Calculations of extreme sea level rise scenarios are strongly dependent on ice sheet model resolution

Check for updates

C. Rosie Williams [1,7] ✉, Pierre Thodoroff[2,7], Robert J. Arthern[1], James Byrne [1], J. Scott Hosking [1,3], Markus Kaiser[2,4], Neil D. Lawrence[2] & Ieva Kazlauskaite[5,6]

The West Antarctic Ice Sheet (WAIS) is losing ice and its annual contribution to sea level is increasing. The future behaviour of WAIS will impact societies worldwide, yet deep uncertainty remains in the expected rate of ice loss. High-impact low-likelihood scenarios of sea-level rise are needed by risk-averse stakeholders but are particularly difficult to constrain. Here, we combine traditional model simulations of the Amundsen Sea sector of WAIS with Gaussian process emulation to show that ice-sheet models capable of resolving kilometre-scale basal topography will be needed to assess the probability of extreme scenarios of sea-level rise. This resolution exceeds many state-of-the-art continent-scale simulations. Our ice-sheet model simulations show that coarser resolutions tend to project a larger range of sea-level contributions than finer resolutions, inflating the tails of the distribution. We therefore caution against relying purely upon simulations 5 km or coarser when assessing the potential for societally important high-impact sea-level rise.

Quantifying the contribution of the Antarctic Ice Sheet (AIS) to global sea level rise over the coming centuries is crucial for policy makers and wider stakeholders due to the huge quantity of ice it contains[1] and the rapid acceleration of ice loss over the last few decades[2,3]. In particular, the West Antarctic Ice Sheet (WAIS) alone contains enough ice to raise mean global sea level by up to 4 metres and it is currently the largest contributor of sea level rise from Antarctica, adding 6.5 mm to mean global sea level between 1992 and 2021[3]. The biggest changes are found in the Amundsen Sea sector of WAIS, which contains two of the most rapidly thinning ice streams, Pine Island Glacier (PIG) and Thwaites Glacier (TG)[4]. These ice streams are particularly vulnerable because the bedrock beneath lies below current sea level and deepens towards the interior of the ice sheet. This can lead to an instability whereby a retreat of the grounding line into deeper water leads to increased ice thickness at the grounding line, which in turn leads to further ice acceleration and increased grounded line retreat (the 'Marine Ice Sheet Instability' (MISI)[5,6]. This process may be irreversible on human timescales once it has begun[7,8], and these ice streams may soon or may already have passed a tipping point[9]. The role of submarine melting within this system is well documented: basal melt below a floating ice shelf causes the shelf to thin, reducing the buttressing of the upstream ice which is grounded on bedrock and causing it to thin, accelerate and retreat[10–13]. Basal melting of the floating ice shelves accounts for approximately half of the mass loss from Antarctica [with iceberg calving approximately accounting for the other half,[14,15]]. In forward regional simulations of the AIS, the amount of basal melting under the floating ice shelves has a large effect of how much ice is lost[7,16–18], and can trigger the MISI.

Despite the clear importance of the ice sheets, ice loss from Antarctica is the largest uncertainty in projections of sea level change over the next century, and deep uncertainty remains[19,20]. This is in part because key physical processes are challenging to incorporate into ice sheet models, and because such models rely on parameter choices that are not well constrained[21,22]. In addition, the future forcing is also unknown and subject to climate model uncertainties and biases[23], and ice-sheet models are sensitive to variations in initialisation procedures[24]. Together, this means large ensembles of different forward simulations are required to make probabilistic projections of sea level contribution (SLC), and the high-impact low-likelihood (HILL) scenarios are particularly difficult to constrain as they occur in the less sampled upper tails of the distributions[19,25]. Furthermore, to incorporate key processes and feedbacks, coupling of different elements of the Earth system such as ice-ocean modelling[26–28] and ice-atmosphere models[27] are required, and physics-based models of processes such as ice fracture need to be incorporated[29,30]. While necessary for making more realistic projections, such modelling strategies massively increase the computational costs and reduce the number of model simulations that can feasibly be produced.

[1]British Antarctic Survey, Cambridge, UK. [2]Department of Computer Science and Technology, University of Cambridge, Cambridge, UK. [3]The Alan Turing Institute, London, UK. [4]Proxima Fusion, Munich, Germany. [5]Department of Engineering, University of Cambridge, Cambridge, UK. [6]Department of Statistics, London School of Economics and Political Science, London, UK. [7]These authors contributed equally: C. Rosie Williams, Pierre Thodoroff. ✉e-mail: chll1@bas.ac.uk

In terms of ice sheet model simulations, much progress has been made in recent years, with continent-scale projections of sea level change produced either by a specific model[31–33] or as part of a multi-model ensemble under certain assumptions[34,35]. In a previously unprecedented effort, the ISMIP6 community generated a multi-model ensemble of sea level projections for Antarctica under common future forcing scenarios, generating over 300 simulations of the Antarctic Ice Sheet[36]. While this is a very valuable multi-model ensemble, the complexity and the large computational cost of running these simulations limits our ability to explore the parameter space and calculate uncertainties. A complimentary approach is to use some form of model emulation to move from a relatively small number of model simulations to a much larger ensemble of predicted model realisations, potentially under new forcings, thus enabling probabilistic projections that account for parametric uncertainty more robustly than could otherwise be achieved. Using this approach, Edwards et al.[20] take the ISMIP6 ensemble generated by Seroussi et al.[36] together with additional simulations and use statistical emulation to produce projections that explore the known uncertainties more fully under new climate forcings without the need to run new computationally expensive simulations. Similarly, Berdahl et al.[37] also use statistical emulation, this time of a specific ice-sheet model, to better quantify uncertainties in Antarctic sea level rise, further highlighting the ability of emulation to produce projections over much larger ensembles than would be possible by direct simulation.

Spatial resolution can also be a limiting factor when generating ensembles of ice sheet simulations. Previous model inter-comparisons on idealised geometries have shown that sufficiently fine resolution is required to capture the physical processes and reproduce either analytical solutions or expected advance and retreat under specific scenarios[38–40]. In particular, even if subgrid interpolation of the grounding line is employed, sufficiently fine resolution (finer than 5 km) is required to resolve the grounding line[39]. Thus, although some models now have sophisticated adaptive mesh schemes[41,42] or sub-grid parameterisation of the grounding line[18,43–45], the requirement to run at high resolution still hampers the construction of sufficiently large ensembles for probabilistic forecasting. This is especially true when these ensembles have to be re-simulated as new forcings are computed from climate models, or as new data become available to refine a likely parameter range. Thus, many ice sheet models are still run at relatively low resolutions (coarser than 4/5 km) when running simulations of the AIS[35,36].

Machine learning techniques have also been used to interpret sea level projections using attribution approaches. Rohmer et al.[46] take the ISMIP6 multi-model ensemble for the Greenland Ice Sheet (GRIS) and assess the effects of various choices of numerical model, model initialisation, ice-sheet processes and environmental forcing on the SLC projections. Within this framework they find that using a coarse spatial resolution influences the projections, and that this may dominate all other modelling assumptions. Other studies also find that without fine resolution, local maxima and minima in areas of complex bed topography are not well represented and simulations of sea-level contribution vary with resolution[16,47,48]. Thus, there is a clear need to run ice sheet models at fine resolutions (at least 4 km), despite the computational challenges this entails. Both Rohmer et al.[46] and Edwards et al.[20] also make a strong argument for using experimental design protocols to improve the generation of ensembles and reduce the uncertainties in SLC in the future[49].

In this study we conduct a systematic evaluation of the effects of spatial model resolution and basal melt rate on probabilistic estimates of sea level contribution, with a particular focus on how model resolution affects the high-impact low-likelihood scenarios (the upper 'tails' of the resulting probability distributions). This is particularly important for risk-averse stakeholders who are especially concerned with these 'worst case' scenarios[25,50,51], because they pose the greatest risks and most expensive adaption challenges, especially when considering the long term commitment to sea level beyond the year 2100 and the associated levels of mitigation needed. To do this, we use a multi-resolution model simulation and emulation approach that explores the effects of basal melt, one of the most important processes leading to sea-level rise in WAIS[52]. Specifically, we combine traditional physics-based modelling with Gaussian process emulation to quantify the effects of model resolution on probabilistic distributions of sea level contribution, in particular focusing on the high-impact low-likelihood scenarios. We deploy our ice sheet model and emulator over a realistic and rapidly changing domain: the Amundsen Sea sector of WAIS. Using a realistic rather than an idealised domain ensures the complex real world behaviours are simulated in the ensemble. We show that the upper tails of the SLC distributions are highly sensitive to the spatial resolution of the ice sheet model. We thus caution against relying only on coarse (coarser than 5 km) resolution simulations when considering extreme scenarios.

Results

Ice sheet model simulations dependent on basal melt and spatial model resolution

We use a state-of-the-art ice sheet model WAVI to produce the simulations of mass loss that are used to build a Gaussian process emulator. This emulator is then used to quantify the effects of model resolution, R, on sea level contribution (SLC). WAVI is a vertically integrated, three dimensional ice sheet model which includes both the membrane stresses in the ice and the effects of vertical shear in order to simulate flow of both grounded and floating ice[18,53,54]. We use the Amundsen Sea sector of the WAIS as the model domain, as shown in Fig. 1.

There are many uncertain parameter values that are required as inputs to ice sheet models, and realistic projections of SLC need to account for as many of these uncertain inputs as possible. Here, we focus on the process of basal melting of the underside of the floating ice shelves. Although coupled ice-ocean models can be used to calculate basal melt rates in forward simulations[26,28], these models are complex and computationally expensive and would massively limit the number of ensemble members we could produce[55]. Thus, we choose to use a well known depth-dependent parameterisation for basal melt, for which a fixed temperature-salinity profile is selected and a calibration coefficient, γ_T, is used to set the average melt rate over all fully-floating points across the ice domain at the start of each simulation, denoted M_{t_0}[56,57], see Methods]. It is this parameter M_{t_0} (directly linked to γ_T) that we choose to vary between simulations. Within each simulation, the melt rate varies with depth such that it is non-constant over the shelves, and the resulting average melt rate varies over time as the geometry of the ice sheet evolves; it is only the parameter γ_T that is held fixed throughout each simulation.

M_{t_0} is varied between 0 to 150 m/a in intervals of 12.5 m/a. This large range with values considerably higher than those observed is used to build the emulator and to ensure improbable scenarios are accounted for, which is important for defining the upper tails of the resulting probability distribution. We chose to vary only this parameter M_{t_0} along with model resolution because this gives sufficient and manageable ensembles on which to analyse the effects of resolution. More realistic forcings could vary the underlying temperature and salinity directly. As such we are exploring these effects in a limited ensemble rather than making projections. However, the method is easily extendable to larger ensembles in which a wider range of parameters are sampled.

We use Bedmachine version 3[58,59] for the bed topography (and initial ice thickness and surface elevation at $t = 0$ years) provided at 500 metre spatial resolution. We run an ensemble of simulations at a range of resolutions $R = 8$ km to $R = 2$ km (Fig. 2 shows example plots of grounding line retreat for two different resolutions). At each R, n different Bedmachine realisations are sampled by choosing points along the diagonal of the square of possibilities (see Methods for details). This sampling method assumes that the bed is isotropic and that the noise is uncorrelated in the bed topography. Due to computational cost increasing considerably with resolution, we sample more Bedmachine realisations for coarser resolution simulations than finer resolution, with a minimum of $n = 8$ (see Methods for details).

Separate initialisations and relaxations are performed for each Bedmachine realisation at each resolution in order to start with an ice sheet configuration that captures the grounding line position, surface velocities
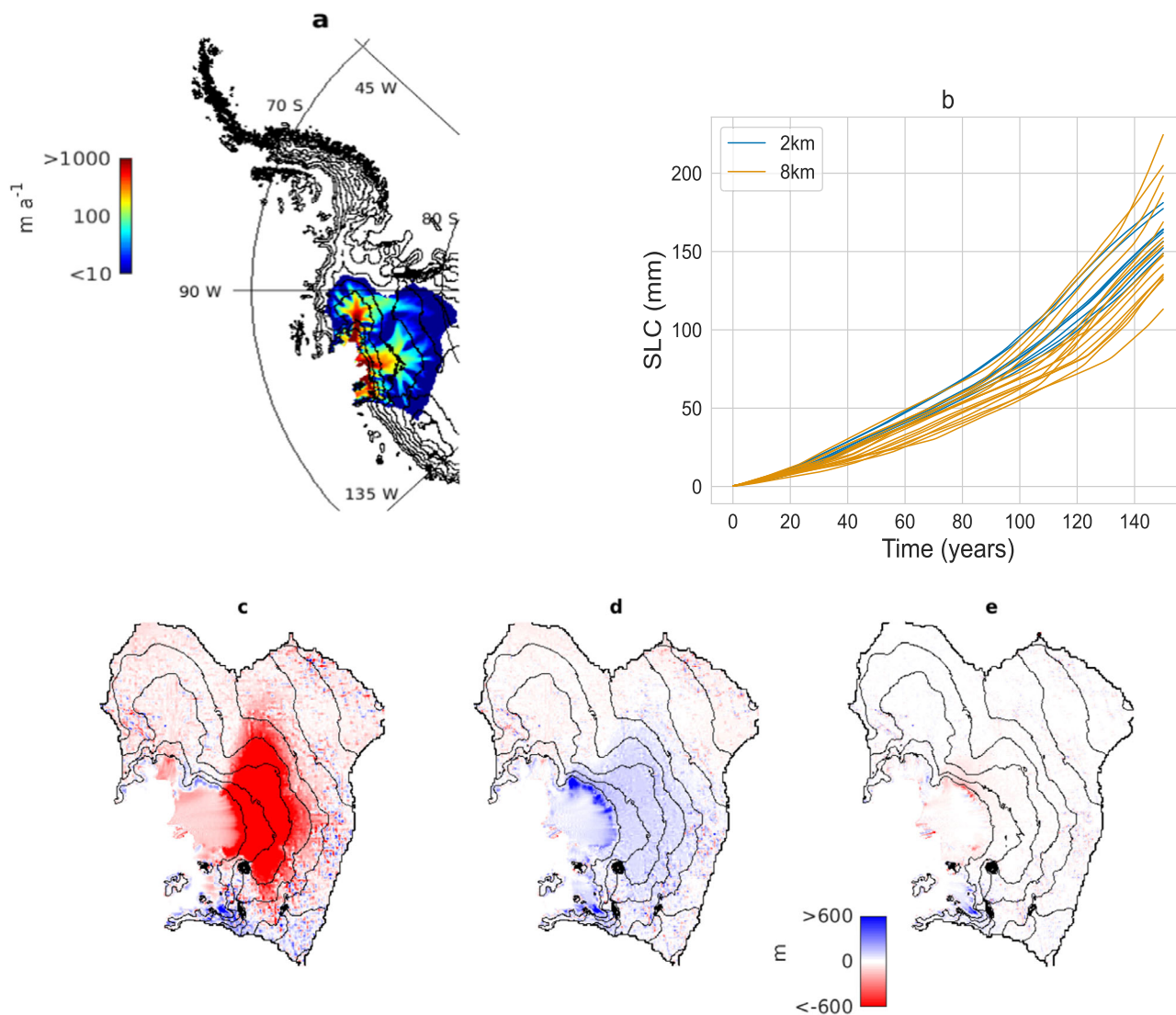
Fig. 1 | Geographical model domain for the Amundsen Sea Sector of West Antarctica and examples of the effects of spatial model resolution on mass loss. **a** Model domain and initial state at 2 km resolution from data assimilation: colours show ice velocities (m a^{-1}) and elevation contours are at 200 m intervals. **b** Example WAVI simulations of SLC over 150 years for $M_{t_0} = 25$ m a^{-1} and $R = 8$ km (orange lines) and $R = 2$ km (blue lines), with each line representing a different Bedmachine realisation at resolution R. Differences in simulated ice thickness at $t = 150$ years between a 2 km and an 8 km (**c**), 5 km (**d**) and 3 km (**e**) resolution Bedmachine realisation for $M_{t_0} = 25$ m a^{-1}. In each case, the coarser resolution thickness is interpolated on to the 2 km grid and thickness at 2 km is subtracted from this.

and rate of thinning of grounded ice for approximately the year 2015 (see Methods and Supplementary Note 1 for full details). WAVI is then run forward for 150 years using a persistence-based forcing in which accumulation remains constant over each run, the ice front does not move, and the thermocline profile and melt tuning parameter γ_T in the depth-dependant melt-rate parameterisation do not change within the forward simulation, but average melt rates over the shelves can vary, as described above (see also Methods). Sea level contributions at every year are averaged over all Bedmachine realisations for each resolution and melt rate, giving a measure of uncertainty on the SLC due to bed topography for a given M_{t_0} and R. This is shown by the shading in Fig. 3, and by the results of the WAVI simulations in in Fig. 4, where each black cross represents the SLC at time $t = 150$ years for a given M_{t_0} and a given Bedmachine realisation at resolution R. It is clear that model resolution has a large affect at all melt rates: for every M_{t_0}, a much larger range of SLCs are found at coarse resolutions (coarser than 5 km) than at fine resolutions i.e. the uncertainties decrease roughly with resolution. The magnitude of this effect is also notable in the example shown in Fig. 1b, where for $M_{t_0} = 25$ m a^{-1} the range of simulated SLC over time is over 3 times larger for $R = 8$ km compared to $R = 2$ km. This is because at finer

resolution each realisation captures more of the fine scale bed topography[48], and there is less variation between different realisations. These plots emphasise the difficulty in taking one Bedmachine realisation at each resolution when using low spatial resolution.

Figure 3 shows the effect of melt M_{t_0} on SLC over time for (a) 8 km resolution and (b) 2 km resolution. As expected, in both cases, higher melt rates lead to more rapid loss of grounded ice and higher SLC. In the no melt case, $M_{t_0} = 0$ m/a, there is substantial sea level rise. While this may indicate a tipping point has passed, SLC may be overestimated in this case due to the lack of an evolving damage field limiting the healing and regrowth of the shelf[28]. When $M_{t_0} \geq \approx 100$ m a^{-1}, the SLC converges. In these cases of extreme melt, the ice shelves are essentially melted so quickly that it becomes an unbuttressed ice sheet and further melt does not increase the rate of grounding line retreat. In these plots the SLCs at each time point are averaged over all bed realisations at the specified resolution and the shaded area corresponds to the uncertainty. Comparison of panels (a) and (b) shows that there are higher uncertainties at coarse resolution (8 km) compared to fine (2 km), which is also indicated by the spread of the SLCs from WAVI in Figs. 1b and 4.
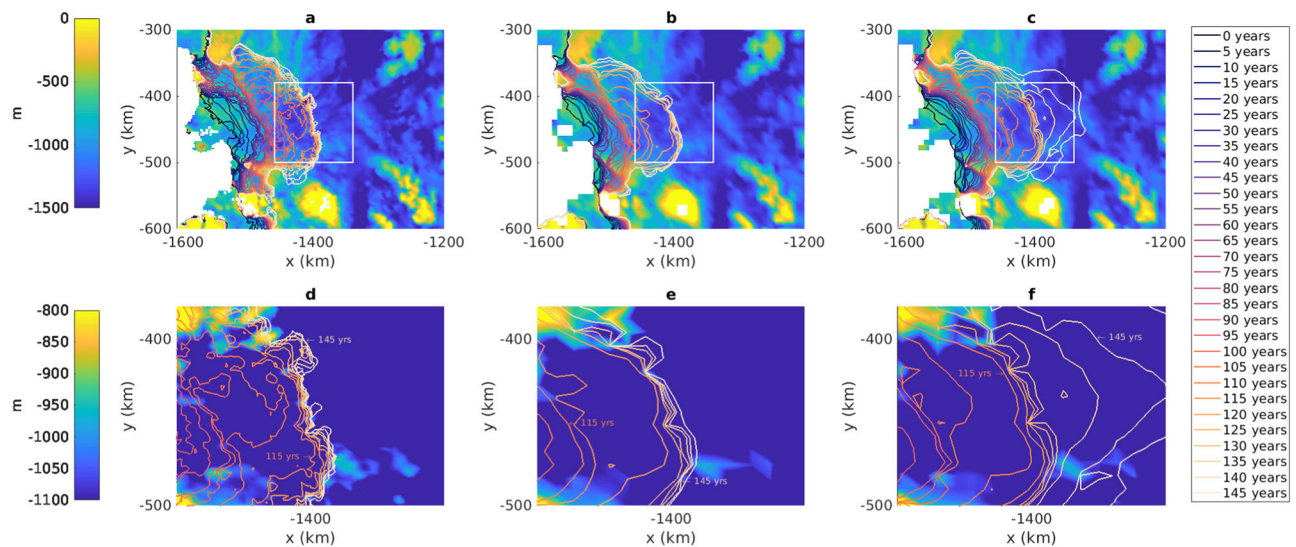
Fig. 2 | Bed topography and grounding line retreat plots for Thwaites Glacier over 150 years. Simulations with M = 25 m a^{-1} are shown in (**a, d**) for a 2 km simulation, and in (**b, e**) and (**c, f**) for two different 8 km simulations. Grounding lines are shown every 5 years, with colours as shown in legend, going from black at $t = 0$ to yellow at $t = 145$ years. **d–f** show a zoom in of the area in the white box in panels (**a–c**).

To explore why coarse resolution can produce such a spread of different SLC for the same M_{t_0}, in Fig. 2 we compare an example of grounding line retreat over bed topography when $M_{t_0} = 25$ m a^{-1} for one 2 km Bed-machine realisation (a,d) and two different realisations at 8 km: one that approximately matches the 2 km mass loss at the end of the simulation (b,e), and one which produced much higher SLC (c,f). Pinning points at $x \approx -1400$ km and -500 km$< y < -400$ km are more pronounced in the 8 km simulation shown in panels (b,e), and the grounding line stays in this area for a longer time for this realisation compared to in panels (c,f), which has fewer and less pronounced pinning points (approximately 30 years compared to 15 years). This leads to substantially different SLC at t = 150 years in these two cases, and explains the divergence between the 2 km and 8 km curves in Fig. 3.

Gaussian process emulation

In order to produce useful information from ice sheet simulations, we need to move from ensembles of WAVI simulations under different conditions to probabilistic predictions of specific scenarios of sea level change. To do this, it is vital to account for known uncertainties in both the parameter choices and the input data of WAVI. Here, this corresponds to the melt average M_{t_0} and the different Bedmachine realisations at different resolutions. Practically, this translates into uncertainty in the predictions of SLC. To provide information for decision makers, one way is to produce probability density functions (PDFs) that describe the probability of observing a specific amount of sea level increase (within a range) at a specific point in time t. Such PDFs can be obtained by defining a prior over the melt-average M_{t_0} and integrating WAVI's prediction (described below). To make this problem computationally tractable, we leverage WAVI simulation runs to fit a probabilistic model using Gaussian Processes (GP) to the distribution of sea level contribution at time t as a function of M_{t_0} (see 'Gaussian Processes' for full details). Here, emulation via Gaussian Processes acts as a surrogate model that predicts the sea level contribution based on a given melt rate and resolution. The predictions are obtained by calculating the mean and covariance matrices for the Gaussian process model. These matrices summarize the relationships between the observed melt rates and sea level contributions. By utilizing these relationships, the surrogate model can make probabilistic predictions of the sea level contribution for new melt rates (see 'Sea level contributions from emulation' for more detail). This offers an effective, efficient, and accurate alternative to exhaustive simulations, allowing us to quickly explore the effects of different assumptions regarding the prior distribution of melt rates. This strategy is similar to the one used by Berdahl et al.[37] (Fig. 1) and allows us to sample the distribution much more than would be possible using WAVI alone, which is especially useful for defining the low-probability high-impact upper tails of the distributions.

Results of the GP emulation are displayed in Fig. 4 at $t = 150$ years. To obtain an estimate of the PDF, we optimize the hyper-parameters of the Gaussian Process (length scale and variance) using a maximum likelihood objective (see 'Maximum likelihood estimation'). The predicted mean is shown along with the results of the WAVI simulations that represent the SLC at time t for a given Bedmachine realisation at resolution R (black crosses). The dependency on resolution of the uncertainty arising from the different samplings (as discussed above) is clear to see: the confidence bands are much tighter when resolution is finer than 4 km, resulting in confidence bands of approximately 200 mm for 8 km and <50 mm for 2 km at $t = 150$ years at 97.5th percentile.

To construct probability density functions (PDF) over the SLC at a set timepoint, a non-uniform prior is used to account for expert knowledge that not all values of the melt parameter are equally likely. The choice of a log-normal prior is standard in such cases, and is described by two parameters, μ and τ, which are the mean and standard deviation, respectively, of the normal distribution that results when the logarithm of the melt rates is taken. Based on recent observations, an average melt rate of 20 m a^{-1} gives an appropriate approximate estimate of present day mean value[15,60]. We explore the sensitivity of the resulting probability density functions to the prior by also selecting $M_{t_0} = 10$ m a^{-1} and $M_{t_0} = 40$ m a^{-1}; higher (lower) mean values for the prior represent the effect of increasing (decreasing) melt rates in the near future. All priors selected account for the fact that the very high end of the range of melt rates used to construct the emulator are not likely to occur in reality (see section 'Gaussian Processes' for full details).

Results for all three priors are shown in Fig. 5: the top row illustrates the probability density function (PDF) over the SLC at $t = 150$ years. These distributions are derived through the integration of Gaussian Process (GP) predictions based on the provided prior of the melt average. As the posterior is not available in closed form for a log-normal prior, we estimate the posterior through numerical methods by initially sampling $N = 1000$ melt averages from the log-normal prior. Subsequently, for each melt average, we determine the approximate normal distribution using the GP and sample $K = 1000$ points from this distribution. We obtain the probability density function by fitting a kernel density estimator[61,62] to the gathered data set, which consists of 1,000,000 data points ($N \times K$).
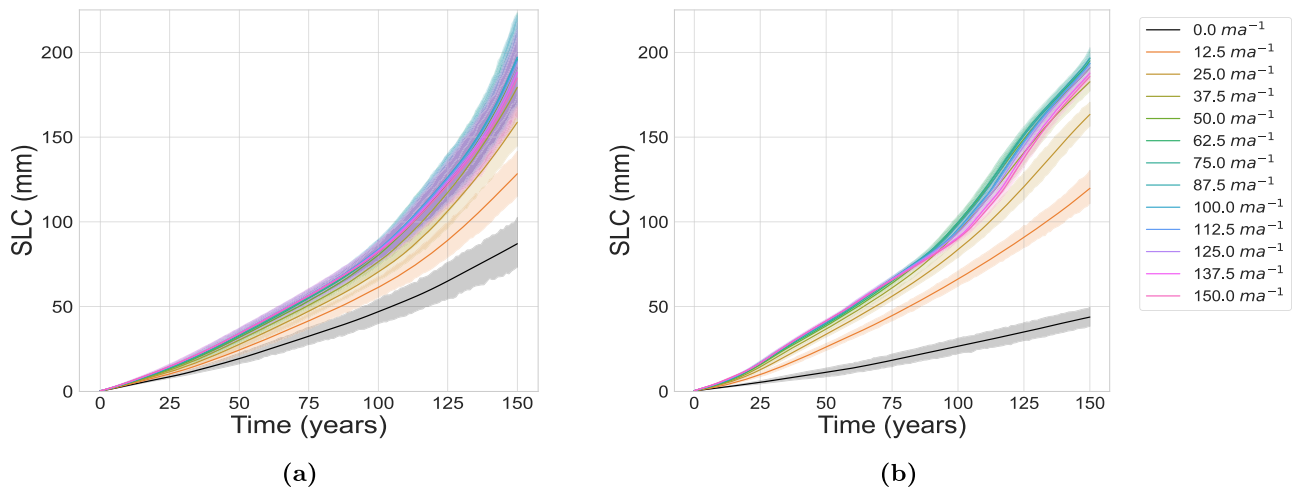
Fig. 3 | Sea level contribution over 150 years for a range of average melt rates. a Sea level contributions (mm) for $M_{t_0} = 0$–150 m a^{-1} with a model resolution of 8 km. **b** Sea level contributions (mm) for $M_{t_0} = 0$–150 m a^{-1} with a model resolution of 2 km. Each solid line represents the average over all Bedmachine realisations at a set resolution and melt rate and the shaded area corresponds to the 95th percentile of the data. Coloured lines represent different average melt rates: $M_{t_0} = 0$ m a^{-1} (black lines), $M_{t_0} = 12.5$ m a^{-1} (orange), $M_{t_0} = 25.0$ m a^{-1} (yellow), with other colors shown in the legend.
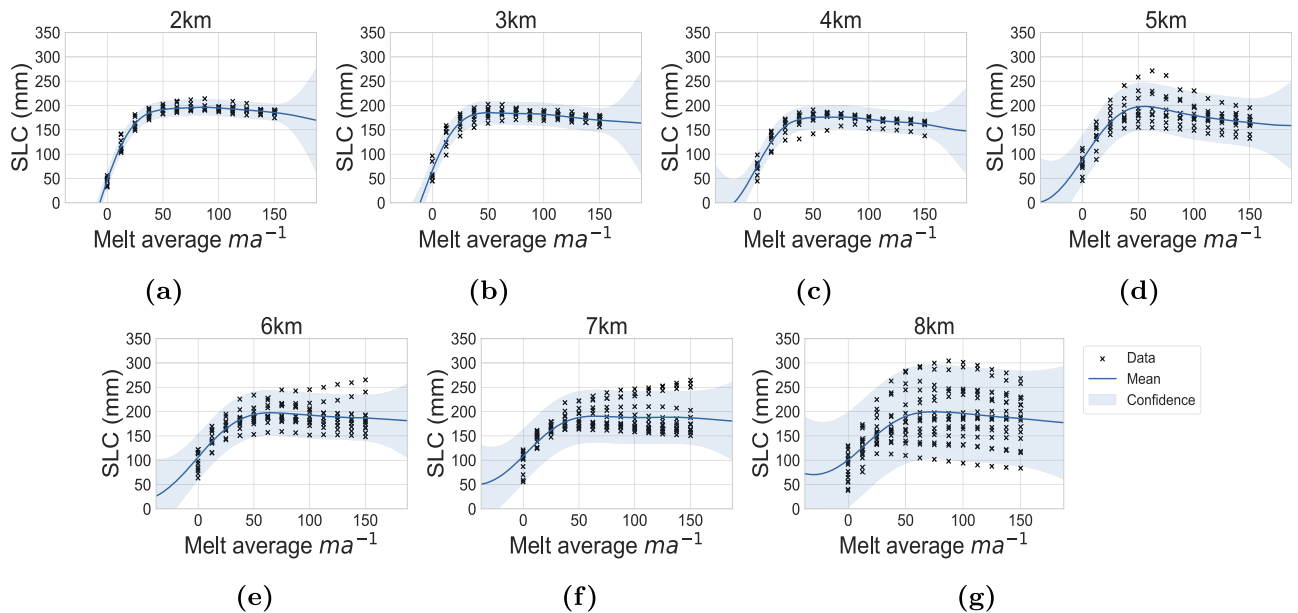

Fig. 4 | Gaussian process (GP) emulation for sea level contribution at $t = 150$ years. a–g SLC computed as a function of the average melt rate at $t = 0$, denoted M_{t_0}: a GP model was fitted separately for each resolution from $R = 2$ km to $R = 8$ km. The black crosses show the WAVI output, the blue line is the predicted mean and the shaded area shows the confidence of the GP model (97.5th percentile of the data).

For the prior distribution with mean $M_{t_0} = 10$ or $M_{t_0} = 20$ m a^{-1}, the coarse resolutions (coarser than 4 km) slightly overestimate the SLC after 150 years, whereas for higher prior distribution mean $M_{t_0} = 40$ m a^{-1}, the means of the distributions are fairly similar. However, for coarse resolutions, the larger range of uncertainties shown in Figs. 3 and 4 results in broader distributions, particularly for prior of 20 m a^{-1} or greater. At resolutions of 5, 6 and 7 km, for $t = 150$ years, the predicted means and overall distributions are similar while 8 km has a lower mean and broader distribution.

The upper tails of PDFs are of particular interest, since they represent high-impact low-likelihood (HILL) scenarios. In Fig. 5, these tails vary considerably with resolution: the high-impact SLC scenarios are assessed to have greater probability if coarse resolution is used (5 km or coarser). For our model, resolution of 4 km appears to be a threshold between fine and coarse resolutions, such that the upper tails for 4 km or finer are relatively similar. At coarse resolution, these broad distributions with long tails are not

only caused by large SLCs under unrealistically high melt M_{t_0}: large SLCs are also predicted for some bed realisations for more realistic values such as $M_{t_0} = 25$ m a^{-1} (e.g. at 8 km resolution with $M_{t_0} = 25$ m a^{-1}, SLC > 200 mm after 150 years). The use of priors with very little weight for high melt scenarios ($M_{t_0} > 100$ m a^{-1}, see section 'Sea level contributions from emulation') causes these extreme scenarios to be sampled much less frequently than the realistic melt scenarios that can also produce large SLCs at coarse resolution. Thus, both high melt and lower, more realistic scenarios of melt contribute to the increased probability of HILL scenarios at coarse resolutions.

The effects of resolution on the tails is amplified over time: the difference in probability between coarse and fine resolution models of a specific high-impact scenario occurring increases between $t = 100$ and $t = 150$ years. Table 1 summarises this information for the prior distribution with mean $M_{t_0} = 20$ m a^{-1}: the dependence on resolution of the probability exceeding
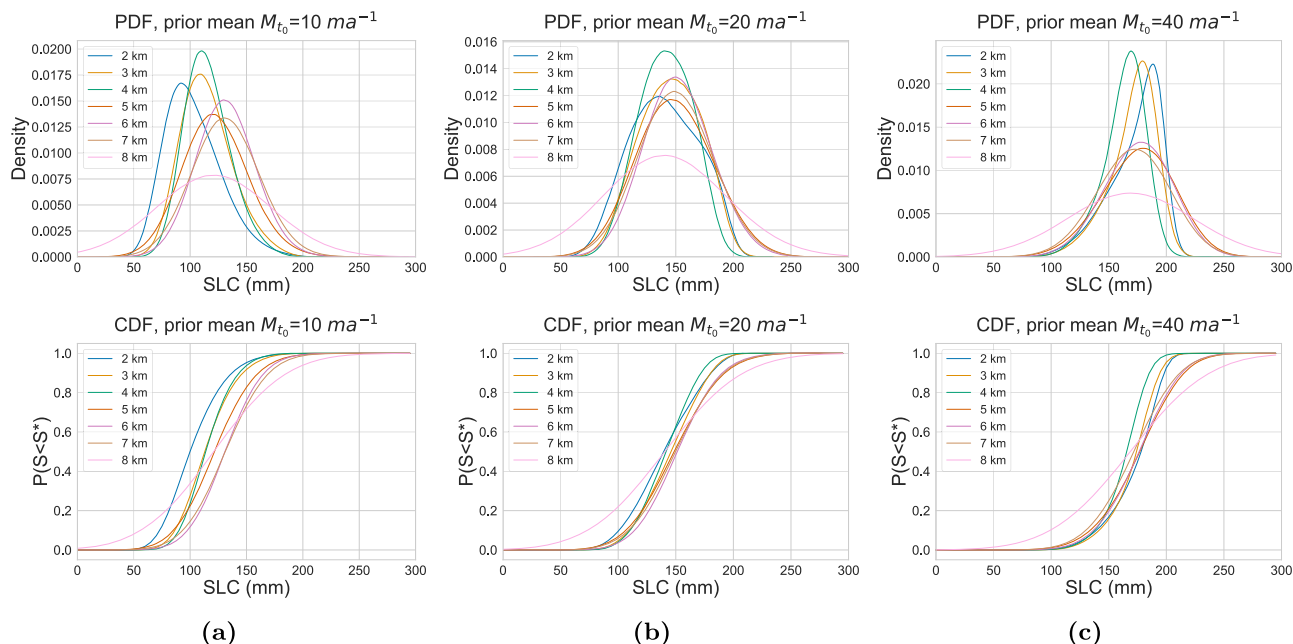
Fig. 5 | Probability density functions (PDFs) and cumulative density function (CDFs) for sea level contribution at $t = 150$ years. Prior distribution means of (**a**) $M_{t_0} = 10$ m a^{-1}, (**b**) $M_{t_0} = 20$ m a^{-1}, and (**c**) mean $M_{t_0} = 40$ m a^{-1} are shown. $P(S < S^*)$ describes the probability of drawing a sample S that is less than a given value (S^*) selected on the x-axis. Colours indicate resolution: 2 km (blue), 3 km (orange), 4 km (green), 5 km (red), 6 km (purple), 7 km (brown), 8 km (pink).

a given threshold is displayed, where the threshold is chosen as the 95th percentile for $R = 2$ km. In this case, at $t = 100/125/150$ years there is a 5% chance of exceeding 107 mm/161 mm/206 mm when $R = 2$ km, and a 2.5–8.2% for $R \geq 5$ km, rising to 6.6–11.2% at $t = 125$ years and 20.8–34.2% at $t = 150$ years (see also Supplementary Fig. 6). For all priors explored, the overestimation of the tails at resolutions coarser than 4 km was observed. Increasing the mean of the prior distribution slightly amplifies the difference in the upper tails between resolutions (Fig. 5). These results indicate that the overestimation of the upper tails of the SLC distributions (the HILL scenarios) at coarse resolutions is not dependent on the prior used.

Figure 6 shows the root mean squared error between the $R = 2$ km and each resolution $R = 3$–8 km for both the PDF and CDF at each year from $t = 0$–150 years, allowing an analysis of the error caused by resolution over time. These errors are calculated over the full distributions and calculating the RMSE error between PDFs puts more emphasis on the higher moments of the distribution compared to calculating the RMSE error between CDFs. Thus, the non-linear effects of resolution in the full PDFs (including means) in Fig. 5 result in some noise in the errors over different resolutions in Fig. 6a (i.e. errors do not increase monotonically with resolution). For the CDF errors (panel b), the clear trend in upper tails seen in Fig. 5 results in a more uniform trend of error with resolution: there is an increasing divergence (within noise) between the SLC distributions at coarse and fine resolution (2 km) over time, indicating the sensitivity of HILL scenarios to resolution. The exception is that for 8 km resolution, the CDF error only starts to increase markedly after 100 years. Inspecting the PDFs and CDFs at t = 100 years (see Supplementary Note 2 and Supplementary Fig. 6), there is relatively high agreement between the 2 km and 8 km in terms of both mean and variance at $t \leq 100$ years, which appears to be coincidental and diverges at later times.

Sea level contribution estimation with limited budget

We have highlighted how fine resolution simulations (finer than 5 km) are critical for obtaining accurate estimates of SLC. However, due to the large run time required for such simulations, it is often not computationally tractable to run a sufficient number of simulations for a range of parameters to fully characterise the probability distributions at fine resolutions. To address this, in this section we examine the efficacy of estimating SLC within the constraints of a finite budget. For example, given a computational budget

of T, you could potentially run either three 2 km simulations or twenty 8 km ones. The key question here is: which of these options will yield a more accurate estimate of sea level contribution?

In practical terms, we vary the available computational budget T (defined in section 'Cost-constrained estimation'), select the number of points this budget allows for at a certain resolution, and compute the discrepancy between the predicted probability distribution function and the PDF calculated at 2 km resolution that incorporates all points (with no computational constraint, hence acting as the ground truth distribution). We estimate the discrepancy by calculating the root mean squared error (RMSE) between the pairs of distributions.

Figure 7 shows the error in estimating the PDF at $t = 150$ years with limited computational budget. Overall this shows that for all priors explored, with sufficiently high computational budget, the error is not decreased by performing a large number of coarse resolution simulations. Thus, with sufficiently high budget it is advisable to run the simulations at the finest resolution possible. However, there is a region of the budget space (up to 10^3 units) in which the RMSE is lower if coarser simulations are selected. This indicates that there are some instances under a constrained budget when it can potentially be advantageous to run more coarser resolutions than fewer finer resolutions simulations. This trend is the same for all priors explored, but the magnitude of the error and, thus, the choice of which resolution to run under a constrained budget is also dependent on the choice of prior. This signifies that the choice of resolution needs to be data-driven, based on collected samples[49], and points to the potential power of a fully multi-fidelity approach, in which the choice of resolution can be guided by consideration of both the computational cost and accuracy of coarse resolution simulations (see also 'Discussion').

Discussion

It has already been documented that coarse resolution simulations do not fully capture the dynamics of ice flow across the grounding line, even if subgrid or parameterisation schemes are used[38,39,44]. However, due to the large computational cost of running fine resolution simulations over large domains such as the Greenland and Antarctic Ice Sheets, and the need for projections of the sea level contribution from these vast areas, low resolution simulations (5–32 km) are often used in ensembles to make probabilistic

Table 1 | The probability of SLC exceeding a given threshold

Year	Probability (%) at various resolutions (km)							Threshold (mm)
	2 km	3 km	4 km	5 km	6 km	7 km	8 km	
100 years	5.0	2.2	0.1	3.7	7.2	8.2	2.5	107 mm
125 years	5.0	2.5	0.2	6.6	11.2	7.6	7.0	161 mm
150 years	5.0	2.5	0.4	23.5	24.3	20.8	34.2	206 mm

Probabilities are shown as percentages. Here, the mean of the prior distribution is 20 m a^{-1} and the threshold is picked as the 95th percentile of the 2 km resolution. This table illustrates the overestimation of SLC by coarse resolution runs.
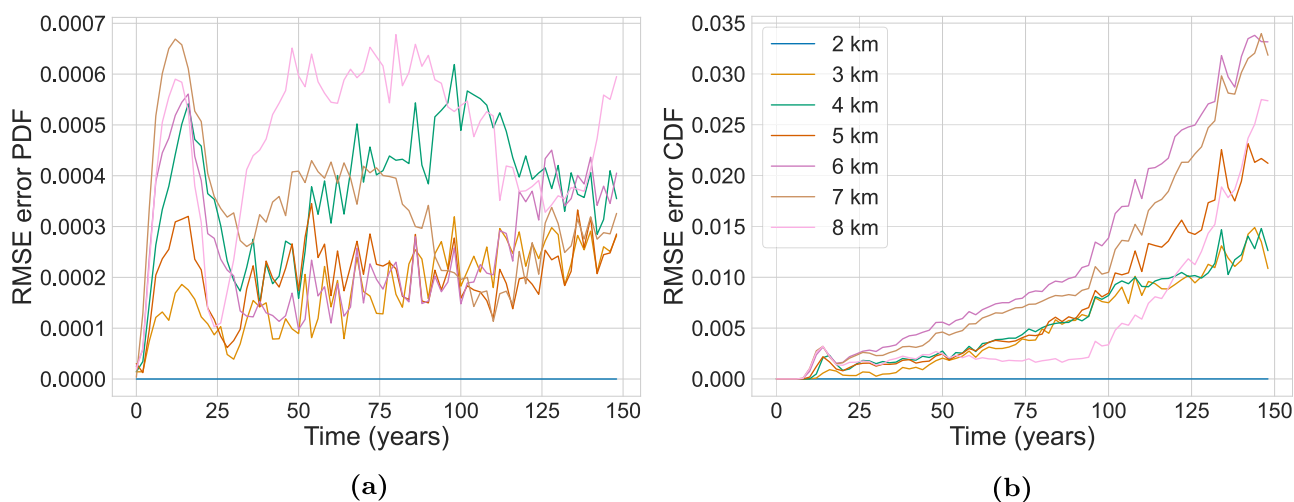

(a)

(b)

Fig. 6 | Error variation with resolution over time. a The dimensionless RMSE between the optimal 2 km PDF and coarser resolutions. **b** The discrepancy between the optimal 2 km CDF and coarser resolutions. In both panels, the melt average is M = 20 ma^{-1}. Colours indicate resolution: 2 km (blue), 3 km (orange), 4 km (green), 5 km (red), 6 km (purple), 7 km (brown), 8 km (pink).

projections[20,36]. In this study, we have shown the specific and considerable effects that coarse spatial resolution (5 km or coarser) has on such probability distributions, in particular in the high-impact low-likelihood upper tails of these distributions. This was shown by using a method in which we used a traditional ice-sheet model to produce simulations over the most rapidly thinning region in Antarctica, the Amundsen Sea sector of the West Antarctic Ice Sheet. These simulations were used to construct a Gaussian process emulator with which to calculate probability density functions of sea level contribution.

For risk-averse end-users, the upper tails of the probability distributions of SLC are hugely important: for example, in coastal risk management, those living in low lying, densely populated regions have a strong desire to protect against not just the mean sea level rise but also the worst-case scenarios, given the huge impact such scenarios would have on lives and livelihoods[50]. If a cost-benefit analysis framework is used, then interpretation of upper tails can lead to strategies of proactively retreating or protecting areas[63]. In the simulations presented here, coarse resolution models assign overly high probabilities to very high sea level contributions: for example, at 150 years, the chance of exceeding 206 mm is approximately 5% at 2 km resolution but over 34% at 8 km. This considerable level of uncertainty based on resolution makes decision-making based on projections from the upper tails extremely risky if sufficiently fine resolution is not used. It could lead to economically inefficient decisions and policies, such as the unnecessary abandonment of an asset or area, or a huge financial waste in unnecessarily protecting against an event that is extremely improbable. If other model processes and uncertain parameters were also included it is possible that fine resolution simulations may exhibit dynamics not found at coarse resolutions, and this may lead to further differences in weight in the upper tails between resolutions. Furthermore, other models have found that fine resolution simulations (2 km) yielded more sea level compared to

coarse resolution (8 km, 4 km)[16], and, thus, using coarse resolution results could underestimate the probable risk of sea level rise. Thus, we expect the effects on model resolution on the upper tails to be model dependent. However, using projections with too little rather than too much weight in the upper tails would also be problematic for planners, as potentially dangerous strategies of insufficient adaption may be implemented. For these reasons, we strongly advise caution when using coarse resolution simulations in decision-making concerning high-impact scenarios, and recommend that modellers assess the impact of resolution and bed sampling/interpolation specifically on these scenarios.

Part of the uncertainty in these estimations relates to the specific Bedmachine realisation, particularly at low resolution: for example, at 8 km resolution, for $M_{t_0} = 50$ m a^{-1}, over all realisations the highest SLC after 150 years is more than double that of the lowest. Previous studies find that small scale features in the bed ranging from a few kilometres to sub-kilometer scale can have a large effect on the current and future stability of the ice sheet as they can form pinning points on which the ice shelf can ground, increasing buttressing and slowing the flow of the ice and the retreat inland[16,64,65]. In accordance with this, the difference in bed topography between different realisations at coarse resolution can result in differences in local maxima and minima (e.g. sub-grid scale pinning points and troughs[48]). Thus, is it not a surprise that the SLC varies so considerably with bed realisation at coarse resolution. For example, it is the timing of the unpinning from relatively high and poorly resolved bedrock shown in Fig. 2 that causes a large spread of SLC for different 8 km bed realisations. Our findings explicitly demonstrate the need to account for the large effects of bed sampling on future projections at coarse resolution when the bed topography is complex, such as the Amundsen Sea Sector. In other regions, such as the interior of Antarctica, if the bed topography is less complex we might expect resolution to have a smaller effect.
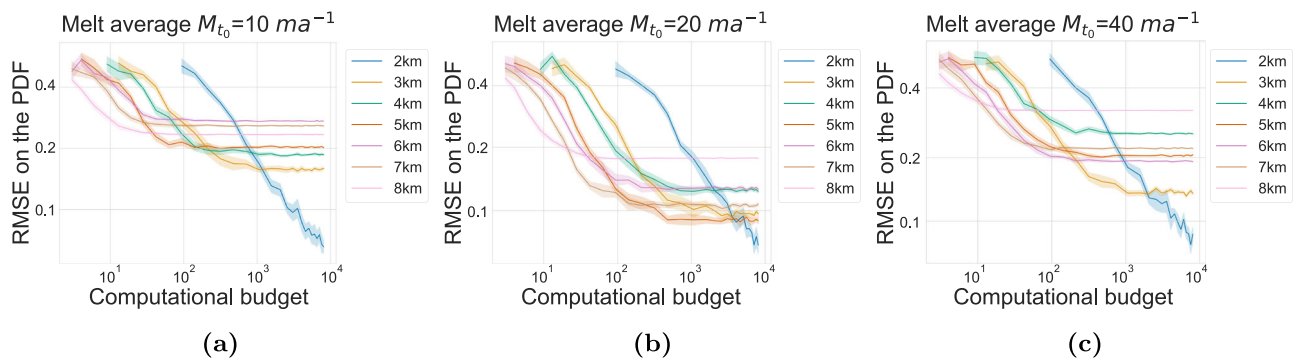
Fig. 7 | The dimensionless RMSE error of the PDF as a function of the computational budget T. This is derived from comparing the predicted PDF at $t = 150$ years under budget T at resolution R to the the full calculated PDF at a $R = 2$ km with all points included (i.e. no computational constraint). Results are shown for a prior distribution mean of (**a**) $M_{t_0} = 10$ m a^{-1}, (**b**) $M_{t_0} = 20$ m a^{-1}, and (**c**) $M_{t_0} = 40$ m a^{-1}. The plots are obtained by calculating the mean error over 500 trials and the shading represents the 95% interval. Colours indicate resolution: 2 km (blue), 3 km (orange), 4 km (green), 5 km (red), 6 km (purple), 7 km (brown), 8 km (pink).

In this study, we produce different Bedmachine realisations by subsampling grid points from the Bedmachine data set, essentially sampling along the diagonal of an Rkm square of possible bed options (with additional samplings off-diagonal for 2 km and 3 km resolutions to ensure a minimum of 8 realisations at each resolution). By choosing this sampling, we are assuming the bed noise is isotropic and uncorrelated, and we construct a non-random cover of the potential space. In reality the Bedmachine data also has an error, which is made up of measurement error (accuracy), sampling length of the input data to Bedmachine, and interpolation error that is introduced when the data is transferred onto the Bedmachine grid. To fully propagate all these sources of errors alongside our sampling strategy errors into the PDF would require multiple realisations of the bed topography product to be constructed, taking in to account the error covariance structure, and to then sample from these as we have done here. Our approach is appropriate if the input data is accurate and well sampled, so that the differences between possible Bedmachine realisations due to data errors is small. However, if the measurement errors are sufficiently large and were also sampled, it is feasible that the difference between resolutions would decrease. Similarly, using a bilinear interpolation scheme may have a small effect on our results, but it could not remove all of the variance we see: the bed would remain unresolved due to features such as troughs and ridges that are smoothed out during interpolation. While there may be sampling strategies that can minimise the misrepresentation of the coarse resolution models, in reality it would be difficult to sufficiently capture the features of a fine resolution bed. This is primarily because the effects of both water depth and buttressing need to be accounted for. Even if the effects of water depth could be captured by, for example, integrating the water depth over a coarse resolution cell based on the high resolution data, the effects of buttressing would be harder to incorporate. Essentially, one would have to use the fine resolution bed to learn a parameterisation for pinning points. This is an active area of research, and is beyond the scope of our study.

In a recent study on the same domain with the same initial state, issues with the bathymetry near the grounding line of Pine Island Glacier were found when compared with data[28]. As a result, the bed was 'dug' in order to better match the fine resolution data and to stop unrealistic advance of the PIG grounding line in a coupled ice-ocean model simulation. Here, we do not alter the bed data set provided from Bedmachinev3, and as a result we see less mass loss from PIG than in Bett et al.[28]. Wernecke et al.[66] also quantify the effect of bed topography uncertainties on SLC projections, in their case under PIG, by using different bed data sets alongside statistically modelled beds that take into account bed topography uncertainties. They conclude that the bedrock uncertainty creates a 5–25% uncertainty in the predicted sea level rise contribution at the year 2100. Spectral analysis of bed roughness from this region shows variance increasing linearly with length of profile[67]. Sampling every 8 km leaves exposure to interpolation errors with four times the variance relative to 2 km. Combined with the strong

dependence of ice flux upon water depth at the grounding line[6], this may explain the large scatter on model results from 8 km resolution. Misrepresentation of grounding line physics at low resolution will also contribute[39]. More broadly, these findings relating to the sensitivity of simulations to the bed topography indicate that small scale features in the bed exert a large influence on simulated future ice loss from this region, as in[48]. Thus, accurate, fine resolution data from underneath ice shelves and grounded ice is essential for making realistic projections of sea level rise over the coming centuries.

Here, we have focused on a simple basal melt parameterisation. In order to use this method for making realistic probabilistic projections, a more physically realistic parameterisation or, ideally, a coupled ice-ocean model is required[28,57,68]. The upper end of the range of basal melt rates we used is extreme, leading to cases with almost total removal of the shelves and essentially an unbuttressed ice sheet, somewhat like ABUMIP, an ice-sheet model inter-comparison that explored these high end scenarios[35]. The simulations we have present are a limited set to explore the methodology and the specific effects of resolution on the upper tails, and do not span the range of uncertainties needed to make realistic projections. In particular, we have neglected uncertainty arising from parameters involved in the model initialisation and in the basal sliding and ice viscosity, and we have not included ice fracture and calving front motion in this model. Although the importance of the marine ice cliff instability (MICI) is still being assessed, it has the potential to have large affects on SLC projections[29]. Future models will need to incorporate this process and assess the affect on SLC projections, adding further to the uncertainties from model parameters.

Despite the caution we advise here, the pressing need to project Antarctica's contribution to sea-level change remains. As models evolve to capture more of the missing processes, and coupled ice-ocean-atmosphere models emerge, the computational cost of running many fine resolution simulations over a range of uncertain parameters and future forcings will also increase. Improvements in computing power will go some way to alleviating this pressure, but producing probability distributions to fully characterise the probability of a particular outcome will remain challenging. Strategic experimental design, as suggested by Rohmer et al.[46] and Edwards et al.[20], and machine learning methodologies have the potential to aid with construction of computationally cheaper probability distributions of sea level. As a first step towards this, similar to Berdahl et al.[37], we used Gaussian processes to emulate the SLC as a function of average melt rate, in order to enable the production of many more samples than are possible using the WAVI model alone and with reduced computational cost. This increases our ability to characterize probabilistic distributions of sea level change from Antarctica, with sufficient sampling in the upper tails to provide information on the high-impact low-likelihood scenarios of societal relevance.

We evaluated strategies for producing PDFs with limited computational budgets, comparing accuracy with our highest resolution simulations.

For low budgets, there are examples where coarser resolution simulations may be preferred, but as budget increases, fine resolution simulations would always be selected. In the future, this approach of devising run strategies using information on cost and accuracy could be a useful tool to optimise large ensemble studies with the aim of making probabilistic SLC projections. Although here we have performed a post-simulation calculation, our findings point the way towards using a multi-fidelity approach in the future[49]. Multi-fidelity experimental design (MFED) serves as a data-driven methodology for automating decisions related to computation, aiming to achieve the most accurate estimates within predefined computational budgets. Practically, given a computational budget, MFED algorithms aim to select the next best point (resolution and parameter) to reduce the uncertainty in the objective given by the user. As the complexity of ice-sheet simulations grows due to the incorporation of additional parameters and processes, the computational costs required for obtaining a probabilistic estimate will increase dramatically. Data-driven computational decision making removes the burden on the researcher to estimate the computational parameters of their simulations. For example, the optimal sampling strategy can vary depending on user requirements; for instance, studies focusing on extreme scenarios may necessitate more frequent use of higher-resolution simulations.

The methodology presented herein is not ice-sheet model specific, and in the future could be deployed for any ice sheet model output and over a larger geographical region such as the whole of AIS. It also has the capacity to include other uncertain model parameters in GP emulation used for producing sea level projections. Additionally, although this study focuses on the issues of model resolution, emulation, and producing SLC projections under limited budget for ice sheet modelling, the techniques and discussion herein will be relevant to other areas in which high resolution models are both required and computationally expensive, such as in climate modelling[69]. In those areas, it is also a huge challenge to run models to convergence with spatial resolution, and strategies will need to be developed in order to make robust model predictions. The findings presented here indicate that it is essential to address the issues surrounding model resolution and computational cost in order to satisfy the pressing need to provide realistic, usable projections of sea level to decision and policy makers.

Methods
The WAVI ice sheet model
Here, we use the Wavelet-based Adaptive-grid Vertically-integrated Ice-sheet model (WAVI). WAVI is based on[53,54], but is rewritten in Julia and can be found at https://rjarthern.github.io/WAVI.jl/. WAVI includes the grounded ice and floating shelves and incorporates stresses to capture the dynamics linking the two regions over the grounding line. The model includes membrane stresses and is vertically integrated but retains the vertical profile of velocity implicitly. WAVI is a finite difference model and the equations are solved numerically over a rectangular uniform mesh. In this study, the same resolution R is used in both horizontal directions. A subgrid parameterisation is used to represent the grounding line movement on finer scale than the grid resolution alone[43,45,70]; full details can be found in Arthern & Williams[18].

Bed topography sampling
The bed topography is provided from the Bedmachine version 3 data set[58,59], which has resolution of 500 metres. To generate grids of bed, ice thickness and surface elevation at a coarser resolutions, we sample this data by selecting points R kilometres apart. At each resolution there are $(2R)^2$ possible choices of bed topography and corresponding thickness and surface elevation. Running simulations for all of these different realisations would be hugely computationally expensive. Instead, we select n different realisations for each R: due to the increasing computation cost with resolution, we select $n = 8$ for $R = 2, 3, 4$ km, $n = 10$ for $R = 5$ km, $n = 12$ for $R = 6$ km, $n = 14$ for $R = 7$ km and $n = 16$ for $R = 8$ km. As previously stated, we do this by choosing points along the diagonal of the square of possibilities (e.g. starting at the grid corner and moving one along in both x and y directions). This

method allows us to ensure we are covering each horizontal and vertical grid column/row once, rather than using a random sample in which not every dimension may be covered. For 2 km and 3 km, 4 and 2 extra realisations were also selected to ensure $n = 8$. In the absence of other information, these were selected along the other diagonal.

Model initialisation and relaxation
For each Bedmachine realisation, a separate initialisation is performed[54]. In brief, two-dimensional spatially-varying fields of basal drag and ice stiffness coefficients are calculated in a data assimilation framework by matching modelled surface velocities with surface velocities [Measures 2014/15[71]], thinning rates[72] and accumulation rates[73]. Englacial temperatures are from Pattyn[74]. The ice geometry is kept fixed during the initialisation. An example of an initial state is shown in Supplementary Fig. 4.

A surface relaxation is then performed for 4000 years for each initial state, during which the accumulation a_{relax} is set as $a_{data} − dh/dt_{data}$, in order to bring the flux divergence into better agreement with observations of thinning and accumulation, usually at a slight loss in agreement between the modelled and observed surface speeds. During the relaxation the grounding line is fixed in place by not allowing thickness change over the shelves or the grounding line area. Ice thickness inland can vary. The basal drag and ice stiffness coefficients are held constant during the relaxation. This leads to an initialised, relaxed state that has the prescribed 2015 grounding line and is thinning at the observed rate i.e. it is initialised into a retreating state, so at the start of the forward simulations the ice sheet is losing mass (it is not at steady state). The use of a separate initialisation and relaxation for each Bedmachine realisation ensures that the same process is followed for each sampling at each resolution and avoids the introduction of interpolation errors. This leads to a total of 78 different initial states. Examples of the match between the initial states both before and after relaxation can be found in Supplementary Note 1 (Supplementary Figs. 1–3) alongside summary statistics for the match between thickness, surface speed and rates of thinning at each resolution (Supplementary Table 1).

Melt rate parameterisation and calibration
We use the depth-dependent melt parameterisation[56,57]. Following the nonmenclature of Favier et al.[57], we use the *quadratic, local dependency* form of the parameterisation, which is widely used[55]. Other parameterisations, such as the non-local form, could alternatively have been used. Here, melt rate M_{t_0} is given by:

$$m = \gamma_T \left(\frac{\rho_{sw} c_{po}}{\rho_i L_i} \right)^2 (T_a - T_f)^2, \tag{1}$$

γ_T is the heat exchange velocity that is usually calibrated [see[57], discussed further below], $\rho_{sw} = 1028.0$ kg m^{-3} and $\rho_i = 918$ kg m^{-3} are the respective densities of ocean water and ice, $c_{po} = 3.974e \times 10^3$ J kg^{-1} K^{-1} is the specific heat capacity of the ocean mixed layer, and $L_i = 3.35 \times 10^5$ J kg^{-1} is the latent heat of fusion of ice. T_f is the melting-freezing point at the interface between the ocean and the ice-shelf basal surface and is defined as

$$T_f = \lambda_1 S_a + \lambda_2 + \lambda_3 z_b, \tag{2}$$

where $\lambda_1 = -5.73 \times 10^{-3}$ °C PSU^{-1}, and $\lambda_2 = 8.32 \times 10^{-4}$ °C and $\lambda_3 = 7.61 \times 10^{-4}$ °C m^{-1} are the liquidus slope, intercept and pressure coefficient respectively, z_b is the ice base elevation. $S_a(z)$ and $T_a(z)$ are the practical salinity and ambient temperature taken from the far field ocean boundary. Here, we take a layered structure for T_a (°C) and S_a (PSU), defining them as piece-wise linear functions of depth such that:

$$T_a(z) = \begin{cases} 1.2 & z < -700m \\ 1.2 - 2.2\frac{z+700}{400} & -700m \leq z \leq -300m \\ -1 & z > -300m \end{cases} \tag{3}$$

$$S_a(z) = \begin{cases} 34.6 & z < -700m \\ 34.6 - 0.6\frac{z+700}{400} & -700m \leq z \leq -300m \\ 34.0 & z > -300m \end{cases} \quad (4)$$

These functions represent profiles with a linear transition between an upper layer of Winter Water and a lower layer of Circumpolar Deep Water (CDW). The values we have used here are roughly consistent with conditions in the Amundsen Sea[75], and were also used in an idealised study[76]. Since we are interested in the effects of resolution and the upper tails of distribution, this provides a sensible choice to demonstrate our methodology. For more realistic projections covering the range of thermocline profiles observed over time[77], these temperature and salinity profiles could be varied directly rather than solely varying the γ_T parameter.

In our case, we run many different Bedmachine realisations over 7 different resolutions. This means that if we want the melt rate M_{t_0} to represent the average melt rate under all fully floating points in our domain at $t = 0$ years, we need to calibrate γ_T separately for each Bedmachine realisation (values can be found in the associated published data set[78]). Once $\gamma_{Tcalibrated}$ is calculated for each resolution as the value required at each spatial resolution to give an average melt rate of $M_{t_0} = 1$ m/a at $t = 0$ years, γ_T in equation (1) is then set by simply scaling $\gamma_{Tcalibrated}$ with the required average melt rate M_{t_0}:

$$\gamma_T = \gamma_{Tcalibrated} \times M_{t_0}. \quad (5)$$

This value of γ_T is then not varied within each simulation. As the grounding line retreats new cavities open underneath the ice shelves, the depth-dependant parameterisation provides the melt rate in these new cavities. Melt rates elsewhere can also vary as the ice shelf thins. Thus, the average melt rate under all shelves varies within a simulation due to differences in evolving geometry.

Previous studies have shown that retreat of the grounding line and, thus, SLC is very sensitive to whether melt is applied to partially grounded cells[18]. While Seroussi & Morlighem[26] recommend not applying melt in these cells, Leguy et al.[79] find otherwise (applying some melt is often beneficial in CISM), and conclude that there is likely not a single best approach for all models. Thus, this model choice should be an informed decision based on testing each model separately. Based on previous sensitivity studies to this choice for WAVI[18], we apply no melt in partially grounded cells in all simulations.

Forward model simulations

WAVI simulations were run forward in time from the relaxed state for 150 years. We use a persistence-based approach in which we prescribe the present day accumulation and keep the calving front fixed. Although unrealistic, as in reality the calving front is likely to change over this time-scale, it is a useful conservative assumption that allows us to investigate the effects of melt rate and resolution on the mean and upper tail of the SLC distributions. No damage evolution model was used, so initial damage from the initialisation remains for the full forward simulation.

The average melt rate at $t = 0$, M_{t_0}, is prescribed using the calibration coefficient γ_T, as described above. This coefficient is then kept the same throughout each full forward simulation. Since the melt rate varies with depth, it is non-constant over the shelves and varies over time as the geometry of the ice sheet evolves: as new cavities open up, the parameterisation applies melt dependent on the bed topography, and as such the average melt rate over the fully floating points varies when $t > 0$ years. It is only the average melt at $t = 0$ years, denoted M_{t_0}, that is prescribing by setting the calibration coefficient γ_T. The range of M_{t_0} used is justified in the 'Results' section.

The same velocity time step was used for 3–8 km, $dt_v = 0.1$ years, but a smaller time step of $dt_v = 0.05$ was required for 2 km resolution. A smaller subtime step dt_h was also used to evolve the ice thickness: at this time step, the thickness and divergence were updated but not the velocity. The melt rates are calculated from the thickness field at each thickness time step dt_h. Since the velocity update is approximately 99% of the cost of a forward time step, this is an efficient way to ensure better resolutions of ice thicknesses without a large computational cost (see Supplementary Methods and Supplementary Fig. 5 for more details). This time step was set as $dt_h < = 0.025$ for all simulations.

Gaussian process emulation

Here, we used Gaussian process emulation to compute the probability density function (PDF) and cumulative distribution function (CDF) describing the SLC based on a prior distribution over melt average. To obtain these quantities, we first fit the Gaussian process to the the SLC predicted by WAVI. We then sampled the melt average from the prior and finally calculated the PDF and CDF using the GP surrogate model. In the sections below, we provide a detailed explanation of all the parameters and methods used for this computation.

Gaussian Processes

A Gaussian Process (GP) is a collection of random variables for which any finite subset has a joint Gaussian distribution. A GP is fully defined by its mean function, often taken as zero, and a covariance function or kernel, which measures the dependence between different points in the function.

GPs are nonparametric models, meaning they do not make strong assumptions about the functional form of the data, offering flexibility. Additionally, as a probabilistic model, a GP not only gives a prediction for an unseen point but also provides a measure of certainty or confidence in that prediction. This feature is crucial in understanding the risk and variability associated with different future scenarios.

Mathematically, a Gaussian Process is defined as:

$$f(x) \sim GP(m(x), k(x, x')) \quad (6)$$

where $f(x)$ is the function we want to estimate, $m(x)$ is the mean function, which we assume to be zero for simplicity as is common practise, and $k(x, x')$ is the covariance function or kernel.

The choice of the covariance function is crucial for GPs, and one of the most commonly used kernels is the squared exponential kernel, also known as the Gaussian or Radial Basis Function (RBF) kernel. The squared exponential kernel is a real-valued function of two variables and is used to measure the similarity between these variables. The squared exponential kernel is defined as:

$$k(x, x') = \sigma^2 \exp\left(-\frac{1}{2l^2}(x - x')^2\right) \quad (7)$$

where x and x' are two points in the input space, σ^2 is the variance that controls the vertical variation, l is the length-scale parameter that controls the horizontal variation or how quickly the function values can change. The kernel function can be interpreted as a measure of similarity: points that are closer in the input space have a higher kernel value and are therefore more similar.

Maximum likelihood estimation

We estimate the parameters of the kernel using Maximum Likelihood Estimation (MLE). The aim is to find the parameter values that maximize the likelihood function, given the SLC predicted by WAVI. The likelihood function measures how probable the obtained SLC are, given a specific parameter value. Practically, finding the maximum likelihood estimate often involves numerical optimization methods as the solution may not be easy to compute analytically. We use L-BFGS optimization algorithm to find the optimal value. All the algorithms used in this paper were developed using GPy[80]. We standardize the inputs (i.e. remove the mean and scale by the variance) to improve the convergence of the MLE estimates.
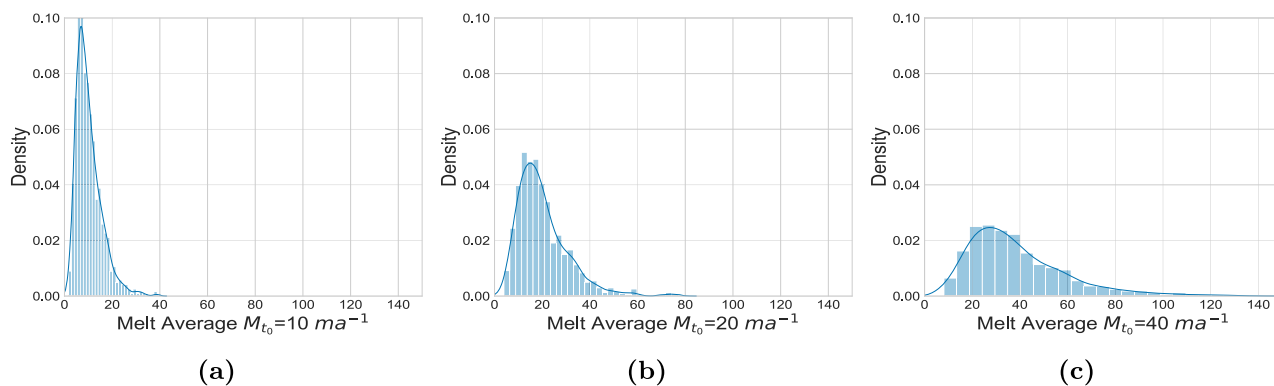
Fig. 8 | Prior distribution of melt averages. These distributions are used to calculate the PDF and CDF of SLC. All are log-normal prior with variance $\tau = 0.5$ and (**a**) mean $\mu = 2.2$ (corresponding to a mean of $M_{t_0} = 10$ m a^{-1}), (**b**) $\mu = 2.87$ ($M_{t_0} = 20$ m a^{-1}), and (**c**) $\mu = 3.52$ ($M_{t_0} = 40$ m a^{-1}). Histograms are shown by the blue bars and the log-normal probability distributions with blue lines.

Emulator validation

To assess the emulator's effectiveness, we validate the model's accuracy using a leave-one-out strategy[20]. With this approach, for every specific melt rate, we leave the SLC predicted by WAVI for the chosen melting rate aside and train the GP model on the remaining SLC predictions. Once the model is trained, we use it to predict the SLC for the excluded melting rate. We then calculate the Root Mean Squared Error (RMSE) by comparing this prediction with the empirical SLC obtained from WAVI. This process helps validate the effectiveness of our model across all ranges of melt rate.

As previously mentioned, we use a squared exponential kernel due to the smoothness of the SLC outputs from WAVI. We optimize the length scale of the kernel by maximizing the evidence of the WAVI SLCs. The RMSE for the GP model without hyper-parameter tuning (default values provided by the library GPy) is 3.12 mm whereas it decreases to 1.95mm when tuning the length scale to an appropriate value. We normalize our experimental data, allowing the library's default hyper-parameters to serve as reasonable initial values for our runs. These experiments validate the effectiveness of the surrogate model as well as the parameter tuning strategy.

To verify the accuracy of the uncertainty bounds of the surrogate model, we calculate the percentage of the SLCs predicted by WAVI that fall within these bounds (defined as the 95% credible interval). Over 97% of the SLC points lie within the set uncertainty bounds of our surrogate model.

Sea level contributions from emulation

The first step to obtaining the PDF over SLC is to sample a dataset of $N = 1000$ melt average points from a log-normal prior displayed in Fig. 8.

Using the GP model from the previous section, we predict for each melt average value the corresponding normal distribution (mean and variance) of SLC. For each normal distribution, we sample $K = 1000$ points which yields a dataset of 1000×1000 points.

To compute the PDF, we use a Kernel Density Estimator on the collected dataset. A Kernel Density Estimator (KDE) is a non-parametric way to estimate the probability density function of a random variable[62], providing a continuous and differentiable estimate of the distribution for visualizations and subsequent analysis. To compute the CDF, we discretize the SLC axis and compute the number of elements from the dataset that fall above the given threshold.

Cost-constrained estimation

In the section Sea level contribution estimation with limited budget, we demonstrate the errors incurred when estimating the PDFs with limited WAVI simulations. To emulate limited WAVI simulations, we calculate the costs of running each simulation based on the runtime required by the resolution (see Supplementary Table 2). We calculate the number of points for a specific resolution by using the formula $\frac{T}{C}$ for costs C within a given

budget T. Supplementary Table 3 displays the corresponding number of points for each resolution.

Through the application of WAVI simulations, we generate a dataset composed of the points m_i, r_i, y_i, with m_i denoting the average melt, r_i representing the resolution, and y_i indicating the resulting SLC.

To extract N points with a budget of T, we initially sample N average melts from the log-normal prior, and identify the closest point derived from the simulation. This is achieved by computing the Euclidean distance between the sampled average melt and the datasets of m_i. These N points are subsequently employed in the computation of the PDF. The error between the computed PDF and the gold standard (a 2 km resolution with no computational limitations) is assessed using the mean squared error metric.

Data availability

The following data sets were used to initialise the ice sheet model: Bedmachine (version 3) (https://doi.org/10.5067/FPSU0V1MWUB6[59]), MEaSUREs (https://doi.org/10.5067/9T4EPQXTJYW9[71]), thinning rates from https://digital.lib.washington.edu/researchworks/handle/1773/45388(accession no. 45388[72]), accumulation[73], and englacial temperatures[74] (accumulation rates and temperatures are also stored in the model output data set[78]). A data set of the model outputs from the WAVI simulations is stored in the UK Polar Data Centre (https://doi.org/10.5285/4de39bc0-fc2b-4232-ac39-3cc1fd723f64[78]). Ensembles of sea level contributions from the WAVI simulation outputs used for the Gaussian process emulation and the production of probabilisitc sea level curves can be found alongside the model codes at https://github.com/pierthodo/multi_resolution_ice_sheet and in Zenodo (https://doi.org/10.5281/zenodo.14422770[81]).

Code availability

The open source WAVI ice sheet model can be found at https://github.com/RJArthern/WAVI.jl. The version used for these simulations is also stored in Zenodo (branch: subtimestepping, commit: 669dc49, https://doi.org/10.5281/zenodo.14393521[82]). The analysis code for the Gaussian process emulation and calculation of computational budgets can be found at https://github.com/pierthodo/multi_resolution_ice_sheet and stored in Zenodo (https://doi.org/10.5281/zenodo.14422770[81]). The model ensembler tool used to efficiently run the WAVI model ensembles is available at https://github.com/jimCircadian/model-ensembler.

References

1. Fretwell, P. et al. Bedmap2: improved ice bed, surface and thickness datasets for Antarctica. *Cryosphere* **7**, 375–393 (2013).

Article

2. IMBIE. Mass balance of the antarctic ice sheet from 1992 to 2017. *Nature* **558**, 219–222 (2018).

3. Otosaka, I. N. et al. Mass balance of the greenland and antarctic ice sheets from 1992 to 2020. *Earth Syst. Sci. Data* **15**, 1597–1616 (2023).

4. Shepherd, A. et al. Trends in Antarctic ice sheet elevation and mass. *Geophys. Res. Lett.* **46**, 8174- 8183 (2019).

5. Weertman, J. Stability of the junction of an ice sheet and an ice shelf. *J. Glaciol.* **13**, 3–11 (1974).

6. Schoof, C. Marine ice-sheet dynamics. part 1. the case of rapid sliding. *J. Fluid Mech.* **573**, 27–55 (2007).

7. Joughin, I., Smith, B. E. & Medley, B. Marine ice sheet collapse potentially under way for the Thwaites Glacier basin, West Antarctica. *Science* **344**, 735–738 (2014).

8. Favier, L. et al. Retreat of Pine Island Glacier controlled by marine ice-sheet instability. *Nat. Clim. Change* **4**, 117–121 (2014).

9. Rosier, S. H. R. et al. The tipping points and early warning indicators for pine island glacier, west antarctica. *Cryosphere* **15**, 1501–1516 (2021).

10. Rott, H., Rack, W., Skvarca, P. & de Angelis, H. Northern Larsen Ice Shelf, Antarctica: further retreat after collapse. *Ann. Glaciol.* **34**, 277–282 (2002).

11. Scambos, T. A., Bohlander, J. A., Shuman, C. A. & Skvarca, P. Glacier acceleration and thinning after ice shelf collapse in the Larsen B embayment, Antarctica. *Geophys. Res. Lett.* **311**, L18402 (2004).

12. Pritchard, H. D., Arthern, R. J., Vaughan, D. G. & Edwards, L. A. Extensive dynamic thinning on the margins of the Greenland and Antarctic ice sheets. *Nature* **461**, 971–975 (2009).

13. Reese, R., Gudmundsson, H., Levermann, A. & Winkelmann, R. The far reach of ice-shelf thinning in antarctica. *Nat. Clim. Change* **8**, 53–57 (2018).

14. Depoorter, M. A. M. et al. Calving fluxes and basal melt rates of Antarctic ice shelves. *Nature* **502**, 89–92 (2013).

15. Rignot, E., Jacobs, S., Mouginot, J. & Scheuchl, B. Ice-shelf melting around antarctica. *Science* **341**, 266–270 (2013).

16. Lipscomb, W. H. et al. Ismip6-based projections of ocean-forced antarctic ice sheet evolution using the community ice sheet model. *Cryosphere* **15**, 633–661 (2021).

17. Nias, I. J., Cornford, S. L. & Payne, A. J. Contrasting the modelled sensitivity of the Amundsen sea embayment ice streams. *J. Glaciol.* **62**, 552–562 (2016).

18. Arthern, R. J. & Williams, C. R. The sensitivity of West Antarctica to the submarine melting feedback. *Geophys. Res. Lett.* **44**, 2352–2359 (2017).

19. Core Writing Team, H. L. & (eds.), J. R. *IPCC, 2023: Climate Change 2023: Synthesis Report. A Report of the Intergovernmental Panel on Climate Change. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (IPCC, Geneva, Switzerland, 2023).

20. Edwards, T. et al. Projected land ice contributions to twenty-first-century sea level rise. *Nature* **593**, 74–82 (2021).

21. Pattyn, F., Favier, L., Sun, S. & Durand, G. Progress in numerical modeling of Antarctic ice-sheet dynamics. *Curr. Clim. Change Rep.* **3**, 174–184 (2017).

22. Pattyn, F. & Morlighem, M. The uncertain future of the antarctic ice sheet. *Science* **367**, 1331–1335 (2020).

23. Forster, P. M., Maycock, A. C., McKenna, C. M. & Smith, C. J. Latest climate models confirm need for urgent mitigation. *Nat. Clim. Change* **10**, 7–10 (2019).

24. Seroussi, H. et al. initmip-antarctica: an ice sheet model initialization experiment of ismip6. *Cryosphere* **13**, 1441–1471 (2019).

25. van de Wal, R. S. W. et al. A high-end estimate of sea level rise for practitioners. *Earth's. Future* **10**, e2022EF002751 (2022).

26. Seroussi, H. & Morlighem, M. Representation of basal melting at the grounding line in ice flow models. *Cryosphere* **12**, 3085–3096 (2018).

27. Siahaan, A. et al. The antarctic contribution to 21st-century sea-level rise predicted by the uk earth system model with an interactive ice sheet. *Cryosphere* **16**, 4053–4086 (2022).

28. Bett, D. T. et al. Coupled ice–ocean interactions during future retreat of west antarctic ice streams in the amundsen sea sector. *Cryosphere* **18**, 2653–2675 (2024).

29. DeConto, R. M. & Pollard, D. Contribution of Antarctica to past and future sea-level rise. *Nature* **531**, 591–597 (2016).

30. Lhermitte, S. et al. Damage accelerates ice shelf instability and mass loss in amundsen sea embayment. *Proc. Natl. Acad. Sci.* **117**, 24735–24741 (2020).

31. Ritz, C. et al. Potential sea-level rise from Antarctic ice-sheet instability constrained by observations. *Nature* **528**, 115–118 (2015).

32. Golledge, N. et al. The multi-millennial antarctic commitment to future sea-level rise. *Nature* **526**, 421–425 (2015).

33. DeConto, R. et al. The paris climate agreement and future sea-level rise from antarctica. *Nature* **593**, 83–89 (2021).

34. Levermann, A. et al. Projecting antarctica's contribution to future sea level rise from basal ice shelf melt using linear response functions of 16 ice sheet models (larmip-2). *Earth Syst. Dyn.* **11**, 35–76 (2020).

35. Sun, S. et al. Antarctic ice sheet response to sudden and sustained ice-shelf collapse (abumip). *J. Glaciol.* **66**, 891–904 (2020).

36. Seroussi, H. et al. Ismip6 antarctica: a multi-model ensemble of the antarctic ice sheet evolution over the 21st century. *Cryosphere* **14**, 3033–3070 (2020).

37. Berdahl, M., Leguy, G., Lipscomb, W. H. & Urban, N. M. Statistical emulation of a perturbed basal melt ensemble of an ice sheet model to better quantify antarctic sea level rise uncertainties. *Cryosphere* **15**, 2683–2699 (2021).

38. Pattyn, F. et al. Results of the marine ice sheet model intercomparison project, mismip. *Cryosphere* **6**, 573–588 (2012).

39. Pattyn, F. et al. Grounding-line migration in plan-view marine ice-sheet models: Results of the ice2sea MISMIP3d intercomparison. *J. Glaciol.* **59**, 410 (2013).

40. Cornford, S. L. et al. Results of the third marine ice sheet model intercomparison project (mismip+). *Cryosphere* **14**, 2283–2301 (2020).

41. Gudmundsson, G. H., Krug, J., Durand, G., Favier, L. & Gagliardini, O. The stability of grounding lines on retrograde slopes. *Cryosphere* **6**, 1497–1505 (2012).

42. Cornford, S. L. et al. Adaptive mesh, finite volume modeling of marine ice sheets. *J. Comput. Phys.* **232**, 529–549 (2013).

43. Gladstone, R. M., Payne, A. J. & Cornford, S. L. Resolution requirements for grounding-line modelling: sensitivity to basal drag and ice-shelf buttressing. *Ann. Glaciol.* **53**, 97–105 (2012).

44. Leguy, G. R., Asay-Davis, X. S. & Lipscomb, W. H. Parameterization of basal friction near grounding lines in a one-dimensional ice sheet model. *Cryosphere* **8**, 1239–1259 (2014).

45. Seroussi, H., Morlighem, M., Larour, E., Rignot, E. & Khazendar, A. Hydrostatic grounding line parameterization in ice sheet models. *Cryosphere* **8**, 2075–2087 (2014).

46. Rohmer, J., Thieblemont, R., Le Cozannet, G., Goelzer, H. & Durand, G. Improving interpretation of sea-level projections through a machine-learning-based local explanation approach. *Cryosphere* **16**, 4637–4657 (2022).

47. Cuzzone, J. K. et al. The impact of model resolution on the simulated holocene retreat of the southwestern greenland ice sheet using the ice sheet system model (issm). *Cryosphere* **13**, 879–893 (2019).

48. Rückamp, M., Goelzer, H. & Humbert, A. Sensitivity of greenland ice sheet projections to spatial resolution in higher-order simulations: the alfred wegener institute (awi) contribution to ismip6 greenland using the ice-sheet and sea-level system model (issm). *Cryosphere* **14**, 3309–3327 (2020).

49. Thodoroff, P. et al. Multi-fidelity experimental design for ice-sheet simulation. *arXiv preprint arXiv:2307.08449* (2023).

50. Hinkel, J. et al. Sea-level rise scenarios and coastal risk management. *Nat. Clim. Change* **5**, 188–190 (2015).

51. Stammer, D. et al. Framework for high-end estimates of sea level rise for stakeholder applications. *Earth's. Future* **7**, 923–938 (2019).

52. Pritchard, H. D. et al. Antarctic ice-sheet loss driven by basal melting of ice shelves. *Nature* **484**, 502–505 (2012).

53. Goldberg, D. A variationally derived, depth-integrated approximation to a higher-order glaciological flow model. *J. Glaciol.* **57**, 157–170 (2011).

54. Arthern, R. J., Hindmarsh, R. C. A. & Williams, C. R. Flow speed within the Antarctic ice sheet and its controls inferred from satellite observations. *J. Geophys. Res.: Earth Surf.* **120**, 1171–1188 (2015).

55. Asay-Davis, X., Jourdain, N. & Nakayama, Y. Developments in simulating and parameterizing interactions between the southern ocean and the antarctic ice sheet. *Curr. Clim. Change Rep.* **3**, 316–329 (2017).

56. Holland, P., Jenkins, A. & Holland, D. The response of ice shelf basal melting to variations in ocean temperature. *J. Clim.* **21**, 2558–2572 (2008). DMH acknowledges support from the United States National Science Foundation Grants OPP-0327664, OPP-0337073, and OCE-0350912.

57. Favier, L. et al. Assessment of sub-shelf melting parameterisations using the ocean–ice-sheet coupled model nemo (v3. 6)–elmer/ice (v8. 3). *Geosci. Model Dev.* **12**, 2255–2283 (2019).

58. Morlighem, M. et al. Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the antarctic ice sheet. *Nat. Geosci.* **13**, 1–6 (2020).

59. Morlighem, M. Measures bedmachine antarctica, version 3 (2022).

60. Milillo, P. et al. Heterogeneous retreat and ice melt of thwaites glacier, west antarctica. *Sci. Adv.* **5**, eaau3433 (2019).

61. Chen, Y.-C. A tutorial on kernel density estimation and recent advances. *Biostat. Epidemiol.* **1**, 161–187 (2017).

62. Terrell, G. R. & Scott, D. W. Variable kernel density estimation. *The Annals Statistics* 1236–1265 (1992).

63. Dietz, S. & Koninx, F. Economic impacts of melting of the antarctic ice sheet. *Nat. Commun.* **13**, 5819 (2022).

64. Thomas, R. H. The dynamics of marine ice sheets. *J. Glaciol.* **24**, 167–177 (1979).

65. Berger, S., Favier, L., Drews, R., Derwael, J.-J. & Pattyn, F. The control of an uncharted pinning point on the flow of an antarctic ice shelf. *J. Glaciol.* **62**, 37–45 (2016).

66. Wernecke, A., Edwards, T. L., Holden, P. B., Edwards, N. R. & Cornford, S. L. Quantifying the impact of bedrock topography uncertainty in pine island glacier projections for this century. *Geophys. Res. Lett.* **49**, e2021GL096589 (2022).

67. Hogan, K. A. et al. Revealing the former bed of thwaites glacier using sea-floor bathymetry: implications for warm-water routing and bed controls on ice flow and buttressing. *Cryosphere* **14**, 2883–2908 (2020).

68. Seroussi, H. et al. Continued retreat of thwaites glacier, west antarctica, controlled by bed topography and ocean circulation. *Geophys. Res. Lett.* **44**, 6191–6199 (2017).

69. Schneider, T. et al. Harnessing ai and computing to advance climate modelling and prediction. *Nat. Clim. Change* **13**, 887–889 (2023).

70. Pattyn, F., Huyghe, A., De Brabander, S. & De Smedt, B. Role of transition zones in marine ice sheet dynamics. *J. Geophys. Res.: Earth Surf.* **111**, F02004 (2006).

71. Mouginot, J., Scheuchl, B. & Rignot., E. Measures annual antarctic ice velocity maps 2005-2017, version 1. [2014/5] (2017).

72. Smith, B. et al. Pervasive ice sheet mass loss reflects competing ocean and atmosphere processes. *Science* **368**, 1239–1242 (2020). Contains maps of thinning rates in Antarctica, demonstrating the significant thinning of WAIS over the satellite period.

73. Arthern, R. J., P. Winebrenner, D. & Vaughan, D. Antarctic snow accumulation mapped using polarization of 4.3-cm wavelength microwave emission. *J. Geophys. Res.* **111**, D06108 (2006).

74. Pattyn, F. Antarctic subglacial conditions inferred from a hybrid ice sheet/ice stream model. *Earth Planet. Sci. Lett.* **295**, 451–461 (2010).

75. Dutrieux, P. et al. Strong sensitivity of pine island ice-shelf melting to climatic variability. *Science* **343**, 174–178 (2014).

76. Bradley, A. et al. A framework for estimating the anthropogenic part of antarctica's sea level contribution in a synthetic setting. *Commun. Earth Environ.* **5**, 121 (2024).

77. Jenkins, A. et al. West Antarctic ice sheet retreat in the Amundsen sea driven by decadal oceanic variability. *Nat. Geosci.* **11**, 733–738 (2018).

78. Williams, C. R., Arthern, R. & Byrne, J. Forward simulations of the amundsen sea sector of west antarctica at different resolutions produced from the ice sheet model wavi (version 1.0) [data set]. *NERC EDS UK Polar Data Centre* https://doi.org/10.5285/4de39bc0-fc2b-4232-ac39-3cc1fd723f64 (2024).

79. Leguy, G. R., Lipscomb, W. H. & Asay-Davis, X. S. Marine ice sheet experiments with the community ice sheet model. *Cryosphere* **15**, 3229–3253 (2021).

80. GPy. GPy: A gaussian process framework in python. http://github.com/SheffieldML/GPy (2012).

81. Thodoroff, P. Pierthodo/multi_resolution_ice_sheet: V1. *Code on Zenodo* https://doi.org/10.5281/zenodo.14422770 (2024).

82. Arthern, R., Bradley, A., Williams, C. R. & Bett, D. WAVI.jl–Williams2024: Release of WAVI used for Williams et al 2024. *Code on Zenodo* https://doi.org/10.5281/zenodo.14393522 (2024).

Author contributions

C. Rosie Williams devised the ice sheet model experiments, initialised the model, ran the simulations and analysed the outputs. Pierre Thodoroff devised and performed the Gaussian process emulation and computational budget calculations. Robert Arthern provided expert advice on the ice sheet modelling and analysis of the model results. James Byrne provided the model ensembling framework and assisted in setting this up for the WAVI model. Scott Hosking helped form the project team and assisted in identifying the goals. Markus Kaiser provided expert advice on machine learning and experimental design during initial stages of the project. Neil Lawrence supervised Pierre Thodoroff's doctoral thesis and provided expertise on the emulation and computational budget work. Ieva Kazlauskaite provided expert advice on machine learning and experimental design throughout the project. C. Rosie Williams, Pierre Thodoroff, Robert Arthern and Ieva Kazlauskaite analysed results and drew conclusions. C. Rosie Williams and Pierre Thodoroff wrote the manuscript, which all authors read and edited.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at
https://doi.org/10.1038/s43247-025-02010-z.

Correspondence and requests for materials should be addressed to C. Rosie Williams.

Peer review information *Communications Earth & Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Shin Sugiyama, Joe Aslin and Aliénor Lavergne. A peer review file is available.

Reprints and permissions information is available at
http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.