

DeepMelt-GL v1: A neural network emulator of ice-shelf melt rates for use in ocean models which partially resolve ice-shelf cavities

Helen Ockenden¹, Clara Burgard^{2,6}, Pierre Mathiot¹, Christoph Kittel^{3,4}, Achille Gellens^{5,6},
Cécile Agosta^{5,6}, and Nicolas C. Jourdain¹

¹Université Grenoble Alpes, CNRS, IRD, INRAE, Grenoble INP, IGE, Grenoble, France,

²Laboratoire d'Océanographie et du Climat Expérimentations et Approches Numériques (LOCEAN), Sorbonne Université, CNRS/IRD/MNHN, Paris, France,

³Laboratory of Climatology, Department of Geography, SPHERES research unit, University of Liège, Liège, Belgium,

⁴Physical Geography Research Group, Department of Geography, Vrije Universiteit Brussel, Brussel, Belgium,

⁵Laboratoire des Sciences du Climat et de l'Environnement (LSCE), CEA, CNRS, UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France,

⁶Institut Pierre-Simon Laplace (IPSL), Paris, France.

Correspondence: Helen Ockenden (helen.ockenden@univ-grenoble-alpes.fr)

Abstract. The spatial pattern of melting beneath Antarctic ice shelves influences ice flow and retreat, and the resulting fresh-water input into the Southern Ocean influences global carbon storage and primary productivity. It is therefore crucial that interactions at the ice-ocean interface are adequately represented in global climate models. However, due to computational limitations, existing climate models are forced to choose between either high resolution or long period simulations, and struggle to resolve melt rates in ice shelf cavities. Here, we show that a simple multilayer perceptron can be used to emulate sub-shelf melt rates in entire small cavities and in the parts of large cavities closest to the grounding zone. We find that the melt rates produced by applying neural network emulators are a good match for melt rates from high resolution simulations, provided that similar conditions were included in the neural network training. We also find that neural networks are particularly sensitive to temperatures and ice-draft slopes outside those used in the training dataset. However, if we train multiple neural networks on the same input data, we demonstrate that the ensemble spread of the neural networks is a good indicator of the reliability of the emulator in any given conditions. The neural network emulator of sub-shelf melt which is presented here, DeepMelt-GL, can be used to improve the representation of ice shelf cavities in both ocean and ~~global~~ climate models.

1 Introduction

The mass loss of the Antarctic ice sheet has led on average to a freshwater input of 2680 Gt yr^{-1} to the Southern Ocean over the period 1997 to 2021 (?). 40 % of this freshwater (1080 Gt yr^{-1}) enters the ocean at a few hundred meters depth, in the form of meltwater originating at the base of the ice shelves, the floating margins of the ice sheet. This meltwater has local and global implications for the ocean, driving overturning circulation, affecting low-latitude primary productivity and influencing the oceanic carbon sink (??), as well as being a driver of mass loss from the grounded Antarctic Ice Sheet (??). As the volume of meltwater from the base of ice shelves is predicted to increase in the future (?????), it is crucial that future

20 climate projections include coupled ocean-ice-sheet models which capture both the response of ocean properties to changing volumes of sub-shelf melt, and the response of sub-shelf melt to changing ocean properties and evolving cavity geometry.

Current modelling approaches which account in some way for sub-shelf ice-ocean interactions include prescribed freshwater fluxes from external climatologies in global ocean models (??), some regional and global coupled ocean-ice-sheet models (???), and sub-shelf melt rates calculated from far-field ocean conditions in continental ice-sheet models (???). Global ocean
25 models used in climate models for multi-decadal and centennial simulations usually work on a grid resolution of approximately 1° (~ 56 km at 60° S, ~ 30 km at 75° S). Many models do not consider sub-ice-shelf cavities, and as a result the freshwater flux into the ocean from sub-ice-shelf melt is prescribed by redistributing a fraction or all of the surface mass balance (precipitation minus sublimation and runoff) of the Antarctic continent into the ocean surrounding the continent (?). This is mostly done to conserve the ocean volume in the climate simulations, but neglects changes in freshwater fluxes due to ice sheet dynamics. It is
30 rare that ocean circulation within ice-shelf cavities is resolved within global ocean models, and it was not in any of the models in the Coupled Model Intercomparison Project Phase 6 (CMIP6) (??). This is one of the reasons why the modelled ocean behaviour around the Antarctic continent often does not match observations of water properties (??). Although regional ocean models that resolve the ocean circulation below ice shelves are becoming more common (?), they are not always designed for multi-centennial or ensemble simulations, due to computational limitations (?).

35 First steps have been made towards the representation of the ocean circulation below ice shelves in $\sim 1^\circ$ ocean models (e.g. ???). However, at 1° resolution, while cavity circulation in the largest ice-shelves (such as Ross, Filchner-Ronne, Larsen C) is represented, smaller ice shelves with high observed melt rates (such as Thwaites, Pine Island and Totten, ?) and all grounding lines remain unresolved or poorly resolved (?). This limits our ability to assess the consequences of both the effect of sub-shelf meltwater on the ocean and the localised effects of sub-shelf melt on the ice sheet dynamics and mass loss.

40 Continental ice-sheet models need sub-shelf melt forcing for all ice shelves. Currently, a range of parameterisations can be applied to infer sub-shelf melt patterns from non-cavity resolving global ocean models (??)(??). However, these parameterisations often fail to reproduce the spatial melt rate patterns (?), especially in warmer ocean conditions representative of future climates (???). Although these parameterisations represent an improvement on the prescribed melt used in ocean models, alternative approaches which fully capture the feedbacks between ocean properties, ice shelf-geometry and sub-shelf melt are
45 needed to facilitate multi-centennial, ensemble, coupled simulations. An option is to use an intermediate system that couples an ice sheet and an ocean model using these parameterisations (??). Another alternative is to leverage machine learning techniques to emulate melt rates from high resolution models.

Supervised machine learning offers the opportunity to use high resolution simulated datasets to improve the parametrisation of the relationship between sub-shelf melting and water properties as resolved in a global ocean model. Using different
50 architectures, ? and ? both showed that simple neural networks can reproduce ice-shelf melting patterns produced in model simulations, and that, in warmer conditions, neural networks have a lower error when compared to reference melt rates than the majority of traditional parameterisations. In both studies, neural networks were used to parametrise the whole of the ice-shelf cavity for use with stand-alone ice-sheet models. ~~Here,~~ Since some ocean models include ice-shelf cavities, here we explore the suitability of using neural networks to parametrise ~~only the part of the cavity that is~~ the parts of cavities which are not resolved

55 in [these](#) ocean models. [For small cavities, this is the entire cavity, whereas for large cavities this means just the part closest to the grounding line.](#) This approach allows us to maintain the full physics of water flow in ice-shelf cavities where resolved, and to ~~use a well-trained neural network to reproduce ice-shelf melt patterns in regions not resolved by the ocean model~~[produce sub-shelf melt patterns for the entire cavity.](#)

60 In this study, we investigate the advantages and limitations of a neural-network emulator of the sub-shelf regions not resolved in the 1° version of the NEMO ocean-sea-ice model (Nucleus for European Modelling of the Ocean, ?), which we train using NEMO simulations with a resolution of 0.25° (4-14 km around Antarctica). Using selected input conditions, we explore the emulator performance in similar conditions to the training data, in conditions outside the training data, and with limited quantities of training data. We produce an emulator which can be used in the NEMO global ocean model, and make recommendations about the conditions in which the neural network emulator can be reliably used.

65 2 Methods

This study aims to explore whether a neural network can be trained to emulate sub-shelf melting in parts of the ice-shelf cavity not resolved by a 1° ocean model, by identifying relationships between the water properties from the ocean model, the geometry of those cavity regions, and the sub-shelf melt rates simulated in relatively higher resolution (0.25°) models. Alongside this manuscript, we provide a neural network trained using all available NEMO 0.25° simulations for application
70 in stand-alone ocean models or coupled systems. However, in order to better understand the conditions in which this trained neural network can be expected to perform well, we additionally trained neural networks on specific individual simulations. Here, we discuss the ocean model simulations that we aim to emulate, the architecture of the neural network used for emulation, and the experimental procedure through which neural network performance is assessed.

2.1 Data: simulated ice-ocean conditions

75 This work aims to produce a neural-network emulator of the sub-ice-shelf regions not resolved in the 1° version of the NEMO ocean-sea-ice model (Nucleus for European Modelling of the Ocean, ?). In order to do this, we employ a selection of NEMO simulations of ocean conditions run on the eORCA025 or eANT025 grids, which are global (?) and regional circum-Antarctic ~~(?)(?)~~ grids with a resolution of 0.25° (4-14 km around Antarctica). We [use simulated melt rates as our “reference” melt rates because there are currently insufficient simultaneous observations of melt-rate and water properties, and because this allows](#)
80 [us to include future climatic conditions which have not yet been observed. We stress that model parameters in the NEMO ocean model are tuned using estimated sub-shelf melt rates and water properties from the last 40 years \(?\).](#) We focus on grid cells which are resolved as ocean points beneath the ice shelf in the 0.25° simulations, but which would be defined as land in the corresponding 1° model (Figure 1). In all simulations, sub-shelf melt is calculated using the widely-used three-equation thermodynamic system that accounts for the heat and salt balances at the ice-ocean interface, and the pressure and temperature
85 dependence of freezing temperature (???)

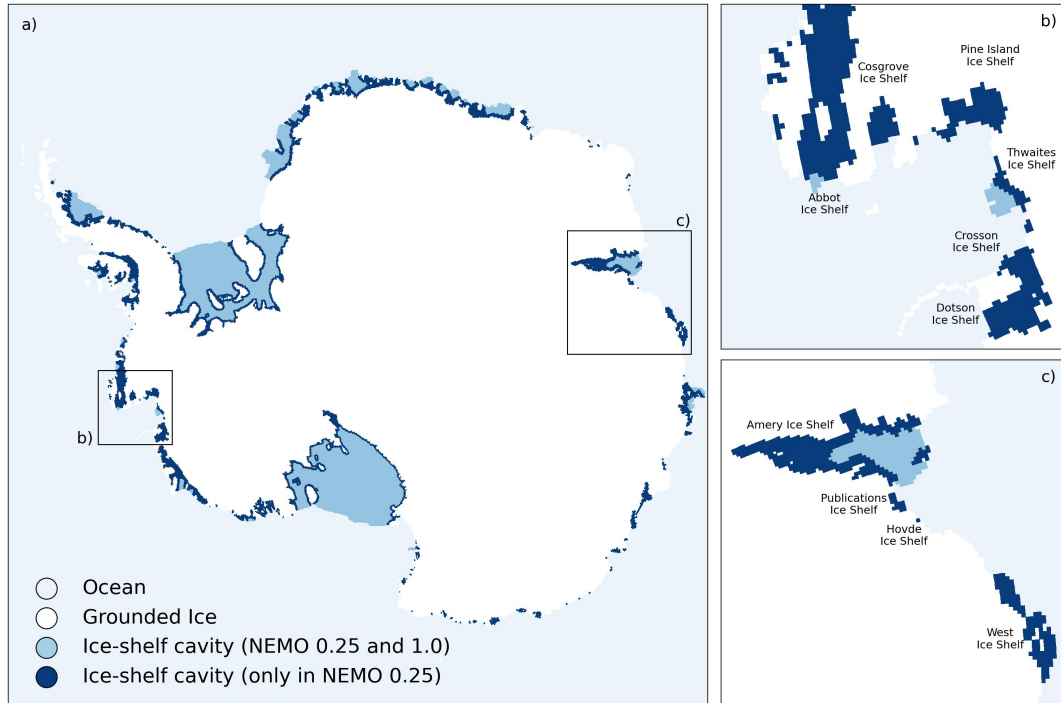


Figure 1. The spatial distribution of ice-shelf cavities which are resolved in the NEMO model at 0.25° and 1° . Panel a) shows the distribution of cavities resolved at 0.25° (dark and light blue) and the ones that would be resolved at 1.0° (light blue only) for the whole Antarctic continent, illustrating that although the largest ice shelves are resolved in the 1.0° model, many important and fast-changing regions require a higher resolution for sub-ice shelf dynamics to be included in ocean models. Panels b) and c) show zoomed-in insets for the Amundsen Sea sector and Amery Ice Shelf as examples. We use the regions which are only resolved in the 0.25° simulations (dark blue) for training the neural network, with ocean properties there propagated from the area that would be resolved at 1° (light blue area).

To assess the performance of neural networks as emulators of sub-shelf melt in these partial cavity regions, we use monthly data from four different NEMO simulations (ANT2000, ANTFGIOM, ANT2100, ANT2300, differences summarised in Table 1). Using monthly data allows us to increase the volume of training data and cover a larger range of input conditions. We use two simulations with present-day forcing, both driven by the JRA55-do atmospheric reanalysis (?) and initialised in 1979 from the climatological WOA2018 conditions (?). In the first present-day simulation, which we call ANT2000, bathymetry and ice shelf draft are a modified version of those used by ?, updated to be closer to Bedmachine Antarctica v2 on the Antarctic continental shelves (?) and with some other small modifications as described in ?. In the second, which we call ANTFGIOM, ice shelf draft is artificially thinned in the Amundsen and Weddell Sea regions and the grounding line retreated, following the MISOMIP2 protocol (Figure A1, ?), to be representative of a possible future geometric configuration, but it is still driven by present-day atmospheric conditions. There are also some small differences in the geometry in other ice shelf basins (Figure 6).

We also use two simulations which simulate plausible future ocean conditions. Both use the same ice-shelf draft and bathymetry as ANT2000. The first of these simulations, which we call ANT2100, uses a climate forcing driven by continuous atmospheric and oceanic anomalies from the CNRM-CM6-1 projections under the SSP5-8.5 scenario (?) until 2100, added to the climatology. ANT2100 is restricted to a regional domain south of 52.5°S. The second simulation, which we call
100 ANT2300, is global and applies a temperature anomaly extracted from the IPSL-CM6A-LR projections under the SSP5-8.5 emissions scenario (??), in order to explore potential extreme conditions at the end of the 23rd century. Further details about the exact parameters used in these simulations can be found in ?.

Additionally, at the end of this study we present a neural network emulator that has been trained on all available 0.25° simulations which we have access to. In addition to the monthly data from the four simulations described above, we also use
105 annual data from five additional simulations. Three of these simulations (WARMROSS, COLDAMU, REALISTIC) were part of the ensemble used for training in ?, and have the same forcing and bathymetry as ANT2000 but vary in the parameters used for oceanic behaviour. Further information about these simulations can be found in ?. The remaining simulations (ANTPRE, ANT-RU2100) use the same regional ocean model configuration as ANT2100. ANTPRE is forced by JRA-55 for the atmosphere and by the present-day ocean simulation of ANT2000 for the ocean. ANT-RU2100 uses the same climate forcing as
110 ANT2100, but with the addition of surface runoff from MAR as an additional freshwater source (?).

2.1.1 Spatial focus

This study aims to develop a neural-network emulator for use in coupled ice-sheet-ocean modelling, which allows as much of the ocean and ice-shelf cavity to be resolved in the ocean model as possible. Therefore, we train the neural network only in regions where the sub-shelf cavity is resolved in the 0.25° model, but which are considered to be grounded ice in a hypothetical
115 equivalent 1° model (Figure 1, dark blue). Regions which would be resolved as sub-shelf cavity in the 1° model are not included. For small ice-shelves, this means emulating the entire ice shelf. However, for the largest ice shelves, we only need to emulate a relatively thin strip around the innermost edge of the ice shelf. In order to have a flexible framework which can adapt to evolving shelf geometries and changing grounding line positions, we divide the coastline of Antarctica into 155 partial ice-shelf basins, based on the locations of the individual outlet glaciers and ice streams (See Appendix A2, ?). Large
120 pinning points (ice rumples and rises) are also separated and form their own basins. These basins are extrapolated into the open ocean using a nearest neighbour approach, and occasionally adjusted manually to ensure that basins which contain an ice-shelf front always have a few points in the open ocean. The resulting partial ice-shelf basins all have similar dimensions, with the exception of basins which include islands and are not part of the main coastline. The margins of the largest ice-shelves fall into several of these partial ice-shelf basins.

125 2.2 Neural network

Supervised machine learning offers the opportunity to use high-resolution simulated datasets to improve the parametrisation of the relationship between water properties and sub-ice-shelf melting. Simple neural networks can leverage computational data structures to identify relationships between ocean conditions and sub-ice-shelf melt rates that would otherwise be difficult

Table 1. Description of the climate forcing and geometry employed in for the simulations used in this study. All simulations have a resolution of 0.25°. WARMROSS, COLDAMU and REALISTIC vary in their internal model parameters (?).

Simulation	Time period	Climate forcing	Geometry
ANT2000 (Present day conditions)	89 years	JRA-55 atmospheric forcing reanalysis initialised from WOA2018 climatology	Modified from Storkey et al. (2018) as described in Mathiot and Jourdain (2023)
ANTFGEOM (Altered geometry)	39 years	JRA-55 atmospheric forcing reanalysis initialised from WOA2018 climatology	Artificially thinned in the Amundsen and Weddell regions, following the MISOMIP2 protocol (De Rydt et al., 2024)
ANT2100 (2100 like conditions)	86 years	CNRM-CM6-1 projections under the SSP5-8.5 emissions scenario until 2100 (Voldoire et al. 2019)	Same as ANT2000, restricted to south of 52.5°
ANT2300 (2300-like conditions)	99 years	IPSL-CM6A-LR projections under the SSP5-8.5 emissions scenario (2300) (Boucher et al., 2020)	Modified from Storkey et al. (2018) as described in Mathiot and Jourdain (2023)
WARMROSS (warm eastern Ross Ice Shelf)	28 years	Same as ANT2000	Same as ANT2000
COLDAMU (relatively cold Amundsenn Sea)	28 years	Same as ANT2000	Same as ANT2000
REALISTIC (similar to present day)	29 years	Same as ANT2000	Same as ANT2000
ANTPRE (present day)	31 years	JRA-55 for the atmosphere and ANT2000 for the ocean	Same as ANT2000, restricted to south of 52.5°
ANT-RU2100 (2100 like conditions + runoff)	86 years	CNRM-CM6-1 projections under the SSP5-8.5 scenario until 2100, plus surface runoff from MAR (Voldoire et al., 2019; Kittel et al., 2021)	Same as ANT2000, restricted to south of 52.5°S

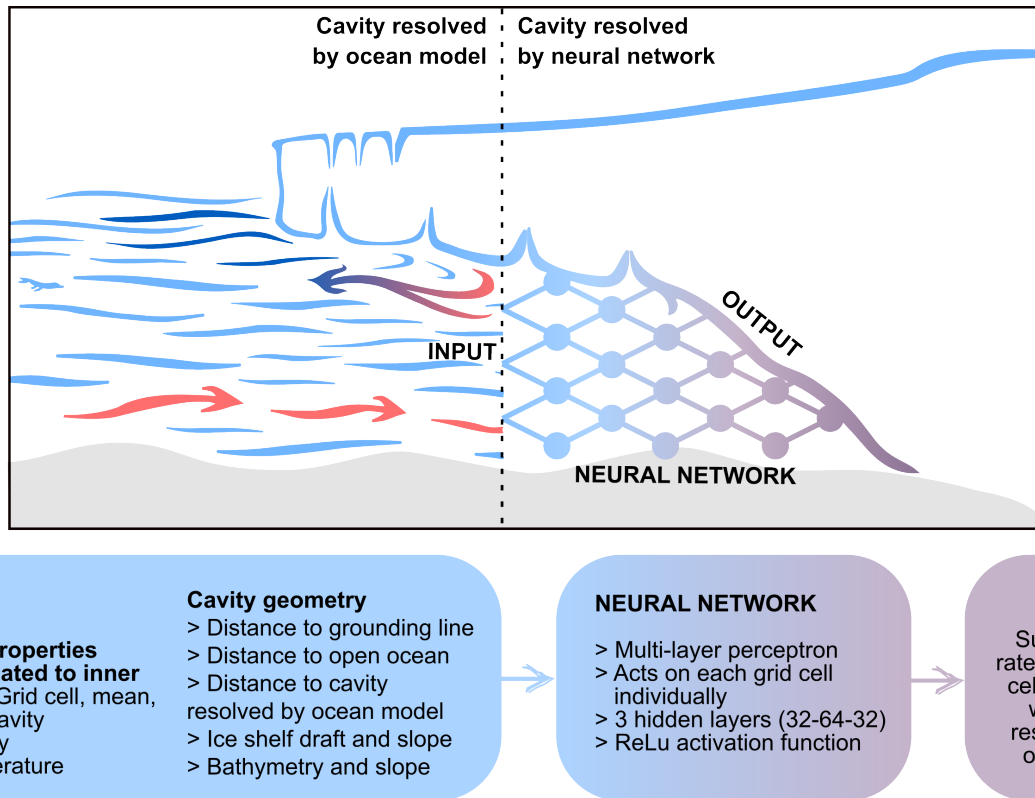


Figure 2. A schematic showing the key inputs and outputs involved in the neural network training and application. The key inputs and outputs involved in the neural network training and application. Water properties and cavity geometry are separate classes of input properties because water properties come from the edge of the cavity as resolved in the 1° ocean model, and cavity geometry comes from the region which is not resolved in the 1° ocean model. In training, cavity geometry is the one prescribed in the NEMO model, but in a coupled ocean-ice-sheet ~~models-model~~ the geometry would come from the ice-sheet model.

to identify in this multi-dimensional data set (??). We use a pointwise neural-network architecture based on the multilayer
 130 perceptron used by ? to predict sub-ice-shelf melt rates using information on the sub-shelf geometry and ocean conditions.

2.2.1 Architecture and hyper-parameters

Following ?, we use a simple multilayer perceptron implemented using the keras package in python (?), which has multiple
 input variables and a single output variable: sub-shelf melt rate. As we consider a slightly different spatial domain, comprising
 a mixture of entire small ice shelves and the peripheral strip of larger ice shelves, we have a slightly different set of input
 135 parameters than ? (See 2.2.2). ~~Input and output variables are connected by 3 hidden layers~~Inputs are processed through three
 sets of computational layers of artificial neurons (hidden layers), which apply weightings and non-linear activation functions,
 allowing the neural network to represent complex non-linear relationships. We use three hidden layers of 32, 64 and 32 neurons

respectively, as this small neural network size was previously found to give good results in cross validation tests over both time and cavity geometry without over-fitting input data (?). During the neural network training process, initially randomised weights are optimised with a gradient descent algorithm (Adam, ?) which is tuned to reduce the mean-squared error (mse) of the reference melt rates from the simulation data relative to the predicted melt rates from the neural network. We use an initial learning rate of 0.001, and a ReLU activation function (?)(??).

To assess the performance of the neural network training, and to avoid over-fitting the data, we use a validation data set containing a randomly chosen 10% of the shuffled training data, which is separated before training begins. We explore the impact of these stochastic processes on the performance of the neural network by applying the same training process initialised with 10 different random seeds, giving us an ensemble of neural networks. Further information on the choice of neural network architecture and hyper-parameters can be found in ?.

2.2.2 Input variables

The neural network is trained and applied to an array of variables calculated at each grid point individually. Similarly to ? we employ two different types of variables: information about the geometry of the ice-shelf cavity in the grid cell, and information about the water properties (temperature and salinity) at the front of the partial ice-shelf cavity. For both training and testing, we compare the output of the neural network to reference sub-shelf melt rates from 0.25° ocean simulations.

Geometrically, we provide the following variables: the distance to the closest point on the grounding line, the closest point in the open ocean, the closest point in the open part of the same cavity (if there is one resolved at 1°), the depth of the ice-shelf draft, and the bathymetry. Unlike ?, who calculated the meridional and zonal slopes, we calculate the slope of the ice-shelf draft and bathymetry both parallel and perpendicular to the front of the closed cavity (resolved at 0.25° but not 1°), in order to better account for ice shelves (or ice-shelf segments) that are not oriented facing north. The orientation of the cavity front is calculated using the mean locations of cavity grid points and the mean locations of ocean points along the front of the cavity.

Hydrographically, some pre-processing of the data is required. In order to have a point-wise neural network architecture, we need information about the temperature and salinity at each point. However, as the aim of the neural network is to emulate these cavities to be eventually used in a NEMO 1° model, the neural network must be trained with water property information which is available at that resolution. As a result, for each of the 155 partial ice-shelf basins, we calculate a mean temperature and salinity profile for the closest ocean points that would be resolved in the corresponding 1° model, which can be either open ocean or open cavities. We identify these points using a kdtree algorithm. We then infer water properties for each grid point in the emulated region by taking these properties from the mean profile at the same depth as the grid point. Additionally, to provide some information about the spatial variability of water properties, we also include the mean and standard deviation of the temperature and salinity across the partial ice-shelf basin. Unlike ??, who used potential temperature and practical salinity, we use conservative temperature and absolute salinity, which improves conservation in the propagation and reduces the data pre-processing needed as these are used in the NEMO ocean model.

170 2.3 Experiments and metrics for evaluation

Using monthly data from the four ocean simulations (ANT2000, ANTFGEOM, ANT2100, ANT2300) and the neural network architecture outlined previously, we carry out several experiments to explore the performance of the neural network emulator. For initial model verification, data from each of the four simulations was used to train the neural network, and model performance was verified for those simulations. In all other experiments, data from one or more simulations is used to train the neural network, and then data from a different simulation is used to assess the performance of the emulator in a range of conditions.

For each experiment, we report a range of metrics for emulator performance. Firstly, we report the root-mean-squared error (RMSE) of the predicted melt rate compared to the reference simulated melt rate, since MSE is the variable optimised during the neural network training, and this is a standard metric in machine learning. However, as we have data which vary both spatially and temporally, we also focus on metrics which allow us to explore those dimensions. Specifically, we look at the integrated melt on annual and monthly timescales, and grouped by regional ice-drainage basins as defined by the Ice sheet Mass Balance Inter-comparison Exercise (IMBIE, ?). In addition to the RMSE, we also report the correlation coefficient (R^2) between the predicted and the reference melt for integrated melt rates.

We carry out three main sets of experiments, in each case training the neural network with one simulation, and testing the quality of the results by applying the trained neural network to an independent simulation. Firstly, we explore the performance of the neural network emulator in different climatic conditions to those used for the training, using the ANT2000, ANT2100 and ANT2300 simulations, which have almost identical geometries but differing water column temperature and salinity. Secondly, we explore the performance of the neural network emulator in different geometric conditions, using the ANT2000 and ANTFGEOM simulations, which have the same climatic forcing, but different geometries. Since all 155 partial ice-shelf basins have different geometries, we would expect that a neural network trained on the ANT2000 data set should give good results on a simulation where the geometries of the basins have changed. Finally, we explore the impact of the quantity of training data on the performance of the neural network, by training neural networks with data from part of the ANT2300 simulation, and testing on the ANT2100 simulation.

All tests which we describe here use monthly data, in order to increase the size of the training data set, and to ensure that it encompasses as wide a range of conditions as possible. Additional training and testing on both annual and monthly data shows that the neural network performance is significantly improved by using monthly data, due to the improved representation of extreme melt values in the training dataset. More information about these additional tests can be found in Appendix A3.

3 Results

3.1 Model verification

We verify that the neural network architecture can return realistic melt rate values by applying a trained neural network model to the same dataset used for training. For all 4four simulations used here, there is a strong correlation between the reference melt and the predicted melt rates for both the non-integrated pointwise melt (R^2 between 0.86 and 0.94) and the spatially

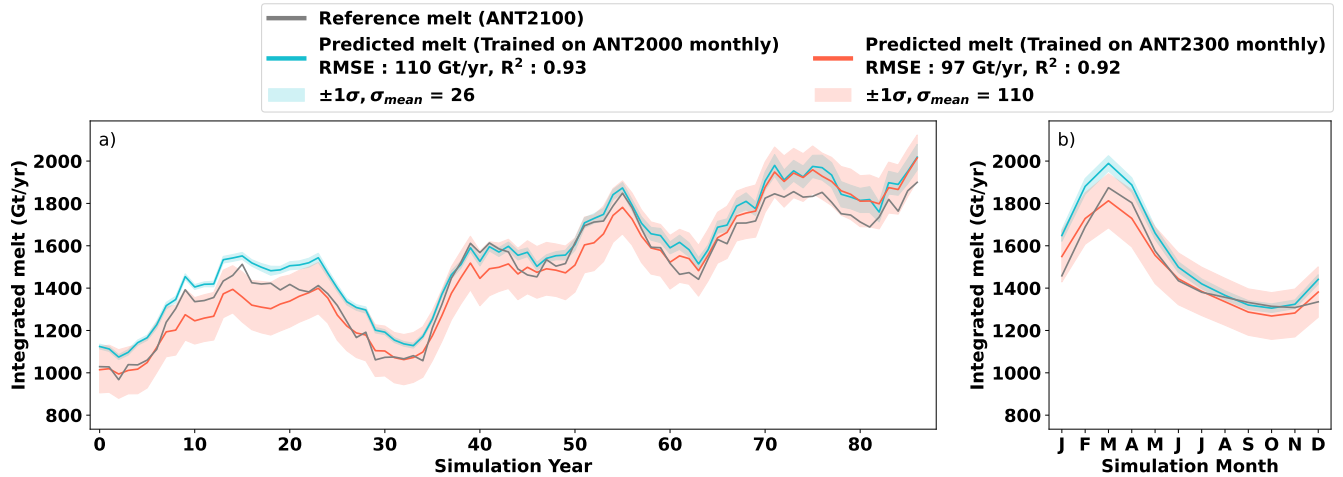


Figure 3. Neural network performance in temperature conditions outside the training dataset. The reference melt rates from the ANT2100 simulation (grey), compared to the predicted melt rates from neural networks trained on monthly data from the ANT2000 (blue) and ANT2300 (red) simulations, grouped into annual time series (a) and monthly-mean seasonal cycle (b) trends. Melt rates are integrated melt rates for the whole of Antarctica in Gt/yr^{-1} , to allow the same scale for annual and monthly trend. The ensemble spread of the 10 neural networks is shown through the error bounds of one standard deviation.

integrated monthly melt rates (R^2 between 0.92 and 0.96). Where the correlation is better for integrated melt rates this is likely due to the cancellation of errors. The full set of metrics including RMSE for the model verification can be found in Figure A5.

3.2 Does the NN parameterisation work in different climatic conditions?

To explore the adaptability of the neural network to different climate scenarios, we trained on the ANT2000, ANT2100 and ANT2300 simulations, which have almost identical geometries, but differing water column temperature and salinity. The maximum grid-cell melt rates are 117 m yr^{-1} for the ANT2000 simulation, 181 m yr^{-1} for the ANT2100 simulation, and 495 m yr^{-1} for the ANT2300 simulation (metres of ice equivalent per year).

Firstly, we consider a neural network trained only with data from the ANT2000 simulation, representing present day conditions. When this neural network is applied to the ANT2100 simulation, which represents warm 2100-like conditions, we find a good match between the integrated melt predicted by the neural network and the reference melt (Figure 3, $\text{RMSE} = 110 \text{ Gt yr}^{-1}$, and $R^2 = 0.93$). The ensemble spread (standard deviation of the ten ensemble members) is 26 Gt yr^{-1} , highlighting a good agreement between ensemble members and a well constrained parameter space. This result shows that the neural network emulator can be applied to climate conditions outside the conditions in the training dataset and still produce realistic melt values.

When the ANT2000-trained neural network is applied to the ANT2300 simulation, which represents a hot 2300-like climate, we see that the neural network consistently under-predicts the melt rate (Figure 4, $\text{RMSE} = 2220 \text{ Gt yr}^{-1}$, $R^2 = 0.70$). We

note that in this case the ensemble spread is 750 Gt yr^{-1} , indicating that the neural network emulator is poorly constrained in this parameter space. These results demonstrate a limitation of these neural networks; if applied to conditions which are too different to the training dataset, the predicted melt rates will not match the expected melt rates.

Next, we consider a neural network trained only with data from the ANT2100 simulation, representing warm 2100-like conditions. When this neural network is applied to the ANT2300 simulation, which represents a hot-2300 like climate, we see an improvement in the results relative to the ANT2000-trained neural network (Figure 4a, $\text{RMSE} = 810 \text{ Gt yr}^{-1}$, $R^2 = 0.93$). Although the ANT2100 neural network still under-predicts the melt rate, especially at the end of the ~~ANT2300~~ ANT2300 simulation, the predicted and reference melt rates diverge later in the simulation and at a higher integrated melt value. This suggests that there is a ~~temperature~~ threshold above which the neural network performance begins to degrade, which is controlled by the training data. Looking at RMSE as a function of conservative temperature and absolute salinity (Figure 5), we see that reducing salinity does not affect RMSE as much as increasing temperature, and so suggest that this threshold is likely to be a temperature threshold.

To assess spatial patterns in the the sub-shelf melt, we also consider the results of applying both the ANT2000 and ANT2100 neural networks to the ANT2300 simulation on a basin scale (Figure 3). We use the glacial drainage basins defined by the Ice-sheet Mass Balance Inter-comparison Exercise (IMBIE). In many basins we see that both neural networks underpredict the melt, but the RMSE is lower for the ANT2100 neural network as it under-predicts by a smaller volume (Ap-B, B-C, E-Ep, Ep-F, G-H, H-Hp, I-Ipp, Ipp-J, J-Jpp). In several other basins, primarily along the coasts of Dronning Maud Land and Victoria Land, the ANT2000 neural network underestimates the melt, but the ANT2100 neural network matches the reference melt (A-Ap, D-Dp, Dp-E, Jpp-K, K-A). Finally, there are some basins that exhibit unique behaviour, such as Hp-I, which contains George VI ice shelf, where the ANT2000 neural network over-predicts the melt and the ANT2100 neural network under-predicts the melt. The George VI ice shelf has a unique geometry within the ice shelves considered here, as it has two ice shelf fronts. Since we consider the mean water properties across the whole ice shelf front, this may cause problems during training. Additionally, there is a known long-standing bias in NEMO simulations in this region, which tends to overestimate melt compared to observations (?).

Finally, we consider a neural network trained only with data from the ANT2300 simulation, representing a hot 2300-like climate. When this neural network is applied to the ANT2100 simulation, which represents warm 2100-like conditions, we see that it reproduces model melt rates (Figure 3, $\text{RMSE} = 97 \text{ Gt yr}^{-1}$, $R^2 = 0.92$). Since the ANT2300 training dataset covers a wider range of conditions than the ANT2100 simulation, this result is as expected.

In order to better understand the origins of mismatches between the predicted and reference sub-shelf melt rates, we look at the RMSE as a function of the neural network input variables. In particular, since these three simulations all have the same geometries, we look at RMSE as a function of conservative temperature and absolute salinity (Figure 5), finding that the highest RMSE values occur when neural networks are applied to grid cells where the temperature is outside the range included in the training dataset for that neural network. The majority of data points in the ANT2000 simulation have a salinity between 33 and 35 g kg^{-1} , and a temperature between -2 and 1°C . In the ANT2100 simulation, the majority of points have salinity between

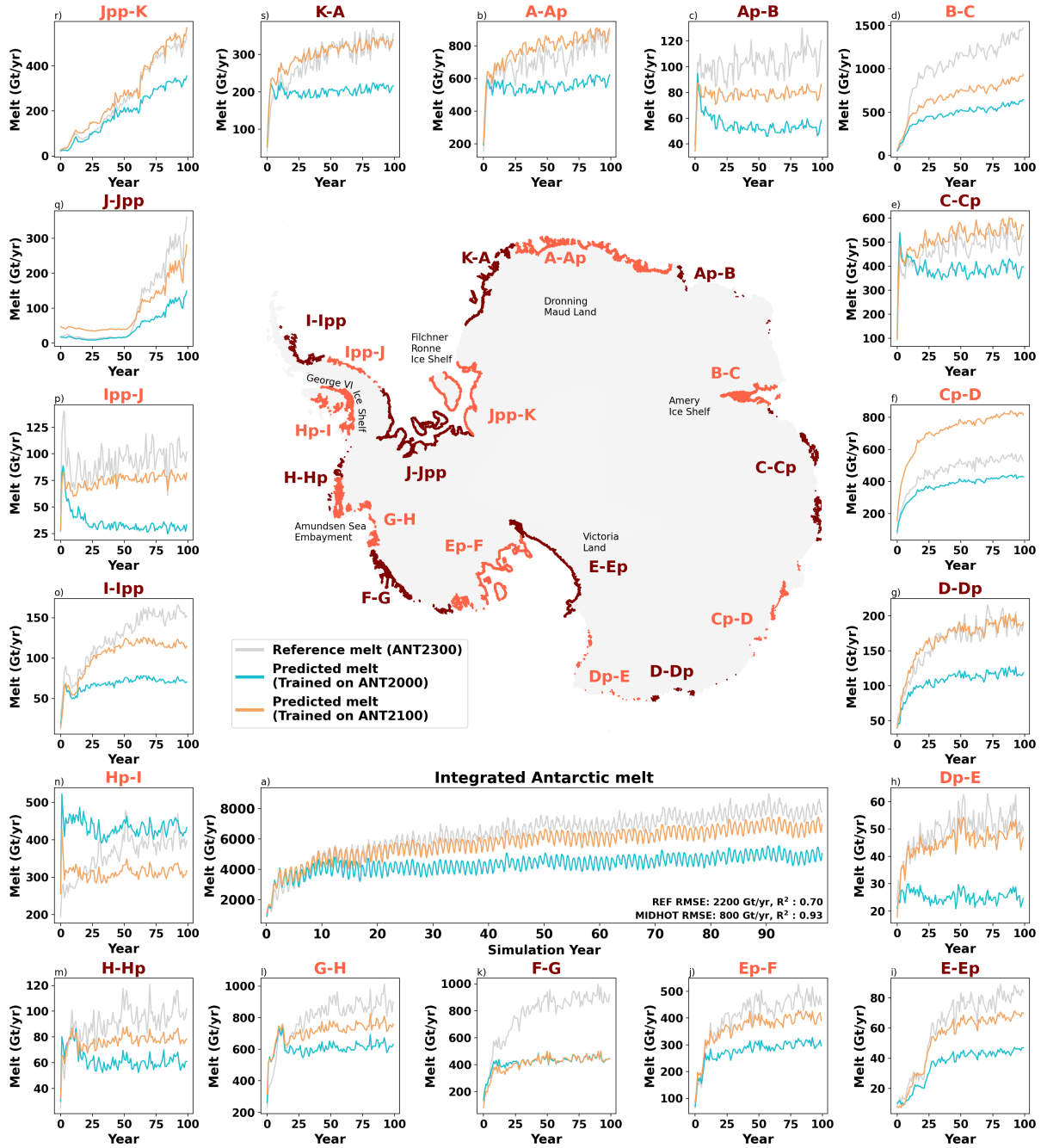


Figure 4. Neural network performance in temperature conditions outside the training dataset and in each of the IMBIE glacial drainage basins. A comparison of the reference melt rate from the ANT2300 simulation (grey) with the predicted melt rate from neural networks trained on monthly data from ANT2000 (blue) and ANT2100 (peach) to each glacial drainage basin in the ANT2300 simulation. a) Integrated melt rates for the whole of Antarctica. b-s) Integrated melt rates for each of the IMBIE glacial drainage basins. The location of each basin is shown in the central inset map. Regions which are discussed in the manuscript are labelled.

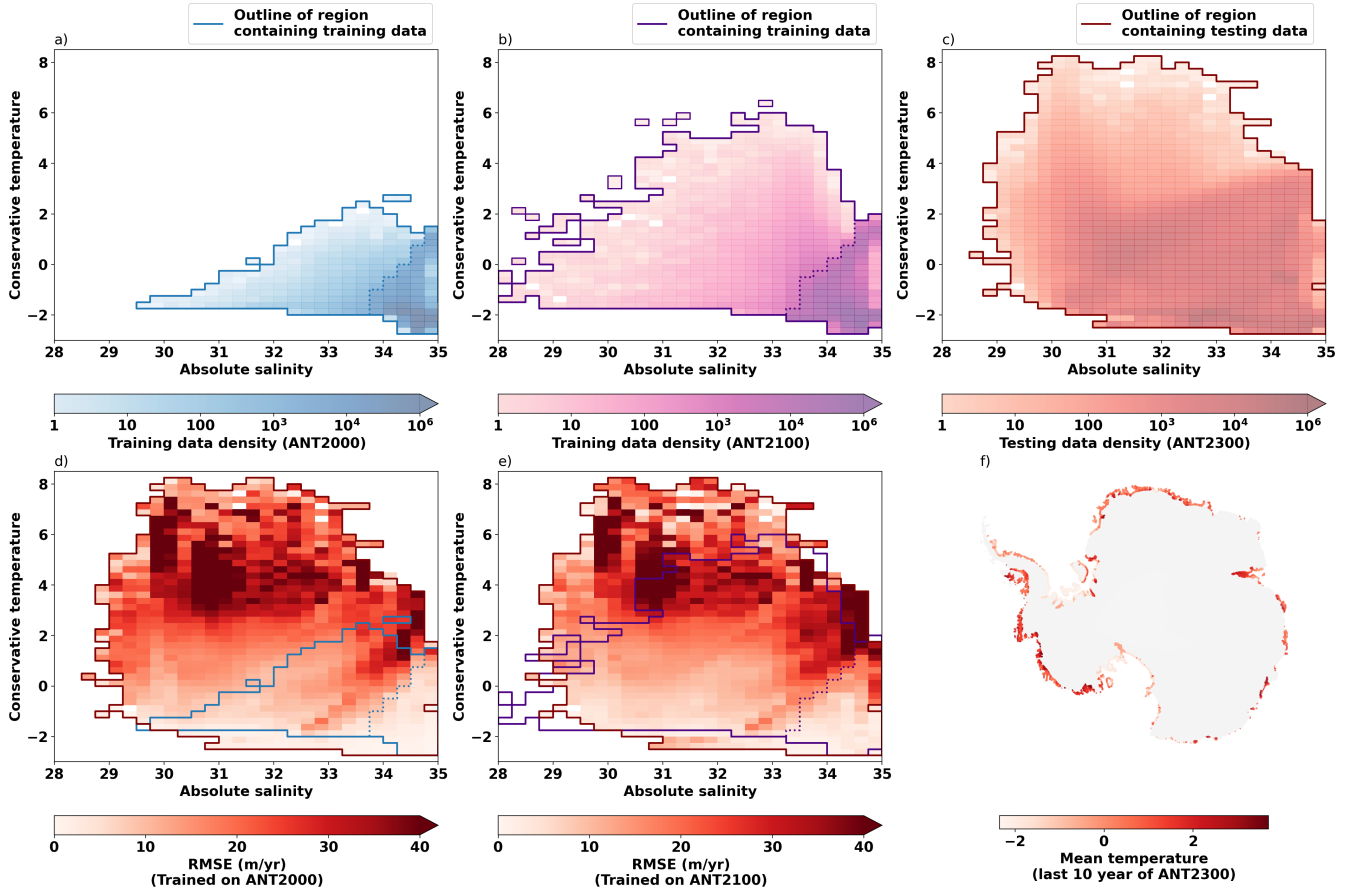


Figure 5. The highest RMSE values are observed for temperatures outside the range of the training dataset. The distribution of temperature and salinity for a) the ANT2000 simulation, b) the ANT2100 simulation and c) the ANT2300 simulation. [For ease of comparison, an outer contour is marked \(solid line\), and an inner contour bounding the region of highest data density \(dotted line\).](#) d) RMSE as a function of temperature and salinity when a neural network trained on the ANT2000 simulation is applied to the ANT2300 dataset. e) RMSE as a function of temperature and salinity when a neural network trained on the ANT2100 simulation is applied to the ANT2300 dataset. f) The mean temperature for the last 10 years of the ANT2300 dataset for each ice shelf grid point, [showing that the highest temperatures are in basins such as B-C \(Amery Ice Shelf\) and F-G \(Marie Byrd Land\), where the neural network results in Figure 4 were poorer.](#)

32 and 35 g kg⁻¹ and temperatures between -2 and 2°C. In contrast, in the ANT2300 dataset, salinities can be as low as 28 g kg⁻¹, and temperatures as high as 8°C.

255 Interestingly, extrapolating to salinities which are outside the range included in the training dataset does not seem to pose as much of a problem as extrapolating to temperatures not included in the dataset (Fig. 5d and e). This is likely why the neural network trained on the ANT2000 simulation performs well when applied to the ANT2100 simulation (Figure 3, RMSE = 110 Gt yr⁻¹), as the majority of grid cells have lower salinities but only marginally higher temperatures. There is also not a clear relationship between the mismatch to the training dataset and the error, suggesting that some of the other input variables also have an influence on RMSE. Despite these complex multivariate relationships, there is a clear need to include simulations
260 with high temperatures in the training when looking to emulate future climate conditions.

3.3 Does the NN emulator work for different ice-shelf geometries?

The ANT2000, ANT2100 and ANT2300 simulations all have almost identical geometries, as the ice shelf draft is fixed and does not vary with time. However, in a coupled ocean-ice sheet model, the geometry would evolve. We therefore also consider data from a fourth simulation, ANTFGIOM, in which some ice shelves have had their geometry synthetically altered. This
265 allows us to assess the impact of changes in geometry on the performance of the neural network.

Figure 6 shows the results of applying a neural network trained on the ANT2000 data to the ANTFGIOM simulation. We find a good match between the integrated melt predicted by the neural network and the reference melt (Figure 6, RMSE = 87 Gt yr⁻¹, R² = 0.90), although the predicted melt is a slight overestimation. The ensemble spread is ~~FILL IN 5.31~~ Gt yr⁻¹ demonstrating that the parameter space is well constrained and all ensemble members give similar results.

270 To assess spatial patterns in the sub-shelf melt, we consider the melt integrated over each of the eighteen IMBIE basins. ~~Melt rates for selected basins are shown in Figure 6b-s), and for all basins in Figure A2a-r.~~ There is significant variability in the quality of the ANT2000 neural network output ~~at this when integrated on a~~ basin scale. In particular, in the ~~western part of the Filchner-Ronne~~ Ronne Ice Shelf (basin J-Jpp, Figure 6qj), the neural network melt prediction is nearly four times the melt rate from the reference ANTFGIOM simulation. If this single basin is removed from the integrated Antarctic melt, then the
275 RMSE falls from 87 Gt yr⁻¹ to 45 Gt yr⁻¹). Since the Filchner-Ronne Basin had the largest changes in geometry between the ANT2000 and ANTFGIOM simulations, it is ~~unsurprising not necessarily surprising~~ that the largest errors are seen in this region. ~~It is more surprising that large errors are also seen in basins where the ice-shelf draft is only slightly different between the two simulations, such as the Amery,~~ although we note that large geometry changes in the Amundsen Sea (basin G-H, Figure 6h) and the Filchner Ice Shelf (basin B-CJpp-K, Figure 6d-i) lead to only moderate errors.

280 In order to better understand the origins of these large errors, we look again at the RMSE as a function of the neural network input variables, focussing on the geometric inputs. Since the sub-shelf cavity was expanded in the ANTFGIOM simulation, there are some points much further from the grounding line than in the dataset used to train the ANT2000 neural network, which we might anticipate to be producing errors. However, analysis does not show any correlation between RMSE and the distance of a grid point to the grounding line or to the ocean. Instead, ~~the~~ the geometric input variable with the strongest relationship to
285 RMSE values is the slope of the ice-shelf draft, and we see the highest RMSE values in grid cells with steep slopes (Figure 7b).

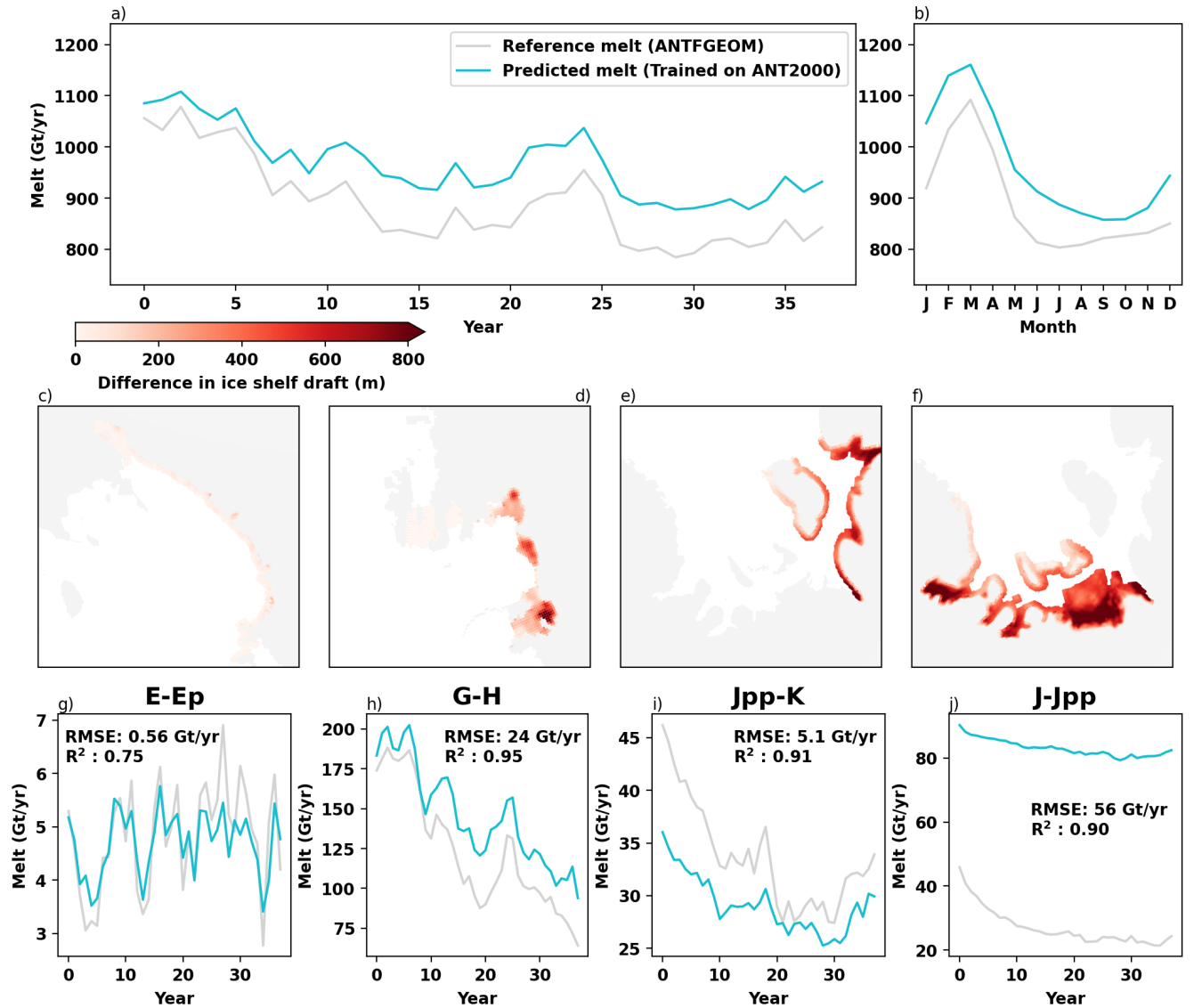


Figure 6. Neural network performance in a different geometry. The reference melt for the ANTFGEOM simulation (grey) and predicted melt for a neural network trained on the ANT2000 simulation (blue) either integrated over the whole of Antarctica (a, [yearly](#), b, [monthly](#)), or for [each-four](#) of the 18 IMBIE basins. [Panels c-f](#) show the difference in the ice shelf draft for basins E-Ep ([b-sRoss Ice Shelf](#)), G-H ([Amundsen Sea](#)), Jpp-K ([Filchner](#)) and J-Jpp ([Ronne](#)), which can be seen in more detail in [Appendix A1](#). [The Panels g-j](#) show the integrated melt rates for each of the four basins, and the root-mean-square error and correlation coefficient (R^2) are given for each basin. [Panel t](#) shows the difference. Similar plots for all 18 basins can be seen in the geometries around Antarctica [Figure A2](#). Specific difference in the Weddell and Amundsen Seas are shown in [Appendix A1](#).

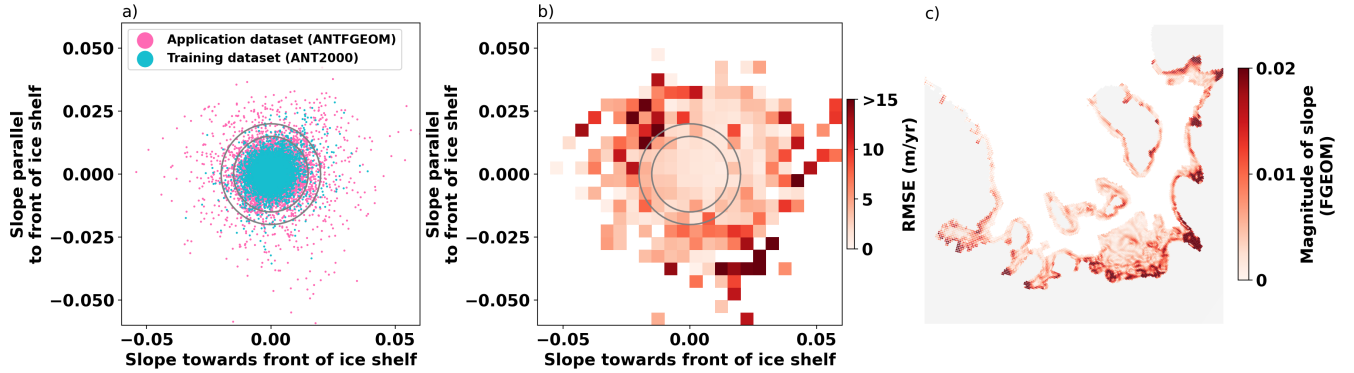


Figure 7. Average RMSE as a function of ice-shelf draft slope. a) The distribution of data in the training dataset, ANT2000 (blue) and the ANTFGEOM simulation (pink), compared to b) the RMSE as a function of the ice-shelf draft slope parallel ($s_{\parallel}$) and perpendicular ($s_{\perp}$) to the ice front when the neural network trained on the ANT2000 dataset is applied to the ANTFGEOM simulation. c) The magnitude of the slope ($\sqrt{s_{\perp}^2 + s_{\parallel}^2}$) of the ice-shelf draft ~~around Antarctica~~ in the Filchner-Ronne region in the FGEOM simulation. The grey circles show where the magnitude of the slope is 0.015 (inner) and 0.02 (outer) respectively.

, which are particularly located very close to the grounding line. Since the ice shelf geometry was artificially thinning of the ice shelf for this NEMO simulation, the slopes in this region may be steeper than would be simulated in a coupled ice-ocean model.

In the training ANT2000 simulation, there are very few points with a slope higher than 0.015 (Figure 7a). We see that when the ANT2000 neural network is applied to the ANTFGEOM simulation, the RMSE is highest where the slope exceeds 0.015 (Figure 7b). The steepest slopes in the ANTFGEOM simulation occur in the Filchner-Ronne Ice Shelf (Figure 6qj), Amery Ice Shelf (Figure 6A2d), and the Amundsen Sea (Figure 6h), which are all regions where the neural network struggles to reproduce melt rates. The geometry was only altered a small amount in the Amery Ice Shelf basin between the two simulations, but the slope of the ice shelf draft there is at the very extreme of the values used in the training dataset, and so the neural network training is not optimised for such slopes.

We note that when a neural network trained on the ANTFGEOM simulation is applied to the ANT2000 dataset, and the ice shelf slopes are all within the limits of the training data, then the RMSE of the integrated melt is only 39 Gt yr^{-1} , or 4% of the total integrated melt (not shown). It is therefore important that a neural network applied in a coupled simulation has been exposed to a range of slopes including steeper geometries in training.

3.4 Understanding neural network performance

In this manuscript we explore the performance of neural networks trained with limited data when applied to other simulations where we can verify the results against high resolution simulation melt rates. As expected, neural network performance is best when applied within the bounds of the training conditions, but we highlight that some input variables (temperature, slope of the ice shelf draft) seem to be particularly sensitive to extrapolation outside the training data. However, these neural networks

305 are designed for use in the NEMO 1 degree ocean model, where we will not be able to verify the results against simulated melt rates. We therefore explore other indicators of neural network performance.

Ensemble spread can be used as an indication of the suitability of the neural network for the conditions in which it is applied. If the ensemble members have a large spread, this indicates that the parameter space has not been well covered during training, and that the random selection of the training and validation data has had a large influence on different ensemble members. 310 Conversely, if the ensemble spread is small, then all ensemble members have converged on a similar solution, and the results are likely to be reliable.

Figure 8b shows the ensemble spread of a selection of neural networks when applied to the ANT2100 simulation. In general, high ensemble spreads are associated with high RMSE values. Sometimes, a high ensemble spread can be ~~assocaited~~associated with a low RMSE, but we do not see any examples of a low ensemble spread and high RMSE value. Therefore, we provisionally 315 suggest that if the ensemble spread is a significant fraction of the total integrated melt, then the neural network output should be considered with caution, and ideally more training data should be provided. However, this fractional threshold may vary depending on the nature of the input variable distributions in the training and application datasets.

We also explore how much training data is required to achieve an acceptable neural network performance. We train neural networks with varying quantities of data from the ANT2300 simulation, and apply them to the ANT2100 simulation, as we 320 know that a neural network trained on the whole ANT2300 performs well (Figure 3, $\text{RMSE} = 97 \text{ Gt yr}^{-1}$). Data is taken from either the beginning or the end of the ANT2300 simulation, or both. If 20 years of simulation data are used, this can either be 20 years from the beginning of the simulation (years 1-20), 20 years from the end of the simulation (years 81-100) or ten years from each of the beginning and the end (years 1-10 and 91-100). The results of applying these neural networks are shown in Figure 8.

325 Firstly, we note that which part of the ANT2300 simulation the data is taken from has a significant effect on the results, and that neural networks trained only with data from the end of the ANT2300 simulation do not perform well, with extremely high RMSE values. As the end of the ANT2300 simulation represents conditions with significantly higher temperatures and melt rates that are never experienced in the ANT2100 simulation, this is not surprising. Even a neural network trained with the last 80 years of data from the ANT2300 simulation does not outperform a neural network trained on just the first 20 years of data. 330 We therefore emphasise the need for training with appropriate data that are matched to the target simulation.

When we consider neural networks trained with only data from the beginning of the ANT2300 simulation, we see that increasing the amount of data included in the training improves the neural network performance up until about 40 years of data is used ($\text{RMSE} = 99 \text{ Gt yr}^{-1}$). After this point, the RMSE of the simulation does not improve significantly. However, we note that the RMSE of neural networks trained with just 10 years of data is still only 157 Gt yr^{-1} , around 10% of the total sub-shelf 335 melt in this simulation. Although training with more data is better, we suggest that a minimum of at least 10 years of relevant simulation data is used in training when applying the neural network to new conditions.

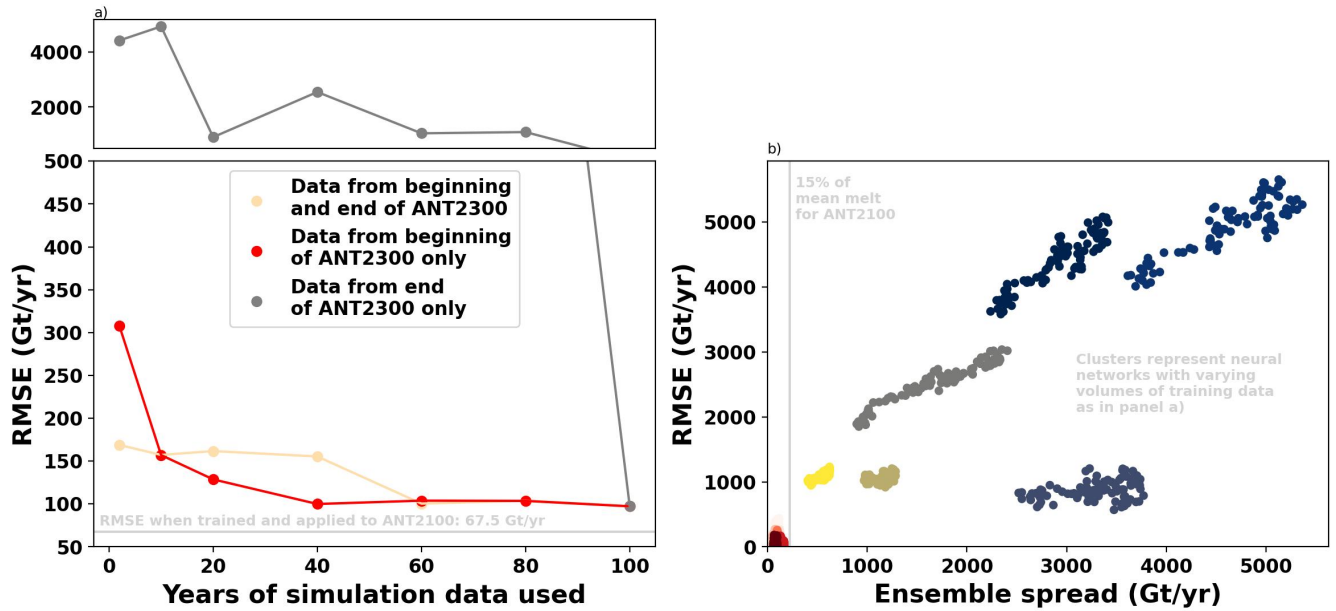


Figure 8. Neural network performance (RMSE) as a function of training data quantity, and ensemble spread. a) The RMSE of the integrated reference melt from the ANT2100 simulation when compared to the predicted melt from neural networks trained with different volumes of data from the ANT2300 simulation. Results are shown when the selected data is either from the beginning of the ANT2300 simulation (red), both the beginning and end of the ANT2300 simulation (light orange), or the end of the ANT2300 simulation (gray), note that the break in the y axis at 500 Gt yr⁻¹, and that it resumes with different scale). The RMSE of the predicted melt from the neural network trained on the ANT2100 simulation and applied to the ANT2100 simulation is also shown in light grey (67.5 Gt yr⁻¹) as a comparison. b) The RMSE of the integrated reference melt when compared to the ensemble spread (the standard deviation of the 10 neural networks), for results grouped by year. Each dot represents one year of data, and the coloured clusters represent are different simulations.

4 Discussion

The neural network presented here produces melt rates that are closely correlated to the melt rates of the NEMO ocean model. This neural network architecture opens the door to a hybrid approach, where melt rates in some parts of the ice shelf are resolved with the full dynamics and thermodynamics of the ocean model, and melt rates in the remaining regions parts of these ice shelves or the entire cavity of small ice shelves can be calculated using the neural network. The output of the neural network is a map of sub-shelf melt rates, which can be integrated to give a total melt flux for use in an ocean model, or combined with melt rates from the ocean-model resolved region and fed directly into an ice sheet model. The neural network emulator can therefore be used in both stand-alone ocean simulations, or in coupled ocean-ice-sheet simulations. If a low-resolution ocean model can-

Low resolution ocean models have started to be coupled to a higher-resolution ice-sheet model, higher-resolution ice sheet models (?), but with unrealistic or no representation of melting in the smallest cavities and near the grounding line, which is

thought to alter the simulated ice sheet dynamics. With coupled models that include the smallest cavities and the grounding zone, it will become possible to run long time-period simulations that include feedbacks between the ocean and the ice-sheet, such as evolving pinning points, changes in ocean stratification, and sea-ice production (???). This will be a significant improvement on models where the ice-ocean boundary is kept static and such feedbacks are neglected, especially as coupled regional models currently do not have a consensus on whether sub-shelf melt provides a positive or negative feedback to climate forcing (??).

We provide as input variables for the neural network point-wise information about the cavity geometry and water properties as close to the cavity as possible. These input variables were chosen because they represent physical quantities which are known to impact sub-shelf melt rates and have been included in other parameterisations (???), including the whole ice-shelf neural network parameterisation of ? which this work builds on. However, there are also additional variables that could have an impact on the sub-shelf melt rates, which are not included here. In particular, ~~we only consider the vertical circulation within an ice-shelf segment, through the ice draft and the distances to the ocean and grounding line. We do not have an input~~
360 ~~variable that accounts~~ our NN considers local ocean conditions, and does not account for the horizontal circulation ~~within an ice-shelf~~ which should connect temperature and salinity properties between grid points, despite the ~~key role that importance of~~ topography and Coriolis forcing ~~play~~ in steering current through the ice-shelf cavity (??). Within the point-wise neural network architecture, we suggest that in future work we could add information about the relative location of the grid point within the cavity or ice-shelf - possibly as simple as a -1 to 1 scale for left and right within the cavity. This may allow the neural network
365 optimisation process to identify simple spatial biases in melt rate if they exist within the training data.

While considering the best neural network set-up, we considered other architectures alongside the point-wise perceptron employed here. Graphical neural networks are designed to accept sparse data on variable domains, and if employed here might have identified different relationships in the input data. However, in order to have good performance with a graphical neural network it is important to design the graph embedding structure correctly. The sub-shelf melt data does not have an obvious
370 graph structure, and so it was unclear whether the additional complexity and expense of a graphical neural network would lead to any performance gains. Convolutional neural networks (CNNs) are widely used to identify patterns in image data, through the application of convolutional filters that consider relationships across multiple pixels. When applied to the parametrisation of sub-shelf melt, they could be used to identify drivers of melt at a wider spatial scale than is currently included. However, CNNs work best with a fixed input image size, and are less flexible for variable ice-shelf sizes (from 1 to 1368 grid points here)
375 and resolutions than the point-wise architecture which we employ. For example, ? used a modified U-Net CNN architecture (ResUNet++, ?) as part of a two-step approach to predicting sub-shelf melt rates. However, their neural network was only trained on synthetic ice-shelves with a fixed image size, and did not demonstrate the adaptability of the point-wise neural network when it comes to applying to a range of real Antarctic ice-shelf geometries on both regular and irregular grids.

A potentially promising approach in the future may be to employ physics informed neural networks. These neural networks
380 use physical laws to constrain neural network training, and have been applied to a wide range of problems in glaciology including studying the flow law of ice-shelves, and inverting for basal friction (??). When applied to the parameterisation of sub-shelf melt, similar thermodynamical constraints to the NEMO model could be used to maintain consistency and act as an

additional regularisation factor. However, the physical equations used in NEMO and somewhat learned by the neural network should ensure some physical consistency in our approach.

385 It is important to note that the neural network emulator will only ever be as good as the NEMO model which is used to train it. Biases in the ocean model will be passed on in training and may be reproduced in the emulator, as we have seen with George VI ice shelf (?). Additionally, since we train with the NEMO 0.25° model as our "~~true~~reference" melt rates, any processes that are neglected at this resolution in NEMO ~~, such as small-scale eddies, or will also not be included in the emulator. This includes processes such as small-scale eddies,~~ tidal fluctuations in the grounding line(~~?~~), ~~will also not be~~
390 ~~included in the emulator,~~ melting in basal crevasses and boundary layer turbulence, which are known to influence melt rates beneath Antarctic Ice shelves from observations (???) or model simulations (??). However, since NEMO melt rates are tuned to be close to estimates from remote sensing, and temperatures and salinity are close to in-situ observations (?), some of these missing processes are implicitly accounted for in NEMO's parameter values. Additionally, one of the strengths of the emulator is that, should there be ~~an~~ a significant update in the model physics which directly influences melt rates, relatively few
395 higher-resolution simulations are required to retrain the emulator, which can then be rapidly employed in the ~~low-resolution~~ low-resolution model. This approach can therefore save valuable computation resources while still facilitating the running of large ensemble coupled ice-sheet ocean models with realistic physics.

5 A neural network for application in the NEMO ocean model

We show here that a neural network can be used to emulate sub-shelf melt rates in the interior regions of ice shelves which
400 are not resolved in the 1° NEMO ocean model, and that the neural network emulator performs best when trained in similar conditions to those which it is applied. In order to produce the most useful neural network for application in the 1° NEMO ocean model, we therefore additionally trained a neural network with all the available 0.25° simulations. We note that as we have trained this neural network with a wide range of temperature conditions, we would expect it to produce realistic melt rates for conditions similar to the present day and up to 2300. However, as we have trained limited variation in geometries and
405 grounding line positions (2 x 155 basins), we would expect good performance in a model with a fixed grounding line, and would be more cautious about performance in an ocean model with evolving ice and bed geometries. To improve the generalisation to changing geometries, we recommend adding simulations with different geometries and grounding line positions when available, and especially a larger range of ice-shelf slopes. Another approach could be to use synthetic ice-shelf geometries such as those used by ? and upcoming high-resolution ocean-ice-sheet coupled simulations (e.g. ?).

410 The resulting neural network is available in the package that accompanies this manuscript. As this emulator was trained with all the available simulations, there is no independent test data to verify the performance of this neural network. To provide an idea of the neural network performance we instead use the validation data set, which is used to tune the hyper-parameters used for training. When applied to the validation data set (11% of the data points used in training), we report an RMSE of 3.29 m yr⁻¹ (ice equivalent), and a correlation coefficient R² of 0.93. As the validation dataset is selected randomly,
415 metrics grouped by basin or year are uninformative. Based on these metrics, we expect this emulator to perform well when

applied under conditions similar to the conditions in the training datasets. However, it will be necessary to do further model validation once the emulator is implemented in the 1° model to assess if it continues to reproduce the melt rates observed in the high-resolution model, as has been done in other studies including for parameterisations of convection (?) and vertical mixing (?).

420 6 Conclusions

We demonstrate that sub-shelf melt rates within the interior segments of ice shelves not resolved in 1° global climate models can be emulated using a simple multilayer perceptron, and present a trained neural network which is ready for use in stand-alone ocean models or coupled ice-sheet-ocean models. This trained neural network leverages all the available continental 0.25° NEMO simulations to cover the maximum possible range of ice shelf conditions. We also explore the performance of neural
425 networks trained with only selected simulations to understand how well the neural network generalises to new conditions. We show that these neural networks are quite sensitive to the range of input parameters, especially if applied to temperatures and ice-draft slopes outside those used in the training dataset. However, we find that the spread of the ensemble neural network results is a good indicator of the match between the training data and the conditions which the neural networks are applied in. If the ensemble spread is high relative to the total integrated melt (we suggest a threshold of >10%), then this indicates
430 that the training data was not suitable for this application, and ideally then additional simulation data would be provided. Our results show that as little as 10 years of targeted simulation data can improve the neural network performance. This manuscript is accompanied by python scripts that allow additional data to be added during neural network training. When trained with adequate simulations and employed in appropriate conditions, which can be verified with a suitable threshold on ensemble spread, the trained neural network presented here has the potential to significantly improve the representation of sub-shelf
435 cavities in global ocean and climate models.

Code and data availability. The code which accompanies this manuscript is archived on Zenodo at <https://doi.org/10.5281/zenodo.17358195>, and is also available on GitHub (?).

The data which accompanies this manuscript is available on Zenodo at <https://doi.org/10.5281/zenodo.17358228>, (?).

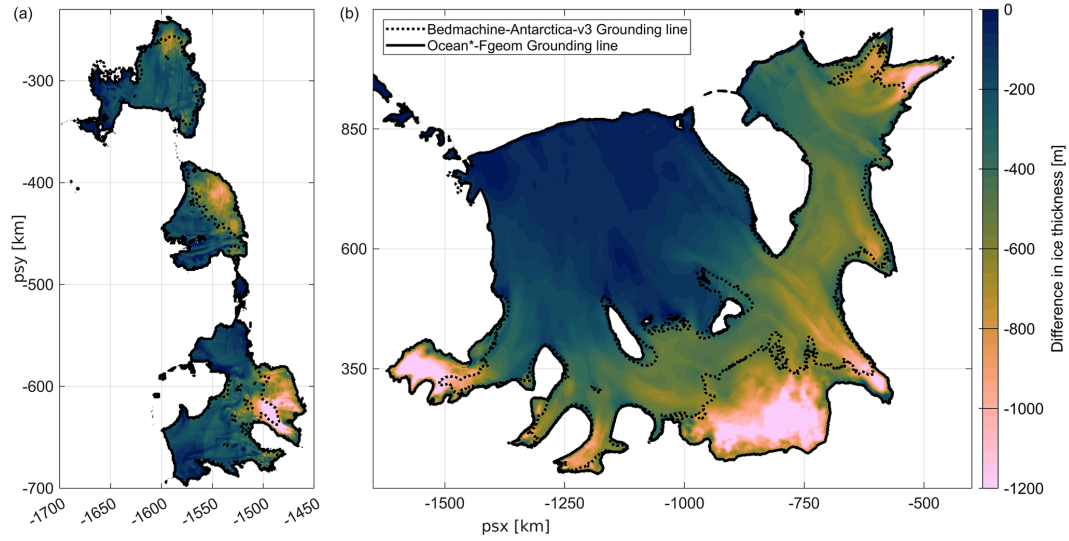


Figure A1. The difference in geometries between the ANT2000 simulation and the ANTFGEOM simulation. The ANT2000 simulation uses the Bedmachine v2 (?) geometry, and the ANTFGEOM simulation is mostly identical but has ice shelf draft is artificially thinned in the Amundsen Sea (a) and the Filchner-Ronne Ice Shelf (b) regions following the MISOMIP2 protocol (?). Figure reproduced from ?.

Appendix A

440 A1 Differences between simulations of ice-ocean conditions

Figure A1 shows the difference in geometries between the ANT2000 simulation and the ANTFGEOM simulation. The ANT2000 simulation uses the Bedmachine v2 (?) geometry, and the ANTFGEOM simulation is mostly identical but has ice shelf draft is artificially thinned in the Amundsen Sea and Filchner-Ronne Ice Shelf regions following the MISOMIP2 protocol (?).

445 A2 Outlines of ice-shelf basins

In order to have a flexible framework which can adapt to evolving shelf geometries and changing grounding line positions, we divide the coastline of Antarctica into 155 partial ice-shelf basins, based on the locations of the individual outlet glaciers and ice streams (Figure A3, ?). Large pinning points (ice rumples and rises) are also separated and form their own basins. These basins are extrapolated into the open ocean using a nearest neighbour approach, and occasionally adjusted manually to ensure
 450 that basins which contain an ice-shelf front always have a few points in the open ocean. The resulting partial ice-shelf basins all have similar dimensions, with the exception of basins which include islands and are not part of the main coastline. The margins of the largest ice-shelves fall into several of these partial ice-shelf basins.

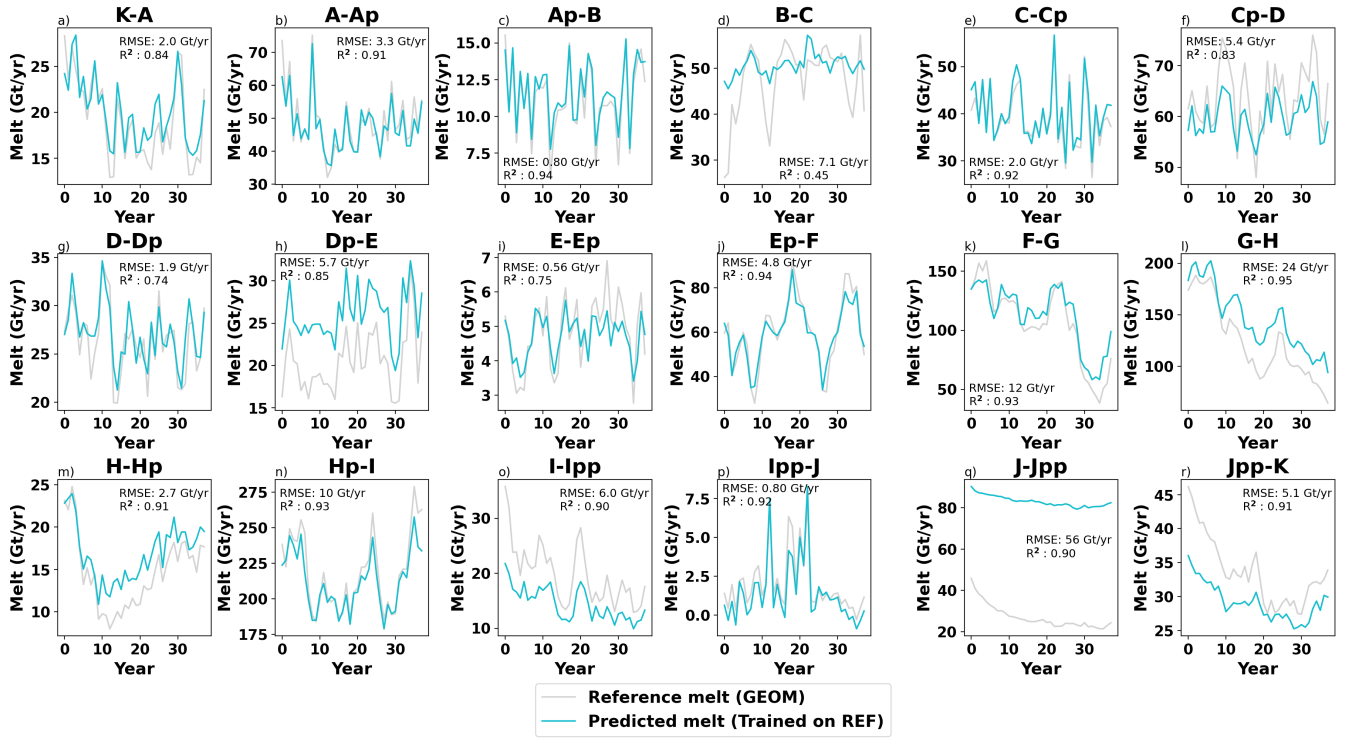


Figure A2. Neural network performance in a different geometry. The reference melt for the ANTGEOM simulation (grey) and predicted melt for a neural network trained on the ANT2000 simulation (blue) for each of the 18 IMBIE basins (a-r). The root-mean-square error and correlation coefficient (R^2) are given for each basin.

A3 Training with annual or monthly data?

Where available, we train the neural networks presented in the main body of the text with monthly data from each of the simulations. Neural network performance is improved by using monthly data, as shown in Figure A4.

In order to assess the usefulness of training with monthly or annual data, we trained neural networks with data from the ANT2000, ANT2100 and ANT2300 simulations, with both annual and monthly data. Firstly, we consider the predicted melt rates from the ANT2000 and ANT2300 neural networks when applied to the ANT2100 simulation, which has intermediate temperature conditions slightly warmer than those in the ANT2000 simulation and cooler than those in the ANT2300 simulation. We expect that a neural network trained using monthly data should optimise the weights based on a wider range of temperature conditions, and would therefore return better results when applied to temperature conditions outside those in which it was trained. Indeed, our results show strongly reduced RMSE and higher correlation coefficients for neural networks trained using monthly data. For neural networks trained on the ANT2000 simulation, we see the RMSE fall from 200 Gt yr⁻¹ to 110 Gt yr⁻¹ when monthly data is used instead of annual. For neural networks trained on the ANT2300 simulation, RMSE decreases from 130 Gt yr⁻¹ to 97 Gt yr⁻¹. Therefore, we suggest that in order to maximise the training potential of a lim-

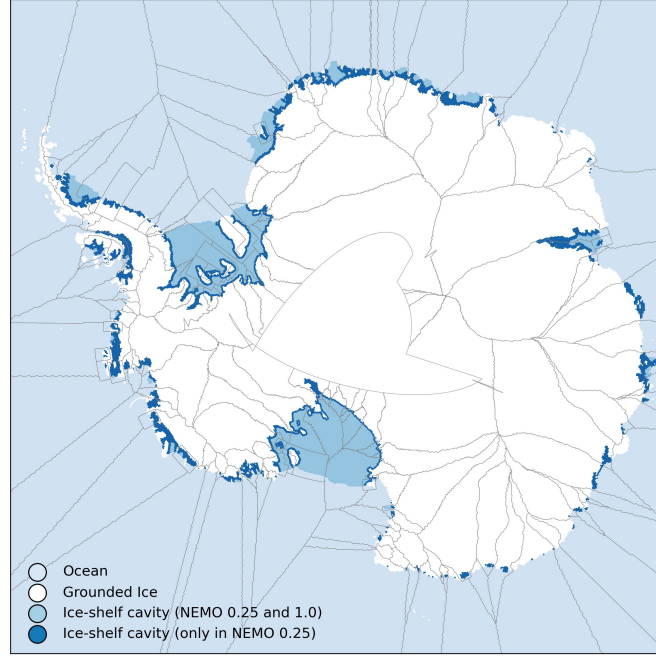


Figure A3. The outlines of the 155 partial ice-shelf basins, based on the locations of the individual outlet glaciers and ice streams (?), are shown by the light grey lines.

ited number of simulations, neural networks emulators of sub-ice shelf melt should be trained using monthly data wherever possible.

A4 Full metrics for simulations presented here

In addition to the metrics discussed in the main text, we report more comprehensive metrics for the main simulations presented
 470 here. In each case, a trained neural network is applied to the data from one of the NEMO simulations. We report the root-mean-squared error (RMSE) and correlation coefficient (R^2) of the reference melt rate from the simulation compared to the melt rate predicted using the neural network. RMSE and R^2 are given for both the pointwise melt in m yr^{-1} (ice equivalent) and for the integrated melt in Gt yr^{-1} . Integrated melt rates are given for the entire Antarctic continent, but also when melt is summed by basin.

475 Figure A5 shows metrics for the model validation experiments, when neural networks are applied to the same data that were used to train them. Figure A6 shows metrics for the testing experiments, when neural networks are applied to simulations which were not used in training to explore the performance of the neural networks in novel conditions. Finally, Figure A7 shows metrics for the neural network which was trained on all the available simulations, when applied to each of these simulations.

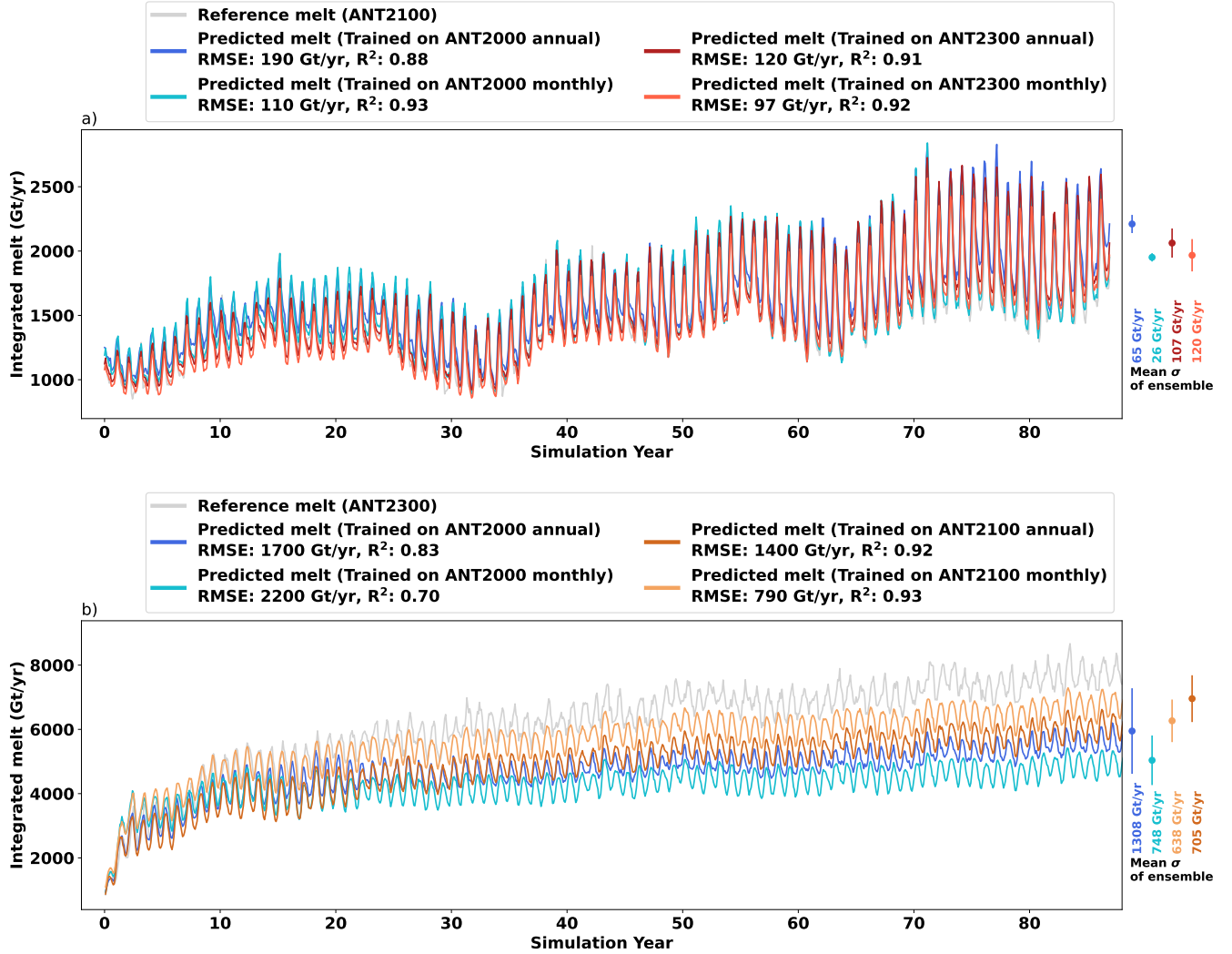


Figure A4. Should neural networks be trained with monthly or annual data? The results of applying neural networks trained on both annual and monthly data from ANT2000 and ANT2300 to the ANT2100 simulation, split into yearly (a) and monthly (b) trends. The reference melt rate for each simulation is show in grey. For all neural network predictions, the root-mean-square error (RMSE) in Gt yr^{-1} and correlation coefficient (R^2) are given in the legend.

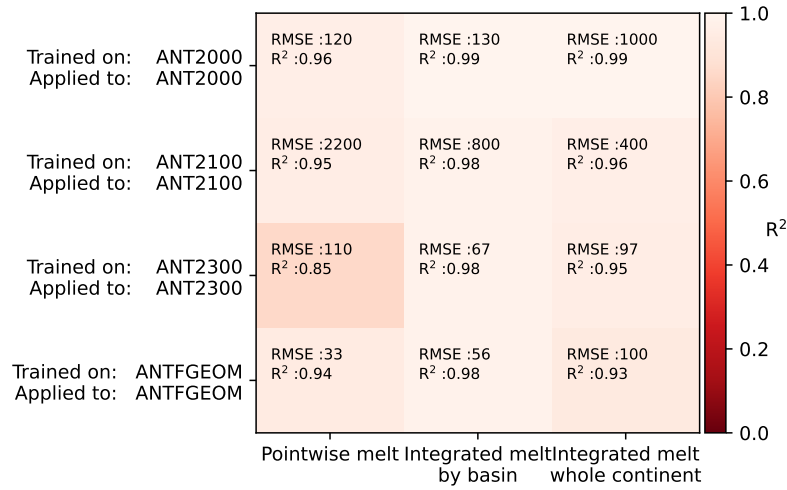


Figure A5. Metrics for the trained neural networks used in model validation. In each case, a trained neural network is applied to one of the simulations. We report the root-mean-squared error (RMSE) and correlation coefficient (R^2) of the reference melt rate from the simulation compared to the melt rate predicted using the neural network, for the pointwise melt in m yr^{-1} (ice equivalent) and for the integrated melt in Gt yr^{-1} (integrated by basin and over the whole continent).

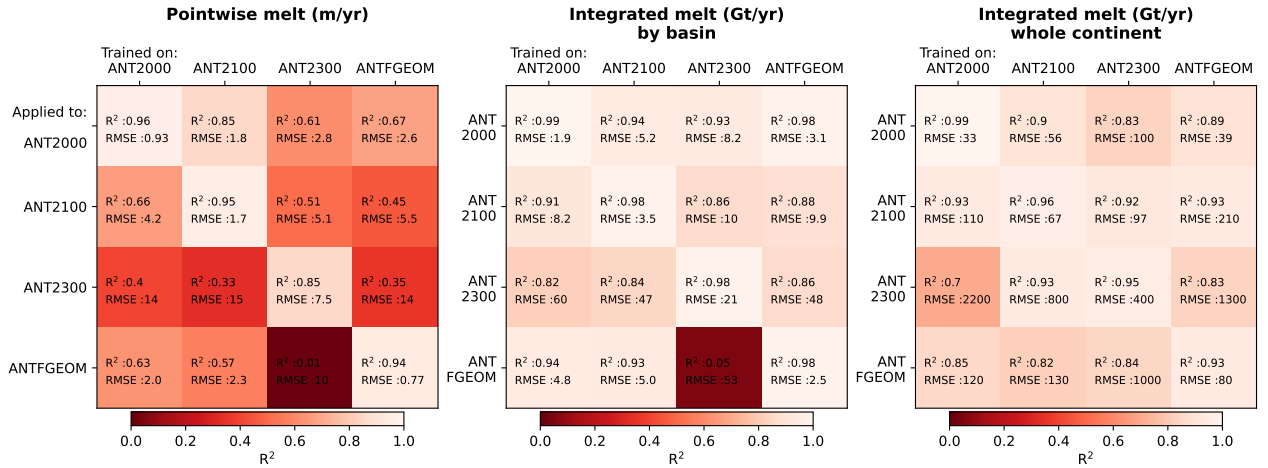


Figure A6. Metrics for the trained neural networks used in testing neural network performance. In each case, a trained neural network is applied to one of the simulations. We report the root-mean-squared error (RMSE) and correlation coefficient (R^2) of the reference melt rate from the simulation compared to the melt rate predicted using the neural network, for the pointwise melt in m yr^{-1} (ice equivalent) and for the integrated melt in Gt yr^{-1} (integrated by basin and over the whole continent).

Neural network trained on ALL simulations

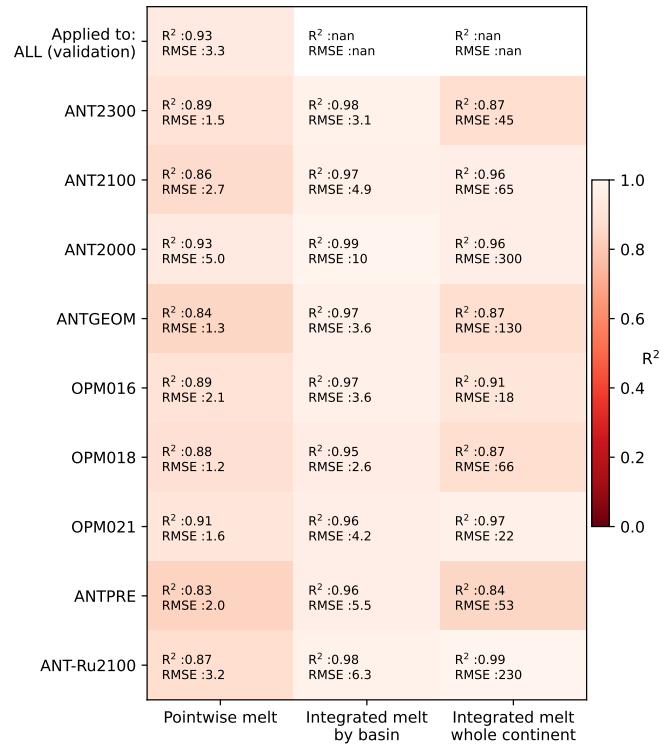


Figure A7. Metrics for neural networks trained using all available NEMO simulations. The neural network is trained with all available simulations, and then applied to each one, as well as to the validation data set from the training. We report the root-mean-squared error (RMSE) and correlation coefficient (R^2) of the reference melt rate from the simulation compared to the melt rate predicted using the neural network, for the pointwise melt in m yr^{-1} (ice equivalent) and for the integrated melt in Gt yr^{-1} (integrated by basin and over the whole continent).

Author contributions. HO, CB, and NCJ developed the original idea of this paper. PM and CK performed the NEMO simulations used to train the neural network. HO carried out the analysis, and wrote the majority of the manuscript. All co-authors contributed to discussions, and to the editing of the manuscript.

Competing interests. The authors declare that they have no competing interests.

Acknowledgements. This work was supported by the French National Research Agency through the AIAI project (ANR-22-CE01-0014), and by the French Government through the France 2030 program managed by ANR (ISCLim, ANR-22-EXTR-0010, and IMPRESSION-ESM, ANR-22-EXTR-0008). C. Kittel is a postdoc fellow of the Fonds de la Recherche Scientifique de Belgique F.R.S–FNRS.

References

- Asay-Davis, X. S., Cornford, S. L., Durand, G., Galton-Fenzi, B. K., Gladstone, R. M., Gudmundsson, G. H., Hattermann, T., Holland, D. M., Holland, D., Holland, P. R., et al.: Experimental design for three interrelated marine ice sheet and ocean model intercomparison projects: MISMIP v. 3 (MISMIP+), ISOMIP v. 2 (ISOMIP+) and MISOMIP v. 1 (MISOMIP1), *Geoscientific Model Development*, 9, 2471–2497, <https://doi.org/10.5194/gmd-9-2471-2016>, 2016.
- Beckmann, A., Hellmer, H. H., and Timmermann, R.: A numerical model of the Weddell Sea: Large-scale circulation and water mass distribution, *Journal of Geophysical Research: Oceans*, 104, 23 375–23 391, <https://doi.org/10.1029/1999JC900194>, 1999.
- Beckmann, J., Reese, R., McCormack, F. S., Cook, S., Bird, L., Gwyther, D., Richards, D., Scheiter, M., Wang, Y., Seroussi, H., et al.: Disentangling uncertainty in ISMIP6 Antarctic sub-shelf melting and 2300 sea level rise projections, *EGU sphere*, 2025, 1–45, <https://doi.org/10.5194/egusphere-2025-4069>, 2025.
- Begeman, C. B., Tulaczyk, S., Padman, L., King, M., Siegfried, M. R., Hodson, T. O., and Fricker, H. A.: Tidal pressurization of the ocean cavity near an Antarctic ice shelf grounding line, *Journal of Geophysical Research: Oceans*, 125, e2019JC015 562, <https://doi.org/10.1029/2019JC015562>, 2020.
- Bett, D. T., Bradley, A. T., Williams, C. R., Holland, P. R., Arthern, R. J., and Goldberg, D. N.: Coupled ice–ocean interactions during future retreat of West Antarctic ice streams in the Amundsen Sea sector, *The Cryosphere*, 18, 2653–2675, <https://doi.org/10.5194/tc-18-2653-2024>, 2024.
- Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., et al.: Presentation and evaluation of the IPSL-CM6A-LR climate model, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002 010, <https://doi.org/10.1029/2019MS002010>, 2020.
- Burgard, C., Jourdain, N. C., Reese, R., Jenkins, A., and Mathiot, P.: An assessment of basal melt parameterisations for Antarctic ice shelves, *The Cryosphere Discussions*, 2022, 1–56, <https://doi.org/10.5194/tc-16-4931-2022>, 2022.
- Burgard, C., Jourdain, N. C., Mathiot, P., Smith, R. S., Schäfer, R., Caillet, J., Finn, T. S., and Johnson, J. E.: Emulating present and future simulations of melt rates at the base of Antarctic ice shelves with neural networks, *Journal of Advances in Modeling Earth Systems*, 15, e2023MS003 829, <https://doi.org/10.1029/2023MS003829>, 2023.
- Burgard, C., Jourdain, N. C., Mosbeux, C., Caillet, J., Mathiot, P., and Kittel, C.: Ocean warming threatens the viability of 60% of Antarctic ice shelves, *Nature*, 0, 0–0, 2025.
- Chen, J.-J., Swart, N. C., Beadling, R., Cheng, X., Hattermann, T., Jüling, A., Li, Q., Marshall, J., Martin, T., Muilwijk, M., et al.: Reduced deep convection and bottom water formation due to Antarctic meltwater in a multi-model ensemble, *Geophysical research letters*, 50, e2023GL106 492, 2023.
- Cheng, G., Krishna, M., and Morlighem, M.: A Python library for solving ice sheet modeling problems using physics-informed neural networks, *PINNACLE v1. 0*, *Geoscientific Model Development*, 18, 5311–5327, <https://doi.org/10.5194/gmd-18-5311-2025>, 2025.
- Chollet, F.: keras, <https://keras.io>, 2015.
- Davis, P. E. and Nicholls, K. W.: Turbulence observations beneath Larsen C ice shelf, Antarctica, *Journal of Geophysical Research: Oceans*, 124, 5529–5550, <https://doi.org/10.1029/2019JC015164>, 2019.
- Davison, B. J., Hogg, A. E., Gourmelen, N., Jakob, L., Wuite, J., Nagler, T., Greene, C. A., Andreasen, J., and Engdahl, M. E.: Annual mass budget of Antarctic ice shelves from 1997 to 2021, *Science Advances*, 9, eadi0186, <https://doi.org/10.1126/sciadv.adi0186>, 2023.

- De Rydt, J., Jourdain, N. C., Nakayama, Y., Van Caspel, M., Timmermann, R., Mathiot, P., Asay-Davis, X. S., Seroussi, H., Dutrieux, P., Galton-Fenzi, B., et al.: Experimental design for the Marine Ice Sheet–Ocean Model Intercomparison Project–phase 2 (MISOMIP2), *Geoscientific Model Development*, 17, 7105–7139, <https://doi.org/10.5194/gmd-17-7105-2024>, 2024.
- 525 DeConto, R. M. and Pollard, D.: Contribution of Antarctica to past and future sea-level rise, *Nature*, 531, 591–597, <https://doi.org/10.1038/nature17145>, 2016.
- DeVries, T.: The oceanic anthropogenic CO₂ sink: Storage, air-sea fluxes, and transports over the industrial era, *Global Biogeochemical Cycles*, 28, 631–647, <https://doi.org/10.1002/2013GB004739>, 2014.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geoscientific Model Development*, 9, 1937–1958, <https://doi.org/10.5194/gmd-9-1937-2016>, 2016.
- 530 Fukushima, K.: Cognitron: A self-organizing multilayered neural network, *Biological cybernetics*, 20, 121–136, <https://doi.org/10.1007/BF00342633>, 1975.
- Goosse, H., Kay, J. E., Armour, K. C., Bodas-Salcedo, A., Chepfer, H., Docquier, D., Jonko, A., Kushner, P. J., Lecomte, O., Massonnet, F., et al.: Quantifying climate feedbacks in polar regions, *Nature communications*, 9, 1919, <https://doi.org/10.1038/s41467-018-04173-0>, 2018.
- 535 Gruber, N., Landschützer, P., and Lovenduski, N. S.: The variable Southern Ocean carbon sink, *Annual review of marine science*, 11, 159–186, <https://doi.org/10.1146/annurev-marine-121916-063407>, 2019.
- Gwyther, D. E., Cougnon, E. A., Galton-Fenzi, B. K., Roberts, J. L., Hunter, J. R., and Dinniman, M. S.: Modelling the response of ice shelf basal melting to different ocean cavity environmental regimes, *Annals of Glaciology*, 57, 131–141, <https://doi.org/10.1017/aog.2016.31>, 2016.
- 540 Heuzé, C.: Antarctic bottom water and North Atlantic deep water in CMIP6 models, *Ocean Science Discussions*, 2020, 1–38, <https://doi.org/10.5194/os-17-59-2021>, 2020.
- Heuzé, C., Heywood, K. J., Stevens, D. P., and Ridley, J. K.: Southern Ocean bottom water characteristics in CMIP5 models, *Geophysical Research Letters*, 40, 1409–1414, 2013.
- 545 Holland, D. M. and Jenkins, A.: Modeling thermodynamic ice–ocean interactions at the base of an ice shelf, *Journal of physical oceanography*, 29, 1787–1800, [https://doi.org/10.1175/1520-0485\(1999\)029<1787:MTIOIA>2.0.CO;2](https://doi.org/10.1175/1520-0485(1999)029<1787:MTIOIA>2.0.CO;2), 1999.
- Holland, P. R., Jenkins, A., and Holland, D. M.: The response of ice shelf basal melting to variations in ocean temperature, *Journal of Climate*, 21, 2558–2572, <https://doi.org/10.1175/2007JCLI1909.1>, 2008.
- 550 Holt, J., Hyder, P., Ashworth, M., Harle, J., Hewitt, H. T., Liu, H., New, A. L., Pickles, S., Porter, A., Popova, E., et al.: Prospects for improving the representation of coastal and shelf seas in global ocean models, *Geoscientific Model Development*, 10, 499–523, 2017.
- Hutchinson, K., Deshayes, J., Éthé, C., Rousset, C., de Lavergne, C., Vancoppenolle, M., Jourdain, N. C., and Mathiot, P.: Improving Antarctic Bottom Water precursors in NEMO for climate applications, *Geoscientific Model Development*, 16, 3629–3650, <https://doi.org/10.5194/gmd-16-3629-2023>, 2023.
- 555 Iwasaki, Y. and Lai, C.-Y.: One-dimensional ice shelf hardness inversion: Clustering behavior and collocation resampling in physics-informed neural networks, *Journal of Computational Physics*, 492, 112 435, <https://doi.org/10.1016/j.jcp.2023.112435>, 2023.
- Jha, D., Smedsrud, P. H., Riegler, M. A., Johansen, D., De Lange, T., Halvorsen, P., and Johansen, H. D.: Resunet++: An advanced architecture for medical image segmentation, in: 2019 IEEE international symposium on multimedia (ISM), pp. 225–2255, IEEE, <https://doi.org/10.1109/ISM46123.2019.00049>, 2019.

- 560 Jourdain, N. C., Asay-Davis, X., Hattermann, T., Straneo, F., Seroussi, H., Little, C. M., and Nowicki, S.: A protocol for calculating basal melt rates in the ISMIP6 Antarctic ice sheet projections, *The Cryosphere*, 14, 3111–3134, <https://doi.org/10.5194/tc-14-3111-2020>, 2020.
- Jourdain, N. C., Mathiot, P., Burgard, C., Caillet, J., and Kittel, C.: Ice shelf basal melt rates in the Amundsen Sea at the end of the 21st century, *Geophysical Research Letters*, 49, e2022GL100629, <https://doi.org/10.1029/2022GL100629>, 2022.
- Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, <https://arxiv.org/abs/1412.6980>, 2017.
- 565 Kittel, C., Amory, C., Agosta, C., Jourdain, N. C., Hofer, S., Delhasse, A., Doutreloup, S., Huot, P.-V., Lang, C., Fichet, T., et al.: Diverging future surface mass balance between the Antarctic ice shelves and grounded ice sheet, *The Cryosphere*, 15, 1215–1236, <https://doi.org/10.5194/tc-15-1215-2021>, 2021.
- Kreuzer, M., Reese, R., Huiskamp, W. N., Petri, S., Albrecht, T., Feulner, G., and Winkelmann, R.: Coupling framework (1.0) for the PISM (1.1. 4) ice sheet model and the MOM5 (5.1. 0) ocean model via the PICO ice shelf cavity model in an Antarctic domain, *Geoscientific*
- 570 *Model Development*, 14, 3697–3714, <https://doi.org/10.5194/gmd-14-3697-2021>, 2021.
- Lambert, E. and Burgard, C.: Brief communication: Sensitivity of Antarctic ice shelf melting to ocean warming across basal melt models, *The Cryosphere*, 19, 2495–2505, <https://doi.org/10.5194/tc-19-2495-2025>, 2025.
- Lambert, E., Jüling, A., van de Wal, R. S. W., and Holland, P. R.: Modelling Antarctic ice shelf basal melt patterns using the one-layer Antarctic model for dynamical downscaling of ice–ocean exchanges (LADDIE v1.0), *The Cryosphere*, 17, 3203–3228, [https://doi.org/10.5194/tc-](https://doi.org/10.5194/tc-17-3203-2023)
- 575 [17-3203-2023](https://doi.org/10.5194/tc-17-3203-2023), 2023.
- Li, D., DeConto, R. M., Pollard, D., and Hu, Y.: Competing climate feedbacks of ice sheet freshwater discharge in a warming world, *Nature Communications*, 15, 5178, <https://doi.org/10.1038/s41467-024-49604-3>, 2024.
- Little, C. M., Gnanadesikan, A., and Oppenheimer, M.: How ice shelf morphology controls basal melting, *Journal of Geophysical Research: Oceans*, 114, <https://doi.org/10.1029/2008JC005197>, 2009.
- 580 Lurton, T., Balkanski, Y., Bastrikov, V., Bekki, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Contoux, C., Cozic, A., et al.: Implementation of the CMIP6 forcing data in the IPSL-CM6A-LR model, *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001940, <https://doi.org/10.1029/2019MS001940>, 2020.
- Madec, G. and the NEMO System Team: NEMO Ocean Engine Reference Manual, <https://doi.org/10.5281/zenodo.1464816>, 2024.
- Mathiot, P. and Jourdain, N. C.: Southern Ocean warming and Antarctic ice shelf melting in conditions plausible by late 23rd century in a high-end scenario, *Ocean Science*, 19, 1595–1615, <https://doi.org/10.5194/os-19-1595-2023>, 2023.
- 585 Mathiot, P., Jenkins, A., Harris, C., and Madec, G.: Explicit representation and parametrised impacts of under ice shelf seas in the z_* coordinate ocean model NEMO 3.6, *Geoscientific Model Development*, 10, 2849–2874, <https://doi.org/10.5194/gmd-10-2849-2017>, 2017.
- Merino, N., Jourdain, N. C., Le Sommer, J., Goosse, H., Mathiot, P., and Durand, G.: Impact of increasing antarctic glacial freshwater release on regional sea-ice cover in the Southern Ocean, *Ocean Modelling*, 121, 76–89, <https://doi.org/10.1016/j.ocemod.2017.11.009>, 2018.
- 590 Morlighem, M., Rignot, E., Binder, T., Blankenship, D., Drews, R., Eagles, G., Eisen, O., Ferraccioli, F., Forsberg, R., Fretwell, P., et al.: Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the Antarctic ice sheet, *Nature geoscience*, 13, 132–137, <https://doi.org/10.1038/s41561-019-0510-8>, 2020.
- Mouginot, J., Scheuchl, B., and Rignot, E.: MEaSUREs Antarctic Boundaries for IPY 2007–2009 from Satellite Radar, Version2, [data set], NASA National Snow and Ice Data Center, <https://doi.org/10.5067/AXE4121732AD>, 2017.
- 595 Naughten, K. A., Holland, P. R., and De Rydt, J.: Unavoidable future increase in West Antarctic ice-shelf melting over the twenty-first century, *Nature Climate Change*, 13, 1222–1228, <https://doi.org/10.1038/s41558-023-01818-x>, 2023.

- Ockenden, H.: Code which accompanies the manuscript "A neural network emulator of ice-shelf melt rates for use in ocean models which partially resolve ice-shelf cavities", Zenodo, <https://doi.org/10.5281/zenodo.17358195>, 2025a.
- Ockenden, H.: Data which accompanies the manuscript "A neural network emulator of ice-shelf melt rates for use in ocean models which partially resolve ice-shelf cavities", Zenodo, <https://doi.org/10.5281/zenodo.17358228>, 2025b.
- 600 Olivé Abelló, A., Mathiot, P., Jourdain, N. C., Kostov, Y., Holland, P. R., Gascoin, S., and Rousset, C.: Iceberg grounding enhances the release of freshwater on the Antarctic continental shelf, *Journal of Geophysical Research: Oceans*, 130, e2025JC022857, <https://doi.org/10.1029/2025JC022857>, 2025.
- Patmore, R. D., Holland, P. R., Vreugdenhil, C. A., Jenkins, A., and Taylor, J. R.: Turbulence in the ice shelf-ocean boundary current and its sensitivity to model resolution, *Journal of Physical Oceanography*, 53, 613–633, <https://doi.org/10.1175/JPO-D-22-0034.1>, 2023.
- Pelle, T., Morlighem, M., and Bondzio, J. H.: Brief communication: PICOP, a new ocean melt parameterization under ice shelves combining PICO and a plume model, *The Cryosphere*, 13, 1043–1049, <https://doi.org/10.5194/tc-13-1043-2019>, 2019.
- Pelle, T., Morlighem, M., Nakayama, Y., and Seroussi, H.: Widespread grounding line retreat of Totten Glacier, East Antarctica, over the 21st century, *Geophysical Research Letters*, 48, e2021GL093213, <https://doi.org/10.1029/2021GL093213>, 2021.
- 610 Pirlet, N., Fichet, T., Vancoppenolle, M., Fraser, A. D., Mathiot, P., Rousset, C., Barthélemy, A., Barriat, P.-Y., Pelletier, C., Madec, G., et al.: Benefits of a landfast ice representation on simulated Antarctic sea ice and coastal polynya dynamics, *Journal of Geophysical Research: Oceans*, 130, e2024JC022032, 2025.
- Pritchard, H., Ligtenberg, S. R., Fricker, H. A., Vaughan, D. G., van den Broeke, M. R., and Padman, L.: Antarctic ice-sheet loss driven by basal melting of ice shelves, *Nature*, 484, 502–505, <https://doi.org/10.1038/nature10968>, 2012.
- 615 Ramachandran, P., Zoph, B., and Le, Q. V.: Searching for activation functions, arXiv preprint arXiv:1710.05941, 2017.
- Richter, O., Gwyther, D. E., King, M. A., and Galton-Fenzi, B. K.: The impact of tides on Antarctic ice shelf melting, *The Cryosphere*, 16, 1409–1429, <https://doi.org/10.5194/tc-16-1409-2022>, 2022.
- Rignot, E.: Ice-shelf changes in Pine Island Bay, Antarctica, 1947–2000, *Journal of Glaciology*, 48, 247–256, <https://doi.org/10.3189/172756502781831386>, 2002.
- 620 Rignot, E., Mouginot, J., Scheuchl, B., van den Broeke, M., van Wessem, M., and Morlighem, M.: Four decades of Antarctic Ice Sheet mass balance from 1979–2017, *Proc. Natl. Acad. Sci. U.S.A.*, 116, 1095–1103, <https://doi.org/10.1073/pnas.1812883116>, 2019.
- Rosier, S. H., Bull, C. Y., Woo, W. L., and Gudmundsson, G. H.: Predicting ocean-induced ice-shelf melt rates using deep learning, *The Cryosphere*, 17, 499–518, <https://doi.org/10.5194/tc-17-499-2023>, 2023.
- Sadai, S., Karmalkar, A. V., Pollard, D., Dong, Y., Lucas, E., Gomez, N., DeConto, R., and Condron, A.: Antarctic meltwater alters future projections of climate and sea level, *Nature Communications*, 16, 9271, <https://doi.org/10.1038/s41467-025-64438-3>, 2025.
- Schmidt, G. A., Mankoff, K. D., Bamber, J. L., Carroll, D., Chandler, D. M., Coulon, V., Davison, B. J., England, M. H., Holland, P. R., Jourdain, N. C., et al.: Datasets and protocols for including anomalous freshwater from melting ice sheets in climate simulations, *EGUsphere*, 2025, 1–40, 2025.
- Seroussi, H., Nakayama, Y., Larour, E., Menemenlis, D., Morlighem, M., Rignot, E., and Khazendar, A.: Continued retreat of Thwaites Glacier, West Antarctica, controlled by bed topography and ocean circulation, *Geophysical Research Letters*, 44, 6191–6199, <https://doi.org/10.1002/2017GL072910>, 2017.
- 630 Seroussi, H., Pelle, T., Lipscomb, W. H., Abe-Ouchi, A., Albrecht, T., Alvarez-Solas, J., Asay-Davis, X., Barre, J.-B., Berends, C. J., Bernales, J., et al.: Evolution of the Antarctic Ice Sheet over the next three centuries from an ISMIP6 model ensemble, *Earth's Future*, 12, e2024EF004561, <https://doi.org/10.1029/2024EF004561>, 2024.

- 635 Smith, R. S., Mathiot, P., Siahayan, A., Lee, V., Cornford, S. L., Gregory, J. M., Payne, A. J., Jenkins, A., Holland, P. R., Ridley, J. K., et al.: Coupling the UK Earth System Model to dynamic models of the Greenland and Antarctic ice sheets, *Journal of Advances in Modeling Earth Systems*, 13, e2021MS002520, <https://doi.org/10.1029/2021MS002520>, 2021.
- Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E. W., Calvert, D., Graham, T., Hewitt, H. T., Hyder, P., et al.: UK Global Ocean GO6 and GO7: A traceable hierarchy of model resolutions, *Geoscientific Model Development*, 11, 3187–3213, <https://doi.org/10.5194/gmd-11-3187-2018>, 2018.
- 640 Timmermann, R. and Goeller, S.: Response to Filchner–Ronne Ice Shelf cavity warming in a coupled ocean–ice sheet model–Part 1: The ocean perspective, *Ocean Science*, 13, 765–776, <https://doi.org/https://doi.org/10.5194/os-13-765-2017>, 2017.
- Timmermann, R. and Hellmer, H. H.: Southern Ocean warming and increased ice shelf basal melting in the twenty-first and twenty-second centuries based on coupled ice-ocean finite-element modelling, *Ocean Dynamics*, 63, 1011–1026, [https://doi.org/10.1007/s10236-013-](https://doi.org/10.1007/s10236-013-0642-0)
- 645 0642-0, 2013.
- Timmermann, R., Wang, Q., and Hellmer, H.: Ice-shelf basal melting in a global finite-element sea-ice/ice-shelf/ocean model, *Annals of Glaciology*, 53, 303–314, <https://doi.org/10.3189/2012AoG60A156>, 2012.
- Tsujino, H., Urakawa, S., Nakano, H., Small, R. J., Kim, W. M., Yeager, S. G., Danabasoglu, G., Suzuki, T., Bamber, J. L., Bentsen, M., et al.: JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do), *Ocean Modelling*, 130, 79–139, <https://doi.org/10.1016/j.ocemod.2018.07.002>, 2018.
- 650 Van Achter, G., Fichefet, T., Goosse, H., Pelletier, C., Haubner, K., and Pattyn, F.: Ocean–ice sheet coupling in the Totten Glacier area, East Antarctica: analysis of the feedbacks and their response to a sudden ocean warming, *Geosciences*, 13, 106, <https://doi.org/10.3390/geosciences13040106>, 2023.
- Voldoire, A., Saint-Martin, D., S  n  si, S., Decharme, B., Alias, A., Chevallier, M., Colin, J., Gu  r  my, J.-F., Michou, M., Moine, M.-P., et al.: Evaluation of CMIP6 deck experiments with CNRM-CM6-1, *Journal of Advances in Modeling Earth Systems*, 11, 2177–2213, <https://doi.org/10.1029/2019MS001683>, 2019.
- 655 Washam, P., Lawrence, J. D., Stevens, C. L., Hulbe, C. L., Horgan, H. J., Robinson, N. J., Stewart, C. L., Spears, A., Quartini, E., Hurwitz, B., et al.: Direct observations of melting, freezing, and ocean circulation in an ice shelf basal crevasse, *Science Advances*, 9, eadi7638, <https://doi.org/10.1126/sciadv.adi7638>, 2023.
- 660 Yung, C. K., Asay-Davis, X. S., Adcroft, A., Bull, C. Y., De Rydt, J., Dinniman, M. S., Galton-Fenzi, B. K., Goldberg, D., Gwyther, D. E., Hallberg, R., et al.: Results of the second Ice Shelf–Ocean Model Intercomparison Project (ISOMIP+), *EGUsphere*, 2025, 1–50, <https://doi.org/10.5194/egusphere-2025-1942>, 2025.
- Yuval, J. and O’Gorman, P. A.: Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions, *Nature communications*, 11, 3295, <https://doi.org/10.1038/s41467-020-17142-3>, 2020.
- 665 Zhu, Y., Zhang, R.-H., Moum, J. N., Wang, F., Li, X., and Li, D.: Physics-informed deep-learning parameterization of ocean vertical mixing improves climate simulations, *National Science Review*, 9, nwac044, <https://doi.org/10.1093/nsr/nwac044>, 2022.