored here.

3.4.1. General Setting

The Bayesian framework of GP emulation is investigated, among others, in O'Hagan (1978), Sacks et al. (1989), and Kennedy and O'Hagan (2001). Further works address more specific questions, such as the use of appropriate covariance functions (Fricker et al., 2012), the use of a nugget term (Andrianakis & Challenor, 2012), or methods to validate the emulator (Bastos & O'Hagan, 2009). In the following, we provide a concise summary of the main steps and formulas leading to the construction of a GP emulator: We refer the reader to the aforementioned literature for a more in-depth overview.

We use lowercase letters for numbers and vectors (plain and bold, respectively), and uppercase letters for matrices (plain). The general setting is the one where the simulator has been run on n design points

$\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)} \in \mathbb{R}^p$, leading to outputs $y_1, \dots, y_n \in \mathbb{R}$. The emulator is a stochastic process η providing probabilistic predictions of the simulator output at any point $\mathbf{x} \in \mathbb{R}^p$, and which interpolates the value y_j at the design point $\mathbf{x}^{(j)}$, $j = 1, \dots, n$.

To build the emulator, a prior mean function $m(\cdot)$ and covariance function $v(\cdot, \cdot)$ must be chosen. Following O'Hagan (1992), we assume

$$m(\mathbf{x}) = \mathbf{h}(\mathbf{x})^\top \boldsymbol{\beta}, \quad \mathbf{x} \in \mathbb{R}^p \quad (4)$$

$$v(\mathbf{x}, \mathbf{x}') = \sigma^2 c(\mathbf{x}, \mathbf{x}'), \quad \mathbf{x}, \mathbf{x}' \in \mathbb{R}^p \quad (5)$$

for some coefficients $\boldsymbol{\beta} \in \mathbb{R}^q$ and $\sigma^2 > 0$. Here, $\mathbf{h}(\mathbf{x})$ is a vector of q functions of the parameter \mathbf{x} , and $c(\mathbf{x}, \mathbf{x}')$ is a covariance function. The functions $\mathbf{h}(\cdot)$ and $c(\cdot, \cdot)$ must be chosen as part of the design of the emulator, while $\boldsymbol{\beta}$ and σ^2 can be estimated from data.

The GP $\eta(\cdot)$ with mean function $m(\cdot)$ and covariance function $v(\cdot, \cdot)$ is conditioned on the observations $\eta(\mathbf{x}^{(j)}) = y_j$ for $j \in \{1, \dots, n\}$. Usually, a "flat" prior distribution (O'Hagan, 1992) is chosen for the pair $(\boldsymbol{\beta}, \sigma^2)$. In this case, under the posterior distribution $\eta(\cdot)$ is a Student's t process (Tracey & Wolpert, 2018) with $n - q$ degrees of freedom. To provide the expression of the posterior mean and covariance functions, we introduce some notation:

$$H = \begin{pmatrix} \mathbf{h}(\mathbf{x}^{(1)})^\top \\ \vdots \\ \mathbf{h}(\mathbf{x}^{(n)})^\top \end{pmatrix} \in \mathbb{R}^{n \times q}, \quad A = \begin{pmatrix} c(\mathbf{x}^{(1)}, \mathbf{x}^{(1)}) & \dots & c(\mathbf{x}^{(1)}, \mathbf{x}^{(n)}) \\ \vdots & \ddots & \vdots \\ c(\mathbf{x}^{(n)}, \mathbf{x}^{(1)}) & \dots & c(\mathbf{x}^{(n)}, \mathbf{x}^{(n)}) \end{pmatrix} \in \mathbb{R}^{n \times n}. \quad (6)$$

Set

$$B = H^\top A^{-1} H \in \mathbb{R}^{q \times q}, \quad (7)$$

$$\mathbf{b} = B^{-1} H^\top A^{-1} \mathbf{y} \in \mathbb{R}^q, \quad (8)$$

$$\mathbf{e} = \mathbf{y} - H\mathbf{b} \in \mathbb{R}^n, \quad (9)$$

where $\mathbf{y} = (y_1, \dots, y_n)^\top$ and, for any input $\mathbf{x} \in \mathbb{R}^p$, define

$$\mathbf{t}(\mathbf{x}) = (c(\mathbf{x}, \mathbf{x}^{(1)}), \dots, c(\mathbf{x}, \mathbf{x}^{(n)}))^\top \in \mathbb{R}^n, \quad (10)$$

$$\mathbf{p}(\mathbf{x}) = \mathbf{h}(\mathbf{x}) - H^\top A^{-1} \mathbf{t}(\mathbf{x}) \in \mathbb{R}^q. \quad (11)$$

Using this notation, the posterior mean of σ^2 is

$$\hat{\sigma}^2 = \frac{\mathbf{e}^\top A^{-1} \mathbf{e}}{n - q - 2}. \quad (12)$$

The posterior mean and covariance function of the process $\eta(\cdot)$ are then given by

$$\hat{m}(\mathbf{x}) = \mathbf{h}(\mathbf{x})^\top \mathbf{b} + \mathbf{t}(\mathbf{x})^\top A^{-1} \mathbf{e} \quad (13)$$

and

$$\hat{v}(\mathbf{x}, \mathbf{x}') = \hat{\sigma}^2 [c(\mathbf{x}, \mathbf{x}') - \mathbf{t}(\mathbf{x})^\top A^{-1} \mathbf{t}(\mathbf{x}') + \mathbf{p}(\mathbf{x})^\top B^{-1} \mathbf{p}(\mathbf{x}')], \quad (14)$$

respectively. In particular, the emulator prediction at a point $\mathbf{x} \in \mathbb{R}^p$ is a Student's t distributed random variable with mean $\hat{m}(\mathbf{x})$, variance $\hat{v}(\mathbf{x}, \mathbf{x})$ and $\nu = n - q$ degrees of freedom: this can be written as

$$\hat{m}(\mathbf{x}) + \sqrt{\frac{\nu - 2}{\nu}} \hat{v}(\mathbf{x}, \mathbf{x}) \cdot T_\nu, \quad (15)$$

where T_ν denotes a standard t distributed random variable with ν degrees of freedom.

3.4.2. Choosing the Mean and Covariance Functions

In the present application, we fit separate emulators for each ice core site. For each site, the design points $\mathbf{x}^{(j)}$ are the eight-dimensional vectors associated with the design morphologies (section 3.2), while the values y_j are the corresponding simulated $\delta^{18}\text{O}$ anomalies.

To help us decide on the form of the function $h(\cdot)$ used in the prior mean (4), we perform a multiple linear regression, describing y_j as linear combination of $x_1^{(j)}, \dots, x_8^{(j)}$. For all the six locations, a linear model proves appropriate, showing unstructured residuals and adjusted coefficient of determination (adjusted R^2) greater than 0.89 for all sites. Given the good fit, we choose

$$m(\mathbf{x}) = \beta_0 + \beta_1 x_1 + \dots + \beta_8 x_8, \quad \mathbf{x} \in \mathbb{R}^8. \quad (16)$$

Using the notation in section 3.4.1, we have $\mathbf{h}(\mathbf{x}) = (1, \mathbf{x}^\top)^\top$, $p = 8$ and $q = p + 1 = 9$.

For the covariance function $c(\cdot, \cdot)$ used in (5) we use the form

$$c(\mathbf{x}, \mathbf{x}') = \exp(-(\mathbf{x} - \mathbf{x}')^\top D^{-1}(\mathbf{x} - \mathbf{x}')) + \mu \delta_{\mathbf{x}, \mathbf{x}'}, \quad \mathbf{x}, \mathbf{x}' \in \mathbb{R}^8, \quad (17)$$

where D is the 8×8 diagonal matrix with diagonal elements d_1^2, \dots, d_8^2 , and $\delta_{x,y}$ equals 1 if $x = y$ and 0 otherwise. The parameters d_1, \dots, d_8 and μ will be estimated from data (section 3.4.3).

The first, squared-exponential term in the definition of $c(\cdot, \cdot)$ models the decay in correlation when the distance of inputs increases. The parameters d_i are known as “correlation lengths.” In addition to the squared-exponential covariance function used in (17), experiments with other covariance functions (Matérn with parameters $\nu = 3/2$ and $\nu = 5/2$; see Rasmussen & Williams, 2006) were carried out while training the emulator. The mean and variance of the emulator predictions, as well as cross-validated estimates of the emulation error, were remarkably similar in the three cases. Given the apparent smooth linear response of the vector y to the eight PC scores, we chose the smoothest of the three covariance functions, that is, the squared exponential.

The second term in the definition of $c(\cdot, \cdot)$ takes observational variance into account. In the presence of this term, the emulator provides non-deterministic predictions (i.e., with nonzero variance) even at the design points. This term is often referred to as “the nugget term” (see, e.g., Andrianakis & Challenor, 2012). In our case, its addition is necessary because the system simulated by HadCM3 is chaotic. Even after averaging the outputs over 50 years, GIS morphologies that are physically equivalent (i.e., with surface elevations differing only by millimeters) result in significantly different $\delta^{18}\text{O}$ simulated anomalies. A numerical benefit of using a nugget term is that it makes the matrix A in (6) better conditioned. It thus helps to achieve numerical stability when computing the emulator predictions, a step that involves the use of the matrix inverse A^{-1} .

3.4.3. Estimation of Correlation Lengths and Nugget Term

The correlation lengths d_i and the nugget term μ appearing in equation (17) can be estimated in different ways. A maximum likelihood approach is often used in the estimation of hyperparameters (see Andrianakis & Challenor, 2012; Berger et al., 2001): however, especially in higher dimensions, the maximization algorithm is often trapped in either local maxima or flat regions. In this work, we use the alternative approach described in this section.

First, we observe that all PCs are normalized and have therefore comparable size. Changes of similar magnitude in any two different components of an input parameter $\mathbf{x} \in \mathbb{R}^8$ therefore yield, overall, changes of similar magnitude in the corresponding ice morphologies. At a single-site level this becomes less true, with similar changes in different PCs affecting differently the surface height at the site. We notice, however, that the simulated $\delta^{18}\text{O}$ response at a location does not depend only on the surface height at the location, but rather on the surface elevation within a larger area around it and, ultimately, on the whole morphology. In the light of this, assuming that similar changes in the components of a vector x have a comparable impact in the simulated $\delta^{18}\text{O}$ at a given location does not appear unreasonable.

Separately, we note that the number of parameters that have to be estimated to build each emulator includes the nine mean regression coefficients β_0, \dots, β_8 , the variance σ^2 , the nugget term μ , and the 8 correlation lengths d_i . Given the limited number of training points ($n = 69$), an independent estimation of all the parameters presents a high risk of overfitting. In light of both this and the previous consideration, we assume $d_1 = \dots = d_8 = d$. The parameter d and the nugget parameter μ are estimated simultaneously, by maximizing a cross-validated estimate of the posterior density of the data, as explained below.

Starting from the data set $(\mathbf{x}^{(j)}, y_j)_{j=1, \dots, n}$, consider the n emulators calibrated on the data sets where the j^{th} observation has been left out, respectively. Let $\rho_{d, \mu}^{(j)}(\cdot)$ be the posterior density function at $\mathbf{x}^{(j)}$ of such an emulator, where the parameters $d = (d_1, \dots, d_n)$ and μ are used as correlation lengths and nugget respectively. From section 3.4.1 we know that $\rho_{d, \mu}^{(j)}$ is the density function of a t distribution with mean and variance as in (13) and (14). Now define

$$f(d, \mu) = \log \left(\prod_{j=1}^n \rho_{d, \mu}^{(j)}(y_j) \right) = \sum_{j=1}^n \log \left(\rho_{d, \mu}^{(j)}(y_j) \right). \quad (18)$$

The function $f(d, \mu)$ is a measure of the goodness of the emulator predictions, when correlation lengths d and nugget μ are used. The parameters d and μ used in our emulators are chosen to maximize $f(\cdot, \cdot)$. We use the Matlab nonlinear solver `fminunc` and repeat the maximization multiple times from different starting points, to reduce the risk of getting stuck in a local extremum.

3.5. Data-Model Comparison

In order to carry out the data-model comparison, we follow the idea of using an implausibility measure, as, for example, in Craig et al. (1997), Vernon et al. (2010), and Williamson et al. (2015). Implausibility measures are used within the context of history matching: This last term, originally used in the oil industry (see the above-mentioned Craig et al., 1997), is now employed in the emulation literature to refer to the process of finding the inputs of a model whose outputs match historical observations. In our case, to select the morphologies compatible with the observed ice core records, we proceed as follows.

Let $\hat{m}_L(\mathbf{x})$ and $\hat{v}_L(\mathbf{x}, \mathbf{x})$ be the mean and the variance of the $\delta^{18}\text{O}$ emulator prediction at a given location L , for the morphology represented by $\mathbf{x} \in \mathbb{R}^8$ (equations (13) and (14)). Let us further denote by R_L the most likely ice core record $\delta^{18}\text{O}$ anomaly at that location, and by R_L^- and R_L^+ the lower and upper bounds associated with it: values for $R_L^- < R_L < R_L^+$ are provided in Table 1. In order to quantify the mismatch, at the location L , between the emulator $\delta^{18}\text{O}$ prediction and the ice core record information, we define the following “implausibility function”:

$$I_L(\mathbf{x}) = \frac{|\hat{m}_L(\mathbf{x}) - R_L|}{\sqrt{\hat{v}_L(\mathbf{x}, \mathbf{x}) + \text{Var}(\text{Rec}_L)}}. \quad (19)$$

The quantity $\text{Var}(\text{Rec}_L)$ is meant to provide a measure of the variability associated with the ice core record. Based on the analogy that the variance of a uniform random variable on an interval of length l is $l^2/12$, we compute $\text{Var}(\text{Rec}_L)$ as follows:

$$\text{Var}(\text{Rec}_L) = \begin{cases} (R_L^+ - R_L)^2/3 & \text{if } \hat{m}_L(\mathbf{x}) > R_L, \\ (R_L - R_L^-)^2/3 & \text{if } \hat{m}_L(\mathbf{x}) < R_L. \end{cases} \quad (20)$$

Notice that, in equation (20), we deal with the two subintervals rather than the whole interval, hence the factor $1/3$ rather than $1/12$ appears. Finally, in order to obtain a quantitative measure of the mismatch between the emulator predictions for the morphology represented by $\mathbf{x} \in \mathbb{R}^8$ and the $\delta^{18}\text{O}$ information at all the six locations, we define

$$I(\mathbf{x}) = \max\{I_{L_1}(\mathbf{x}), \dots, I_{L_6}(\mathbf{x})\}, \quad (21)$$

where L_1, \dots, L_6 are the six locations in question.

We classify a morphology as RC, if the condition $I(\mathbf{x}) \leq 2$ holds, and at least 95% of the morphology cell heights are within 2 standard deviations of the average height of the 14 original morphologies at that cell. This ensures that RC morphologies pass a general physical plausibility test, in addition to being compatible with the data records. The possibility of relaxing the first criterion ($I_L(\mathbf{x}) \leq 2$ for $L = L_1, \dots, L_6$) to the five sites with lowest implausibility measure has also been considered, but the choice leads to neglect in most cases the DYE3 constraint (in line with the high $\delta^{18}\text{O}$ anomaly recorded at the site). This is considered at odds with both the aim of the three-scenario analysis and the fact that DYE3 is the only southern site available in this study.

We conclude with a comment. In the classical emulation literature, implausibility measures are generally used in combination with a several-wave simulation approach: The latter sequentially rules out the parts of

Table 3

Location by Location, the Values of the Correlation Length (d) and the Nugget Term (μ) That Are Used to Build the Corresponding Emulator

	NEEM	NGRIP	GRIP	GISP2	Camp Century	DYE3
Correlation length (d)	208.14	309.40	397.58	322.93	3,117.3	513.40
Nugget term (μ)	0.45	0.35	0.28	0.53	2.54×10^{-4}	0.09

Note. The estimation is carried out through the cross-validation procedure explained in section 3.4.3. Higher values of d denote a stronger correlation between the $\delta^{18}\text{O}$ outputs associated with different morphologies. For a given d , higher values of μ indicate higher variance of the emulator predictions at the design points, alongside a generally smoother mean prediction.

the input space for which the condition $I(x) > c$ holds true and builds a new emulator on the “Not Ruled Out Yet” part of the space. The constant c is often chosen to be equal to 3 following Pukelsheim’s 3σ rule (Pukelsheim, 1994). In this work, resource constraints prevented us from performing several iterations of the procedure. Hence, the implausibility measure $I(x)$, alongside the physical plausibility test described in the previous paragraph, is used to directly select morphologies, rather than rejecting them. For this reason we use a stricter criterion ($c = 2$) than the one more commonly found in the literature: The approach allows us to identify morphologies with a greater potential to match the ice core data. Properties of these morphologies are explored in the next section.

4. Results

4.1. Fit of the Emulator

One emulator for each location is built according to the procedure described in section 3.4. The estimated values of the correlation lengths and nugget are shown in Table 3. With the only exception of CC, the estimated d are of comparable magnitude to the standard deviations σ_i of the first PCs, see Table 2. In the case of CC, the estimation returns instead a relatively large d . After considerations, we decided to retain such value of d : indeed, it can be seen that for very large d the emulator converges to a simpler Bayesian linear regression model, which would still represent an appropriate limit case since the data at this site is well-described by a linear model, with an adjusted R^2 of 0.951.

To assess the quality of the fitted emulators, we consider the cross-validated residuals for each site: That is, for each simulation, we fit an emulator leaving out this simulation and use this emulator to predict the simulation output. This is done, in turn, for all n simulations. Figure 5 shows the differences between actual and predicted model outputs, at the six sites. The figure shows that nearly all simulation results lie within two emulator standard deviations of the prediction. Thus, both the values and the uncertainties predicted by the emulator seem appropriate. In particular, the emulator does not only make accurate predictions but also can accurately assess the uncertainty of its predictions.

4.2. Comparison Between Emulator Predictions and Ice Core Records

The six emulators allow us to efficiently predict, at the different locations, the $\delta^{18}\text{O}$ response of the climate simulator to any ice sheet morphology. Using the methodology of section 3.5, emulator outputs can be compared to the available $\delta^{18}\text{O}$ ice core information to decide whether a given GIS morphology is RC.

In order to carry out the data-model comparison detailed in section 3.5, values of R_L^-, R_L and R_L^+ must be specified at all the locations where compatibility with records is to be imposed. We use the values in Table 1. However, as explained in section 2, lower bounds for the $\delta^{18}\text{O}$ anomalies at CC and DYE3 cannot be easily inferred from the records. In order to acknowledge this uncertainty, we carry out the analysis in the case of three different scenarios, according to how close these lower bounds are set to the most likely values $R_{CC}^- = +2.5\%$ and $R_{DYE3}^- = +4.7\%$. The scenarios we investigate are as follows:

1. Loose scenario (only imposing nonnegative anomalies):

$$R_{CC}^- = 0\%, \quad R_{DYE3}^- = 0\%. \quad (22)$$

2. Middle scenario:

$$R_{CC}^- = 1\%, \quad R_{DYE3}^- = 2\%. \quad (23)$$

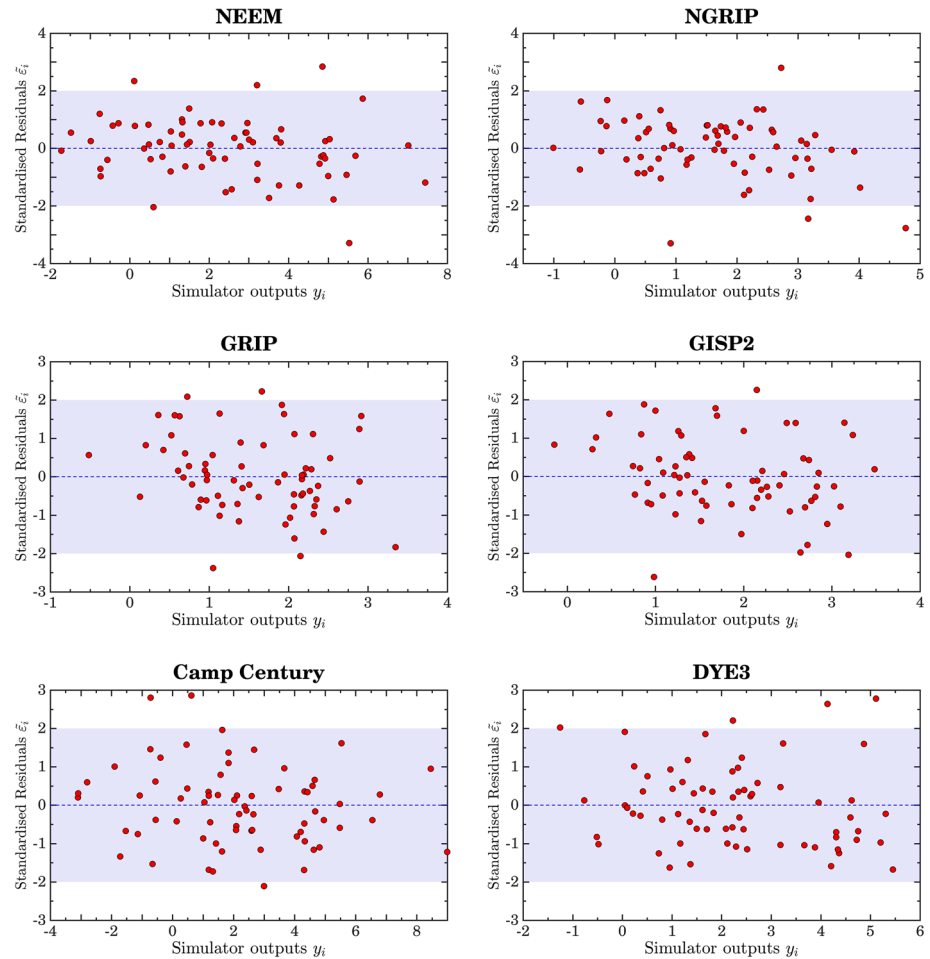


Figure 5. Cross-validated residuals of the emulators at the different sites. Each red dot in a panel corresponds to one of the $n = 69$ simulations: The x component shows the simulator output; the y component shows the difference between the emulator prediction and the simulator output, measured in multiples of the standard deviation associated with the emulator prediction. The shaded bands highlight y values between -2 and 2 , where around 95% of the standardized predictions are expected to lie. This is the case for all the six emulators used in this work.

3. Tight scenario ($R_L^- = R_L - 1\%$):

$$R_{CC}^- = 1.5\%, \quad R_{DYE3}^- = 3.7\%. \quad (24)$$

For our analysis, we represent the prior distribution on all possible ice sheet morphologies by a set of $N = 10^7$ morphologies, sampled randomly from the prior distribution described in section 3.1. Each of these N morphologies is classified as either being RC or not, according to the criterion described in section 3.5. The resulting RC morphologies then form a sample from the posterior distribution, incorporating the constraints from the data. In the loose scenario case, where the imposition of the record compatibility at CC and DYE3 has little effect, the RC morphologies represent 5.75% of the morphologies sampled from the prior distribution. This percentage approximately halves (2.86%) in the middle scenario, and further decreases to 0.73% in the tight scenario.

In each of the three scenarios, the percentage of morphologies that satisfy the condition $I(x) < 2$ but do not pass the physical plausibility test (end of section 3.5) is as follows: 16.5% (loose), 17.3% (middle), and 27.0% (tight). Perhaps not surprisingly, this shows that the condition of matching the six ice core records, as simulated by the HadCM3 model and later estimated via the emulator, is alone not sufficient to identify physical morphologies.

Figure 6 shows how the posterior distribution compares to the prior, in the two cases of the loose and tight scenarios. Two-dimensional cross sections of the subspace spanned by the first three PCs are shown. The

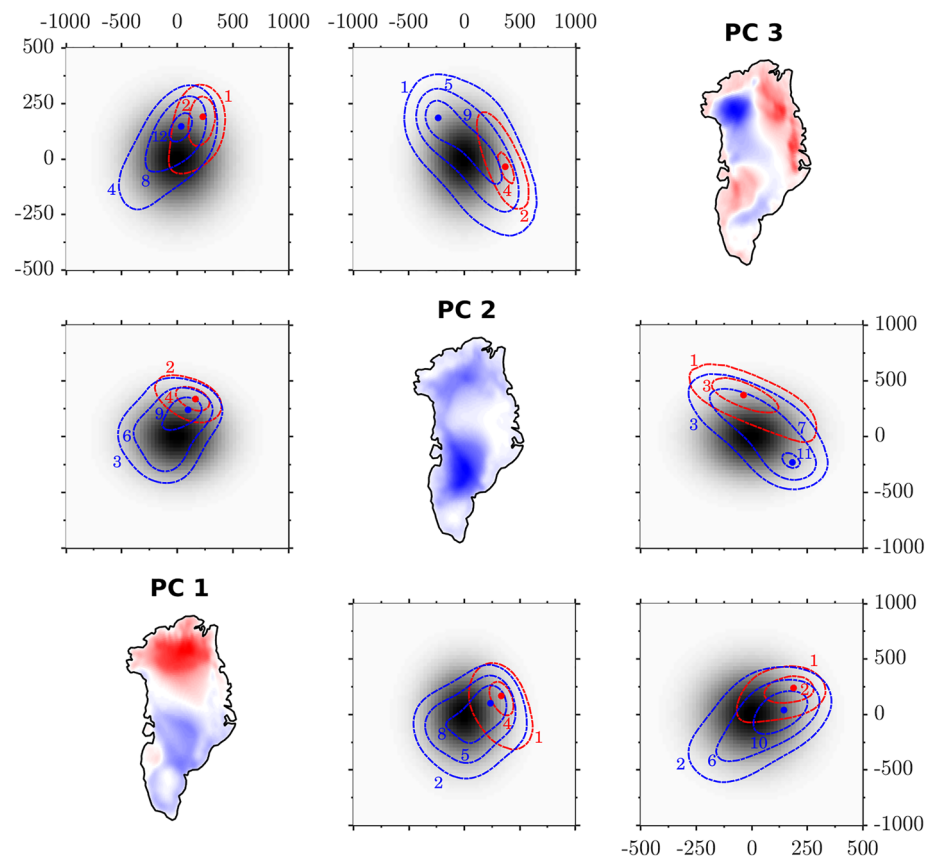


Figure 6. Two-dimensional sections of the space of PC scores, for the first three PCs. On the diagonal, illustration of the PCs. The shaded gray background in the off-diagonal plots represents the density of the prior Gaussian distribution on the morphologies. The contour lines represent contours of the posterior densities of RC morphologies, in the case of the loose (blue) and tight (red) scenarios. Labels indicate the percentage of morphologies that are compatible with the records along the specified contour. Different contours are shown in symmetric plots.

colored lines represent contours of the posterior densities (red for the tight scenario, blue for the loose one), and the labels indicate the percentage of RC morphologies that are selected along the corresponding contour. The gray shaded background instead shows the Gaussian prior distribution.

From Figure 6, it can be appreciated that the posterior distribution of RC morphologies looks closer to the prior in the loose scenario than in the tight scenario, as expected. However, the particular directions in which the shifts happen, alongside the PC shapes, can provide insight on the patterns displayed by the RC morphologies. Especially in the tight scenario case, we can see that areas characterized by a higher density of RC morphologies tend to have a positive score in the first PC (more evident in the subplot of PC1 and PC3). In light of the “North-South” pattern shown by the first PC, this seems to suggest the following: Imposing the compatibility with records leads to morphologies with lower surface height (i.e., less or no ice) in the south and higher surface height in the north, compared to the average morphology generated through the prior.

Similarly, the two subplots in Figure 6 concerning the second and third PCs reveal critical information. The RC morphology scores for the second PC are mostly negative in the loose scenario (blue) and positive in the tight scenario (red). Notice that the second PC has negative values almost everywhere, particularly in the southwestern block of Greenland. The strong constraint on DYE3 and CC that is imposed in the tight scenario seems therefore to induce a clear loss of ice in the southwestern block. However, when a loose constraint is imposed at the two locations, the typical surface elevation in the south would increase, reflecting the presence of more ice.

4.3. Shape and Uncertainty of RC Morphologies

The main characteristics displayed by RC morphologies are illustrated in Figure 7. Mean and standard deviation are shown for the prior (left), and for the posterior in the three scenarios (right). Most of the variability,

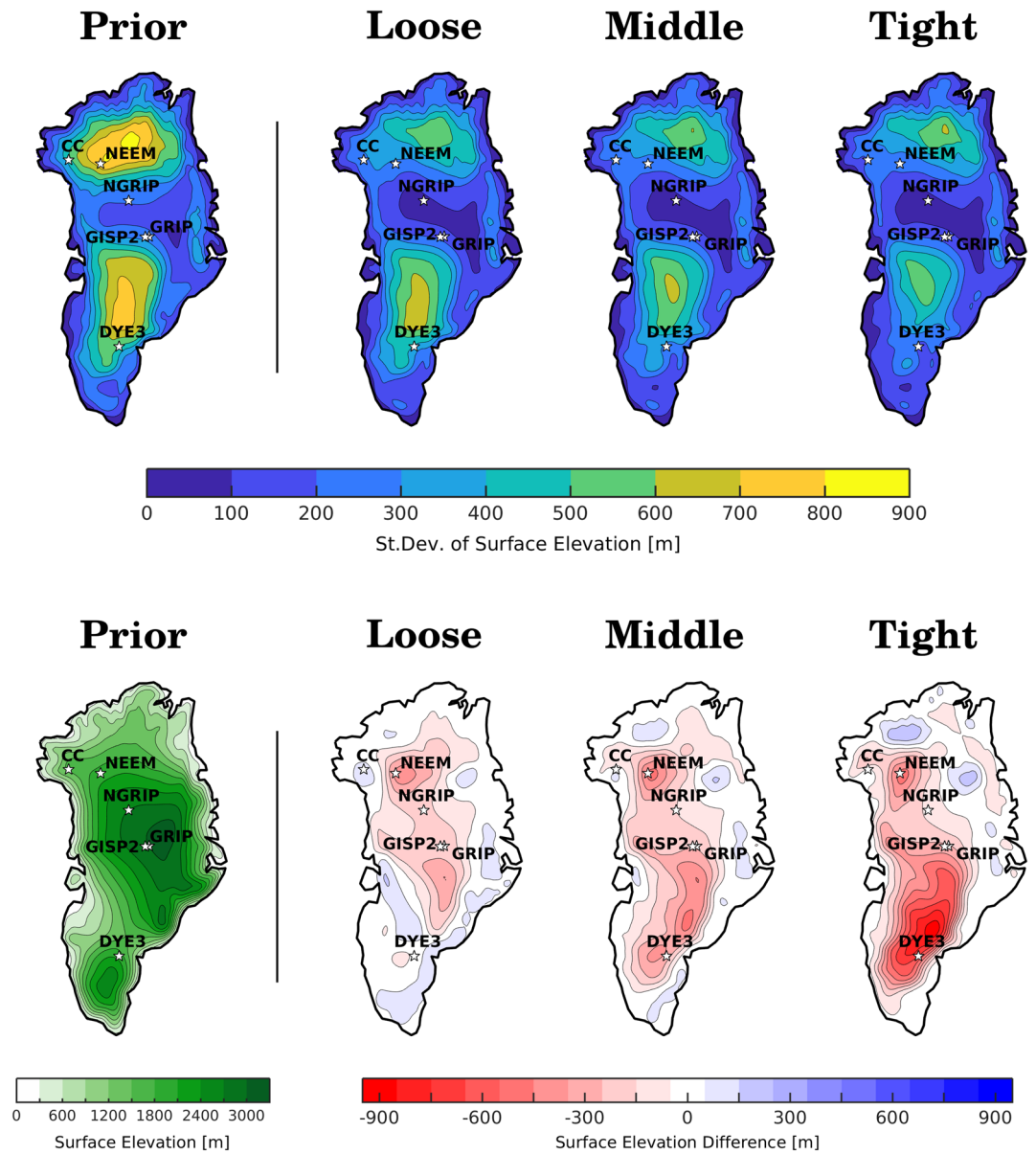


Figure 7. Plots represent the cellwise standard deviation (top part) and average (bottom part) of the surface height of different sets of morphologies. $N = 10^7$ morphologies sampled from the prior distribution are used in the plots on the left; only RC morphologies are used for the posterior in the three scenarios on the right. To ease visual interpretation, in the bottom right plots we show the difference between the posterior and prior mean. The “Loose” scenario corresponds to the best current data, the “Middle” and “Tight” scenarios illustrate results which could be achieved, if better constraints for DYE3 and Camp Century were available.

and thus of the uncertainty, is displayed within two regions, one in the north and one in the center south (top row). The uncertainty in surface elevation in these two areas is significantly reduced, once the constraints from the data are taken into account. However, while little difference is present in the north between the three scenarios, the imposition of a stronger constraint on the DYE3 and CC records induces a clear reduction in the uncertainty of the southern surface elevation.

The prior and posterior mean morphologies are shown in the bottom row of Figure 7. To ease the interpretation, we show the prior mean on the left and the difference between posterior and prior means on the right. Progressively from left to right, a distinct decrease of surface height in the south appears. From the loose to the middle scenario, the increase of 2% on the lower bound of R_{DYE3}^- induces a decrease of the average surface height in the area surrounding DYE3 of up to 400 m. Moreover, by further tightening the R_{DYE3}^-

constraint of 1.7‰ (middle to tight scenario), we observe an additional decrease in the same area of up to 450–500 m. The other area in Greenland where a similar phenomenon arises, albeit less prominently, is the one around CC and in the immediate north of it, where the decrease is of approximately 150–200 m in each of the two steps (loose to middle and middle to tight scenarios).

Also observe that, in the case of the loose scenario, the areas displaying surface heights lower than the prior morphologies are concentrated around the four locations where record constraints play an active role (NEEM, NGRIP, GRIP, and GISP2), leaving the DYE3 and CC areas mostly unaffected.

We remark that our approach tightens simultaneously the DYE3 and CC records, hence it does not allow us to ascribe with certainty a given posterior feature to the constraint imposed on one or the other location. The geographical position of the two sites and the size of the constraint on their R_L^- nonetheless favor the interpretation that the DYE3 constraint is responsible for the southern surface elevation decrease of tighter scenarios, as implied above. Inspection of the land ice masks for morphologies compatible with the tight scenario shows that most of these morphologies present a two-dome structure, with DYE3 sitting at the northern top of the southern dome.

We conclude with a note. The authors have carried out further experiments aiming to assess the sensitivity of the results to changes in the minimum NGRIP anomaly, whose value is also uncertain as acknowledged in section 2. The experiments reveal that no tangible change to the shape of the RC morphologies can be appreciated when the minimum $\delta^{18}\text{O}$ NGRIP anomaly is decreased by 0.5‰, with Figures 6 and 7 remaining essentially unchanged. The percentage of morphologies which result RC increases to approximately 7.5%, 4%, and 1%, in the loose, middle, and tight scenarios, respectively.

4.4. Record Compatibility in the Wider Sense

Within the ice sheet modeling community, Greenland ice age information from ice cores is used to attempt to constrain LIG GIS results. The presence of ice at NGRIP and at the Summit region through the (late) LIG period has previously been used to reject particular simulations (e.g., Robinson et al., 2011; Stone et al., 2013). CC and DYE3 sites are not generally used in the same way because the age of the (possible) LIG period ice is considerably more ambiguous. At DYE3, ice may or may not have covered the site during the LIG (Alley et al., 2010; Willerslev et al., 2007). At CC some ice at the base occurs that has been interpreted as being pre-LIG in nature (Dansgaard et al., 1982), however, as at DYE3, this is ambiguous. Without better ice age information there will not be community agreement on whether DYE3, or CC, was definitely ice-covered during the LIG.

As well as ice age information, ice core-based elevation change reconstructions, based on $\delta^{18}\text{O}$, have been used to infer LIG GIS changes (e.g., Johnsen & Vinther, 2007). However, this cannot be used as an additional constraint here, because this $\delta^{18}\text{O}$ information is already accounted for by our analytical approach. Beyond $\delta^{18}\text{O}$ elevation interpretations, air content measurements have also been used as an independent constraint. In particular, the NEEM elevation change reconstruction, which specifies a NEEM surface elevation difference of $+45 \pm 350$ m at 126 ka compared to present day, is derived from air content measurements (NEEM community members, 2013). This data is fully independent of our approach.

If we impose the NEEM elevation constraint on our resulting RC morphologies, in the loose and middle scenarios cases around 3.7% of the RC morphologies are compatible with this additional NEEM data. This percentage increases in the tight scenario, where 6.5% of our $\delta^{18}\text{O}$ compatible morphologies are in agreement. The relatively small overlap between air content and $\delta^{18}\text{O}$ compatible morphologies could possibly reflect uncertainties in the air content elevation reconstruction methodology. The method relies on calculating corrections to air content due to insolation and temperature (Eicher et al., 2016; Raynaud et al., 2007) alongside secular changes in surface pressure and winds (Krinner et al., 2000; Martinerie et al., 1994). Surface melting also affects NEEM air content measurements between 127 and 118.3 ka (NEEM community members, 2013). Thus, the interpretation of elevation from air content measurements is rather uncertain. However, in spite of these uncertainties, results from the two independent approaches ($\delta^{18}\text{O}$ and air content) do overlap.

5. Conclusions

Malmierca-Vallet et al. (2018) demonstrated that an isotopic-enabled climate model cannot correctly simulate the LIG $\delta^{18}\text{O}$ measurements at DYE3, GRIP, and Renland sites via sea ice changes alone. Here, our

analysis using the emulator establishes that outputs from the isotope-enabled HadCM3 model can be reconciled with the best current LIG $\delta^{18}\text{O}$ measurements from ice cores. We have shown that this requires a careful parameterization of ice sheet morphologies and a thorough approach to the characterization of uncertainty.

The major source of uncertainty, tackled here via the construction of suitable emulators, arises from the impossibility of running the simulator on a high number of morphologies. Additional uncertainties stem from the fact that, both in reality and in simulations, the isotopic response in the ice is dependent not only on the ice sheet morphology but also on other parameters such as evaporation conditions, transport pathway effects, and Arctic sea ice variations (Sime et al., 2013). In this regard, we note that the nugget term used to model the prior emulator covariance, alongside reflecting the chaotic behavior of the simulator, accounts as well for the variability in simulator outputs which is due to inputs not explicitly included in the emulator: This is referred to as residual variability in Kennedy and O'Hagan (2001). Last but not least, we stress once again that the available ice-core-based estimates of $\delta^{18}\text{O}$ anomalies are themselves uncertain. The case of DYE3 and CC is a prototypical example of this.

In this work we aim to account for all the previous unavoidable sources of uncertainty within a unified framework, joining mathematical parameterization of the space of morphologies, Gaussian process emulation and available ice core records. The scenario-based approach we undertake demonstrates the importance of well-constrained records. In common with previous studies, we show that dated ice from around DYE3 is crucial for certain inferences about Greenland (Sime et al., 2019). If more certain records at DYE3 and CC confirmed that the high $\delta^{18}\text{O}$ values measured near the bottom of the ice cores were from the LIG, our results show that this would be compatible with morphologies characterized by low surface height in the south: Roughly, a decrease of 200 m has been observed corresponding to an increase of 1‰ in the minimal $\delta^{18}\text{O}$ anomaly. We have also shown that the variability in reconstructed surface elevation would decrease in such tighter scenarios, with glimpses that the data would then favor a two-dome structure of the LIG ice sheet.

While this study's aims are to provide a methodology of wide applicability to infer LIG GIS changes and to stress the importance of statistical emulators in similar expensive tasks, we believe that further steps may be included to ensure wider data-model comparisons. In light of the above, we encourage the search of additional data at DYE3 and CC and suggest that ice, sea ice, and isostatic changes are all accounted for within a joint framework. This may enable a useful quantitative assessment of how Greenland ice sheet changes affected global sea levels during the LIG.

Data Availability Statement

The data generated and employed in this study are available at the repository Domingo et al. (2020). The data sets contain (a) the values of the PCs V_i and of the average morphology \bar{M} used in equation (1) and (b) the PC scores identifying the design morphologies and the associated simulated $\delta^{18}\text{O}$ anomalies at the six ice core sites.

Acknowledgments

The authors thank the EPSRC-funded Past Earth Network (Grant EP/M008363/1), which funded the feasibility study giving rise to the present work, and acknowledge NERC funding for this work through Grants NE/P009271/1, NE/P013279/1, and NE/J004804/1 and through Irene Malmierca-Vallet's PhD studentship. Emilie Capron acknowledges financial support from the ChronoClimate project, funded by the Carlsberg Foundation. The authors thank Bo M. Vinther for fruitful discussions regarding the Greenland ice core data set. The authors also thank the two referees for their attentive reading of the manuscript and the valuable suggestions provided during the review phase.

References

- Alley, R. B., Andrews, J., Brigham-Grette, J., Clarke, G., Cuffey, K., Fitzpatrick, J., et al. (2010). History of the Greenland Ice Sheet: Paleoclimatic insights, 29, 1728–1756. <https://doi.org/10.1016/j.quascirev.2010.02.007>
- Andrianakis, I., & Challenor, P. G. (2012). The effect of the nugget on Gaussian process emulators of computer models. *Computational Statistics and Data Analysis*, 56(12), 4215–4228. <https://doi.org/10.1016/j.csda.2012.04.020>
- Bastos, L. S., & O'Hagan, A. (2009). Diagnostics for Gaussian process emulators. *Technometrics*, 51(4), 425–438.
- Bazin, L., Landais, A., Lemieux-Dudon, B., Toyé Mahamadou Kele, H., Veres, D., Parrenin, F., et al. (2013). An optimized multi-proxy, multi-site Antarctic ice and gas orbital chronology (AICC2012): 120–800 ka. *Climate of the Past*, 9, 1715–1731. <https://doi.org/10.5194/cp-9-1715-2013>
- Berger, J. O., de Oliveira, V., & Sansó, B. (2001). Objective Bayesian analysis of spatially correlated data. *Journal of the American Statistical Association*, 96(456), 1361–1374. <https://doi.org/10.1198/016214501753382282>
- Born, A., & Nisancioglu, K. H. (2012). Melting of Northern Greenland during the last interglaciation. *The Cryosphere*, 6, 1239–1250. <https://doi.org/10.5194/tc-6-1239-2012>
- CAPE Last Interglacial Project Members (2006). Last Interglacial Arctic warmth confirms polar amplification of climate change. *Quaternary Science Reviews*, 25, 1382–1400. <https://doi.org/10.1016/j.quascirev.2006.01.033>
- Calov, R., Robinson, A., Perrette, M., & Ganopolski, A. (2015). Simulating the Greenland ice sheet under present-day and palaeo constraints including a new discharge parameterization. *The Cryosphere*, 9, 179–196. <https://doi.org/10.5194/tc-9-179-2015>
- Colville, E. J., Carlson, A. E., Beard, B. L., Hatfield, R. G., Stoner, J. S., Reyes, A. V., & Ullman, D. J. (2011). Sr-Nd-Pb isotope evidence for ice-sheet presence on southern Greenland during the last interglacial. *Science*, 333, 620–623. <https://doi.org/10.1126/science.1204673>

- Craig, P. S., Goldstein, M., Seheult, A. H., & Smith, J. A. (1997). Pressure matching for hydrocarbon reservoirs: A case study in the use of bayes linear strategies for large computer experiments. *Case studies in Bayesian statistics* (pp. 37–93). New York, NY: Springer New York. <https://doi.org/10.1007/978-1-4612-2290-3&urlscore:2>
- Cuffey, K. M., & Marshall, S. J. (2000). Substantial contribution to sea-level rise during the last interglacial from the Greenland ice sheet. *Nature*, *404*, 591–594. <https://doi.org/10.1038/35007053>
- Dansgaard, W., Clausen, H. B., Gundestrup, N., Hammer, C. U., Johnsen, S. F., Kristinsdottir, P. M., & Reeh, N. (1982). A new Greenland deep ice core. *Science*, *218*, 1273–1277. <https://doi.org/10.1126/science.218.4579.1273>
- DeConto, R. M., & Pollard, D. (2016). Contribution of Antarctica to past and future sea-level rise. *Nature*, *531*, 591–597. <https://doi.org/10.1038/nature17145>
- Domingo, D., Malmierca-Vallet, I., Sime, L., & Voss, J. (2020). Surface elevation of 69 Greenland Ice Sheet morphologies and associated $\delta^{18}\text{O}$ anomalies (with respect to Pre-industrial) simulated by HadCM3 [Data set]. UK Polar Data Cen735 tre, Natural Environment Research Council, UK Research & Innovation, <https://doi.org/10.5285/35AED839-1634-4692-B6D6-4D6312953EB5>, <https://ramadda.data.bas.ac.uk/repository/entry/show?entryid=35aed839-1634-4692-b6d6-4d6312953eb5>
- Dutton, A., Carlson, A. E., Long, A. J., Milne, G. A., Clark, P. U., DeConto, R., et al. (2015). Sea-level rise due to polar ice-sheet mass loss during past warm periods. *Science*, *349*(6244), aaa4019. <https://doi.org/10.1126/science.aaa4019>
- Eicher, O., Baumgartner, M., Schilt, A., Schmitt, J., Schwander, J., Stocker, T. F., & Fischer, H. (2016). Climatic and insolation control on the high-resolution total air content in the NGRIP ice core. *Climate of the Past*, *12*, 1979–1993. <https://doi.org/10.5194/cp-12-1979-2016>
- Fricker, T. E., Oakley, J. E., & Urban, N. M. (2012). Multivariate Gaussian process emulators with nonseparable covariance structures. *Technometrics*, *55*, 47–56. <https://doi.org/10.1080/00401706.2012.715835>
- Fyke, J. G., Weaver, A. J., Pollard, D., Eby, M., Carter, L., & Mackintosh, A. (2011). A new coupled ice sheet/climate model: Description and sensitivity to model physics under Eemian, Last Glacial Maximum, late Holocene and modern climate conditions. *Geoscientific Model Development*, *4*, 117–136. <https://doi.org/10.5194/gmd-4-117-2011>
- Gierz, P., Werner, M., & Lohmann, G. (2017). Simulating climate and stable water isotopes during the Last Interglacial using a coupled climate-isotope model. *Journal of Advances in Modeling Earth Systems*, *9*, 2027–2045. <https://doi.org/10.1002/2017MS001056>
- Gordon, C., Cooper, C. A., Banks, H., Gregory, J. M., Johns, T. C., et al. (2000). The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, *16*, 147–168. <https://doi.org/10.1007/s003820050010>
- Govin, A., Capron, E., Tzedakis, P. C., Verheyden, S., Ghaleb, B., Hillaire-Marcel, C., et al. (2015). Sequence of events from the onset to the demise of the last interglacial: Evaluating strengths and limitations of chronologies used in climatic archives. *Quaternary Science Reviews*, *129*, 1–36. <https://doi.org/10.1016/j.quascirev.2015.09.018>
- Helsen, M. M., van de Berg, W. J., van de Wal, R. S. W., van den Broeke, M. R., & Oerlemans, J. (2013). Coupled regional climate-ice-sheet simulation shows limited Greenland ice loss during the Eemian. *Climate of the Past*, *9*, 1773–1788. <https://doi.org/10.5194/cp-9-1773-2013>
- Holloway, M. D., Sime, L. C., Allen, C. S., Hillenbrand, C., Bunch, P., Wolff, E., & Valdes, P. J. (2017). The spatial structure of the 128 ka Antarctic sea ice minimum. *Geophysical Research Letters*, *44*, 11,129–11,139. <https://doi.org/10.1002/2017GL074594>
- Holloway, M. D., Sime, L. C., Singarayer, J. S., Tindall, J. C., & Valdes, P. J. (2016). Reconstructing paleosalinity from $\delta^{18}\text{O}$: Coupled model simulations of the Last Glacial Maximum, Last Interglacial and Late Holocene. *Quaternary Science Reviews*, *131*, 350–364. <https://doi.org/10.1016/j.quascirev.2015.07.007>
- Johnsen, S. J., Dahl-Jensen, D., Gundestrup, N., Steffensen, J. P., Clausen, H. B., Miller, H., et al. (2001). Oxygen isotope and palaeotemperature records from six Greenland ice-core stations: Camp Century, Dye-3, GRIP, GISP2, Renland and NorthGRIP. *Journal of Quaternary Science*, *16*(4), 299–307. <https://doi.org/10.1002/jqs.622>
- Johnsen, S., & Vinther, B. (2007). Ice core records—Greenland stable isotopes. In S. A. Elias (Ed.), *Encyclopedia of quaternary science* (pp. 1250–1258). Oxford: Elsevier.
- Jolliffe, I. T. (2002). *Principal component analysis*, Springer Series in Statistics (2nd ed.). Berlin, Heidelberg: Springer.
- Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society, B*, *63*, 425–464.
- Kocis, L., & Whiten, W. J. (1997). Computational investigations of low-discrepancy sequences. *ACM Transactions on Mathematical Software*, *23*(2), 266–294. <https://doi.org/10.1145/264029.264064>
- Koerner, R. M., & Fischer, H. (2002). Ice-core evidence for widespread Arctic glacier retreat in the Last Interglacial and the early Holocene. *Annals of Glaciology*, *35*, 19–24. <https://doi.org/10.3189/172756402781817338>
- 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. <https://doi.org/10.1038/nature08686>
- Krinner, G., Raynaud, D., Doutriaux, C., & Dang, H. (2000). Simulations of the Last Glacial Maximum ice sheet surface climate: Implications for the interpretation of ice core air content. *Journal of Geophysical Research*, *105*, 2059–2070. <https://doi.org/10.1029/1999JD900399>
- Landais, A., Chappellaz, J., Delmotte, J., Blunier, T., Bourq, C., Caillon, N., et al. (2003). A tentative reconstruction of the last interglacial and glacial inception in Greenland based on new gas measurements in the Greenland Ice Core Project (GRIP) ice core. *Geophysical Research Letters*, *108*(D18), 4563. <https://doi.org/10.1029/2002JD003147>
- Langebroek, P. M., & Nisancioglu, K. H. (2016). Moderate Greenland ice sheet melt during the last interglacial constrained by present-day observations and paleo ice core reconstructions. *The Cryosphere Discussions*, 1–35. <https://doi.org/10.5194/tc-2016-15>
- Lhomme, N., Clarke, G. K. C., & Marshall, S. J. (2005). Tracer transport in the Greenland ice sheet: Constraints on ice cores and glacial history. *Quaternary Science Reviews*, *24*, 173–194. <https://doi.org/10.1016/j.quascirev.2004.08.020>
- MacGregor, J. A., Fahnestock, M. A., Catania, G. A., Paden, J. D., Gogineni, S. P., Young, S. K., et al. (2015). Radiocarbon and age structure of the Greenland Ice Sheet. *Journal of Geophysical Research: Earth Surface*, *120*, 212–241. <https://doi.org/10.1002/2014JF003215>
- Malmierca-Vallet, I., Sime, L. C., Valdes, P. J., Capron, E., Vinther, B. M., & Holloway, M. D. (2018). Simulating the Last Interglacial Greenland stable water isotope peak: The role of Arctic sea ice changes. *Quaternary Science Reviews*, *198*, 1–14. <https://doi.org/10.1016/j.quascirev.2018.07.027>
- Martinerie, P., Lipenkov, V. Y., Raynaud, D., Chappellaz, J., Barkov, N. I., & Lorius, C. (1994). Air content paleo record in the Vostok ice core (Antarctica): A mixed record of climatic and glaciological parameters. *Journal of Geophysical Research*, *99*, 10,565–10,576. <https://doi.org/10.1029/93JD03223>
- Masson-Delmotte, V., Braconnot, P., Hoffmann, G., Jouzel, J., Kageyama, M., L. A., et al. (2011). Sensitivity of interglacial Greenland temperature and $\delta^{18}\text{O}$: Ice core data, orbital and increased CO_2 climate simulations. *Climate of the Past*, *7*, 1041–1059. <https://doi.org/10.5194/cp-7-1041-2011>
- Montgomery, D. C. (2017). *Design and analysis of experiments* (9th ed.). USA: John Wiley & Sons, Inc.

- NEEM community members (2013). Eemian interglacial reconstructed from a Greenland folded ice core. *Nature*, 493, 489–494. <https://doi.org/10.1038/nature11789>
- NGRIP Project Members (2004). High-resolution record of northern hemisphere climate extending into the last interglacial period. *Nature*, 431, 147–151. <https://doi.org/10.1038/nature02805>
- O'Hagan, A. (1978). Curve fitting and optimal design for prediction. *Journal of the Royal Statistical Society, B*, 40(1), 1–42.
- O'Hagan, A. (1992). Some Bayesian numerical analysis. In J. M. Bernardo, J. O. Berger, A. P. Dawid, & A. F. M. Smith (Ed.), *Bayesian statistics 4*. Proceedings of the Fourth Valencia International Meeting (pp. 345–363). Oxford: Oxford University Press.
- Otto-Bliessner, B. L., Marsha, S. J., Overpeck, J. T., Miller, G. H., Hu, A. X., & CAPE Last Interglacial Project members (2006). Simulating Arctic climate warmth and icefield retreat in the last interglaciation. *Science*, 311(5768), 1751–1753. <https://doi.org/10.1126/science.1120808>
- Pukelsheim, F. (1994). The three sigma rule. *The American Statistician*, 48(2), 88–91.
- Quiquet, A., Ritz, C., Punge, H. J., & Melia, D. S. y. (2013). Greenland ice sheet contribution to sea level rise during the last interglacial period: A modelling study driven and constrained by ice core data. *Climate of the Past*, 9, 353–366. <https://doi.org/10.5194/cp-9-353-2013>
- Rasmussen, C. E., & Williams, C. K. I. (2006). *Covariance functions, Gaussian processes for machine learning, Adaptive computation and machine learning* (pp. 79–104). Cambridge, MA, London, UK: MIT Press.
- Raynaud, D., Lipenkov, V., Lemieux-Dudon, B., Duval, P., Loutre, M.-F., & Lhomme, N. (2007). The local insolation signature of air content in Antarctic ice. A new step toward an absolute dating of ice records. *Earth and Planetary Science Letters*, 261, 337–349. <https://doi.org/10.1016/j.epsl.2007.06.025>
- Robinson, A., Calov, R., & Ganopolski, A. (2011). Greenland ice sheet model parameters constrained using simulations of the Eemian Interglacial. *Climate of the Past*, 7, 381–396. <https://doi.org/10.5194/cp-7-381-2011>
- Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical Science*, 4(4), 409–435.
- Salter, J. M., Williamson, D. B., Scinocca, J., & Kharin, V. (2019). Uncertainty quantification for computer models with spatial output using calibration-optimal bases. *Journal of the American Statistical Association*, 0, 1–15. <https://doi.org/10.1080/01621459.2018.1514306>
- Sime, L. C., Hopcroft, P. O., & Rhodes, R. H. (2019). Impact of abrupt sea ice loss on Greenland water isotopes during the Last Glacial period. *Proceedings of the National Academy of Sciences*, 116(10), 4099–4104. <https://doi.org/10.1073/pnas.1807261116>
- Sime, L. C., Risi, C., Tindall, J. C., Sjolte, J., Wolff, E. W., Masson-Delmotte, V., & Capron, E. (2013). Warm climate isotopic simulations: What do we learn about interglacial signals in Greenland ice cores? *Quaternary Science Reviews*, 67, 59–80. <https://doi.org/10.1016/j.quascirev.2013.01.009>
- Sjolte, J., Hofmann, G., & Johnsen, S. J. (2014). Modelling the response of stable water isotopes in Greenland precipitation to orbital configurations of the previous interglacial. *Tellus B: Chemical and Physical Meteorology*, 66, 22,872. <https://doi.org/10.3402/tellusb.v66.22872>
- Stone, E. J., Lunt, D. J., Annan, J. D., & Hargreaves, J. C. (2013). Quantification of the Greenland ice sheet contribution to Last Interglacial sea level rise. *Climate of the Past*, 9, 621–639. <https://doi.org/10.5194/cp-9-621-2013>
- Sutter, J., Gierz, P., Grosfeld, K., Thoma, M., & Lohmann, G. (2016). Ocean temperature thresholds for Last Interglacial West Antarctic Ice Sheet collapse. *Geophysical Research Letters*, 2675–2682. <https://doi.org/10.1002/2016GL067818>
- Suwa, M., & Bender, M. L. (2008). O₂/N₂ ratios of occluded air in the GISP2 ice core. *Journal of Geophysical Research*, 113, D11119. <https://doi.org/10.1029/2007JD009589>
- Suwa, M., von Fischer, J. C., Bender, M. L., Landais, A., & Brook, E. J. (2006). Chronology reconstruction for the disturbed bottom section of the GISP2 and the grip ice cores: Implications for termination ii in greenland. *Journal of Geophysical Research*, 111, D02101. <https://doi.org/10.1029/2005JD006032>
- Tarasov, L., & Peltier, W. R. (2003). Greenland glacial history, borehole constraints, and Eemian extent. *Journal of Geophysical Research*, 108(B3), 2143. <https://doi.org/10.1029/2001JB001731>
- Tindall, J., Flecker, R., Valdes, P. J., Schmidt, D. N., Markwick, P., & Harris, J. (2010). Modelling the oxygen isotope distribution of ancient seawater using a coupled ocean-atmosphere GCM: Implications for reconstructing early Eocene climate. *Earth and Planetary Science Letters*, 292, 265–273. <https://doi.org/10.1016/j.epsl.2009.12.049>
- Tindall, J. C., Valdes, P. J., & Sime, L. C. (2009). Stable water isotopes in HadCM3: Isotopic signature of El Niño Southern Oscillation and the tropical amount effect. *Journal of Geophysical Research*, 114, D04111. <https://doi.org/10.1029/2008JD010825>
- Tracey, B. D., & Wolpert, D. (2018). Upgrading from Gaussian processes to Student's-t Processes. In *2018 AIAA non-deterministic approaches conference*, Kissimmee, Florida. <https://doi.org/10.2514/6.2018-1659>
- Veres, D., Bazin, L., Landais, A., Toyé Mahamadou Kele, H., Lemieux-Dudon, B., Parrenin, F., et al. (2013). The Antarctic ice core chronology (AICC2012): An optimized multi-parameter and multi-site dating approach for the last 120 thousand years. *Climate of the Past*, 9, 1733–1748. <https://doi.org/10.5194/cp-9-1733-2013>
- Vernon, I., Goldstein, M., & Bower, R. G. (2010). Galaxy formation: A Bayesian uncertainty analysis. *Bayesian Analysis*, 5(4), 619–669. <https://doi.org/10.1214/10-BA524>
- Willerslev, E., Cappellini, E., Boomsma, W., Nielsen, R., Hebsgaard, M. B., Br, T. B., et al. (2007). Ancient biomolecules from deep ice cores reveal a forested southern Greenland. *Science*, 317, 111–114. <https://doi.org/10.1126/science.1141758>
- Williamson, D., Blaker, A. T., Hampton, C., & Salter, J. (2015). Identifying and removing structural biases in climate models with history matching. *Climate Dynamics*, 45, 1299–1324.
- Yau, A. M., Bender, M. L., Robinson, A., & Brook, E. J. (2016). Reconstructing the last interglacial at Summit, Greenland: Insights from GISP2. *Proceedings of the National Academy of Sciences*, 113, 9710–9715. <https://doi.org/10.1073/pnas.1524766113>