Calibrating calving parameterizations using graph neural network emulators: Application to Helheim Glacier, East Greenland

Younghyun Koo[1,2,3], Gong Cheng[4], Mathieu Morlighem[4], and Maryam Rahnemoonfar[1,2]

[1]Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA 18015, USA
[2]Department of Civil and Environmental Engineering, Lehigh University, Bethlehem, PA 18015, USA
[3]National Snow and Ice Data Center (NSIDC), Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado Boulder, Boulder, CO 80309, USA
[4]Department of Earth Sciences, Dartmouth College, Hanover, NH 03755, USA

Correspondence: Maryam Rahnemoonfar (maryam@lehigh.edu)

Abstract. Calving is responsible for the retreat, acceleration, and thinning of numerous tidewater glaciers in Greenland. An accurate representation of this process in ice sheet numerical models is critical to better predict the future response of the ice sheet to climate change. While traditional numerical models have been used to simulate ice dynamics and calving under specific parameterized conditions, the computational demand of these models makes it difficult to efficiently fine-tune these

5   parameterizations, adding to the overall uncertainty in future sea level rise. In this study, we adopt three standard graph neural network (GNN) architectures, including graph convolutional network, graph attention network, and equivariant graph convolutional network (EGCN), to develop surrogate models for finite-element simulations from the Ice-sheet and Sea-level System Model. GNNs are particularly well-suited for this problem as they naturally capture the representation of unstructured meshes used by finite-element models. When these GNNs are trained with numerical simulations of Helheim Glacier, Greenland,

10   for different calving stress thresholds, they successfully reproduce the observed evolution of ice velocity, ice thickness, and ice front migration between 2007 and 2020. Moreover, these emulators exhibit uncertainties of less than 10-20 % when extrapolating to out-of-sample calving parameterization cases. Among the three GNN architectures, EGCN outperforms the others by preserving the equivariance of graph structures. By leveraging the GPU-based GNN emulators, which are 30-34 times faster than traditional numerical simulations, we determine the temporal variations of the optimal calving threshold that

15   minimizes the misfit between modeled and observed ice fronts. This fine-tuned calving parameterization, enabled by GNN emulators, can enhance the reliability of numerical models in capturing glacier mass loss driven by calving.

1   Introduction

Over the past three decades, the Greenland ice sheet has experienced an average annual loss of 170 Gt of ice, resulting in a global mean sea level rise exceeding 15 mm (Otosaka et al., 2023). This trend of mass loss has intensified in recent years.

20   Between 1990 and 2000, the annual mass loss hovered around 40 Gt a^{-1}, but in the 2010s, it surged to approximately 280 Gt a^{-1} (Otosaka et al., 2023; Mouginot et al., 2019). This escalating mass loss can be attributed to two primary processes: (1) the change in surface mass balance driven by enhanced surface melt and (2) calving and submarine melting of marine-terminating

glaciers, commonly referred to as frontal ablation (King et al., 2020; Choi et al., 2021). In specific regions and seasons, ice discharge may be responsible for more than 50% of the total ice sheet mass loss. (Mouginot et al., 2019; King et al., 2020; Choi
25 et al., 2021; Aschwanden et al., 2019).

Enhanced calving significantly impacts ice dynamics, often resulting in ice flow acceleration and subsequent thinning (Bondzio et al., 2017; Cheng et al., 2022; Lippert et al., 2024). Given that calving is sensitive to climate conditions (Greene et al., 2024; Wood et al., 2021), it is important to understand how future calving rates would impact ice sheet mass balance and sea-level rise (Choi et al., 2021). Numerous studies have utilized ice sheet numerical models to explore the applicability of
30 various calving laws to Greenland and Antarctic ice sheets and identified the set of model parameters that minimize the misfit between model and observations (Wilner et al., 2023; Choi et al., 2018). According to Choi et al. (2018), the von Mises calving law (Morlighem et al., 2016, VM) best replicates observed terminus positions of nine outlet glaciers in Greenland compared to other existing laws.

Although numerical models can provide reliable solutions for ice flow when ice extent is kept constant, capturing the precise
35 impacts of spatiotemporally varying calving rates on terminus migration and ice flow in numerical models remains challenging. Additionally, assimilating remote sensing observations into numerical models, required to infer certain model parameters, is both complex to implement and computationally intensive (Choi et al., 2023). Furthermore, the integration of calving as a boundary condition in numerical models introduces significant complexity since calving directly alters the ice geometry and model domain during simulations. Consequently, identifying the optimal calving parameterizations consistent with observations is difficult and
40 time-consuming, thereby limiting our ability to project the future mass balance of ice sheets under various parameter settings (Choi et al., 2018; Edwards et al., 2021; Morlighem et al., 2020).

To address the computational demands of numerical models, various statistical approaches have emerged relying on faster machine learning models in lieu of ice sheet numerical models. While traditional numerical models often necessitate using high-performance computing clusters to solve partial differential equations (PDEs) on central processing units (CPUs), machine
45 learning emulators offer the advantage of operating on lighter computational resources, leveraging the parallel processing capabilities of graphic processing units (GPUs). For instance, Downs et al. (2023) used a Gaussian Process (GP) emulator to infer the sensitivity of time-independent model parameters to the frontal ablation of Helheim Glacier in Southeast Greenland. Although they were able to identify the best set of calving threshold parameters in the VM calving law, their GP approach did not account for spatial relationships or interactions between neighboring nodes. Additionally, their emulator focused on
50 matching the observed and modeled terminus positions along a central flow line rather than the entire glacier system.

Given that ice dynamics and calving are affected by the glacier geometry and underlying bed topography, it is important for the emulators to learn the spatial context across the entire glacier domain. To account for spatial relationships between nodes for emulating ice sheet dynamics, graph neural networks (GNNs) have emerged as an effective neural network architecture for handling irregular non-Euclidean data structures such as molecular structures, point clouds, social networks, and natural
55 language (Zhang et al., 2019). GNNs are adaptable to any type of data structure organized as graphs, comprising *nodes* (i.e., data points) and *edges* (i.e., the connections between nodes), which make them suitable to manifest meshes of numerical simulations (Pfaff et al., 2021). GNNs make predictions by utilizing pairwise message-passing between nodes, wherein information exchange

occurs, updating individual node features through interactions with connected nodes. Inspired by the resemblance of the mesh structure in finite-element analysis to a graph structure, numerous studies have investigated the training of emulators of finite-element numerical simulations using GNNs (Fu et al., 2023; Shivaditya et al., 2022; Black and Najafi, 2022; Perera et al., 2022; Salehi and Giannacopoulos, 2022; Maurizi et al., 2022; Jiang and Chen, 2023).

In this study, we investigate whether GNNs can be used as the backbone architecture for statistical mapping between input and output variables of finite-element ice sheet modeling, while embedding the spatial connections between nodes. GNN emulators take direct advantage of unstructured meshes of the Ice-sheet and Sea-level System Model (ISSM), ~~such as~~ allowing flexible spatial resolution and efficient allocation of computational resources, and can speed up model parameter search. Here, as an illustration, we use GNN emulators to determine how calving parameterization affects ice dynamics, while maintaining the accuracy and precision of ISSM simulations. Our focus is on Helheim Glacier, chosen as the target site for training and evaluating our GNN emulators (Fig. 1a). Helheim Glacier is one of the largest outlet glaciers in Greenland, and it has been shown that its ice velocity is closely correlated to the position of its terminus (Cheng et al., 2022). Despite considerable advancements in observation and modeling techniques over the past decades, the mechanisms governing Helheim Glacier's ice front position remain elusive (Bevan et al., 2015; Miles et al., 2016; Cheng et al., 2022). A key contributing factor to this knowledge gap is the computational intensity of numerical models incorporating calving, impeding the fine-tuning of calving parameters to accurately reflect observations. As the main parameter that determines the terminus positions, the temporal variations of calving parameters and their impacts on calving should be explored (Downs et al., 2023). In this study, we train GNN models using simulation data derived from numerical models and evaluate their fidelity and computational efficiency in modeling the dynamics and calving front migration of Helheim Glacier. We assess the potential of GNN architectures as statistical emulators for numerical finite-element ice sheet models to represent spatial features of ice sheet dynamics and calving across the entire glacier domain.

The remainder of the paper is structured as follows. Section 2 provides a comprehensive review of relevant literature about calving parameterization in ice sheet models, other machine-learning emulators employed in ice sheet modeling, and the use of GNNs as emulators for finite-element models. In Section 3, we describe the training data collected from numerical simulations, while Section 4 outlines the specifics of GNN architectures. Section 5 presents the accuracy and computational efficiency of these models in replicating the dynamics of Helheim Glacier. In addition, we demonstrate the utility of GNN emulators in optimizing calving parameterization to match simulation results with observations. Finally, we discuss the scientific implications of calving calibration and the limitations of GNN emulators in section 6.

2 Background

2.1 von Mises calving law

In the VM calving law, the calving rate, c, is assumed to be proportional to both the tensile stress and the magnitude of ice velocity, formulated as follows (Morlighem et al., 2016):

$$c = ||\boldsymbol{v}|| \frac{\tilde{\sigma}}{\sigma_{\max}} \qquad (1)$$
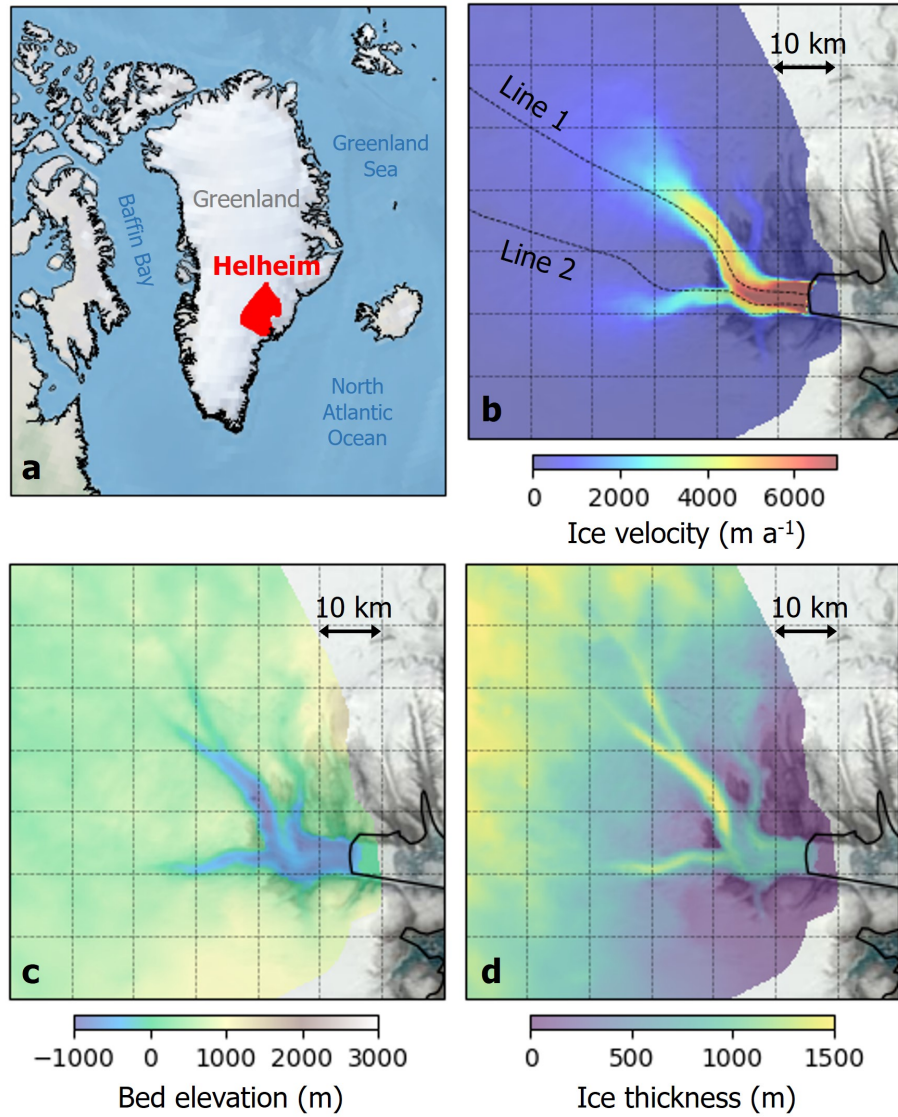
Figure 1. (a) Location of Helheim Glacier, Greenland; (b) ice velocity, (c) bed elevation, and (d) ice thickness near the calving front of Helheim glacier in 2007. The background image is the MODIS true color image from June 20, 2023.

where $\tilde{\sigma}$ is a scalar quantity representing the effective tensile stress of the ice, $\sigma_{\max}$ is a stress threshold that needs to be calibrated on a glacier-by-glacier basis, and $\boldsymbol{v}$ is the ice flow velocity at the ice front. The migration rate of the ice front is then determined from the calving rate c using the following equation (Morlighem et al., 2016):

$$\boldsymbol{v}_f = \boldsymbol{v} - (c + \dot{M})\boldsymbol{n} \tag{2}$$

where $\boldsymbol{v}_f$ is the ice front migration rate, $\dot{M}$ is the melting rate on the calving front, and $\boldsymbol{n}$ is a unit normal vector pointing outward from the ice domain. In the ISSM, the numerical ice sheet model we use in this study, moving boundaries of ice sheets are represented implicitly using the level set method (Osher and Sethian, 1988; Bondzio et al., 2016; Morlighem et al., 2016; Cheng et al., 2024). The level set method defines a scalar field $\varphi(\boldsymbol{x}, t)$: φ is negative in the location of $\boldsymbol{x}$ that are ice covered at time t, positive where there is no ice, and the zero contour of φ defines the ice boundary. With a given initial condition $\varphi(\boldsymbol{x}, t_0)$, the level set φ is updated by solving an advection equation using the migration rate $\boldsymbol{v}_f$ that

$$\frac{\partial \varphi}{\partial t} + \boldsymbol{v}_f \cdot \nabla \varphi = 0. \tag{3}$$

In VM, $\sigma_{\max}$ is the only parameter that needs to be calibrated (Morlighem et al., 2016). A lower value of $\sigma_{\max}$ correlates to weaker ice and consequently, a larger calving rate. According to Morlighem et al. (2016), $\sigma_{\max}$ is consistent with the range of ice tensile strength, which is typically around 1 MPa, but may be as low as 0.7 MPa and not exceed 3 MPa. In most of the previous studies, this $\sigma_{\max}$ value is simply assumed to remain constant over time within a predefined range. However, environmental conditions around the ice front (e.g., temperature, mélange, sea ice, ocean wave, tide, etc.) can vary over time, potentially influencing this calving parameter in response and leading to substantial changes in calving patterns (de Juan et al., 2010; Todd and Christoffersen, 2014; Kondo and Sugiyama, 2023; Xie et al., 2019; Wehrlé et al., 2023). Therefore, fine calibration of $\sigma_{\max}$ is essential for accurately reproducing observed glacier evolution.

2.2 Machine learning emulator for ice sheet modeling

Machine learning techniques have been extensively employed as statistical emulators for numerical ice sheet models and further for calibration of model parameters. Tarasov et al. (2012) applied Markov Chain Monte Carlo methods to artificial neural networks (ANNs) for the Bayesian calibration of model parameters. This approach allows for obtaining posterior distributions for different model parameters given observational datasets. Similarly, Brinkerhoff et al. (2021) used an ANN surrogate model to infer the posterior distributions for parameters that govern the behavior of sliding law and hydrological model. Downs et al. (2023) employed the GP for the calving parameterization associated with the VM calving law in Helheim Glacier. Nevertheless, the ANNs of these previous studies are simple feed-forward networks without spatial embeddings between nodes. Given that ice dynamics are determined by the topographical features, it is essential for the surrogate models to fully capture the interactions between neighboring elements to accurately represent realistic ice sheet behavior.

In replicating spatial characteristics of glaciers, convolutional neural networks (CNN) have been predominantly adopted as the main architectures based on their advantages in capturing the spatial variations in topographical features of glaciers (Jouvet et al., 2022; Jouvet and Cordonnier, 2023; Jouvet, 2023; Verjans and Robel, 2024). For instance, the CNN of the Instructed Glacier

Model (IGM) (Jouvet et al., 2022) reproduced the ice dynamics from the Parallel Ice Sheet Model (Winkelmann et al., 2011, PISM) and CfsFlow models (Jouvet et al., 2008). Jouvet (2023) extended this CNN emulator to address inversion problems, inferring optimal ice thickness distribution, ice flow velocity, and ice surface elevation to match both a Stokes model and observational data. Another CNN emulator introduced by Jouvet and Cordonnier (2023) employed a physics-informed loss function to minimize the energy associated with ice-flow equations during training. However, CNNs cannot take full advantage of finite-element ice sheet models on their native grid because they rely on regular grids (Zhang et al., 2019). Given the GNN's capacity to account for dynamic interactions between nodes and edges (Satorras et al., 2022), it can be a promising tool for predicting the dynamic behavior of ice. Indeed, Koo and Rahnemoonfar (2025) demonstrated the fidelity and computational efficiency of GNN architecture in emulating a finite-element ice sheet numerical model. This capability of GNN in preserving the mesh structure from the numerical models can have significant advantages in accurately describing ice front migration.

2.3 Graph neural networks for finite-element analysis

GNNs have been broadly used to emulate finite-element numerical models due to the similarities between ~~a computational mesh~~ computational meshes and graphs. Perera et al. (2022) developed a GNN framework to simulate fracture and stress evolution in brittle materials, training their model with data generated from a finite-element method fracture solver. Similarly, Shivaditya et al. (2022) proposed a GNN surrogate model for finite-element simulations of metal forging processes, demonstrating superior performance compared to other machine learning models and achieving a tenfold reduction in processing time. Salehi and Giannacopoulos (2022) also developed PhysGNN, a GNN framework tailored for simulating soft tissue deformation, and Maurizi et al. (2022) utilized GNNs to predict stress, strain, and deformation across various material systems, including fiber and stratified composites, and lattice metamaterials. Fu et al. (2023) proposed a boundary-oriented graph embedding (BOGE) approach within the GNN framework for solving finite-element cantilever beam problems, incorporating both boundary elements and local neighbor elements. Jiang and Chen (2023) introduced a novel graph attribute representation for triangular meshes in finite-element von Mises stress problems, effectively capturing geometry and boundary conditions to mitigate over-smoothing issues associated with deep GNNs. Black and Najafi (2022) introduced a multi-fidelity GNN for the cantilever beam problem, leveraging low-fidelity projections to inform high-fidelity modeling across arbitrary subdomains of subgraphs. However, despite the prevalence of finite-element analysis in ice sheet modeling (Larour et al., 2012; dos Santos et al., 2021; Gagliardini et al., 2013), to the best of our knowledge, GNNs have yet to be adopted exclusively for modeling calving dynamics and calibrating parameters involved.

3 Data

3.1 Ice Sheet numerical simulation

To generate training datasets for the GNN emulators, we conduct transient simulations of ice dynamics and calving of Helheim Glacier between 2007 and 2020, using ISSM (Larour et al., 2012). The Shelfy-Stream Approximation (SSA; MacAyeal, 1989)

is used for describing ice flow. The SSA, which assumes depth-independent horizontal velocity and negligible vertical shear stresses, is appropriate for fast-flowing glaciers controlled by basal sliding such as Helheim Glacier (Cheng et al., 2022; Choi et al., 2018).

The model setup is identical to the one described in Cheng et al. (2022). A two-dimensional unstructured mesh is constructed with a spatial resolution ranging from 100 m in the fast-flowing ice front to 1,500 m in the inland domain, ultimately comprising 46,434 elements and 23,466 vertices (nodes). The transient simulations run forward in time with a time step of 1.825 days (0.005 years), and we output the state of the model every 10 time steps ($\sim$18 days). Consequently, each transient simulation generates a total of 261 outputs between 2007 and 2020. Basal friction is calibrated and held constant using the surface velocities from satellite interferometry (Mouginot et al., 2017, 2019) (Fig. 1b). Bed topography and the initial ice thickness are from BedMachine Greenland v6 (Morlighem et al., 2017) (Fig. 1c and 1d). Surface mass balance (SMB) is from the Regional Atmosphere Model (Tedesco and Fettweis, 2020), and the ocean thermal forcing is from Wood et al. (2021). The melting rate at the calving front (i.e., $\dot{M}$ in Eq. 2) is parameterized based on Rignot et al. (2016). To examine the sensitivity of ice dynamics to $\sigma_{\max}$ of the VM calving law, we run transient solutions for 9 different $\sigma_{\max}$ values (i.e., 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.00, 1.05, and 1.10 MPa) based on the values proposed by Choi et al. (2018).

The ISSM simulations provide the solutions of ice velocity, ice thickness, and a mask of ice-covered region every 10 time steps. We convert the triangular mesh from ISSM into a data structure that aligns with the input and output requirements of deep learning architectures. Specifically, for GNN architectures, we convert the meshes into graph nodes and edges by extracting adjacency matrices that represent the connectivity between nodes. In a triangular mesh, each element consists of three nodes that are interconnected by edges (Fig. 2). Using the nodes and elements of the mesh ensures that the resolution of this graph matches exactly with the finite-element mesh used in ISSM simulations.

3.2 Observations

To determine the best $\sigma_{\max}$ that aligns with real observations, we compare the numerical simulation results with remote-sensing derived terminus positions and ice velocities for the same periods as the numerical simulations (2007-2020). First, we use surface ice velocities with a spatial resolution of 150 m (Mouginot et al., 2017, 2019), which have a precision better than 20 m a^{-1} (Mouginot et al., 2017). Second, we use a time series of calving front positions of Helheim Glacier from Greene et al. (2024). During the targeted time frame from 2007 to 2020, we use monthly averaged ice front positions, yielding a total of 156 distinct ice front positions for analysis.

4 Method

Ice sheet modeling can be regarded as a node-regression problem within graph structures, where the output features of individual nodes are derived from the input features of nodes. The unstructured meshes of ISSM can be represented as graph structures, with node connectivity expressed via adjacency matrices. Based on the graph structures of the ISSM meshes, we develop three GNN architectures: graph convolutional network (GCN), graph attention network (GAT), and equivariant graph convolutional
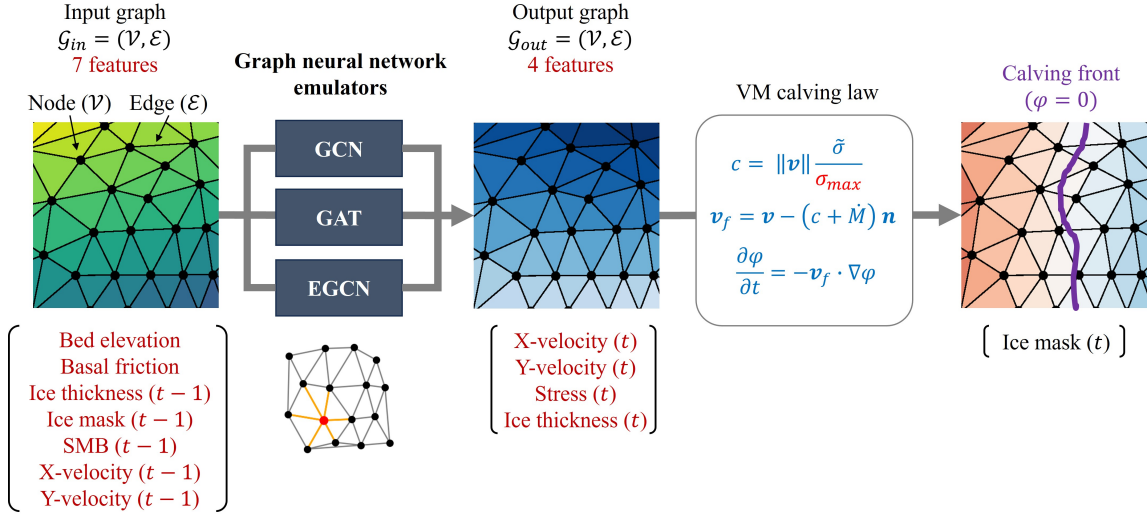
Figure 2. Schematic workflow of using GNN emulators to predict calving front.

network (EGCN). Typical GNN architectures update graph nodes iteratively through message-passing processes between neighboring nodes, and the way to achieve this message-passing determines the specific type of GNN architecture. For the undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with N nodes $v_i \in \mathcal{V}$, edges $(v_i, v_j) \in \mathcal{E}$, and an adjacency matrix $A \in \mathbb{R}^{N \times N}$, lth GNN layer receives a set of node features $\mathbf{h}^{(l)} = \{h_1^{(l)}, h_2^{(l)}, ..., h_N^{(l)}\}, h_i^{(l)} \in \mathbb{R}^{F_l}$, as the input and produces a new set of node features, $\mathbf{h}^{(l+1)} = \{h_1^{(l+1)}, h_2^{(l+1)}, ..., h_N^{(l+1)}\}, h_i^{(l+1)} \in \mathbb{R}^{F_{l+1}}$, for the next $l + 1$th layer. F_l and F_{l+1} is the number of features in each node at lth layer and $l + 1$ layer, respectively. The GCN, GAT, and EGCN operate on graph structures but use different message-passing approaches in updating $\mathbf{h}^{(l+1)}$ from $\mathbf{h}^{(l)}$. By comparing three representative GNN architectures, we evaluate what approach is more effective in replicating ice sheet dynamics and calving from the ISSM simulations.

4.1 Graph Convolutional Network

First, we employ a GCN proposed by Kipf and Welling (2017). We design a GCN with one input layer, five graph convolutional hidden layers, and one output layer (Fig. 2). The number of hidden layers is determined after conducting trial and error experiments with several options: we tested 1, 2, 5, and 10 hidden layers, and 5 hidden layers showed the best accuracy. The graph convolutional hidden layers are inspired by the localized first-order approximation of spectral graph convolutions on graph-structured data (Kipf and Welling, 2017). For each graph convolutional layer, the number of features is set to 128. Similar to the hidden layers, the number of hidden features was determined from trial and error with 4 options: 32, 64, 128, and 256.

The weights of graph convolutional layers are updated via the layer-wise propagation rule as follows:

$$h_i^{(l+1)} = \sigma \left(\sum_{j \in \mathcal{N}(i)} \frac{1}{c_{ij}} \mathbf{W}^{(l)} h_j^{(l)} \right) \tag{4}$$

where $\mathcal{N}(i)$ is the set of neighbors of node i, c_{ij} is an appropriately chosen normalization constant for the edge (v_i, v_j) defined as the product of the node degrees (i.e., $c_{ij} = \sqrt{|\mathcal{N}(j)|} \sqrt{|\mathcal{N}(i)|}$), and $\mathbf{W}^{(l)} \in \mathbb{R}^{F_{l+1} \times F_l}$. $\mathbf{W}^{(l)}$ is a layer-specific trainable weight matrix ($\mathbf{W}^{(l)} \in \mathbb{R}^{F_{l+1} \times F_l}$), and $\sigma(\cdot)$ is an activation function; we use the Leaky ReLU activation function with 0.01 negative slope in this study.

4.2 Graph Attention Network

Since the original GCN filters merely depend on graph structures and node connectivity (Eq. 4), a model trained with a certain graph structure can have limitations in general applicability to different graph structures. The GAT is proposed to address such shortcomings by adding masked self-attention layers (Veličković et al., 2018). This architecture assigns different weights to different nodes in a neighborhood by inserting a self-attention mechanism in a hidden layer, which can allow better generalizability for different ice conditions or topographies. The propagation process in graph attention layers can be expressed by the following equation (Veličković et al., 2018):

$$h_i^{(l+1)} = \sigma \left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(l)} \mathbf{W}^{(l)} h_j^{(l)} \right) \tag{5}$$

where $\alpha^{(l)}$ is the attention score between node i and node j defined as follows:

$$\alpha_{ij}^{(l)} = \text{softmax}_j \left(e_{ij}^{(l)} \right) = \frac{\exp(e_{ij}^{(l)})}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik}^{(l)})}$$

$$e_{ij}^{(l)} = a \left(\mathbf{W}^{(l)} h_i^{(l)}, \mathbf{W}^{(l)} h_j^{(l)} \right) \tag{6}$$

where $a : \mathbb{R}^{F_{l+1}} \times \mathbb{R}^{F_{l+1}} \rightarrow \mathbb{R}$ is a self-attention mechanism to compute attention coefficient $e_{ij}^{(l)}$. This attention mechanism a is a single-layer feedforward neural network parameterized by a weight vector $\mathbf{a} \in \mathbb{R}^{2F_{l+1}}$, normalized by LeakyReLU function afterward. The graph structure is applied to this attention mechanism by computing $e_{ij}^{(l)}$ for only nodes $j \in \mathcal{N}(i)$ where $\mathcal{N}(i)$ is set of neighbors of node i. We execute three independent attention mechanisms of Equation 5 and average them for the final graph attention layer (Veličković et al., 2018). Similar to the GCN, the GAT consists of one input layer, five graph attention hidden layers with 128 features, and one output layer (Fig. 2).

4.3 Equivariant Graph Convolutional Network

Another graph neural network we adopt is EGCN, which is designed to conserve equivariance to rotations, translations, reflections, and permutations in a graph structure (Satorras et al., 2022). Since our emulator is intended to predict ice front

migration, we anticipate that the preservation of equivariance to rotations and translations on spatial coordinates via the EGCN structure guarantees sufficient generalizability to various graph structures of dynamics systems. An equivariant graph convolutional layer can be expressed by the following equations:

$$m_{ij} = \phi_e(h_i^{(l)}, h_j^{(l)}, ||x_i^{(l)} - x_j^{(l)}||^2, a_{ij}) \tag{7}$$

$$v_i^{(l+1)} = \phi_v(h_i^{(l)})v_i^0 + C \sum_{j \neq i}(x_i^{(l)} - x_j^{(l)})\phi_x(m_{ij}) \tag{8}$$

$$x_i^{(l+1)} = x_i^{(l)} + v_i^{(l+1)} \tag{9}$$

$$m_i = x_i^{(l)} + C \sum_{j \neq i} m_{ij} \tag{10}$$

$$h_i^{(l+1)} = \phi_h(h_i^{(l)}, m_i) \tag{11}$$

where a_{ij} is the edge attributes, x_i and x_j are the 2D coordinate embeddings for node i and j, respectively, and C is a constant for normalization computed as $1/|\mathcal{N}(i)|$. For the edge attributes, we use five attributes that can be extracted from the connecting nodes: distance, surface slope, base slope, and gradient of the x and y components of the ice velocity. ϕ_e, ϕ_x, and ϕ_h are the edge, position, and node operations, respectively, which are approximated by single-layer neural networks with 128 hidden features. Herein, the x- and y-components of ice velocities are represented as velocity embeddings (v_i), which induce the displacement of coordinate embeddings (x_i) (Eq. 9), and ice thickness is represented as hidden embeddings ($h_i^{(l)}$). More details about the EGCN architecture are described in Satorras et al. (2022).

4.4 Model training

From the ice sheet simulation results, we collect 2,349 graphs (261 sets of results per transient simulations $\times$ 9 $\sigma_{\max}$ values). Each graph consists of 23,466 nodes and 139,798 edges. Individual nodes have numerical input and output features from the ISSM simulations and are connected to each other via edges, which are determined by the node connections of the finite elements. The input features include 7 variables that determine the ice sheet dynamic conditions: bed elevation, basal friction, ice thickness at time $t-1$, ice mask at time $t-1$, surface mass balance at time $t-1$, and x and y components of ice velocity at time $t-1$. The output includes 4 variables that represent ice dynamics at time t: the two components of ice velocity, ice thickness, and tensile stress $\tilde{\sigma}$ (Fig. 2). Our emulators are trained to establish a statistical mapping between input and output

features: how the input variables at time $t-1$ change the output variables at time t in numerical modeling. Previous machine learning studies attempted similar types of mapping via various deep learning architecture (Brinkerhoff et al., 2021; Downs et al., 2023; Tarasov et al., 2012; Jouvet and Cordonnier, 2023; Jouvet et al., 2022; Jouvet, 2023). All input and output features are normalized to the range [-1, 1] for stable learning using the nominal maximum and minimum values of these variables. After the GNNs predict ice velocity, ice thickness, and tensile stress $\tilde{\sigma}$, the ice mask at t is computed from these outputs using a given σ_{max} using Equations 1, 2, and 3 (Fig. 2).

All 2,349 graph structures are divided into training and test datasets based on the σ_{max} values to assess if our emulators can be generalized for different σ_{max} values. We conduct four sets of sensitivity tests to evaluate whether the emulator can be reliably generalized to unseen σ_{max} values (Table 1). In set A, the training dataset consists of samples with σ_{max} values of 0.75, 0.90, and 1.05 MPa, while the remaining σ_{max} values are used for testing. This set evaluates whether the model can represent the full range of σ_{max} values when trained with evenly sampled σ_{max} values. In set B, the training dataset includes low σ_{max} values (0.70, 0.75, and 0.80 MPa) to test whether the model can generalize to high-value samples when trained only with low-value samples. In set C, we use medium σ_{max} values (0.85, 0.90, and 0.95 MPa) as the training dataset to check whether the model can represent high or low-value samples when trained only with medium-value samples. Finally, set D contains only high σ_{max} values (1.00, 1.05, and 1.10 MPa) to check the model performance when extrapolating to lower values. For each sensitivity test, the training and test datasets contain 783 and 1,566 graph structures, respectively. The model is optimized using Adam stochastic gradient descent algorithm with the mean square error (MSE) loss function, over 500 epochs, and a learning rate of 0.001.

Table 1. Division of training/test σ_{max} in four different sets.

Set	Training σ_{max} (MPa)	Test σ_{max} (MPa)	Note
A	[0.75, 0.90, 1.05]	[0.70, 0.80, 0.85, 0.95, 1.00, 1.10]	Even-value samples
B	[0.70, 0.75, 0.80]	[0.85, 0.90, 0.95, 1.00, 1.05, 1.10]	Low-value samples
C	[0.85, 0.90, 0.95]	[0.70, 0.75, 0.80, 1.00, 1.05, 1.10]	Medium-value samples
D	[1.00, 1.05, 1.10]	[0.70, 0.75, 0.80, 0.85, 0.90, 0.95]	High-value samples

4.5 Model evaluation

We evaluate the ability of our emulators to reproduce ice velocity, ice thickness, and calving front migration by comparing them to ISSM simulation results. For this evaluation, we calculate two metrics: (i) root mean square error (RMSE) and (ii) binary calving front error (BE); RMSE is used to evaluate ice velocity and ice thickness predictions, and BE is used to evaluate the calving front delineation. These metrics are calculated using the following equations:

$$\mathrm{RMSE}(\hat{y}, y) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2} \qquad (12)$$

$$\text{BE}(\hat{y}, y) = 1 - \frac{1}{N} \sum_{i=1}^{N} \mathbf{I}(\hat{y}_i = y_i) \tag{13}$$

where $\hat{y}$ denotes predicted values, y denotes true values, N is the number of data points, and $\mathbf{I}(\hat{y}_i = y_i)$ denotes the indication function that returns 1 if $\hat{y}_i = y_i$ and returns 0 if $\hat{y}_i \neq y_i$. RMSE represents the difference between the prediction and reference; lower RMSE corresponds to better fidelity. In calculating BE, we convert the ice mask into binary values, representing ice nodes as 1 and non-ice nodes as 0. Therefore, BE is a metric indicating the proportion of incorrectly predicted nodes relative to the total number of nodes; lower BE corresponds to better fidelity.

5   Results

5.1   Model Fidelity

Our GNN models are trained with four different training/test sets, as outlined in Table 1. The training and test errors of ice velocity, ice thickness, and calving front position for the GNN emulators compared to ISSM simulation results are shown in Figure 3. When the GNN emulators are trained with even-value samples (set A), the training and test errors are almost the same for ice velocity, ice thickness, and calving front. Indeed, t-test results confirm no significant difference between training and test errors for all models (Table A1). This result indicates that the emulators trained with even-value samples can effectively represent the full range of $\sigma_{\max}$ values.

In the case of low-value samples (set B), the test error is occasionally even lower than the training error, particularly in ice velocity prediction, which suggests that the emulators trained only with low-value samples can represent high $\sigma_{\max}$ value cases with $< 10\ \%$ of uncertainty in most cases (Table A1). Although the test errors are slightly greater than training errors in sets C and D, their differences are primarily within a 5 % uncertainty range for ice velocity and ice thickness and within a 20 % uncertainty range for the calving front (Table A1). The results of sets C and D demonstrate that the GNN emulators trained with medium or high $\sigma_{\max}$ values can generally have $< 10\text{-}20\ \%$ of uncertainty in representing the out-of-sample $\sigma_{\max}$ values. Overall, the sensitivity tests with four different sets show that GNN emulators trained with selected $\sigma_{\max}$ values can be generalizable for all other $\sigma_{\max}$ values with $< 10\text{-}20\ \%$ of uncertainties even if those $\sigma_{\max}$ values are not included in the training dataset. While the GP-based emulator by Downs et al. (2023) was limited to interpolating within the range of training samples, our GNN emulators can extrapolate to unseen $\sigma_{\max}$ values. This capability in extrapolation is particularly valuable for optimizing calving parameterization beyond the scope of training $\sigma_{\max}$ values.

Among the three GNN architectures, the EGCN exhibits notably lower RMSE values: $\sim$50-60 m a^{-1} of ice velocity RMSE and $\sim$30 m of ice thickness RMSE (Fig. 3). That is, the EGCN reduces ice velocity RMSE by 40-150 m a^{-1} and ice thickness RMSE by 15-20 m compared to the other GNN architectures (see Table A2 for statistical significance). On the other hand, the EGCN and GAT yield similar ice front predictions with BE ranging from 0.60 to 0.75 %. The relatively lower errors achieved by the EGCN support that its equivariance architecture is beneficial in predicting ice dynamics. As mentioned in section 4,
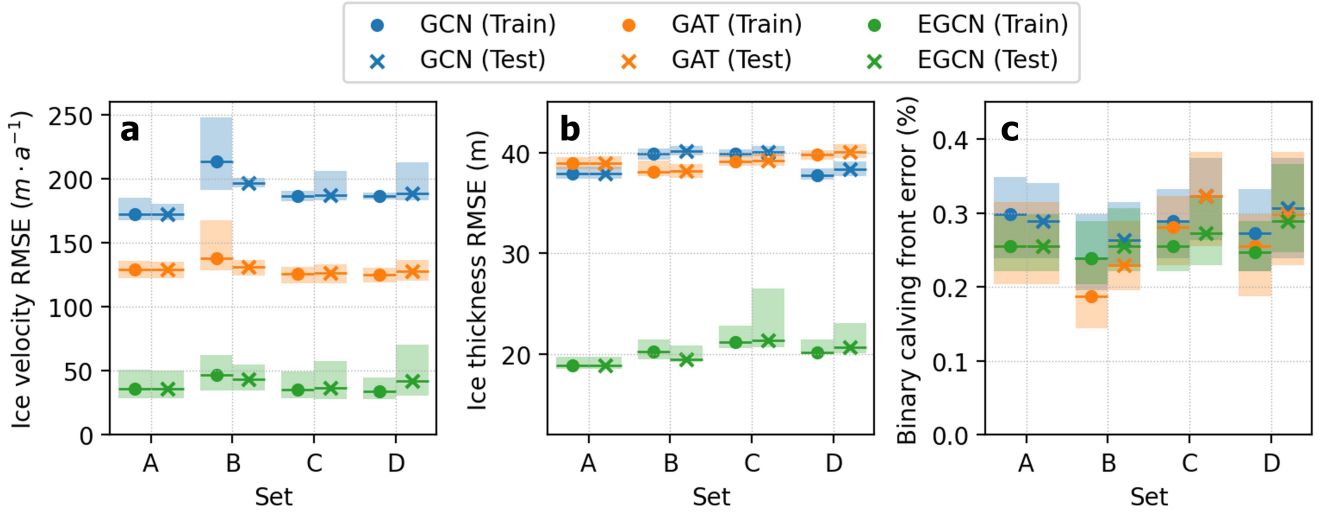
Figure 3. (a) RMSE of ice velocity, (b) RMSE of ice thickness, and (c) binary calving front error (BE) for four different sets. For each set, nine different σ_{max} datasets are divided into training (three σ_{max} values) and test (six σ_{max} values) datasets as described in Table 1. The circle markers indicate the median errors from training datasets, and the crossing markers indicate the median errors from test datasets. The shaded areas indicate the interquartile ranges (i.e., 25 % quantile to 75 % quantile) of errors.

the GCN, GAT, and EGCN have distinct characteristics in message-passing approaches: the GCN simply uses the adjacency status between neighboring nodes to determine the weights during the propagation process (Eq. 4); the GAT uses additional self-attention mechanisms to evaluate the relative importance of neighboring nodes (Eq. 6); the EGCN uses the message passing from all nodes to preserve the equivariance of the entire graph (Eq. 7, 8, and 9). We conjecture that the equivariance architecture

310 of the EGCN contributes significantly to the improvement of model fidelity.

Figures 4 and 5 show the maps of ice velocity and thickness on two σ_{max} test cases of 0.70 MPa and 1.10 MPa predicted by the GNN emulators trained with the training set A. Overall, all GNN emulators replicate the spatial and temporal patterns of ice velocity and thickness successfully. Additionally, they reproduce the trend of ice front retreat at lower σ_{max} and advance at higher σ_{max}. In terms of ice velocity (Fig. 4), the EGCN shows the lowest errors of approximately < 100 m a^{-1} across the

315 glacier domain, followed by GAT and GCN. In terms of ice thickness (Fig. 5), the EGCN generally shows an average RMSE $<$ 40 m across the glacier domain, but the GCN and GAT have higher RMSEs of around > 50 m in some areas.

5.2 Computational time

The most significant advantage of GNN emulators is their ability to reduce this computation time by leveraging GPUs while maintaining finite element structures, which is not completely available via other non-graph neural network architectures (Koo

320 and Rahnemoonfar, 2025). When we run the ISSM numerical model, it takes 948 seconds ($\sim$15.8 minutes) on average to
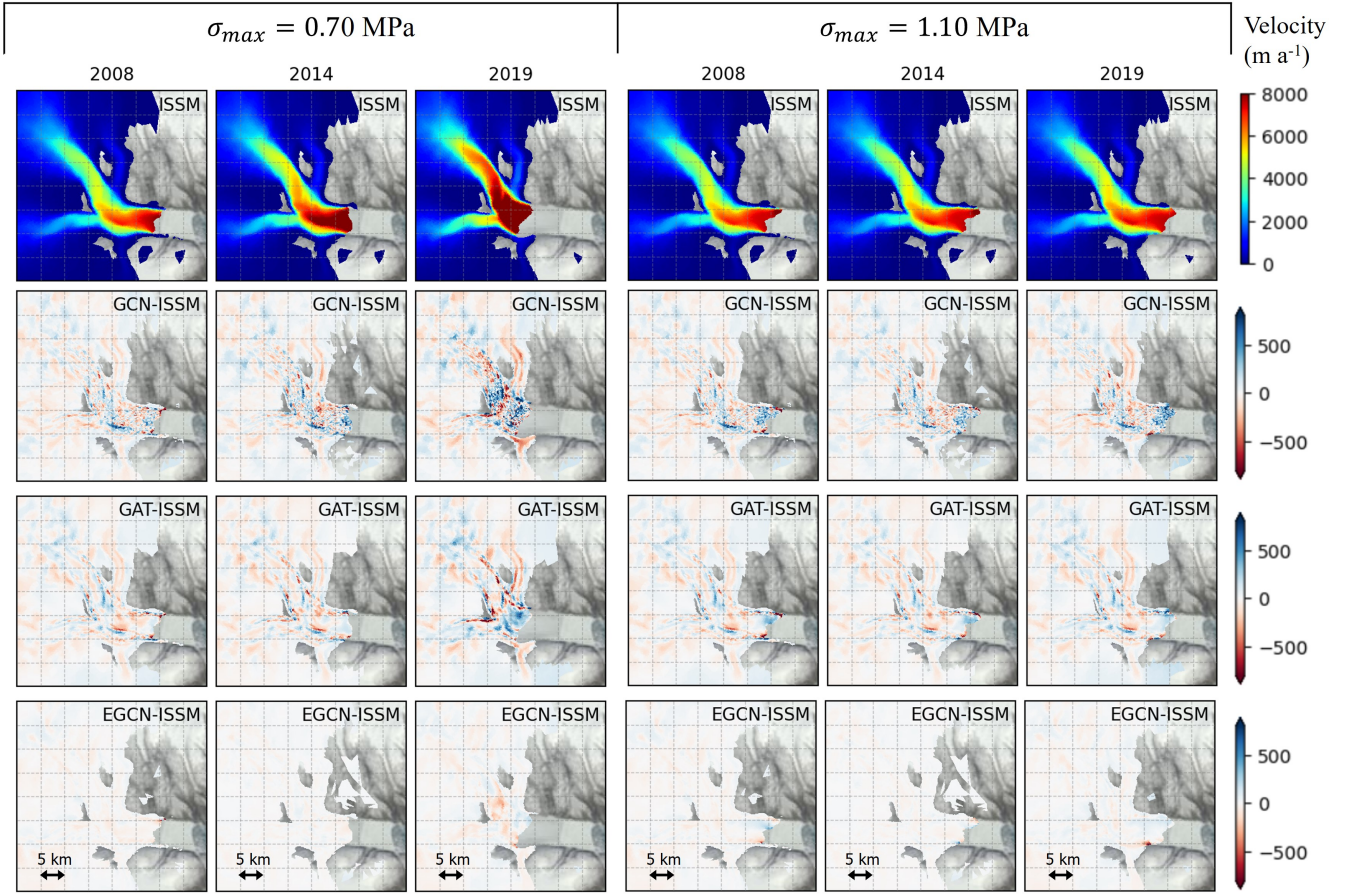
Figure 4. Maps of ISSM-simulated ice velocity, and difference of ice velocity between GNN emulators and ISSM simulations (GCN, GAT, and EGCN from top to bottom) for 0.70 MPa and 1.10 MPa test σ_{max} values of set A.

complete one 13-year transient simulation for a given σ_{max} value. This computation time of ISSM reflects the total elapsed time on a single node of the Texas Advanced Computing Center (TACC) Frontera supercomputing cluster, which is equipped with 56 cores of Intel 8280 Cascade Lake CPUs (192 GB memory). To evaluate the computational efficiency of the GNN emulators, we also record the time required to generate a single transient simulation for a given σ_{max} value with our GNN emulators (Table 2). The GNN computations are performed on a CPU (Intel(R) Core(TM) i7-11700F; 32 GB memory) and a GPU (NVIDIA GeForce RTX 3070; 24GB memory) within the same desktop (Lenovo Legion T5 26IOB6). We observe a dramatic speed-up with GPU-based deep learning emulators, achieving computation times 30-34 times faster than ISSM simulations. The GCN shows the highest speed-up, around 34 times faster than ISSM. Among the three GNN architectures, the GAT shows the longest computational time due to the complexity of its attention mechanisms. The EGCN also requires a longer computational time than the GCN because it operates on all graph nodes to preserve the equivariance in the entire graph, whereas the GCN operates
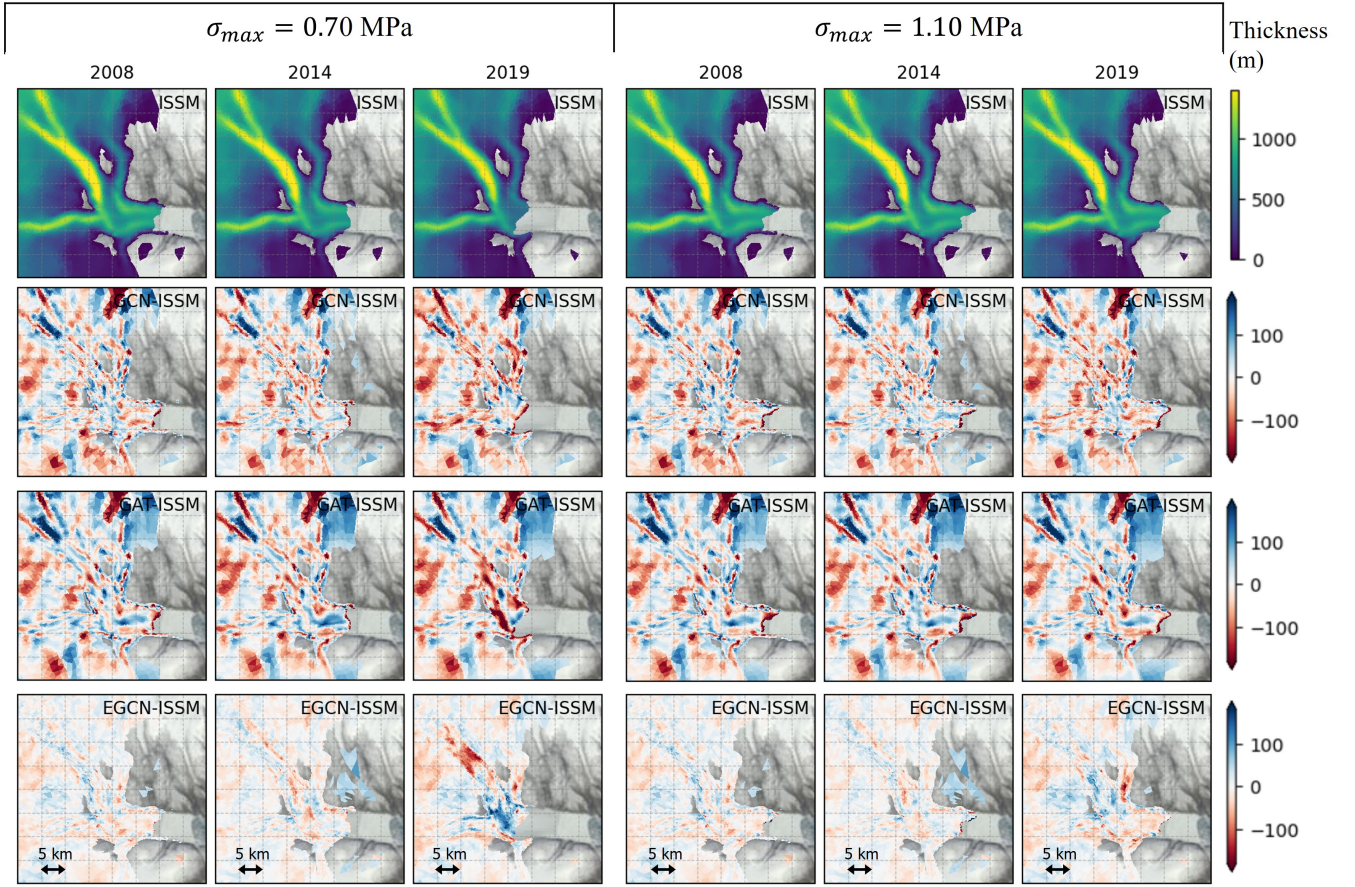
14

Figure 5. Maps of ISSM-simulated ice thickness, and difference of ice thickness between GNN emulators and ISSM simulations (GCN, GAT, and EGCN from top to bottom) for 0.70 MPa and 1.10 MPa test σ_{max} values of set A.

only on the adjacent nodes. It is also noteworthy that using GPUs reduces the computation time of deep learning emulators by up to ~4 times compared to using CPUs.

This experiment is also promising because, while ISSM computation is performed on a supercomputer, GNNs can be run on a personal desktop. By leveraging GPU-based GNN emulators, computationally intensive simulations of ice dynamics and calving can be quickly reproduced on personal desktops without the need for high-performance computing systems. We expect that these types of emulators will facilitate the efficient parameterization of different ice conditions, ice properties, and external climate forcings.

However, although GNNs can successfully replicate the finite-element structure of ISSM simulations and reduce the computational time by leveraging GPUs, it is important to consider the upstream costs for training deep learning emulators to assess their whole-process efficiencies. Table 2 presents the number of learnable parameters and model training time for the GCN, GAT, and EGCN. This training time is recorded from the 500-epoch training of each model on a multiple-GPU system

equipped with 8 NVIDIA RTX A5000 GPUs. The GCN takes the least training time, followed by EGCN and GAT. Herein, these training times are based on the training datasets of set A; however, the training times of four different sets (Table 1) are similar to each other because they have the same number of training samples. It should also be noted that collecting various

345 ISSM numerical simulations with various σ_{max} values is required to train GNN emulators: in our case, approximately 8,500 seconds (~2 hours 22 minutes) are required to collect all 9 simulation results with different σ_{max} values. Nevertheless, given that the emulators show sufficient performances only with 3 training σ_{max} values, fewer simulations would be sufficient for training. Furthermore, we emphasize that these GNN emulators can be efficiently applied to this target glacier with a short implementation time once they are trained, which can accelerate the search for optimal parameterization settings that align the

350 numerical model with observations.

Table 2. Computation time for running one 13-year transient simulation from ISSM and GNN emulators and upstreaming computational time for training deep learning models

Model	Averaged elapsed time for one transient run (speed-up compared to ISSM)		Training time	Number of learnable parameters
	CPU	GPU		
GCN	45 s ($\times$21.0)	**28 s ($\times$33.8)**	**617 s (10.3 min)**	67,588
GAT	135 s ($\times$7.0)	31 s ($\times$30.6)	1,315 s (21.9 min)	205,572
EGCN	65 s ($\times$14.6)	29 s ($\times$32.7)	814 s (13.6 min)	104,066

5.3 Comparison with real observations

We use the EGCN emulator, trained with set A training samples, to determine the appropriate σ_{max} values that accurately capture the migration of Helheim's calving front, as the EGCN shows the best performances in predicting ice velocity, ice thickness, and calving front positions (Fig. 3). The satellite-derived ice velocity and ice front observations at each time step are used as

355 input for the EGCN emulator to predict the ice front at the next time step. Then, we identify the optimal σ_{max} value for each time step that minimizes BE of ice front location between prediction and observation (Eq. 13). Given that previous studies (Morlighem et al., 2016; Wilner et al., 2023; Choi et al., 2018) suggest that σ_{max} should fall between 0.7 to 1.1 MPa, we vary σ_{max} incrementally by 0.01 MPa within this range and determine the σ_{max} value with the lowest BE.

 Figure 6 shows the temporal variation of optimal σ_{max} derived from the EGCN spanning from 2007 to 2020. Interestingly, there

360 is a significant temporal correlation between the optimal σ_{max} and the terminus positions with coefficient > 0.5 (p-value < 0.01). This result indicates that σ_{max} is an important parameter that determines the calving front migration and should be tuned carefully in calving models. However, we should also note that the temporal variations of optimal σ_{max} and observed calving front do not agree well during the summers of 2017 and 2019. Given that these two periods coincide with prolonged plume-polynya events
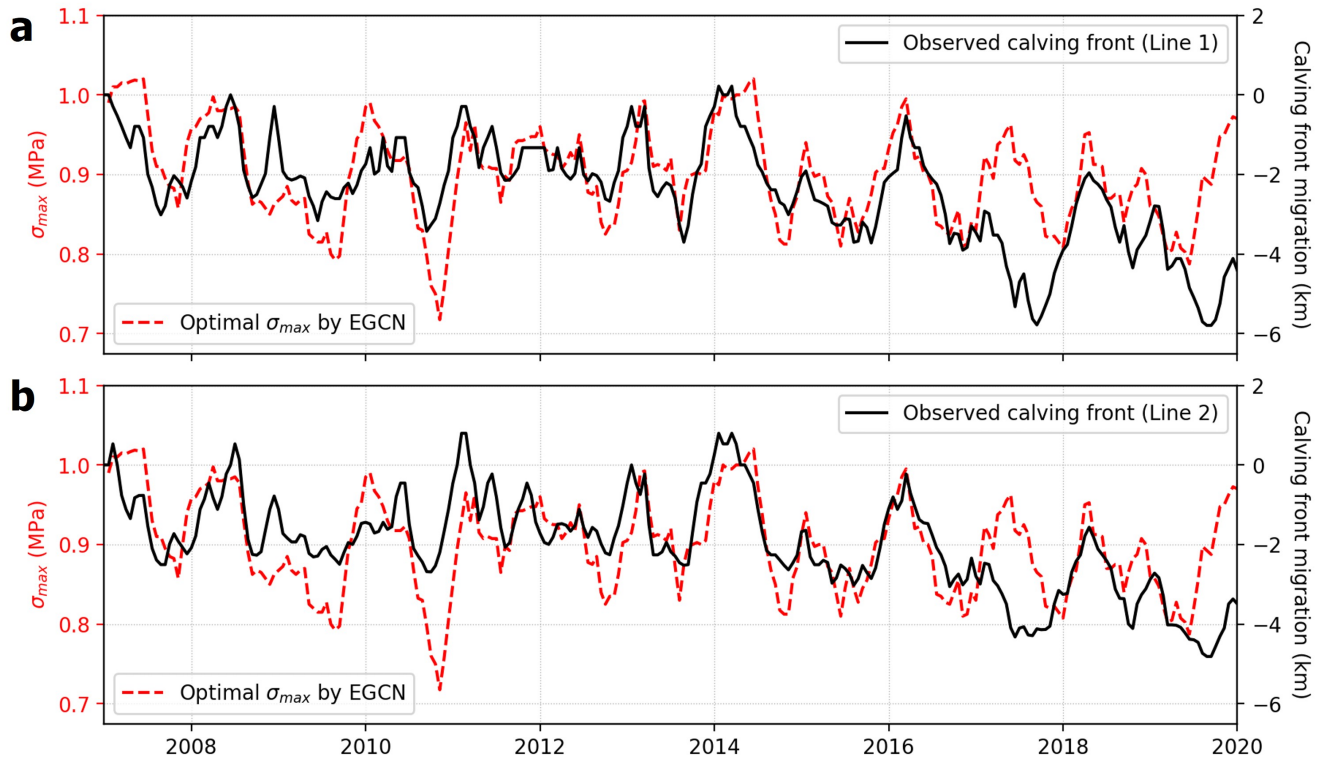
Figure 6. Temporal variations of optimal $\sigma_{\max}$ derived by EGCN (red dashed line) and the temporal movement of terminus of flow line 1 (a) and flow line 2 (b) in Figure 1b.

lasting more than 20 days in the Helheim Glacier (Melton et al., 2022), we conjecture that buoyant subglacial discharge plumes
365 may have influenced the calving mechanisms during these periods, introducing additional uncertainties to the VM calving law (Everett et al., 2021).

Year-to-year migrations of ice front from observations and numerical models with different $\sigma_{\max}$ are shown in Figure 7. It should be noted that a temporarily constant $\sigma_{\max}$ does not represent the observed calving front ~~mitigation~~ migration well. If the $\sigma_{\max}$ value is set too high (> 0.8 MPa), the calving front does not shift or slightly advances after 2017; on the contrary, if $\sigma_{\max}$ is
370 set too low (≤ 0.8 MPa), the calving front retreats too much such that the modeling results ~~do not accompany with~~ no longer match the observations. On the other hand, the temporally varying optimal $\sigma_{\max}$ reproduces the long-term trend of calving front movement along the flow line 1 better than constant $\sigma_{\max}$ settings, although the short-term variations do not perfectly match (Fig. 8a). However, we should also note that the calving front migration of flow line 2 is not as well predicted as flow line 1 (Fig. 8b). Although some small advance or retreat along flow line 2 is predicted by numerical simulations with the optimal $\sigma_{\max}$ settings,
375 the calving front remains almost the same for 13 years of simulation. This difference between flow line 1 and 2 suggests the
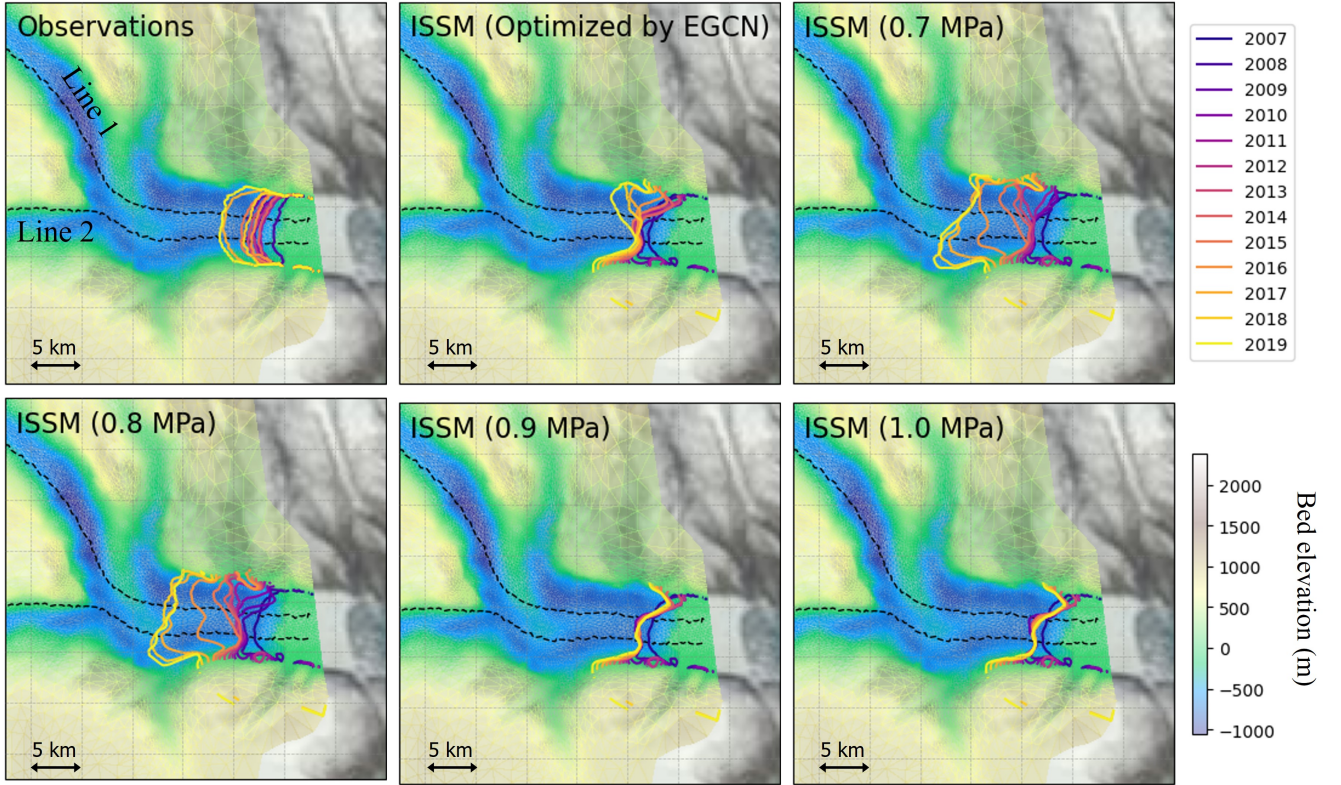
Figure 7. Annual movement of ice front observed by satellite imagery and modeled by ISSM with different σ_{max} parameterizations: optimal σ_{max} values from EGCN and constant σ_{max} values (0.7, 0.8, 0.9, and 1.0 MPa).

need for spatially varying σ_{max} parameterization in addition to the temporally varying settings, which is discussed in detail in section 6.

6 Discussion

In this section, we discuss the scientific implications of σ_{max} variation and the significant role of GNN emulators in improving the physical modeling of glacier calving. Our results reveal that the fine-tuning of temporally varying σ_{max} leads to better performance in calving prediction compared to traditionally used constant σ_{max} parameterization. As a parameter indicating how difficult ~~calving can~~ it is for calving to happen, a greater σ_{max} means a lower rate of calving, and a lower σ_{max} means a higher rate of calving. Previous studies have shown that the physical condition of mélange around the ice front, such as ice thickness and concentration, changes the frequency of calving events; a weaker or disappearing mélange area leads to an increase in calving events, which corresponds to lower σ_{max} (Wehrlé et al., 2023; Xie et al., 2019; Meng et al., 2025). Although this study
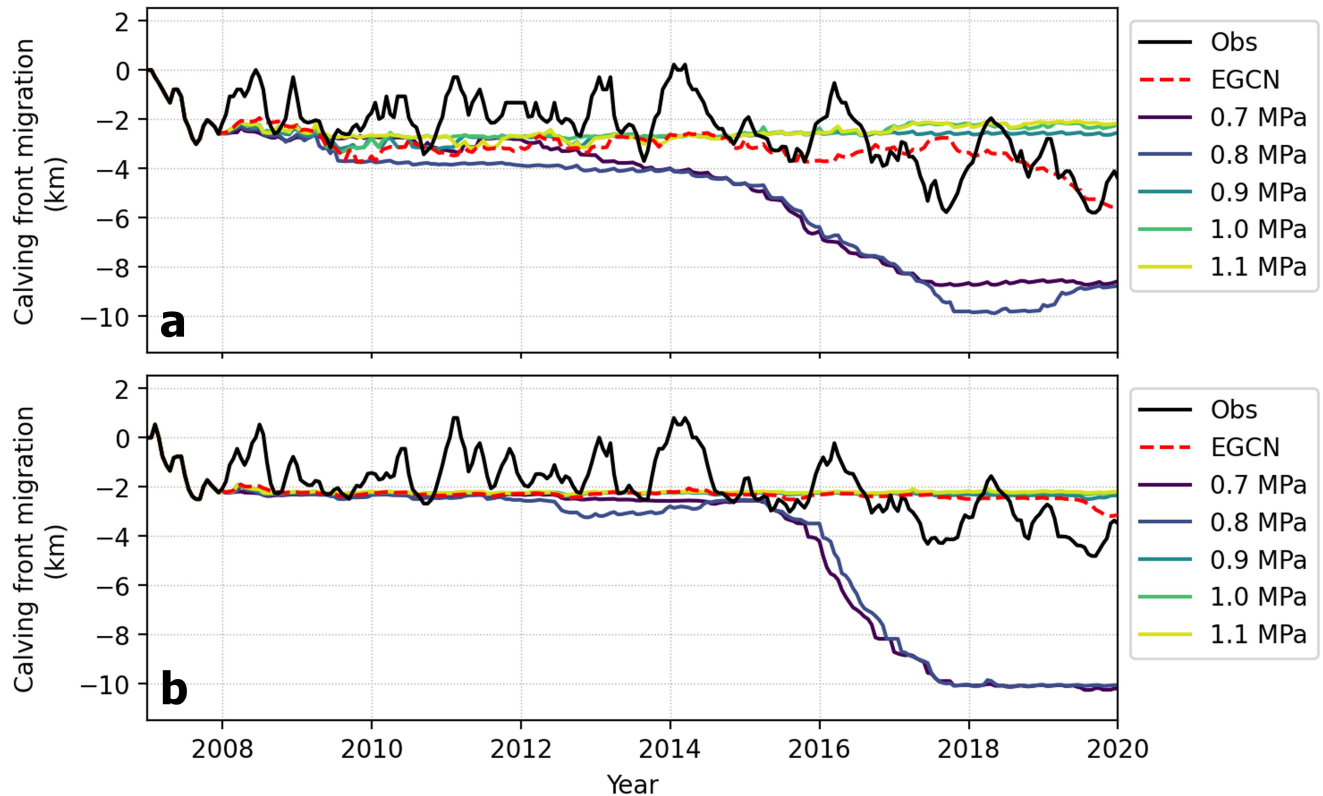
18

Figure 8. Movement of ice front location for (a) flow line 1 and (b) flow line 2 predicted by numerical models with different σ_{max} parameterizations; optimal σ_{max} values from EGCN and constant σ_{max} values (0.7, 0.8, 0.9, 1.0, and 1.1 MPa). The observed ice front movement is drawn as black solid line.

is focused on finding optimal σ_{max} values using GNNs, exploring the linkage between optimal σ_{max} and mélange conditions can contribute to the accurate parameterization of σ_{max} for calving models. Additionally, given that ~~this~~ mélange and sea ice conditions can vary along the ice front, assigning spatially varying σ_{max} can address the discrepancies in calving front migration for two flow lines shown in Figures 7 and 8. To set spatially varying σ_{max}, we can consider finding the best σ_{max} for multiple flow lines and interpolate these values at the ice front. It would also be valuable to investigate how σ_{max} interacts with other environmental factors, including tidal forcing and deep water circulation, which are known as significant drivers for calving events (O'Neel et al., 2003; Bassis and Jacobs, 2013; Slater and Straneo, 2022). Consequently, finding the optimal calving parameterization setting and connecting it to external environmental drivers will contribute to the accurate modeling of glacier mass loss and sea level rise beyond the conventional constant calving calibration.

19

Given the scientific implication of temporally varying σ_{max}, the significance of this study lies in the (i) fidelity and (ii) computational efficiency of using GNN architectures, especially EGCN, in emulating numerical ice sheet models operating on unstructured meshes to find the optimal calving parameterization. By applying GNNs directly on raw unstructured meshes, we can keep high resolution around the fast ice area to delineate the calving front accurately. In particular, EGCN shows the best fidelity in predicting ice thickness due to its equivariance concept throughout the graph structures. Moreover, the use of GNN emulators results in a dramatic increase in computational efficiency, with speeds 30-34 times faster than the ISSM simulations. This speed-up allows us to find the temporally varying optimal σ_{max} values quickly. Given that running numerical simulations every time to determine the optimal σ_{max} setting is extremely time-consuming, our fast GNN emulators reduce the workload associated with calving calibration.

Besides these advantages of GNN architectures, we emphasize that GNN emulators have significant potential for further improvement, particularly through integration with various architectures, including recurrent neural networks (RNNs) (Wu et al., 2021). For example, while our GNN emulator only implements a mapping between a previous time step $(t-1)$ to the next time step (t), embedding recurrent units into the GNN architecture can make it possible to find the sequential relationships between ice sheet dynamics and calving parameters (i.e., mapping between multiple time steps). This recurrent GNN architecture will be able to predict how the historical context of calving parameters affects ice dynamics and ice front migration.

However, it is also worth noting several limitations of our approach. Firstly, despite the computational efficiency of GNN emulators, they should be trained using numerical simulations. Since the model performance is highly dependent on the quality of training datasets, simulation data should be collected carefully with appropriate parameters. ~~Additionally~~Second, the collection of training datasets from various climatological scenarios can be helpful for better generalizability and reliability of emulators to predict future glacier behaviors; however, this process can be exceptionally time-consuming. ~~Additionally~~This generalizability issue can also be associated with the applicability of our GNN emulators for other glaciers beyond Helheim Glacier. To assess the fidelity and transferability of our GNN emulators trained on Helheim Glacier, we apply them to the Pine Island Glacier (PIG), Antarctica, which has different ice velocity and thickness distributions. The ice velocity RMSE in PIG ranges from 200 to 400 m a^{-1} (2-4 times greater RMSE than Helheim glacier), and the ice thickness RMSE ranges from 120 to 200 m (5-10 times greater RMSE than Helheim glacier) (Fig. S1, Fig. S2, and Table S1). These results suggest the importance of training GNN emulators on datasets encompassing a diverse range of ice velocity and thickness distributions to ensure broad generalizability and applicability across different glacier settings. Finally, although we employ the VM calving law to determine calving front migration, it is important to recognize that calving mechanisms are not yet fully understood and may be more complex than those represented by the VM law. While the VM method has been validated for many glaciers in Greenland and Antarctica (Choi et al., 2018; Wilner et al., 2023), detailed calving mechanisms remain elusive. Thus, although our GNN emulators can provide valuable insights into how to select appropriate σ_{max} for VM calving law, they rely on our current imperfect physical understanding of calving.

7 Conclusions

This study develops three graph neural networks – GCN, GAT, and EGCN – as surrogate models to reproduce finite-element ice dynamics and calving retrieved from the ISSM. After training these GNNs with 13-year transient simulations of Helheim Glacier, they demonstrate significant spatiotemporal agreement with ISSM simulations in predicting ice velocity, ice thickness, and ice front location. The GNN emulators successfully reproduce the retreat of the ice front for a lower calving stress threshold σ_{max}, as well as the stable condition of the ice front for a higher σ_{max}. Additionally, the GNN emulators show only $<$10-20 % of uncertainties in predicting out-of-sample σ_{max} values, implying a better capability in extrapolation and generalizability than previous statistical emulators. Among the three GNN architectures, the EGCN shows the best robustness in the prediction of ice thickness and velocity by preserving the equivariance of graph structures. By leveraging 30-34 times faster computational time of the GPU-based GNN emulators compared to numerical simulations, we efficiently find temporally varying optimal σ_{max} parameterization for the von Mises calving law. By applying satellite-derived observations to the EGCN, we find that the temporal variations of optimal σ_{max} have a significant correlation with the retreat and advance of the calving front from 2007 to 2020. Additionally, the numerical simulations based on the fine-tuned σ_{max} calibration show a better agreement with observations compared to conventional constant σ_{max} setting. Therefore, this result highlights the importance of setting σ_{max} values appropriately to improve the reliability of numerical models. As the first attempt to use GNNs for calibrating calving parameters, GNN emulators can further contribute to improving the prediction accuracy of ice sheet mass loss and resulting sea level rise.

Appendix A: T-tests for training and test datasets

We conduct statistical t-tests to examine (i) whether test error is significantly different from training error (Table A1) and (ii) whether a certain GNN architecture shows significantly better performance than others (Table A2).

Table A1. Reduction in test error relative to training error in percentage (%) for each training/test set and GNN emulator. The positive values mean lower error in test datasets than in training datasets. If the mean difference is not statistically significant (p-value > 0.01; i.e., statistically have the same mean), we highlight it in bold.

Set	Ice velocity RMSE			Ice thickness RMSE			BE		
	GCN	GAT	EGCN	GCN	GAT	EGCN	GCN	GAT	EGCN
A	**0.2**	**-0.2**	**0.1**	**0.0**	**-0.1**	**0.1**	**2.9**	**0.0**	**0.0**
B	7.8	4.9	7.0	-0.8	-0.4	3.8	-10.7	-22.7	-7.1
C	-0.5	-0.8	-2.9	-0.4	-0.2	-1.2	-11..8	-15.1	-6.7
D	-1.3	-2.3	-24.3	-1.7	-0.5	-2.5	-12.5	-16.7	-17.2

Table A2. Performance improvement by GAT and EGCN relative to GCN (i.e., error reduction relative to GCN error) in percentage (%). The positive values mean a better performance by GAT or EGCN compared to GCN. If the mean difference between the two models is not statistically significant (p-value > 0.01; i.e., statistically have the same mean), we highlight it in bold.

Set	Ice velocity		Ice thickness		Calving front	
	GAT	EGCN	GAT	EGCN	GAT	EGCN
A	24.9	78.9	-2.7	50.3	11.8	11.8
B	33.3	77.7	4.8	51.4	12.9	**3.2**
C	32.3	80.4	**-2.2**	46.5	**0.0**	15.8
D	32.1	77.7	4.3	46.1	**2.8**	**5.5**

450

References

Aschwanden, A., Fahnestock, M. A., Truffer, M., Brinkerhoff, D. J., Hock, R., Khroulev, C., Mottram, R., and Khan, S. A.: Contribution of the Greenland Ice Sheet to sea level over the next millennium, Science Advances, 5, eaav9396, https://doi.org/10.1126/sciadv.aav9396, 2019.

Bassis, J. N. and Jacobs, S.: Diverse calving patterns linked to glacier geometry, Nature Geoscience, 6, 833–836, https://doi.org/https://doi.org/10.1038/ngeo1887, 2013.

Bevan, S. L., Luckman, A., Khan, S. A., and Murray, T.: Seasonal dynamic thinning at Helheim Glacier, Earth and Planetary Science Letters, 415, 47–53, https://doi.org/10.1016/j.epsl.2015.01.031, 2015.

Black, N. and Najafi, A. R.: Learning finite element convergence with the multi-fidelity graph neural network, Computer Methods in Applied Mechanics and Engineering, 397, 115 120, https://doi.org/10.1016/j.cma.2022.115120, 2022.

Bondzio, J., Morlighem, M., Seroussi, H., Kleiner, T., Ruckamp, M., Mouginot, J., Moon, T., Larour, E., and Humbert, A.: The mechanisms behind Jakobshavn Isbræ's acceleration and mass loss: A 3-D thermomechanical model study, Geophys. Res. Lett., 44, 6252–6260, https://doi.org/10.1002/2017GL073309, 2017.

Bondzio, J. H., Seroussi, H., Morlighem, M., Kleiner, T., Rückamp, M., Humbert, A., and Larour, E. Y.: Modelling calving front dynamics using a level-set method: application to Jakobshavn Isbræ, West Greenland, The Cryosphere, 10, 497–510, https://doi.org/10.5194/tc-10-497-2016, 2016.

Brinkerhoff, D., Aschwanden, A., and Fahnestock, M.: Constraining subglacial processes from surface velocity observations using surrogate-based Bayesian inference, Journal of Glaciology, 67, 385–403, https://doi.org/10.1017/jog.2020.112, 2021.

Cheng, G., Morlighem, M., Mouginot, J., and Cheng, D.: Helheim Glacier's Terminus Position Controls Its Seasonal and Inter-Annual Ice Flow Variability, Geophysical Research Letters, 49, e2021GL097 085, https://doi.org/10.1029/2021GL097085, 2022.

Cheng, G., Morlighem, M., and Gudmundsson, G. H.: Numerical stabilization methods for level-set-based ice front migration, Geoscientific Model Development, 17, 6227–6247, https://doi.org/10.5194/gmd-17-6227-2024, 2024.

Choi, Y., Morlighem, M., Wood, M., and Bondzio, J. H.: Comparison of four calving laws to model Greenland outlet glaciers, The Cryosphere, 12, 3735–3746, https://doi.org/10.5194/tc-12-3735-2018, 2018.

Choi, Y., Morlighem, M., Rignot, E., and Wood, M.: Ice dynamics will remain a primary driver of Greenland ice sheet mass loss over the next century, Communications Earth & Environment, 2, 26, https://doi.org/10.1038/s43247-021-00092-z, 2021.

Choi, Y., Seroussi, H., Morlighem, M., Schlegel, N.-J., and Gardner, A.: Impact of time-dependent data assimilation on ice flow model initialization and projections: a case study of Kjer Glacier, Greenland, The Cryosphere, 17, 5499–5517, https://doi.org/10.5194/tc-17-5499-2023, 2023.

de Juan, J., Elósegui, P., Nettles, M., Larsen, T. B., Davis, J. L., Hamilton, G. S., Stearns, L. A., Andersen, M. L., Ekström, G., Ahlstrøm, A. P., Stenseng, L., Khan, S. A., and Forsberg, R.: Sudden increase in tidal response linked to calving and acceleration at a large Greenland outlet glacier, Geophysical Research Letters, 37, L12 501, https://doi.org/10.1029/2010GL043289, 2010.

dos Santos, T. D., Morlighem, M., and Seroussi, H.: Assessment of numerical schemes for transient, finite-element ice flow models using ISSM v4.18, Geoscientific Model Development, 14, 2545–2573, https://doi.org/10.5194/gmd-14-2545-2021, 2021.

Downs, J., Brinkerhoff, D., and Morlighem, M.: Inferring time-dependent calving dynamics at Helheim Glacier, Journal of Glaciology, 69, 381–396, https://doi.org/10.1017/jog.2022.68, 2023.

Edwards, T., Nowicki, S., Goelzer, H., Seroussi, H., Marzeion, B., Smith, C. E., Jourdain, N. C., Slater, D., McKenna, C., Simon, E., Abe-Ouchi, A., Gregory, J., Hock, R., Larour, E., Lipscomb, W., Payne, A., Shepherd, A., Agosta, C., Alexander, P., Albrecht, T., Anderson,

B., Asay-Davis, X., Aschwanden, A., Barthel, A., Calov, R., Chambers, C., Golledge, N., Greve, R., Hatterman, T., Hoffman, M., Humbert, A., Huss, M., Huybrechts, P., Immerzeel, W., Kleiner, T., Kraaijenbrink, P., Le Clec'h, S., Lee, V., Leguy, G., Little, C., Lowry, D., Malles, J.-H., Maussion, F., Morlighem, M., Nias, I., Pattyn, F., Pelle, T., Price, S., Quiquet, A., Radic, V., Reese, R., Rounce, D., Rückamp, M., Sakai, A., Schlegel, N., Trusel, L. D., Van Breedam, J., Van De Wal, R. S. W., van den Broeke, M., Winkelmann, R., Zhao, C., Zhang, T., Zekollari, H., Zwinger, T., Champollion, N., Choi, Y., Cullather, R., Cuzzone, J., Dumas, C., Felikson, D., Fettweis, X., Fujita, S., and Gladstone, R.: Quantifying uncertainties in the land ice contribution to sea level rise this century, Nature, 593, 74–82, https://doi.org/10.1038/s41586-021-03302-y, 2021.

Everett, A., Murray, T., Selmes, N., Holland, D., and Reeve, D. E.: The Impacts of a Subglacial Discharge Plume on Calving, Submarine Melting, and Mélange Mass Loss at Helheim Glacier, South East Greenland, Journal of Geophysical Research: Earth Surface, 126, e2020JF005 910, https://doi.org/10.1029/2020JF005910, 2021.

Fu, X., Zhou, F., Peddireddy, D., Kang, Z., Jun, M. B.-G., and Aggarwal, V.: An finite element analysis surrogate model with boundary oriented graph embedding approach for rapid design, Journal of Computational Design and Engineering, 10, 1026–1046, https://doi.org/10.1093/jcde/qwad025, 2023.

Gagliardini, O., Zwinger, T., Gillet-Chaulet, F., Durand, G., Favier, L., de Fleurian, B., Greve, R., Malinen, M., Martín, C., Råback, P., Ruokolainen, J., Sacchettini, M., Schäfer, M., Seddik, H., and Thies, J.: Capabilities and performance of Elmer/Ice, a new-generation ice sheet model, Geoscientific Model Development, 6, 1299–1318, https://doi.org/10.5194/gmd-6-1299-2013, 2013.

Greene, C. A., Gardner, A. S., Wood, M., and Cuzzone, J. K.: Ubiquitous acceleration in Greenland Ice Sheet calving from 1985 to 2022, Nature, 625, 523–528, https://doi.org/10.1038/s41586-023-06863-2, 2024.

Jiang, C. and Chen, N.-Z.: Graph Neural Networks (GNNs) based accelerated numerical simulation, Engineering Applications of Artificial Intelligence, 123, 106 370, https://doi.org/10.1016/j.engappai.2023.106370, 2023.

Jouvet, G.: Inversion of a Stokes glacier flow model emulated by deep learning, Journal of Glaciology, 69, 13–26, https://doi.org/10.1017/jog.2022.41, 2023.

Jouvet, G. and Cordonnier, G.: Ice-flow model emulator based on physics-informed deep learning, Journal of Glaciology, 69, 1941–1955, https://doi.org/10.1017/jog.2023.73, 2023.

Jouvet, G., Picasso, M., Rappaz, J., and Blatter, H.: A new algorithm to simulate the dynamics of a glacier: theory and applications, Journal of Glaciology, 54, 801–811, https://doi.org/10.3189/002214308787780049, 2008.

Jouvet, G., Cordonnier, G., Kim, B., Lüthi, M., Vieli, A., and Aschwanden, A.: Deep learning speeds up ice flow modelling by several orders of magnitude, Journal of Glaciology, 68, 651–664, https://doi.org/10.1017/jog.2021.120, 2022.

King, M. D., Howat, I. M., Candela, S. G., Noh, M. J., Jeong, S., Noël, B. P., van den Broeke, M. R., Wouters, B., and Negrete, A.: Dynamic ice loss from the Greenland Ice Sheet driven by sustained glacier retreat, Communications Earth & Environment, 1, 1, https://doi.org/10.1038/s43247-020-0001-2, 2020.

Kipf, T. N. and Welling, M.: Semi-Supervised Classification with Graph Convolutional Networks, https://doi.org/10.48550/arXiv.1609.02907, 2017.

Kondo, K. and Sugiyama, S.: Calving, ice flow, and thickness of outlet glaciers controlled by land-fast sea ice in Lützow-Holm Bay, East Antarctica, Journal of Glaciology, 69, 1751–1763, https://doi.org/10.1017/jog.2023.59, 2023.

Koo, Y. and Rahnemoonfar, M.: Graph convolutional network as a fast statistical emulator for numerical ice sheet modeling, Journal of Glaciology, 71, e15, https://doi.org/10.1017/jog.2024.93, 2025.

Larour, E., Seroussi, H., Morlighem, M., and Rignot, E.: Continental scale, high order, high spatial resolution, ice sheet modeling using the Ice Sheet System Model (ISSM), Journal of Geophysical Research: Earth Surface, 117, F01 022, https://doi.org/10.1029/2011JF002140, 2012.

Lippert, E., Morlighem, M., Cheng, G., and Khan, S.: Modeling a century of change: Kangerlussuaq Glacier's mass loss from 1933 to 2021, Geophysical Research Letters, 51, e2023GL106 286, https://doi.org/10.1029/2023GL106286, 2024.

MacAyeal, D. R.: Large-scale ice flow over a viscous basal sediment: Theory and application to ice stream B, Antarctica, Journal of Geophysical Research: Solid Earth, 94, 4071–4087, https://doi.org/10.1029/JB094iB04p04071, 1989.

Maurizi, M., Gao, C., and Berto, F.: Predicting stress, strain and deformation fields in materials and structures with graph neural networks, Scientific Reports, 12, 21 834, https://doi.org/10.1038/s41598-022-26424-3, 2022.

Melton, S. M., Alley, R. B., Anandakrishnan, S., Parizek, B. R., Shahin, M. G., Stearns, L. A., LeWinter, A. L., and Finnegan, D. C.: Meltwater drainage and iceberg calving observed in high-spatiotemporal resolution at Helheim Glacier, Greenland, Journal of Glaciology, 68, 812–828, https://doi.org/10.1017/jog.2021.141, 2022.

Meng, Y., Lai, C.-Y., Culberg, R., Shahin, M. G., Stearns, L. A., Burton, J. C., and Nissanka, K.: Seasonal changes of mélange thickness coincide with Greenland calving dynamics, Nature Communications, 16, 1–16, https://doi.org/10.1038/s41467-024-55241-7, 2025.

Miles, V. V., Miles, M. W., and Johannessen, O. M.: Satellite archives reveal abrupt changes in behavior of Helheim Glacier, southeast Greenland, Journal of Glaciology, 62, 137–146, https://doi.org/10.1017/jog.2016.24, 2016.

Morlighem, M., Bondzio, J., Seroussi, H., Rignot, E., Larour, E., Humbert, A., and Rebuffi, S.: Modeling of Store Gletscher's calving dynamics, West Greenland, in response to ocean thermal forcing, Geophysical Research Letters, 43, 2659–2666, https://doi.org/10.1002/2016GL067695, 2016.

Morlighem, M., Williams, C. N., Rignot, E., An, L., Arndt, J. E., Bamber, J. L., Catania, G., Chauché, N., Dowdeswell, J. A., Dorschel, B., Fenty, I., Hogan, K., Howat, I., Hubbard, A., Jakobsson, M., Jordan, T. M., Kjeldsen, K. K., Millan, R., Mayer, L., Mouginot, J., Noël, B. P. Y., O'Cofaigh, C., Palmer, S., Rysgaard, S., Seroussi, H., Siegert, M. J., Slabon, P., Straneo, F., van den Broeke, M. R., Weinrebe, W., Wood, M., and Zinglersen, K. B.: BedMachine v3: Complete Bed Topography and Ocean Bathymetry Mapping of Greenland From Multibeam Echo Sounding Combined With Mass Conservation, Geophysical Research Letters, 44, 11,051–11,061, https://doi.org/10.1002/2017GL074954, 2017.

Morlighem, M., Rignot, E., Binder, T., Blankenship, D., Drews, R., Eagles, G., Eisen, O., Ferraccioli, F., Forsberg, R., Fretwell, P., Goel, V., Greenbaum, J. S., Gudmundsson, H., Guo, J., Helm, V., Hofstede, C., Howat, I., Humbert, A., Jokat, W., Karlsson, N. B., Lee, W. S., Matsuoka, K., Millan, R., Mouginot, J., Paden, J., Pattyn, F., Roberts, J., Rosier, S., Ruppel, A., Seroussi, H., Smith, E. C., Steinhage, D., Sun, B., Broeke, M. R. v. d., Ommen, T. D. v., Wessem, M. v., and Young, D. A.: Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the Antarctic ice sheet, Nat. Geosci., 13, 132–137, https://doi.org/10.1038/s41561-019-0510-8, 2020.

Mouginot, J., Rignot, E., Scheuchl, B., and Millan, R.: Comprehensive annual ice sheet velocity mapping using Landsat-8, Sentinel-1, and RADARSAT-2 data, Remote Sensing, 9, 364, https://doi.org/10.3390/rs9040364, 2017.

Mouginot, J., Rignot, E., Bjørk, A. A., van den Broeke, M., Millan, R., Morlighem, M., Noël, B., Scheuchl, B., and Wood, M.: Forty-six years of Greenland Ice Sheet mass balance from 1972 to 2018, Proceedings of the National Academy of Sciences, 116, 9239–9244, https://doi.org/10.1073/pnas.1904242116, 2019.

Osher, S. and Sethian, J. A.: Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations, Journal of Computational Physics, 79, 12–49, https://doi.org/10.1016/0021-9991(88)90002-2, 1988.

Otosaka, I. N., Shepherd, A., Ivins, E. R., Schlegel, N.-J., Amory, C., van den Broeke, M. R., Horwath, M., Joughin, I., King, M. D., Krinner, G., Nowicki, S., Payne, A. J., Rignot, E., Scambos, T., Simon, K. M., Smith, B. E., Sørensen, L. S., Velicogna, I., Whitehouse, P. L., A, G.,

Agosta, C., Ahlstrøm, A. P., Blazquez, A., Colgan, W., Engdahl, M. E., Fettweis, X., Forsberg, R., Gallée, H., Gardner, A., Gilbert, L., Gourmelen, N., Groh, A., Gunter, B. C., Harig, C., Helm, V., Khan, S. A., Kittel, C., Konrad, H., Langen, P. L., Lecavalier, B. S., Liang, C.-C., Loomis, B. D., McMillan, M., Melini, D., Mernild, S. H., Mottram, R., Mouginot, J., Nilsson, J., Noël, B., Pattle, M. E., Peltier, W. R., Pie, N., Roca, M., Sasgen, I., Save, H. V., Seo, K.-W., Scheuchl, B., Schrama, E. J. O., Schröder, L., Simonsen, S. B., Slater, T., Spada, G., Sutterley, T. C., Vishwakarma, B. D., van Wessem, J. M., Wiese, D., van der Wal, W., and Wouters, B.: Mass balance of the Greenland and Antarctic ice sheets from 1992 to 2020, Earth System Science Data, 15, 1597–1616, https://doi.org/10.5194/essd-15-1597-2023, 2023.

O'Neel, S., Echelmeyer, K. A., and Motyka, R. J.: Short-term variations in calving of a tidewater glacier: LeConte Glacier, Alaska, U.S.A., Journal of Glaciology, 49, 587–598, https://doi.org/10.3189/172756503781830430, 2003.

Perera, R., Guzzetti, D., and Agrawal, V.: Graph neural networks for simulating crack coalescence and propagation in brittle materials, Computer Methods in Applied Mechanics and Engineering, 395, 115 021, https://doi.org/10.1016/j.cma.2022.115021, 2022.

Pfaff, T., Fortunato, M., Sanchez-Gonzalez, A., and Battaglia, P. W.: Learning Mesh-Based Simulation with Graph Networks, https://arxiv.org/abs/2010.03409, 2021.

Rignot, E., Xu, Y., Menemenlis, D., Mouginot, J., Scheuchl, B., Li, X., Morlighem, M., Seroussi, H., den Broeke, M. v., Fenty, I., et al.: Modeling of ocean-induced ice melt rates of five west Greenland glaciers over the past two decades, Geophysical Research Letters, 43, 6374–6382, 2016.

Salehi, Y. and Giannacopoulos, D.: PhysGNN: A Physics–Driven Graph Neural Network Based Model for Predicting Soft Tissue Deformation in Image–Guided Neurosurgery, Advances in Neural Information Processing Systems, 35, 37 282–37 296, https://doi.org/10.5555/3600270.3602972, 2022.

Satorras, V. G., Hoogeboom, E., and Welling, M.: E(n) Equivariant Graph Neural Networks, https://doi.org/10.48550/arXiv.2102.09844, 2022.

Shivaditya, M. V., Alves, J., Bugiotti, F., and Magoulès, F.: Graph Neural Network-based Surrogate Models for Finite Element Analysis, in: 2022 21st International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), pp. 54–57, IEEE, https://doi.org/0.1109/DCABES57229.2022.00035, 2022.

Slater, D. and Straneo, F.: Submarine melting of glaciers in Greenland amplified by atmospheric warming, Nature Geoscience, 15, 794–799, https://doi.org/10.1038/s41561-022-01035-9, 2022.

Tarasov, L., Dyke, A. S., Neal, R. M., and Peltier, W.: A data-calibrated distribution of deglacial chronologies for the North American ice complex from glaciological modeling, Earth and Planetary Science Letters, 315-316, 30–40, https://doi.org/10.1016/j.epsl.2011.09.010, sea Level and Ice Sheet Evolution: A PALSEA Special Edition, 2012.

Tedesco, M. and Fettweis, X.: Unprecedented atmospheric conditions (1948–2019) drive the 2019 exceptional melting season over the Greenland ice sheet, The Cryosphere, 14, 1209–1223, https://doi.org//10.5194/tc-14-1209-2020, 2020.

Todd, J. and Christoffersen, P.: Are seasonal calving dynamics forced by buttressing from ice mélange or undercutting by melting? Outcomes from full-Stokes simulations of Store Glacier, West Greenland, The Cryosphere, 8, 2353–2365, https://doi.org/10.5194/tc-8-2353-2014, 2014.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y.: Graph Attention Networks, https://doi.org/10.48550/arXiv.1710.10903, 2018.

Verjans, V. and Robel, A.: Accelerating Subglacial Hydrology for Ice Sheet Models With Deep Learning Methods, Geophysical Research Letters, 51, e2023GL105 281, https://doi.org/10.1029/2023GL105281, e2023GL105281 2023GL105281, 2024.

Wehrlé, A., Lüthi, M. P., and Vieli, A.: The control of short-term ice mélange weakening episodes on calving activity at major Greenland outlet glaciers, The Cryosphere, 17, 309–326, https://doi.org/10.5194/tc-17-309-2023, 2023.

605   Wilner, J. A., Morlighem, M., and Cheng, G.: Evaluation of four calving laws for Antarctic ice shelves, The Cryosphere, 17, 4889–4901, https://doi.org/10.5194/tc-17-4889-2023, 2023.

Winkelmann, R., Martin, M. A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK) Part 1: Model description, The Cryosphere, 5, 715–726, https://doi.org/10.5194/tc-5-715-2011, 2011.

Wood, M., Rignot, E., Fenty, I., An, L., Bjørk, A., Broeke, M. v. d., Cai, C., Kane, E., Menemenlis, D., Millan, R., Morlighem, M., Mouginot,
610   J., Noël, B., Scheuchl, B., Velicogna, I., Willis, J. K., and Zhang, H.: Ocean forcing drives glacier retreat in Greenland, Science Advances, 7, eaba7282, https://doi.org/10.1126/sciadv.aba7282, 2021.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Yu, P. S.: A Comprehensive Survey on Graph Neural Networks, IEEE Transactions on Neural Networks and Learning Systems, 32, 4–24, https://doi.org/10.1109/TNNLS.2020.2978386, 2021.

Xie, S., Dixon, T. H., Holland, D. M., Voytenko, D., and Vaňková, I.: Rapid iceberg calving following removal of tightly packed pro-glacial
615   mélange, Nature communications, 10, 3250, https://doi.org/10.1038/s41467-019-10908-4, 2019.

Zhang, S., Tong, H., Xu, J., and Maciejewski, R.: Graph convolutional networks: a comprehensive review, Computational Social Networks, 6, 1–23, https://doi.org/10.1186/s40649-019-0069-y, 2019.