Retrieving climatic insights from the Last Glacial Maximum in the Alps using an inverted glacier model

Kejdi Lleshi[1] (ID), Guillaume Jouvet[1] (ID), Sarah Kamleitner[1,2], Tancrede Leger[1], Frédéric Herman[1] and Samuel James Cook[1,3] (ID)

[1]Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland; [2]Department of Geography, University of Zurich, Zurich, Switzerland and [3]Institute of Geography, FAU, Erlangen, Germany

Abstract

The climatic conditions, particularly the sources of precipitation that enabled extensive glacial growth during the Last Glacial Maximum (LGM) in the European Alps, remain poorly constrained. Here, we apply an inversion method to reconstruct equilibrium line altitude (ELA) fields using glacier footprints, such as the moraines deposited by Alpine glaciers during the LGM. By employing a machine-learning emulator trained on outputs from a glacier-evolution model, we predict glacier maximal thickness. The emulator is integrated into a gradient-based inversion scheme to derive ELA fields consistent with LGM footprints. The results show that the reconstructed ELA fields align with those from previous studies, validating the robustness of our approach. Unlike existing inversion methods, our approach is more general and avoids restrictive assumptions. Notably, by incorporating the transient response of glaciers to climate variability (we do not assume steady state), we show that the cold spell period is crucial for interpreting the reconstructed climate patterns—an aspect previously overlooked. Our findings provide new insights into climatic variability during the LGM, particularly concerning the interaction between precipitation patterns and the cold spell period. Furthermore, the computational efficiency of our method makes it applicable to large-scale paleoclimate reconstructions based on glacier footprints.

1. Introduction

Understanding past natural climate variations during the Quaternary is crucial for understanding current and future climate change. Specifically, the period of the Last Glacial Maximum (LGM), around 24–25 ka BP, remains a topic of debate regarding the changes in the sources and directions of moisture supply over the European Alps. The Alps are positioned as a barrier to meridional moisture transport, recording shifts in the North Atlantic storm track, which significantly influenced regional climate dynamics (Florineth and Schlüchter, 2000; Luetscher and others, 2015; Monegato and others, 2017). Geomorphic evidence from the LGM, such as moraines and trimlines, is utilized to reconstruct the paleo-glacier's extent and thickness (Bini and others, 2009; Ehlers and others, 2011). The reconstructed paleo-glacier can, in turn, be used to extract paleo-climate information (Ohmura and others, 1992; Heyman and others, 2013; Albrecht and others, 2020; Martin and others, 2020; Višnević and others, 2020; Rettig and others, 2023).

1.1. State-of-the-art

One simple inversion method to retrieve climate information from paleo-glacier reconstructions is to seek the glacier equilibrium line altitude (ELA), the altitude where accumulation and ablation of ice due to precipitation and melt are zero, assuming a constant accumulation area ratio (Pellitero and others, 2015; Martin and others, 2020). The accumulation area ratio is the ratio of the area above the equilibrium line to the total area of the glacier (Porter, 1970). The ELA is a fundamental parameter, primarily controlled by the temperature and precipitation regimes of the climate. However, the accumulation area ratio approach to finding the ELA is too simplistic, as it does not account for glacier dynamics and assumes a stationary state of glaciers. In fact, it is well established that glaciers are rarely stationary due to the inherent variability of the climate and the presence of internal glacial processes (e.g. surges) (Jóhannesson and others, 1989; Lüthi, 2009).

Another approach is manually searching for a climate solution consistent with glacier footprints (Heyman and others, 2013; Albrecht and others, 2020). This involves sampling an ensemble of plausible past climate scenarios and performing simulations with a glacier-evolution model. The output of each glacier-evolution model (predicted glacier extent) is

then compared with the reconstructed palaeo-glacier extent. The climate scenarios that yield results most compatible with observations are retained. Given the computational expense of the glacier-evolution model, an iterative strategy is advisable. This method systematically updates the climate inputs to converge to a climate scenario compatible with the reconstructed paleo-glacier extent. An alternative approach was introduced by Višnević and others (2018), who proposed an automatic, iterative inversion method powered by graphics processing units.

Using this method, Višnević and others (2020) were able to reconstruct a spatially varying ELA field compatible with the reconstructed palaeo-glacier extent. Although promising, this method yields high computational costs and employs a heuristic descent approach rather than the more efficient steepest descent method for minimizing the cost function. The cost function is the misfit between the predicted and observed glacier extents. Additionally, Višnević and others (2020) assume that glaciers are in a steady state at the time of maximum extent, a condition rarely met in the dynamic nature of real-world glaciers that frequently respond to climatic changes with a level of delay and inertia that depends on the ice mass and the duration/magnitude of climatic change (Jóhannesson and others, 1989; Lüthi, 2009). They also employ the shallow ice approximation (Hughes, 1984), simplifying the ice mechanics, which is particularly problematic in steep mountainous regions where vertical shearing predominantly influences ice flow (Le Meur and others, 2004).

One promising way to reduce the high computational costs of glacier-evolution models—such as that of Višnević and others (2020)—and to address their inherent limitations is through machine learning-based emulation. The latter consists of using a machine learning model to replicate the outputs of a complex model with high fidelity but at a fraction of the computational cost. Based on this approach, Jouvet (2022) was able to speed up significantly an ice flow model. Additionally, using automatic differentiation in conjunction with the emulator enables gradients to be computed, providing an efficient and precise method for solving complex inverse problems (Paszke and others, 2017; Jouvet, 2022). To the best of our knowledge, combining machine learning-based emulators and automatic differentiation for paleo-glacier reconstructions has never been attempted, despite its potential for both accuracy and lower computational costs.

A further difficulty in paleo-climate reconstruction is the non-uniqueness of solutions: the same glacier extent can be explained by multiple combinations of temperature, precipitation and model parameters. This uncertainty underscores the importance of running multiple inversions to thoroughly explore the parameter space and refine our understanding of the climate–glacier interaction. By dramatically lowering computational costs, machine learning-based emulators pave the way for extensive modelling for parameter exploration, thereby helping constrain solutions more effectively than is possible with traditional, more expensive simulation frameworks.

1.2. Proposed solution

In this paper, we address the limitations mentioned earlier by applying a gradient descent inversion method that uses a machine learning-based emulator to efficiently reconstruct paleo-climate from glacier footprints and apply it to reconstruct spatially varying ELA fields across the European Alps during the LGM. While the

ELA is conventionally represented as a single value characterizing an entire glacier (Braithwaite and Raper, 2009), we aim to analyse its spatial variability as a climate proxy. Furthermore, to ensure consistency with the methodology adopted by Višnević and others (2018, 2020), we retain the term *ELA field* to denote the spatially varying equilibrium altitudes at which the climatic mass balance is zero.

In the following, we outline the construction of the machine learning emulator, beginning with data generation using a glacier-evolution model and progressing to the training of the emulator. We then elaborate on the inversion scheme. Finally, we present our results that assess the quality of both the emulator and the inversion scheme, as well as our outcomes for the climate prevailing at LGM in the Alps.

2. Methods

2.1. Data generation with a glacier-evolution model

Examining a range of climate forcings, glaciation durations and other glacier-related physical parameters is essential for building the dataset used to train our emulator. To accomplish this, we employ the instructed glacier model (IGM) (Jouvet and others, 2022) as the glacier-evolution model that couples ice flow, surface mass balance (SMB) and mass conservation. The specificity of IGM is that it uses an emulator informed by model realization runs to model the ice dynamics, offering a significant speed-up compared to traditional solvers with fairly limited loss of accuracy (Jouvet and others, 2022). In our study, we required 4200 simulations of the Alps, with each simulation covering an average of 2500 years. Given these high computational demands, IGM's efficiency is essential for feasibility while still capturing the key physical processes governing glacier evolution.

Each IGM simulation is initialized with ice-free conditions over the present-day Alpine bedrock $b(x, y)$. The decision to initialize the simulations as ice-free was made due to the lack of data concerning glacier extents in the Alps prior to the LGM. The bedrock is defined by a 451 by 301 raster grid, using a 2×2 km resolution.

We use the Python package *FyeldGenerator* to generate two-dimensional Gaussian random ELA fields (Fig. 3) for investigating various climate forcings, represented as $z_{ELA_0}(x, y)$. These fields exhibit a mean altitude ranging between 1000 and 2000 m. By systematically generating new fields, we obtain glaciers with diverse geometries, thereby enhancing the emulator's ability to accurately capture the relationship between ELA variations and glacier morphology.

Luetscher and others (2015) document oscillatory climatic behaviour during the LGM (27–21 ka BP; Fig. 1). To mimic a transient cooling and warming of the climate, we designed the simulations such that the ELA field follows one period of a cosine function (T) as shown in Figure 1 and :

$$z_{ELA}(x, y, t) = z_{ELA_0}(x, y, t_0) + 500 \times cos(2\pi \times t/T), \quad (1)$$

where $z_{ELA}(x, y, t)$ is the ELA that evolves with time, $z_{ELA_0}(x, y)$ is the ELA at the initial time ($t = 0$), 500 is the amplitude of the cosine function in meters, t is the time index and T is the period of the cosine.

We model the climate's impact on glacier dynamics through a spatially and temporally varying SMB function, expressed as:
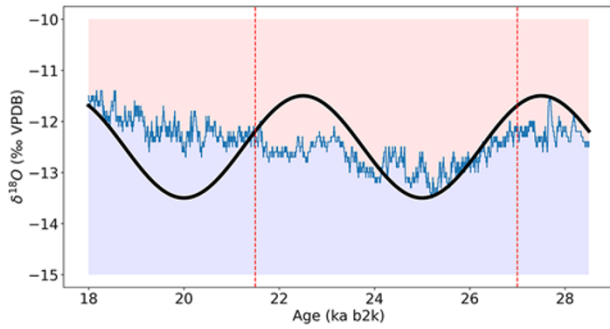
Figure 1. The ELA field constantly changes during the simulations following a cosine with a period of 5000 years (black line). The blue line is the Oxygen isotope ratio from Luetscher and others (2015). Red lines are at 21 500 and 27 000 years BP.

$$\text{smb}(x, y, t) = \min\left(\gamma \cdot (s(x, y, t) - z_{\text{ELA}}(x, y, t)), c\right),$$

$$\gamma = \begin{cases} \beta_{\text{abl}} & \text{if } s(x, y, t) - z_{\text{ELA}}(x, y, t) < 0, \\ \beta_{\text{acc}} & \text{if } s(x, y, t) - z_{\text{ELA}}(x, y, t) \geq 0, \end{cases} \quad (2)$$

where β_{abl} is the mass balance gradient for the ablation area, β_{acc} is the mass balance gradient for the accumulation area = 0.0005 a^{-1}, $s(x, y, t)$ is the surface elevation, $z_{ELA}(x, y, t)$ is the spatially and temporally variable ELA defined in Eqn (1) and c is the maximum ice accumulation rate = 2 m a^{-1} (Meier, 1962).

This approach enables the emulator to learn the dynamic relationship between climate variations and glacier morphology. It has been established that glaciers of varying sizes react differently to a given climate scenario (Zekollari and Huybrechts, 2015). Smaller glaciers generally respond more quickly to climate change, while larger glaciers require more time to react.

We have carried out 4200 simulations, each with a unique initial ELA field ($z_{ELA_0}(x, y)$). In addition, simulations have different run periods (T), ranging from 1000 to 5000 years and physical parameters whose values are shown in Table 2. The physical parameters include the rate factor in Glen's flow law (Glen, 1955) that controls the ice shearing from the cold-ice case (low A) to the temperate ice case (A), the sliding coefficient (s) that controls the strength of basal motion and mass balance gradient β_{abl} from Eqn (2). After each simulation, we define the glacier maximal thickness (GMT) $\mathcal{H}(x, y)$ as the maximal ice thickness $h(x, y, t)$ over the run time from $t = 0$ to $t = T$ for each pixel. We calculated the GMT because, in cases where paleo-glacier thickness can be reconstructed from trimlines, it typically represents the glacier's maximal thickness.

$$\mathcal{H}(x, y) = \max_t h(x, y, t) \quad (3)$$

Having described the data generation process through IGM simulations, we now detail how these data form the basis for constructing the emulator.

2.2. Emulator

Our emulator is a machine learning model that aims to replicate the behaviour of a glacier-evolution model with high accuracy and a minimized computational cost. We build the emulator (E) to map an input to an output, Eqn (4). The input data to train the emulator are the bedrock topography $b(x, y)$, the initial ELA field ($z_{ELA_0}(x, y)$), the period (T) and physical parameters such as

the sliding coefficient s, the mass balance gradient β_{abl} from Eqn (2) and the rate factor from Glen's law (A). On the other hand, the output data for the emulator is $\mathcal{H}(x, y)$. The input and output are two-dimensional fields defined by raster grids with the same dimensions, as shown in Eqn (4).

$$E : \{b(x, y), z_{ELA_0}(x, y), T, A, \beta_{\text{abl}}, S\} \to \{\mathcal{H}(x, y)\}$$
$$\mathbb{R}^{N_x \times N_y \times 6} \to \mathbb{R}^{N_x \times N_y \times 1}. \quad (4)$$

Our emulator is based on a convolutional neural network, a specialized neural network architecture particularly effective for processing grid-like data, such as images. This makes convolutional neural networks well-suited for analysing two-dimensional fields representing glaciers, topography and climatic conditions. Further details on convolutional neural networks are provided in and in the following references O'Shea and Nash (2015); Jouvet (2022); Cong and Zhou (2023).

We allocate 80% of our data collection for training, allowing the emulator to learn the complex spatial relationships between climate conditions and glacier attributes. The remaining 20% is used to validate the model's predictive accuracy and generalizability. To optimize performance, we train multiple emulators with varying parameters, including the number of hidden layers, kernel size, batch size and overall architecture (Appendix A). We monitor each model's performance across different parameter configurations using TensorBoard's hparams plug-in (Abadi and others, 2016). The best-performing emulator, requiring no further training, is selected as our final emulator E.

Finally, we assess the emulator's generalization ability using a previously unseen dataset. This dataset is generated independently using the same data-generation pipeline, ensuring a robust evaluation of the emulator's predictive capabilities.

2.3. Inversion

This subsection presents an inversion method that utilizes a two-dimensional reconstruction of the Alpine ice field mask $\mathcal{M}_{obs}(x, y)$ (it equals 1 in ice-covered regions, and 0 in ice-free regions) during the LGM (Ehlers and others, 2011) to infer a compatible ELA field.

The task of finding ELA fields that can explain the observations($\mathcal{M}_{obs}$) can be framed as an optimization problem. The goal is to seek an ELA field that minimizes $\mathcal{C}$:

$$\mathcal{C}(ELA) = \| \mathcal{M}_{obs}(x, y) - \mathcal{M}(E\{b(x, y), z_{ELA_0}(x, y), T, A, \beta_{\text{abl}}, S\}) \|_2^2 + \mathcal{R}(z_{ELA}(x, y)). \quad (5)$$

Here, $\mathcal{C}$ represents the cost function, and $\mathcal{M}_{obs}(x, y)$ denotes the observed glacier mask extent (Ehlers and others, 2011). The term $E\{b(x, y), zELA_0(x, y), T, A, \beta_{\text{abl}}, S\}$ represents the glacier maximal ice thickness predicted by the emulator (E), and the function $\mathcal{M}$ keeps its sole mask (i.e. equal to 1 in ice-covered regions, and 0 elsewhere) to be compared with observations. Additionally, $\mathcal{R}(z_{ELA}(x, y))$ is a regularization term for the inferred ELA field. This term plays a crucial role in mitigating overfitting and ensuring the physical plausibility of our solution by preventing extreme variations (Tikhonov, 1963). The regularization term is mathematically defined as follows:

$$\mathcal{R}(z_{ELA}(x, y)) = \lambda \int (\nabla z_{ELA}(x, y))^2 dx dy, \quad (6)$$
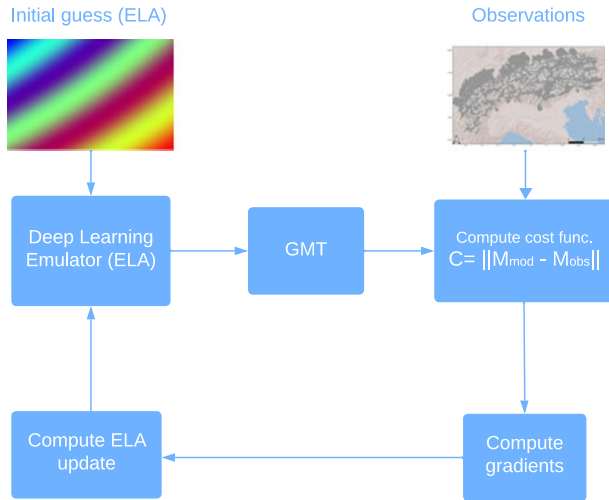
Figure 2. Schematic representation of our iterative optimization approach. The emulator is provided with input variables from Table 2. The initial ELA field follows a gradient from northwest to southeast, ranging from 1300 to 2000 m. The emulator then predicts the GMT, which is subsequently compared to the observations. Based on the cost function, gradients are computed, and the ELA field is iteratively updated to find the optimal solution.

Table 1. Comparison of computational times for the IGM (Jouvet and others, 2022) and the emulator (E), highlighting the efficiency of the emulator *E*. Computations are done with a GPU NVIDIA RTX A3000 12GB.

Time [yrs]	IGM run time [s]	*E* run time [s]
5000	169.31	0.06
4000	128.01	0.06
3000	106.63	0.06
2000	53.85	0.06
1000	26.04	0.06

where λ is the regularization coefficient dictating the scale of smoothness enforced on the ELA field. We minimize the functional (5) within its parameter space using an iterative gradient-based approach. The optimization scheme selected for this study is the Adaptive Moment Estimation (ADAM) algorithm (Kingma and Ba, 2014).

As illustrated in Fig. 2, we start with an initial estimate of the ELA field ($z_{ELA_0}(x, y)$) while keeping the other input variables of the emulator unchanged during the inversion process. Given all the input variables, the emulator E predicts a glacier extent four orders of magnitude faster than a glacier-evolution model would (Table 1). We then compute the 'misfit' between the predicted glacier extent and the 'observed' glacier extent using the cost function $\mathcal{C}$ defined by Eqn (5).

We obtain the gradient (g_i) of the cost function (5) with respect to its parameters, using AD. The gradient, $g_i = \nabla_{ELA}\mathcal{C}_i(ELA)$, is used by the ADAM optimizer to identify the descent direction (p_i) and the step length (α_i) that minimizes the function (5). The descent direction (p_i) ensures that the cost function is reduced after each iteration, facilitating the algorithm's convergence to a minimum. The step length (α_i) is a positive scalar scaled by the gradient.

In summary, the ELA field is updated as follows:

$$z_{ELA}^{i+1} = z_{ELA}^i + \alpha^i p^i. \tag{7}$$

In the subsequent iteration, the updated ELA field (z_{ELA}^{i+1}) is given to the emulator to predict the new glacier extent ($E\{z_{ELA}^{i+1}\}$), which should be closer to the observed glacier extent ($\mathcal{M}_{obs}(x, y)$). This process is repeated until the cost function converges or shows insufficient improvement. Upon completion, the final ELA field is considered the solution of the inversion scheme.

3. Results

This section delineates the findings in two main segments. Initially, we explore the emulator, assessing its precision in emulating a glacier-evolution model. Subsequently, we present the inversion results to extract palaeoclimate information from reconstructed glacier extent in the Alps during the LGM. The inversion method is employed multiple times to explore combinations of ELA fields with the period T, rate factor (A) and β_{abl}. A sensitivity analysis is then conducted to test the impact of climatic or glacial parameters on the results.

3.1. Emulation capabilities

We assess the emulator's proficiency using data that were not used during training. The capability to generate results comparable to those produced by the IGM (Jouvet and others, 2022) is then evaluated. These assessments, aimed at understanding the emulator's capability, are reported in Figs 3 and A2. Moreover, Table 1 delineates the computational efficiency of the emulator compared to a glacier-evolution model such as the IGM.

As detailed in the Methods section, the test dataset was generated by IGM simulations. The input data are then fed to the emulator, and the predicted GMT is compared to the GMT in the dataset, which the IGM generated. The 'True-Predicted' subfigure in Fig. 3 illustrates a close alignment between the emulator's outputs and the IGM-simulated GMT. Nevertheless, minor discrepancies are observed in the glacier lobes, particularly where the glacier extends into lowland areas. These regions are known to be highly sensitive to small parameter variations and are particularly challenging for palaeo-glacier reconstructions (Seguinot and others, 2018; Jouvet and others, 2023).
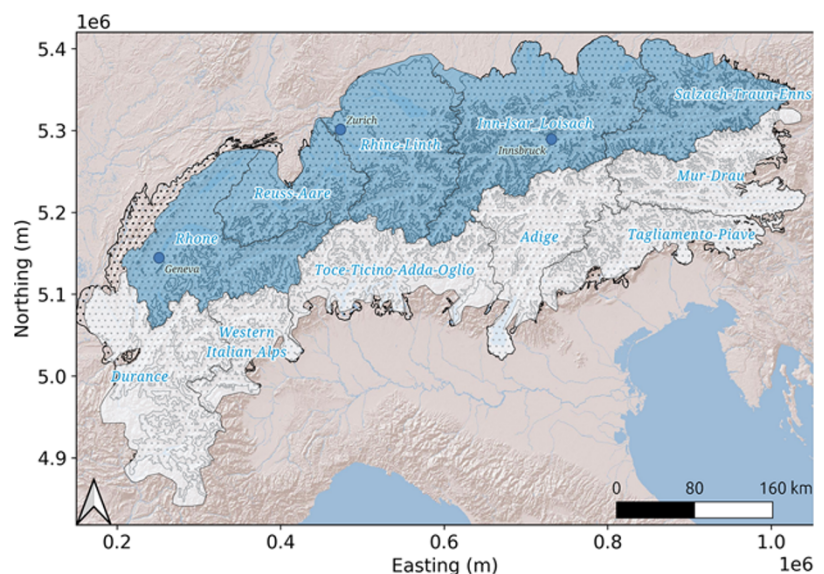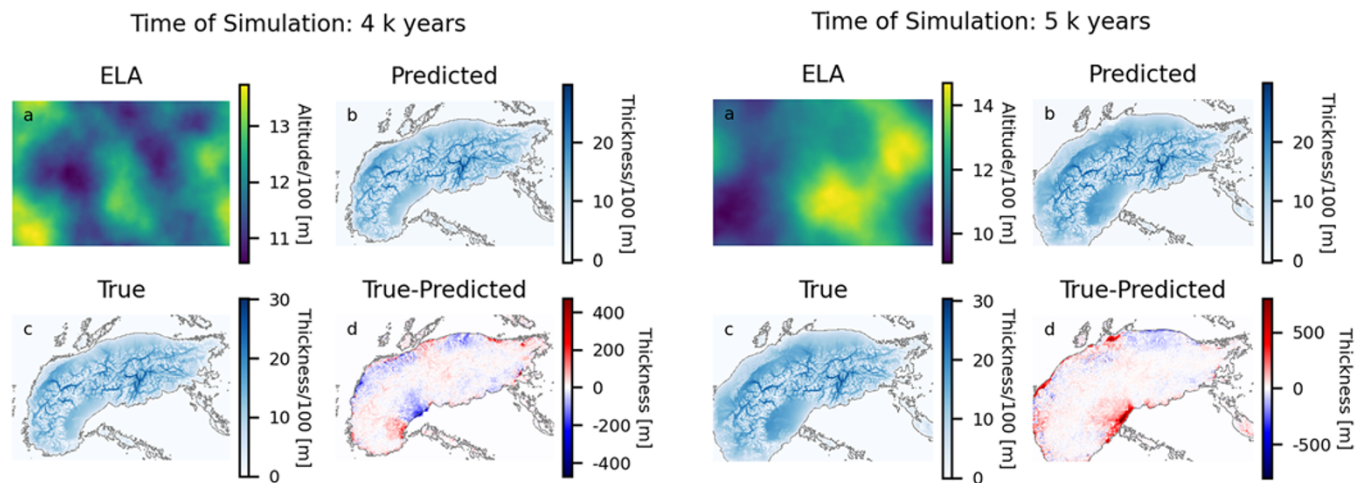
Table 1 compares the computational time required to derive GMT using both the IGM and the emulator, executed on the same computational platform, an NVIDIA RTX A3000 12GB GPU. The results indicate that the emulator operates at three to four orders of magnitude faster than the IGM. This efficiency is particularly crucial during the inversion process, where numerous forward modelling predictions are needed. Such performance underscores the substantial computational advantages of employing the emulator over conventional glacier-evolution models.

3.2. Inverting paleo-glaciers from the LGM

The inversion scheme applied here derives a two-dimensional ELA field for the Alps during the LGM by inverting the observed glacier extent (Fig. 4). To facilitate the interpretation of inferred ELA results, we delineated 11 regions within the Alps. These regions are thought to represent interconnected glacier systems during the LGM. We calculate the mean ELA value for each region from the inverted ELA field, as shown in Fig. 5.

To fully harness the capabilities of our inversion scheme, we applied it to the observed glacier extent under three distinct

Figure 3. Emulator evaluation at two distinct run times (4 and 5 ka) to illustrate its predictive capabilities. (a) ELA field used as input for the IGM and the emulator. (b) Emulator-predicted GMT. (c) GMT from the IGM. (d) Difference between true and predicted GMT (blue: overestimation; red: underestimation). Overall, the emulator closely replicates the IGM outputs.

Figure 4. Reconstructed glacier extent from observations (Ehlers and others, 2011), delineated by black lines with a dotted pattern inside the boundaries. The regions shaded in light blue and white are defined based on current hydrological catchment maps.

periods T: 3000, 4000 and 5000 years, while keeping other physical parameters fixed. The physical parameters, denoted in bold in Table 2, are selected based on a literature review (Cohen and others, 2018; Višnević and others, 2020; Jouvet, 2022). All results were compiled into Fig. 6.

Figure 6 illustrates two main findings. Firstly, the delineated regions are clustered into two groups: regions north of the Alpine divide, with lower ELA values, and southern regions with higher ELA values. The glacier footprints indicate the presence of large glacier lobes in the northern Alps foreland, corresponding with lower ELA values. In contrast, the higher ELA values in the southern Alps resulted in smaller glacier lobes. Secondly, there is a clear relationship between the climate and the period T. The Alpine ice field could have formed under lower ELAs, indicating a cold climate or high precipitation, during a short period T (i.e. 3000 years). Conversely, it could have been built under higher ELAs during an extended period T (i.e. 5000 years).

Figure 5. The ELA map obtained from the inversion scheme using the parameters: $T = 4000$ a, $A = 78$ MPa^{-3} a^{-1}, $\beta = 0.008$ a^{-1}, $s = 2.1$ km MPa^{-3} a^{-1} and $\lambda = 0.1$. The ELA map is delineated by predefined catchments (Fig. 4). The map reveals a distinct gradient, with ELA values increasing from northwest to southeast.

Table 2. List of input parameters for the emulator, where $b(x, y)$ is bedrock topography, $z_{ELA_0}(x, y)$ is the initial ELA map, T is the period of the cosine (Fig. 1) and time of simulation, A is Glen's law rate factor, β_{abl} is the mass balance gradient and S is the sliding coefficient. Bold values are selected as default parameters.

Parameter	Tested values	Units
$b(x, y)$	Alps topography	m
$z_{ELA_0}(x, y)$	Initial ELA field	m
T	3000, **4000**, 5000	a
A	39, **78**, 156	MPa^{-3} a^{-1}
β_{abl}	0.004, **0.008**, 0.016	a^{-1}
S	**2.1**	km MPa^{-3} a^{-1}
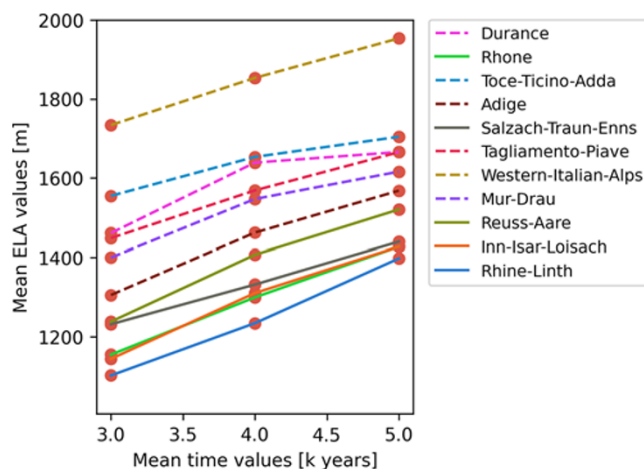
Figure 6. ELA values averaged over region-of-interest (Fig. 4) plotted against the period T. The figure illustrates the automatic spatial clustering of the ELA values into two primary groups and the correlation between the period T and the climate scenario. Each line in the figure, represented by a unique colour, corresponds to a specific region. The solid lines represent the northern regions, while the dashed lines represent the southern regions.

3.3. Sensitivity analysis

Due to the inherent uncertainties and our assumptions in the rate factor in Glen's flow law (A) and the mass balance gradient (β_{abl}), we have conducted a sensitivity analysis to investigate the robustness of our results concerning these two parameters.

We investigate the role of the mass balance gradient (β_{abl}) by halving and doubling the standard reference value 0.008 $a^{(-1)}$ (Table 2) while maintaining all other input parameters at their standard values as detailed in the Methods section. For each β_{abl} value, we perform the inversion three times, each with a different period T, as specified in Table 2. This process results in nine distinct ELA fields ($3 \times 3 = 9$).

The corresponding results are presented in Fig. 7a, where the mean ELA values obtained using the standard parameters are shown in bold colours while those derived from the halved and doubled β_{abl} are depicted with reduced opacity. Figure 7a illustrates the sensitivity of our method to variations in β_{abl}. A greater separation between plots of the same colour indicates increased sensitivity of the inversion to changes in β_{abl}.

Subsequently, we examine the sensitivity to Glen's law rate factor (A), by halving and doubling the standard reference value 78 MPa^{-3} a^{-1} (Table 2), again holding all other input parameters constant. The same methodology applied to β_{abl} is used to assess the sensitivity of our method to variations in Glen's law rate factor, as depicted in Fig. 7b.

Figure 7 reveals that the inversion scheme exhibits greater sensitivity to changes in the mass balance gradient (β_{abl}). These findings underscore the importance of carefully interpreting our results concerning β_{abl}. At maximum, the β_{abl} parameter alone can, within the tested values, modify the ELA by up to 100 m (Western Italian Alps region, 5 ka T period). In contrast, the method demonstrates less sensitivity to the rate factor A variations. Regardless of the parameter selected, our method consistently delineates regions into two distinct clusters, north and south of the Alpine divide. Additionally, a strong positive correlation between ELAs and the period T remains.

4. Discussion

4.1. LGM climate reconstruction

A key open question concerning the Alpine ice field during the LGM is the prevailing precipitation patterns. Florineth and Schlüchter (2000) and Luetscher and others (2015) proposed that during the LGM, weather patterns in the Alps were dominated by southerly airflow. The hypothesized mechanism is a southward displacement of the oceanic polar front and associated storm tracks to the latitude of Spain and Italy, making the North Atlantic Ocean play a smaller role in moisture sourcing during this time. Moreover, meridional temperature gradients reconstructed by Haeberli and Penz (1985) suggest cold and arid conditions in the northern Alps during the peak of the LGM, while the southern Alps experienced relatively warmer and more humid conditions.

On the other hand, inverting a glacier-evolution model, Višnević and others (2020) found an increasing ELA from West to East and from North to South of the mountain range, which suggests a westerly dominated moisture advection pattern originating from the North Atlantic, thus relatively similar to today. These findings challenged earlier studies (Florineth and Schlüchter, 2000; Luetscher and others, 2015) and aligned more closely with recent Regional-Climate-Model-derived climate reconstructions (Russo and others, 2024). However, it is important to note that Russo and others (2024) reported higher uncertainties in reconstructing LGM precipitation patterns compared to temperatures. When using LGM climate snapshots produced by Jouvet and others (2023); Russo and others (2024) found a clear improvement in the agreement between their model and empirical data on the former LGM extent of the Alpine ice field.

In agreement with Višnević and others (2020), the resulting ELA fields from our study (Fig. 6) show a clear north-to-south gradient, with the northern catchments showing lower ELAs. However, one must note that numerous factors influence the ELA. While temperature and precipitation are the primary determinants of the ELA, other factors such as ice surface slope, debris cover, glacier orientation (whether southern or northern exposure), avalanching and shielding effects from surrounding topography also play significant roles (Anderson and Mackintosh, 2012; Kneib and others, 2024). These factors can lead to variations in ELA that are not solely attributable to direct climate conditions, making it a complex indicator to interpret in paleoclimate reconstructions.

Even when all these factors are meticulously accounted for, identifying a precise paleoclimate state from the reconstructed glacier footprints remains difficult. For instance, we could theoretically reconstruct the Alpine ice field by simulating a scenario with a cold climate and high precipitation, resulting in a lower ELA, over a short period T. Alternatively, a milder climate with lower precipitation, leading to a higher ELA, could produce similar glacier extents
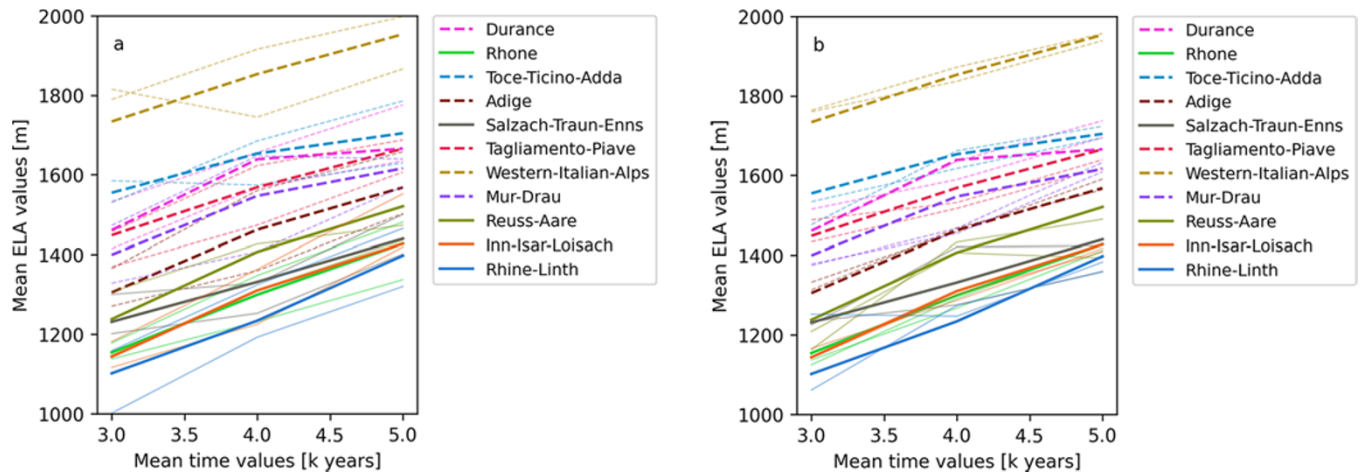
Figure 7. (a) Sensitivity of mean ELA values to variations in the mass balance gradient β_{abl}. (b) Sensitivity of mean ELA values to variations in the rate factor in Glen's flow law (A). Mean values from standard parameters are in bold colours, while halved and doubled parameters are shown with reduced opacity. Solid lines indicate northern regions; dashed lines represent southern regions.

if the period T was longer. The actual climate conditions during the LGM could have been intermediate between these two extremes, further complicating the reconstruction process. This underscores the inherent complexity and the need for a cautious interpretation of our results, particularly considering the period T.

4.2. Temporal and spatial climate variability during the LGM

Studies on the LGM climate often decouple spatial (Russo and others, 2024) and temporal (Luetscher and others, 2015) climate variability. On the one hand, climate modelling studies (e.g. Russo and others, 2024) can hardly afford to obtain multiple snapshots of spatially variable climate using regional climate models due to computational expenses. Moreover, the sparse and non-uniform distribution of proxies (e.g. Luetscher and others, 2015) reconstruct temporally variable climate proxies commonly referred to as climate signals but they provide limited spatial information. Glacier modelling studies by Jouvet and others (2023) or Seguinot and others (2018) applied global climate signals (i.e. Antarctic and Greenland ice core records) as forcing, disregarding local variations of the cold spell that can significantly impact glacier behaviour. In revisiting earlier hypotheses (Florineth and Schlüchter, 2000; Luetscher and others, 2015), our results could suggest a spatially heterogeneous period T across the Alps. The primary driver of such heterogeneity was likely variations in precipitation patterns during the LGM rather than temperature changes.

In the Tagliamento morainic amphitheatre, Monegato and others (2007) found a two-phased glacial maximum. The older, more extensive advance was dated to 26.5 and 23 ka, while the younger advance was between 24 and 21 ka. This formation could be a combination of shorter, repeated cold spells and variable alpine catchment geomorphologies. The southern Alps have narrower and shorter valleys, allowing for the development of smaller glaciers, which react more to short-term changes in climate conditions. We illustrated this in Appendix B by modelling a spatially constant ELA, therefore analysing the effects of the Alps's geomorphology on glacier formation. Even if the ELA was constant across the Alps, using the signal from Luetscher and others (2015) indicates that the glaciers in the North will grow larger than those in the South (see Appendix B).

As a result, the ELA values in our study (Fig. 6) could be interpreted in different ways: for the northern regions, the ELA values correspond to those on the longer periods T, while in the southern catchments, they align with the shorter, more episodic periods T. This interpretation suggests that differences in absolute climatic conditions (temperature, precipitation) between the northern and southern Alps, may not have been as pronounced as previously thought (Višnević and others, 2020; Jouvet and others, 2023). Instead, an alternative hypothesis would be that the larger glaciers in the north result from differences in valley and catchment geomorphology and longer cold spells.

4.3. Method limitations and future improvements

Inversion methods, including glacier dynamics for deriving past climate insights from observed glacier footprints, represent a significant advancement over the Accumulation Area Ratio approach (Martin and Monnier, 2014). Although Višnević and others (2018) pioneered this method, their approach required considerable assumptions, which are essential for its effectiveness. Firstly, they assume glaciers are in a steady state at the time of the glacier footprint formation, a condition rarely met in the dynamic nature of glacier-climate interactions. Secondly, they employ the SIA (Hughes, 1984), simplifying the ice mechanics, which is particularly problematic in mountainous regions where vertical shearing predominantly influences ice flow (Le Meur and others, 2004). Lastly, they use a heuristic descent direction not based on the gradient of the cost function, a limitation that results in slower convergence of the inversion process and restricts the method to a simplistic SMB.

Our new approach overcomes these assumptions. However, it is important to acknowledge certain limitations in the current study, many of which can nonetheless be addressed using the same methodology. Firstly, the current approach uses a relatively simplistic parameterization of SMB, which may not fully capture the complexity of SMB processes effectively as more sophisticated models, such as the positive degree day (Braithwaite, 1984; Harding, 1989) or energy SMB (Gardner and others, 2023). Furthermore, the bedrock topography, the resolution of our model, and other parameters governing glacier mechanics are fixed once

the emulator is trained. This current rigidity limits the capability of the emulator, and therefore the possibility of exploring other settings in the search for solutions.

This study demonstrated the capacity of our new method to reconstruct a plausible ELA pattern that explains the Alpine ice field geometry during the LGM. This proof of concept has the potential to be applied to other formerly glaciated regions, though it may not be suitable for the largest ice sheets due to the complex two-way coupling between climate and glaciation. Future research will aim to develop a more generic emulator that is not limited by a specific topography, SMB, or resolution. The computational efficiency of the emulator holds significant promise for broader regional applications. We plan to apply our inversion method to reconstructed glacial maximum extents from the LGM to wider areas, such as the Euro-Asian continent. Such efforts could provide new insights into the climatic dynamics that have shaped the Earth's history, contributing to a more nuanced understanding of global climate evolution during the Quaternary period.

5. Conclusion

We introduced a new inversion scheme designed to reconstruct the spatially and temporally variable ELA across the Alps during the LGM from glacier footprints. In agreement with earlier studies (Višnević and others, 2020; Russo and others, 2024), our results show a distinct spatial pattern: ELAs in northern catchments are consistently lower than those in southern catchments. However, our new method permits us to explore the duration of the cold spell period T, in addition to ELAs. We found a clear relationship between the period T and ELA across all catchments, with longer periods T correlating with higher ELAs. Building on our primary results, in agreement with paleoclimate modelling studies, incorporating the cold spell period T into our analysis allowed us to revisit the hypothesis that LGM weather patterns were dominated by southerly airflow (Florineth and Schlüchter, 2000; Ivy-Ochs and others, 2008; Luetscher and others, 2015). For this hypothesis to be consistent with glacier models, it suggests that the northern Alps would have experienced a more stable climate, while more episodic climatic conditions would have characterized the southern Alps. These nuanced interpretations underscore the complexity of glacial dynamics and highlight the importance of considering temporal variability in paleoclimate reconstructions. Our computationally efficient method has a high potential for being applied at a larger scale.

References

Abadi M and **46 others** (2016) TensorFlow: Large-scale machine learning on heterogeneous distributed systems. Available at https://arxiv.org/abs/1603.04467

Albrecht T, Winkelmann R and Levermann A (2020) Glacial-cycle simulations of the Antarctic Ice Sheet with the Parallel Ice Sheet Model (PISM) – part 2: Parameter ensemble analysis. *The Cryosphere* **14**(2), 633–656. https://doi.org/10.5194/tc-14-633-2020

Anderson B and Mackintosh A (2012) Controls on mass balance sensitivity of maritime glaciers in the Southern Alps, New Zealand: The role of debris cover. *Journal of Geophysical Research: Earth Surface* **117**(F1) https://doi.org/10.1029/2011JF002064

Bini A and **10 others** (2009) *Switzerland During the Last Glacial Maximum1:500 000.* Wabern: Bundesamt für Landestopografie. Available at https://www.geocat.ch/geonetwork/srv/api/records/75c2f212-413c-4d55-8a93-c021c8bbfc95

Braithwaite RJ (1984) Calculation of degree-days for glacier-climate research. *Zeitschrift für Gletscherkunde und Glazialgeologie* **20**(1984), 1–8. Available at: https://www.academia.edu/103555237

Braithwaite R and Raper S (2009) Estimating Equilibrium-Line Altitude (ELA) from glacier inventory data. *Annals of Glaciology* **50**(53), 127–132. https://doi.org/10.3189/172756410790595930

Cohen D, Gillet-Chaulet F, Haeberli W, Machguth H and Fischer UH (2018) Numerical reconstructions of the flow and basal conditions of the Rhine glacier, European Central Alps, at the Last Glacial Maximum. *The Cryosphere* **12**(8), 2515–2544. https://doi.org/10.5194/tc-12-2515-2018

Cong S and Zhou Y (2023) A review of convolutional neural network architectures and their optimizations. *Artificial Intelligence Review* **56**(3), 1905–1969, https://doi.org/10.1007/s10462-022-10213-5

Ehlers J, PL G and PD H (2011) Quaternary glaciations–Extent and chronology: A closer look, volume 15 of *Developments in Quaternary Science.* Elsevier, Amsterdam

Florineth D and Schlüchter C (2000) Alpine evidence for atmospheric circulation patterns in Europe during the Last Glacial Maximum. *Quaternary Research* **54**(3), 295–308. https://doi.org/10.1006/qres.2000.2169

Gardner AS, Schlegel NJ and Larour E (2023) Glacier Energy and Mass Balance (GEMB): S model of firn processes for cryosphere research. *Geoscientific Model Development* **16**(8), 2277–2302. https://doi.org/10.5194/gmd-16-2277-2023

Glen JW (1955) The creep of polycrystalline ice. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **228**(1175), 519–538. https://www.jstor.org/stable/99642

Haeberli W and Penz U (1985) An attempt to reconstruct glaciological and climatological characteristics of 18 ka bp ice age glaciers in and around the Swiss Alps. *Zeitschrift für Gletscherkunde und Glazialgeologie* **21**, 351–361. available at: https://www.research-collection.ethz.ch/handle/20.500.11850/487610

Harding, RJ (1989) *Glacier Fluctuations and Climatic Change.* https://doi.org/10.1002/joc.3370110109

Heyman BM, Heyman J, Fickert T and Harbor JM (2013) Paleo-climate of the central European uplands during the last glacial maximum based on glacier mass-balance modeling. *Quaternary Research* **79**(1), 49–54. https://doi.org/10.1016/j.yqres.2012.09.005

Hughes TJ (1984) Kolumban Hutter. *Theoretical Glaciology: Material Science of Ice and the Mechanics of Glaciers and Ice Sheets*. Dordrecht: D. Reidel Publishing Co.; Tokyo: Terra Scientific Publishing Co., 1983. xxxii+510 pp. (Mathematical Approaches to Geophysics). *J. Glaciol.* **30**(105), 254–255. https://doi.org/10.3189/S0022143000006055

Ivy-Ochs S and **7 others** (2008) Chronology of the last glacial cycle in the European Alps. *Journal of Quaternary Science: Published for the Quaternary Research Association* **23**(6-7), 559–573. https://doi.org/10.1002/jqs.1202

Jóhannesson T, Raymond C and Waddington E (1989) Time-scale for adjustment of glaciers to changes in mass balance. *Journal of Glaciology* **35**(121), 355–369. https://doi.org/10.3189/S002214300000928X

Jouvet G (2022) Inversion of a Stokes glacier flow model emulated by deep learning. *Journal of Glaciology*, 1–14. https://doi.org/10.1017/jog.2022.41

Jouvet G and **11 others** (2023) Coupled climate-glacier modelling of the last glaciation in the Alps. *Journal of Glaciology*, 1–15. https://doi.org/10.1017/jog.2023.74

Jouvet G, Cordonnier G, Kim B, Lüthi M, Vieli A and Aschwanden A (2022) Deep learning speeds up ice flow modelling by several orders of magnitude. *Journal of Glaciology* **68**(270), 651–664. https://doi.org/10.1017/jog.2021.120

Khan A, Sohail A, Zahoora U and Qureshi AS (2020) A survey of the recent architectures of deep convolutional neural networks. *Artificial intelligence review* **53**, 5455–5516. https://doi.org/10.1007/s10462-020-09825-6

Kingma DP and Ba J (2014) Adam: A method for stochastic optimization. Available at https://arxiv.org/abs/1412.6980

Kneib M and **15 others** (2024) Distributed surface mass balance of an avalanche-fed glacier. *EGUsphere* **2024**, 1–35. https://doi.org/10.5194/egusphere-2024-1733

Le Meur E, Gagliardini O, Zwinger T and Ruokolainen J (2004) Glacier flow modelling: A comparison of the Shallow Ice Approximation and the full-Stokes solution. *Comptes Rendus Physique* **5**(7), 709–722. https://doi.org/10.1016/j.crhy.2004.10.001

Luetscher M and **8 others** (2015) North Atlantic storm track changes during the Last Glacial Maximum recorded by Alpine speleothems. *Nature Communications* **6**(1), 6344, https://doi.org/10.1038/ncomms7344

Lüthi MP (2009) Transient response of idealized glaciers to climate variations. *Journal of Glaciology* **55**(193), 918–930. https://doi.org/10.3189/002214309790152519

Martin LC and **15 others** (2020) Antarctic-like temperature variations in the Tropical Andes recorded by glaciers and lakes during the last deglaciation. *Quaternary Science Reviews* **247**, 106542. https://doi.org/10.1016/j.quascirev.2020.106542

Martin N and Monnier J (2014) Adjoint accuracy for the full Stokes ice flow model: Limits to the transmission of basal friction variability to the surface. *The Cryosphere* **8**(2), 721–741. https://doi.org/10.5194/tc-8-721-2014

Meier MF (1962) Proposed definitions for glacier mass budget terms. *Journal of Glaciology* **4**(33), 252–263. https://doi.org/10.3189/S0022143000027544

Monegato G, Ravazzi C, Donegana M, Pini R, Calderoni G and Wick L (2007) Evidence of a two-fold glacial advance during the last glacial maximum in the Tagliamento end moraine system (eastern Alps). *Quaternary Research* **68**(2), 284–302. https://doi.org/10.1016/j.yqres.2007.07.002

Monegato G, Scardia G, Hajdas I, Rizzini F and Piccin A (2017) The Alpine LGM in the boreal ice-sheets game. *Scientific reports* **7**(1), 2078. https://doi.org/10.1038/s41598-017-02148-7

Ogundokun RO, Maskeliunas R, Misra S and Damaševičius R (2022) Improved CNN based on batch normalization and Adam optimizer. In *International Conference on Computational Science and Its Applications*, 593–604, Springer. https://doi.org/10.1007/978-3-031-10548-7_43

Ohmura A, Kasser P and Funk M (1992) Climate at the equilibrium line of glaciers. *Journal of Glaciology* **38**(130), 397–411. https://doi.org/10.3189/S0022143000002276

O'Shea K and Nash R (2015) An introduction to convolutional neural networks. https://doi.org/10.48550/arXiv.1511.08458

Paszke A and **9 others** (2017) Automatic differentiation in PyTorch. *OpenReview*. available at https://openreview.net/pdf?id=BJJsrmfCZ

Pellitero R and **7 others** (2015) A GIS tool for automatic calculation of glacier equilibrium-line altitudes. *Computers & Geosciences* **82**, 55–62. https://doi.org/10.1016/j.cageo.2015.05.005

Porter SC (1970) Quaternary glacial record in Swat Kohistan, west Pakistan. *Geological Society of America Bulletin* **81**(5), 1421–1446. https://doi.org/10.1130/0016-7606(1970)81[1421:QGRISK]2.0.CO;2

Rettig L, Monegato G, Spagnolo M, Hajdas I and Mozzi P (2023) The Equilibrium Line Altitude of isolated glaciers during the Last Glacial Maximum – New insights from the geomorphological record of the Monte Cavallo Group (south-eastern European Alps). *CATENA* **229** 107187. https://doi.org/10.1016/j.catena.2023.107187

Russo E and **8 others** (2024) High-resolution LGM climate of Europe and the Alpine region using the regional climate model WRF. *Climate of the Past* **20**(3) 449–465, https://doi.org/10.5194/cp-20-449-2024

Seguinot J, Ivy-Ochs S, Jouvet G, Huss M, Funk M and Preusser F (2018) Modelling last glacial cycle ice dynamics in the Alps. *The Cryosphere* **12**(10) 3265–3285, https://doi.org/10.5194/tc-12-3265-2018

Tikhonov AN (1963) On the regularization of ill-posed problems. *Doklady Akademii Nauk* **153**(1) 49–52. https://doi.org/10.1137/1.9780898717921

Višnević V, Herman F and Prasicek G (2020) Climatic patterns over the European Alps during the LGM derived from inversion of the paleo-ice extent. *Earth and Planetary Science Letters* **538**, 116185. https://doi.org/10.1016/j.jpgl.2020.116185

Višnjević V, Herman F and Podladchikov Y (2018) Reconstructing spatially variable mass balances from past ice extents by inverse modeling. *Journal of Glaciology* **64**(248) 957–968. https://doi.org/10.1017/jog.2018.82

Zekollari H and Huybrechts P (2015) On the climate–geometry imbalance, response time and volume–area scaling of an alpine glacier: insights from a 3-D flow model applied to Vadret da Morteratsch, Switzerland. *Annals of Glaciology* **56**(70) 51–62. https://doi.org/10.3189/2015AoG70A921

Appendix A.

Hyperparameters are variables set before the training process that govern the behaviour and performance of machine learning models. Unlike model parameters, which are learned during training, hyperparameters must be manually tuned. Examples include the number of epochs, batch size and learning rate. Finding optimal hyperparameter values is crucial as they directly influence the convergence speed, model accuracy and generalization to new data.

In this study, we conducted a grid search to explore the hyperparameter space, aiming to improve training efficiency and model performance. The model was trained for 100 epochs, with 80% of the dataset used for training and 20% for validation.

Each hidden layer in our CNN consists of a convolutional layer, ReLU activation and a pooling layer:
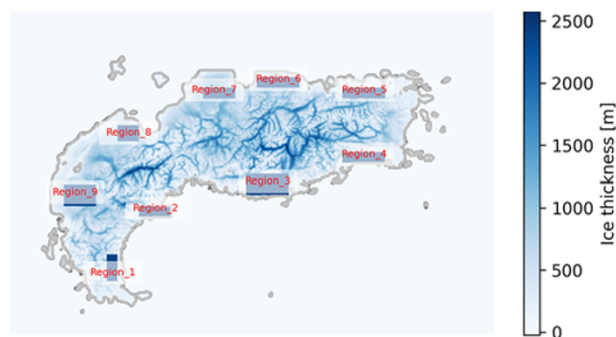
Figure A1. Geographical representation of the nine regions over the Alps used for evaluating the emulator's probable systematic bias.
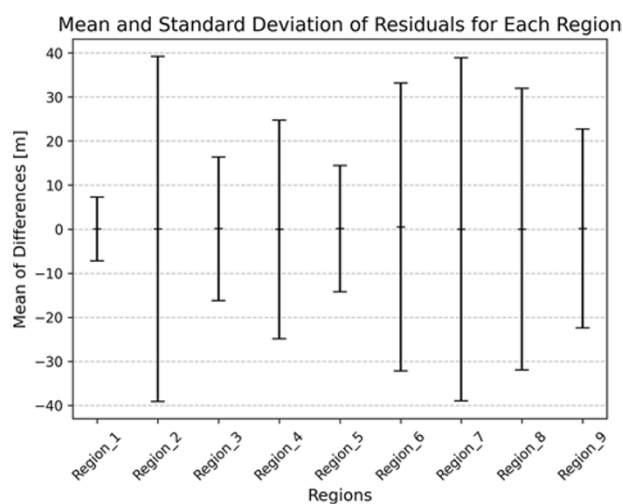

Figure A2. Mean and standard deviation of residuals for each region. The mean values are close to zero, indicating no systematic bias in the emulator's predictions.

- The convolutional layer uses a kernel (filter) to scan the input, creating feature maps. Smaller kernels capture finer details, while larger kernels capture more abstract, global features. The kernel size affects the computational cost and the network's ability to capture patterns.
- Batch size defines the number of samples processed in each training iteration. Smaller batch sizes allow the model to update weights more frequently, which can accelerate learning but introduce noise, making training less stable. Larger batch sizes produce more stable gradient estimates but require more memory and result in slower convergence (Ogundokun and others, 2022).
- The learning rate controls the step size in updating model weights. A large learning rate may cause the model to overshoot the optimal solution, while a small learning rate can slow down convergence or cause the model to get stuck in local minima.
- The CNN architecture determines its complexity. Simpler models may not capture complex patterns, while deeper networks, such as U-Net, have higher learning capacity but are prone to overfitting, especially when the data is limited (Khan and others, 2020) and are computationally more expensive.

To assess whether the emulator exhibits systematic bias that could influence our version scheme—specifically, by consistently overestimating or underestimating glacier thickness in a particular region—we design a controlled experiment. Figure A1 delineates nine regions across the Alps, within which we evaluate the emulator's performance by computing the residual between the true and predicted GMT. For each region (1–9), we determine the mean residual. Subsequently, across all data points in the test dataset, we calculate both the mean and the standard deviation (Fig. A2). The results indicate that, for all regions, the mean residual is centered near zero, suggesting that the emulator does not introduce systematic spatial bias.

Appendix B.

As an experiment to demonstrate the importance of climate-forcing signals, we designed three simulations over the Alps. Each simulation was initiated
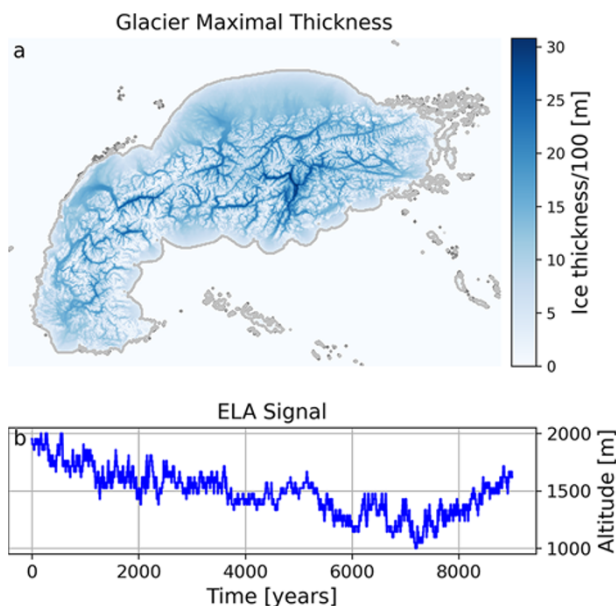
Figure B1. Panel a shows the GMT across the domain, while panel b illustrates the temporal evolution of the spatially uniform ELA, which follows the shift pattern described by Luetscher and others (2015). Despite the ELA being spatially constant, glaciers in the northern regions grow larger over time.
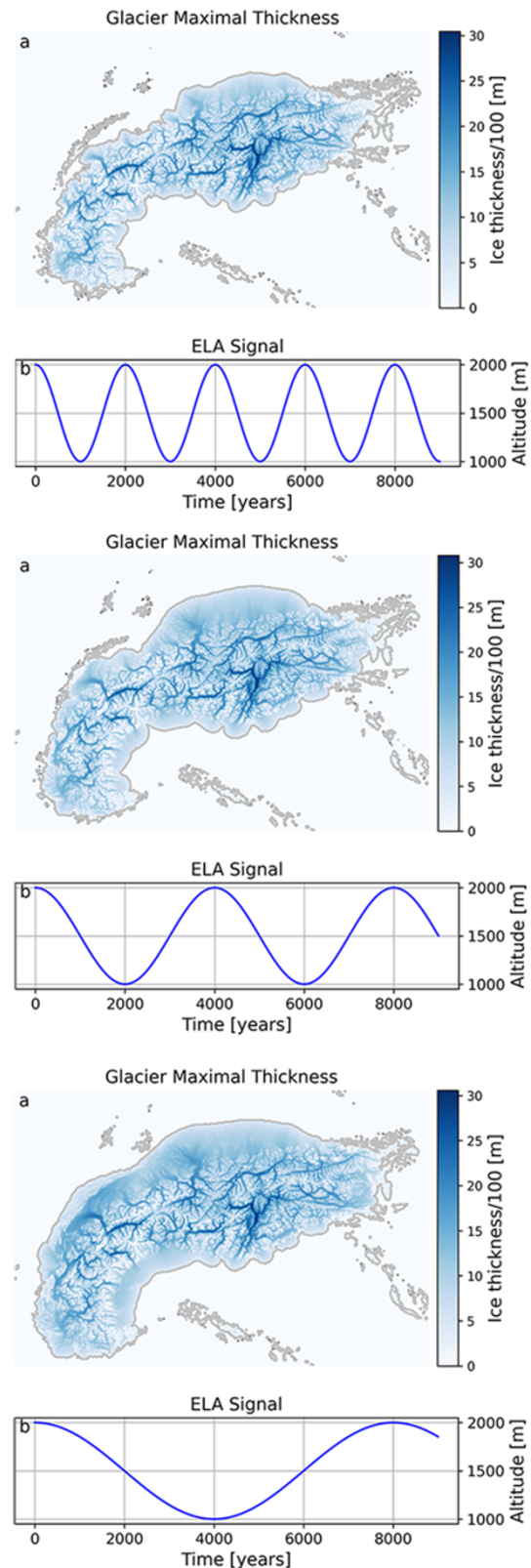

Figure B2. Evaluation of glacier response to periodic ELA forcing using a cosine signal with different periods ($T = 2000$, 4000 and 8000 years). Although all ELA values oscillate between 1000 and 2000 m, the resulting GMT fields (panel a in each subfigure) vary significantly depending on the cosine period. Longer periods allow more time for glaciers to grow and adjust, leading to a larger GMT.

with ice-free topography and a constant ELA. During the simulations, different ELA signals were applied (bottom panels in Figs B1 and B2), and the glacier maximum thickness was computed for each scenario. Despite all experiments having a minimum ELA of 1000 m, the resulting glacier maximum thickness varied significantly. We observed that when climate follows the signal from Luetscher and others (2015), glaciers in the northern Alps grow larger than those in the south (Fig. B1). Conversely, when climate transitions are slower, glaciers expand uniformly across the region. This variation is attributed to the geomorphology of the Alps, where the northern valleys are wider and longer, while the southern valleys are narrower and steeper.

Appendix C.

To determine the optimal regularization coefficient (λ) in Eqn (6), we evaluated six different values and constructed an L-curve (Fig. C1). Based on this analysis, we selected $\lambda = 10\,000$ as it provides an optimal trade-off between ensuring a smooth Equilibrium Line Altitude (ELA) field and maintaining good convergence in the inversion scheme.
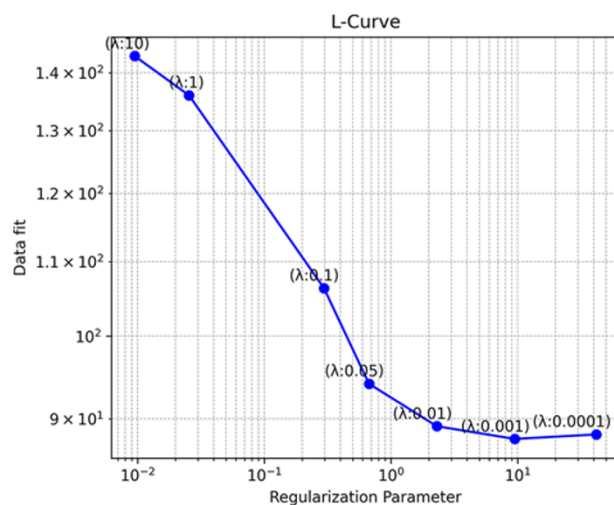
Figure C1. L-curve used to determine the optimal regularization parameter λ. The point of maximum curvature, corresponding to the optimal trade-off between model fit and regularization, is identified at $\lambda = 0.05$.