

Emulating present and future simulations of melt rates at the base of Antarctic ice shelves with neural networks

C. Burgard¹, N.C. Jourdain¹, P. Mathiot¹, R.S. Smith², R. Schäfer³, J.
Caillet¹, T.S. Finn⁴, J.E. Johnson¹

¹Univ. Grenoble Alpes, IRD, CNRS, INRAE, Grenoble INP, IGE, 38000 Grenoble, France

²NCAS/Department of Meteorology, University of Reading, Reading, UK

³Physikalisch-Technische Bundesanstalt, Braunschweig, Germany

⁴CÉREA, École des Ponts and EDF R&D, Île-de-France, France

Key Points:

- We show that simple neural networks can produce reasonable basal melt rates by emulating circum-Antarctic cavity-opening ocean simulations.
- Predicted melt rates for present and warmer conditions are similar or closer to the reference simulation than traditional parameterisations.
- We show that neural networks are suited to be used as basal melt parameterisations for century-scale ice-sheet projections.

Corresponding author: Clara Burgard, clara.burgard@univ-grenoble-alpes.fr

Abstract

Melt rates at the base of Antarctic ice shelves are needed to drive projections of the Antarctic ice sheet mass loss. Current basal melt parameterisations struggle to link open ocean properties to ice-shelf basal melt rates for the range of current sub-shelf cavity geometries around Antarctica. We present a novel parameterisation based on deep learning. With a simple feedforward neural network, or multilayer perceptron, acting on each grid cell separately, we emulate the behavior of circum-Antarctic cavity-resolving ocean simulations. We explore different neural network sizes and find that, in all cases containing at least one hidden layer, this kind of emulator produces reasonable basal melt rates for our training ensemble, closer to the reference simulation than traditional parameterisations. For testing, we use an independent ensemble of simulations that was produced with the same ocean model but with different model parameters, different cavity geometries and different forcing. In this challenging test, traditional and neural network parameterisations yield similar results on present conditions. In much warmer conditions than the training ensemble, both traditional parameterisations and neural networks struggle, but the neural networks tend to produce basal melt rates closer to the reference than a majority of traditional parameterisations. These neural networks are therefore suitable for century-scale Antarctic ice-sheet projections.

Plain Language Summary

A warmer ocean around Antarctica leads to higher melting of the floating ice shelves, which influence the ice loss from the Antarctic ice sheet and therefore sea-level rise. In computer simulations of the ocean, these ice shelves are often not represented. For simulations of the ice sheet, so-called parameterisations are used to link the oceanic properties in front of the shelf and the melt at their base. We show that this link can be emulated with a simple neural network, which performs at least as well as traditional physical parameterisations both for present and much warmer conditions. This study also proposes several potential ways of further improving the use of deep learning to parameterise basal melt.

1 Introduction

The contribution of the Antarctic Ice Sheet to sea-level rise has been increasing in past decades and this increase is projected to continue with increasing greenhouse gas emissions (Fox-Kemper et al., 2021). Most of the mass loss is occurring at the margins of the ice sheet through faster ice flow from the grounded ice sheet to the ocean, mainly in West Antarctica (Mouginot et al., 2014; Rignot et al., 2014; Scheuchl et al., 2016; Khazendar et al., 2016; Shen et al., 2018; The IMBIE Team, 2018). This is because the floating ice shelves at the margins of the ice sheet, which usually buttress the ice flow, are rapidly thinning and retreating due to ocean-induced melt at their base (Rignot et al., 2013; Paolo et al., 2015; Adusumilli et al., 2020). In some bedrock configurations, increased ocean-induced melt can even trigger marine ice sheet instabilities (Weertman, 1974; Schoof, 2007; Gudmundsson et al., 2012), which have the potential to strongly increase Antarctic mass loss, on timescales below a century (Fox-Kemper et al., 2021). This makes ocean-induced sub-shelf melt, or *basal melt*, one of the main sources of uncertainty for future projections of sea-level rise.

Basal melt is a result of warm ocean water coming into contact with the base of the ice shelf. Which water masses reach the ice-ocean interface depends on the circulation of the water, not only in front of the ice shelf, but also after entering the ice-shelf cavity (Dinniman et al., 2016). As a consequence, to simulate the properties of the water at the ice-ocean interface accurately, both the ocean circulation around Antarctica and the circulation in the cavities below the ice shelves need to be simulated accurately. A few global or circum-Antarctic ocean models already include ice-shelf modules (Losch,

2008; Timmermann et al., 2012; Dinniman et al., 2015; Mathiot et al., 2017; Comeau et al., 2022), but such ocean models are expensive to run on long timescales or for large ensembles. Instead, a majority of the global climate models used until now in the Coupled (CMIP) or Paleoclimate (PMIP) Model Intercomparison Projects still poorly represent the ocean dynamics along the Antarctic margins and do not include ice-shelf cavities (Beadling et al., 2020; Heuzé, 2021). Getting the right water masses in the right place around Antarctica is a matter for global and regional ocean modelling and will not be the focus of this study. In this study, we focus on the circulation within the ice-shelf cavities and the resulting melt.

To infer the basal melt forcing for projections of the Antarctic contribution to sea-level rise, ice-sheet models commonly rely on parameterisations linking hydrographic properties in front of the ice shelves, given by observations or oceanic output from global climate models, and the basal melt (Jourdain et al., 2020). Due to different assumptions and simplifications concerning the circulation in the cavities, the range of existing basal melt parameterisations leads to widely differing melt patterns and associated contributions to sea-level rise (Favier et al., 2019; Burgard et al., 2022). The magnitude of the resulting uncertainty contribution is similar, or even larger, than the choice of emission scenario used to force the projections (Seroussi et al., 2020; Edwards & the ISMIP6 Team, 2021).

Emulating the three-dimensional ocean circulation within the cavity in simplified physical parameterisations is challenging and calls for exploring alternative approaches. We suggest that deep learning can be one tool to tackle this challenge. In recent years, the amount of ocean simulation output including ice-shelf cavities has increased and tools that make the application of deep learning techniques easily accessible have been developed, opening up the possibility of developing a neural network parameterisation for basal melt. If trained with high-resolution model output, a neural network parameterisation could implicitly include more intrinsic information about the system than a traditional physical parameterisation. This approach has been applied promisingly in several areas of Earth System Sciences in the form of multilayer perceptrons applied on the grid-cell level (e.g. Gentine et al., 2018; Rasp et al., 2018), convolutional neural networks applied on multidimensional fields (e.g. Bolton & Zanna, 2019; Rosier et al., 2023) or random forests (e.g. Yuval & O’Gorman, 2020).

Deep learning has also been explored for basal melt parameterisations. Rosier et al. (2023) performed promising experiments that showed that a cavity-resolving ocean model can be emulated with a convolutional neural network in a variety of idealised ice-shelf geometries. In the present study, we choose a different deep learning approach to developing such a *deep emulator*, or *surrogate model*, which differs on two fundamental points. On the one hand, we train on the circum-Antarctic cavity-resolving ocean simulations with realistic geometries used in Burgard et al. (2022). On the other hand, we use a multilayer perceptron architecture applied to each grid cell, as preliminarily explored in Bouissou et al. (2022). In the following, we present a proof of concept for a multilayer perceptron, which takes in hydrographic properties in front of the ice shelf and the geometric information at each grid point. In Sec. 2, we present the training and testing data, the neural network architecture, and the evaluation procedure. In Sec. 3, we show that the multilayer perceptron can successfully emulate cavity-resolving ocean simulations and produce integrated basal melt and patterns at least as close as but generally closer to the reference than traditional parameterisations in conditions similar to present. In Sec. 4 we explore the applicability of such a neural network to an independent set of simulations produced with a few different model parameters, slightly different geometries and in warmer oceanic conditions. Finally, in Sec. 5, we discuss the lessons learned from our study and give an outlook on possible directions to explore further in the future.

2 Data and Methods

The goal of this study is to explore if and how a neural network, in the form of a multilayer perceptron, can emulate the link between hydrographic properties in front of an ice shelf, geometric characteristics of the cavity, and the melt rates at its base as simulated by a cavity-resolving ocean model. In the following, we present the ocean model used and the set of simulations used for training, validation and testing the neural network; the neural network, its architecture, and its input variables; and the training and testing procedure.

2.1 Data

We choose to emulate a cavity-resolving version of the 3-D primitive-equation coupled ocean–sea-ice model NEMO (Nucleus for European Modelling of the Ocean, NEMO Team, 2019) run on the eORCA025 horizontal grid (Storkey et al., 2018). This grid has a resolution of 0.25° in longitude on average, i.e. a resolution of 4 to 14 km in the Antarctic seas and below the ice shelves, which is sufficient to capture the basic ocean circulation below multiple Antarctic ice shelves (Mathiot et al., 2017; Bull et al., 2021).

For the training phase, we use the same ensemble of simulations as used for the assessment of traditional basal melt parameterisations in Burgard et al. (2022). The ensemble is composed of four ocean simulations spanning 30 to 40 years, depending on the simulation, between 1979 and 2018. They were run with a standalone version of NEMO and forced with atmospheric forcing from JRA55-do version 1.4 (Tsujino et al., 2018). The Antarctic continental shelf bathymetry and ice shelf draft are constant and based on Bedmachine Antarctica version 2 (Morlighem, 2020; Morlighem et al., 2020). The simulations in the ensemble differ in a small number of parameters which are not directly related to the physics driving the ocean circulation and melt within the ice-shelf cavities but rather lead to a variety of hydrographic properties all around Antarctica. A more detailed description of the exact model configuration and differences in parameters can be found in Burgard et al. (2022).

For the testing phase, we use two simulations independent from the ensemble used for training. In this case, NEMO was run in coupled mode as the oceanic component of the Earth System Model UKESM1.0-ice (Smith et al., 2021), which couples the UK Earth System Model (UKESM1, Sellar et al., 2019) to an adapted version of the ice-sheet model BISICLES (Cornford et al., 2013). In this coupled configuration, the cavities below the ice shelves are open and the ice-shelf melt is computed with the same approach as in the training ensemble (as proposed by Mathiot et al., 2017). This means that a z^* coordinate is used for depth and the three equations are used to parameterise the ice-shelf melt in the ice-ocean boundary layer. Due to the coupled setup, the ice-shelf draft evolves according to the simulated evolution of the ice sheet. Note that the position of the ice front at the surface remains fixed by ice-sheet model design. More details about the configuration of NEMO in this model setup can be found in Smith et al. (2021). The two test simulations differ in their atmospheric forcing. In the first one, which we will call "REPEAT1970", UKESM1.0-ice was run for several decades under constant 1970 greenhouse gas and other forcings. In the second one, which we will call "4xCO2", UKESM1.0-ice was run for several decades under instantaneously quadrupled 1970 CO_2 concentrations. In our study, we use 60 years of simulation, from year 10 to year 70, for both runs.

The training and the testing dataset result from NEMO simulations. Nevertheless, next to differences in forcing from the atmosphere and the ice and bed geometry, the training and testing ensembles also differ in several technical aspects of NEMO. The training simulations were run with the version of 4.0.4. of NEMO (NEMO Team, 2019), including the sea-ice model SI³, while the test simulations were run with the version 3.6 of NEMO (Madec & NEMO Team, 2017) and version 5.1 of the Community Ice Code (CICE, Hunke et al., 2015). In addition, a few different parameter choices may affect the

170 link between hydrographic properties in front of the ice shelf and the melt at the base
 171 of the ice shelf. The training ensemble was computed on 121 vertical levels (represent-
 172 ing 20 m at 600 m depth), while the testing ensemble was computed on 75 vertical lev-
 173 els (representing 60 m at 600 m depth). In both ensembles, the thickness of the top bound-
 174 ary layer is bound at 20 m but can differ locally due to the different vertical resolutions.
 175 In the training ensemble, the thermal Stanton number is set to 7×10^{-4} while in the test-
 176 ing ensemble the thermal Stanton number is set to 1.45×10^{-3} . In the training ensem-
 177 ble, the top tidal velocity varies locally based on the CATS2008 dataset (Padman et al.,
 178 2008; Howard et al., 2019), while it is fixed to 5 cm/s in the testing ensemble. In con-
 179 clusion, this means that the testing ensemble is a slightly different model than the model
 180 which the neural network is trained to emulate and therefore represents a demanding test-
 181 ing experiment.

182 The training and testing ensembles cover a range of states that do not necessar-
 183 ily match observational estimates of hydrographic properties and basal melt rates. In
 184 both standalone and coupled mode, eORCAO25 configurations are prone to biases in the
 185 ocean circulation around Antarctica (Smith et al., 2021). Nevertheless, in Burgard et al.
 186 (2022), we showed that, if the forcing and parameters were carefully chosen to reproduce
 187 realistic ocean conditions in the Southern Ocean, the resulting basal melt rates were in
 188 agreement with observational estimates from Rignot et al. (2013). The physical link be-
 189 tween the hydrographic properties in front of the ice shelves and the basal melt rates is
 190 therefore reasonable. Based on this assumption, biases in the input properties should not
 191 affect the credibility of the training and evaluation procedure and the resulting neural
 192 network. On the contrary, a large variety of states is even beneficial because it provides
 193 more cases for our neural network to train on than only using the very limited sample
 194 of observations.

195 On a more technical note, for this study, the NEMO output was interpolated bi-
 196 linearly to a stereographic grid of 5 km spacing, as ice-sheet models and basal melt pa-
 197 rameterisations are commonly run on a stereographic grid. All pre-processing, training,
 198 testing, and analysis is conducted using this regridded data. From this regridded data,
 199 we cut out the different ice shelves according to latitude and longitude limits defined on
 200 the present geometry (details found in Burgard (2022)) and then apply a routine to adapt
 201 this mask to slightly different geometries, like the ones resulting from the fully coupled
 202 UKESM1.0-ice runs. Of these ice shelves, we only keep the largest ice shelves. The ef-
 203 fective resolution of physical ocean models, i.e. the resolution below which the circula-
 204 tion might not be resolved well, is typically 5 to 10 times the grid spacing (Bricaud et
 205 al., 2020). We empirically choose a cutoff at an area of 2500 km^2 (i.e. $6.25 \Delta x$) to be
 206 in this range while keeping a sufficiently large number of ice shelves. Due to different ge-
 207 ometries in the training and testing ensemble, this results into a slightly different ensem-
 208 ble of resolved ice shelves in these two ensembles (as listed in the figures of Appendix
 209 A).

210 2.2 Neural network

211 We design our neural network to predict the basal melt rates based on information
 212 about the ocean temperature and salinity in front of the ice shelf and about the ice-shelf
 213 geometry (Fig. 1). To link the input to the prediction, we use a multilayer perceptron,
 214 which is applied to each grid cell independently. A multilayer perceptron is the simplest
 215 form of a neural network and is a composition of functions (also called hidden layers),
 216 which takes an input array containing any number of variables and outputs a prediction.
 217 Specifying its number of neurons, each hidden layer is characterised by its parameters
 218 – the weights and biases, that connect each layer to its previous layer and shift the val-
 219 ues in the hidden layer, respectively. An activation function in the hidden layer intro-
 220 duces non-linearities in the relationship between input and output. In this study, we ex-
 221 plore different numbers of layers and numbers of neurons per layer. As activation func-

222 tion, we use the rectified linear unit (ReLU, Fukushima, 1975; Nair & Hinton, 2010). The
 223 multilayer perceptron is implemented in Python with the package Keras (Chollet et al.,
 224 2015).

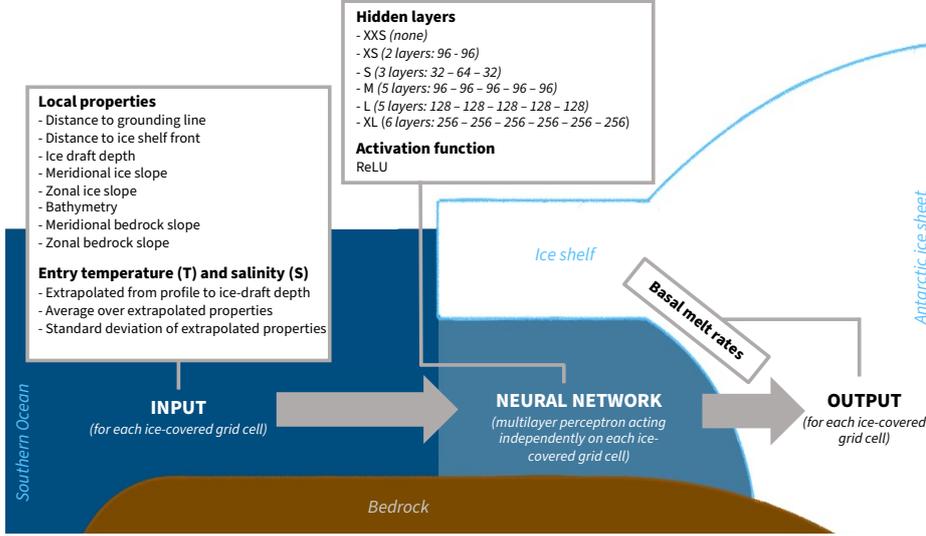


Figure 1. Schematic of the workflow around our neural network.

225 The strength of a neural network, and supervised machine learning techniques in
 226 general, is that it can reproduce complex non-linear relationships without being given
 227 the driving equations behind the data. Instead, its performance is driven by the super-
 228 vised training phase, which determines the weights and biases of each neuron in the net-
 229 work. During training, the loss, describing the averaged distance of the network predic-
 230 tions to a given target output, is backpropagated to the weights of the network. The weights
 231 are then optimised with stochastic gradient descent. The training dataset is randomly
 232 split up into batches, over which the optimisation is looped. A complete pass through
 233 the batches defines an epoch, and the weights and biases are optimised over several such
 234 epochs. In parallel to the training, the neural network is applied to a validation dataset
 235 to monitor its performance on data that has not been used for the training. After train-
 236 ing, the final performance of the neural network is estimated by applying it to a previ-
 237 ously unseen testing dataset.

238 In this study, to train the neural network, the loss which we reduce is the mean-
 239 squared-error over all ice-covered points between the predicted (m_{NN}) and target (m_{ref})
 240 basal melt rates,

$$MSE = \frac{\sum_i^{N_{pts}} \sum_t^{N_{years}} (m_{NN}[i, t] - m_{ref}[i, t])^2}{N_{pts} N_{years}} \quad (1)$$

241 where N_{pts} is the number of ice-covered grid points and N_{years} is the number of years
 242 used in the training. In Burgard et al. (2022), we argued that tuning on the grid-cell level
 243 would give too much weight to the larger ice shelves, as they cover a larger area. We still
 244 agree with this statement for traditional parameterisations because they already intrin-
 245 sically contain assumptions about the physics of the circulation and the melt before tun-
 246 ing and have only one or two tuneable parameters. In the case of our neural network,
 247 the relationship between the properties in front of the ice shelf and the melt is learnt
 248 from scratch, and it contains a larger number of parameters to adjust. We therefore argue that
 249 training on the grid-cell level is more sensible.

250 The neural network is optimised with Adam (Kingma & Ba, 2014), an initial learn-
 251 ing rate of 0.001, $\beta_1=0.9$ and $\beta_2=0.999$. We split the training dataset in batches with
 252 a size of 512 samples and optimise the neural network for at most 100 epochs. If the val-
 253 idation loss is not improved for 5 epochs, we reduce the learning rate by a factor of 2.
 254 If the validation loss is not improved for 10 epochs, we stop the training early. After early
 255 stopping, the model weights with the lowest validation loss are restored.

256 2.3 Input variables

257 The multilayer perceptron takes an array of variables as input for each grid cell in-
 258 dependently. In our case, the input array contains information about the geometrical prop-
 259 erties of the grid cell and the hydrographic forcing (Fig. 1).

260 For the geometrical properties, the input contains the following information: the ice
 261 draft depth, the local meridional and zonal slopes of the ice draft, the bathymetry,
 262 the local meridional and zonal slopes of the bedrock, and the distance of the grid cell
 263 to the nearest grounding line cell and the distance to the nearest ice front cell. All these
 264 variables are defined on the same horizontal plane and domain as the output array, the
 265 basal melt rates.

266 For the hydrographic forcing, more pre-processing is needed. To map the hydro-
 267 graphic forcing to the same grid cells as the other input variables, we proceed in the same
 268 manner as for traditional simple parameterisations in Burgard et al. (2022). First, we
 269 convert the conservative temperature and absolute salinity given by NEMO into poten-
 270 tial temperature and practical salinity with the GSW oceanographic toolbox (Firing et
 271 al., 2021). Second, we average the potential temperature and practical salinity, respec-
 272 tively, over the continental shelf within 50 km of the front of each ice shelf. The conti-
 273 nental shelf is defined as grid cells where the depth of the bathymetry is shallower than
 274 1500 m. The 50 km criterion imitates CMIP-type global ocean models that have reso-
 275 lutions around 1° (Heuzé, 2021), corresponding to a distance of between 38 km (70°S)
 276 and 56 km (60°S) in longitude. Third, we extrapolate the temperature and salinity from
 277 these mean profiles in front of the ice shelf to the local ice-draft depth, resulting in one
 278 local temperature and local salinity value per grid cell in the ice-shelf domain. Fourth,
 279 we also compute, for each time step, the average and standard deviation of these extrap-
 280 olated temperature and salinity fields and use them as additional input variables for each
 281 grid cell.

282 2.4 Training, validation and testing methodology

283 In a first step, we explore different neural network sizes using the method of cross
 284 validation on our training ensemble. In a second step, we choose a subsample of the neu-
 285 ral networks to explore their performance on the testing dataset.

286 We conduct two variations of leave-one-block-out cross validation to estimate the
 287 validation loss (MSE as defined in Eq. 1), one on the ice shelf dimension and one on the
 288 time dimension, like in Burgard et al. (2022). This approach consists of dividing the dataset
 289 into N blocks, training the neural network to minimise the training loss on $N-1$ blocks
 290 and using the left-out block to compute the validation loss (Wilks, 2006; Roberts et al.,
 291 2017). The procedure is re-iterated N times, leaving out each of the N blocks succes-
 292 sively, so that, in the end, each N -th block has been left out of training once. All pre-
 293 dictions for the left-out blocks, using the separately trained neural networks, are then
 294 concatenated to form a "synthetically independent" evaluation dataset. Applying an eval-
 295 uation metric on this evaluation dataset, we assess how well the neural network gener-
 296 alises to data "unseen" during training. We use $N=35$ for the cross validation over ice
 297 shelves. For the cross validation over time, we divide the years into blocks of approxi-
 298 mately 10 years (ten 10-year blocks and three 9-year blocks) to reduce the effect of au-

Table 1. Neural network size of the different variations explored in the cross validation.

Neural network configuration	Number of hidden layers	Number of neurons
XXS	0	0
XS	2	96/96
S	3	32/64/32
M	5	96/96/96/96/96
L	5	128/128/128/128/128
XL	6	256/256/256/256/256/256

299 tocorrelation, which is typically 2 to 3 years in our input temperatures. This results in
300 $N=13$ for the cross validation over time.

301 Before training, we normalise the training sample to put each of the 14 input vari-
302 ables (listed in Fig. 1) as well as the output variable on a similar order of magnitude and
303 avoid potential problems of gradient explosion. We do so by subtracting the mean and
304 dividing by the standard deviation of the training sample. To avoid that validation data
305 leaks into the training, this normalisation is reiterated for each iteration of the cross val-
306 idation.

307 We use the framework of cross validation to evaluate not only one but several neu-
308 ral networks to estimate the effect of their size on their performance. We sample differ-
309 ent sizes ranging from an extra-extra small (XXS) neural network, with no hidden layer,
310 and thus corresponding to a linear regression, to an extra-large (XL) neural network, with
311 six hidden layers, each containing 256 neurons. The different sizes are listed in Table 1.

312 To evaluate the resulting basal melt rates, we use the same metrics as in Burgard
313 et al. (2022), namely: (1) the root-mean-squared error (RMSE) of the yearly integrated
314 melt on the ice-shelf level and (2) the RMSE of the mean melt near the grounding line
315 for each ice shelf. For the former, we compute the RMSE between the simulated and em-
316 ulated yearly integrated melt (M) of the individual ice shelves [in Gt/yr] as follows:

$$RMSE_{\text{int}} = \sqrt{\frac{\sum_k^{N_{\text{isf}}} \sum_t^{N_{\text{years}}} (M_{\text{NN}}[k, t] - M_{\text{ref}}[k, t])^2}{N_{\text{isf}} N_{\text{years}}}} \quad (2)$$

317 where the subscript NN stands for neural network, N_{isf} is the number of ice shelves and
318 N_{years} the number of simulated years, and the integrated melt M of ice shelf k [in Gt/yr]
319 is:

$$M[k] = \rho_i \times 10^{-12} \sum_j^{N_{\text{grid cells in } k}} m_j a_j \quad (3)$$

320 where ρ_i is the ice density, m_j is the melt [in m ice per year] in grid cell j , and a_j is the
321 area of grid cell j . For the latter, we compute the RMSE between the simulated and em-
322 ulated yearly mean melt rate near the grounding line [in m ice per year]:

$$RMSE_{\text{GL}} = \sqrt{\frac{\sum_k^{N_{\text{isf}}} \sum_n^{N_{\text{simu}}} (m_{\text{GL,NN}}[k, n] - m_{\text{GL,ref}}[k, n])^2}{N_{\text{isf}} N_{\text{simu}}}} \quad (4)$$

323 where N_{simu} is the number of simulations in the ensemble and where m_{GL} for ice shelf
 324 k and simulation n is:

$$m_{\text{GL}}[k, n] = \frac{1}{N_{\text{years in } n}} \sum_t^{N_{\text{years in } n}} \frac{\sum_j^{N_{\text{grid cells near GL in } k}} (m_j a_j)}{\sum_j^{N_{\text{grid cells near GL in } k}} a_j} \quad (5)$$

325 The domain "near the grounding line" is the area covered by the first box prepared for
 326 the box parameterisation, when considering a maximum amount of five boxes, and is equiv-
 327 alent to approximately 10 % of the shelf area.

328 After cross validation, we choose a subsample of these neural networks to do fur-
 329 ther evaluation on a completely independent dataset. To do so, we reiterate the train-
 330 ing of the subsample of neural networks over the whole training dataset and choose to
 331 work with a deep ensemble (Lakshminarayanan et al., 2017). The final weights and bi-
 332 ases of neural networks depend on the initialisation of the weights before the first train-
 333 ing iteration (Goodfellow et al., 2016). To account for this uncertainty and gain a more
 334 robust performance from the neural networks, we reiterate the training of the subsam-
 335 ple of neural networks ten times with ten different random initialisations. We then ap-
 336 ply this deep ensemble of ten neural networks to the independent testing input and com-
 337 pute an ensemble mean over the ten resulting melt rates. Note that we only investigate
 338 a small sample of neural network sizes for exploration in this study and do not claim that
 339 the best performing neural network here is the best performing neural network for the
 340 problem. This study is rather a proof of concept to encourage further research in this
 341 direction.

342 3 Training and cross validation

343 3.1 Integrated melt and mean melt near the grounding line

344 The two evaluation metrics for the cross validation of the different neural network
 345 sizes are shown in Fig. 2. In addition, to compare the performance to traditional param-
 346 eterisations, we show the evaluation metrics for a subset of existing parameterisations:
 347 the quadratic local parameterisation using a constant Antarctic slope (e.g. Holland et
 348 al., 2008) and using a local slope (e.g. Favier et al., 2019; Jourdain et al., 2020), the plume
 349 parameterisation proposed by Lazeroms et al. (2019), the box parameterisation with the
 350 same box amount as in Reese et al. (2018), and the PICOP parameterisation from Pelle
 351 et al. (2019). The parameterisations are used as presented and tuned in Burgard et al.
 352 (2022).

353 Corresponding to a linear regression, the XXS neural network leads to a RMSE of
 354 a similar order as traditional parameterisations in the cross validation over time and, for
 355 the melt near the grounding line, in the cross validation over ice shelves as well. For the
 356 integrated melt, the cross validation over ice shelves leads to a comparably high RMSE.
 357 In the further course of this study, we therefore focus on neural networks that include
 358 hidden layers.

359 For both metrics, the RMSE for the cross validation over time is considerably re-
 360 duced when using a neural network with hidden layers compared to traditional param-
 361 eterisations and the XXS neural network. The RMSE for the cross validation over ice
 362 shelves is higher than for the cross validation over time but remains on the lower end
 363 of the range of RMSEs given by traditional parameterisations.

364 The RMSE_{int} of the cross validation over time is very similar between neural net-
 365 work sizes and spans between 6 Gt/yr (XL) and 11 Gt/yr (S). It remains well below the
 366 mean reference integrated melt on the ice-shelf level of 39 Gt/yr. The RMSE_{int} of the

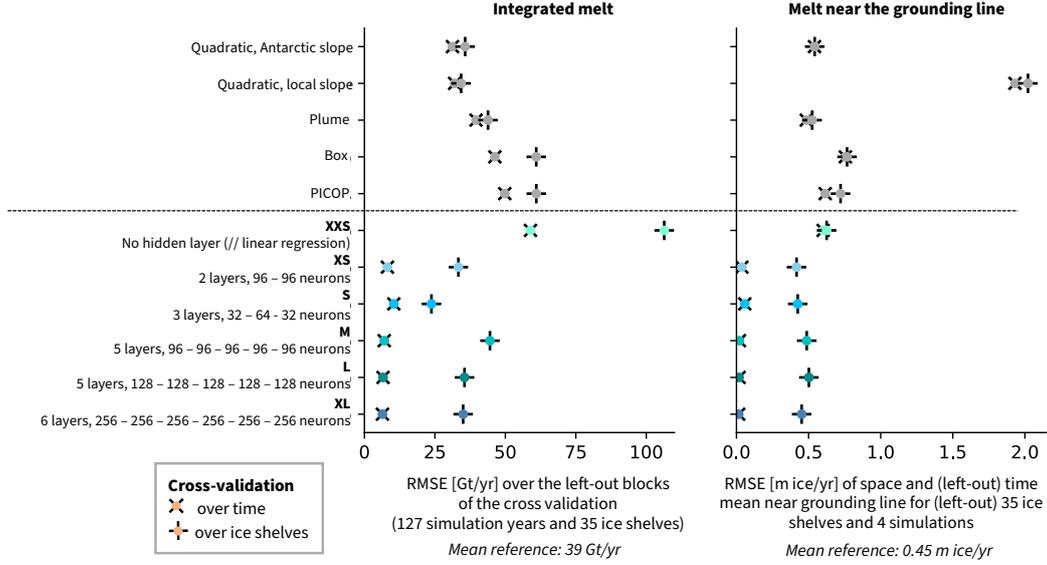


Figure 2. Summary of the RMSE of the integrated melt ($RMSE_{int}$) for the cross validation over time (\times) and for the cross validation over ice shelves ($+$) for a selection of traditional parameterisations (as shown in Burgard et al., 2022) [in Gt/yr] (left) and summary of the RMSE of the melt rate averaged over time and space near the grounding line ($RMSE_{GL}$) [in m ice/yr] (right). The colors represent the different parameterisation approaches: traditional parameterisations (grey), neural network (shades of blue). The RMSE is computed following Eq. (2), left panel, and Eq. (4), right panel, on the synthetically independent evaluation dataset.

367 cross validation over ice shelves varies more and is higher, between 24 (S) and 45 Gt/yr
 368 (M). The performance does not correlate with the neural network size. On the contrary,
 369 the lowest $RMSE_{int}$ of the cross validation over ice shelves is found for a comparably small
 370 neural network (S).

371 For the melt near the grounding line, the $RMSE_{GL}$ does not vary much in both cross
 372 validations between neural network sizes. The cross validation over time leads to a very
 373 low RMSE, varying from 0.02 m/yr (M,L,XL) to 0.06 m/yr (S). The cross validation over
 374 ice shelves leads to a RMSE between 0.42 m/yr (XS,S) and 0.50 m/yr (L), on the same
 375 order as the mean reference melt near the grounding line on the ice-shelf level, which is
 376 0.45 m ice/yr.

377 The neural networks have more difficulties generalising to unseen ice shelves than
 378 generalising to unseen time periods. This means that one of the obstacles for the neural
 379 networks' performance is the application to unknown cavity geometries. Some of the
 380 cavity geometries are so different from the rest of the ensemble that they force the neural
 381 networks to extrapolate far from their training domain. However, if they have seen
 382 a given geometry at least once during training, they perform well on this geometry for
 383 another time step. This aspect is encouraging, as this means that the neural networks
 384 adapt well to temperature and salinity variations across the training ensemble.

385 3.2 Spatial patterns

386 To add on the metrics at the ice-shelf level, we analyse the spatial patterns result-
 387 ing from the XS, S and L neural networks (Fig. 3) for the training ensemble member clos-

388 est to realistic conditions (called REALISTIC in Burgard et al., 2022). For the cross val-
 389 idation over time, the patterns of XS, S and L are nearly indistinguishable from the refer-
 390 ence for Filchner-Ronne, Pine Island, Fimbul, and Totten ice shelves. For Ross ice shelf,
 391 all patterns are close to the reference, but the S pattern contains more widespread melt-
 392 ing.

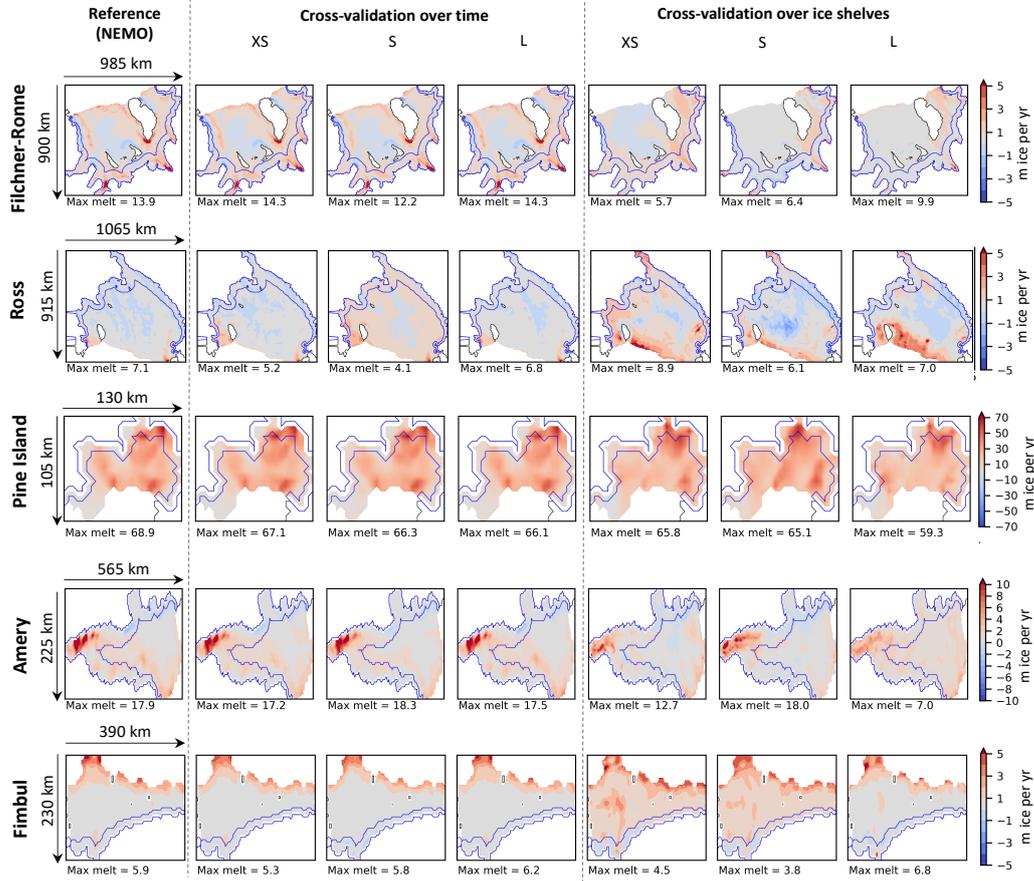


Figure 3. Subset of ice shelves for a visual evaluation of the melt patterns. This is the time average for the training ensemble member closest to real conditions (39 years) where the melt for each timestep has been computed with the neural network trained on the dataset leaving out that timestep (cross validation over time, columns 2 to 4) and where the melt of each ice-shelf has been computed with the neural network trained on the dataset leaving out that ice shelf (cross validation over ice shelves, columns 5 to 7). The blue line indicates the region used to evaluate the melt rate near the grounding line (which is defined as the first box in the 5-box setup of the box parameterisation).

393 For the cross validation over ice shelves, the patterns are not matching in as much
 394 detail as in the cross validation over time. In particular for the two largest ice shelves,
 395 Filchner-Ronne and Ross, it becomes clear that if the neural network has been trained
 396 without one of them, it will mimic the spatial pattern of the other because they are the
 397 only ones to share given ranges in the input variables, such as for example large distances
 398 to the ice front and grounding line. For Filchner-Ronne and Ross, the result of the cross
 399 validation over ice shelves does not match the reference in any of the neural networks.

400 For Pine Island and Amery, the XS and S patterns match the reference better than the
 401 L pattern, while, for Fimbul and Totten, the L pattern is a little better.

402 The low RMSE in the cross validation over time suggests an overfit on the geom-
 403 etry, which is fixed over time in the training dataset. The patterns very close to the ref-
 404 erence in the cross validation over time show that, even if our neural networks are ap-
 405 plied on each grid-cell separately, the location of the grid cell is more or less encoded in
 406 one or more input variables. However, as our problem is not necessarily well constrained
 407 with the input variables given, we suggest that this overfit can be used to our advantage.
 408 Our hypothesis is that, if the neural network has seen each ice shelf once, it has captured
 409 the variety of geometries and will be able to generalise to future changes in these "known"
 410 ice shelves. We do not expect new and completely different ice shelves to appear in the
 411 next centuries. To assess this idea, we need to investigate how well the neural network
 412 will perform on a geometry which is similar to but not identical to the training.

413 In the following, we investigate further if the neural networks are suitable for evol-
 414 ving ice-shelf geometries that are close to existing geometries and to temperature and salin-
 415 ity input properties outside the training range. We choose to continue with (1) the S size,
 416 because it has the lowest RMSE in the cross validation over ice shelves, (2) the XS size
 417 because it has similarly low RMSE to the larger sizes but remains very small and sim-
 418 ple, and (3) the L size to include a larger neural network and explore potential differ-
 419 ences during the testing compared to its behavior in the cross validation.

420 4 Testing on independent simulations

421 We apply our subsample of neural network sizes on two independent datasets, one
 422 representing 60 years of constant 1970-forcing (REPEAT1970), and one representing warmer
 423 conditions, i.e. 60 years of abrupt 4xCO₂ forcing (4xCO₂), from Smith et al. (2021). The
 424 REPEAT1970 simulation has a relatively steady ice-sheet geometry, similar (but not iden-
 425 tical) to the training geometry and is useful to assess the sensitivity of the neural net-
 426 works to different near-present-day atmospheric conditions (from the UKESM atmosphere
 427 component), to different parameters used in NEMO, and to slightly different geometries.
 428 The 4xCO₂ simulation experiences larger changes in ice-sheet geometry and much warmer
 429 conditions, which is useful to test the neural networks far outside of their training range.
 430 As a consequence, this evaluation is demanding and permits to evaluate the limits of the
 431 neural networks.

432 For evaluation, we divide the 4xCO₂ run into two 30-year blocks to capture poten-
 433 tial differences with warming in time. As explained in Sec. 2.4, we train the XS, S and
 434 L neural networks ten times each, with ten different random initialisations. In the fol-
 435 lowing, the results shown are averages over the predictions of the ten ensemble members
 436 for each neural network size.

437 4.1 Integrated melt and melt near the grounding line

438 The neural networks reproduce well the REPEAT1970 melt rates integrated over
 439 individual ice shelves, with a RMSE_{int} of 16 to 19 Gt/yr (Fig. 4a, left). This error is slightly
 440 larger than in the cross validation over time (see Fig. 2), and becomes similar to the quadratic
 441 and plume parameterisations. It should be noted that the RMSE_{int} of these parameter-
 442 isations is lower than in the cross validation, likely because of the overall lower melt rates
 443 in this simulation (24 Gt/yr compared to 39 Gt/yr in the training ensemble). The neu-
 444 ral networks still clearly outperform the box and PICOP parameterisation (RMSE_{int} \approx 35 Gt/yr).

445 For the melt near the grounding line, all parameterisations are uncertain, with RMSE_{GL}
 446 close to the reference mean melt near the grounding line of 0.34 m/yr (Fig. 4a, right).
 447 The neural networks and the traditional parameterisations yield similar RMSE_{GL}, be-

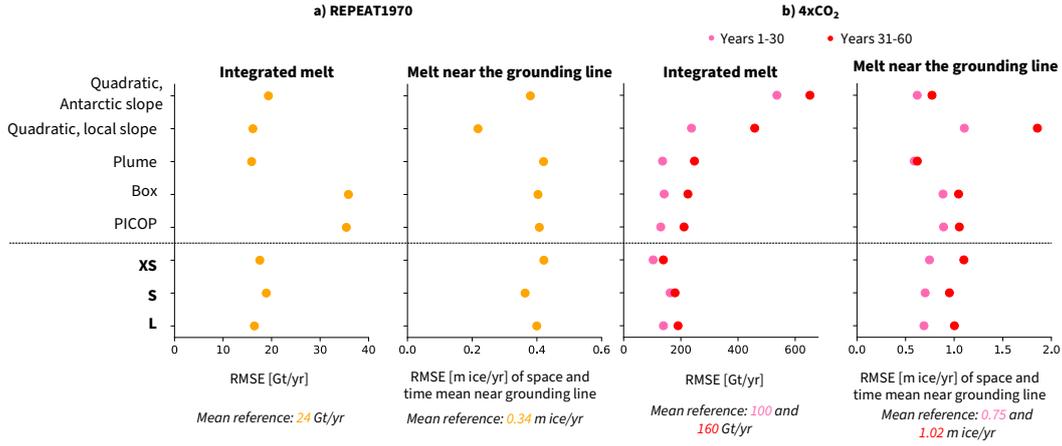


Figure 4. Summary of the RMSE of the integrated melt ($RMSE_{int}$) [in Gt/yr] and of the RMSE of the melt rate averaged over time and space near the grounding line ($RMSE_{GL}$) [in m ice/yr] for a selection of traditional parameterisations and a subsample of neural networks for the application on REPEAT1970 (a) and $4xCO_2$ (b). Note the change in x-axis between the (a) and (b) panels.

448 between 0.36 and 0.42 m/yr, except the quadratic using a local slope, which leads to a slightly
 449 lower RMSE, on the order of 0.22 m/yr.

450 For the warmer conditions ($4xCO_2$), all parameterisations struggle to reproduce
 451 the integrated melt on the ice-shelf level, with high spread in performance between the
 452 parameterisations (Fig. 4b, left). The $RMSE_{int}$ is multiplied by more than 10 for the neural
 453 networks and reaches nearly 650 Gt/yr for the quadratic parameterisation using an
 454 Antarctic slope in the second period. While this jump in RMSE can be explained by a
 455 higher mean reference integrated melt (100 Gt/yr for the first period and 159 Gt/yr for
 456 the second period, see also Fig. A3), it is probably also a result of forcing unseen during
 457 training such as much warmer and less saline ocean conditions (Figs. A1 and A2).
 458 Over both periods, the neural networks remain at the lower range of the difference to
 459 the reference melt rates. While neural networks, plume, box and PICOP parameterisation
 460 have comparable RMSEs for the first warm period (between 103 and 163 Gt/yr),
 461 the RMSE increases more for the plume, box and PICOP parameterisation (between 211
 462 and 248 Gt/yr) than for the neural networks (between 138 and 191 Gt/yr) in the even
 463 warmer second period.

464 For the melt near the grounding line, the parameterisations perform differently than
 465 for the integrated melt, pointing to potential challenges outside the domain near the ground-
 466 ing line. The neural networks perform in a similar uncertain manner as in the REPEAT1970
 467 case (Fig. 4b, right). Their $RMSE_{GL}$ (0.69-0.75 m/yr in the first period and 0.95-1.10 m/yr
 468 in the second period) is close to the reference mean melt near the grounding line (0.75 m/yr
 469 for the first period and 1.02 m/yr for the second period). In the first period, only the
 470 quadratic local parameterisation using an Antarctic slope and the plume parameterisation
 471 have lower $RMSE_{GL}$ (0.62 and 0.59 m/yr respectively), while in the second period
 472 only the quadratic parameterisation using a local slope performs clearly worse than the
 473 other parameterisations. For all, the RMSE increases with warmer conditions but the
 474 gap between the periods depends on the parameterisation, ranging from a difference of
 475 0.04 m/yr for the plume parameterisation to a difference of 0.76 m/yr for the quadratic
 476 parameterisation using a local slope.

477 From this demanding application on an independent testing dataset, several con-
 478 clusions can be drawn. First, the neural networks apply reasonably well to data inde-
 479 pendent from training in present conditions. This means that, if they have seen all ge-
 480 ometries of the main circum-Antarctic ice shelves, they can adapt to slightly different
 481 geometries. This is even more encouraging as the testing simulations were conducted with
 482 a slightly different version of NEMO than the neural networks were trained on. Second,
 483 none of the neural networks seems to constantly be the one with the best performance
 484 for all metrics. Third, the RMSE of the neural networks is higher when applied to warmer
 485 conditions, but, in comparison with the traditional parameterisations, it performs at least
 486 as well or even better.

487 4.2 Spatial patterns

488 Looking at the spatial patterns averaged over the last 10 years of the 4xCO₂ run,
 489 it becomes clear that all parameterisations, both neural networks and traditional ones,
 490 struggle with warmer conditions and different geometries to the training ensemble (Fig. 5).
 491 The maximum melt rates remain far below the maximum melt rates of the reference for
 492 all of them except the quadratic parameterisation using the local slope, which largely
 493 overestimates the maximum melt rates (as seen already in Burgard et al., 2022). Look-
 494 ing at the general patterns, the neural networks tend to overestimate the melt on wide
 495 areas of Filchner-Ronne and Ross but underestimate it over the whole ice shelf for smaller
 496 ones. The quadratic parameterisations (both using Antarctic and local slope) and, in some
 497 cases, the plume parameterisation, tend to overestimate the melt over wide areas, in par-
 498 ticular for the Ross and Filchner-Ronne ice shelves. The box parameterisation under-
 499 estimates the melt for all ice shelves, completely missing regions of strong melt.

500 5 Discussion

501 In this study, we showed that a simple multilayer perceptron can emulate melt rates
 502 as simulated by the cavity-resolving ocean model NEMO. This result is encouraging for
 503 further development because, as it is applied on a grid-cell level, it allows larger amounts
 504 of training data to be used than architectures containing convolutions such as MELT-
 505 NET (Rosier et al., 2023) or, more generally, U-Nets (Ronneberger et al., 2015), which
 506 take spatial domains as inputs. In addition, this architecture is independent of the do-
 507 main size and is therefore directly applicable to any ice shelf around Antarctica. In the
 508 following, we discuss insights from this study and possible further improvements to this
 509 approach.

510 5.1 Variable importance

511 One argument that is often made against the use of neural networks is that they
 512 remain statistical emulators of the training data and do not contain any physical con-
 513 straints. The performance when applied to a slightly different model and to different con-
 514 ditions (see Sec. 4) already gives us a sense that the neural networks can reasonably adapt
 515 to conditions outside of training. In addition, we now perform a sanity check to verify
 516 that the neural network is doing "the right thing for the right reasons". This sanity check
 517 also gives insight into the importance of the different input variables and could help fu-
 518 ture development of deep learning parameterisations as well as physical parameterisa-
 519 tions to focus on these variables.

520 To assess the importance of the different variables on the performance of the neu-
 521 ral networks, we apply two variations of the permute-and-predict approach. In the permute-
 522 and-predict approach, one of the variables is shuffled randomly and used as input for the
 523 neural network alongside the other variables that remain in the original order. In the first
 524 variation (Fig. 6a), we shuffle the input variables within the REPEAT1970 sample to eval-

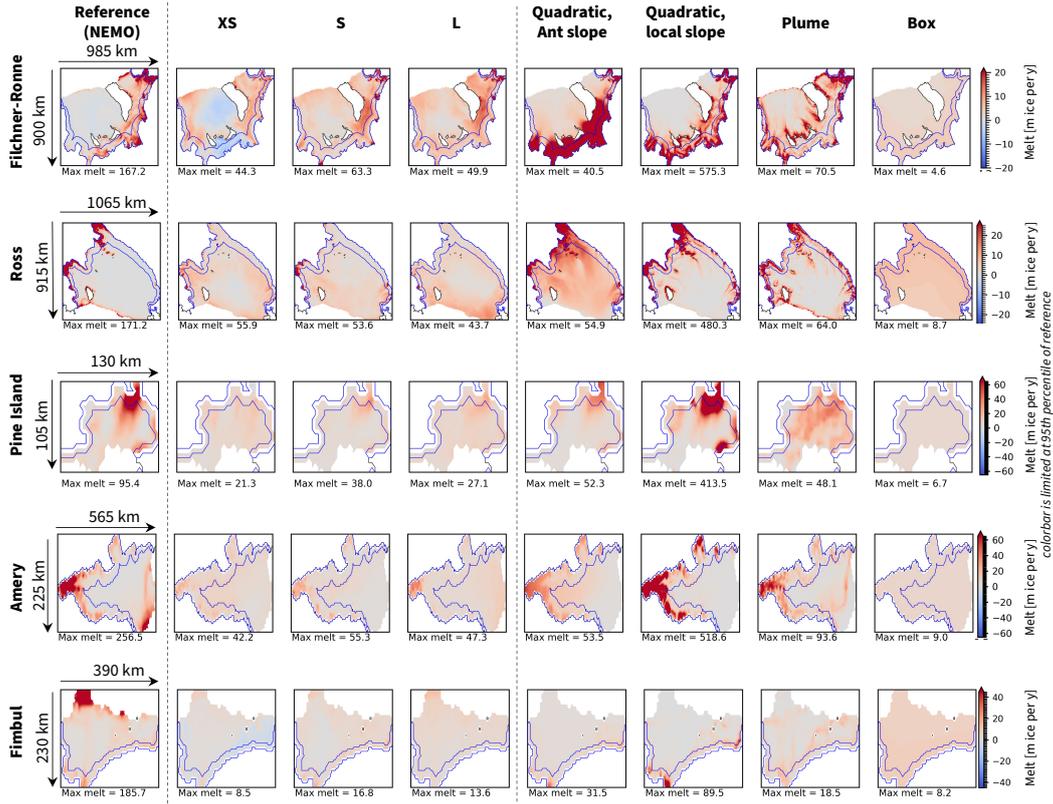


Figure 5. Subset of ice shelves for a visual evaluation of the melt patterns. This is the time average for the last 10 years of the 4xCO₂ run. The colorbar is limited to the 95th percentile of the NEMO reference. The blue line indicates the region used to evaluate the melt rate near the grounding line (which is defined as the first box in the 5-box setup of the box parameterisation).

525 uate the importance of the different variables in a situation close to the training condi-
 526 tions. In the second variation (Fig. 6b), we use a random sample from the 4xCO₂ input
 527 for the shuffled variable and run the neural network using all other original input vari-
 528 ables from the REPEAT1970 run to evaluate the importance of different variables in much
 529 warmer conditions. The shuffling is reiterated for each variable separately. In addition,
 530 we also shuffle blocks of potentially correlated variables simultaneously to gain insight
 531 on the effect of correlation on the shuffling results.

532 For the shuffling within the REPEAT1970, the geometric properties dominate the
 533 performance of all three neural networks for the integrated melt (Fig. 6a, left). For the
 534 XS version, the ice-shelf size, for which the distance to the ice front could be seen as a
 535 proxy, and the water column height, through ice-draft depth and bathymetry, have the
 536 highest importance. For the S and L version, the bathymetry is less important but the
 537 distance to the ice front and the ice-draft depth remain the most important variables,
 538 with an effect on the RMSE decreasing from S to L. The shuffling of the temperature
 539 and salinity variables have a smaller effect when shuffled separately, which can be ex-
 540 plained by the correlation between these variables. However, when shuffled by group, the
 541 temperature information gains in importance, leading to a similar increase in RMSE as
 542 the distance to the ice front in the L version. The bedrock and ice slopes are not impor-
 543 tant for the performance on the integrated melt. For the melt near the grounding line
 544 (Fig. 6a, right), many variables are not important, the RMSE is reduced when they are

	Integrated melt [Gt/yr]			Melt near grounding line [m ice/yr]		
	XS	S	L	XS	S	L
Original RMSE (REPEAT 1970)	17.6	18.9	16.5	0.42	0.36	0.40
(a) Difference in RMSE to original after shuffling within REPEAT1970						
Distance GL	2.5	2.2	-0.4	-0.05	-0.04	-0.06
Distance IF	15.4	15.5	11.7	0.03	0.06	0.05
Ice draft depth	20.4	18.8	10.5	0.02	-0.04	-0.02
Bathymetry	16.3	2.3	3.8	0.04	0.01	0.01
Slope bed lon	0.3	0.6	-0.2	-0.01	-0	-0.01
Slope bed lat	0.3	-0.2	0.1	0	0.01	0.01
Slope ice lon	0.4	1	0.5	0.02	0.05	0.03
Slope ice lat	0.1	0	-0	0.01	0.02	0.01
Temperature	4.7	5.2	5.2	0.09	0.11	0.11
Salinity	9.4	8.2	1.3	-0.03	-0.01	0
Temperature mean	3.3	5.3	4.4	0.06	0.1	0.09
Salinity mean	4.9	3.2	3.6	0.01	0.02	0.03
Temperature std	0.7	0.9	0.5	0	-0.02	0.05
Salinity std	2.2	0.4	1.4	0.02	0.05	0.04
Position	14.2	19	13	-0.02	0.01	-0.01
Water column	14.7	18.9	6.9	-0.03	-0.01	-0.01
Slopes bed	0.6	0.2	0.1	-0.01	0	-0
Slopes ice	1.1	1.1	1	0.05	0.07	0.05
Temperature info	10.2	13	12.9	0.14	0.18	0.17
Salinity info	3.3	5.1	2.6	0.05	0.06	0.08
(b) Difference in RMSE to original after inserting random sample from 4xCO₂ into REPEAT1970						
Distance GL	2.5	2.1	-0.4	-0.05	-0.03	-0.07
Distance IF	14.9	15	11.3	0.05	0.06	0.06
Ice draft depth	25.4	15.5	12.7	0.02	-0.04	-0.01
Bathymetry	16.7	2.4	4	0.04	0.01	0.01
Slope bed lon	0.3	0.5	-0.2	-0.01	0	-0.01
Slope bed lat	0.3	-0.1	0.1	0	0.01	0.01
Slope ice lon	0.4	1	0.5	0.02	0.04	0.03
Slope ice lat	0.2	0.2	-0.1	0.01	0.02	0.01
Temperature	179.7	151.7	85	-0.06	-0.01	-0.04
Salinity	51.1	115.2	10.1	0.08	0.04	0.05
Temperature mean	92.5	127.1	91.1	0.1	-0.06	-0.09
Salinity mean	120.9	377.4	55.3	-0.01	0.01	-0.01
Temperature std	12.9	1.9	13.2	-0	0.01	0.02
Salinity std	29.6	11.9	7.9	0.02	0.02	0.01
Position	13.9	18.6	13	-0.01	0.02	0
Water column	15.9	16.3	7.1	-0.03	-0	-0.01
Slopes bed	0.5	0.2	0.1	-0	0	0
Slopes ice	1.1	1.2	0.9	0.04	0.07	0.05
Temperature info	330.6	307.1	266.5	0.21	0.14	-0.03
Salinity info	20.7	95.8	3.2	0.07	0	0.06

Figure 6. Difference in RMSE between an application using a random sample for the given variable of the REPEAT1970 input (a) and of the 4xCO₂ input (b) and the original application on the REPEAT1970 input using the XS, S and L deep ensemble. The original RMSE when applied to REPEAT1970 is indicated above each column. The upper part of the tables shows the results when shuffling the variables individually while the lower part is for variables that have been shuffled as a group. "Temperature" and "Salinity" are the ocean properties extrapolated to the ice-draft depth, "Temperature mean" and "Salinity mean" are their average over each cavity, and "Temperature std" and "Salinity std" their standard deviation over each cavity. In the block *Position* we group the distance to the grounding line and to the ice front, in the block *Water column* we group the ice-draft depth and the bathymetry, in the block *Slopes bed* and *Slopes ice* we group the meridional and zonal slope of the bedrock and ice respectively, in the block *Temperature info* and *Salinity info* we group the local value, the average and the standard deviation of temperature and salinity respectively.

545 shuffled. The strongest effect is seen when shuffling the temperature variables as a group.
 546 The salinity variables, the ice slopes, and the distance to the ice front are the second most
 547 important group.

548 When inserting random samples of $4xCO_2$ input, the importance of the ice front,
 549 the ice-draft depth and the bathymetry remains of a similar order of magnitude for the
 550 integrated melt as in the REPEAT1970 shuffling (Fig. 6b, left). However, the effect of
 551 the temperature increases drastically and leads to increases in the RMSE of more than
 552 300 Gt/yr. For the XS and S, the importance of the grouped salinity information increases
 553 as well. This result reflects the difficulty for neural networks to extrapolate outside of
 554 the training range. Looking at the distribution of the input variables, the geometrical
 555 conditions in the $4xCO_2$ run are in a similar range as the training ensemble, despite an
 556 involving ice-shelf geometry, while the temperature and salinity variables are clearly out-
 557 side of the distribution (Fig. A4). For the melt near the grounding line (Fig. 6b, right),
 558 introducing variables from warmer conditions does not affect the RMSE very differently
 559 than in the REPEAT1970 case.

560 Several conclusions can be drawn from this experiment. First, this experiment shows
 561 that the geometry, in particular the distance to the ice front and the ice-draft depth, are
 562 key variables for the neural networks to infer reasonable integrated melt when applied
 563 on variables close to the training range, closely followed by the temperature. Ice-draft
 564 depth and temperature already are an integral part of existing parameterisations (Burgard
 565 et al., 2022). However, the distance to the ice-shelf front or the ice-shelf size are currently
 566 only partly considered, and only in the more complex parameterisations such as the plume
 567 and box parameterisations (Lazeroms et al., 2019; Reese et al., 2018).

568 Second, when applied to much warmer conditions, the distribution of geometric vari-
 569 ables remains close to their distribution in the training ensemble. In contrast, the tem-
 570 perature and salinity, well outside the training range, clearly affect the resulting inte-
 571 grated melt. This suggests that training the neural networks on simulations of warmer
 572 conditions could already improve their performance. Even more promising, the low ef-
 573 fect of geometry changes on integrated melt in warmer conditions suggests that coupled
 574 ice-ocean simulations of warmer conditions are not necessarily needed for training and
 575 that cavity-opening ocean simulations with fixed geometry could already be sufficient.

576 Third, for the melt near the grounding line, the position of the grid cell is (maybe
 577 surprisingly) less important than for the integrated melt and the key variable is the tem-
 578 perature information, both near the training range and in warmer conditions. While the
 579 ice slope does not affect the integrated melt, it has some effect on the melt near the ground-
 580 ing line. This suggests that including ice slopes is necessary for a good performance near
 581 the grounding line. However, the way it is currently included in simple parameterisations
 582 is not successful as we showed in Burgard et al. (2022) that it leads to a clear overesti-
 583 mation of the melt in this region.

584 Fourth, the effect of the shuffling on the RMSE is generally lower for the L size of
 585 the neural networks. This could suggest an overfit as it could mean that the neural net-
 586 work is not following variations in the input variables as much as the other neural net-
 587 work sizes and is therefore less flexible. This possible overfit would also explain why we
 588 did not see an increase in the performance during the cross-validation with increasing
 589 network size in Sec. 3.

590 5.2 Possible improvements

591 While the results of our neural networks are encouraging, a variety of further im-
 592 provements can be conducted in the future. The most obvious conclusion from this study
 593 is that predicting warmer conditions, similar to climate change conditions, is challeng-
 594 ing for this particular neural network architecture because these conditions were not con-

595 tained during training and neural networks are known to struggle with extrapolation prob-
596 lems. We therefore suggest, when possible, to introduce a set of simulations containing
597 high-end future scenarios in the training dataset to make the neural network more ro-
598 bust for future projections. At the same time, we saw that the traditional parameter-
599 isations struggle to represent future conditions as well. How to tune melt parameteri-
600 sations to be applicable in both present and future conditions is therefore a problem that
601 is not limited to deep learning approaches.

602 Another possible improvement is the treatment of the largest ice shelves. When
603 looking at the cross-validation results into more detail, i.e. at the scale of each ice shelf
604 (not shown), the total RMSE over all ice shelves is strongly influenced by the high RMSE
605 for the Ross ice shelf and, to a smaller extent, by the relatively high RMSE for the Filchner-
606 Ronne ice shelves. These two ice shelves have an area which is much larger than the other
607 ice shelves around Antarctica. Their cavities are so large that they develop their own
608 internal circulation (e.g. Gerdes et al., 1999; Naughten et al., 2021) and the residence
609 time of water masses reaches several years (Michel et al., 1979; Nicholls & Østerhus, 2004).
610 It is therefore not too surprising that parameterisations, which use input temperature
611 and salinity averaged over thousands of kilometers at the front of the ice shelves and do
612 not represent horizontal circulation explicitly, struggle with the representation of melt
613 in these cavities. If we remove these two from the RMSE in the 4xCO₂ case for exam-
614 ple, we find that the RMSE is clearly reduced for both neural networks and traditional
615 parameterisations (Fig. 7 compared to Fig. 4b). It would therefore be worth consider-
616 ing whether these rather simple parameterisations are appropriate for the application
617 on the Ross and Filchner-Ronne ice shelves and if it would not be wiser to push efforts
618 towards the opening of these two cavities in ocean models, even at the lower resolution
619 of 1°, as was already done for NEMO in Smith et al. (2021) or Hutchinson et al. (2023).
620 On the same line, we suggest it is worth thinking about tuning the parameterisations
621 on the smaller ice shelves, and tuning the parameters and neural networks differently on
622 the larger ice shelves.

623 There is also space for improvement in the definition of input temperatures and
624 salinities. Like in Burgard et al. (2022), the input profiles of temperature and salinity
625 are here averaged over a given domain in front of the ice shelf. Then, we extrapolate the
626 properties to the ice-draft depth. To give the neural network more information about
627 the whole profile, we also gave it the mean and standard deviation of these extrapolated
628 temperature and salinity. However, machine learning gives us the opportunity to think
629 bigger than traditional statistics when representing information about a given domain.
630 One direction that could be explored in further development is the encoding of the im-
631 portant information about the water masses in front of the ice shelf using a machine learn-
632 ing technique. Ideally, this technique would take in a three-dimensional (horizontal plane
633 and depth), or even a four-dimensional (taking also time as input to account for lags and
634 residence time), field of temperature and salinity in front of the ice shelf and encode in-
635 formation about this field in a format to be given to the neural network. Such encod-
636 ing might contain more information about the spatial distribution of the properties in
637 front of the ice shelf and therefore potentially encode changes in the ocean circulation
638 which might change the circulation within the cavities, as expected to happen in warmer
639 conditions for the Filchner-Ronne ice shelf (Naughten et al., 2021).

640 Rosier et al. (2023) showed that a convolutional architecture can also be used to
641 infer basal melt rates from hydrographic and geometric properties. A convolutional ar-
642 chitecture, often U-Nets, is the preferred choice in many current studies exploring the
643 application of machine learning to Earth System Sciences (e.g. Ebert-Uphoff & Hilburn,
644 2020; Andersson et al., 2021; Finn et al., 2023). In the case of basal melt and the ocean
645 circulation in the cavity, such architectures clearly make sense as they can capture spa-
646 tial patterns and correlations. However, these architectures require much more simula-
647 tion data for training as they take each time step as one training sample while our ap-

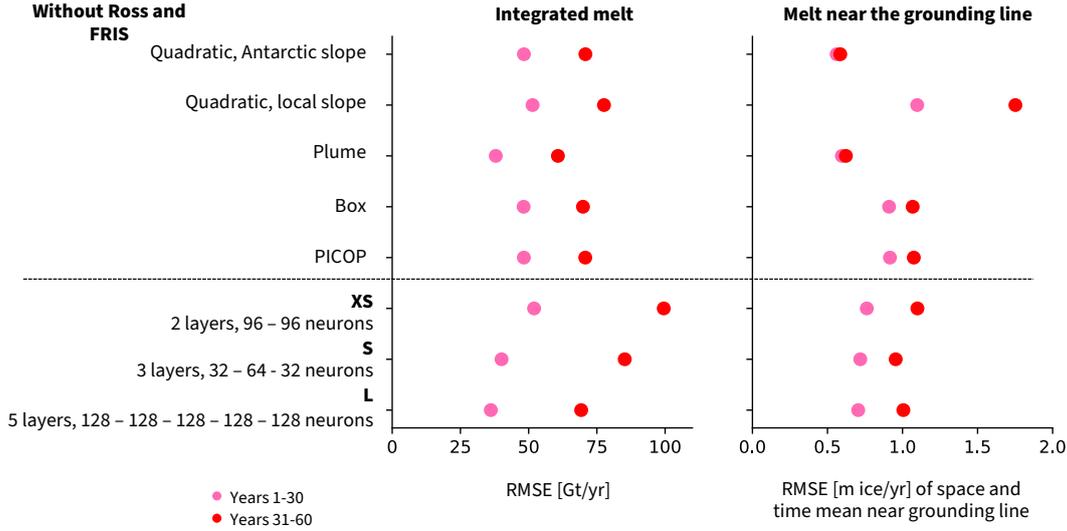


Figure 7. Summary of the RMSE of the integrated melt ($RMSE_{int}$) [in Gt/yr] and of the RMSE of the melt rate averaged over time and space near the grounding line ($RMSE_{GL}$) [in m ice/yr] computed on all ice shelves except Ross and Filchner-Ronne ice shelves for a selection of traditional parameterisations and a subsample of neural networks for the application on a simulation with $4xCO_2$ forcing. The lighter colors represent the first 30 years of simulation and the darker colors the last 30 years of simulation.

648 proach takes each time step and grid cell as one training sample. Also, Rosier et al. (2023)
 649 demonstrate the performance of their MELTNET in a fixed domain and have not yet
 650 shown how to apply it to larger ice shelves than this domain. MELTNET remains how-
 651 ever a promising approach and we are looking forward to its further development.

652 Finally, this study has focussed on the emulation of one ocean model at a given res-
 653 olution. We acknowledge that NEMO’s simulation of basal melt rates is not a perfect
 654 reflection of reality. Therefore, an interesting further direction to follow would be to train
 655 a neural network to emulate NEMO at other resolutions and also to emulate other cavity-
 656 resolving ocean models. In this context, to ensure that the relationship remains sensi-
 657 ble, we suggest training separate emulators and using them as an ensemble. This would
 658 provide an ensemble of emulators to be used as a variety of basal melt parameterisations,
 659 in addition to physics-based parameterisations. In a context where basal melt remains
 660 one of the main sources of uncertainty in projections of the Antarctic contribution to sea-
 661 level rise, a wide sample of this uncertainty in the form of a higher variety of parame-
 662 terisations is welcome.

663 6 Conclusions

664 In conclusion, we show that a rather simple neural network architecture can be used
 665 to emulate a cavity-resolving ocean model. Our multilayer perceptrons are designed to
 666 be rather simply usable as a basal melt parameterisation for ice-sheet modellers. They
 667 use input properties needed for the traditional parameterisations already and can be ap-
 668 plied on the grid-cell level, similarly to most traditional parameterisations. While they
 669 struggle nearly as much as traditional parameterisations to generalise to ice shelves un-
 670 seen during tuning, the neural networks generalise much better on time blocks unseen

671 during training and the patterns are clearly better represented. In the demanding test-
 672 ing phase, on a dataset produced with different NEMO parameters, geometry pertur-
 673 bations unseen during training and different forcing, they still perform at least as well
 674 or even better than traditional parameterisations, both in historical and much warmer
 675 conditions.

676 These results are promising as neural networks and machine learning in general are
 677 topics that have been gaining lots of traction lately and efforts are done in many disci-
 678 plines of the Earth System Sciences to explore their application. In this study, we
 679 provide guiding thoughts for further exploration and refinement of this approach, while this
 680 first proof of concept can already be used as an additional parameterisation in the ice-
 681 sheet modelling landscape.

682 **Appendix A Distributions of variables of interest in the training and**
 683 **testing ensemble**

Temperature profiles over 50 km in front of the ice shelf
 for the different simulations of the ensemble

TRAINING DATA REPEAT1970 4xCO₂
 (127 years) (60 years) (60 years)

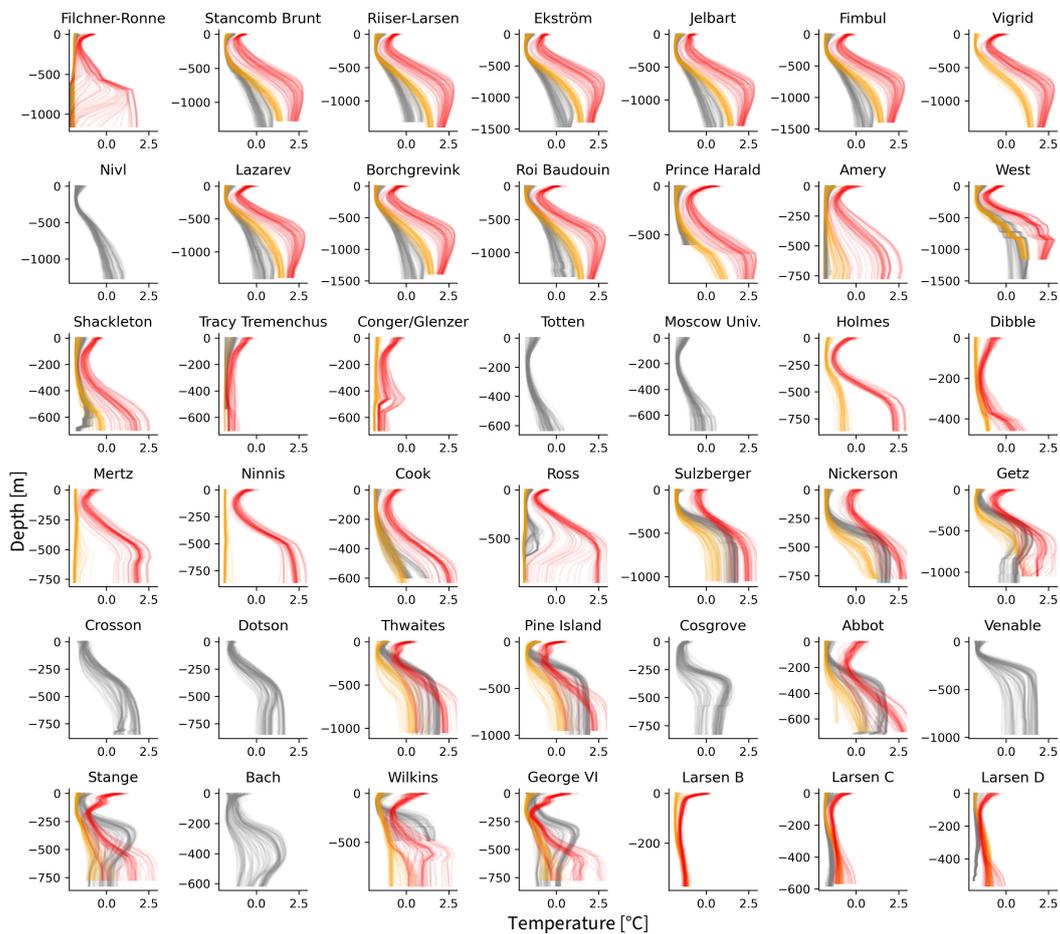


Figure A1. Input profiles of temperature for the different ice shelves. Profiles of the training ensemble are shown in grey, profiles for the REPEAT1970 run in orange and profiles for the 4xCO₂ run in red.

Salinity profiles over 50 km in front of the ice shelf
for the different simulations of the ensemble

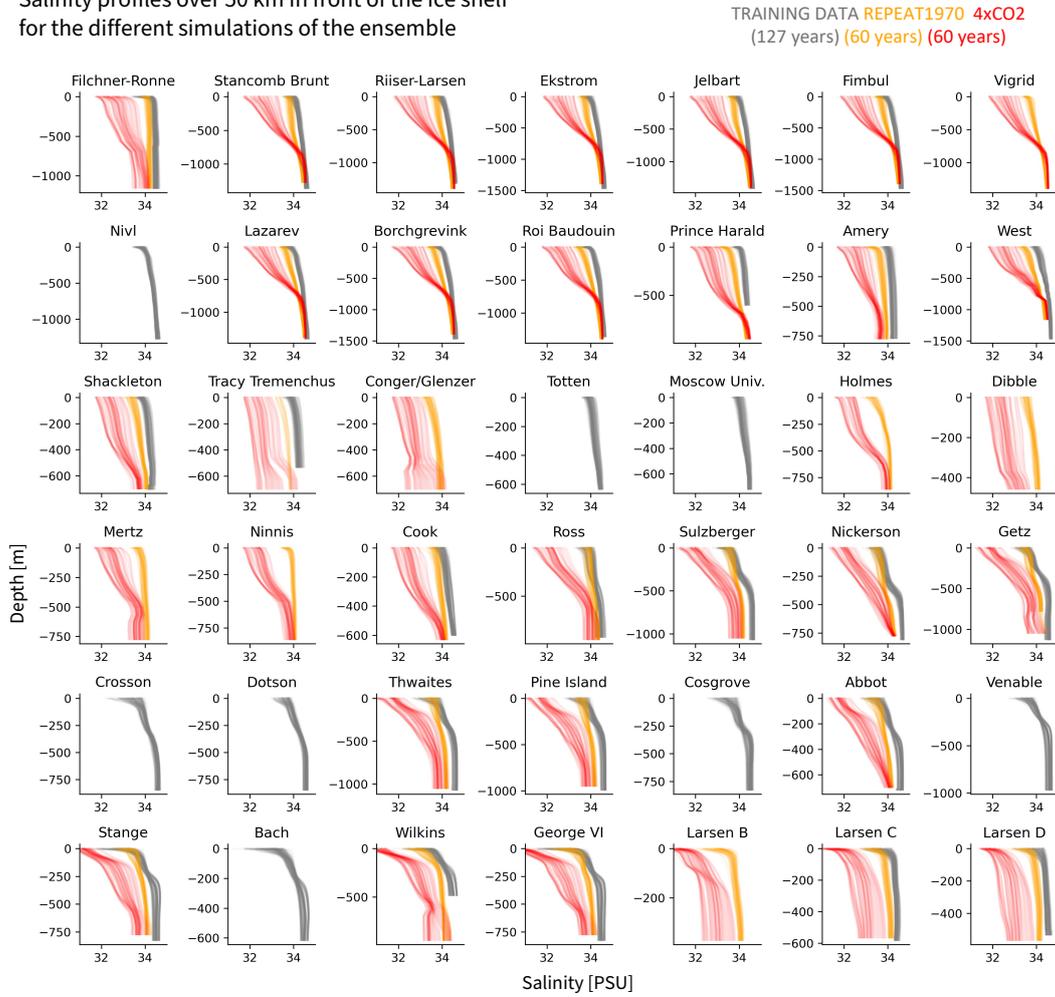


Figure A2. Input profiles of salinity for the different ice shelves. Profiles of the training ensemble are shown in grey, profiles for the REPEAT1970 run in light blue and profiles for the 4xCO₂ run in dark blue.

Integrated melt over time

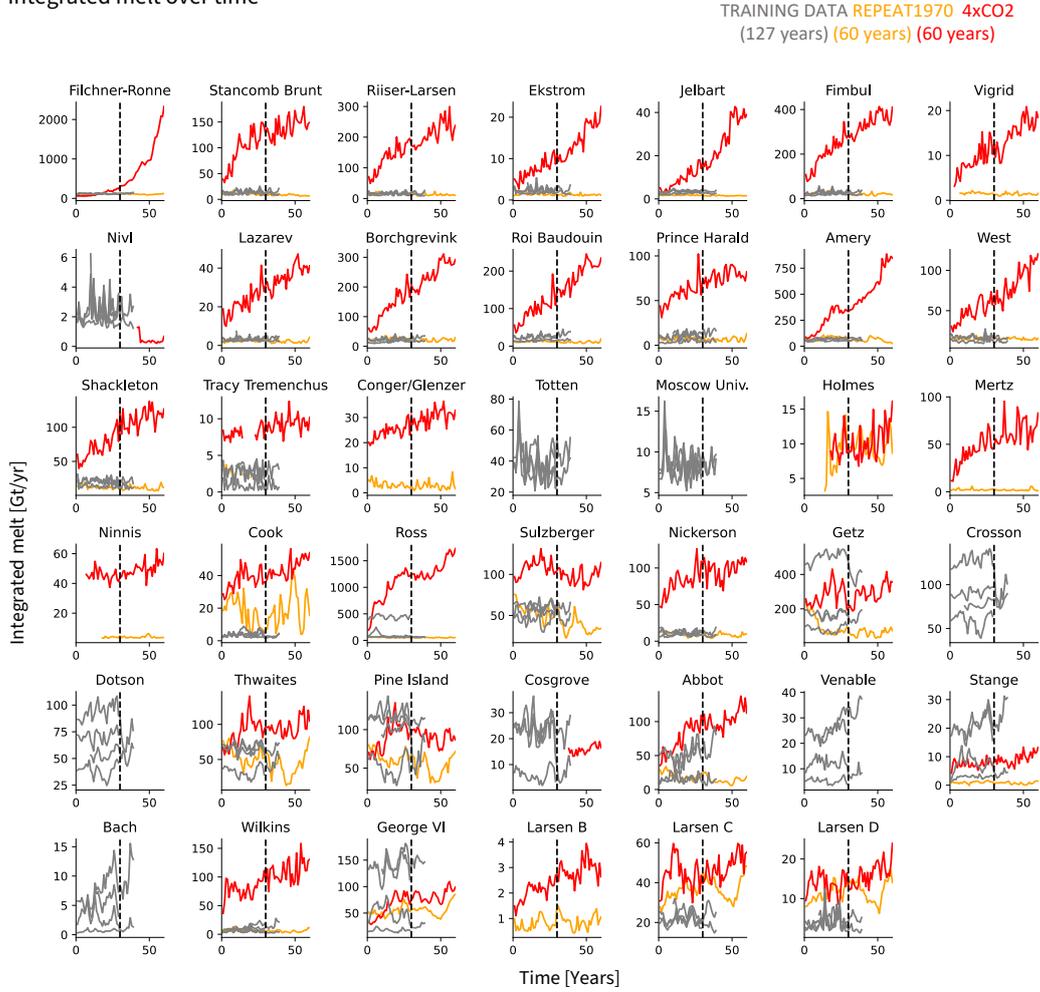


Figure A3. Timeseries of the integrated melt for the different ice shelves. The training ensemble is shown in grey, the REPEAT1970 run in orange and the 4xCO₂ run in red. The black dashed line limits the first and second 30-year block used in Sec. 4 for the 4xCO₂ run

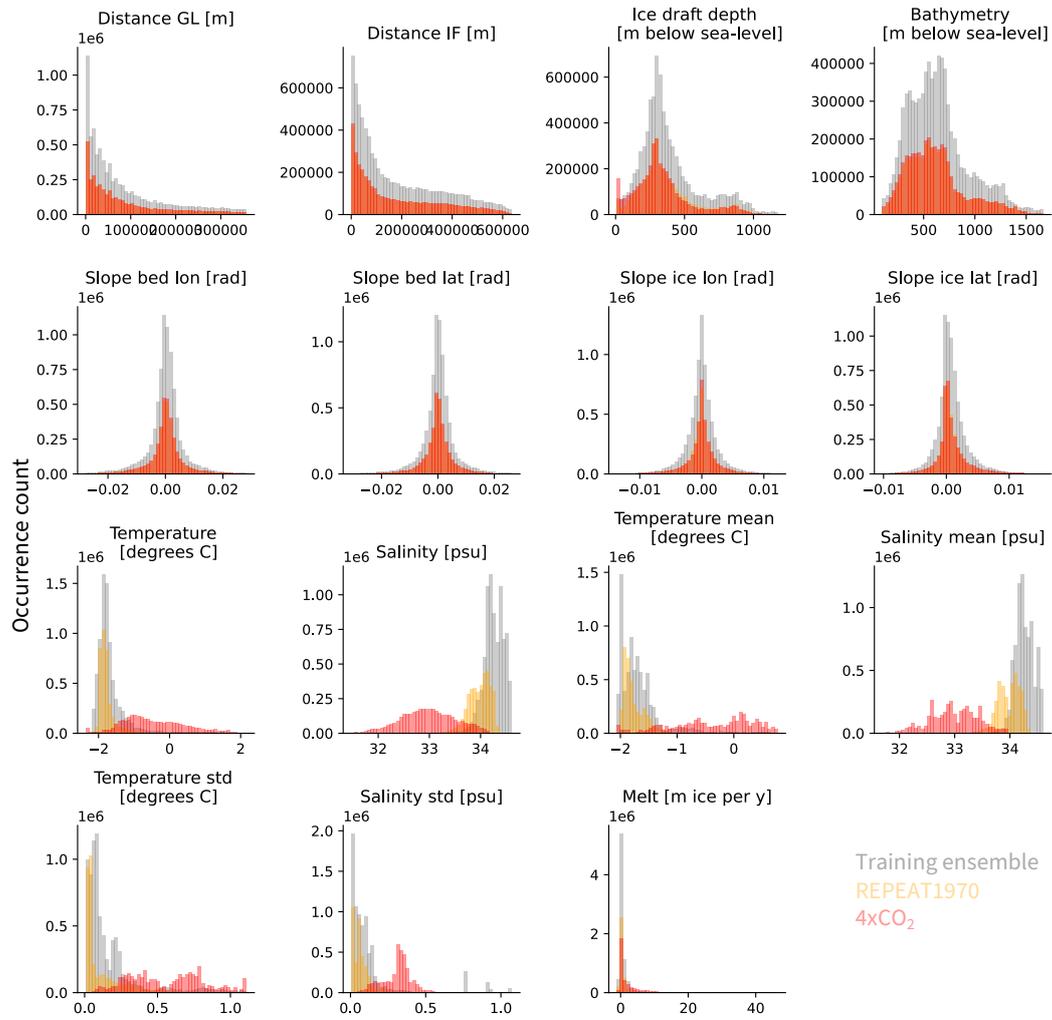


Figure A4. Distribution (occurrence count) of the different input variables and the melt over the training ensemble (grey), the REPEAT1970 run (orange) and the 4xCO₂ run (red).

684 **Open Research**

685 The simulation data from Burgard et al. (2022) used for the training ensemble can
686 be found on Zenodo: <https://doi.org/10.5281/zenodo.7308352>. The simulation data
687 from (Smith et al., 2021) used for the testing ensemble will be uploaded on Zenodo as
688 soon as possible. All code to train the neural networks and produce the figures can be
689 found on Github: https://github.com/ClimateClara/basal_melt_neural_network
690 and will be uploaded to Zenodo upon paper acceptance.

691 **Acknowledgments**

692 We thank Dani Jones, Paul Holland, Tom Andersson, Alex Bradley, Simon Thomas, Anna
693 Vaughan, and many others at the British Antarctic Survey for interesting discussions and
694 exchange on ocean, ice, and machine learning. Most of the computations presented in
695 this paper were performed using the GRICAD infrastructure ([https://gricad.univ-grenoble
696 -alpes.fr](https://gricad.univ-grenoble-alpes.fr)), which is supported by Grenoble research communities. The NEMO sim-
697 ulations were performed using HPC resources from GENCI-CINES (MISOCS project,
698 allocations A0080106035 and A0100106035). This research was mainly conducted through
699 the DEEP-MELT project (IRGA Pack IA 2021-2022), which is supported by MIAI @
700 Grenoble Alpes (ANR-19-P3IA-0003). This research was also supported by the Euro-
701 pean Union’s Horizon 2020 research and innovation programme under grant agreements
702 no. 869304 (PROTECT), 820575 (TiPACCs) and 101003536 (ESM2025), as well as by
703 the French National Research Agency through the AIAI project (ANR-22-CE01-0014).

704 CB and NCJ developed the original idea of this paper. CB carried out all analy-
705 ses and wrote the manuscript. PM carried out the NEMO simulations used for train-
706 ing and RSS carried out the UKESM simulations. NCJ, RS and JC provided valuable
707 help and code for the definition of the ice-shelf masks when the ice shelves evolve over
708 time. TSF provided methodological input on the training of neural networks and JEJ
709 provided useful input about how to think about machine learning. CB, NCJ, PM, RS,
710 RSS, JC, TSF, JEJ contributed to discussions.

References

- 711
- 712 Adusumilli, S., Fricker, H., Medley, B., Padman, L., & Siegfried, M. (2020). Inter-
 713 annual variations in meltwater input to the Southern Ocean from Antarctic ice
 714 shelves. *Nature Geoscience*, *13*, 616-620. doi: 10.1038/s41561-020-0616-z
- 715 Andersson, T., Hosking, J., Pérez-Ortiz, M., Paige, B., Elliott, A., Russell, C.,
 716 ... Shuckburgh, E. (2021). Seasonal Arctic sea ice forecasting with
 717 probabilistic deep learning. *Nature Communications*, *12*, 5124. doi:
 718 10.1038/s41467-021-25257-4
- 719 Beadling, R., Russell, J., Stouffer, R., Mazloff, M., Talley, L., Goodman, P., ...
 720 Pandde, A. (2020). Representation of Southern Ocean Properties across Cou-
 721 pled Model Intercomparison Project Generations: CMIP3 to CMIP6. *Journal*
 722 *of Climate*, *33*(15), 6555-6581. doi: 10.1175/JCLI-D-19-0970.1
- 723 Bolton, T., & Zanna, L. (2019). Applications of Deep Learning to Ocean Data In-
 724 ference and Subgrid Parameterization. *Journal of Advances in Modeling Earth*
 725 *Systems*, *11*(1), 376-399. doi: 10.1029/2018MS001472
- 726 Bouissou, B., Burgard, C., & Jourdain, N. (2022). Parameterising ocean-induced
 727 melt of an idealised Antarctic ice shelf using deep learning. *ECCOMAS22 Con-*
 728 *ference proceedings*. doi: 10.23967/eccomas.2022.216
- 729 Bricaud, C., Le Sommer, J., Madec, G., Calone, C., Deshayes, J., Ethe, C., ...
 730 Levy, M. (2020). Multi-grid algorithm for passive tracer transport in
 731 the NEMO ocean circulation model: a case study with the NEMO OGCM
 732 (version 3.6). *Geoscientific Model Development*, *13*(11), 5465-5483. doi:
 733 10.5194/gmd-13-5465-2020
- 734 Bull, C., Jenkins, A., Jourdain, N., Vaňková, I., Holland, P., Mathiot, P., ... Sallée,
 735 J. (2021). Remote control of filchner-ronne ice shelf melt rates by the
 736 antarctic slope current. *Journal of Geophysical Research: Oceans*, *126*. doi:
 737 10.1029/2020JC016550
- 738 Burgard, C. (2022). Multimelt, a python framework to apply existing basal melt
 739 parameterisation. *Python Package Index - PyPI*, [https://pypi.org/
 740 project/multimelt/](https://pypi.org/project/multimelt/).
- 741 Burgard, C., Jourdain, N., Reese, R., Jenkins, A., & Mathiot, P. (2022). An
 742 assessment of basal melt parameterisations for Antarctic ice shelves. *The*
 743 *Cryosphere*, *16*(12), 4931-4975. doi: 10.5194/tc-16-4931-2022
- 744 Chollet, F., et al. (2015). *Keras*. <https://keras.io>.
- 745 Comeau, D., Asay-Davis, X., Begeman, C., Hoffman, M., Lin, W., Petersen, M., ...
 746 Turner, A. (2022). The DOE E3SM v1.2 Cryosphere Configuration: Descrip-
 747 tion and Simulated Antarctic Ice-Shelf Basal Melting. *Journal of Advances in*
 748 *Modeling Earth Systems*, *14*, e2021MS002468. doi: 10.1029/2021MS002468
- 749 Cornford, S., Martin, D., Graves, D., Ranken, D., Le Brocq, A., Gladstone, R.,
 750 ... Lipscomb, W. (2013). Adaptive mesh, finite volume modeling of ma-
 751 rine ice sheets. *Journal of Computational Physics*, *232*, 529-549. doi:
 752 10.1016/j.jcp.2012.08.037
- 753 Dinniman, M., Asay-Davis, X., Galton-Fenzi, B., Holland, P., Jenkins, A., & Tim-
 754 mermann, R. (2016). Modeling Ice Shelf/Ocean Interaction in Antarctica: A
 755 Review. *Oceanography*, *29*(4), 144-153. doi: 10.5670/oceanog.2016.106
- 756 Dinniman, M., Klinck, J., Bai, L.-S., Bromwich, D., Hines, K., & Holland, D. (2015).
 757 The Effect of Atmospheric Forcing Resolution on Delivery of Ocean Heat to
 758 the Antarctic Floating Ice Shelves. *Journal of Climate*, *28*, 6067-6085. doi:
 759 10.1175/JCLI-D-14-00374.1
- 760 Ebert-Uphoff, I., & Hilburn, K. (2020). Evaluation, Tuning, and Interpretation of
 761 Neural Networks for Working with Images in Meteorological Applications. *Bul-*
 762 *letin of the American Meteorological Society*, *101*, E2149-E2170. doi: 10.1175/
 763 BAMS-D-20-0097.1
- 764 Edwards, T., & the ISMIP6 Team. (2021). Projected land ice contributions
 765 to twenty-first-century sea level rise. *Nature*, *593*(7857), 74-82. doi:

- 10.1038/s41586-021-03302-y
- 766 Favier, L., Jourdain, N., Jenkins, A., Merino, N., Durand, G., Gagliardini, O., ...
 767 Mathiot, P. (2019). Assessment of sub-shelf melting parameterisations using
 768 the ocean–ice-sheet coupled model NEMO(v3.6)–Elmer/Ice(v8.3). *Geoscientific*
 769 *Model Development*, *12*(6), 2255–2283. doi: 10.5194/gmd-12-2255-2019
- 770 Finn, T., Durand, C., Farchi, A., Bocquet, M., Chen, Y., Carrassi, A., & Dansereau,
 771 V. (2023). Deep learning of subgrid-scale parametrizations for short-term
 772 forecasting of sea-ice dynamics with a Maxwell-Elasto-Brittle rheology. *EGU-*
 773 *sphere*. doi: 10.5194/egusphere-2022-1342
- 774 Firing, E., Fernandes, F., Barna, A., & Abernathey, R. (2021). Teos-10/gsw-python:
 775 v3.4.1.post0. *Zenodo*. ([used version 3.6.16]) doi: 10.5281/zenodo.5214122
- 776 Fox-Kemper, B., Hewitt, H., Xiao, C., Adalgeirsdóttir, G., Drijfhout, S., Edwards,
 777 T., ... Yu, Y. (2021). Ocean, Cryosphere and Sea Level Change [Chapter]. In
 778 V. Masson-Delmotte et al. (Eds.), *Climate Change 2021: The Physical Science*
 779 *Basis. Contribution of Working Group I to the Sixth Assessment Report of the*
 780 *Intergovernmental Panel on Climate Change* (chap. 9). Cambridge, United
 781 Kingdom and New York, NY, USA: Cambridge University Press.
- 782 Fukushima, K. (1975). Cognitron: A self-organizing multilayered neural network. *Bi-*
 783 *ological Cybernetics*, *20*(3), 121–136. doi: 10.1007/BF00342633
- 784 Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., & Yacalis, G. (2018). Could Ma-
 785 chine Learning Break the Convection Parameterization Deadlock? *Geophysical*
 786 *Research Letters*, *45*(11), 5742–5751. doi: 10.1029/2018GL078202
- 787 Gerdes, R., Determann, J., & Grosfeld, K. (1999). Ocean circulation beneath
 788 Filchner-Ronne Ice Shelf from three-dimensional model results. *Journal of*
 789 *Geophysical Research: Oceans*, *104*, 15827–15842. doi: 10.1029/1999JC900053
- 790 Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
 791 (<http://www.deeplearningbook.org>)
- 792 Gudmundsson, G., Krug, J., Durand, G., Favier, L., & Gagliardini, O. (2012). The
 793 stability of grounding lines on retrograde slopes. *The Cryosphere*, *6*(6), 1497–
 794 1505. doi: 10.5194/tc-6-1497-2012
- 795 Heuzé, C. (2021). Antarctic Bottom Water and North Atlantic Deep Water in
 796 CMIP6 models. *Ocean Science*, *17*(1), 59–90. doi: 10.5194/os-17-59-2021
- 797 Holland, P., Jenkins, A., & Holland, D. (2008). The Response of Ice Shelf Basal
 798 Melting to Variations in Ocean Temperature. *Journal of Climate*, *21*(11),
 799 2558–2572. doi: 10.1175/2007JCLI1909.1
- 800 Howard, S. L., Padman, L., & Erofeeva, S. (2019). *Cats2008: Circum-antarctic tidal*
 801 *simulation version 2008*. Retrieved from [https://www.usap-dc.org/view/](https://www.usap-dc.org/view/dataset/601235)
 802 [dataset/601235](https://www.usap-dc.org/view/dataset/601235) doi: 10.15784/601235
- 803 Hunke, E., Lipscomb, W., Turner, A., Jeffery, N., & Elliott, S. (2015). *CICE: The*
 804 *Los Alamos sea ice model, documentation and software, version 5.1 la-cc-06-*
 805 *012 (Computer software manual No. LA-CC-06-012)*.
- 806 Hutchinson, K., Deshayes, J., Ethé, C., Rousset, C., de Lavergne, C., Vancoppenolle,
 807 M., ... Mathiot, P. (2023). Improving Antarctic Bottom Water precursors in
 808 NEMO for climate applications. *EGUsphere*. doi: 10.5194/egusphere-2023-99
- 809 Jourdain, N., Asay-Davis, X., Hattermann, T., Straneo, F., Seroussi, H., Little, C.,
 810 & Nowicki, S. (2020). A protocol for calculating basal melt rates in the IS-
 811 MIP6 Antarctic ice sheet projections. *The Cryosphere*, *14*(9), 3111–3134. doi:
 812 10.5194/tc-14-3111-2020
- 813 Khazendar, A., Rignot, E., Schroeder, D., Seroussi, H., Schodlok, M., Scheuchl, J.,
 814 B.and Mouginot, ... Velicogna, I. (2016). Rapid submarine ice melting in the
 815 grounding zones of ice shelves in West Antarctica. *Nature Communications*,
 816 *7*(1), 13243. doi: 10.1038/ncomms13243
- 817 Kingma, D., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv*
 818 *preprint*. doi: 10.48550/ARXIV.1412.6980
- 819 Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and Scalable
- 820

- 821 Predictive Uncertainty Estimation Using Deep Ensembles. In *Proceedings of*
 822 *the 31st International Conference on Neural Information Processing Systems*
 823 (p. 6405–6416). Red Hook, NY, USA: Curran Associates Inc.
- 824 Lazeroms, W., Jenkins, A., Rienstra, S., & van de Wal, R. (2019). An Ana-
 825 lytical Derivation of Ice-Shelf Basal Melt Based on the Dynamics of Melt-
 826 water Plumes. *Journal of Physical Oceanography*, *49*(4), 917-939. doi:
 827 10.1175/JPO-D-18-0131.1
- 828 Losch, M. (2008). Modeling ice shelf cavities in a z coordinate ocean general circu-
 829 lation model. *Journal of Geophysical Research: Oceans*, *113*, C08043. doi: 10
 830 .1029/2007JC004368
- 831 Madec, G., & NEMO Team. (2017). Nemo ocean engine (v3.6-patch). *Notes du Pôle*
 832 *de modélisation de l'Institut Pierre-Simon Laplace (IPSL)*, *Zenodo*, *27*. doi: 10
 833 .5281/zenodo.3248739
- 834 Mathiot, P., Jenkins, A., Harris, C., & Madec, G. (2017). Explicit representation
 835 and parametrised impacts of under ice shelf seas in the z^* coordinate ocean
 836 model nemo 3.6. *Geoscientific Model Development*, *10*(7), 2849-2874. doi:
 837 10.5194/gmd-10-2849-2017
- 838 Michel, R., Linick, T., & Williams, P. (1979). Tritium and carbon-14 distributions
 839 in seawater from under the Ross Ice Shelf Project ice hole. *Science*, *203*(4379),
 840 445–446.
- 841 Morlighem, M. (2020). *MEaSURES BedMachine Antarctica, Version 2*. (Boulder,
 842 Colorado USA. NASA National Snow and Ice Data Center Distributed Active
 843 Archive Center.) doi: 10.5067/E1QL9HFQ7A8M
- 844 Morlighem, M., Rignot, E., Binder, T., Blankenship, D., Drews, R., Eagles, G., ...
 845 Young, D. (2020). Deep glacial troughs and stabilizing ridges unveiled beneath
 846 the margins of the antarctic ice sheet. *Nature Geoscience*, *13*, 132-137. doi:
 847 10.1038/s41561-019-0510-8
- 848 Mouginot, J., Rignot, E., & Scheuchl, B. (2014). Sustained increase in ice discharge
 849 from the Amundsen Sea Embayment, West Antarctica, from 1973 to 2013.
 850 *Geophysical Research Letters*, *41*(5), 1576-1584. doi: 10.1002/2013GL059069
- 851 Nair, V., & Hinton, G. (2010). Rectified Linear Units Improve Restricted Boltz-
 852 mann Machines. In *Proceedings of the 27th International Conference on Inter-
 853 national Conference on Machine Learning* (p. 807–814). Madison, WI, USA:
 854 Omnipress. doi: 10.5555/3104322.3104425
- 855 Naughten, K., De Rydt, J., Rosier, S., Jenkins, A., Holland, P., & Ridley, J. (2021).
 856 Two-timescale response of a large Antarctic ice shelf to climate change. *Nature*
 857 *Communication*, *12*, 1991. doi: 10.1038/s41467-021-22259-0
- 858 NEMO Team. (2019). Nemo ocean engine. *Scientific Notes of Climate Modelling*
 859 *Center*, *27*. doi: 10.5281/zenodo.1464816
- 860 Nicholls, K. W., & Østerhus, S. (2004). Interannual variability and ventilation
 861 timescales in the ocean cavity beneath Filchner-Ronne Ice Shelf, Antarc-
 862 tica. *Journal of Geophysical Research: Oceans*, *109*(C4), C04014. doi:
 863 10.1029/2003JC002149
- 864 Padman, L., Erofeeva, S., & Fricker, H. (2008). Improving antarctic tide models by
 865 assimilation of icesat laser altimetry over ice shelves. *Geophysical Research Let-
 866 ters*, *35*, L22504. doi: 10.1029/2008GL035592
- 867 Paolo, F., Fricker, H., & Padman, L. (2015). Volume loss from Antarctic ice shelves
 868 is accelerating. *Science*, *348*(6232), 327-331. doi: 10.1126/science.aaa0940
- 869 Pelle, T., Morlighem, M., & Bondzio, J. (2019). Brief communication: PI-
 870 COP, a new ocean melt parameterization under ice shelves combining
 871 PICO and a plume model. *The Cryosphere*, *13*(3), 1043-1049. doi:
 872 10.5194/tc-13-1043-2019
- 873 Rasp, S., Pritchard, M., & Gentine, P. (2018). Deep learning to represent subgrid
 874 processes in climate models. *Proceedings of the National Academy of Sciences*,
 875 *115*(39), 9684–9689. doi: 10.1073/pnas.1810286115

- 876 Reese, R., Albrecht, T., Mengel, M., Asay-Davis, X., & Winkelmann, R. (2018).
 877 Antarctic sub-shelf melt rates via PICO. *The Cryosphere*, *12*(6), 1969-1985.
 878 doi: 10.5194/tc-12-1969-2018
- 879 Rignot, E., Jacobs, S., Mouginot, J., & Scheuchl, B. (2013). Ice-shelf melting around
 880 Antarctica. *Science*, *341*(6143), 266-270. doi: 10.1126/science.1235798
- 881 Rignot, E., Mouginot, J., Morlighem, M., Seroussi, H., & Scheuchl, B. (2014).
 882 Widespread, rapid grounding line retreat of Pine Island, Thwaites, Smith, and
 883 Kohler glaciers, West Antarctica, from 1992 to 2011. *Geophysical Research*
 884 *Letters*, *41*(10), 3502-3509. doi: 10.1002/2014GL060140
- 885 Roberts, D., Bahn, V., Ciuti, S., Boyce, M., Elith, J., Guillera-Arroita, G., ... Dor-
 886 mann, C. (2017). Cross-validation strategies for data with temporal, spa-
 887 tial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913-929. doi:
 888 10.1111/ecog.02881
- 889 Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for
 890 biomedical image segmentation. In *Medical Image Computing and Computer-*
 891 *Assisted Intervention—MICCAI 2015: 18th International Conference, Munich,*
 892 *Germany, October 5-9, 2015, Proceedings, Part III 18* (pp. 234–241). doi:
 893 doi.org/10.1007/978-3-319-24574-4_28
- 894 Rosier, S., Bull, C., Woo, W., & Gudmundsson, G. (2023). Predicting ocean-induced
 895 ice-shelf melt rates using deep learning. *The Cryosphere*, *17*(2), 499-518. doi:
 896 10.5194/tc-17-499-2023
- 897 Scheuchl, J., B.and Mouginot, Rignot, E., Morlighem, M., & Khazendar, A. (2016).
 898 Grounding line retreat of Pope, Smith, and Kohler Glaciers, West Antarctica,
 899 measured with Sentinel-1a radar interferometry data. *Geophysical Research*
 900 *Letters*, *43*(16), 8572-8579. doi: 10.1002/2016GL069287
- 901 Schoof, C. (2007). Ice sheet grounding line dynamics: Steady states, stability, and
 902 hysteresis. *J. Geophys. Res.*, *112*(F3), F03S28. doi: 10.1029/2006JF000664
- 903 Sellar, A., Jones, C., Mulcahy, J., Tang, Y., Yool, A., Wiltshire, A., ... Zerroukat,
 904 M. (2019). UKESM1: Description and Evaluation of the U.K. Earth System
 905 Model. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4513-4558.
 906 doi: 10.1029/2019MS001739
- 907 Seroussi, H., Nowicki, S., Payne, A., Goelzer, H., Lipscomb, W., Abe-Ouchi, A.,
 908 ... Zwinger, T. (2020). ISMIP6 Antarctica: a multi-model ensemble of the
 909 Antarctic ice sheet evolution over the 21st century. *The Cryosphere*, *14*(9),
 910 3033–3070. doi: 10.5194/tc-14-3033-2020
- 911 Shen, Q., Wang, K., Shum, C., Jiang, L., Hsu, H., & Dong, J. (2018). Re-
 912 cent high-resolution Antarctic ice velocity maps reveal increased mass loss
 913 in Wilkes Land, East Antarctica. *Scientific Reports*, *8*(1), 4477. doi:
 914 10.1038/s41598-018-22765-0
- 915 Smith, R., Mathiot, P., Siahaan, A., Lee, V., Cornford, S., Gregory, J., ... Jones,
 916 C. (2021). Coupling the U.K. Earth System Model to Dynamic Models of the
 917 Greenland and Antarctic Ice Sheets. *Journal of Advances in Modeling Earth*
 918 *Systems*, *13*, e2021MS002520. doi: 10.1029/2021MS002520
- 919 Storkey, D., Blaker, A., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E., ...
 920 Sinha, B. (2018). Uk global ocean go6 and go7: a traceable hierarchy of
 921 model resolutions. *Geoscientific Model Development*, *11*, 3187-3213. doi:
 922 10.5194/gmd-11-3187-2018
- 923 The IMBIE Team. (2018). Mass balance of the Antarctic Ice Sheet from 1992 to
 924 2017. *Nature*, *558*(7709), 219-222. doi: 10.1038/s41586-018-0179-y
- 925 Timmermann, R., Wang, Q., & Hellmer, H. (2012). Ice-shelf basal melting in a
 926 global finite-element sea-ice/ice-shelf/ocean model. *Geoscientific Model Devel-*
 927 *opment*, *5*3, 303-314. doi: 10.3189/2012AoG60A156
- 928 Tsujino, H., Urakawa, S., Nakano, H., Small, R., Kim, W., Yeager, S., ... Yamazaki,
 929 D. (2018). Jra-55 based surface dataset for driving ocean–sea-ice models
 930 (jra55-do). *Ocean Modelling*, *130*, 79-139. doi: 10.1016/j.ocemod.2018.07.002

- 931 Weertman, J. (1974). Stability of the Junction of an Ice Sheet and an Ice Shelf.
932 *Journal of Glaciology*, *13*(67), 3-11. doi: 10.3189/S0022143000023327
- 933 Wilks, D. (2006). *Statistical methods in the atmospheric sciences* (2nd ed.). Amster-
934 dam Paris: Elsevier.
- 935 Yuval, J., & O’Gorman, P. (2020). Stable machine-learning parameterization of sub-
936 grid processes for climate modeling at a range of resolutions. *Nature communi-
937 cations*, *11*(1), 1–10. doi: 10.1038/s41467-020-17142-3