

Modelling Glacier Thickness and Evolution with Machine Learning

Evaluating the Instructed Glacier Model (IGM) on the Batysh-Sook
Glacier through Inverse and Forward Modelling

Master's thesis

Robbe BULS

Supervisor: Prof. Dr. P. Huybrechts
Vrije Universiteit Brussel

Cosupervisor: *Dr. L. Van Tricht*
Vrije Universiteit Brussel; ETH Zürich

Master thesis submitted in fulfillment
of the requirements for the degree in
Master of Science in Geography

Academic year 2024-2025

© Copyright by KU Leuven

Without written permission of the promoters and the authors it is forbidden to reproduce or adapt in any form or by any means any part of this publication. Requests for obtaining the right to reproduce or utilize parts of this publication should be addressed to KU Leuven, Faculteit Wetenschappen, Geel Huis, Kasteelpark Arenberg 11 bus 2100, 3001 Leuven (Heverlee), Telephone +32 16 32 14 01. A written permission of the promoter is also required to use the methods, products, schematics and programs described in this work for industrial or commercial use, and for submitting this publication in scientific contests.

© Copyright by Vrije Universiteit Brussel

Without written permission of the promoters and the authors it is forbidden to reproduce or adapt in any form or by any means any part of this publication. Requests for obtaining the right to reproduce or utilize parts of this publication should be addressed to Vrije Universiteit Brussel, Faculteit Wetenschappen en Bio-ingenieurswetenschappen, Pleinlaan 2, 1050 Brussel, Telephone +32 (0)2 629 33.57. A written permission of the promoter is also required to use the methods, products, schematics and programs described in this work for industrial or commercial use, and for submitting this publication in scientific contests.

Contribution Statement

This master thesis was written by Robbe Buls under the supervision of Prof. Dr. Philippe Huybrechts and Dr. Lander Van Tricht.

Fieldwork for this thesis was conducted by Lander Van Tricht in the summer of 2023. This fieldwork resulted in ice thickness and elevation data. Mass balance measurements were done by Ruslan Kenzhebaev.

The processing of the radio measurements were done by Robbe Buls. The processing of the elevation data into an elevation model was done by Lander Van Tricht. All other processing and analysis steps described in the thesis were performed by Robbe Buls.

Generative AI was used to correct and debug coding for the analysis and visualisations. Everytime GenAI was used for coding, the code snippets were inspected by Robbe Buls himself.

Acknowledgements

Writing a thesis can be compared to riding a cycling race. A course with long, steady climbs that burn in your legs, stretches of cobblestones that rattle your focus, and days when the rain falls sideways and the headwind won't let up. But it also includes rolling roads, breathtaking views, and glorious stretches of downhill. Whoever wants to descend, first has to climb. Fortunately, I was able to rely on strong riders who tucked me into their slipstream. Just like in a race, writing a thesis may be an individual effort, but without teammates, you won't get far.

A good team depends on wise leadership. Riders must trust the team car blindly — those who read the race like a book, know every turn of the road by heart, and guide their riders with calm precision to the finish line. I consider myself lucky to have had such an excellent team behind me. My sincere thanks go to Prof. Philippe Huybrechts, whose many years of research experience lifted my thesis to a higher level. I also want to express my deep gratitude to Lander Van Tricht for the solid bidon collé he handed me. Without his support and his quick replies to all my questions (even the silly ones), I surely would have finished outside the time limit. A heartfelt thanks to both of you as well for the high-altitude training camp in Kyrgyzstan. The fieldwork was an unforgettable experience, and taught me more than I could've imagined about glaciological research.

Of course, every good rider needs solid mechanics. At times, this journey felt like a stretch of Paris-Roubaix: no flat tires, but plenty of rattling cobbles, near-crashes, and code-induced cramps. A special shoutout to Joachim Piret, who gave me a strong push when it came to everything IGM-related. The folks at the VUB helpdesk and Chloë Paice also deserve my deepest respect. When I got lost using Hydra, they showed me the way and got me back on track. Without them, I might have ended up with a DNF behind my name.

My sponsors deserve a heartfelt mention as well. Not only did they support my incredible five years in Leuven (and a bit in Brussels), they also backed the eighteen years before that. My most loyal supporters — who not only funded my long ride but also cheered at every feed zone, stitched up the crashes, and kept the morale high on the worst of days. They were my lifelong soigneurs, keeping me fed, focused, and on the road — even on the toughest stages.

Finally, I want to thank all my teammates and training buddies. Despite spending our first year behind facemasks and computer screens, we grew into a tight wolfpack — the strongest team I can imagine. That goes not only for my classmates, but also for the many geographers from the years before and after us. Beyond the classroom climbs and fieldwork flats, we trained together on winding trails through Scotland, the deceptively scenic peaks of Aosta (where the water hit harder than the altitude) and the long road to Barcelona. Not every stage went smoothly, but with wraps stuffed with choco, a bottle of after-sun, and a trusty

Walker biscuit, together we always kept going (as soon as Lukas arrived). Now that our team time trial is nearly over and the road ahead stretches out before us, remember: the peloton's never far behind. If ever you need a wheel to follow or someone to break the wind, I'll be more than happy to take the lead for you.

This thesis marks the sprint of a five-year ride that has taken me to beautiful places, introduced me to wonderful people, and brought unforgettable moments. Choosing to study geography was, without a doubt, the best decision I could have made. It's been the ride of a lifetime — a route full of twists, turns, and unexpected detours, but all worth it. So, a sincere thank you to everyone who made that possible. Now, with the flamme rouge overhead and the finish line in sight, all that's left is the booming voice over the speaker, the click of pedals, and the rush of adrenaline — like in the Ronde van Vlaanderen: "Thesis, vertrekt!"

Samenvatting

De afvoer van smeltwater uit de gletsjers van het Tien Shan-gebergte in Kirgizië is van cruciaal belang voor de samenleving en economie in de stroomafwaarts gelegen vlaktes. In het droge Centraal-Azië is de landbouw sterk afhankelijk van irrigatie met dit water. De gletsjers functioneren als watertorens: ze herverdelen de schaarse neerslag over de seizoenen, wat grootschalige landbouw mogelijk maakt. Door klimaatverandering trekken de gletsjers zich echter snel terug, waardoor de inwoners van de vlaktes hun belangrijkste inkomensbron dreigen te verliezen.

Het doel van deze thesis is het evalueren van het Instructed Glacier Model (IGM). Dit model gebruikt een neurale netwerk als emulator om ijsstroming te beschrijven, wat computationele voordelen oplevert en het mogelijk maakt het model zowel invers (interpoleren van ijsdiktemetingen) als voorwaarts (voorspellen van gletsjerevolutie onder klimaatscenario's) te gebruiken. De prestaties worden vergeleken met klassieke ruimtelijke interpolatiemethoden, zoals de vloeispanningsmethode, en met een hoger-ordemodel dat de gletsjerevolutie simuleert. De evaluatie vindt plaats op de Batysh-Sook-gletsjer in de Tien Shan.

De resultaten tonen dat het geïnverteerde IGM consistente, nauwkeurige en gladde ijsdiktevelden genereert. Vooral in schaars bemonsterde zones presteert het model beter dan de conventionele vloeispanningsmethode. Voor de gletsjerevolutie laten zowel IGM als het hoger-ordemodel zien dat de Batysh-Sook-gletsjer gevoeliger is voor opwarming dan veel omliggende gletsjers, waarschijnlijk door haar kleine huidige omvang. IGM volgt de trends van het hoger-ordemodel goed, al simuleert het in de historische periode iets minder smelt, waardoor het gletsjervolume groter blijft.

Deze bevindingen tonen aan dat IGM een veelbelovend instrument is voor zowel ijsdikteinterpolatie als projectie, vooral in gebieden met beperkte veldmetingen. Betere modellering van gletsjers zoals de Batysh-Sook kan leiden tot nauwkeurigere voorspellingen van toekomstige waterbeschikbaarheid, wat van groot belang is voor wetenschappelijk onderzoek en voor duurzaam waterbeheer in Centraal-Azië.

Summary

The runoff of meltwater from the glaciers of the Tien Shan mountains in Kyrgyzstan is of crucial importance to the society and economy of the downstream plains. In the arid regions of Central Asia, agriculture relies heavily on irrigation with this water. The glaciers act as natural water towers: they redistribute scarce precipitation over the seasons, enabling large-scale agriculture. However, due to climate change, the glaciers are retreating rapidly, putting the main source of income for the inhabitants of the plains at risk.

The aim of this thesis is to evaluate the Instructed Glacier Model (IGM). This model uses a neural network as an emulator to describe ice flow, providing computational advantages and allowing the model to be used both inversely (interpolating ice thickness measurements) and forward (predicting glacier evolution under climate scenarios). Its performance is compared with classical spatial interpolation methods, such as the flowline method, and with a higher-order model that simulates glacier evolution. The evaluation focuses on the Batysh-Sook Glacier in the Tien Shan.

The results show that the inverted IGM produces consistent, accurate, and smooth ice thickness fields. In sparsely sampled areas in particular, the model outperforms the conventional flowline method. For glacier evolution, both IGM and the higher-order model indicate that the Batysh-Sook Glacier is more sensitive to warming than many surrounding glaciers, likely due to its small current size. IGM follows the trends of the higher-order model well, although it simulates slightly less melt during the historical period, resulting in a larger retained glacier volume.

These findings demonstrate that IGM is a promising tool for both ice thickness interpolation and glacier projection, especially in regions with limited field measurements. Improved modelling of glaciers such as the Batysh-Sook can lead to more accurate predictions of future water availability, which is of great importance for scientific research as well as for sustainable water management in Central Asia.

Glossary

α albedo.

ASL basal sliding parameter ($m \cdot s^{-1}$).

$A(T)$ flow rate factor ($Pa^{-n} \cdot s^{-1}$).

c velocity of light in vacuum ($m \cdot s^{-1}$).

χ weighted average of the accuracy and the smoothness.

δ electrical conductivity of ice ($S \cdot m^{-1}$).

d flux divergence ($m \cdot yr^{-1}$).

E enhancement factor.

$\dot{\epsilon}$ strain rate (s^{-1}).

$\dot{\epsilon}_e$ effective strain rate (s^{-1}).

ε_{ij} strain.

ε_r permittivity of ice ($F \cdot m^{-1}$).

ϵ_t error on the travel time (s).

η effective viscosity ($Pa \cdot s$).

γ spatial variability / semivariogram values.

γ_T temperature lapse rate ($K \cdot km^{-1}$).

h ice thickness (m).

H_{ref} elevation of the meteostation (m).

$I_{i,j}$ discrete Laplacian.

J cost function inverse IGM.

$l_{i,j}$ distance to the nearest glacier margin (m).

λ_i kriging weight on point i .

M ablation ($kg \cdot m^2$).

μ Lagrange multiplier in semivariogram.

n creep exponent.

ω angular frequency ($rad \cdot s^{-1}$).

P cryostatic pressure (Pa).

P_s solid precipitation (mm).

ϕ surface slope (mm^{-1}).

P_{ratio} coefficient to calculate local precipitation based on meteorostation.

P_{ref} precipitation at the meteorostation (mm).

Ψ energy Balance (J).

Q activation energy for creep (J).

Q_{in} incoming shortwave radiation ($W \cdot m^2$).

R universal gas constant ($J \cdot mol^{-1} K^{-1}$).

R_w part of the meltwater that stays within the snowpack (mm).

s surface elevation (m).

τ stress (Pa).

τ_0 yield stress (Pa).

τ_b basal shear stress (Pa).

τ' deviatoric stress (Pa).

τ_e effective stress (Pa).

T_m temperature at the pressure melting point (K).

T^* temperature adjusted for pressure melting point (K).

D Kolmogorov Smirnov test statistic.

I Moran's I test statistic.

u velocity component in the x-direction ($m \cdot yr^{-1}$).

u_b basal sliding velocity ($m \cdot s^{-1}$).

v velocity component in the y-direction ($m \cdot yr^{-1}$).

v_{ice} propagation velocity of a wave through ice ($m \cdot s^{-1}$).

φ transmissivity of the atmosphere.

$\vec{v}$ vertically averaged horizontal velocity ($m \cdot yr^{-1}$).

w velocity component in the z-direction ($m \cdot yr^{-1}$).

ζ dimensionless vertical coordinate.

Acronyms

ANN Artificial Neural Network.

CNN Convolutional Neural Network.

CPU Central processing Units.

DEM Digital Elevation Model.

ELA Equilibrium Line Altitude.

GCM Global Circulation Models.

GCPs Ground Control Points.

GHG Greenhouse Gas.

GLOF Glacial Lake Outburst Floods.

GPU Graphical Processing Units.

HOM Higher Order Model.

IGM Instructed Glacier Model.

IPCC Intergovernmental Panel on Climate Change.

ITMIX Ice Thickness Model Intercomparison eXperiment.

KS Kolmogorov Smirnov test.

LIA Little Ice Age.

PDP Particle Dependence Plot.

PINN Physics-Informed Neural Network.

PISM Parallel Ice Sheet Model.

PMP Pressure Melting Point.

PSO Particle Swarm Optimisation.

RES Radio Echo Sounding.

RGI Randolph Glacier Inventory.

RMSE Root Mean Square Error.

RWV Radio Wave Velocity.

SHAP Shapely Additive exPlanations.

SIA Shallow Ice Approximation.

SMB Specific Mass Balance.

SSA Shallow Shelf Approximation.

SSP Shared Socioeconomic Pathways.

SW Shapiro Wilk test.

UAV Unmanned Aerial Vehicle.

VPINN Variational Physics-Informed Neural Network.

WGMS World Glacier Monitoring Service.

List of Figures

1.1	Digital elevation model of Kyrgyzstan and the surrounding countries. The locations of the meteo-stations of Bishkek and Taldy-Suu (Талды-Суу) are indicated.	2
1.2	Climate charts for Bishkek (left) and Taldy Suu (right) (Climate-Data.org, 2024)	3
1.3	Schematic illustration of the changes in runoff from a glacierized basin in response to continuous atmospheric warming (Huss and Hock, 2018)	6
1.4	Flowchart of the research process.	9
2.1	Schematic representation of a RES measurement. The transmitter sends a radiowave which is reflected by the bedrock. The reflected wave is measured by the receiver at a distance d from the transmitter. The time the wave needs to reflect represents the ice thickness h	11
2.2	Locations and ice thickness values of Radio Echo Sounding (RES) measurements	12
2.3	Locations of the different subsets on a Digital Elevation Model (DEM). In the left figure, the red dots represent the bottom 50% measurements and the blue dots represent the top 50% measurements. The second figure represents the random 50% measurements, the third figure represents the random 90% measurements and the figure on the right depicts the measurements along the central flowline.	14
2.4	Theoretical semivariogram with the spatial variability γ on the vertical axis and the lag on the horizontal axis (ESRI, 2023b). The indicated parameters are the nugget, the range and the sill.	16
2.5	Theoretical diagram of the strain rate with the stress for different creep relations ($n = 1$ and $n = 3$). The yield stress is indicated with a vertical dashed line. . .	18
2.6	Surface velocity based on Millan et al. (2022).	21
2.7	Mean specific mass balance based on stake measurements of Ruslan Kenzhebaev for the period 2011-2024	27
2.8	Mass balance profiles based on stake measurements for the period 2011-2024 .	28
2.9	Figure taken from Van Tricht et al. (2021b). Historic reconstruction of the temperature (a) and the precipitation (c) for the Kumptor-Tien Shan Meteo-station (3659 m). Plots (b) and (d) represent the anomalies from the mean 1821-1850 climate	30
2.10	Example of the measurement of the length of the glacier on satellite images, here for a Landsat image of 25/09/2025. The points are used to determine the length of the glacier.	36

2.11	Example of identification process of the crevasses. On the satellite image from 2016 (left), a crevasse was identified with a red line. The same crevasse is identified on the satellite image from 2017 (right). The distance with which the crevasse moved indicates the surface velocity of the glacier.	37
2.12	Flowchart of IGM modules (Jouvet et al., 2024)	38
3.1	Ice thickness fields for the different interpolation methods based on the different subsets.	42
3.2	Residual maps of the different interpolation techniques. The residuals are depicted for all RES-measurements, both the points used for the interpolation and the unused points.	43
3.3	Histograms of the residuals of the different interpolation techniques. The residuals are calculated for the RES-measurements that were not used for the interpolation. The vertical red line represents the null-line.	45
3.4	QQ-plots of the residuals. On the horizontal axis, the theoretical quantiles of a normal distribution are placed. The quantiles of the observed distribution are plotted on the vertical axis. The dashed red line depicts the -1:1 line, representing the normal distribution.	46
3.5	RMSE of the interpolation techniques	47
3.6	Moran's I of the residuals	48
3.7	Kolmogorov-Smirnov test statistics of the interpolation techniques. Lower values mean more normally-distributed residuals.	49
3.8	Parameter importance of the Instructed Glacier Model (IGM) parameters. The green barplots represent the feature importance for the Root Mean Square Error (RMSE) (top) and the Laplacian (bottom). The yellow barplots represent the permutation importance for the same two indicators.	50
3.9	Partial dependence plots of the IGM-parameters for the RMSE. Lower values correspond to better model performance (see methods section).	50
3.10	Partial dependence plots of the IGM-parameters for the laplacian	51
3.11	Summary plot Shapely Additive exPlanations (SHAP) values for the IGM interpolations. The vertical axis ranks the parameters by importance, from most to least influential. In the plot, low parameter values are coloured blue, and high values are coloured red. A negative SHAP value indicates that the parameter value pushes the prediction lower; a positive SHAP value means it pushes the prediction higher.	52
3.12	Maps of the interpolated yield stresses along the central flowlines of the different subsets	53
3.13	Boxplots of the interpolated yield stresses along the central flowlines of the different subsets	54
3.14	Heatmap of specific mass balance RMSE compared to point stake observations	54
3.15	Heatmap of mean specific mass balance RMSE compared to the geodetic mean specific mass balance for the period 2000-2010	55
3.16	Heatmap of mean specific mass balance RMSE compared to the geodetic mean specific mass balance for the period 2010-2020	55
3.17	Heatmap of mean specific mass balance RMSE compared to the geodetic mean specific mass balance for the period 2000-2020	56

3.18	Profile of the multi-year mean modelled and measured Specific Mass Balance (SMB) for C0=18, C1=-179	56
3.19	Heatmap of specific mass balance RMSE based on the elevation profiles of mass balances	57
3.20	Heatmap of specific mass balance RMSE based on the elevation profiles of mass balances	57
3.21	Comparison of the modelled and measured mean specific mass balance	58
3.22	Comparison of the modelled and measured cumulative mean specific mass balance	58
3.23	Comparison of the yearly profiles of the measured and modelled SMB	59
3.24	Dynamic calibration: The temperature is adjusted by adding a certain amount of °C. The precipitation is summed with a certain fraction of the original precipitation.	60
3.25	Glacier length modelled by Higher Order Model (HOM) and by IGM compared with the glacier length measured on Landsat Images.	61
3.26	Heatmap of the ice thickness RMSE	62
3.27	Heatmap of the ice thickness RMSE	62
3.28	Visualisation of the ϵ -constraint method to determine the optimal values of A0 and ASL.	63
3.29	Comparison glacier area (left column) and volume (right column) evolution for HOM (top line) and IGM (bottom line). The black line represents the historic period. The coloured lines represent different Shared Socioeconomic Pathways (SSP)s, with the standard deviations shaded.	64
3.30	Future evolution of the area (left panel) and the volume (right panel) for both HOM (full lines) and IGM (dashed lines) for SSPs 1-1.9 (blue), 2-4.5 (orange) and 5-8.5 (red). The black dot represents the interpolation of the ice thickness from the first part of this thesis.	64
3.31	Ice thickness maps of the HOM simulations for SSP1-1.9, SSP2-4.5 and SSP5-8.5 in the years 2050, 2075 and 2100.	66
3.32	Ice thickness maps of the IGM simulations for SSP1-1.9, SSP2-4.5 and SSP5-8.5 in the years 2050, 2075 and 2100.	67
3.33	Boxplots of the ice thicknesses for SSP1-1.9 (upper panel), SSP2-4.5 (middle panel) and the SSP5-8.5 (lower panel). The depicted years are, from left to right, 1900, 1950, 2000, 2050 and 2100. The blue boxes represent the HOM result, the red boxes are the IGM-interpolations	68
3.34	Cross-section in the year 2023. The RES-measurements are placed on the cross-section as dots. The black line is the interpolation created in the first part of this thesis. The cross-sections of the models in 2023 (HOM in blue and IGM in red) are added to the plot.	69
3.35	Cross-sections of the simulations for SSP1-1.9, SSP2-4.5 and SSP5-8.5 in the years 2050, 2075 and 2100 for both the HOM (blue) and IGM (red).	70
3.36	Cross-sections over time for HOM (upper row) and IGM (bottom row) simulations for SSP1-1.9, SSP2-4.5 and SSP5-8.5. The colour of the surface represents the year.	71
3.37	Annual stacked barplots representing the accumulation (blue) and ablation (red) area for SSP1-1.9, SSP2-4.5 and SSP5-8.5 for both HOM (left column) and IGM (right column).	72

3.38	Surface velocity maps of the Batysh-Sook glacier on different moments in time for both HOM (first column) and IGM (second column) under SSP2-4.5. In the third column, the difference map between the HOM and IGM fields are depicted.	73
3.39	Evolution of the glacier-wide mean surface velocity of the Batysh-Sook glacier for HOM and IGM.	74
3.40	Scatterplot of the annual runoff for the different SSPs concerning the HOM. The trendline is the moving average over 11 years. The vertical line indicates the moment of the peakrunoff as predicted for SSP2-4.5.	75
3.41	Scatterplot of the annual runoff for the different SSPs concerning IGM. The trendline is the moving average over 11 years. The vertical line indicates the moment of the peakrunoff as predicted for SSP2-4.5.	75
B.1	Parameter importance of the kriging parameters. The green barplots represent the feature importance for the RMSE (top) and the Laplacian (bottom). The yellow barplots represent the permutation importance for the same two indicators.	97
B.2	Parameter importance of the ANUDEM parameters. The green barplots represent the feature importance for the RMSE (top) and the Laplacian (bottom). The yellow barplots represent the permutation importance for the same two indicators.	98
B.3	Partial dependence plots of the kriging parameters for the RMSE. Lower values correspond to better model performance (see methods section).	98
B.4	Partial dependence plots of the ANUDEM parameters for the RMSE. Lower values correspond to better model performance (see methods section).	99
B.5	Partial dependence plots of the kriging parameters for the Laplacian. Lower values correspond to better model performance (see methods section).	99
B.6	Partial dependence plots of the ANUDEM parameters for the Laplacian. Lower values correspond to better model performance (see methods section).	99
B.7	Summary plot SHAP values for the kriging interpolations. The vertical axis ranks the parameters by importance, from most to least influential. In the plot, low parameter values are coloured blue, and high values are coloured red. A negative SHAP value indicates that the parameter value pushes the prediction lower; a positive SHAP value means it pushes the prediction higher.	100
B.8	Summary plot SHAP values for the ANUDEM interpolations. The vertical axis ranks the parameters by importance, from most to least influential. In the plot, low parameter values are coloured blue, and high values are coloured red. A negative SHAP value indicates that the parameter value pushes the prediction lower; a positive SHAP value means it pushes the prediction higher.	100

List of Tables

1.1	Characteristics of the glaciers studied by Van Tricht and Huybrechts (2023) and of Batysh-Sook. The characteristics represent the situation at the moment of measurements: Ashu-Tor 2019, Bordu 2017, Golubin 2019, Grigoriev 2021, Kara-Batkak 2017, Sary-Tor 2013, Batysh-Sook 2023.	4
2.1	Description of kriging parameters used.	16
2.2	Description of ANUDEM parameters used.	17
2.3	Description of IGM parameters used in the optimization process.	22
2.4	Ranges of parameter values within which the PSO algorithm was performed.	25
2.5	Geodetic mass balance of the glacier by Hugonnet et al. (2021)	28
2.6	Comparison of HOM and IGM model characteristics.	40
3.1	Summary of projected glacier area and volume for each SSP and model (HOM, IGM). For scenarios where the glacier does not vanish, the remaining area and volume in 2100 are shown.	65
4.1	Projected percentage of glacier area and volume remaining relative to the current situation (2021, except for Batysh-Sook) values under different SSPs by 2100 concerning Van Tricht and Huybrechts (2023)	80
4.2	Projected peak runoff years for the glaciers studied by Van Tricht and Huybrechts (2023) and Batysh-Sook).	81
A.1	Performance metrics for different interpolation methods and training sets: the RMSE, the Moran's I, the Kolmogorov Smirnov test (KS) test statistic and the p-value of that KS. The p-values tested to be significant at a significance level of 0.05 are coloured.	95

Contents

Contribution	v
Acknowledgements	i
Samenvatting	iii
Summary	v
Glossary	vii
Acronyms	xi
List of Figures	xii
List of Tables	xvi
Contents	xvii
1 Introduction	1
1.1 Study Area	1
1.2 Climate change	5
1.3 Social relevance	5
1.4 Instructed Glacier Model	7
1.5 Research questions	8
2 Methods	11
2.1 Ice thickness	11
2.2 Mass balance model	26
2.3 Glacier evolution models	31
3 Results	41
3.1 Ice thickness	41
3.2 Mass balance	54
3.3 Glacier evolution	60
3.4 Runoff	74
4 Discussion	77
4.1 Ice thickness	77
4.2 Glacier evolution	79

5	Conclusions	85
5.1	Batysh Sook evolution	85
5.2	Performance of IGM for ice thickness interpolations	85
5.3	Performance of IGM for modelling glacier evolution	85
	Bibliography	87
	Appendices	93
A	Validation interpolation methods	95
B	Sensitivity analysis spatial interpolation methods	97

Chapter 1

Introduction

According to the Randolph Glacier Inventory (RGI), planet Earth consists of 274 531 different glaciers, good for 706 744 km² area. 75 613 of those glaciers are located in Central Asia with an area of 50 344 km² (RGI 7.0 Consortium, 2023). These glaciers have a leading role in the hydrology of the downstream areas to sustain large-scale irrigation-dependent agriculture. Climate change places this water supply under great pressure. The enhanced melting and possible disappearance of these glaciers can disturb the livelihoods of people living in the region.

Since the late nineteenth century, early glaciologists have carried out research in the Tien Shan mountain range (Kick, 1960). In the 1980s, the Soviets started glaciological programmes, carrying out accurate mass balance measurements for several glaciers in the mountain range. After the collapse of the USSR, the monitoring projects suddenly stopped. Since 2010, new projects to monitor different glaciers in Central Asia have started (Hoelzle et al., 2017b). This thesis aims to contribute to this reinitiating of the glaciological research in the area by investigating one of the monitored glaciers in detail. More specifically, the Batysh-Sook (Suek) glacier in Kyrgyzstan will be the subject of this research.

While many recent efforts have focused on developing regional-scale models to simulate the evolution of glacier systems across the Tien Shan (Van Tricht et al., 2023b), the present study takes a complementary approach. Although regional models are useful for capturing large-scale trends and estimating glacier runoff on a broader scale, they inevitably rely on simplifications. By contrast, modelling individual glaciers allows for higher accuracy and a more detailed understanding of glacier-specific processes. This study therefore focuses on the Batysh-Sook glacier as a case study, aiming to enhance process-level understanding and support future regional modelling efforts through improved local calibration.

1.1 Study Area

Tien Shan mountain range

The Tien Shan is a wide and high mountain range in Central Asia on the territories of Kyrgyzstan, China, Uzbekistan and Kazakhstan. The highest point, Jengish Chokusu (China), reaches an elevation of 7439 m above mean sea level (figure 1.1).

Figure 1.2 depicts a climate chart for Bishkek, located in the lowlands, and for Taldy Suu

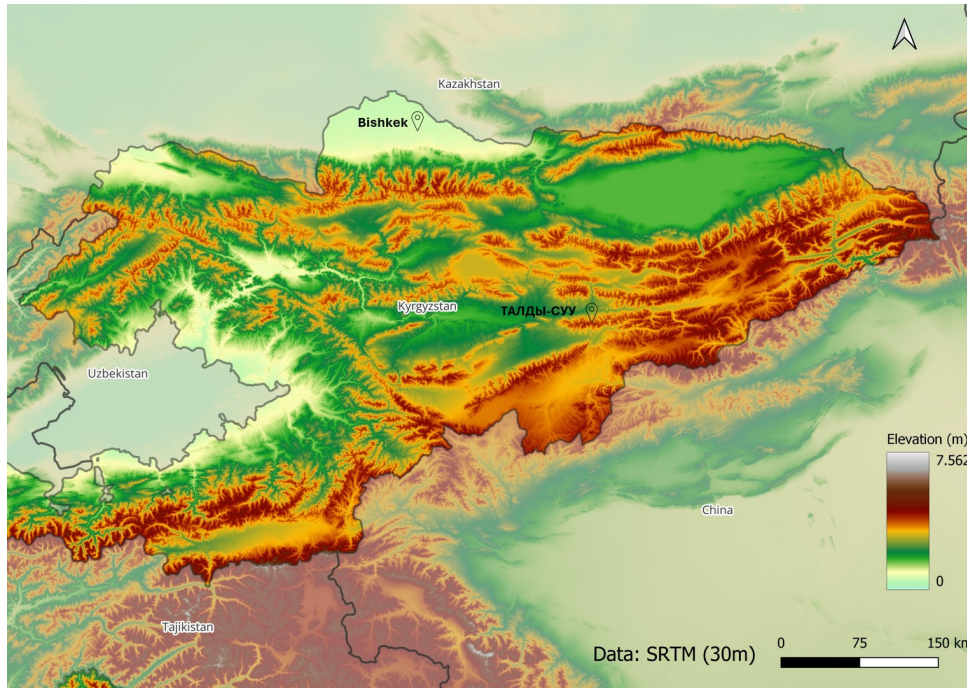


Figure 1.1: Digital elevation model of Kyrgyzstan and the surrounding countries. The locations of the meteo-stations of Bishkek and Taldy-Suu (Талды-Суу) are indicated.

(Талды-Суу), located in the Tien Shan in the upstream part of the Big Naryn basin. Temperatures reach a maximum between June and August. The temperature is mainly determined by the altitude: the mean temperature in Bishkek (750 m) is 9.8°C while in Taldy Suu (3067 m) a mean temperature of -4.0°C is noted. Due to the continental location of the mountain range and the influence of the Siberian high, Kyrgyzstan has a dry climate. Precipitation patterns in the Tien Shan are largely influenced by orographic effects. During winter, westerly winds bring relatively moist air to the region. The northern part receives significantly more precipitation than the southern parts, due to the blocking effect of the mountain range itself. The western and northern part of the Tien Shan experience two seasons with more precipitation: one between March and May, and one in October and November. In these regions, the summer is the driest period of the year. The southern and eastern parts, which are located in the rain shadow of the mountain range, only experience a rainy season in summer (Aizen et al., 1997). In spring and summer, the upper parts are often confronted with showers and thunderstorms, caused by convection and the formation of thermal lows, which are typically for subtropical areas (Johnson, 2003).

Kyrgyzstan used to be part of the former Soviet Union. In the 1960s, the Soviet government decided to organise a large-scale programme to ensure food and cotton supply for the country. Given the dryness of the region, large amounts of irrigation are required for viable agriculture. For that purpose, large dams were built on several rivers draining Central Asia. This had severe consequences for the water supply in the downstream areas. The most well-known consequence is the dramatic drying up of the Aral Sea (Levintanus, 1992). Despite the large disadvantages for the local environment, the agriculture (mainly the cotton cultivation) indeed expanded rapidly. The countries in Central Asia became one of the largest producers of cotton in the world during the 1970s and 1980s. After the collapse of the Soviet Union, the governance

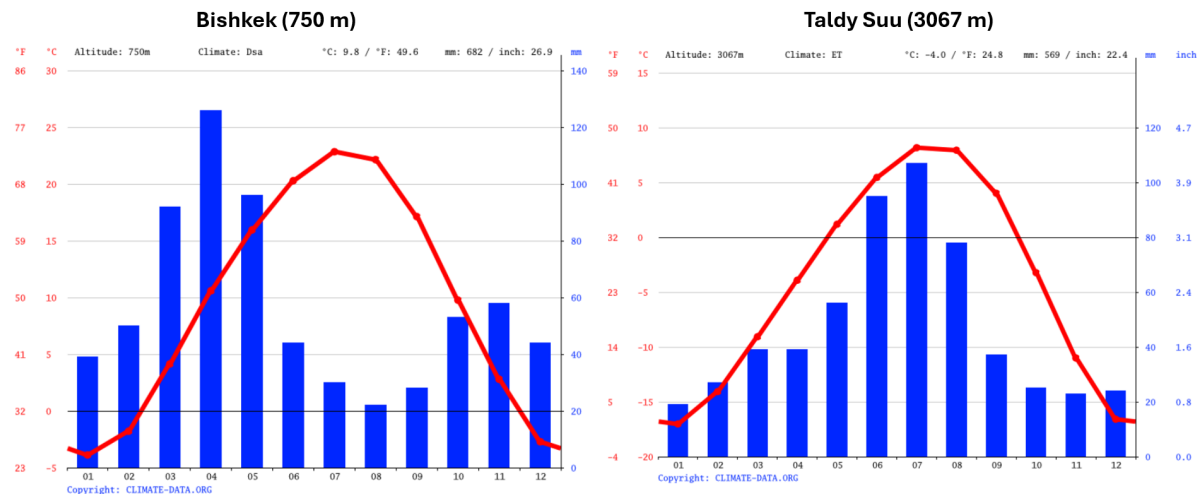


Figure 1.2: Climate charts for Bishkek (left) and Taldy Suu (right) (Climate-Data.org, 2024)

of the central authority in Moscow disappeared. As a consequence, the water policy, that was already problematic, totally lost control. Disagreements between the countries and a serious lack of funds caused bad maintenance of infrastructure and serious overconsumption of irrigation water. These human interventions cause desertification, leading to dust storms, decline in agricultural yields, etc. (Sehring, 2007).

The historic land use causes today a water dependence for Kyrgyzstan. Agriculture counts for 44% of export for the whole of Kyrgyzstan (The Observatory of Economic Complexity, Accessed 2024). This path dependence makes the Kyrgyz economy highly dependent on the cotton cultivation, and therefore also on the irrigation, which mainly comes from melt water of the glaciers and the snowpack in the Tien Shan and other mountain ranges (Chen et al., 2019; Sehring, 2007). However, the extent of this dependence varies regionally and seasonally. Studies in the northern and central Tien Shan show that while glacier melt contributes significantly to river runoff (12–59% during peak melt periods), snowmelt (22–49%) and base-flow (30–75%) are also major sources (Chen et al., 2019). This suggests that Kyrgyzstan's agricultural water security is not solely tied to glacial melt, but also to snow hydrology and groundwater recharge. Nevertheless, as climate change accelerates glacier retreat, long-term water availability may decline, exacerbating risks for cotton-dependent regions. Thus, while irrigation systems currently rely on a mix of meltwater sources, future shifts in hydrology could destabilize the agricultural sector unless adaptive measures are implemented.

Besides the need for the melt water to irrigate the agriculture fields, it also plays a crucial role in the energy supply for Kyrgyzstan and the surrounding countries at the foot of the mountains. The damming of water in upstream countries like Kyrgyzstan or Tajikistan for hydropower makes the downstream countries (Uzbekistan and Kazakhstan) face water shortages for their irrigation. Meltwater management is a major source of geopolitical tension between the downstream countries. Military tensions and potential water wars pose threats that circle above Central Asia (Bernauer and Siegfried, 2012).

Recently, a number of glaciers in the area has been modelled by Van Tricht and Huy-

brechts (2023). Table 1.1 contains the characteristics of these glaciers, together with the characteristics of Batysh-Sook glacier, measured as described in the methods section.

Table 1.1: Characteristics of the glaciers studied by Van Tricht and Huybrechts (2023) and of Batysh-Sook. The characteristics represent the situation at the moment of measurements: Ashu-Tor 2019, Bordu 2017, Golubin 2019, Grigoriev 2021, Kara-Batkak 2017, Sary-Tor 2013, Batysh-Sook 2023.

Glacier	Length (km)	Area (km ²)	Volume (km ³)	Max. Thickness (m)	Mean Altitude (m)	Orientation
Ashu-Tor	4.81	4.76	0.43	201	4200	SSW
Bordu	4.24	4.89	0.33	148	4300	NW
Golubin	4.95	4.64	0.30	154	3850	NNW
Grigoriev	4.53	7.55	0.39	114	4300	mainly SSW
Kara-Batkak	3.46	2.27	0.10	113	3950	N
Sary-Tor	3.83	2.69	0.15	159	4350	NW
Batysh-Sook	2.10	0.80	0.055	185	4200	N

Batysh-Sook glacier

This thesis will investigate the evolution of the Batysh-Sook (Suek) glacier in Kyrgyzstan. In previous research, this glacier was also called 'Suyok', 'Syek Zapadny' or 'Glacier No. 419'. In this thesis, the glacier will be referred to as 'Batysh-Sook'. Its glacier ID at the World Glacial Monitoring Service is G077749E41787N. The RGI ID is RGI2000-v7.0-G-13-26329. The glacier's altitude lies between 3,950 and 4,450 metres above sea level. The glacier covered an area of 1.2 km² in 2005 (Hoelzle et al., 2017a). In 2010, its volume was estimated at 0.033 km³ $\pm$ 0.0021 km³ (Hagg et al., 2013a).

The glacier drains towards the Naryn river, which is an important tributary of the Syr Darya river, which is together with the Amu Darya the most important river draining into the Aral Sea. In 2007, 507 glaciers were observed in the basin, covering an area of 471 km² (RGI 7.0 Consortium, 2023). Batysh-Sook is located in the Sook range, where Hagg et al. (2013b) in 2007 counted 44 glaciers, corresponding to a total area of 30.9 km² and a volume of 1.2 km³. Using data from the Tien Shan meteorological station (3659 m) between 1997 and 2014, the mean annual air temperature was -6.0 °C, with monthly averages ranging from -21.7 °C in January to 4.4 °C in July. To estimate the conditions at the glacier surface (3950–4450 m), a standard lapse rate of 0.0065 °C/m was applied, resulting in a corrected mean annual temperature between approximately -7.9 °C and -11.1 °C. The mean annual precipitation in the same period was 360 mm, of which two-thirds fell during summer showers (Kutuzov and Shahgedanova, 2009).

Batysh-Sook Glacier was previously included in glaciological monitoring programmes by the USSR. For this purpose, mass balance measurements were made in 1970–1971, 1983–1984 and 1988–1991. When new measurement projects began in Kyrgyzstan in 2010, Batysh-Sook was selected to be one of the monitored glaciers (Kenzhebaev et al., 2017). Annual snow

pit and stake measurements have been conducted since 2010, which are used to calibrate a mass balance model (Hoelzle et al., 2017b). The value of the mean mass balance fluctuates between -0.4 and -0.8 m w.e.yr⁻¹ (Hagg et al., 2013a). Kenzhebaev et al. (2017) attempted to reconstruct the mass balances between 2004 and 2010 (when there were no measurements) using three models. That research resulted in an average mass balance of -0.39 m w.e.yr⁻¹. Between 1956 and 2007, the glacier lost 12.1% of its volume and 19.8% of its area (Hagg et al., 2013a).

1.2 Climate change

The annual temperatures in the Tien Shan have been increasing since the 1930s (Aizen et al., 1997). The rate of warming is accelerating (Bolch et al., 2022). For the precipitation, the trend is less distinct. Although a small increase in precipitation can be found, it is not clear whether this is caused by climate change or that it can be attributed to cyclic changes. Overall, the interannual differences in precipitation are large (Aizen et al., 1997). At higher altitudes, it is suggested that environmental temperatures are rising more rapidly as a consequence of global warming (Mountain Research Initiative EDW Working Group, 2015). The decrease in albedo as a result of the melt of glaciers is at least part of the explanation. Temperatures in the Tien Shan are projected to continue increasing significantly. Compared with the mean between 1960 and 1990, temperatures will rise by $+0.7$ °C (SSP1-1.9) to $+5.7$ °C (SSP5-8.5) by the end of the century. For precipitation, the models do not predict a significant change (Van Tricht and Huybrechts, 2023).

As in many other warming regions in the world, mountain glaciers in Central-Asia and the Tien Shan are rapidly decreasing. According to Farinotti et al. (2015), the total glacier mass loss in the Tien Shan between 1961 and 2012 is estimated to 27 ± 15 %, while the decrease of area is 18 ± 6 %. Dating end moraines indicates a lateral retreat of 989 ± 540 m since the Little Ice Age, while the elevation of the fronts rose 151 ± 105 m for the whole Tien Shan. However, some differences exist for different regions. For the Inner Tien Shan, where Batysh-Sook is situated, the average retreat of the front is 866 m and its elevation rise is 140 m (Solomina et al., 2004). For the Big Naryn basin, 23.4 % of the area is lost since the middle of the 20th century. According to Hagg et al. (2013b), this corresponds to a total volume loss of 6.6 to 8.4 km³, representing 20 % of the total volume. Even under a very ambitious scenario of a global temperature rise of 1.5 °C, models project that 36 ± 7 % of all ice masses in Central Asia will be lost by the end of the century (Kraaijenbrink et al., 2017).

1.3 Social relevance

Use of meltwater

The runoff produced by glaciers serves the downstream areas for many purposes. Irrigation for agriculture might be the most important one in the Tien Shan. Bolch et al. (2022), for example, estimated that the irrigated area of Aksu river basin in the central Tien Shan mountain range has increased from 4074 km² in 1994 to 8232 km² in 2020. In the same period, the annual water demand per area has only increased from 623 965 m³/km² to 696 682 m³/km². This intensification of agriculture and the expansion of cultivated land is mainly linked to the

cotton cultivation. The runoff due to meltwater has a seasonal pattern. During the summer months, when temperatures are positive, the precipitation is highest in the mountains (figure 1.2). For the agriculture, this has the advantage of compensating for dry years or seasons in the downstream area (figure 1.2) by providing a relatively stable amount of runoff (Huss and Hock, 2018).

In a warming climate with unchanged precipitation patterns, rising temperatures cause glaciers to melt more rapidly, leading to an initial increase in runoff. This increase continues until a maximum is reached, known as peak runoff. Beyond this point, as the glacier shrinks and its ice reserves diminish, the annual runoff begins to decline. Eventually, once the glacier has disappeared entirely, the annual runoff returns to its pre-glacier level, which only reflects precipitation input.

However, the timing of runoff changes significantly. With less glacier meltwater available during the melt season, runoff becomes lower during this critical period. This seasonal shift is particularly concerning because the melt season in glacier-fed basins often coincides with the dry season downstream. As a result, reduced meltwater availability during the growing season could exacerbate drought stress and increase the risk of failed harvests (Huss and Hock, 2018). The term *watertowers* is used to express the glaciers' capacity to redistribute the precipitation over the seasons (Immerzeel et al., 2020).

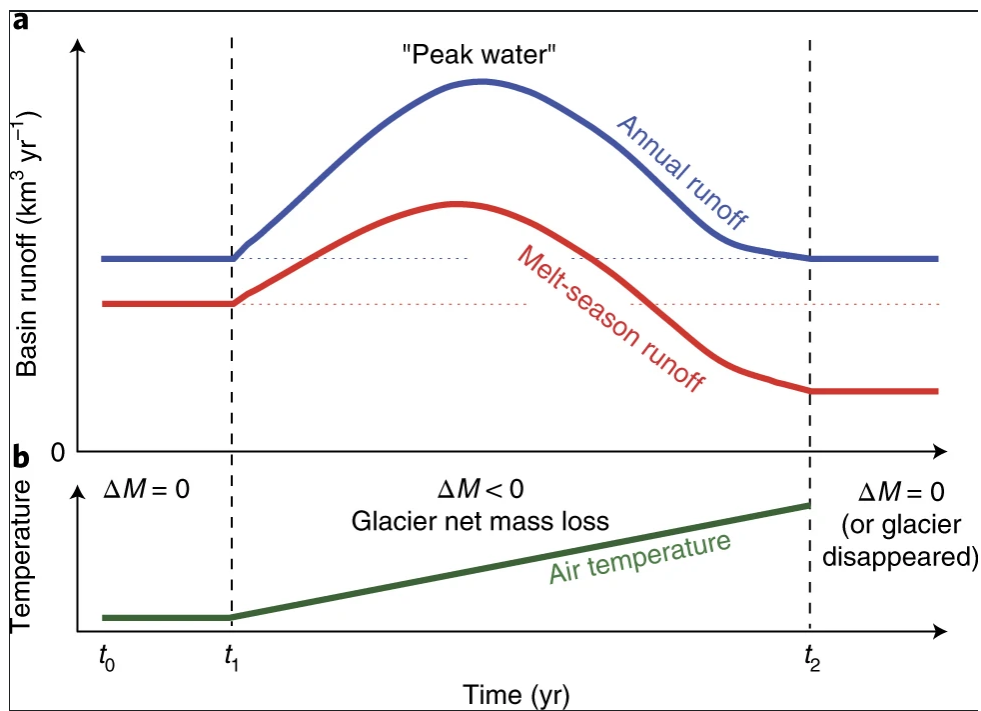


Figure 1.3: Schematic illustration of the changes in runoff from a glacierized basin in response to continuous atmospheric warming (Huss and Hock, 2018)

Agriculture is not the only sector that would be affected by this change in hydrological regime. Also for industrial and domestic use (and human consumption) the melting of the glaciers could have severe consequences. It is estimated that approximately 800 million people depend on the meltwater of the high mountain glaciers in Central Asia (Pritchard, 2019).

The total amount of glacier runoff plays a crucial role for downstream communities. Studying individual glaciers not only improves our understanding of glacier dynamics and their response to climate change, but also helps to constrain regional estimates of runoff and glacier behaviour. Detailed case studies can thus support the refinement of regional models and improve projections of future water availability.

Natural hazards

The enhanced melting of glaciers (as long as the peak discharge is not reached) or the more concentrated runoff during summer (when peak discharge passed) entails hazards for the people living in the proximity of these glaciers.

A major concern is that of an increasing number and a growing intensity of Glacial Lake Outburst Floods (GLOF). Glacial lakes, typically formed at the glacier side of an end moraine, are fed with meltwater coming from the glacier. Before the peak runoff, more water reaches the lake. This can cause more dam breaches, leading to rock avalanches, debris and mud flows and floods. Those events can cause damage to infrastructure, destroy agricultural yields, disturb fauna and flora, lead to lives losses and it has economic consequences for individuals and for the society as a whole (Emmer et al., 2022).

1.4 Instructed Glacier Model

Glacier evolution models are essential tools for simulating the temporal behaviour of glaciers, particularly under changing climatic conditions. At regional to global scales, large-scale models like OGGM, GloGEM(flow) or PyGEM are used to project glacier change, but they necessarily rely on various simplifications to remain computationally feasible. To assess the validity of these simplifications and to gain a deeper understanding of glacier processes, detailed ice-dynamical glacier evolution studies (IDGES) are carried out at the scale of individual glaciers. These models integrate surface mass balance calculations with physically based ice flow modelling, allowing for improved representation of processes such as debris cover, frontal ablation, and subglacial dynamics. Over recent decades, glacier modelling has evolved from simple one-dimensional flowline models to advanced three-dimensional simulations that account for complex stress regimes and use energy balance-based surface mass balance models (Zekollari et al., 2022). Such detailed studies are crucial not only for scientific understanding but also for applications such as hazard forecasting and water resource planning.

As available datasets grow, physical representations become more detailed, and required resolutions increase, computational demands have risen sharply. To address this, Jouvét (2023) developed the IGM, which combines physics-based ice flow equations with a deep learning emulator. Implemented as a Convolutional Neural Network (CNN), the emulator is trained to reproduce the output of a higher-order ice dynamics solver. This hybrid approach greatly accelerates simulations by exploiting the parallel processing capabilities of Graphical Processing Units (GPU)s, enabling high-resolution and large-domain modelling at a fraction of the computational cost.

IGM is modular in design, with two functionalities central to this thesis:

- Inverse modelling: interpolating sparse ice thickness observations into a continuous, physically consistent glacier geometry.

- Forward modelling: simulating glacier evolution over time in response to climate forcing.

In this thesis, IGM is applied to the Batysh-Sook Glacier in the Tien Shan for both inverse and forward modelling. Its performance is compared to classical spatial interpolation methods and to a higher-order glacier model without an emulator, enabling a critical evaluation of the benefits and limitations of this approach.

1.5 Research questions

This thesis investigates a range of research questions aimed at understanding the dynamics of the Batysh-Sook glacier. The research will address both practical, real-world concerns and methodological questions that arise from the need for accurate modelling and prediction of glacier behaviour. The research questions can be broadly divided into two categories: questions related to the evolution of the Batysh-Sook glacier and questions concerning the performance of the modelling techniques used to simulate this evolution.

The first set of questions is focused on understanding the current state and future evolution of the Batysh-Sook glacier under different climate scenarios. These questions aim to fill in the gaps in knowledge about the glacier's behaviour and the factors driving its changes:

- **What is the current ice thickness distribution of the Batysh-Sook glacier?**
This question seeks to establish a baseline for understanding the glacier's current state, an essential first step for any predictive modelling.
- **How will the Batysh-Sook glacier evolve under different climate scenarios?**
This question aims to predict the future behaviour of the glacier under various potential climate trajectories, taking into account warming temperatures, precipitation changes, and other relevant climatic factors.
- **When will the peak runoff of Batysh-Sook occur under different climate scenarios? How will the seasonal distribution of meltwater runoff change?**
This question relates the evolution of the glacier to the consequences for the society. Although the peak runoff of a single glacier might not be that meaningful for human applications, it is interesting to look at the behaviour of an individual glacier as well to look for patterns that might predict or explain the peak runoff on a larger, catchment-wide area.

In addition to the physical aspects of glacier evolution, this research also addresses several key methodological questions. These questions pertain to the application and evaluation of advanced modelling techniques, which are critical for simulating and forecasting glacier dynamics with accuracy. These methodological questions are important because they determine the reliability and accuracy of the models used to understand and predict glacier behaviour. By addressing these questions, the research will not only enhance the understanding of the Batysh-Sook glacier, but also contribute to the broader field of glacier modelling, improving the tools available for future research and policy-making.

- **How well does IGM perform in interpolating ice thickness?** This question explores the reliability and accuracy of IGM as a tool for estimating ice thickness distributions, an essential step in understanding glacier structure.

- **How well does IGM forward modelling perform in simulating the glacier's evolution over time?** This question seeks to evaluate the predictive power of IGM in capturing the complex physical processes driving glacier change, such as ice flow dynamics, melt rates, and accumulation processes.

Research flowchart

An overview of the research workflow is shown in figure 1.4. Each component of the process is described in more detail in the methods section.

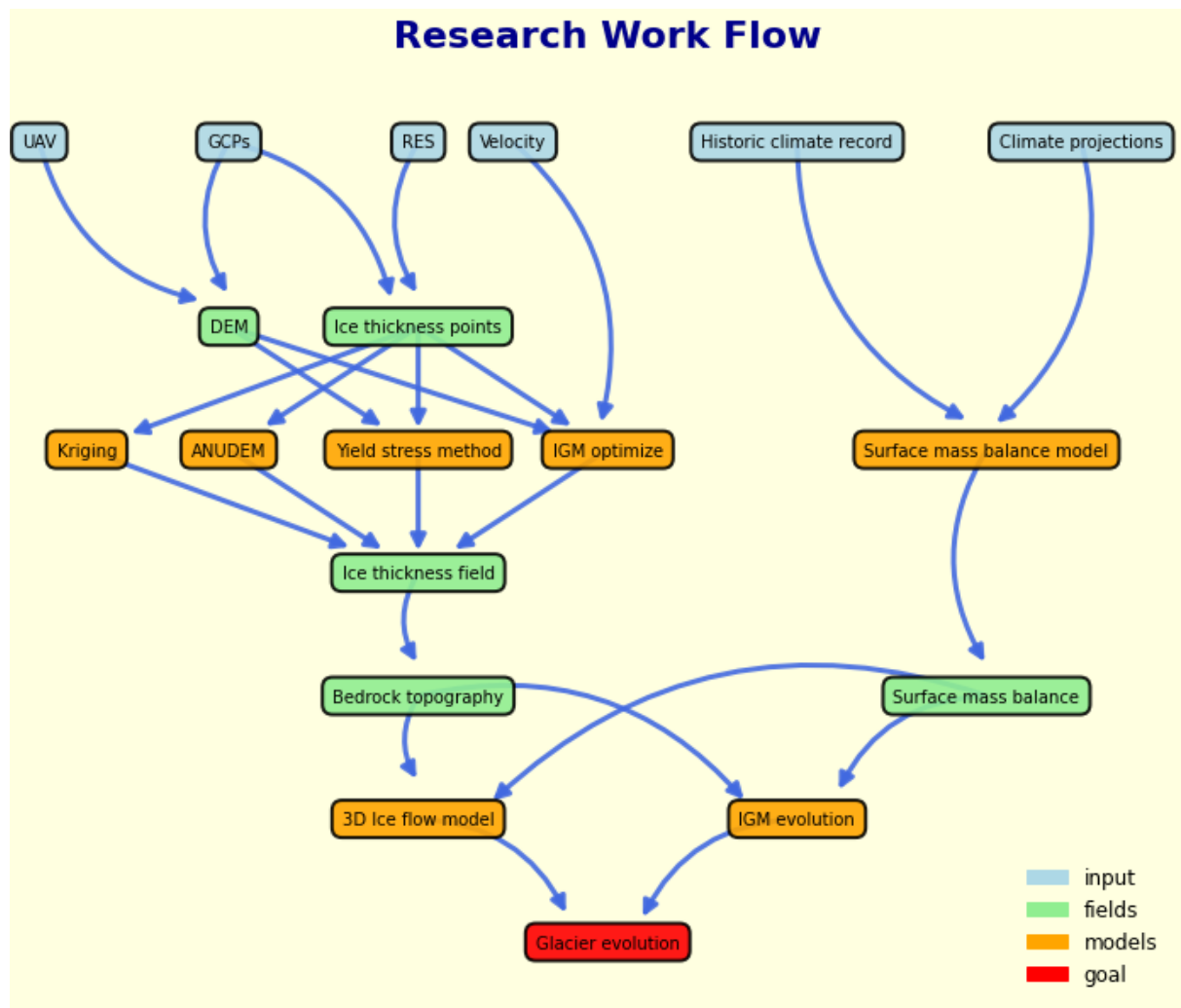


Figure 1.4: Flowchart of the research process.

Some abbreviations are introduced here for clarity: Unmanned Aerial Vehicle (UAV) refers to drone imagery, Ground Control Points (GCPs) are ground control points with known coordinates, and RES represents radar-based ice thickness measurements.

Chapter 2

Methods

Modelling a glacier's evolution requires an ice thickness field and a mass balance. In the first part of this section, the measurement and interpolation of ice thicknesses is described. In the second part, a brief introduction to the used mass balance model is given. Finally, the glacier evolution models are discussed.

2.1 Ice thickness

Ice thickness measurements

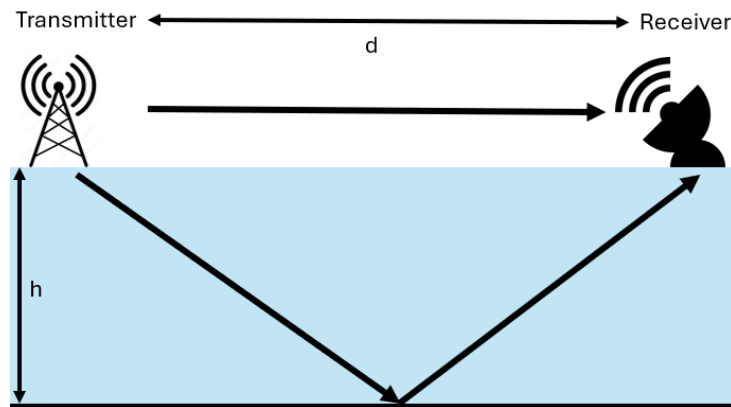


Figure 2.1: Schematic representation of a RES measurement. The transmitter sends a radiowave which is reflected by the bedrock. The reflected wave is measured by the receiver at a distance d from the transmitter. The time the wave needs to reflect represents the ice thickness h

Ice thickness can be measured using Radio Echo Sounding (RES). The concept is explained in figure 2.1. A transmitter sends radiowaves (1 - 1000 GHz) into the ice. Once the waves meet the bedrock, portion of it is reflected back to the surface. The distance between the transmitter and the receiver is d . A receiver measures the time the wave needed to be reflected. This time can be transferred into an ice thickness following equation 2.1.

$$\Delta t = \frac{2\sqrt{h^2 + \left(\frac{d}{2}\right)^2}}{v_{ice}} \quad (2.1)$$

h stand for the ice thickness. v_{ice} represents the propagation velocity of the waves through the medium, in this case ice. This velocity is calculated as in equation 2.2.

$$v_{ice} = \frac{c}{\sqrt{\frac{1}{2}\varepsilon_r((\delta\omega\varepsilon_r)^2 + 2)}} \quad (2.2)$$

In equation 2.2, c stands for the velocity of light in vacuum, ε_r is the permittivity of the ice, ω is the angular frequency ($\omega = 2\pi f$, where f is the frequency of the wave) and δ is the electrical conductivity of the ice (Plewes and Hubbard, 2001).

For all places where a RES is performed, the locations of both the transmitter and the receiver are measured by using a GPS-system. In figure 2.2, the individual measurements are depicted. The image on the background is a 10x10 m satellite image of the Sentinel-2 satellite, dating from 3/09/2024. Only the measurements that are used for the ice thickness interpolations are depicted. In total, 163 measurements were performed, of which 107 measurements could be used. All other measurements were too much scattered or unrealistically high or low. The measurements were performed with a Narod system.

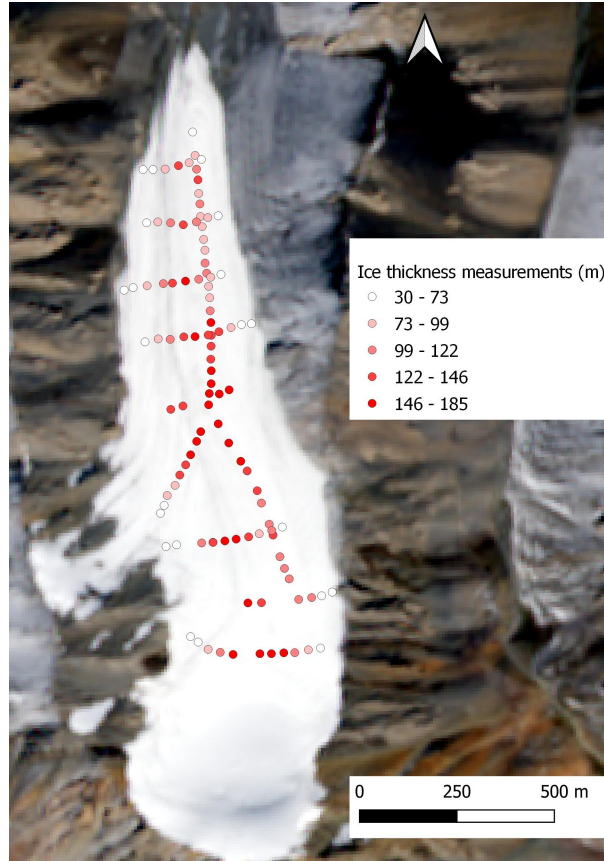


Figure 2.2: Locations and ice thickness values of RES measurements

The error in ice thickness measurements is a combination of the error on the ice thickness (the RES-measurements) itself and the horizontal error on the position measurement (the GPS-measurements) (Lapazaran et al., 2016a). The error on the RES arises from both the Radio Wave Velocity (RWV) and the travel time.

Δt must be corrected for for the normal moveout, which is the fact that receivers on different locations receive the reflected waves at different times. The correction is applied as described in equation 2.3 (Lapazaran et al., 2016a).

$$\Delta t = \sqrt{\Delta t_{measured}^2 - \left(\frac{d}{v_{ice}}\right)^2} \quad (2.3)$$

Concerning the error on the RWV, the dielectric properties (permittivity, conductivity, etc.) of the material are of large influence. Therefore, it would be better to use a different RWV depending on the material (snow/firn/ice, temperature, amount of impurities, etc.) (Glen and Paren, 1975). For practical reasons however, a constant RWV is used for all measurements. Lapazaran et al. (2016a) suggest to use a correction for the presence of firn, which usually leads to an overestimation of the ice thickness in the ablation area and an underestimation of the ice thickness in the accumulation area.

The error on the timing ϵ_t is related to the frequency of the emitted wave, as can be noted in equation 2.4.

$$\epsilon_t = 1/f \quad (2.4)$$

The emitted frequency is subject to a trade-off. Following equation 2.4, a higher frequency leads to a higher vertical resolution. On the other hand, a low frequency provides a deeper penetration. In addition to enabling deeper penetration, lower frequencies also experience less disturbance from meltwater, resulting in reduced scattering. During the fieldwork campaign, a frequency of 5 MHz was used. Another source of errors are off-nadir reflections. Errors can also arise from the identification of the bedrock reflection (Lapazaran et al., 2016a).

In addition to GPS measurement errors, other sources location error must be considered. At the boundary of the glacier, the thickness is assumed to be zero. Since the boundary is manually delineated using an (aerial) photo, an error can be made by drawing the boundary. Assuming the boundary is accurately drawn, the uncertainty on the boundary equals the resolution of the photo (Lapazaran et al., 2016b).

Ice thickness modelling

For Batysh-Sook, RES measurements were obtained during fieldwork in the summer of 2023. As previously described, performing RES measurements is labour-intensive, expensive, and in some areas, both difficult and dangerous. Moreover, the measurements require both inter- and extrapolation. Given the importance of the meltwater the glaciers produce, their total volume has got scientific attention before. It is impossible to perform RES measurements on all glaciers in an area.

Interpolation of ice thicknesses is required to derive an ice thickness field. Numerous spatial interpolation techniques have been developed for various types of geographic research, such as interpolating precipitation data, creating a DEM, and analysing hydrological data. Commonly used methods include spline interpolation, inverse distance weighting, and Delaunay triangulation, which rely solely on the spatial distribution of measured points. Since ice dynamics are often influenced by ice thickness, some equations can be inverted to estimate ice thickness through interpolation. As a result, physics-based ice thickness models have been

developed. Some of these models were tested on individual glaciers where RES measurements were performed, among others by the Ice Thickness Model Intercomparison eXperiment (IT-MIX) project (Farinotti et al., 2009, 2017; Frey et al., 2014; Fürst et al., 2017; Li et al., 2012; Linsbauer et al., 2012; Paul and Linsbauer, 2012; Van Tricht et al., 2021a). Recently, deep learning algorithms have garnered increasing interest for ice thickness interpolation. In this thesis, the 'Optimize' module of IGM (Jouvet, 2023) is examined and compared to some spatial interpolation techniques and compared to a popular physics-based ice thickness model.

In addition to the chosen interpolation technique, the spatial distribution and number of measurement points also influence the final ice thickness field. In order to assess these effects, all interpolation techniques are applied to a number of subsets of the data. The subsets are the 50% highest located RES measurements, the 50% lowest located RES measurements, 50% randomly distributed over the glacier, 90% randomly distributed over the glacier and all points in the line perpendicular on the profiles. This last subset is referred to as the 'central flowline', since these measurements follow, roughly, the central flowline from the accumulation to the ablation area. For each subset, a 'mirror subset' is created as well that contains the data points that are not included in the subset. These mirror subsets are used to study the accuracy of the interpolations. The subsets are shown in figure 2.3, with the DEM that was created by UAV measurements during the same field work when the RES measurements were performed. The aerial pictures taken by a DJI Phantom 4 RTK UAV are converted into a DEM by using stereoscopy.

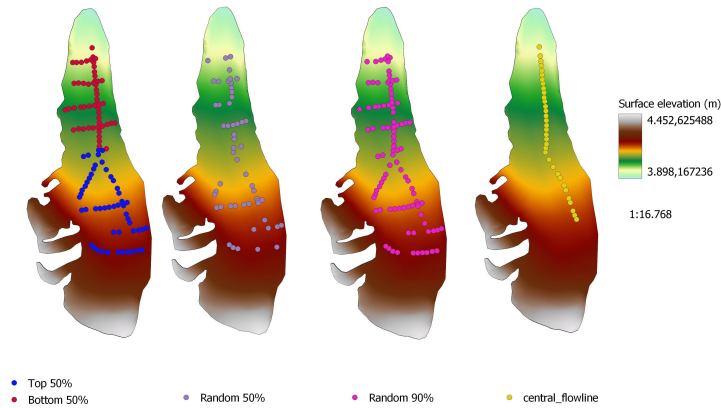


Figure 2.3: Locations of the different subsets on a DEM. In the left figure, the red dots represent the bottom 50% measurements and the blue dots represent the top 50% measurements. The second figure represents the random 50% measurements, the third figure represents the random 90% measurements and the figure on the right depicts the measurements along the central flowline.

Spatial interpolation: Kriging

Kriging is a popular spatial interpolation technique in many spatial and geostatistical analysis. The basic idea is that the value of the variable (in this case the ice thickness h) on a certain location x , is related to the values of this variable on other locations (where the value of $h(x)$ is known). The value of $h(x)$ is calculated as the weighted sum of the known values.

$$h(x) = \sum_{i=1}^N \lambda_i h(x_i) \quad (2.5)$$

In equation 2.5, λ_i represents the weight of all points i . The spatial pattern that declares the relation between the different locations is modelled by using a semivariogram. The semivariogram describes the spatial variability $\gamma(\Delta x)$ of the ice thickness h as a function of the distance Δx between locations. It indicates how values become more or less similar as the distance between locations increases. Typically, values are more similar in the vicinity, while with an increasing distance Δx , the spatial correlation usually reduces. The semivariogram is expressed as in equation 2.6 (Oliver and Webster, 2014).

$$\gamma(\Delta x) = \frac{1}{2N(\Delta x)} \sum_{i=1}^{N(\Delta x)} [h(x_i) - h(x_i + \Delta x)]^2 \quad (2.6)$$

In this equation, $N(\Delta x)$ is the number of point pairs on a distance Δx from each other. The semivariogram is used to make a covariance matrix which describes the spatial correlation between known points and the unknown location. System 2.7 must be solved to calculate the weights.

$$\begin{bmatrix} \gamma(x_1, x_1) & \gamma(x_1, x_2) & \dots & \gamma(x_1, x_n) & 1 \\ \gamma(x_2, x_1) & \gamma(x_2, x_2) & \dots & \gamma(x_2, x_n) & 1 \\ \vdots & \vdots & \ddots & \vdots & 1 \\ \gamma(x_n, x_1) & \gamma(x_n, x_2) & \dots & \gamma(x_n, x_n) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma(x_1, x_0) \\ \gamma(x_2, x_0) \\ \vdots \\ \gamma(x_n, x_0) \\ 1 \end{bmatrix} \quad (2.7)$$

In the system above, $\gamma(x_i, x_j)$ are the semivariogram values between the locations x_i and x_j . μ is a Lagrange multiplier to have this constraint: $\sum_{i=1}^n \lambda_i = 1$. The semivariogram can be modelled in different ways: spherical (using a cut-off after which the semivariance stays constant), gaussian (a more smooth transition towards the constant part), linear (without flattening of the curve) or exponential (with a quick decrease in correlation at small distances).

Kriging can be applied in different ways depending on the assumptions about the spatial process. Two commonly used methods are ordinary kriging and universal kriging. Ordinary kriging assumes that the spatial process has an unknown but constant mean across the entire study area. This means that no large-scale trends are present in the data, and the variability is captured entirely by the semivariogram. In contrast, universal kriging is designed for situations where a spatial trend exists in the data. It assumes that the mean is not constant but varies smoothly across space according to a predefined model. This trend is typically expressed as a linear combination of basic functions.

$$\mu(x) = \sum_{k=1}^p \beta_k f_k(x) \quad (2.8)$$

In equation 2.8, β_k are the coefficients to be estimated, and $f_k(x)$ are the basic functions describing the trend. Universal kriging separates the observed variability into a deterministic trend and stochastic residuals. The semivariogram is then used to model the spatial correlation of these residuals (Oliver and Webster, 2014).

In this thesis, the kriging interpolation is applied using the 'Kriging' function of the Arcpy-package in python. It was opted for a universal kriging, since a certain trend is suspected. The functions $f_k(s)$ has to be chosen, since universal kriging requires them. Glaciers have often a more or less parabolic shape. Therefore, the drift term is chosen to be quadratic ('QUADRATICDRIFT'). The semivariogram values $\gamma(x_i, x_j)$ are the range ('majorRange') and the sill ('partialSill'). The range determines the distance h above which the correlation between points becomes practically zero. At that distance, $\gamma(h)$ reaches the sill, or the maximal value at larger distances. Another parameter is called the nugget ('nugget'), which captures the small-scale variability like measuring errors. Finally, the lag ('lagSize') is the value on the x-axis of the semivariogram. It is, in other words, the distance between the different measurement points (ESRI, 2023b). Figure 2.4 is a clear, though theoretical, illustration of a variogram and its parameters.

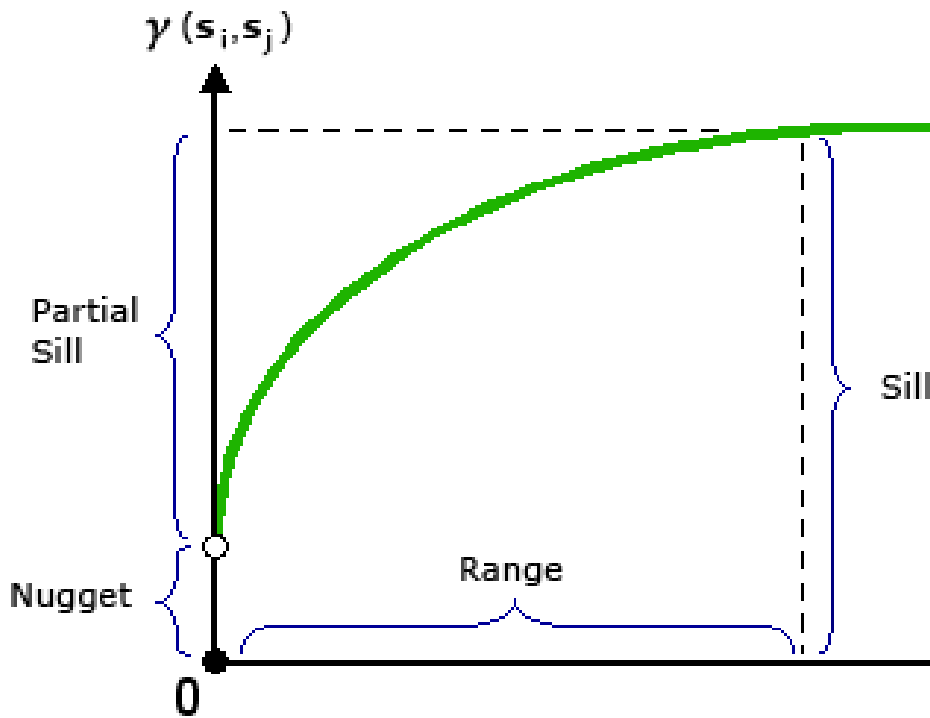


Figure 2.4: Theoretical semivariogram with the spatial variability γ on the vertical axis and the lag on the horizontal axis (ESRI, 2023b). The indicated parameters are the nugget, the range and the sill.

Table 2.1: Description of kriging parameters used.

Name	Meaning
nugget	The discontinuity at the origin
distance	Distance within which input sample points are used for interpolation
majorrange	Distance beyond which there is little or no correlation
minnumberofpoints	Minimum number of points used for interpolation
partialsill	The difference between the nugget and the sill
lagsize	Distance-unit (used to group semivariance computations)

Spatial interpolation: ANUDEM

An interpolation technique that is often used to create DEMs from point observations is the ANUDEM algorithm developed by Hutchinson (1988). The main idea behind ANUDEM is to create DEMs that have realistic drainage properties. Elevation measures can be supplied to the model as points, lines and contours. On a regular grid, ANUDEM interpolates the elevation so that it minimizes roughness elements, but drainage features like rivers, ridges, cliffs and valleys are preserved. ANUDEM ensures the drainage structure of the terrain to be preserved. Unrealistic sinks, where water could accumulate, are removed, although lake boundaries are detected. The drainage enforcement algorithm connects the points between the ridges ("saddle points") to lower sinks, which ensures a logical flow of water. ANUDEM accepts various types of input data. Besides point locations with elevation information, also points that represent areas where water could accumulate, lines representing streams or rivers, data representing the boundaries of regions, lines representing elevation changes on maps, data defining the boundaries of lakes and data specifying areas of cliffs can be provided to the model. The algorithm reads the input data and processes it at various grid resolutions, starting from a coarse resolution and progressively refining it. That allows ANUDEM to handle both large and detailed datasets. For each grid resolution, the program assigns input data points to grid cells and uses the Gauss-Seidel iteration method to adjust the grid values, ensuring they adhere to drainage constraints, including ridge and stream representation. The program detects remaining sinks after the grid fitting process and optionally outputs them for further analysis (Hutchinson, 2011).

In this thesis, the implementation of this algorithm in the Python package Arcpy called TopoToRaster was used. The outline of the glacier is given as boundary. As with the kriging interpolation, the datasets are supplemented with points on the boundary that represent an ice thickness of five meters, and outside the glacier outline of zero meters. Four optional parameters are provided to the model. Many more parameters can be included, but only these four parameters are optimized and deviate from the default settings. The *discrete error factor* adjusts the amount of smoothing. A smaller value results in a less smooth interpolation. The *roughness penalty* is described as 'the integrated squared second derivative as a measure of roughness'. The *vertical standard error* parameter simulates measurement uncertainty by adding a small random noise to the ice thickness values, mimicking real-world observational error. In contrast, the *tolerance 1* parameter defines how closely the interpolated surface is allowed to follow the input data. A smaller value enforces stricter adherence to the input measurements, effectively increasing the influence of their spatial density and presumed accuracy. All other parameters are left on their default value (ESRI, 2023a).

Table 2.2: Description of ANUDEM parameters used.

Name	Meaning
<code>discrete_error_factor</code>	Adjusts smoothing
<code>roughness_penalty</code>	Roughness measure
<code>tolerance_1</code>	Accuracy of the elevation points
<code>vertical_standard_error</code>	Random error in the z -values of the input data

Inverse modelling interpolation: Yield stress method

In classical (forward) modelling, inputs are given to the model and this computes the outputs. Inverse modelling does the inverse: it infers the inputs from given output fields. Inverse modelling often involves iterative techniques to minimize the error between observed outputs and model-predicted outputs (Neto and da Silva Neto, 2012).

Glen's flow law (equation 2.9) describes the creep relation: how much the ice deforms when a certain amount of forces is applied. The deformation of the ice is described by the strain ε_{ij} . The rate at which the strain changes over time is called the strain rate $\dot{\varepsilon}$ (yr^{-1}). In case there is no water at the bottom of the glacier and the bedrock is not deforming under the pressure, deformation of the ice is the main mechanism in which a glacier flows. The forces that act on the ice are described as stresses (τ).

$$\dot{\varepsilon}_e = A(T)\tau_e^n \quad (2.9)$$

In equation 2.9, the effective strain rate $\dot{\varepsilon}_e$ and the effective stress τ_e are linked. These are scalars representing the magnitude of respectively the strain rate $\dot{\varepsilon}$ and the stress τ . Creep exponent n indicates the prevailing creep relation. For glacier ice, this exponent has been identified to vary between 1.5 and 4.2, and is usually taken to be 3. This indicates a non-linear viscous behaviour. If n were equal to one, the ice would deform linearly with the applied stress (Cuffey and Paterson, 2010). This relation is visualised in figure 2.5.

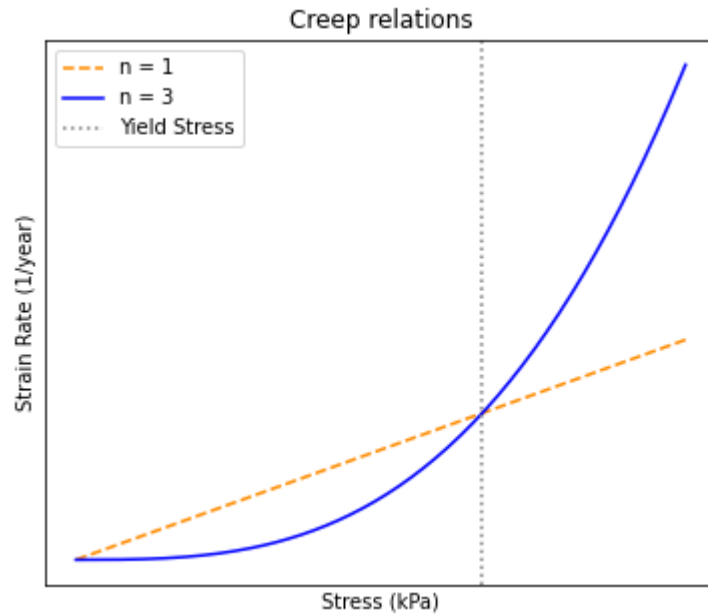


Figure 2.5: Theoretical diagram of the strain rate with the stress for different creep relations ($n = 1$ and $n = 3$). The yield stress is indicated with a vertical dashed line.

Graph 2.5 indicates the yield stress. Often, a perfect plasticity assumption is used, where the ice is assumed not to deform until the yield stress is reached. Once that threshold is reached, the material deforms or flows, but the stress does not increase further (Cuffey and Paterson, 2010).

By using the concept of the yield stress, the interpolation of the ice thickness observations can be done more informed than in the spatial interpolations. The basal shear stress τ_b , which is the component of the stress at the interface between the bedrock and the glacier ice parallel to the direction of movement, is assumed to be equal to the yield stress (because the glacier is moving). The glacier moves under the stress of its own weight. The basal shear stress τ_b can be described by using equation 2.10.

$$\tau_b = \rho g h \sin \phi \quad (2.10)$$

The basal shear stress τ_b is determined by the ice thickness h and the slope of the surface ϕ . Since the surface slope is relatively low, the sine is close to the tangent. The tangent of the surface slope corresponds to the magnitude in surface gradient ∇s (dimensionless). g stands for the acceleration of gravity and ρ is the density of the ice. Inverting equation 2.10 and filling in the surface gradient, the ice thickness can be calculated as in equation 2.11.

$$h = \frac{\tau_0}{\rho g |\nabla s|} \quad (2.11)$$

In equation 2.11, τ_0 is the yield stress. This equation can be used to interpolate the ice thickness observations. In this thesis, the method was based on the described method by Zekollari et al. (2013).

In the first step, a 25 m-resolution grid is filled with the observations by the RES. Just as for the spatial interpolation methods, the last cells inside the glacier outline are assigned 5 meters. This approach uses a central flowline, since in most cases glacier ice is thickest along the centreline of glacier flow. To derive the central flowlines, the distance to the nearest glacier margin $l_{i,j}$ is calculated for each gridcell. This creates a two dimensional distance field, which is high in the middle of the glacier and zero at the margins. On this distance field, a discrete Laplacian operator $I_{i,j}$ is applied. Equation 2.12 describes the operator, which is actually a second derivative in two dimensions.

$$I_{i,j} = \frac{1}{4} (l_{i+1,j} + l_{i-1,j} + l_{i,j+1} + l_{i,j-1}) - l_{i,j} \quad (2.12)$$

At the peaks of $l_{i,j}$, the discrete Laplacian operator is strongly negative, which indicates concave-down regions in the distance field. These areas are the central parts of the glacier and assumed to be located on the central flowline. Only the cells with a discrete Laplacian less than -5 m are selected. Some of those central flowline points have direct RES measurements. These points are used to calculate the yield stress τ_0 . These yield stress values are interpolated over the entire central flowline. Equation 2.11 is used to derive the ice thicknesses on the central flowline.

Ice thicknesses at the remainder of the glacier were interpolated from the central flowlines to the glacier margins using an inverse quadratic distance weighting method, applied at regular cross-sectional intervals. The inverse distance interpolation step was omitted in areas where direct field measurements existed to preserve empirical values. Gaps remaining after this process were filled using bilinear interpolation. Finally, a linear smoothing filter is applied with a 100 m moving window to reduce small-scale noise in the reconstructed thickness field.

Inverse modelling interpolation: IGM optimize module

The Instructed Glacier Model (IGM) developed by Jouvét et al. (2024) is used later on in this thesis to simulate the glaciers evolution. However, it also contains a pre-processing module 'optimize' which performs the inverse modelling of the glacier fields. The description of the model itself (which in this part is used in the inverse version) can be found further in this method section. Here, only the optimisation algorithm that goes along with the inverse modelling is described.

The optimize module from IGM adjusts the initial values of glacier fields to make them consistent with observations and physical laws. This is done by solving an optimisation problem that minimizes a cost function. The cost function quantifies the difference between modelled and observed data and adds regularization to ensure the physical consistency of the model. The user chooses which variables are optimised (e.g. ice thickness, surface elevation, etc.) and which cost components are included. This cost function is defined in equation 2.13 (Jouvét, 2023).

$$J(h, c, A, s) = C_u + C_h + C_s + C_d + R_h + R_b + R_A + P_h \quad (2.13)$$

In this cost function J , there are observation terms (C) that look for the misfit between the modelled and the observed field, regularisation terms (R) that improve smoothness and physical plausibility and a penalty term (P) that prevents negative ice thicknesses. There are observation terms for the surface velocity u , ice thickness profiles h , surface elevation s and the flux divergence d . The equations used for these observation terms are given below (Jouvét, 2023).

$$\begin{cases} C_u = \int_{\Omega} \frac{1}{2\sigma_u^2} \|u_{s,obs} - u_s\|^2 \\ C_h = \int_{\Omega} \frac{1}{2\sigma_h^2} |h_{obs} - h|^2 \\ C_s = \int_{\Omega} \frac{1}{2\sigma_s^2} \|s - s_{obs}\|^2 \\ C_d = \int_{\Omega} \frac{1}{2\sigma_d^2} |\nabla \cdot (h\bar{\mathbf{u}}) - d|^2 \end{cases} \quad (2.14)$$

In these expressions, Ω stands for the glacier domain. The regularisation terms enforce anisotropic smoothness and convexity of ice thickness and the bedrock (R_b), a smooth sliding coefficient (R_c) and a smooth Arrhenius factor (R_A). For a more detailed discussion of the cost terms, we refer to Jouvét (2023).

The optimize module minimizes J with a gradient-based algorithm. Therefore, the open source software library TensorFlow is used. This library was developed for deep learning approaches. TensorFlow computes the gradients of the cost function with automatic differentiation, which is both accurate and efficient (TensorFlow, 2024).

The optimisation process stops when a predetermined criterion is reached, for example a threshold for change in J . The main advantage of this approach is its efficiency. Thanks to TensorFlow's automatic differentiation and GPU parallelization, the optimisation is performed quickly. In addition, users can adjust variables, cost components, and regularization parameters as needed. Due to physical constraints and regularization, unrealistic solutions are avoided. Since the ice flow is modelled with a deep learning algorithm, the accuracy of the model depends on the quality of the training. The quality and completeness of observations determine the success of the optimisation. Areas without data can result in larger uncertainties. In

addition, setting the right parameters for regularization and data integration is sensitive to biases (Jouvet, 2023).

The user is allowed to remove cost variables from the cost function 2.47. IGM can calculate the variable fields (surface elevation, ice thickness, velocity) starting from zero and only base itself on the observations. On the other side, IGM also allows to offer an initial field for the variables and then adjust this field based on the observations. The surface elevation s is measured using the UAV and provides both an initial field as the observations used in the cost function. The ice thickness h observations are used in the cost function. No initial ice thickness field is used. No velocity measurements were performed, but Millan et al. (2022) developed a remote sensing method to calculate the ice thickness and ice flow velocity of glaciers on a global scale at a 50 by 50 m resolution. In this thesis, IGM inverse modelling has been used with and without the Milan-ice flow velocity field as initial field. In figure 2.6, the result of the Milan method for Batysh-Sook is depicted.

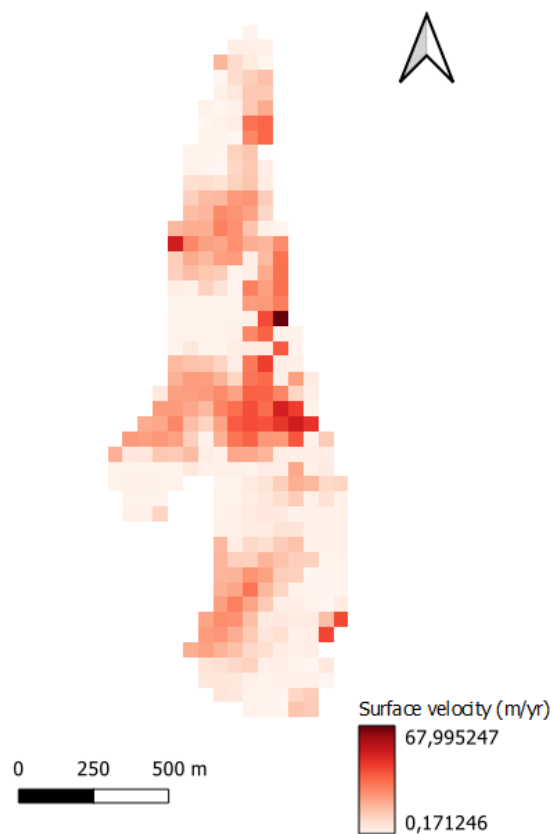


Figure 2.6: Surface velocity based on Millan et al. (2022).

The velocity field has a remarkable pattern that not meets the expectations. The highest values are predicted to occur at the side of the glacier, which is unrealistic. Besides that, the field is extremely scattered and has very high values. The field cannot be used in the cost function, because of its unrealistic behaviour. The field will be used as initial field in one of the experiments, while keeping in mind its poor result.

In order to fit the model for the glacier, some parameters have to be tuned. The first

parameters to vary are the confidence levels for the input variables. These parameters quantify how well the interpolated fields have to match the given input. These confidence levels are σ_u (surface velocity), σ_h (ice thickness), σ_s (surface elevation) and σ_d (divergence of the flux). Since the divergence of the flux and the velocity are not included in the cost function, σ_d and σ_u are ignored.

In a second phase, the regularization parameters can be changed. These parameters, 4 in total, are part of the expressions of the cost terms in equation 2.13. Increasing α_h and α_A control, improves the smoothness of the ice thickness and the flow fields. β determines how the smoothing has to happen, isotropic or anisotropic. A lower β causes a more anisotropic smoothing in the flow direction of the glacier. Parameter γ adds more convexity when it is lower.

Eventually, also some parameters for the numerical approximation can be adapted. The number of iterations and the iteration step size can be changed. However, during this research, these parameters are kept constant at respectively 1000 and 0.001. This choice was made, since the fields reach their final state already within a relatively low number of iterations.

Table 2.3: Description of IGM parameters used in the optimization process.

Symbol	Name	Meaning	Optimized?
σ_u	opti_velsurfobs_std	Confidence level: surface velocity observations	Yes
σ_h	opti_thkobs_std	Confidence level: ice thickness observations	Yes
σ_s	opti_usurfobs_std	Confidence level: surface elevation observations	No
σ_d	opti_divfluxobs_std	Confidence level: flux divergence observations	No
R_h	opti_regu_param_thk	Regularization term: ice thickness	Yes
R_c	opti_regu_param_slidingco	Regularization term: sliding coefficient	No
R_A	opti_smooth_anisotropy_factor	Regularization term: smoothing isotropic or anisotropic	Yes
–	opti_convexity_weight	Regularization term: adding convexity	Yes
–	opti_nbitmax	Number of iterations	No

The inverse modelling approach of IGM has been used before to model or interpolate the ice thickness and bedrock. Jouvét (2023) tested the model on ten large glaciers in the Swiss Alps. They found that the proposed CNN emulator-based inverse modelling approach offers significant improvements in terms of efficiency, robustness, and accuracy compared to traditional methods, and it closely matches observed measurements. Not many others have used IGM inverse module to calculate ice thicknesses. Although van Pelt and Frank (2025) and Frank and van Pelt (2024) use IGM, the inverse modelling in those studies did not make use of the optimize module.

Accuracy measurement

In order to compare the different methods and subsets, the RMSE is used. This metric compares the measured value of a variable with the modelled or interpolated value of that variable. The RMSE is defined as in equation 2.15.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{h}_i - h_i)^2}{N}} \quad (2.15)$$

In this equation h is the variable for which the interpolation has been done (in this case the ice thickness). h_i is the interpolated value of h and $\hat{h}_i$ is the measured value of h . The units of the RMSE are the units of the variable, so meters in the case of ice thicknesses.

If the residuals do not originate from a normal distribution, this could point to problems with the implementation of the model, heteroscedasticity or large problems in the measurements. Therefore, a statistical test to check the normality of the residuals is performed. The choice fell on the KS. Compared with the Shapiro Wilk test (SW), the KS test is less suitable for small datasets, but KS is more sensitive for deviations in the middle of the distribution while SW depends more on deviations in the tails of the distribution (Burt et al., 2009). Since the RES measurements are highly sensitive for basal water, the presence of outliers is not unthinkable. Therefore, KS is preferred over SW. The KS test statistic D is calculated as described in equation 2.16.

$$D = \max_i |S(h_i) - F(h_i)| \quad (2.16)$$

The KS searches for the maximal discrepancy between the cumulative frequency of the residuals $S(h_i)$ and the cumulative frequency of a normal distribution $F(h_i)$ (Burt et al., 2009). A large D casts doubt on the null hypothesis, which is that the residuals come from a normal distribution. An uncertainty level of 5 % is used.

The residuals might show a spatial dependence. A spatial dependence could indicate there are regions where the interpolation is significantly wrong. Moran's I was calculated to check for spatial autocorrelation among the residuals. The formula is given in equation 2.17.

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (h_i - \bar{h}) (h_j - \bar{h})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \sum_{i=1}^n (h_i - \bar{h})^2} \quad (2.17)$$

I is the Moran's I statistic, n is the number of observations, h_i is the observed ice thickness on location i and h_j on location j . $\bar{h}$ is the overall average of h . w_{ij} is a matrix consisting of spatial weights. The used metric is inverse distance weighing, where the weight is equal to the inverse of its Euclidean distance. A small number (e^{-10}) is added to each weight to ensure numerical stability. The weights at the diagonal are replaced by zeros so that self-correlation is avoided. The more the Moran's I approaches 1, the stronger the positive spatial autocorrelation, and the closer to -1, the stronger the negative spatial autocorrelation (Burt et al., 2009).

Optimisation algorithm

All interpolation methods, except for the yield stress method, have to be calibrated to the data. That implies adapting the model parameters in order to fit the specific data as well as possible. In this calibration, usually a trade off exists between fitness (the interpolation being as close as possible to the input data) and the smoothness of the interpolated field. The fitness is calculated via the root mean squared error (see section about accuracy measurement). The smoothness is summarised by using the Laplacian of the interpolated field. The vector Laplacian ∇^2 , not to be confused with the discrete Laplacian used for the yield stress method, is a differential operator applied to a vector field. In two dimensional Cartesian coordinates, it is defined as in equation 2.18.

$$\nabla^2 h = (\nabla h_x, \nabla h_y) \quad (2.18)$$

In equation 2.18, h_x and h_y are the components of the vector field of the ice thickness. ∇ stands for the scalar Laplacian of h , which can be calculated as in equation 2.19.

$$\nabla h = \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} \quad (2.19)$$

The second derivatives in both directions are numerically approached as defined in equation 2.20.

$$\begin{cases} \frac{\partial^2 h}{\partial x^2} [i, j] = \frac{h[i, j+1] + h[i, j-1] - 2h[i, j]}{\Delta x^2} \\ \frac{\partial^2 h}{\partial y^2} = \frac{h[i+1, j] + h[i-1, j] - 2h[i, j]}{\Delta y^2} \end{cases} \quad (2.20)$$

Based on these numerical approximations, the Laplacian is measured. The smoothness parameter is defined as in equation 2.21.

$$smoothness = \sum_{i,j} (\nabla^2 h_{ij})^2 \quad (2.21)$$

A low smoothness indicates a smoother interpolated field.

Since a trade-off must be made between smoothness and fitness, the optimisation cannot simply optimise the RMSE or the smoothness. A weighted average is used and is defined in equation 2.22. Both RMSE and laplacian are first normalized before they were used in equation 2.22.

$$\chi = RMSE + \lambda \cdot smoothness \quad (2.22)$$

In equation 2.22, λ is a parameter that determines to which extent the smoothness has to be taken into account compared with the RMSE. λ is a user defined input parameter. During the optimisation, χ is minimised.

The used optimisation algorithm is Particle Swarm Optimisation (PSO). The algorithm looks for the optimal solution by using a determined amount of combinations of parameter values ("particles"). The particles adapt their direction and velocity based on their own experience and on the experience of the other particles. This movement is described in equations 2.23.

$$\begin{cases} v_i = v_i + c_1 r_1 (P_{best,i} - x_i) + c_2 r_2 (G_{best} - x_i) \\ x_i = x_i + v_i \end{cases} \quad (2.23)$$

In equation 2.23, v_i and x_i are the velocity and the location of particle i , $P_{best,i}$ is the best performance (in our case the lowest χ) of particle i , G_{best} is the global best performance of all particles, c_1 and c_2 are learning factors and r_1 and r_2 are random numbers between 0 and 1. After sufficient iterations, most particles will end up in a situation where χ is optimal, although the risk of directing towards a local optimum is not ruled out.

Besides the weighing factor λ that determines the trade-off between fitness to the data and smoothness, the user has to provide ranges for each parameter in which it can occur. A list of all optimised parameters and their tested ranges is given in table 2.4. During the optimisations,

both factors (smoothness and accuracy) are assumed to be equally important. Therefore, λ is assumed to be 1. To avoid effects that are related to units, both the smoothness and RMSE are normalised.

Table 2.4: Ranges of parameter values within which the PSO algorithm was performed.

Interpolation	Parameter	Boundaries
Kriging	lagSize	10.0 – 100.0
	majorRange	50.0 – 500.0
	partialSill	0.01 – 5.0
	nugget	0.0 – 1.0
	distance	25.0 – 300.0
	minNumberOfPoints	0.0 – 5.0
ANUDEM	discrete_error_factor	0.1 – 6.0
	tolerance_1	5.0 – 15.0
	roughness_penalty	0.05 – 0.5
	vertical_standard_error	0.0 – 0.5
IGM	opti_velsurfobs_std	0.1 – 10.0
	opti_thkobs_std	0.1 – 10.0
	opti_regu_param_thk	1.0 – 20.0
	opti_smooth_anisotropy_factor	0.1 – 0.9
	opti_convexity_weight	0.0 – 1000.0
	iflo_init_slidingco	0.01 – 0.1

Sensitivity analysis

Sensitivity analysis is often used to determine the impact of (small) variations in input variables and parameters on a models outcome. In this thesis, the impact of the input variables will not be investigated, although the different amounts and spatial distributions of the ice thickness measurements can be seen as a kind of sensitivity analysis. The model parameters of ANUDEM, kriging and IGM will undergo a sensitivity analysis. However, no extra runs were performed specifically for this sensitivity analysis. During the optimisation process, the output of the intermediate iterations is used to access the responsiveness of the model for perturbations in the parameters.

To determine the importance of each parameter, the feature importance and the permutation importance are calculated. The feature importance is a measure that indicates the importance of each parameter for the model performance (e.g. RMSE for the accuracy or the Laplacian for the smoothness). To calculate this measure, a gradient boosting method is used. In gradient boosting algorithms, a random forest method is sequentially performed, so that each decision tree is trained in order to reduce the errors of the previous one. It is important to notice that the model is not run in the gradient boosting algorithm itself. Instead, the gradient boosting uses the output of the model runs in the optimisation loop, as well as the associated RMSE values. The permutation importance on the other hand is used to assess

the impact of a single parameter on the model performance. The permutation importance is calculated by changing the parameter values randomly and then looking at the change in performance. Feature importance tells how much a feature contributes to the model, while permutation importance provides a robust measure of the change in model performance when a feature is changed or removed.

For each parameter, a Particle Dependence Plot (PDP) is made. A PDP visualises how the model performance changes with a varying value of a certain parameter. It is determined by varying one specific parameter and taking the average of all others. The partial dependence plots show the marginal effect of a single parameter on the target metric, in this case the RMSE, while averaging out the influence of all other parameters. In other words, it isolates how changes in one parameter affect the model's performance, assuming all other parameters remain fixed. The horizontal axis shows the range of tested values for the parameter. The vertical axis represents the partial dependence, which here reflects how the RMSE changes as the parameter varies. A positive slope indicates that increasing the parameter value increases RMSE, while a negative slope suggests a decrease in RMSE and therefore a better performance. The risk of these PDPs is that non-linear effects remain undetected. In addition, also correlation between different parameters could lead to a misleading PDP.

A commonly used alternative are SHAP values. This method is based on game theory and assumes every parameter to be a player in a cooperative game. The SHAP value indicates how much a specific player contributes by calculating the marginal change of each parameter. A SHAP value around zero indicates a low contribution of that parameter, while a strong positive or negative value indicates a strong positive or negative effect, respectively, of that certain parameter. These SHAP values are often visualised in a 'summary plot', where all parameters are ranked (the most important one is depicted the highest on the vertical axis). For each parameter, a scatterplot is made against the SHAP values on the horizontal axis. The colour of each point in the scatterplot represents the value of the parameter.

2.2 Mass balance model

The specific mass balance rate, often abbreviated as specific mass balance, defined as the net gain or loss of glacier mass per unit surface area over time (typically in m w.e. yr^{-1}), is a key measure of glacier health. Assuming a constant density allows the specific mass balance to be expressed as a thickness of ice or water (Cuffey and Paterson, 2010). The mean specific mass balance is the specific mass balance per unit area, averaged over the entire glacier extent (Cogley et al., 2010). The mean specific mass balance is the result of all accumulation and ablation processes on the complete glacier during a (mass balance) year. It has a close connection to the prevailing climate. Being a good indicator of climate change, the mean specific mass balance is crucial for predicting a glacier's evolution.

Mass balance measurements

Mass balance measurements are done in several ways, ranging from digging snow pits to remote sensing estimates based on satellite images. The most popular method uses stakes that are drilled into the ice at the end of the ablation season. The SMB is calculated by comparing the change in glacier surface elevation to the fixed position of the stake's top. It

represents the difference between the solid precipitation and the melted ice. Since these stake measurements can only be performed in the ablation area, they are often combined with snow pits in the accumulation area where a hole is dug until the surface of last summer is reached. The SMB is measured after 1 year ('mass balance year'), although it is also done in other moments of the year to measure for example the winter mass balance (measured at the end of spring, when no ablation took place yet). Measuring the SMB by using stakes and pits is referred to as the 'glaciological method' (Cogley et al., 2010). The main shortcomings are the large amount of labour required, the sparse distributed points (requiring a relatively large amount of interpolation) and the absence of points in (part of) the ablation area due to the high number of crevasses and snowbridges. Recently, progress has been made by using UAV in order to determine the 'geodetic mass balance'. Van Tricht et al. (2023a) performed UAV measurements on a number of glaciers in the Tien Shan region, indicating a close agreement with the glaciological method. Hugonnet et al. (2021) developed a method to make an estimate for all glaciers based on satellite remote sensing. Mass balance data using this method are freely available.

The World Glacier Monitoring Service (WGMS) consists of a continuous list of SMB-measurements since 2010 performed by the Central Asian Institute for Applied Geosciences. Via personal contact with Mr. Ruslan Kenzhebaev, access was given to the SMB per stake. In addition, the mass balances of 1970, 1983, 1988, 1989 and 1990 are also available. Besides the annual SMB, also the change in front variation was measured by the university of Fribourg for the years 1975, 1977, 1994, 1998-2002 and 2006-2015. Finally, also the ice thickness change was measured for the years 1999, 2004, 2009, 2014 and 2019 by the university of Toulouse.

Figure 2.7 depicts the mean specific mass balance, as measured using stakes. The stake measurements are averaged over the glacier by using elevation bins of 50 meters and considering the proportion of the glacier area within each elevation bin.

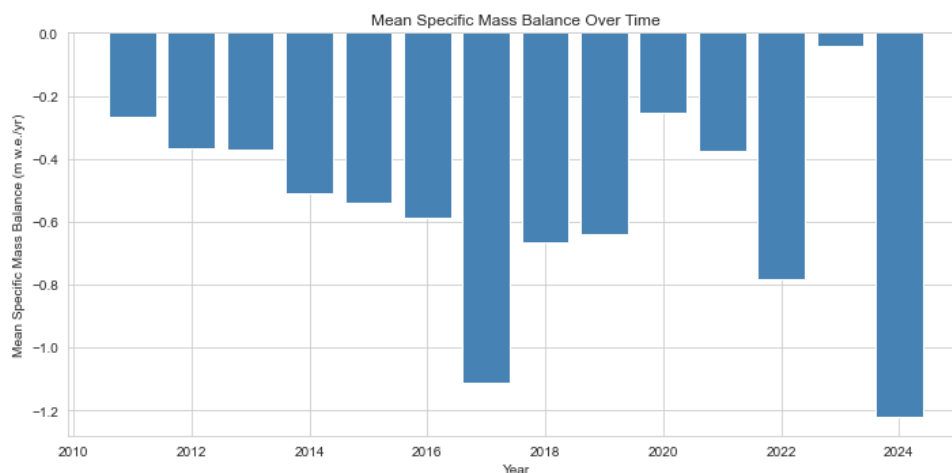


Figure 2.7: Mean specific mass balance based on stake measurements of Ruslan Kenzhebaev for the period 2011-2024

All mean specific mass balances since 2011 have been negative, ranging from $-0.043 \text{ m w.e. yr}^{-1}$ in 2022-2023 to $-1.220 \text{ m w.e. yr}^{-1}$ in 2023-2024. Figure 2.8 shows the SMB value for every elevation bin of 50 m over the different measurement years. The bold line represents the average SMB over all years.

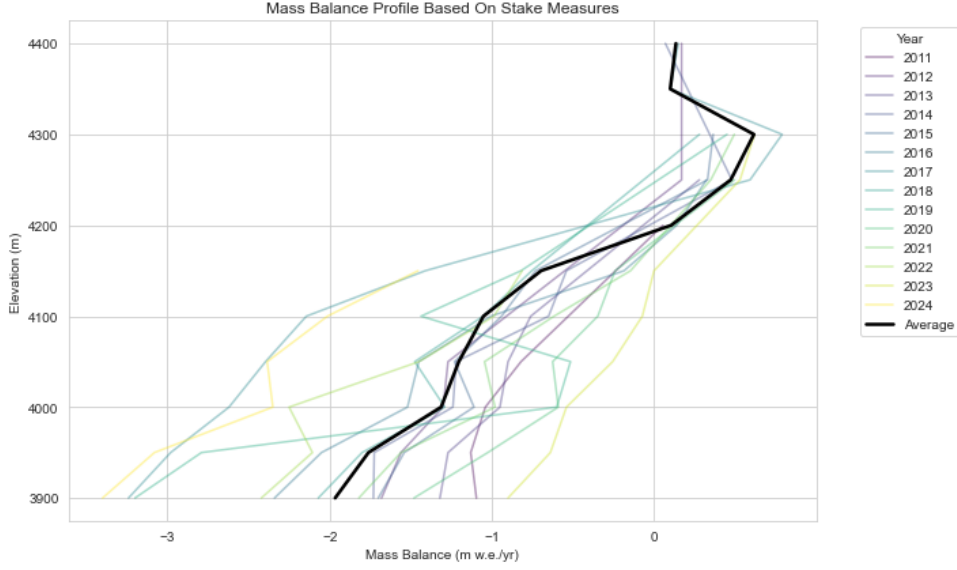


Figure 2.8: Mass balance profiles based on stake measurements for the period 2011-2024

All mass balance profiles have roughly the same shape: a fairly negative SMB in the lowest regions, which is increasing (becoming less negative) up to the Equilibrium Line Altitude (ELA), above which the mass balance becomes slightly positive. From an altitude of 4300 m onwards, the SMB decreases again. Note that measurements above 4300 m were only performed in two years. However, this observation led to the introduction of a second precipitation rate in the model, which is meant to take this decrease above 4300 m into account. Possible reasons for this decrease might be the mean elevation of the clouds base. If the cloud base is located around 4300 m, less precipitation will fall above. Another possible explanation is the removal of snow due to winds.

The geodetic mass balance as measured by Hugonnet et al. (2021) is described in table 2.5.

Table 2.5: Geodetic mass balance of the glacier by Hugonnet et al. (2021)

Period	2000–2010	2010–2020	2000–2020
Mass balance (m w.e.yr ⁻¹)	−0.867	−0.615	−0.741

Mass balance model

The glacier evolution models, however, do not use mass balance observations. Therefore, the SMB is computed by using the simple mass balance model Van Tricht et al. (2021b) used. This model uses the coupling between the SMB (equation 2.24) and the energy balance (equation 2.25).

$$SMB_{t+1} = SMB_t + P_{s,t} + (M_t - R_w) \quad (2.24)$$

$$\Psi = (1 - \alpha) \cdot \varphi \cdot Q_{in} + C_0 + C_1 \cdot T_{air} \quad (2.25)$$

The mass balance SMB_{t+1} at a moment $t+1$ in time is equal to the balance at the previous timestep (SMB_t) plus the accumulation of solid precipitation P_s and the ablation M (negative value). Since part of the meltwater stays in the snow (R_w), this should be subtracted from the mass balance.

The energy balance Ψ (equation 2.25) is determined by the incoming shortwave radiation (Q_{in}), corrected for the albedo (α) and the transmissivity of the atmosphere (φ), the longwave radiation and the sensible heat flux. C_0 and C_1 are parameters that have to be calibrated. The albedo is parametrised as suggested by Oerlemans (1998):

$$\alpha = \alpha_{snow} - (\alpha_{ice} - \alpha_{snow}) * e^{-\frac{d}{d^*}} \quad (2.26)$$

In expression 2.26, d is the snow depth and d^* is the characteristic snow depth. The albedo of snow is a compound of the albedo of snow and firn, which is defined as in equation 2.27.

$$\alpha_{snow} = \alpha_{firn} - (\alpha_{frnsnow} - \alpha_{firn}) * e^{-\frac{t_{snow}}{t^*}} \quad (2.27)$$

In equation 2.27, t_{snow} is the time since the last snowfall event expressed in days. t^* stands for the time it takes for snow to develop to an albedo of firn after it has fallen.

The model has to be calibrated in order to assign a value to C_0 and C_1 . The calibration is done manually by changing the values of the two parameters and calculating the difference with the observed values. Two sources of observations are used to compare: the stake measurements and the geodetic mass balance calculated as by Hugonnet et al. (2021). During the calibration runs, the surface elevation was forced not to change. Calibration happened since 1990 (when climate data are available). The pointwise SMB and the mean specific mass balance are compared with the modelled mass balances. Besides that, also the profile of the modelled and the observed SMB are compared over the different observation years and as multi-year mean.

The different measures point different optimal combinations of C_0 and C_1 . The ϵ -constraint method helps to find the best compromise between the different mass balance measurements by focusing on optimizing one main goal while keeping the others within acceptable limits. One measure must be chosen to be most important to get right. Since the stake-measurements are likely to be the most precise since they are not interpolated and they are of yearly precision. A reasonable maximum error limits for the other measurements like the geodetic balances and elevation profiles is set based on the accuracy of the measurements. For each combination of C_0 and C_1 , it is checked if all the measurements stay within these limits. From all the combinations that meet these requirements, the combination that gives the smallest error for the stake measurements is selected.

Historical mass balance

A continuous historical dataset of mean specific mass balances since the Little Ice Age (LIA) can be obtained by feeding the model with historical meteorologic data. Van Tricht et al. (2021b) reconstructed the Kumtor-Tien Shan meteorological dataset. A measuring station at an elevation of 3659 m has monthly temperatures and monthly total precipitation data since 1930. Since 2005, the temperature and precipitation are measured every three hours. The distribution of the three hourly data since 2005 is applied to the monthly averages since the beginning of the measurements in 1930. To deal with a replacement of the measuring

station in the summer of 1996, a correction was applied based on the correlation with another measuring station in the neighbourhood. In order to reconstruct the climate for a longer period back, it is not possible to only rely on the direct measurements. Van Tricht et al. (2021b) managed to create a historical climate dataset up to 1750 by using tree ring reconstructions and the CRUTEM4 (Climatic Research Unit Temperature) dataset. The latter is a grid-based dataset of historical air temperatures over land based on station data (Osborn et al., 2020). These data are, just like the in situ measurements, rescaled to data per three hours based on the distribution between 2005 and 2020. A more detailed description of the reconstruction can be found in Van Tricht et al. (2021b). The reconstructed temperature and precipitation, together with their anomalies from the mean 1821-1850 climate, are depicted in figure 2.9.

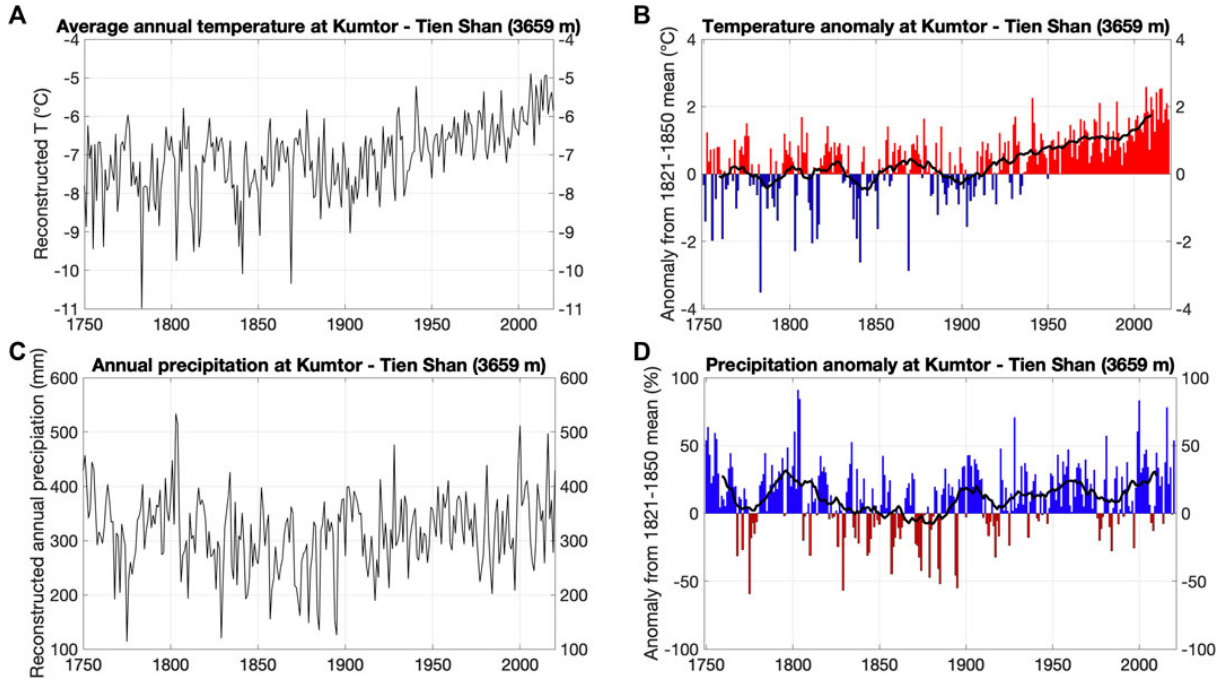


Figure 2.9: Figure taken from Van Tricht et al. (2021b). Historic reconstruction of the temperature (a) and the precipitation (c) for the Kumtor-Tien Shan Meteostation (3659 m). Plots (b) and (d) represent the anomalies from the mean 1821-1850 climate

To reconstruct the historical mass balance, the meteorologic data from the described dataset have to be adapted for the glaciers location. For every gridcell covering the glacier, the temperature at a certain time can be approached as is written in equation 2.28.

$$T_{air} = T_{ref} + \gamma_T * (H_{sur} - H_{ref}) \quad (2.28)$$

In equation 2.28, H_{ref} stands for the elevation of the meteorological station where the measurements were done and γ_T is the lapse rate of the temperature with the elevation. The precipitation at a certain point in space and time can be calculated by multiplying the reference precipitation with a coefficient (P_{ratio}) that can be calculated as shown in equation 2.29.

$$P_{ratio} = \frac{P_{ref} + \gamma_p * (H_{sur} - H_{ref})}{P_{ref}} \quad (2.29)$$

Note that the mass balance does not require the total precipitation, which is the multiplication of P_{ref} and P_{ratio} , but the solid precipitation P_{ratio} . Whether the precipitation is solid (snow) or liquid (rain) depends on the air temperature. For a temperature below 0.5 °C, almost all precipitation is snow, where above 2.5 °C almost all precipitation is liquid. For the area in between, equation 2.30 is used.

$$P_s = P * \left(\frac{T_s - T_{air} + 1}{2} \right) \quad (2.30)$$

Future mass balance

For the future climate, five SSP of the Sixth Assessment Report (AR6) of the Intergovernmental Panel on Climate Change (IPCC) are used. These SSP's represent different ways in which the global society and economy will evolve in the (near) future. That determines the evolution of the global greenhouse gas emissions and therefore also how climate will change. The used scenarios are SSP1-1.9, SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5. The numbers indicate the expected level of radiative forcing by the year 2100. For example SSP1-1.9 is expected to have a radiative forcing of 1 to 1.9 W/m^2 by the year 2100. SSP1-1.9 and SSP1-2.6 are the scenarios with low greenhouse gas emissions (becoming net negative in respectively 2050 and 2075), SSP2-4.5 has intermediate emissions and the two others have high emissions (having a doubling or a tripling in Greenhouse Gas (GHG) emissions before the end of the century) (IPCC, 2023).

Van Tricht and Huybrechts (2023) used the SSPs described above to feed a number of Global Circulation Models (GCM) that are considered in the sixth phase of the Coupled Model Intercomparison Project (CMIP6). The historical and future model output are downloaded for the gridcells closest to the Tien Shan Kumtor meteo station. The gridcell resolution varies between 0.93° and 2.50°. The combination of all models leads to a multi-model mean time series for monthly precipitation and temperature. These data are downscaled to hourly resolution based on observations of the Tien Shan Kumtor meteo station. For a more detailed description of the assemblage of the future meteorological datasets, Van Tricht and Huybrechts (2023) can be consulted.

2.3 Glacier evolution models

Physics

Many physical processes determine a glacier's behaviour, such as ice flow, thermodynamics, subglacial hydrology, and their interactions with the atmosphere, lithosphere, and ocean. Combining equations for these processes, results in glaciological models. However, the increasing complexity of these models leads to increasing computational demands that recent advances in numerical methods and computing power cannot fully meet.

The motion of glaciers is governed by the Stokes equations, which express the conservation of momentum in a fluid. Ice, while solid, behaves as a very slow-moving, viscous fluid over long timescales. Combining the Stokes equations with an expression for the strain rate results in a set of three dimensional equations that describe the flow of a glacier.

Stokes equations

The basis of glacier evolution models is the conservation of mass, which can be expressed as the divergence of the velocity field being zero when ice is assumed to be incompressible (equation 2.31).

$$\nabla \cdot \vec{v} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad (2.31)$$

In equation 2.31, u , v and w are the velocity components in the x , y and z direction, respectively. This condition ensures that flow divergence is balanced by convergence elsewhere, i.e., no mass is lost or gained internally. If equation 2.31 is integrated between the bed ($z = 0$) and the glacier surface ($z = h$), the vertically averaged continuity equation for ice thickness h as a function of time t is found. In equation 2.32, this vertically averaged continuity equation is combined with a SMB.

$$\frac{\partial h}{\partial t} = -\nabla \cdot (h \vec{v}) + SMB \quad (2.32)$$

$\vec{v}$ is the vertically averaged horizontal velocity ($\vec{v} = (\bar{u}, \bar{v})$). This equation forms the cornerstone of most glacier flow models, providing a direct link between ice flux and thickness evolution. The SMB provides the climatologic boundary condition.

To compute velocities needed in equation 2.32, the conservation of momentum, derived from Newton's second law for a slow, viscous, and incompressible fluid is used (equation 2.33).

$$\rho \frac{\partial \vec{v}}{\partial t} = \nabla \cdot \tau + \rho g \quad (2.33)$$

Since in most cases ice is flowing relatively slow, the acceleration term can be neglected ($\frac{\partial \vec{v}}{\partial t} = 0$). The total stress tensor τ can be split into two parts: the cryostatic pressure P , which acts equally in all directions and arises from the weight of the overlying ice, and the deviatoric stress tensor τ' , which captures shear stresses that drive deformation without changing volume. Because glacier ice is considered incompressible, changes in pressure do not cause volume changes and therefore do not contribute directly to the flow. Instead, it are the deviatoric stresses (those that change the shape of the ice) that are responsible for its deformation and movement. This is why we reformulate the momentum balance in terms of τ' , leading to equation 2.34.

$$\nabla \cdot \tau' - \nabla P + \rho g = 0 \quad (2.34)$$

To link stress to deformation, we use Glen's flow law (equation 2.9). This relation can be also described in terms of deviatoric stress (τ') and strain rate ($\dot{\epsilon}$) as in 2.35.

$$\dot{\epsilon}_{ij} = A(T) \tau_e^{n-1} \tau'_{ij} \quad (2.35)$$

In equation 2.35, $\dot{\epsilon}$ is the strain rate tensor, τ' the deviatoric stress components, n is the flow law exponent (commonly taken as 3), and $A(T)$ is the temperature-dependent flow rate factor. τ_e is the effective stress, defined from the second invariant of the stress tensor. Since n is not equal to 1, this law is nonlinear and captures how ice deforms faster under higher stress and warmer temperatures (see figure 2.5). Glen's flow law enables the calculation of

strain rates, and thus velocity gradients, based on the local stress state, which is essential for dynamic ice flow modelling. The flow rate factor $A(T)$ is a pre-exponential constant dependent on the material properties of ice. In equation 2.36, the expression used to approximate $A(T)$ is provided.

$$A(T) = E \cdot a \cdot \exp\left(-\frac{Q}{R \cdot T^*}\right) \quad (2.36)$$

Equation 2.36 has its base in the Arrhenius equation. This correction lowers the melting point of ice under pressure, simulating pressure melting at glacier bases. E is an enhancement factor which is tuned. Factor a is a pre-exponential factor that accounts for the material. In the case of ice, this is temperature dependent. Q is the activation energy for creep, representing the energy barrier for ice molecules to rearrange under stress. Also Q is temperature dependent. R is the universal gas constant. T^* is the temperature adjusted for the Pressure Melting Point (PMP) as in equation 2.37.

$$T^* = T - T_m + T_0 = T - (PMP \cdot \zeta \cdot h) + T_0 \quad (2.37)$$

In equation 2.37, T is the temperature expressed in $^{\circ}\text{C}$, PMP is the pressure melting point lowering ($8.7 \cdot 10^{-4} \text{K m}^{-1}$), T_m is the pressure melting temperature, T_0 is 273.15K (to convert $^{\circ}\text{C}$ to K), h is the ice thickness and ζ is a dimensionless vertical coordinate (scaled between 0 and 1).

From Glen's law, the effective viscosity η can be defined as in equation 2.38.

$$\eta = \frac{1}{2} A(T)^{-\frac{1}{n}} \dot{\epsilon}_e^{\frac{1-n}{n}} \quad (2.38)$$

In equation 2.38, $\dot{\epsilon}_e$ stands for the effective strain rate, calculated from the second invariant of the strain rate tensor. The effective strain rate is given in equation 2.39.

$$\dot{\epsilon}_e = \sqrt{\frac{1}{2} \left(\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 + \left(\frac{\partial w}{\partial z} \right)^2 \right) + \frac{1}{4} \left(\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right)^2 \right)} \quad (2.39)$$

This scalar value allows us to evaluate viscosity and stress without handling full tensors. Inserting viscosity into the momentum balance and expressing everything in terms of velocity gradients yields the full 3D Stokes equations for glacier ice (equation 2.40).

$$\begin{cases} \frac{\partial}{\partial x} (2\eta \frac{\partial u}{\partial x}) - \frac{\partial P}{\partial x} + \frac{\partial}{\partial y} \left(\eta \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \right) + \frac{\partial}{\partial z} \left(\eta \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right) \right) = 0 \\ \frac{\partial}{\partial x} \left(\eta \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \right) + \frac{\partial}{\partial y} (2\eta \frac{\partial v}{\partial y}) - \frac{\partial P}{\partial y} + \frac{\partial}{\partial z} \left(\eta \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right) \right) = 0 \\ \frac{\partial}{\partial x} \left(\eta \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right) \right) + \frac{\partial}{\partial y} \left(\eta \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right) \right) + \frac{\partial}{\partial z} (2\eta \frac{\partial w}{\partial z}) - \frac{\partial P}{\partial z} = \rho g \end{cases} \quad (2.40)$$

The full Stokes equations for ice flow are derived by combining the momentum balance (Eq. 2.34) with Glen's flow law (Eq. 2.35) and expressing them in terms of velocity gradients. This yields a system of three equations representing force balance in the x , y , and z directions. These equations form a system of nonlinear partial differential equations that describe how ice flows under its own weight, driven by gravity and resisted by internal deformation and basal drag.

Blatter-Pattyn approximation

Already since the 1950s, ice has been modelled as a viscous, non-Newtonian fluid with the 'Full Stokes' equations (Glen, 1953). These equations describe all shears and stresses within the ice and therefore provide the most accurate representation. The downside of this accurate model, are the extremely expensive computational costs that come with it, which makes them difficult to apply on larger scales. A zeroth order approach that is often used is the Shallow Ice Approximation (SIA). SIA assumes that ice flow is dominated by vertical shear stresses (Cuffey and Paterson, 2010). The Shallow Shelf Approximation (SSA) on the other side, is used in areas where horizontal stresses are more important, like fast flowing glaciers or floating ice. A hybrid model that applies SIA and SSA on different places tries to strike a balance between accuracy and efficiency, simplifying stress balance to 2-D equations and enabling long-term simulations. The Parallel Ice Sheet Model (PISM), for example, uses this approach and can therefore be used to simulate for example glacial cycles (Seguinot et al., 2018).

Despite their overall good performance, these approximations are probably too simple for a glacier, where the ratio between the width and the depth is not that high. This makes them less suitable to model the evolution of a glacier, although in many cases their result is rather satisfactory. Therefore, higher order models are required that simplify the computationally expensive full Stokes equations, but that are on the other side more accurate than SIA or SSA. In this thesis, the Blatter/Pattyn approximation is used (Blatter, 1995; Pattyn, 2003).

The main assumption of this Blatter/Pattyn approximation is that the vertical stress balance is dominated by the ice weight. Therefore, the vertical component of the momentum balance is approximated in cryostatic equilibrium. In other words: vertical resistive stresses are neglected, as shown in equation 2.41.

$$\frac{\partial \tau'_{zz}}{\partial z} - \frac{\partial P}{\partial z} = \rho g \quad (2.41)$$

By applying this, the pressure is removed from the mass balance. Therefore, the full Stokes equations simplify to a system with only the two horizontal velocity components as unknown variables. Physically, this means that stress transfer between different parts of the ice (bridging effects) is neglected. A second assumption is that horizontal gradients of the vertical velocity are small compared to the vertical gradient of the horizontal velocity (equation 2.42).

$$\frac{\partial w}{\partial x} \ll \frac{\partial u}{\partial z} \quad \text{and} \quad \frac{\partial w}{\partial y} \ll \frac{\partial v}{\partial z} \quad (2.42)$$

In this thesis, two models are used to simulate the evolution of the glacier: the higher-order model and Instructed Glacier Model (IGM). Both models implement a Blatter/Pattyn approximation, but in a different way.

Higher order model

One approach to simulating glacier dynamics is through a HOM. The HOM was developed and described by Fürst et al. (2011), and later amongst others used by Zekollari et al. (2014) and Van Tricht and Huybrechts (2023). The HOM solves the full 3D momentum equations for the horizontal velocity components, while neglecting vertical resistive stresses. This allows

the model to account for longitudinal and transverse stress gradients, making it suitable for regions with complex topography or strong flow variations.

The model requires boundary conditions. At the ice surface (so the ice-atmosphere boundary), a stress free boundary condition is used. This assumes that the air above the glacier exerts negligible stress on the ice surface, which is a valid assumption since atmospheric pressure is small compared to internal ice stresses.

At the bottom of the glacier, the choice of boundary condition depends on whether there is sliding at the base or not. For cold-based glaciers or parts of a glacier where the temperature is below the pressure melting point, it is assumed that the ice is frozen to the bed and no basal sliding occurs. For temperate glaciers, a Weertman-type sliding law is applied, which allows for basal sliding driven by the shear stress at the base (the so called basal drag). The boundary conditions are summarised in system 2.43.

$$\begin{cases} \vec{n} \cdot (\tau_{ij} \cdot \vec{n})_s = 0 & \text{for } z = s \\ \vec{t} \cdot (\tau_{ij} \cdot \vec{t})_b = 0 & \text{for } z = s - h \text{ and } T < T_m \\ \vec{t} \cdot (\tau_{ij} \cdot \vec{t})_b = \tau_b & \text{for } z = s - h \text{ and } T \geq T_m \end{cases} \quad (2.43)$$

In this system, $\vec{n}$ is the outward pointing unit vector normal to the surface of the ice, while $\vec{t}$ is the tangent vector to the surface elevation. s is the surface elevation and h is the ice thickness. The subscripts b and s stand for respectively the base and the surface. T_m is the temperature at the pressure melting point. The basal shear stress τ_b is described by equation 2.44.

$$\tau_b = -\rho g h (\nabla s) = -\rho g h \sqrt{\left(\frac{\partial s}{\partial x}\right)^2 + \left(\frac{\partial s}{\partial y}\right)^2} \quad (2.44)$$

To convert the basal shear stress into sliding, a Weertman-type of parametrisation was used as described by equation 2.45.

$$u_b = basl \cdot \tau_b \quad (2.45)$$

In this equation, the basal sliding velocity u_b is related to the basal shear stress τ_b by mean of a geometrically dependent sliding coefficient $basl$, which is calculated in equation 2.46.

$$basl = ASL \cdot (\rho g)^2 \cdot h \cdot |\nabla s|^2 \quad (2.46)$$

The HOM requires a maximum extent of the glacier during the LIA. The model starts from the LIA geometry. To avoid glacier drift, it is assumed the glacier experienced steady-state at the LIA. The model is forced to describe the retreat of the glacier since this period. Therefore, a hillshade image of the DEM is used to detect the end moraine. The end moraine is assumed to be the maximal extent of the glacier during the LIA. The maximum extent the glacier is allowed to grow is therefore placed at the end moraine. Besides a lateral growth, the glacier might have been larger at the sides as well. Since there are no sidemoraines visible on the orthophoto or on the DEM, it is assumed to have an accumulation area with a similar shape as today. The shape of the current ablation zone is also used as a baseline, under the assumption that lateral spreading in this region was limited during the LIA. In the absence

of geomorphological evidence for a significantly broader tongue, the present glacier width is taken as representative. This maximum extent provides the baseline geometry.

First, the model is fed by the LIA-climate (average between 1821 and 1850) to spin up until it reaches the LIA-outline. From there on, the HOM was fed by the climate data until the mass balance year 2023-2024 and further with the projections concerning the different SSP's.

It is important for the glacier to match the outline in the present day, as well as in the LIA. Since both the historical climate data before 1990 and the mass balance model contain errors, a dynamic calibration of the glacier is performed. To match the present day length of the glacier, the climate forcing is perturbed. For a too short glacier, the precipitation is enhanced and/or the temperature is lowered. For a too long glacier, the opposite is done. The modelled glacier length is compared with the observed glacier length in different moments in time, based on Landsat images. The LIA outline is matched by adding or removing mass. This process is referred to as *dynamic calibration*. Once calibrated, the model can be run for the different SSPs.

The calculation of the glacier length on satellite images at the different moments in time was done in an automated way. All Landsat images of the glacier since 1975 were downloaded via Google Earth Engine. Only the images in the months September and October were collected in order to sample only at the end of the melt season. A cloud cover threshold of 20% was set. A straight line was drawn between the highest and the lowest point of the glacier in the LIA-outline. Along this line, points were generated. For all satellite images, the length was calculated by the number of points located on the glacier. For the simulations by the models, the same approach was used: the points that have an ice thickness value larger than zero determine the length of the glacier in that year.



Figure 2.10: Example of the measurement of the length of the glacier on satellite images, here for a Landsat image of 25/09/2025. The points are used to determine the length of the glacier.

Two parameters of the flow model have to be calibrated: the enhancement factor E (equation 2.36) and the sliding parameter ASL . The enhancement factor is spatially uniform and acts as a tuning parameter to adjust the softness of the ice across the domain. ASL controls how efficiently basal shear stress is converted into sliding. As for C0 and C1 in the mass balance model, the HOM was run for different combinations of E and ASL . For this calibration runs, a steady-state average (baseline) climate between 1961 and 1990 was used to spin up the glacier. From 1990 onwards until 2023, the HOM was run with the corresponding climate. The ice thickness observations (with the RES) from 2023 were compared with the modelled ice thicknesses to validate which combination of E and ASL is the best.

As additional validation, the surface velocity could be used, as for example was done by Zekollari et al. (2013). Since no direct velocity measurements are available, satellite images of Sentinel-2 were used to detect large scale crevasses. Although this small glacier has relatively few crevasses that are small and therefore hard to detect, some crevasses could be identified in consecutive years. By measuring the distance the crevasses have travelled, an indication of the surface velocity could be derived. The identification of individual crevasses in multiple years is extra difficult due to the occurrence of clouds and by snowfall that covers the crevasses. An example of an identified crevasse in two years is depicted in figure 2.11.

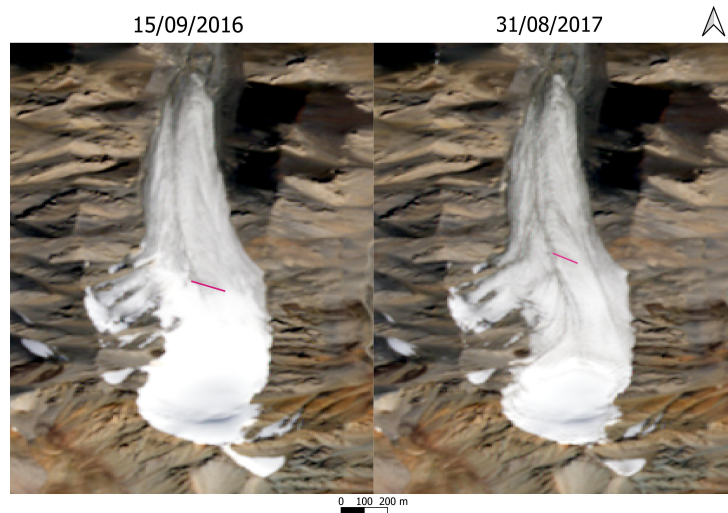


Figure 2.11: Example of identification process of the crevasses. On the satellite image from 2016 (left), a crevasse was identified with a red line. The same crevasse is identified on the satellite image from 2017 (right). The distance with which the crevasse moved indicates the surface velocity of the glacier.

Just as for the calibration of the mass balance model, the final choice of E and ASL has been made by using a ϵ -constraint method.

Instructed Glacier Model

To solve higher-order models, improving computer architecture can help a hand. Numerical approximations of higher-order models can be transferred from central processing units (Central processing Units (CPU)s) on graphical processing units (GPUs), although the large amount of processes that have to run parallel is a complex task for GPUs as well (Jouvet et al., 2022).

The Instructed Glacier Model is a glacier evolution model developed by Jouvét et al. (2024). It integrates physics-driven equations with deep learning techniques to improve computational efficiency and accuracy. Deep learning surrogate models might be a solution for the high computational costs. The idea is that a Convolutional Neural Network (CNN) can substitute the complex Glen-Stokes like equations. CNNs are a type of Artificial Neural Network (ANN) that use kernels to analyse only parts of the data, which makes them useful for spatial data. In addition, ANNs are known to perform very efficient on GPUs. However, as with all statistical models, they are extremely dependent on the training data, which causes poor performance outside the range of the training data. An alternative that has recently developed, is a Physics-Informed Neural Network (PINN). PINNs integrate physical laws directly into the learning process, rather than relying on datasets produced by physical models. The basic forms minimize the residuals of the equations and boundary conditions. A Variational Physics-Informed Neural Network (VPINN) on the other side, uses the minimization form of the problem as their loss function, offering the benefit of requiring lower-order derivatives than residual-based methods (Jouvét et al., 2022).

IGM is built on modules, divided in pre-processing, processing and post-processing modules, as depicted in figure 2.12. From the pre-processing modules, the optimize module has already been discussed before, since it is used to interpolate the ice thickness measurements. To read in the measured data, in this thesis the `load_ncdf` module is used. This module requires input in the form of a NetCDF file. Other options are loading separate TIF files (`load_tif`) or taking data from the Open Glacial Global Model (OGGM) by using the `oggm_shop` module (Jouvét et al., 2024).

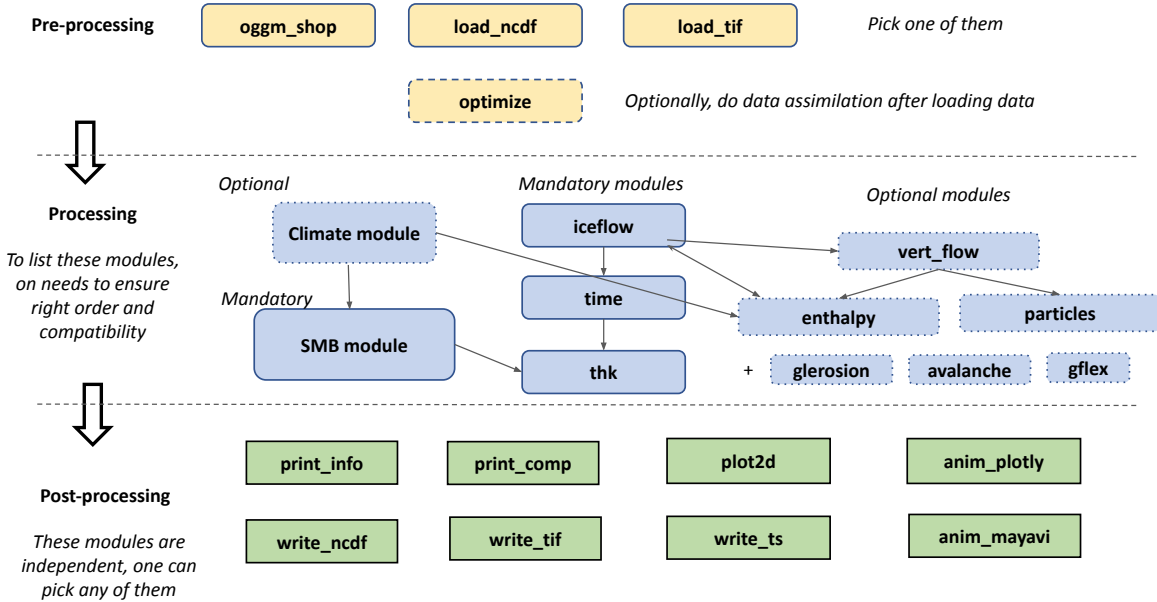


Figure 2.12: Flowchart of IGM modules (Jouvét et al., 2024)

The processing modules in IGM are responsible to update the glacier fields. They implement the physical models within a time iteration loop. The `climate` module generates the climate forcings: air temperature, precipitation and air temperature variability. These fields are used

to force the *SMB* and the optional *enthalpy* module. The *SMB* modules update the climatic surface climate mass balance. Some of these modules start from given parameters, others require climatic inputs (Jouvet et al., 2024). Although IGM includes a *SMB*-model, it is chosen to use the previously explained mass balance model as implemented in the HOM. Therefore, the *climate* and *SMB* modules are not used. Also the same LIA-geometry reconstructed via the dynamical calibration of the HOM has been used as input of IGM. By using the same mass balance and the same LIA-geometry for both models, an easier comparison between both model implementations (with and without emulator) is possible.

It is the *ice flow* module where the 3D CNN based on a physics informed neural network comes to play. Inputs for the CNN are surface topography s , ice thickness h , ice flow parameters $A(T)$ and sliding coefficient c . On a grid of spatial resolution H_H , the horizontal ice flow (u, v) is predicted. The function tries to minimize the energy represented by the Blatter-Pattyn equations through the cost function in equation 2.47.

$$J(v) = \int_V 2 \frac{A^{-\frac{1}{n}}}{1 + \frac{1}{n}} |D(v)|^{1+\frac{1}{n}} dV + \int_\Gamma \frac{c}{1+m} |v|_M^{1+m} dS + \rho g \int_V (\nabla s \cdot \mathbf{v}) dV - \int_{\partial\Omega} \mathbf{F}(h) \cdot \mathbf{v} dS \quad (2.47)$$

In this cost function, V is the ice volume and Γ is the interface with the bedrock.

The *time* module calculates the timestep based on a pre-defined maximal timestep and the frequency of saving times. To ensure stability in the transport scheme, a Courant-Friedrichs-Lewy condition is used. The ice thickness *thk* module uses the conservation of mass (equation 2.31) to calculate the ice thickness in each timestep. An explicit first-order upwind finite volume scheme is used to simulate mass transport between cells. This scheme is mass-conserving and fully parallelizable, but requires strict observance of the CFL condition. Since the *iceflow* module only calculates horizontal velocities, the vertical velocity is calculated by the vertical flow module. This module uses the incompressibility equation (equation 2.48).

$$\nabla \cdot \mathbf{u} = 0 \quad (2.48)$$

These are the only modules used for this thesis. Remark that no thermodynamic model was used. The *enthalpy* module contains the thermodynamics for IGM. The enthalpy model describes the thermodynamic state of the ice, including temperature and water content. It solves an advection-diffusion equation using two time-splitting methods. The module also calculates the basal melting rate, water discharge and shear coefficient. Since the *enthalpy* module requires input from the *climate* module, which is unused since the *SMB* was provided, no thermodynamics are included during this experiment. The flow rate factor $A(T)$ is taken constant and does not changes in space or in time, in contrast to the HOM-experiment. The used value of $A(T)$ is determined during the optimisation in the first part of this thesis.

Table 2.6: Comparison of HOM and IGM model characteristics.

Category	HOM	IGM
Ice Flow		
Solver type	Physics-based	Physics-Informed Convolutional Neural Network
Solver speed	CPU-based, slower	GPU-accelerated via TensorFlow
Numerical Framework		
Language	Fortran + Bash	Python
Grid	Regular 3D finite-difference grid	2D raster grid + vertical layers
Parallelisation	Limited	GPU-parallelised
Time stepping	–	Explicit schemes with CFL-condition enforcement
Thermal Model		
Temperature	Temperature only	Full enthalpy model (includes water content)
Meltwater refreezing	Approximate	Modelled explicitly (via water content)
Energy conservation	Temperature only	Full energy-conserving enthalpy scheme
Basal Conditions		
Subglacial hydrology	Not included	Till-water based model
Basal sliding	Weertman-type	Based on basal water pressure

Chapter 3

Results

3.1 Ice thickness

Interpolations

Figure 3.1 presents an overview of the interpolated ice thickness fields. Spatial interpolations based on the bottom and top 50% of the data yield unrealistically low ice thickness values in regions lacking observational data. In contrast, interpolations using randomly selected 50% and 90% subsets, as well as those based on central flowline data, exhibit a less pronounced dependence on data coverage. Nevertheless, in the highest parts of the accumulation area (located in the far south), the interpolations still indicate thin ice due to the absence of RES measurements. Additionally, these spatial interpolation methods show a relatively high sensitivity to individual data points.

The yieldstress method produces a thicker ice field in the unmeasured areas for the bottom and top 50% measurements. For the bottom 50%, the highest part of the glacier has a relatively uniform ice thickness with thinner ice only at the margins. The interpolation based on the top 50% creates a very thick and uniform lower part of the glacier with an extremely thin tongue. The interpolations with random 50 and 90 % of the observations are relatively similar to the spatial interpolations, except that they are more smooth. The individual measurements are less prominent than in the spatial interpolations. The interpolation based on the central flowline looks very similar to the random 50 and 90 % and performs considerably better than the spatial interpolations using the central flowline measurements. The highest places in the accumulation area show an ice thickness field that is likely to be too thick, but the interpolation is much better than the accumulation areas of the spatial interpolation techniques. For all interpolations based on the yieldstress method, the steep parts at the side of the glacier are assumed to be connected via (very thin) ice, while for all other interpolation techniques, these parts are not interpolated to be ice covered. A possible explanation for that might be the explicit ice mask (zero outside the margins) for those techniques, while for the yieldstress method only the last pixel within the outline is assigned to be 5 meters thick. In those areas, where the distance between the different parts of the glacier is very small, the ice is interpolated between the different pixels with a value of 5 meters.

In contrast, the interpolations using IGM produce a significantly smoother ice thickness field. Individual measurement points are not discernible in the resulting distribution, and the overall pattern appears more physically realistic. Compared to conventional spatial interpola-

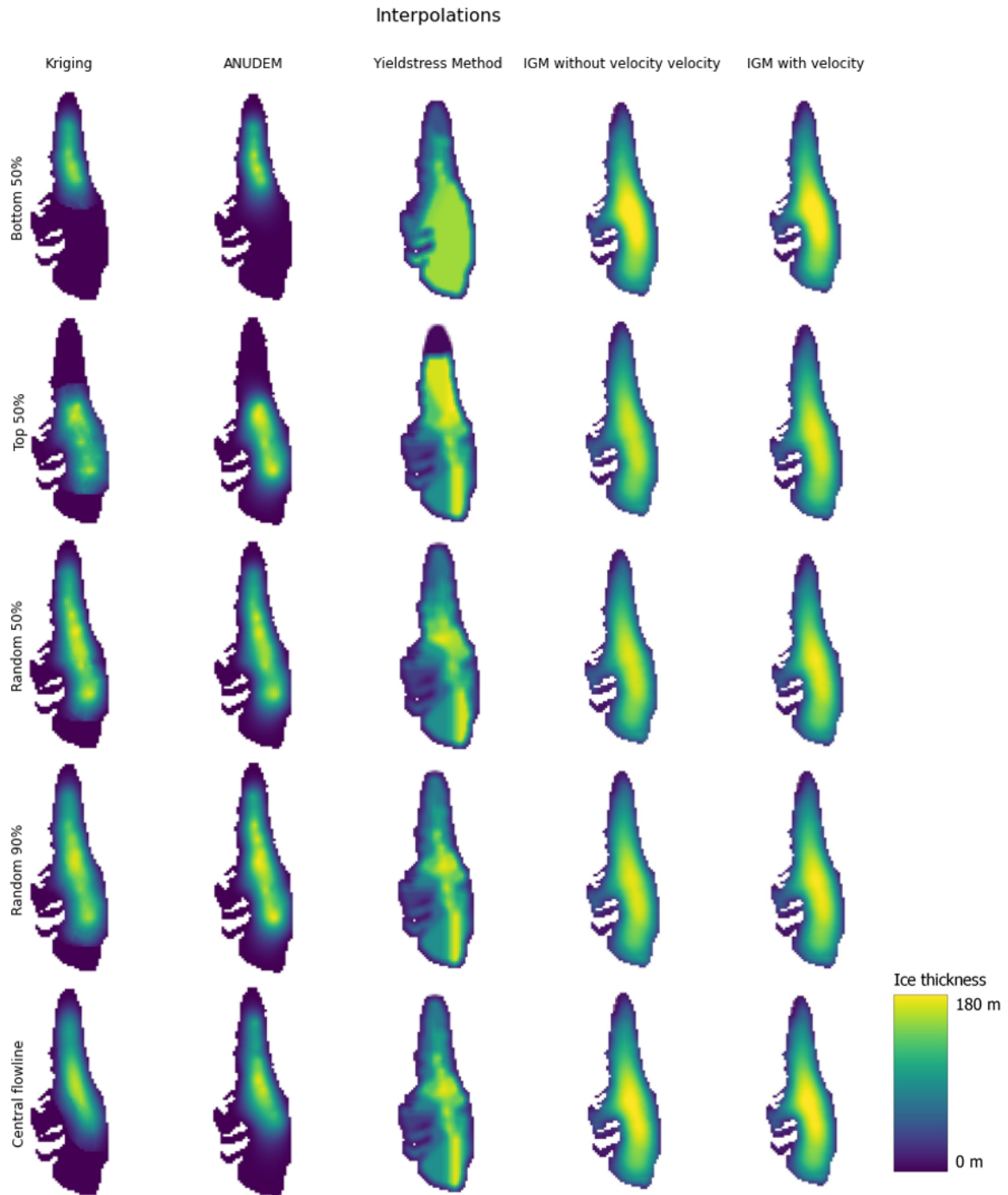


Figure 3.1: Ice thickness fields for the different interpolation methods based on the different subsets.

tion techniques, IGM shows much less dependence on the spatial configuration of the data. Furthermore, the IGM-based interpolations using the bottom and top 50% do not artificially assign low ice thickness values to regions lacking observations.

Accuracy analysis

Figure 3.2 shows the residuals of all ice thickness observations, including both used and unused data points. The depicted residuals are all the observation points, both the used and the non-used points for the interpolation.

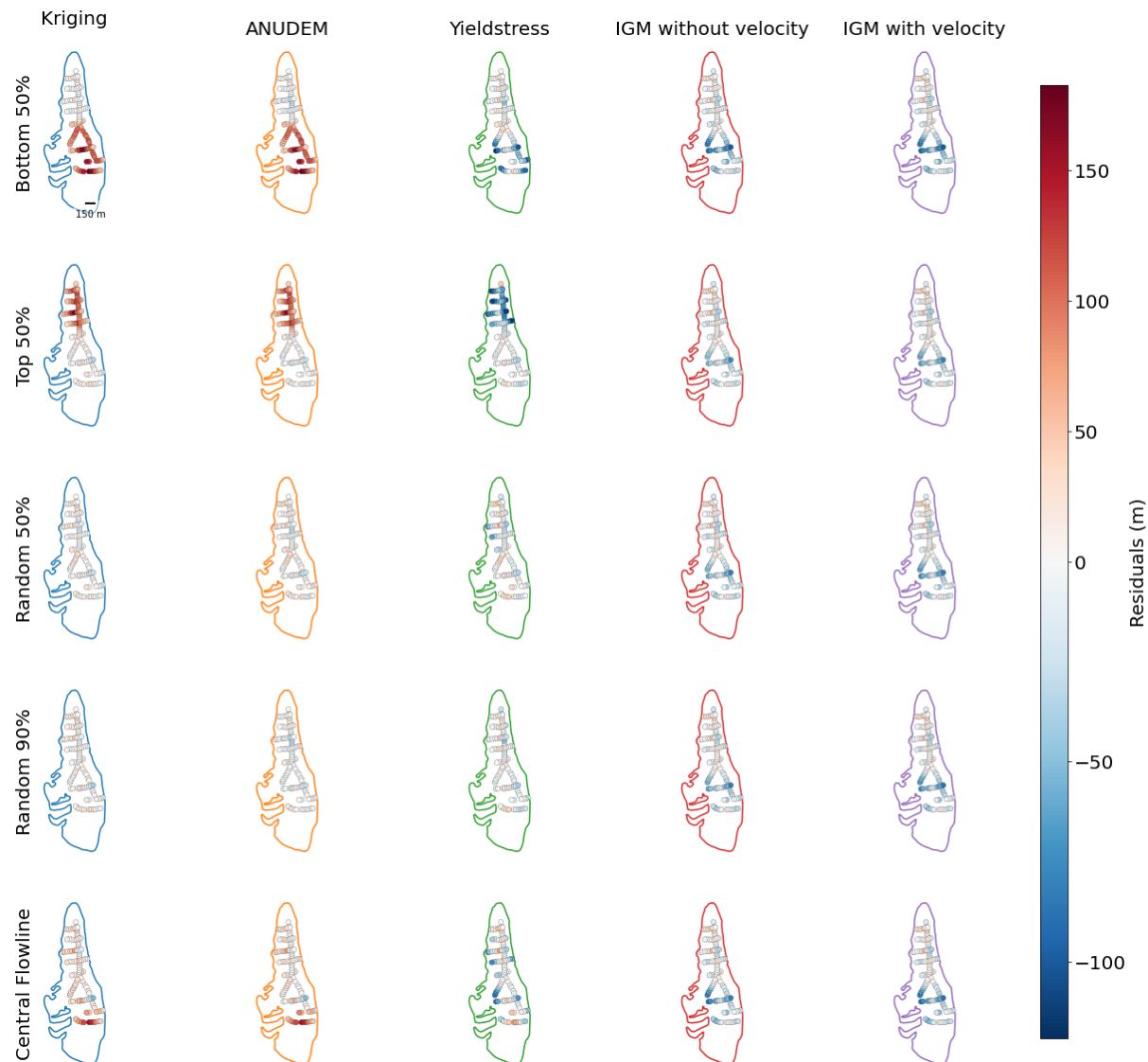


Figure 3.2: Residual maps of the different interpolation techniques. The residuals are depicted for all RES-measurements, both the points used for the interpolation and the unused points.

The spatial interpolation maps and the yield stress method based on the bottom and top 50% of the data exhibit relatively high residuals in regions lacking observational coverage, while residuals remain comparatively low in areas where measurements are available. In contrast, interpolations using IGM demonstrate significantly lower residuals across the domain. Spatial interpolation techniques tend to overestimate ice thickness in data-sparse regions, whereas the yield stress method underestimates the ice thickness in well measured areas. IGM produces fields that slightly underestimate ice thickness in the upper regions and slightly overestimate it in the lower regions, regardless of the input dataset.

Interpolations based on randomly distributed observations result in considerably lower residuals overall. In these cases, spatial interpolation techniques yield smaller residuals than IGM. The yield stress method has residuals in between for the interpolation based on 50% of the data. When 90% of the data was provided, the yield stress method approaches the residuals of the spatial interpolation methods. However, the resulting ice thickness fields derived from kriging, ANUDEM and, to some extent, the yield stress method, show signs of overfitting. When interpolations are based on datasets along the central flowline, residuals from spatial interpolation techniques increase, especially in the upper regions near the accumulation area, where ice thickness tends to be overestimated. In contrast, IGM continues to produce lower (negative) residuals in these regions, indicating better performance than the spatial methods. The yield stress method follows the pattern of the IGM interpolations.

The residual distribution offers important insights into the performance of each interpolation method. To illustrate this, histograms of the residuals are provided in Figure 3.3, with the dotted red line indicating zero residuals.

The histograms of residuals from the spatial interpolation techniques based on the top and bottom 50% of the data exhibit a wide spread and significant deviation from a normal distribution. These distributions are highly skewed, with pronounced right tails. In contrast, the residuals from the IGM-based interpolations and the yield stress method are less skewed and exhibit a shape more closely resembling a Gaussian distribution. While a slight left tail is still present, it is considerably smaller than the heavy tails observed in the spatial interpolation results.

Interpolations based on randomly distributed observations yield residual distributions that more closely approximate a normal distribution. This improvement is particularly evident in the case of the spatial interpolation methods. The yield stress method also results in more normally distributed residuals, especially for the interpolation based on random 90% of the observations. For IGM-based interpolations, the enhancement in normality is less pronounced, though the previously observed left tail largely disappears. Among all techniques, the residual histograms of the IGM interpolations align most closely with a normal distribution for the top and bottom 50%, outperforming both kriging and ANUDEM. For the random distributions, the yield stress method becomes the method which has the most normal distribution. In contrast, the spatial interpolation methods produce distributions with a narrow central peak and extended tails, indicating a leptokurtic shape. For the interpolations based on the central flowline, the histograms deviate more from the gaussian shape than the random distributed datasets. The largest deterioration occurs for the spatial interpolation methods, which again show the right tails, and the yield stress method, with the left tail. The IGM-interpolations again show the strongest consistence.

To facilitate visual assessment of the residuals' normality, quantile-quantile (QQ) plots were generated for each interpolation method and are presented in figure 3.4. The red line in each plot represents the expected behaviour under a theoretical normal distribution.

The QQ-plots of the spatial interpolations based on the top and bottom 50% of the data lie above the theoretical quantiles on the left side of the graph, indicating a heavier left tail than expected under a normal distribution. In other words, these interpolations produce more extreme negative residuals than would be predicted by a Gaussian distribution. At the same time, the central portion of the residual distribution is more tightly clustered around zero, suggesting reduced variability in the centre. This combination of a peaked centre and heavy

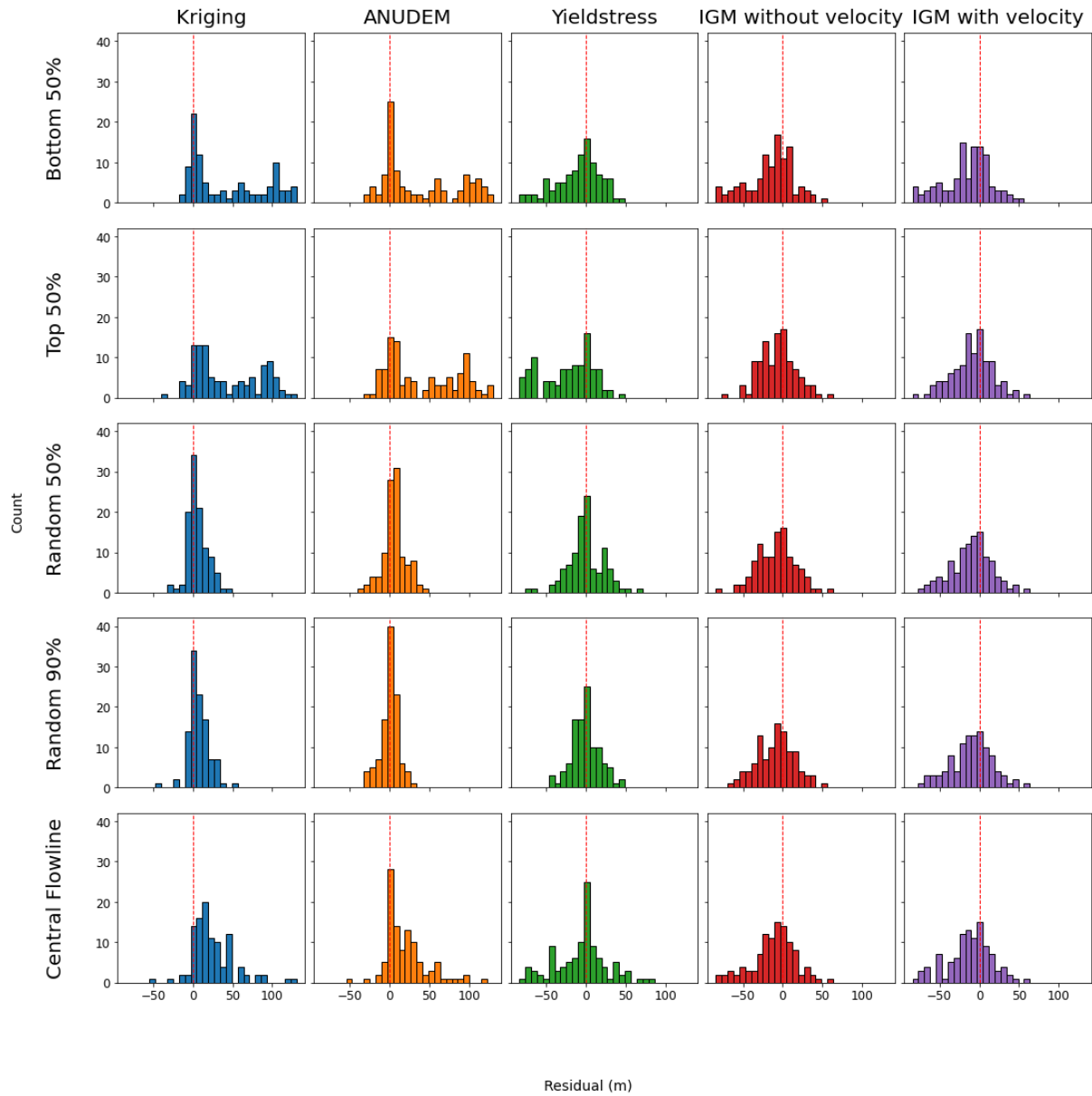


Figure 3.3: Histograms of the residuals of the different interpolation techniques. The residuals are calculated for the RES-measurements that were not used for the interpolation. The vertical red line represents the null-line.

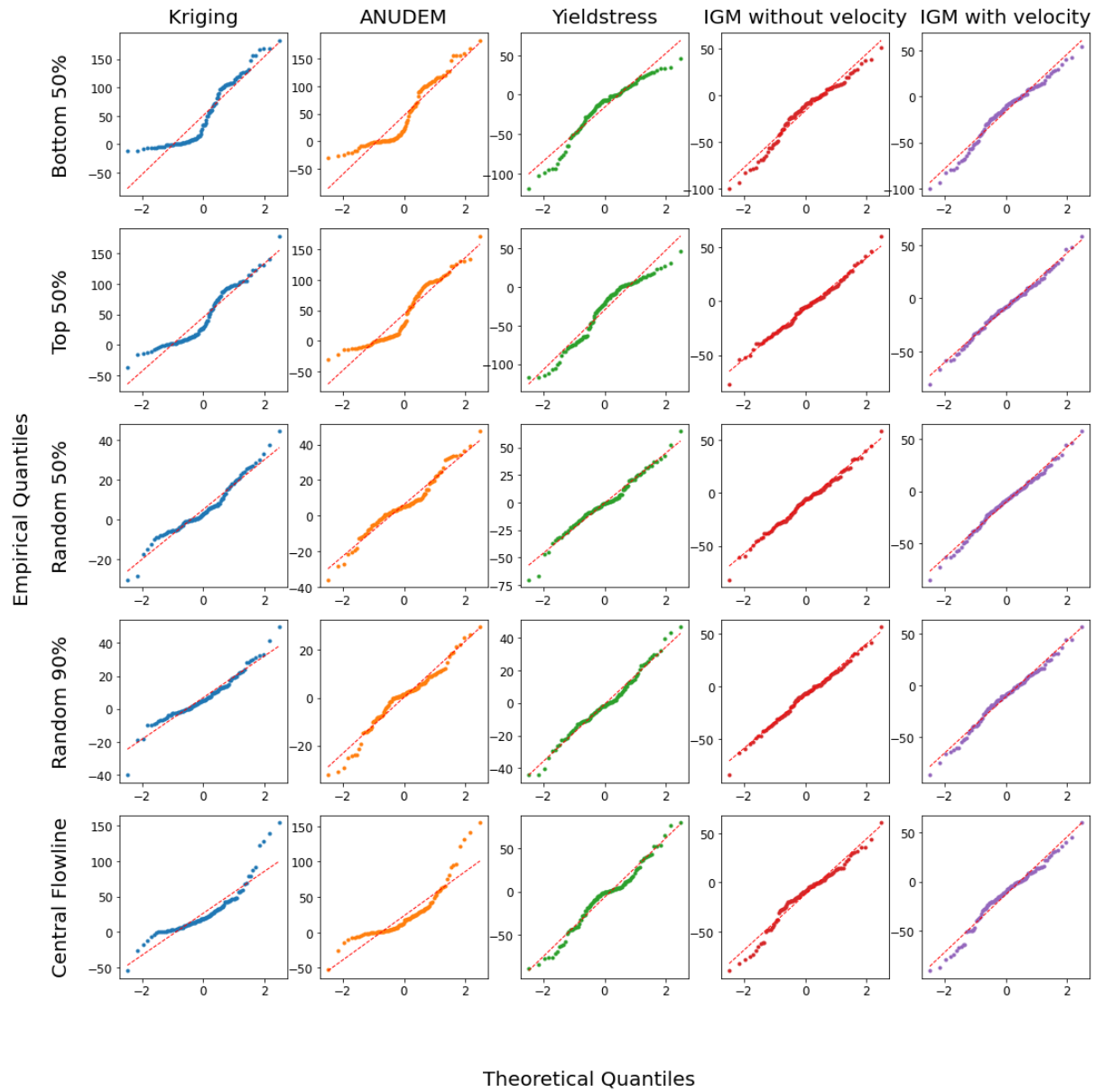


Figure 3.4: QQ-plots of the residuals. On the horizontal axis, the theoretical quantiles of a normal distribution are placed. The quantiles of the observed distribution are plotted on the vertical axis. The dashed red line depicts the -1:1 line, representing the normal distribution.

tails is characteristic of a leptokurtic distribution. On the right side of the QQ-plots, the residuals approximately follow the expected normal trend. In contrast, the residuals from the yield stress method and the IGM-based interpolations align closely with the theoretical normal distribution across the entire range. However, these graphs show a small platykurtic behaviour, meaning they have lighter tails than a normal distribution. The data is more clustered around the centre, making the distribution appear sharper in the middle. It must be repeated that this effect is rather small. The pattern is slightly more pronounced for the yield stress method than for the IGM interpolations.

The QQ-plots for the spatial interpolations using randomly selected 50% and 90% subsets display a curved pattern in the left tail, indicating a deficiency in extreme negative residuals but an excess of moderate negative values compared to a Gaussian distribution. The right tail is slightly heavier than expected under normality. Overall, these deviations are notably smaller than those observed in interpolations based on the top or bottom 50% of the data. The difference between the random 50% and random 90% datasets is minimal. The yield stress method results in heavier tails when the data is randomly distributed and therefore approximate the normal distribution more closely. The IGM-based interpolations using these randomly distributed points continue to produce residual distributions that more closely approximate normality than those from spatial interpolation techniques. However, the improvement over the IGM results from the top or bottom 50% datasets is relatively minor.

Spatial interpolations based on measurements aligned with the central flowline exhibit pronounced deviations from normality, with very heavy tails and an underrepresentation of residuals near the mean value of zero meters. This effect is, to a much smaller extent, also visible in the yield stress method based interpolation. The IGM-based interpolations on the other side, create lighter tails when they are fed with the data points along the central flowline.

The average RMSE of the complete interpolations are compared in the bar graphs of figure 3.5.

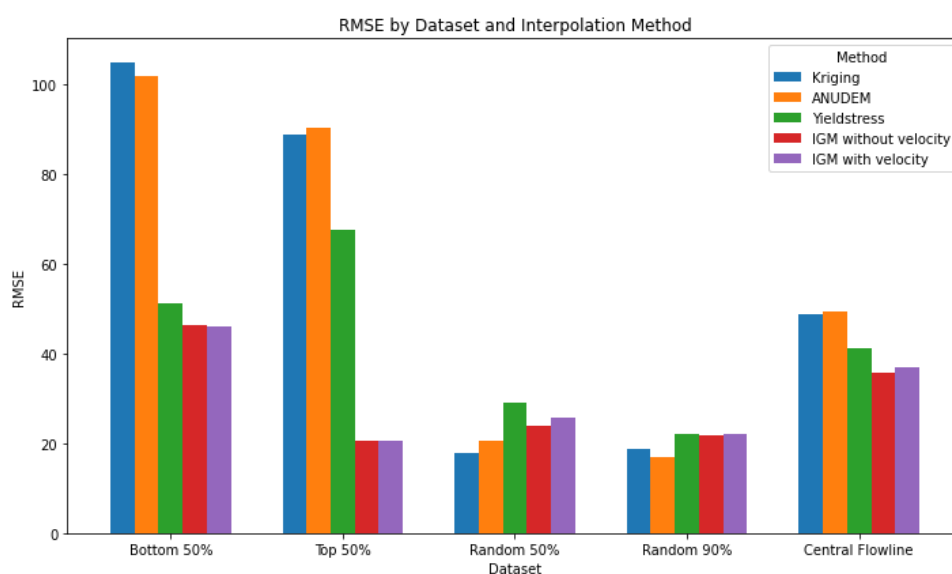


Figure 3.5: RMSE of the interpolation techniques

For the interpolations based on the bottom and top 50 % observations, the IGM interpolations have clearly the lowest RMSE, the same is true for the interpolations based on the observations near the central flowline. The yield stress method results in a similar RMSE as the IGM interpolations for the bottom 50% and the central flowline. For the top 50%, the yield stress method has a higher RMSE than the IGM interpolations, although it is lower than the RMSE of the spatial interpolation techniques. For the random 50 and 90 %, the RMSE is slightly higher for the IGM interpolations than the RMSE of the spatial interpolation methods. The method based on the yield stress method has an RMSE which is comparable to that of the IGM interpolations when random 90% of the observations is used, but when 50% or the complete dataset with RES measurements is used, the yield stress method has a slightly higher RMSE than IGM. For the interpolations based on the RES-measurements near the central flowline, the interpolations methods result in higher RMSE-values than the randomly distributed datasets. For these interpolations, the physics-based methods outperform the spatial interpolation methods again. The IGM interpolation with the Milan velocity field as input has in every case a very similar RMSE as the interpolation without the initial velocity field. The difference in RMSE between both IGM-interpolations is negligibly small.

In order to have an idea about the spatial autocorrelation between the residuals, the Moran's I of the residuals was calculated and depicted in figure 3.6.

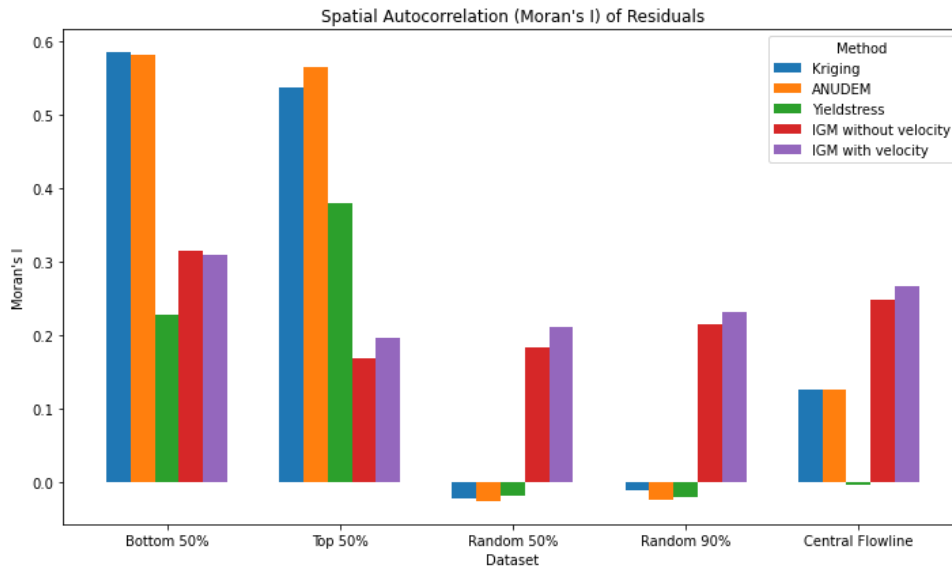


Figure 3.6: Moran's I of the residuals

The interpolations based on the bottom and upper 50%, the spatial interpolation methods have the largest Moran's I, with values above 0.5, indicating residuals with similar values are found in the vicinity of each other. That points at some spatial patterns that are left unexplained by the interpolation. For the bottom 50%, the yieldstress method produces the residuals with the least spatial autocorrelation. For the upper 50% interpolations, IGM produces residuals with the lowest Moran's I values. The interpolations based on the randomly distributed subsets are characterised by very low (even slightly negative) Moran's I values for kriging, ANUDEM and the yieldstress method. The IGM-interpolations have significantly higher Moran's I values than the other methods. However, the values are very similar to the Moran's I values of the IGM-interpolations with the top and bottom 50%, although the

bottom 50% has slightly higher values. The interpolations based on the points along the central flowline have a lower spatial autocorrelation for the spatial interpolations than for the IGM-interpolations, which have similar Moran's I values as in the other interpolations. The yieldstress method clearly outperforms the other methods for this subset.

To quantify the normality of the residuals, the Kolmogorov-Smirnov test statistic is calculated and presented in figure 3.7.

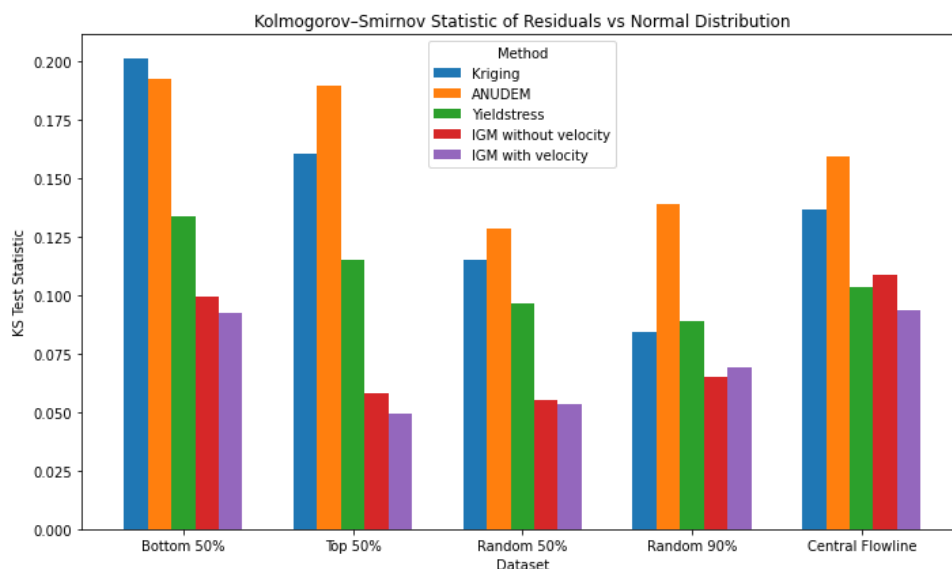


Figure 3.7: Kolmogorov-Smirnov test statistics of the interpolation techniques. Lower values mean more normally-distributed residuals.

For all datasets, the IGM-interpolations have the lowest KS test statistic compared to the spatial interpolation techniques, indicating that the distributions of these residuals are closer to a normal distribution than the residuals of the spatial interpolations. The yield stress method has intermediate values for most datasets, except for the central flowline dataset, where it has similar values as the IGM-interpolations. For the 90% dataset, the KS value of the kriging-interpolation is in the same order of magnitude as that of the yieldstress method.

Sensitivity analysis

IGM was assessed for both feature importance and permutation importance. These results are visualized in Figure 3.8, where the green bars represent feature importance and the yellow bars indicate permutation importance.

For IGM, the opti-smooth-anisotropy-factor is highlighted as most influential. However, due to the previously discussed chessboard effect, this parameter was fixed at 1000 for all interpolations. Aside from this, the opti-thkobs-std is identified as the next most important factor. Its permutation importance is slightly exceeded by that of opti-regu-params-thk. In contrast, iflo-init-slidingco and opti-convexity-weight contribute very little.

To further understand the relationships between parameters and model outputs, partial dependence plots were generated for RMSE and Laplacian values. The partial dependence plots of the IGM-parameters to the RMSE are shown in figure 3.9.

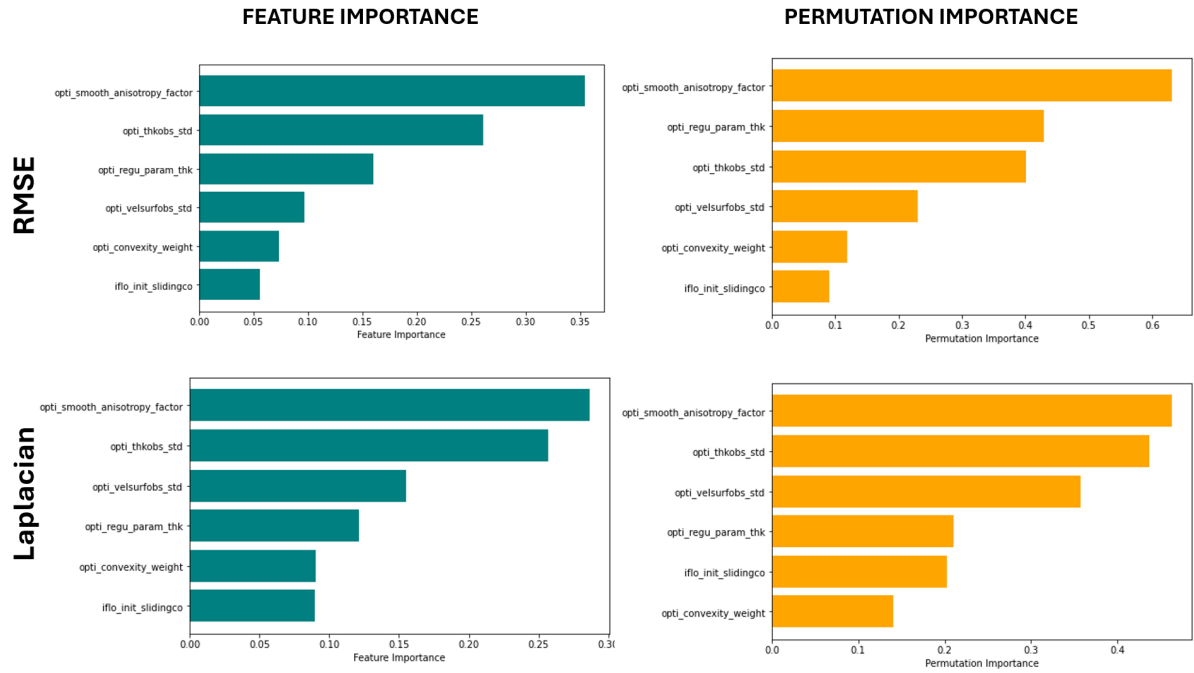


Figure 3.8: Parameter importance of the IGM parameters. The green barplots represent the feature importance for the RMSE (top) and the Laplacian (bottom). The yellow barplots represent the permutation importance for the same two indicators.

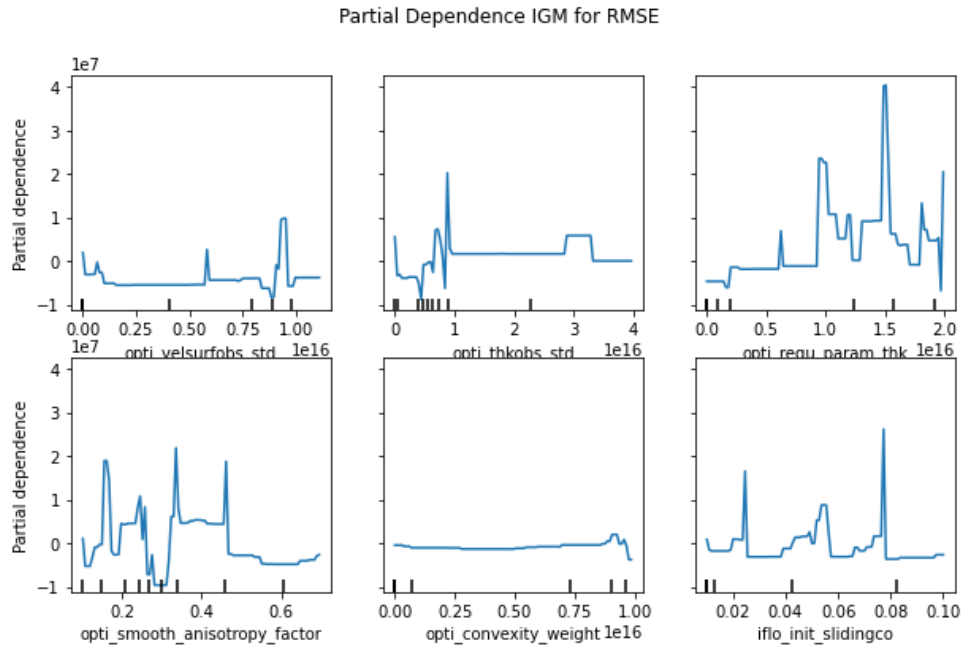


Figure 3.9: Partial dependence plots of the IGM-parameters for the RMSE. Lower values correspond to better model performance (see methods section).

opti-velsurfobs-std shows a relatively flat dependence for most of its range, but with noticeable fluctuations at higher values, suggesting the parameter slightly affects RMSE at extreme values. *opti-thkobs-std* exhibits spikes in RMSE at specific intervals, indicating that certain ranges of this parameter strongly influence performance. *opti-regu-param-thk* has a complex, non-linear impact on RMSE, with several peaks suggesting instability or sensitivity around specific values. *opti-smooth-anisotropy-factor* demonstrates multiple sharp increases and decreases, implying that careful tuning of this parameter is necessary. Also *iflo-init-slidingco* shows some localized peaks and valleys, suggesting that RMSE is sensitive to variations of this parameter in certain ranges.

opti-convexity-weight shows a nearly flat profile, indicating little to no effect on RMSE across its range. However, many parameter values were observed to result in a chessboard pattern. This is likely the result of numerical instability. A personal communication with the developers of the model confirmed this as to be a known artifact that is currently under investigation. Therefore, the value was set to $1e16$, independent on the results of the optimization algorithm.

To assess how parameters influence the smoothness of the interpolations, partial dependence plots were also created for the Laplacian. Figure 3.10 contains the partial dependence plots of the IGM-parameters for the Laplacian.

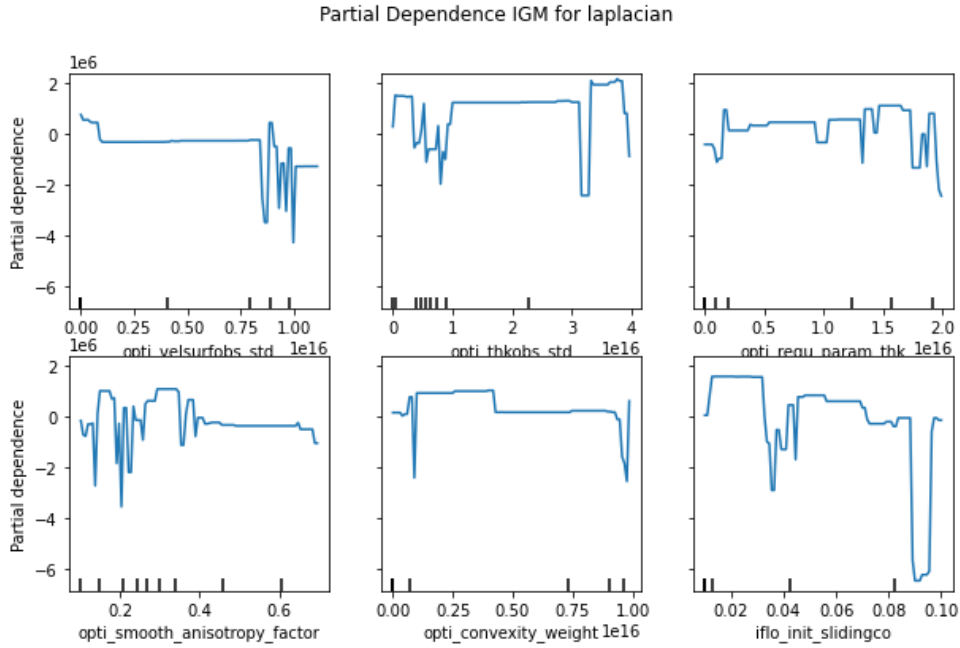


Figure 3.10: Partial dependence plots of the IGM-parameters for the Laplacian

opti-velsurfobs-std and *opti-thkobs-std* display sharp changes in Laplacian for certain parameter values, suggesting that these parameters have strong influence on surface smoothness. *opti-regu-param-thk* and *iflo-init-slidingco* show multiple non-linear jumps, indicating sensitive regions where small parameter changes can significantly affect smoothness. *opti-smooth-anisotropy-factor* strongly affects the Laplacian, which aligns with its role in controlling anisotropic smoothing in the model. *opti-convexity-weight*, while relatively flat for RMSE,

shows noticeable variations for Laplacian, meaning this parameter influences smoothness more than accuracy. As discussed before, this is the effect of the chessboard pattern.

Figure 3.11 shows the SHAP-summary plot for the IGM interpolations.



Figure 3.11: Summary plot SHAP values for the IGM interpolations. The vertical axis ranks the parameters by importance, from most to least influential. In the plot, low parameter values are coloured blue, and high values are coloured red. A negative SHAP value indicates that the parameter value pushes the prediction lower; a positive SHAP value means it pushes the prediction higher.

The standard deviation of the thickness observations is the most important parameter, which is logical given its direct connection to the interpolation uncertainty. High standard deviation values have a strong influence, pushing predictions both positively and negatively, while low standard deviations result in SHAP values close to zero. The standard deviation of the surface velocity observations shows an opposite pattern: low standard deviations tend to lead to positive SHAP values, while high standard deviations are associated with negative SHAP values. The regularization term for the ice thickness exhibits a similar pattern to the thickness standard deviation: higher parameter values are associated with a stronger influence (both positive and negative) on the predictions. In contrast, the anisotropy factor does not display a clear trend, as both high and low values appear across the SHAP value range. A low initial sliding coefficient is generally associated with more positive SHAP values, whereas higher sliding coefficients tend to result in negative SHAP values. Finally, for the convexity weight, low values yield SHAP values around zero, while higher convexity weights result in a broader spread toward both positive and negative SHAP values.

As for IGM, the sensitivity analysis could be performed on ANUDEM and Kriging as well, since for these methods some parameters were optimized as well via the PSO. Since this thesis has as goal to compare IGM with established methods, performing a sensitivity analysis on the depreciated spatial interpolation methods would be outside the scope of the research. For completeness, the figures of the parameter importance, the partial dependence plots for both the RMSE and the Laplacian, as well as the summary plots for the SHAP values are placed in the appendix. Although they were used during the optimization and for some experiments that were not implemented in the thesis, they will not be discussed further.

For the yield stress method, no sensitivity analysis was performed. However, having a look at the distribution of the interpolated yield stresses can be instructive. Therefore, figure 3.12 depicts the interpolated yield stresses along the central flowline for each subset of the data. The orthophoto created from the UAV images is shown at the background.

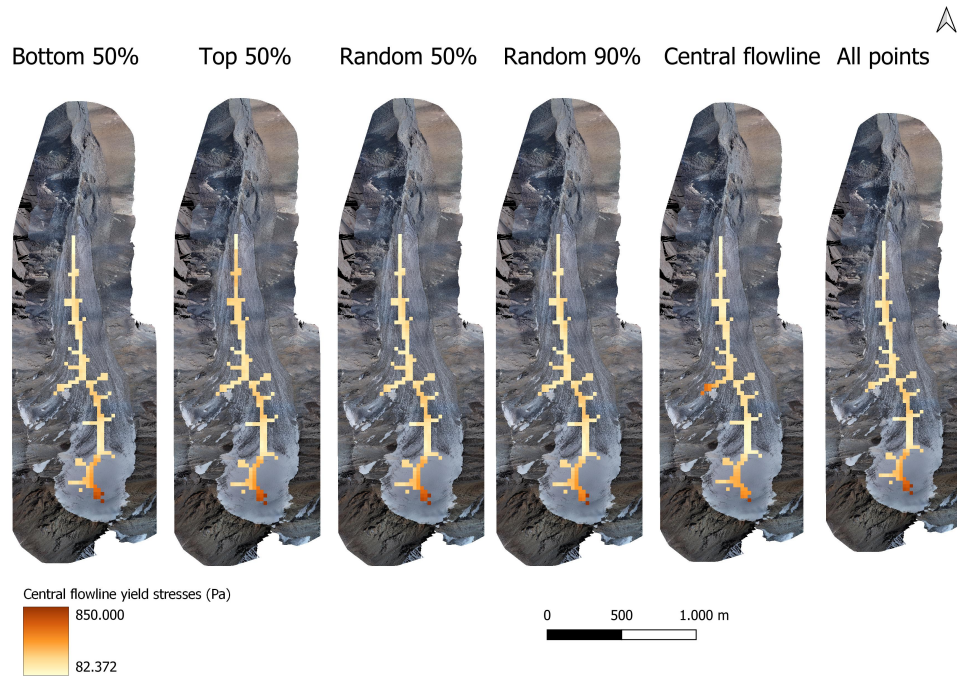


Figure 3.12: Maps of the interpolated yield stresses along the central flowlines of the different subsets

The differences between the different datasets are rather small. For all observations, the yield stresses in the upper part (where no RES measurements are performed) increase significantly. Especially the last pixels are high. Note that these pixels are extrapolated values and not interpolated. For the interpolation based on the central flowline, also the points in the side branch have a relatively high yieldstress. This does not surprise, since the points on the central flowline are located relatively far from this side branch. Therefore, these points require relatively many extrapolation as well. Figure 3.13 provides boxplots of the interpolated yield stresses along the central flowline for each subset of data.

Overall, the distributions of the interpolated yield stresses are relatively equal to each other. The means vary between 23.9 MPa (central flowline) and 30.0 MPa (random 90%). The interquartile ranges as well are very similar. Except for the dataset based on the bottom 50%, also the whiskers are relatively similar. What stands out is the positive outliers in all distributions.

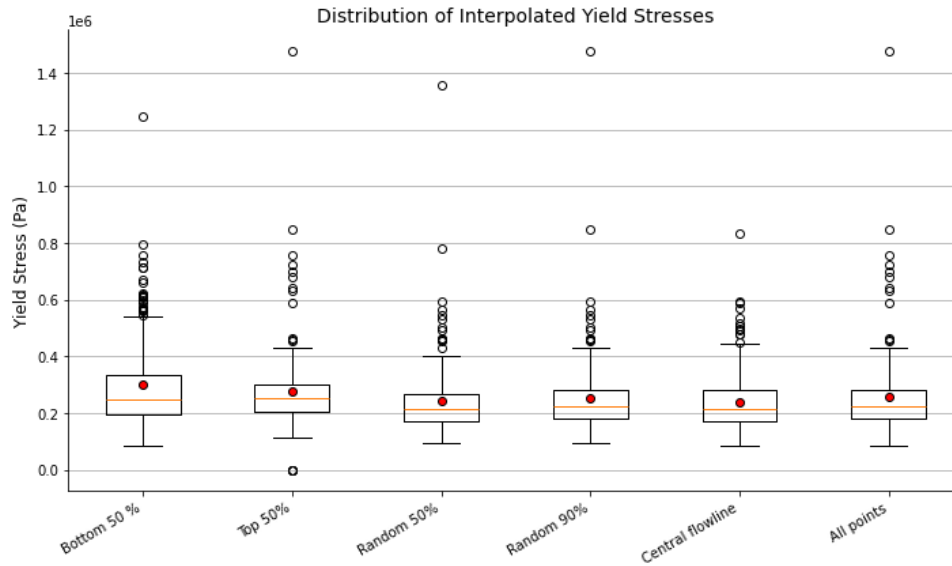


Figure 3.13: Boxplots of the interpolated yield stresses along the central flowlines of the different subsets

3.2 Mass balance

The calibration of the mass balance model is done by comparing the observed and the modelled SMB for different values of C_0 and C_1 . The first measure compared is the stake measurements. In figure 3.14, a heatmap of the mean RMSE between the modelled and measured SMB at the stake locations is presented.

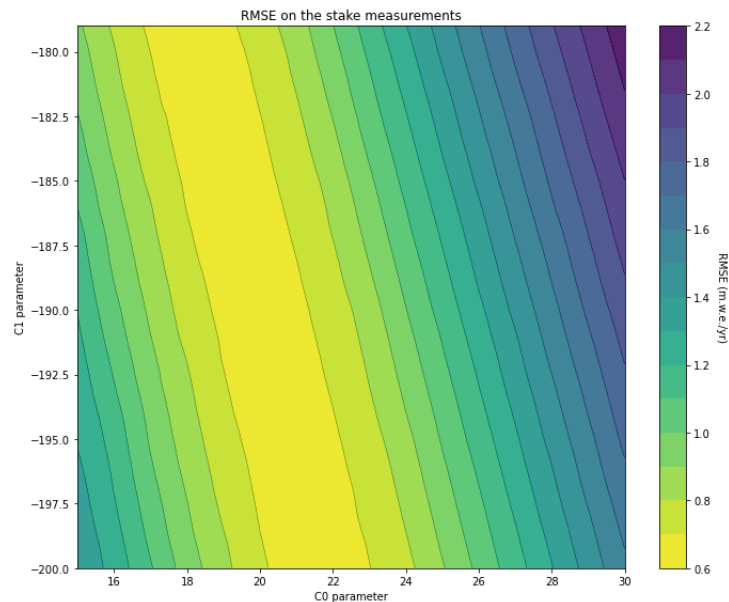


Figure 3.14: Heatmap of specific mass balance RMSE compared to point stake observations

The RMSE has an area of minimal value in the tested ranges of C_0 and C_1 . A relation

exists between C0 and C1 for the same RMSE. Since it is not only the local value of the mass balance that is important, also the value of the complete glacier, so the mean specific mass balance, is important to reproduce. In figures 3.15 and 3.16, the same heatmap of the RMSE is presented, but now the modelled mass balance is compared with the geodetic mean specific mass balance from Hugonnet et al. (2021). This geodetic mean SMB is available for the periods 2000-2010 and 2010-2020. Figure 3.15 shows the heatmap for the period between 2000-2010, figure 3.16 for the period 2010-2020 and figure 3.17 for the period 2000-2020.

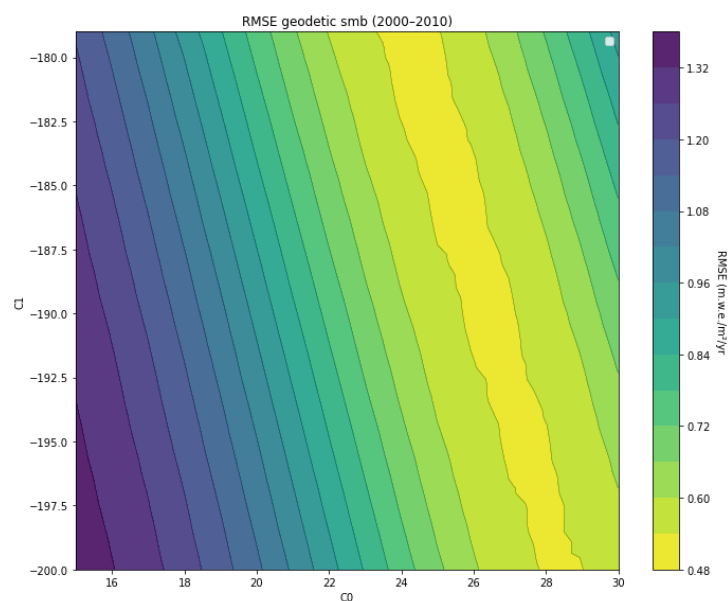


Figure 3.15: Heatmap of mean specific mass balance RMSE compared to the geodetic mean specific mass balance for the period 2000-2010

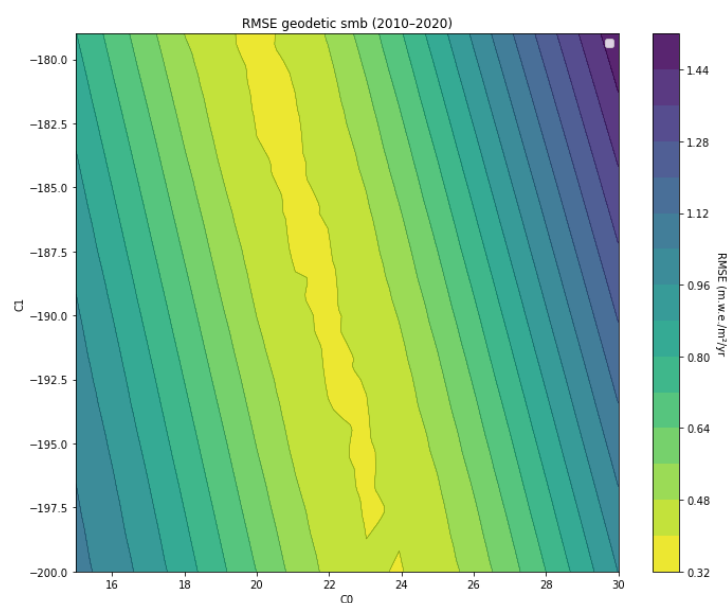


Figure 3.16: Heatmap of mean specific mass balance RMSE compared to the geodetic mean specific mass balance for the period 2010-2020

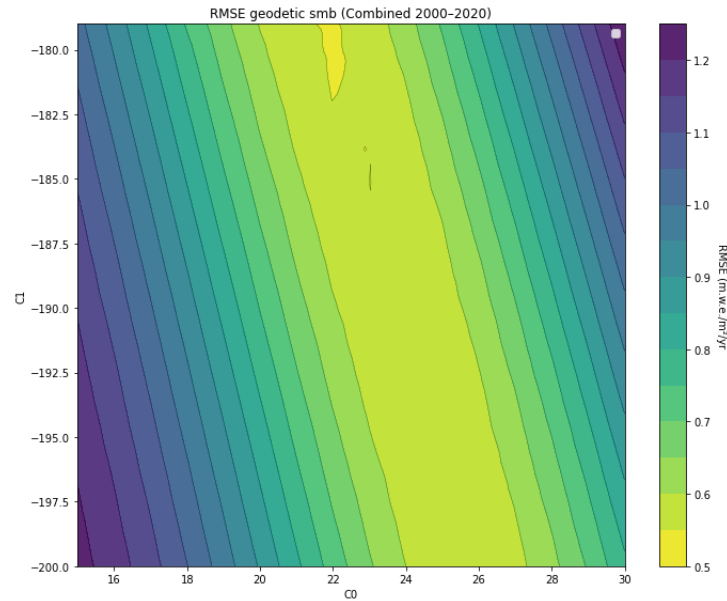


Figure 3.17: Heatmap of mean specific mass balance RMSE compared to the geodetic mean specific mass balance for the period 2000-2020

The comparison with the geodetic mass balance also shows a strip of optimal combinations of parameters. This optimal band is slightly different for both periods. Besides reproducing both the mean specific mass balance and the average specific mass balance values, as well as the reproducing the mass balance pattern over the elevation range of the glacier is important. In figure 3.18, the SMB observations are binned per 100 meters and plotted together with the corresponding modelled results for $C0=18$ and $C1=-179$, which below is evaluated to be the optimal combination.

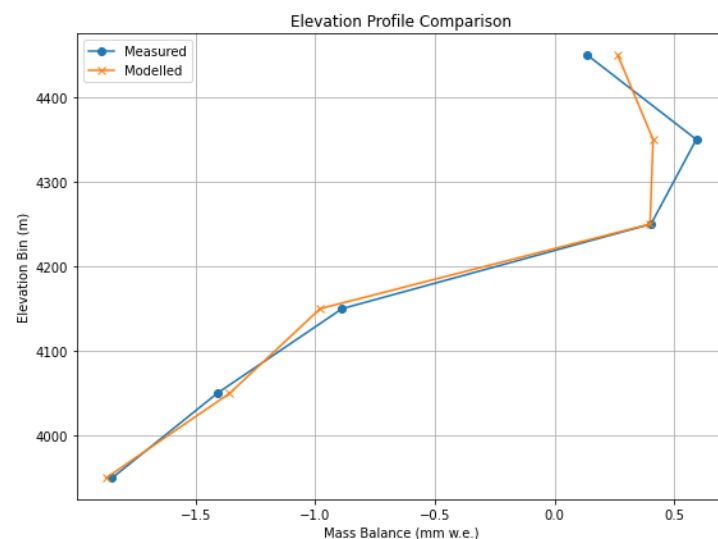


Figure 3.18: Profile of the multi-year mean modelled and measured SMB for $C0=18$, $C1=-179$

The mass balance profiles over the elevation of the glacier matches relatively well with

the observations. A similar profile can be made for each combination of C_0 and C_1 . The RMSE can also be calculated based on this elevation-binned difference. The heatmap of this elevation-binned RMSE is presented in figure 3.19.

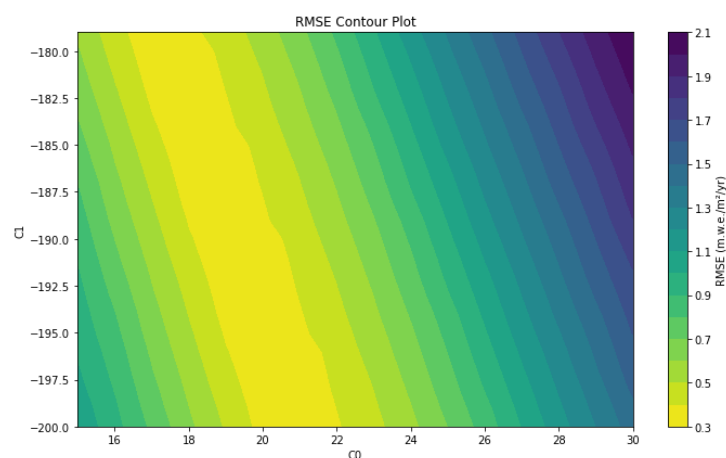


Figure 3.19: Heatmap of specific mass balance RMSE based on the elevation profiles of mass balances

To make a final decision on the optimal values of C_0 and C_1 , the ϵ -constraint method is used. Figure 3.20 contains the plots of the RMSE of the different SMB measures against the RMSE of the stake measurements.

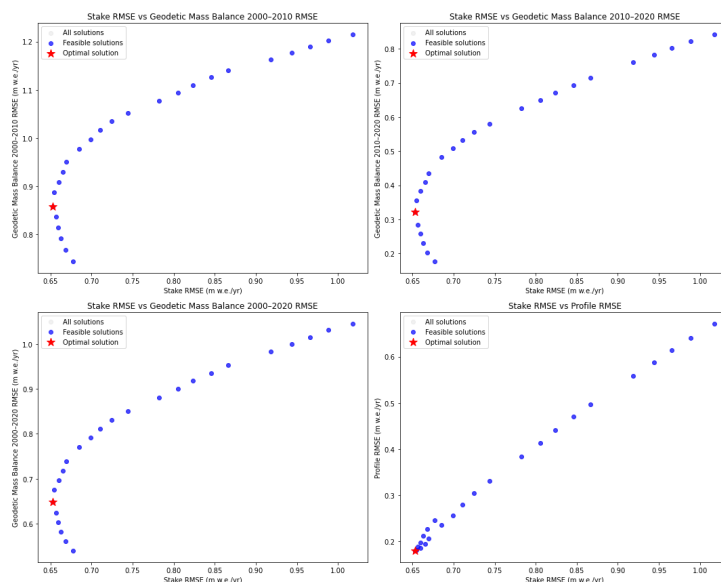


Figure 3.20: Heatmap of specific mass balance RMSE based on the elevation profiles of mass balances

The chosen combination of C_0 and C_1 performs relatively well for the multi-year mean values of the mean specific mass balance, the point-based RMSE and the profile-based RMSE. However, it is important to also look at annual performances. Figure 3.21 shows a comparison between the measured and the modelled mean specific mass balance between 2011 and 2024.

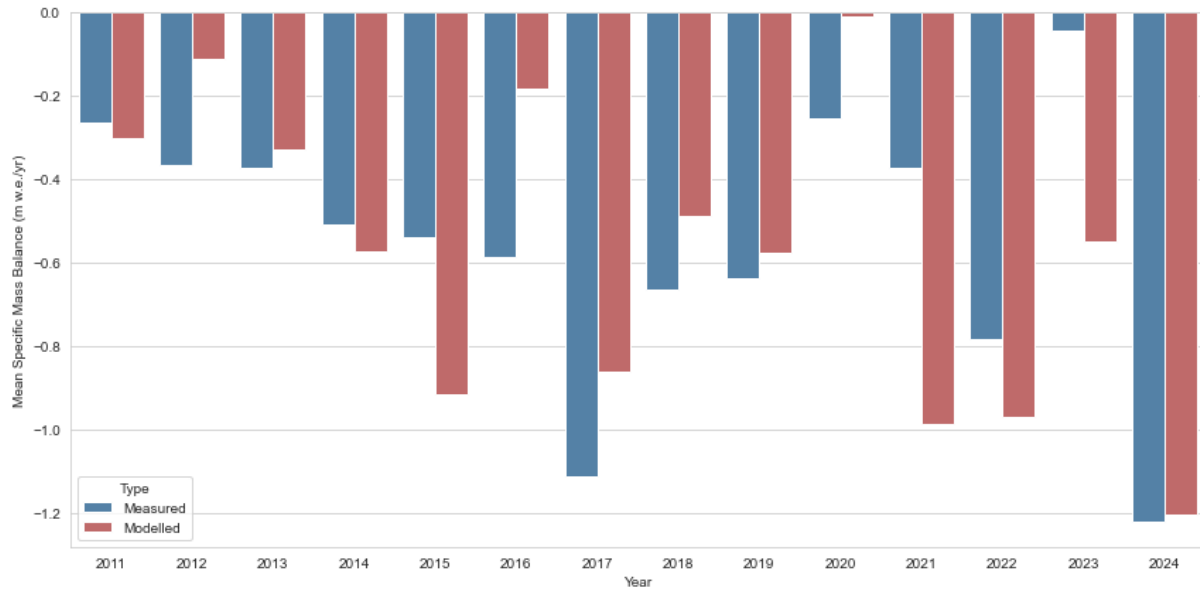


Figure 3.21: Comparison of the modelled and measured mean specific mass balance

A large number of mass balance years have a strong agreement between the modelled and the measured values. However, there are years in which the model significantly underestimates the melting (2011-2012, 2015-2016, 2016-2017, 2019-2020) and there are some years where the model overestimates the melting (2014-2015, 2020-2021, 2022-2023). Over- or underestimating does not seem to depend on the time or on the magnitude of the SMB, since they are spread randomly across the years and the magnitudes. Figure 3.22 shows the measured and the modelled cumulative mean specific mass balance.

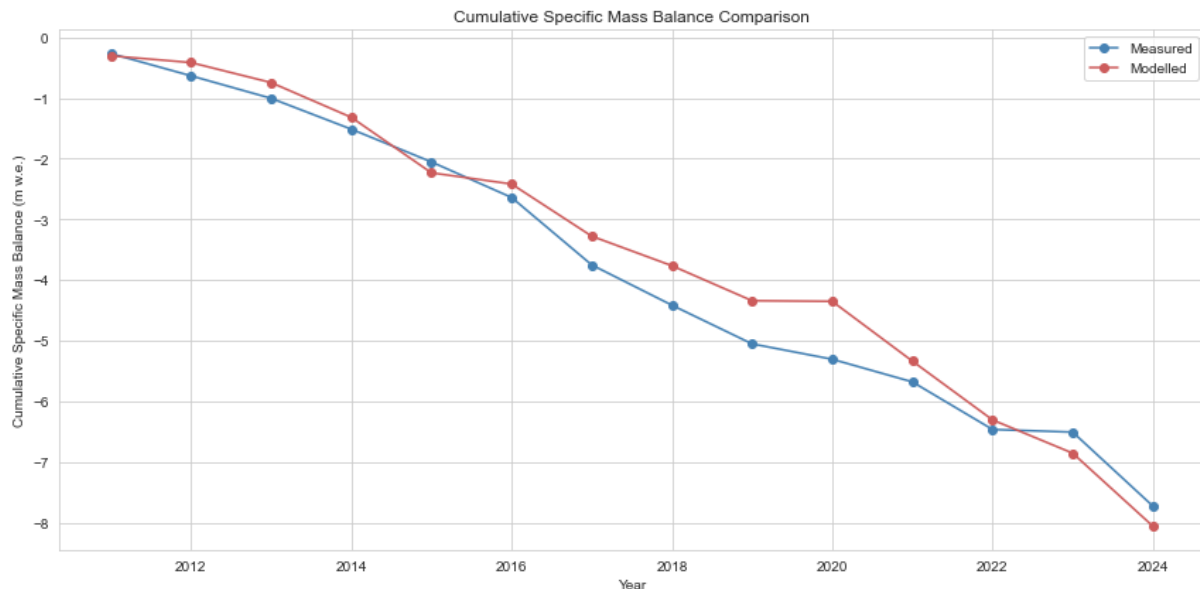


Figure 3.22: Comparison of the modelled and measured cumulative mean specific mass balance

The modelled cumulative specific mass balance matches the measured one relatively well.

Only in 2020, the error exceeds 1 m.w.e.. Figure 3.23 depicts a comparison of the measured and modelled mass balance profiles for all years separate.

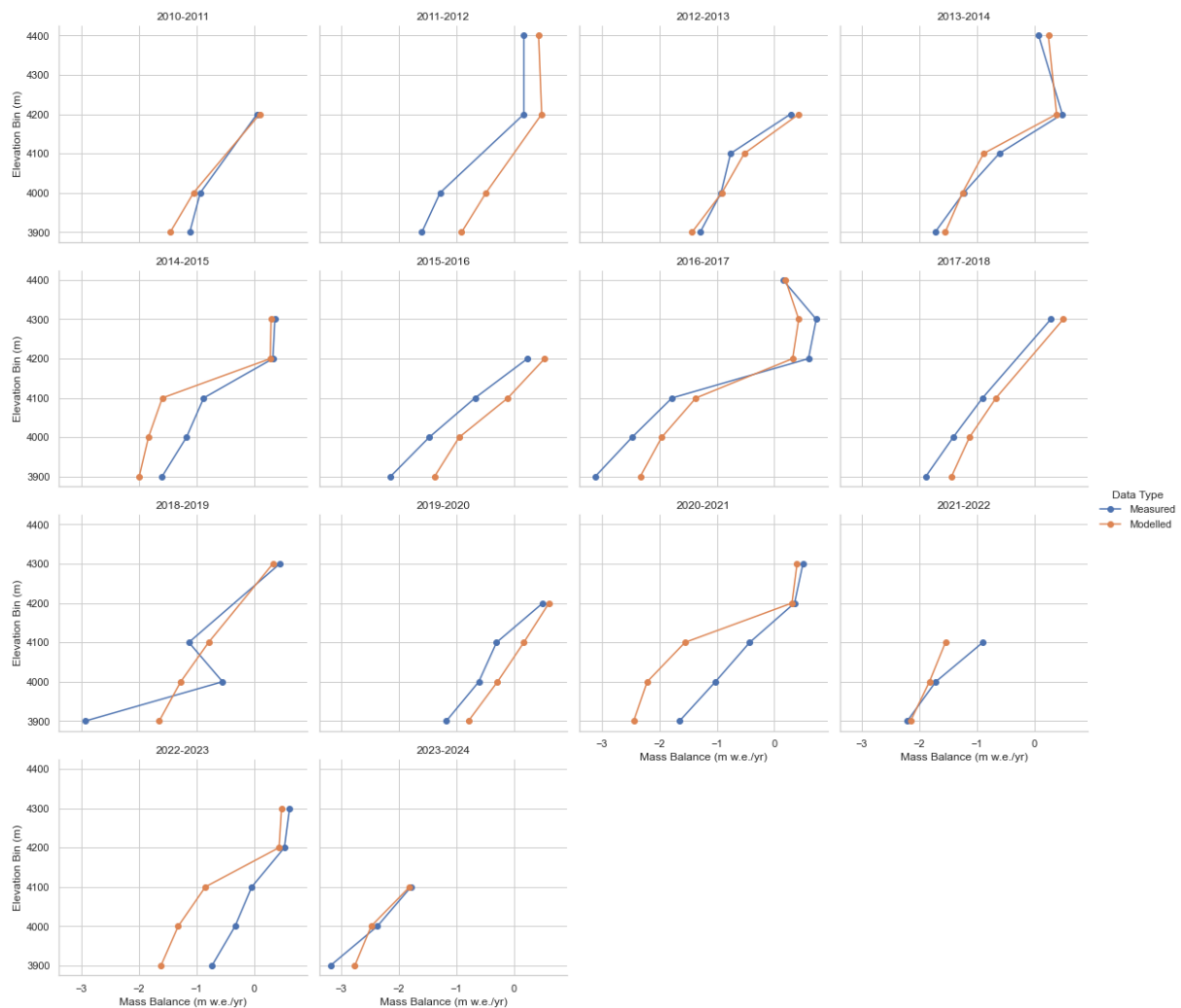


Figure 3.23: Comparison of the yearly profiles of the measured and modelled SMB

The years that were identified to over- or underestimate the mean specific mass balance, can easily be identified in figure 3.23. For the overestimations, the modelled graph is at all elevations under the measured graph (also in the accumulation area). The underestimations show exactly the opposite. Overall, the modelled profiles follow the shape of the measured profiles relatively well. Since over- and underestimations alternate, the mass balance model can be considered to be sufficient.

3.3 Glacier evolution

Calibration of the flow model

During the dynamic calibration, the modelled length along the flowline is compared with the length of the glacier observed in the Landsat images. Both temperature and precipitation are adjusted during the calibration process in order to make the measured length. This mass balance adjustments are visualised in figure 3.24.

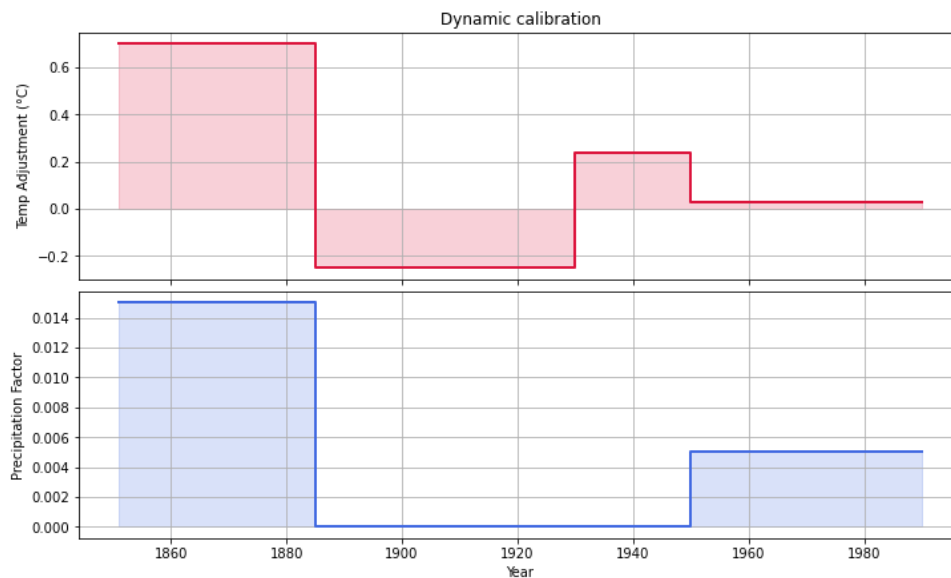


Figure 3.24: Dynamic calibration: The temperature is adjusted by adding a certain amount of °C. The precipitation is summed with a certain fraction of the original precipitation.

The adaptations as depicted in 3.24 are applied before 1900. After 1900, no adjustments have been made to the climatic input, since this information is based on the measurements of a meteorological station, for which the uncertainty is lower.

The length of the glacier is shown in figure 3.25. It is important to note that the HOM is calibrated in order to match this length. IGM was provided exactly the same mass balance as the one calibrated for HOM.

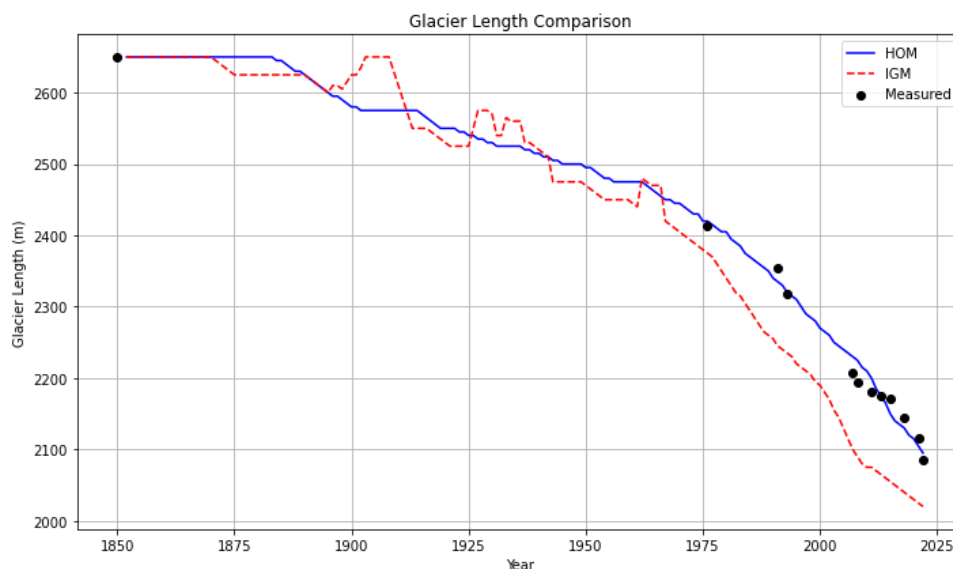


Figure 3.25: Glacier length modelled by HOM and by IGM compared with the glacier length measured on Landsat Images.

Between 1850 and 1880, the HOM kept the length it has at the LIA. The profile shows a very straight line, indicating the glacier wants to grow outside the glacier mask, but the model does not allow that. Warming the climate in this period could help the glacier fit inside its boundaries. However, after a large number of tests, no satisfying warming could be reached. In the same period, IGM produces a similar field, but for a shorter period. In the period between 1880 and 1900, there is a rapid decrease of the glacier length modelled by HOM. This retreat is likely to be too fast, but it compensates for the stable period since 1850. The state of the glacier in 1900, which is the result of the two unnatural periods, is assumed to be good. During this period, IGM produces a more varying glacier length. However, the length never deviates more than 50 meters from the length modelled by HOM. No structural difference between both models is observed in this period: sometimes IGM is longer, sometimes HOM is longer. The retreat of the glacier since the 1970 concerning HOM is modelled relatively accurately based on the Landsat images. During this last period, IGM produces a glacier that is a little shorter than the then one produced by HOM, and the closely approximated measurements. The deviation between both models is relatively small and does not become larger over time.

To calibrate parameters E and ASL , a heatmap of the calculated RMSE compared with the RES-measurements in 2023 is presented in figure 3.26.

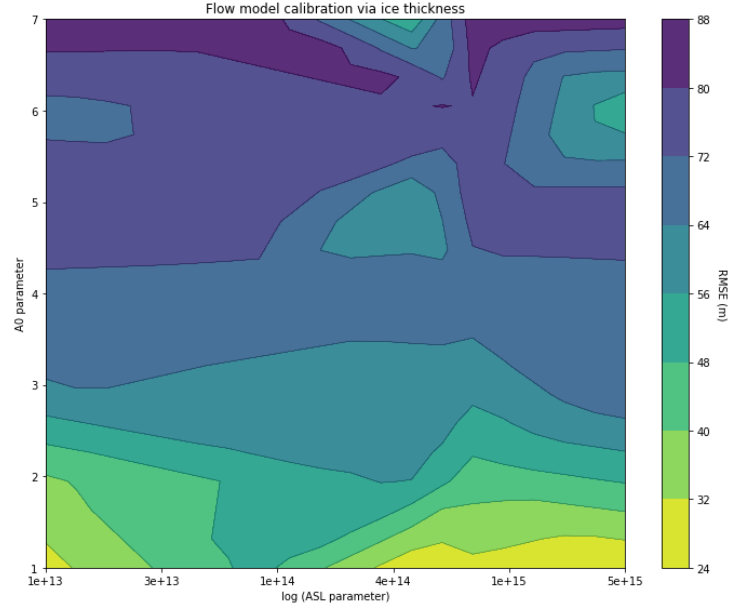


Figure 3.26: Heatmap of the ice thickness RMSE

The heatmap shows an erratic behaviour with not much structure in it. Overall, the accuracy increases with a declining enhancement factor (E). Another measure that is important to match is the velocity as measured by the crevasses. The RMSE of this velocity is depicted in figure 3.27.

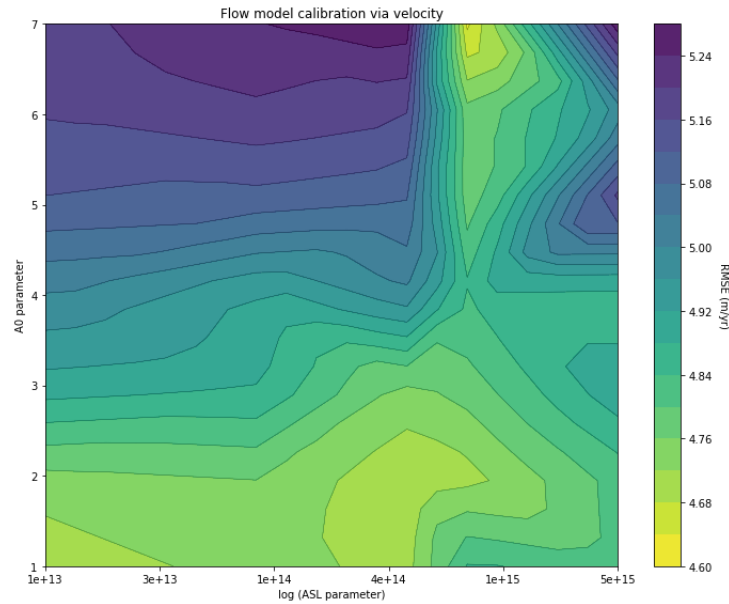


Figure 3.27: Heatmap of the ice thickness RMSE

For the velocity as well, there is an overall increase in accuracy for a decreasing E , except for ASL values around $1e+15$, where the accuracy increases with an increasing E . As done for the calibration of the mass balance model, the ϵ -constraint method is used to make a final decision of the values of E and ASL . This is depicted in figure 3.28.

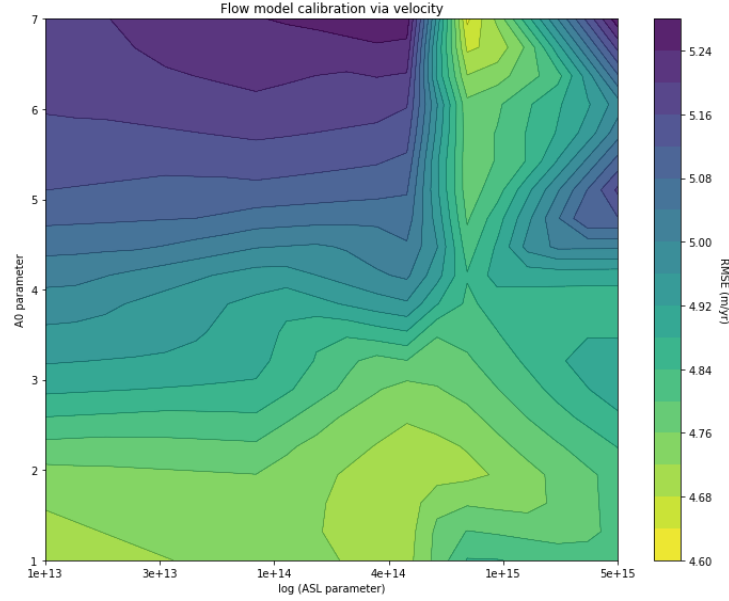


Figure 3.28: Visualisation of the ϵ -constraint method to determine the optimal values of A_0 and ASL .

The optimal values are $E = 1$ and $ASL = 5e15$. The scattered patterns of the individual heatmaps also pops up on this scatterplot. The scattered behaviour of the different parameters indicates no stable trend or systematic direction of improvement. Numerical instability might be a possible explanation. The results revealed a high degree of sensitivity, but no clear pattern or trend in RMSE, indicating that both parameters interact in a non-linear way with the glacier geometry and thermodynamics. Despite the physical expectations that higher E would generally increase internal deformation and that higher ASL would enhance basal sliding, the results showed multiple parameter combinations yielding comparable RMSE values. This suggests a counteracting effect between the two processes: increased sliding can offset reduced internal deformation and vice versa.

Evolution

The modelled evolution of the glacier under the different SSP's is described in figure 3.29. The black line represents the glacier area for the historic period, where climate data or proxy's were available. For the future period, each SSP has received another colour. The shaded areas represent the confidence interval.

Both the area and the volume decrease over almost the entire historical period. Around 1900, there a small and short during increase in both volume and area is predicted. This increase is more pronounced for the HOM. Between 1900 and 1970, the melt of the glacier is predicted to slow down. A similar historic evolution is found by Van Tricht and Huybrechts (2023) for other glaciers and ice caps in the Tien Shan. However, it is important to take into account the perturbations to the climate during the dynamic calibration. This might impact the historic evolution to some extent.

The area and volume decrease in the future for all SSPs for both models, with a decreasing speed from 2070 onwards. For the HOM-simulations, a relatively sharp change towards a more

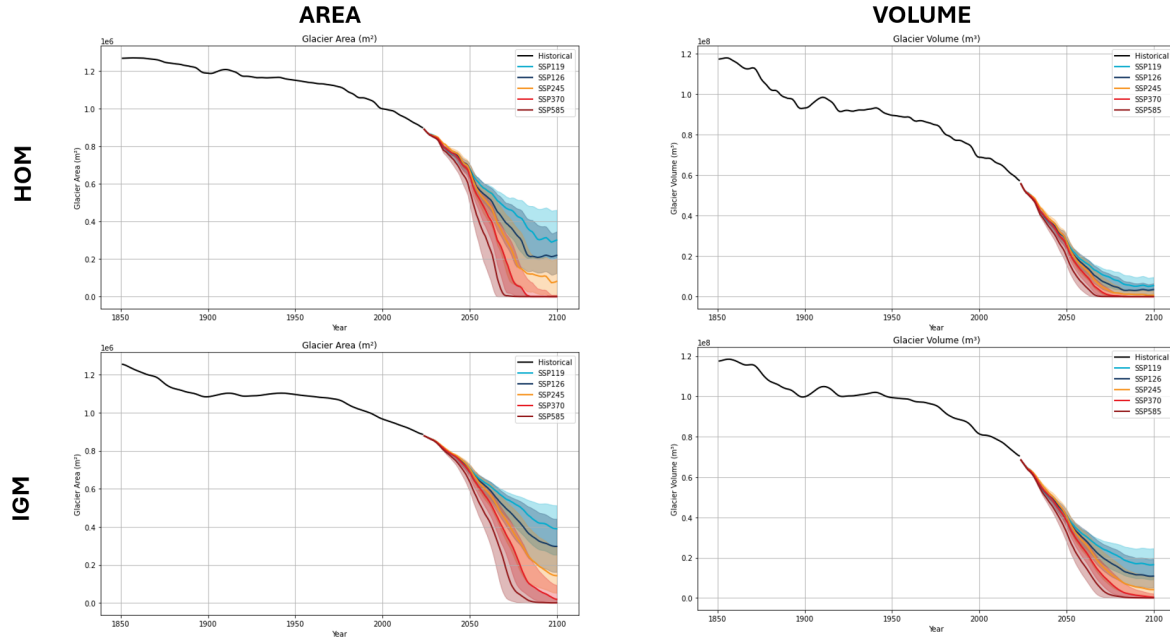


Figure 3.29: Comparison glacier area (left column) and volume (right column) evolution for HOM (top line) and IGM (bottom line). The black line represents the historic period. The coloured lines represent different SSPs, with the standard deviations shaded.

or less stable area occurs. The decrease in volume is more fluently.

For a better comparison of HOM and IGM, three SSPs are selected and the future period is visualised in figure 3.30.

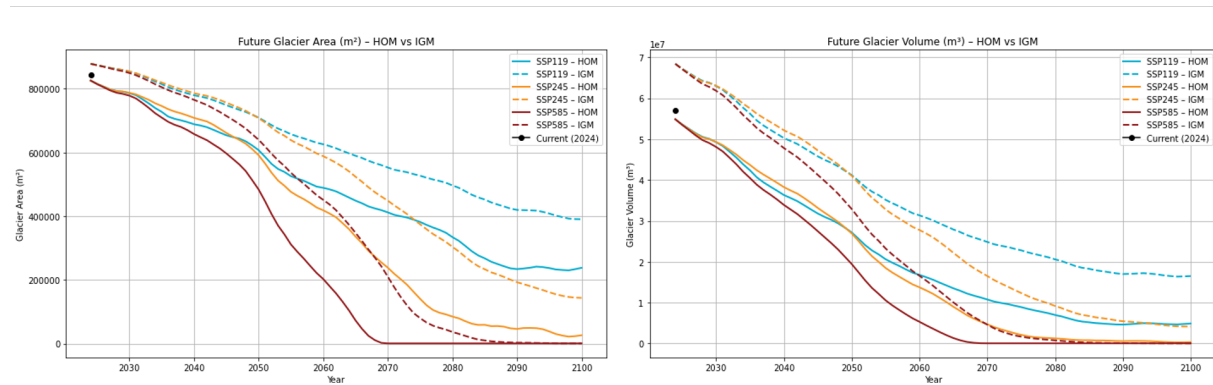


Figure 3.30: Future evolution of the area (left pannel) and the volume (right pannel) for both HOM (full lines) and IGM (dashed lines) for SSPs 1-1.9 (blue), 2-4.5 (orange) and 5-8.5 (red). The black dot represents the interpolation of the ice thickness from the first part of this thesis.

In 2024, the area modelled by HOM is slightly lower than the observed area, while IGM models the glacier slightly too large. The volume is slightly underestimated by HOM, but IGM

significantly overestimates the volume. While the area was modelled comparatively good for IGM as for HOM, IGM has modelled a too thick glacier in 2024.

The difference between both models in 2024 remains more or less constant over time. The lines for the same SSP are parallel, indicating a very similar evolution for the future is predicted. The time by which the glacier is modelled to dissapear, or the volume and area remaining in 2100 are summarized in table 3.1.

Table 3.1: Summary of projected glacier area and volume for each SSP and model (HOM, IGM). For scenarios where the glacier does not vanish, the remaining area and volume in 2100 are shown.

SSP	Model	Zero Year (Area)	Remaining Area (m ²)	Remaining Volume (m ³)
SSP1-1.9	HOM	Not by 2100	237,875	4,874,635
SSP1-2.6	HOM	Not by 2100	167,500	3,023,371
SSP2-4.5	HOM	Not by 2100	26,125	320,775
SSP3-7.0	HOM	2077	—	—
SSP5-8.5	HOM	2069	—	—
SSP1-1.9	IGM	Not by 2100	390,500	16,493,862
SSP1-2.6	IGM	Not by 2100	297,250	10,811,459
SSP2-4.5	IGM	Not by 2100	143,500	4,128,359
SSP3-7.0	IGM	Not by 2100	19,250	400,064
SSP5-8.5	IGM	2085	—	—

For SSP5-8.5, HOM predicts the glacier to disapear in 2069. Concerning IGM, no more ice will be available in 2085. HOM also predicts that there will be no more ice left under SSP3-7.0 from 2077 onwards, while IGM models that in 2100 a small patch of 19 000 m² will be left. The differences in remaining area and volume for all other SSPs are significant. However, figure 3.30 shows the cause of this difference is mainly caused by the historic period.

Figure 3.31 shows maps of the ice thickness at different moments in time for three different SSPs as predicted by HOM. Figure 3.32 shows exactly the same as figure 3.31 for the model results of IGM. The distribution of the ice thicknesses for both models is depicted in boxplots in figure 3.33.

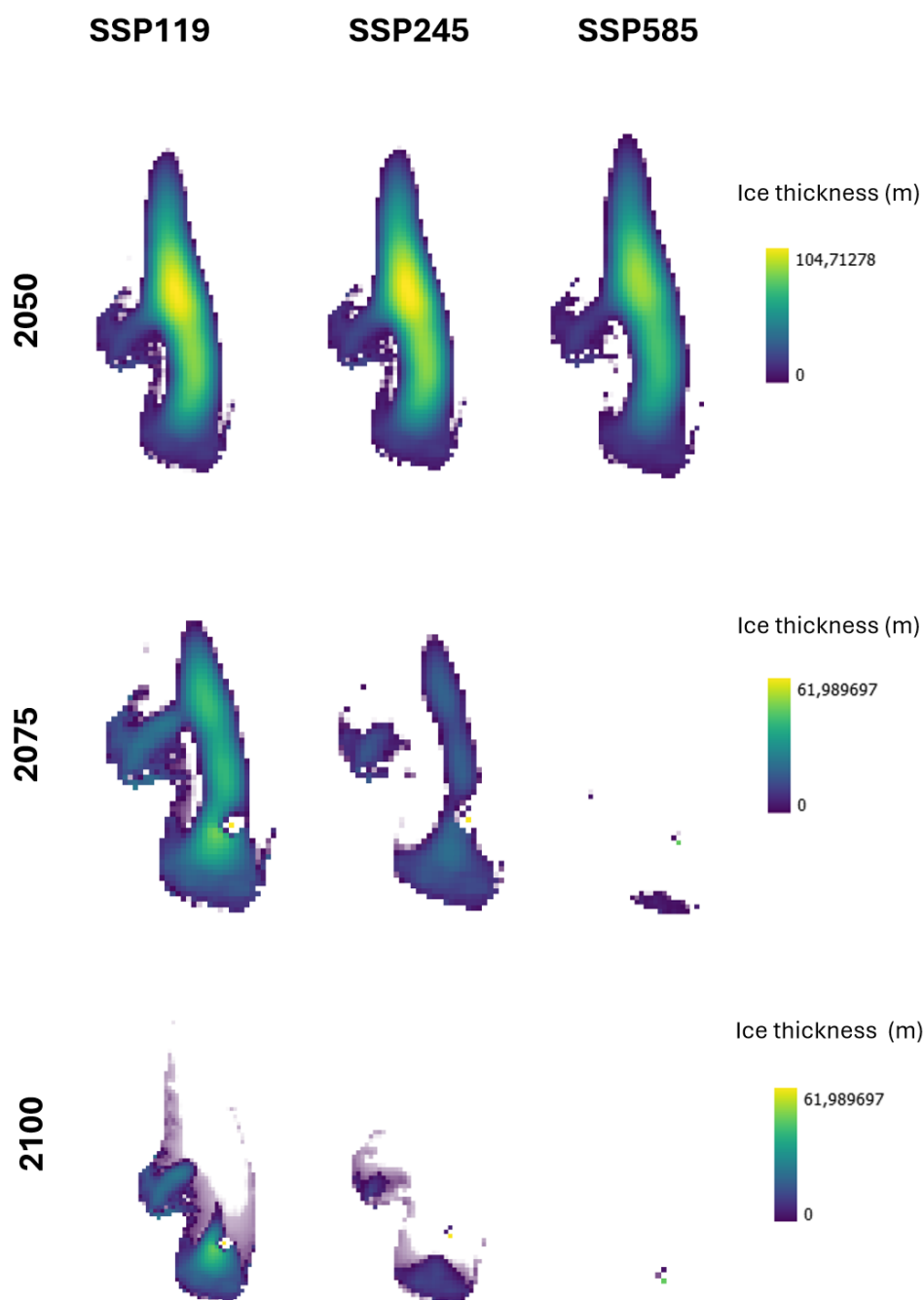


Figure 3.31: Ice thickness maps of the HOM simulations for SSP1-1.9, SSP2-4.5 and SSP5-8.5 in the years 2050, 2075 and 2100.

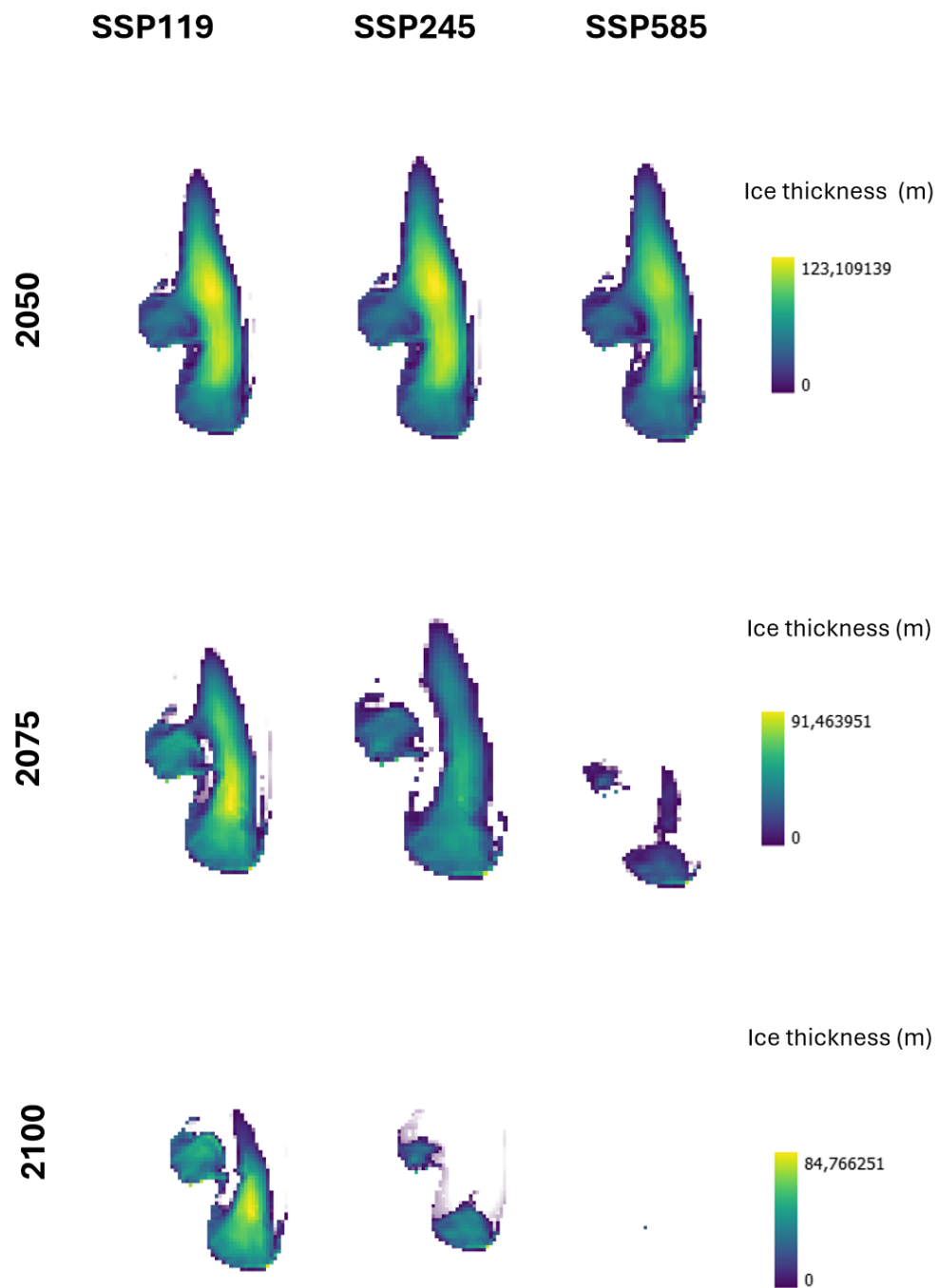


Figure 3.32: Ice thickness maps of the IGM simulations for SSP1-1.9, SSP2-4.5 and SSP5-8.5 in the years 2050, 2075 and 2100.

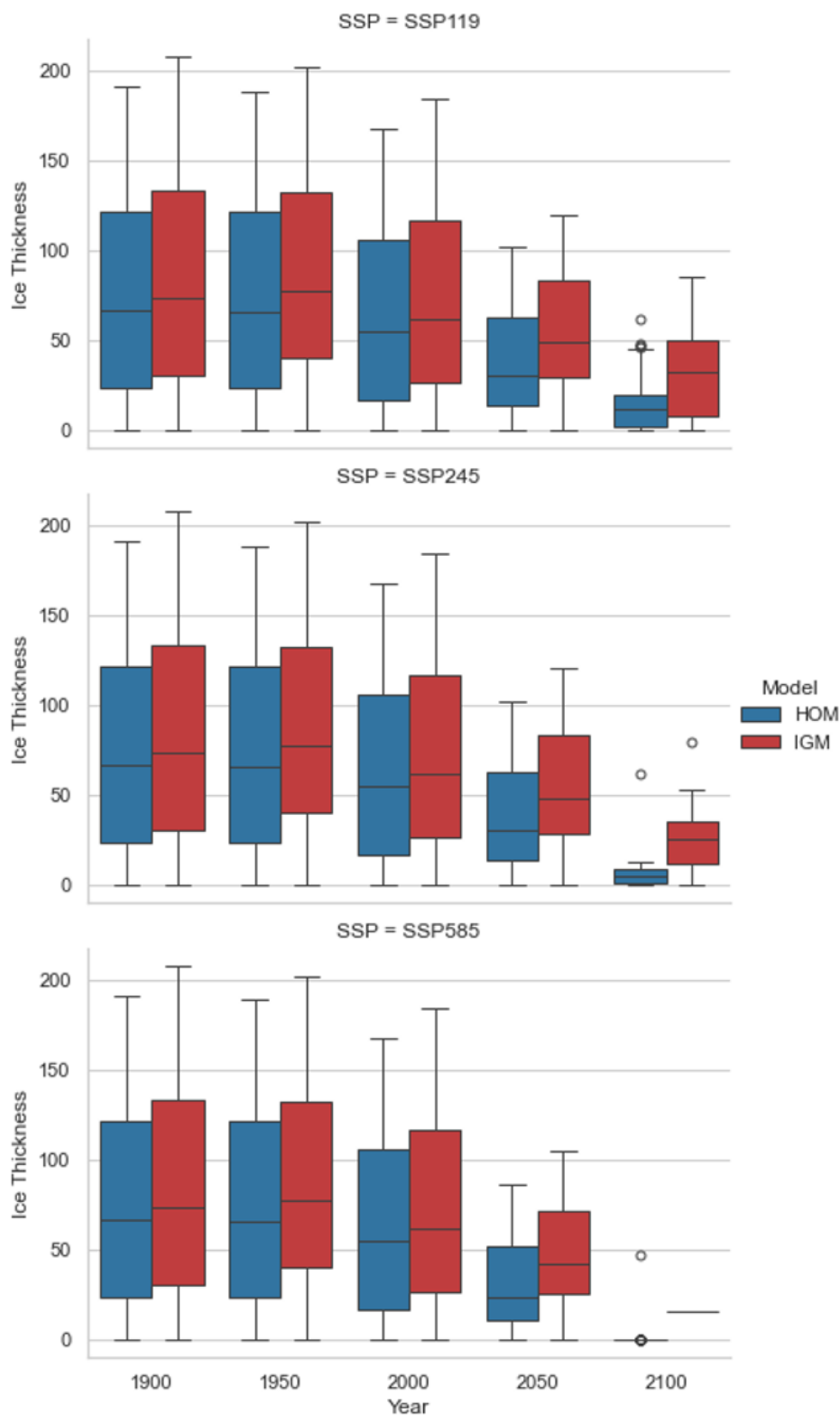


Figure 3.33: Boxplots of the ice thicknesses for SSP1-1.9 (upper panel), SSP2-4.5 (middle panel) and the SSP5-8.5 (lower panel). The depicted years are, from left to right, 1900, 1950, 2000, 2050 and 2100. The blue boxes represent the HOM result, the red boxes are the IGM-interpolations

For HOM, the side branch gets decoupled from the main glacier, except for SSP1-1.9 it happens before 2075 under all SSPs. Under SSP1-1.9 and SSP2-4.5, ice remains in both the main glacier and in the side branch. A remarkable feature occurs in 2075 and 2100 under SSP1-1.9. A hole occurs in the middle of the glacier. This is likely due to numerical instability. For the two other SSPs, the ice has been removed on this place before this hole could occur.

Although ice thickness are overall higher than modelled by HOM, the side branch still gets decoupled when IGM is used. Except for the magnitude of the ice thickness, the areas where ice is removed are relatively similar to HOM. This indicates IGM predicts a similar retreat as HOM, only delayed because of its slower melting in the historic period.

For all moments in time over all SSPs, the mean ice thickness calculated by IGM is higher than the mean ice thickness calculated by HOM. Except for the year 2100, the boxes have a very similar length, indicating a comparable variability for both models. Also the whiskers seem to have very similar lengths. This all points to very similar distributions of the ice thicknesses, except for the fact that IGM produces thicker ice.

In figure 3.34, a cross-section in the year 2023 is presented along the flowline presented before.

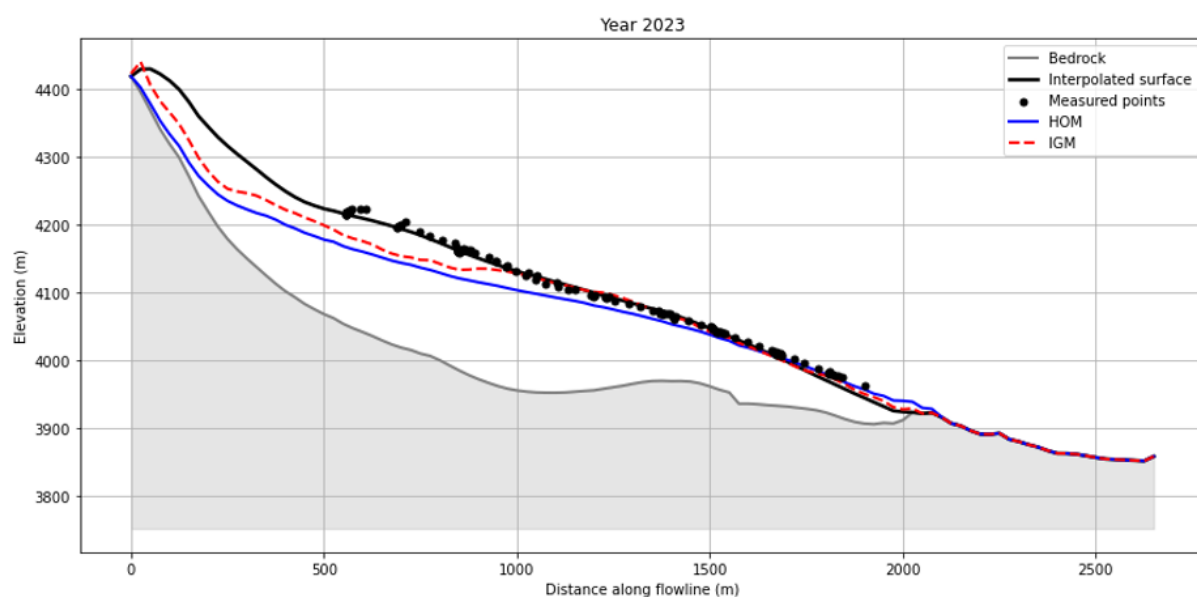


Figure 3.34: Cross-section in the year 2023. The RES-measurements are placed on the cross-section as dots. The black line is the interpolation created in the first part of this thesis. The cross-sections of the models in 2023 (HOM in blue and IGM in red) are added to the plot.

On the cross-section, the interpolation clearly aligns very close with the observations. However, it must be noted that there are clearly no observations in the accumulation area (the upper part of the glacier). In the lower parts of the glacier, the modelled surfaces are relatively close to the interpolated surface. Towards the higher parts of the glacier, the difference between the interpolation and the modelled surfaces increase when the elevation increases. The lack in RES-measurements however makes it uncertain what the surface elevation is in these highest parts. A possible explanation for the underestimation in the highest parts of the

glacier are the occurrence of too low accumulation above the ELA in the mass balance model or a too high compaction of snow. The surface modelled by IGM is at every place higher than the surface modelled by HOM. In the lowest part of the glacier, the difference is negligible. From 4050 m onward, the surface modelled by HOM is lower than the surface modelled by IGM. The difference never becomes larger than 30 m.

2023 is the only year when a comparison with the observations is possible. A comparison between the surfaces modelled by IGM and HOM at other moments in time is depicted in figure 3.35.

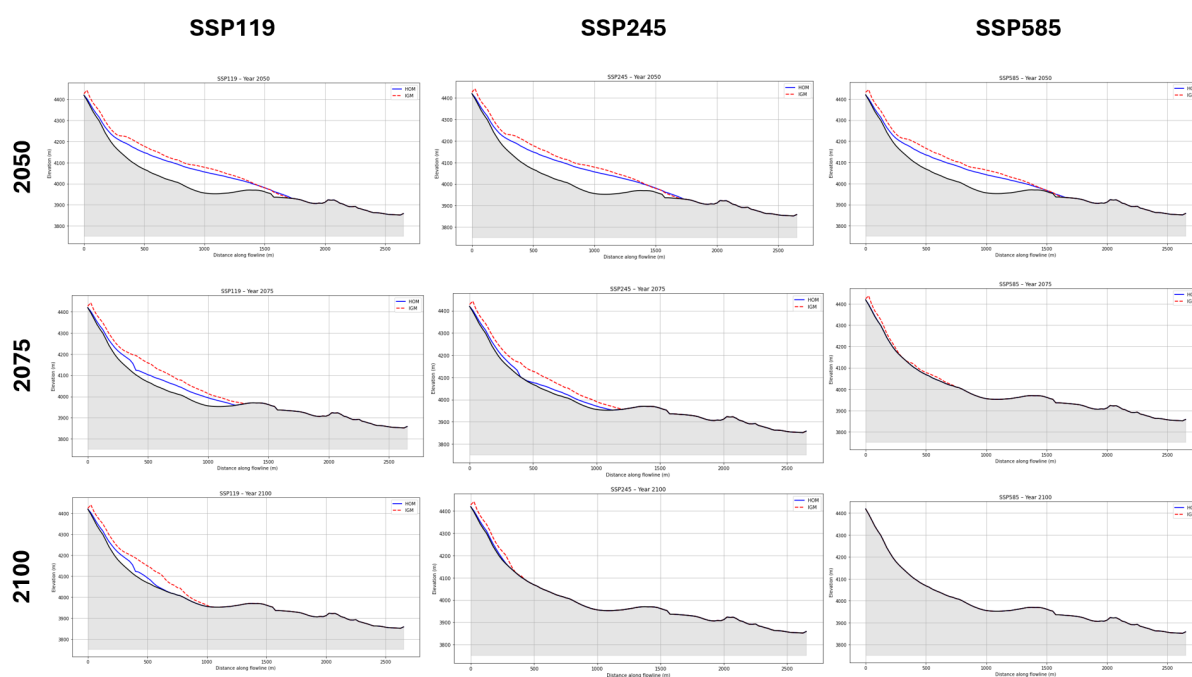


Figure 3.35: Cross-sections of the simulations for SSP1-1.9, SSP2-4.5 and SSP5-8.5 in the years 2050, 2075 and 2100 for both the HOM (blue) and IGM (red).

Overall, the IGM surface is always higher than the HOM surface, which aligns with the previously done observations that IGM predicts a thicker and (slightly) wider glacier. Only in 2050, all SSPs predict the HOM surface to be (a very little bit) higher than the IGM surface below 4000 m.

HOM predicts in 2075 around 4100 a place where the ice thickness becomes zero. This occurs in for SSP1-1.9 and SSP2-4.5 (SSP5-8.5 has no ice anymore in 2075). IGM does not predict such behaviour to happen. When looking at figure 3.31, the glacier has around this elevation a very narrow part that is remaining. The flowline passes just along this remaining patch, causing the cross-section to go to zero over a short distance.

In figure 3.36, the evolution of the glacier is depicted over time by using the colour.

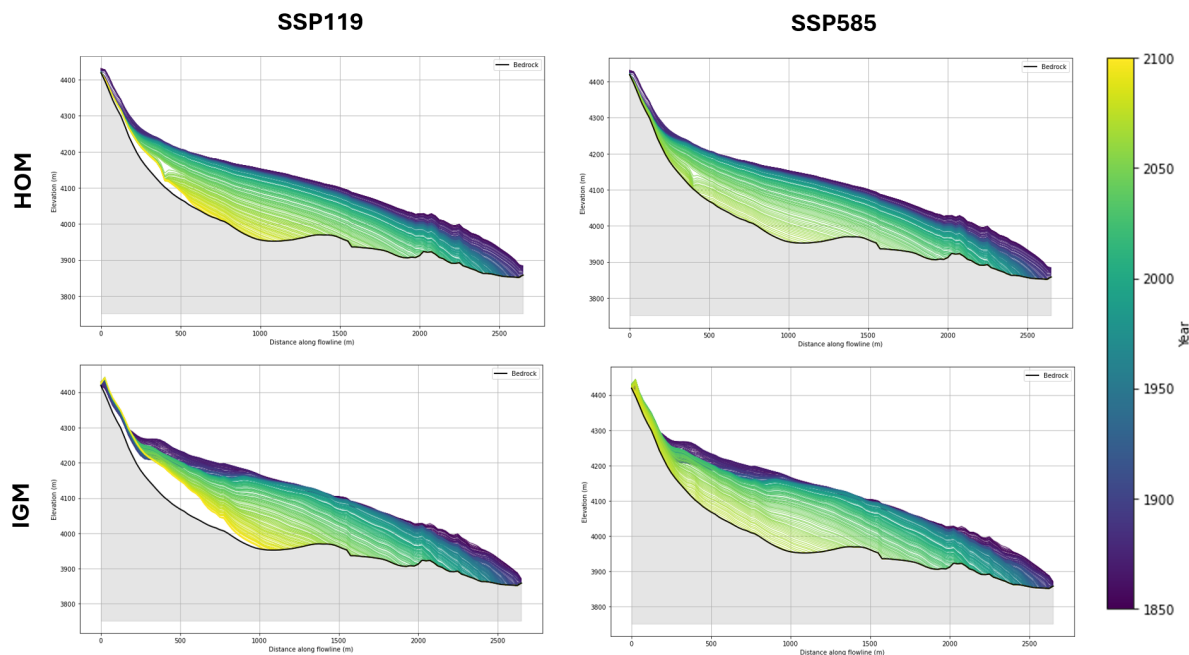


Figure 3.36: Cross-sections over time for HOM (upper row) and IGM (bottom row) simulations for SSP1-1.9, SSP2-4.5 and SSP5-8.5. The colour of the surface represents the year.

The most remarkable pattern is the thick blue part above 4150 m for both IGM figures. This points at a rapid melting in the highest parts of the glacier. For SSP1-1.9, the glacier even grows in the highest parts after 2000, after an strong thinning before 1900. However, the IGM-glacier seems much thicker in this part before 1900. The glacier starts too thick in this part and quickly melts. For SSP1-1.9, IGM melts much slower than HOM after 2050.

Figure 3.37 represents a stacked barplot for each year. The red part represents the ablation area, the blue part the accumulation area.

Overall, ablation areas are growing in proportion. HOM shows a more pronounced shift toward ablation dominance after 2050, especially SSP5-8.5. IGM tends to show earlier and more severe loss of accumulation area under higher emissions. For SSP1-1.9, accumulation area remains substantial for both models. Accumulation and ablation area are relatively well balanced. For SSP1-2.6, the accumulation area shrinks clearly after 200, while ablation becomes dominant. The evolution concerning HOM shows more lingering accumulation than the evolution concerning IGM. For SSP5-8.5, ablation dominates strongly by 2080. By 2100, almost no accumulation area remains in either model. IGM shows an earlier total collapse than HOM.

In order to gain further insights in the model results, an examination of the surface velocity is depicted by figure 3.38

The velocity field as modelled by HOM flows in 1850 the fastest in the ablation area north of the confluence of the body of the glacier and the small side branch at the western side. From 1900 onwards, also this steep ice patch and the highest part of the accumulation area contain fast flowing ice. From 2000 onwards, the velocity of the snout of the glacier significantly decreases. IGM models the same parts (glacier tongue, the side branch and the highest part)

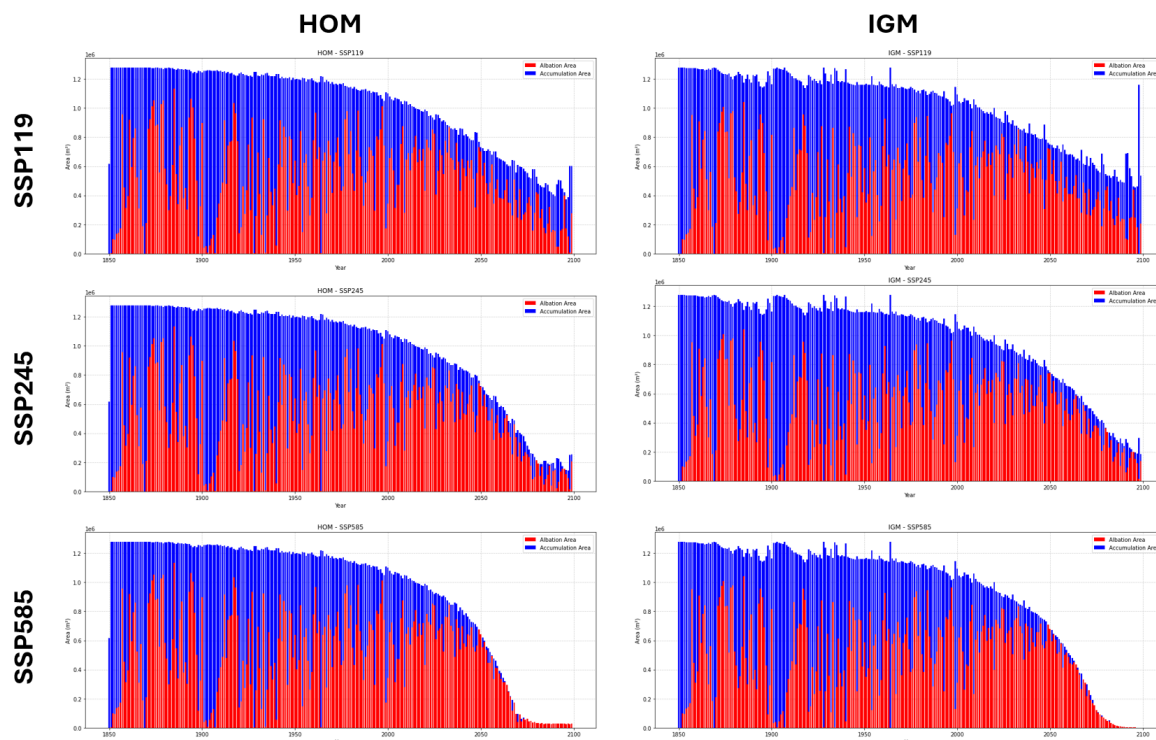


Figure 3.37: Annual stacked barplots representing the accumulation (blue) and ablation (red) area for SSP1-1.9, SSP2-4.5 and SSP5-8.5 for both HOM (left column) and IGM (right column).

to be the fast flowing parts already in 1850. From 1900 onwards, the velocity has decreased significantly, although the described regions remain (slightly) faster than the other parts.

In 1850, IGM produces a much faster field than HOM, except for the glacier tongue. Remarkably, in the middle of the accumulation area, there is a thin strip where HOM has a slightly faster field than IGM. In 1900, the difference between both fields is much smaller. Near the central flowline, IGM produces a faster flowing field. In the upper parts, the side branch and at the sides of the glacier, HOM predicts a faster flowing field. At the side walls in the east of the glacier, there is a part where IGM flows significantly faster. In the side branch and in the upper part, the difference between HOM and IGM is the most pronounced. From 1950 onwards, the field produced by HOM is higher than the field by IGM, except for the very thin strip along the side walls in the east. Mark the largest differences can be found on the fastest flowing parts: the highest area, the side branch and the middle of the tongue. In 2050, the difference between fields is reduced, except for the side branch.

In figure 3.39, the glacier-wide mean velocity as calculated by the two models is depicted.

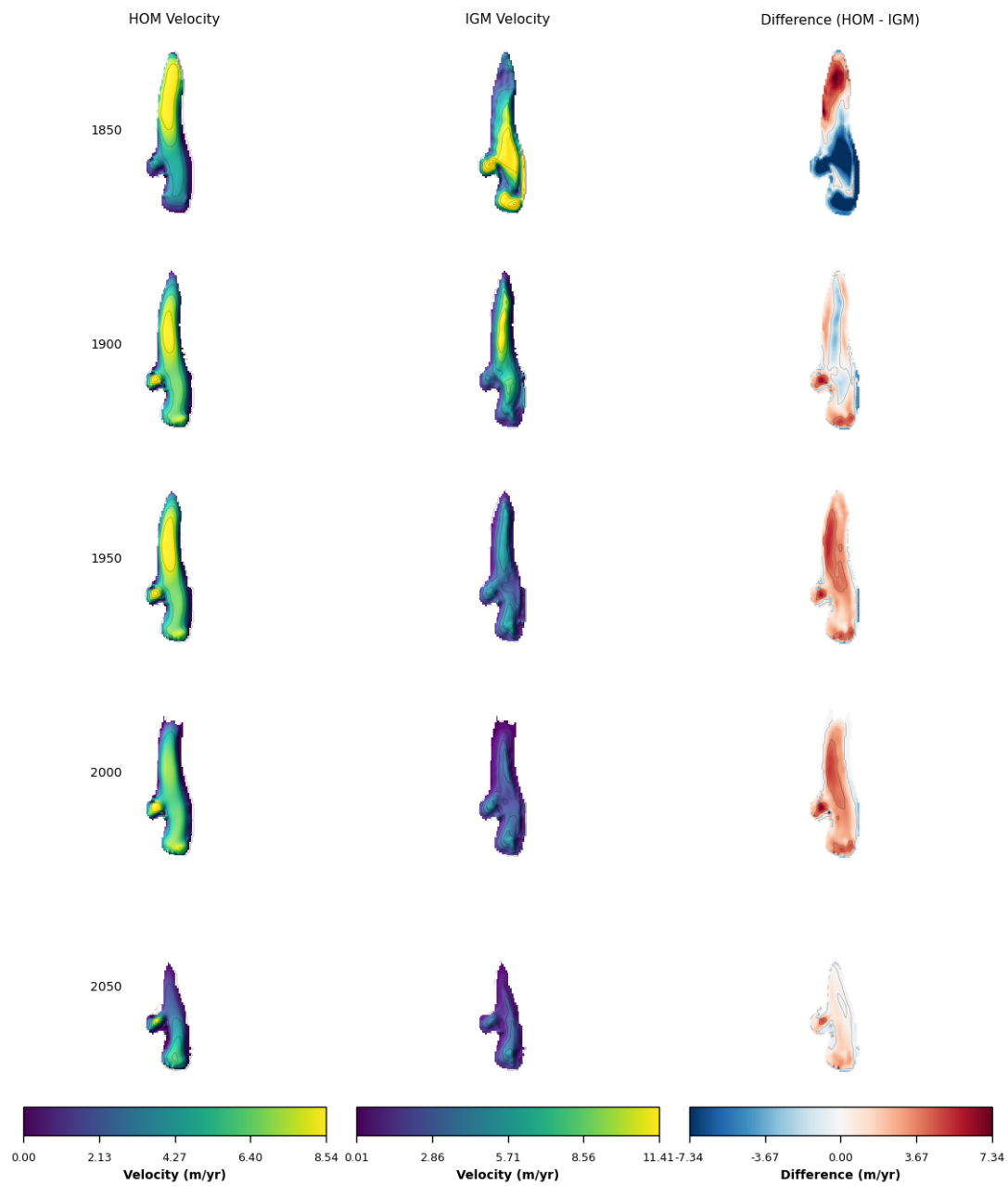


Figure 3.38: Surface velocity maps of the Batysh-Sook glacier on different moments in time for both HOM (first column) and IGM (second column) under SSP2-4.5. In the third column, the difference map between the HOM and IGM fields are depicted.

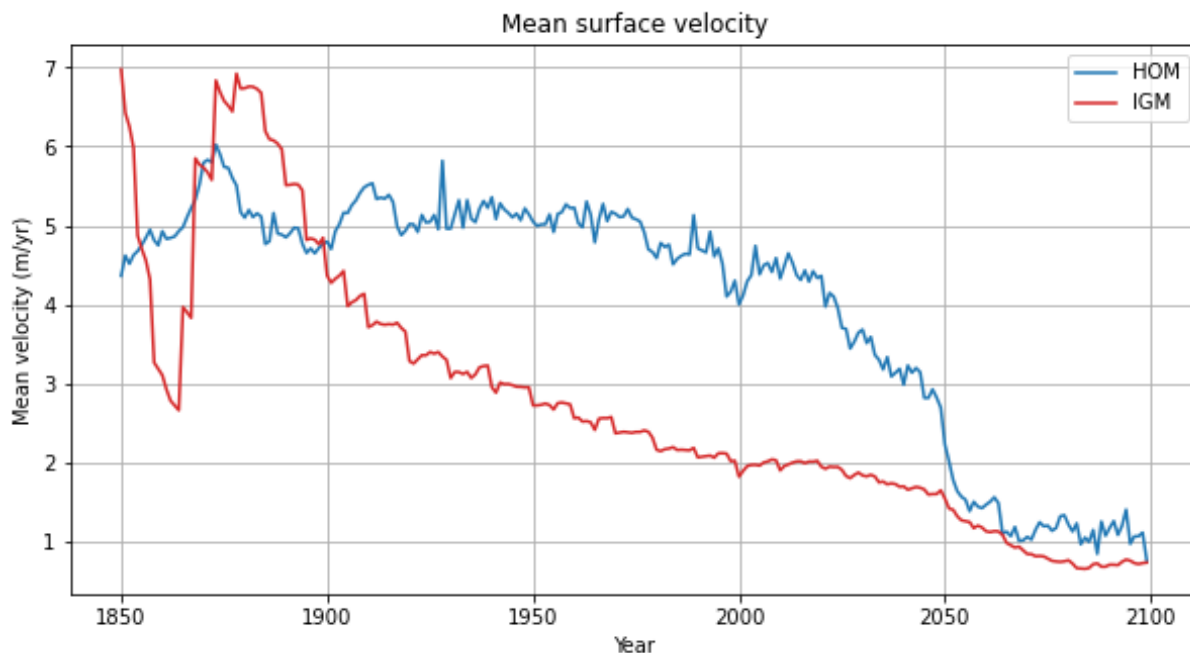


Figure 3.39: Evolution of the glacier-wide mean surface velocity of the Batysh-Sook glacier for HOM and IGM.

HOM predicts a more or less constant mean surface velocity up to 1970 and then a constant decline until 2060, from whereon the surface velocity remains constant. IGM shows a rapid decrease in the first years, after which a fast increase in velocity occurs. From 1880 onwards, the velocity is modelled to decrease parabolic: very fast in the beginning, slowing down later on. Since 1900, the mean surface velocity of IGM is lower than the velocity modelled by HOM. From 2050 onwards, the velocities of both models converge and are more alike. Until then, there is a significant difference between the models, both in magnitude and in shape.

3.4 Runoff

Figure 3.40 shows a scatterplot of the runoff as predicted by the HOM.

For SSP2-4.5, the peak runoff is predicted around 2052. However, the lower emission scenarios seem to reach their peak around the current period, while the higher emission scenarios are predicted to have reached a plateau by now. In figure 3.41, the same graph is made based on IGM.

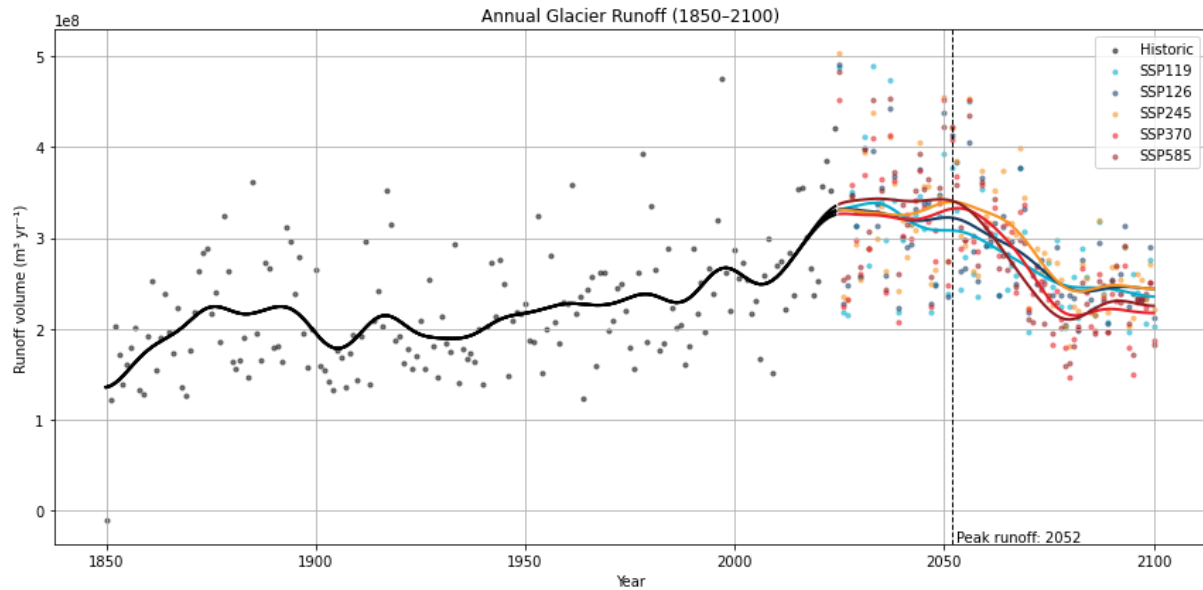


Figure 3.40: Scatterplot of the annual runoff for the different SSPs concerning the HOM. The trendline is the moving average over 11 years. The vertical line indicates the moment of the peakrunoff as predicted for SSP2-4.5.

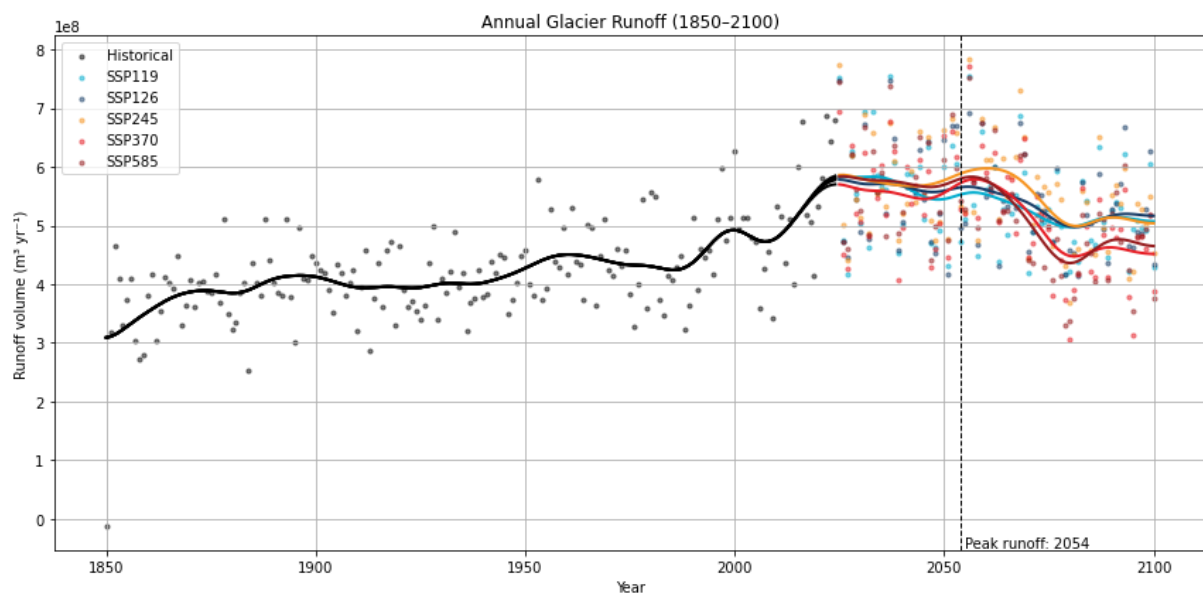


Figure 3.41: Scatterplot of the annual runoff for the different SSPs concerning IGM. The trendline is the moving average over 11 years. The vertical line indicates the moment of the peakrunoff as predicted for SSP2-4.5.

IGM predicts to reach the peak runoff around 2054, which is similar as the HOM. Also for IGM, a plateau is predicted to be reached at the moment for the higher emission scenarios (from SSP2-4.5 onward). The lower emission scenarios are predicted to reach the peak runoff by now. In absolute values, the runoff predicted by IGM is significantly higher than those by HOM, which does not surprise since IGM predicts slower retreat than HOM.

Chapter 4

Discussion

4.1 Ice thickness

Comparison of Interpolation Methods

The analysis of the residuals revealed clear differences between the different interpolations. Across most data subsets, IGM produces smoother and more coherent ice thickness fields than the spatial interpolation techniques, kriging and ANUDEM, and the yield stress method. Based on the histograms, QQ-plots and Kolmogorov-Smirnov tests, the IGM-residuals are overall closer to normal, less skewed, and less kurtotic than the spatial interpolation techniques. These results demonstrate that IGM consistently outperforms spatial interpolation techniques in producing smooth and physically plausible ice thickness fields, particularly in regions with sparse data. This suggests that IGM is better suited for extrapolating ice thickness in sparsely measured regions, as it avoids the sharp, unrealistic gradients seen in the spatial methods, particularly when only partial datasets (e.g. top or bottom 50%) are used. The slightly higher RMSE of IGM for random 50% and 90% subsets, compared to kriging and ANUDEM, is not unexpected given that the latter methods perform best when data coverage is dense and homogeneously distributed. However, IGM still provides residual distributions that are closer to Gaussian, which suggests less systematic bias and a more robust generalization in regions of sparse data.

The spatial autocorrelation analysis via Moran's I further confirmed that IGM-generated residuals reflect much less spatially structured deviations when the input is unevenly distributed, indicating unresolved spatial trends. The spatial interpolation techniques might result in (slightly) smaller RMSEs and lower Moran's I values when the RES-measurements are spatially even distributed. However, no measurements could be performed in the highest and steepest parts of the accumulation area, which also has to be taken into account when trying to interpret the RMSE values. The relatively poor performance of spatial interpolation techniques is not surprisingly and is consistent with earlier findings like Morlighem et al. (2011), where purely spatial methods struggled to represent ice thickness in accumulation zones with limited radar measurements.

The yield stress method exhibits intermediate performance, outperforming spatial interpolations in data-sparse regions but remaining more sensitive to the dataset distribution than IGM. The fields are smoother than kriging and ANUDEM, but less smooth than IGM. The fact that the residuals are closer to normal than the residuals of IGM, indicates strong performance

in sufficiently measured areas. However, the tests with the bottom and top 50% points out that the yield stress model relies on the data more than IGM. IGM retains higher Moran's I values than the yield stress method for randomly distributed subsets. This could indicate that IGM tends to preserve broader-scale structural patterns, whereas the yield stress method may be more localized. The extrapolation of thin connections between disconnected ice margins, unique to the yield stress method, can be attributed to its lack of an explicit ice mask.

Because the highest and steepest parts of the accumulation area are often difficult to reach, unmeasured areas are common. The spatial interpolations do not use any information from the surface elevation DEM, while IGM and the yield stress method are able to use this information. Many of the common methods to deduce ice thickness from the surface characteristics use the surface slope. Since IGM and the yield stress method have the steepness information too, the relation between the surface slope and the ice thickness (via the yield stress) can be used by those models, in contrast to the spatial interpolation techniques. In areas without thickness observation, the model still has the slope information, while the spatial interpolation techniques have no information at all. The input of the digital elevation model also brings resistance against poor measurements. In this case, a relatively high amount of meltwater disturbed the RES measurements. IGM likely resulted in a more stable ice thickness field than the spatial interpolations, which only use the (largely fluctuating) ice thickness observations.

The smoother and more physically plausible fields generated by IGM are preferable for modelling applications, such as ice flow simulations or mass balance assessments. While spatial methods may offer lower RMSE in well-sampled areas, their tendency to overfit limits their use in predictive contexts. The yield stress method creates an ice thickness field with intermediate smoothness. In densely measured areas, the yield stress method likely produces more realistic fields than IGM, which is possibly too smooth to be 100% realistic. However, in many areas where measurements are scarce or even absent, the yield stress method is likely to result in a less accurate field than the field produced by IGM. The results suggest that incorporating physical constraints, even in a simplified form as done in IGM, improves generalizability and robustness. This has broader implications for cryospheric science, where observational data are often sparse and unevenly distributed.

The optimisation algorithm built into the IGM Optimize module allows for passing initial fields (velocity, surface elevation, ice thickness) to the model. During all performed tests, the surface elevation was given as initial field (and as cost function factor too). In half of the tests, also the initial velocity field of Millan et al. (2022) was given as initial field. Barely any difference between the interpolations with and without initial flow field are observed. Since the ice velocity estimates of Millan et al. (2022) are the result of a global model based on rough satellite data, they are known to perform rather poor on individual glaciers. Therefore, it does not surprise the cost function creates a velocity field entirely different from the initial one, and the original field does not improve the interpolation.

Sensitivity analysis

The sensitivity analysis underscores the importance of the standard deviation of thickness observations and the regularization term for IGM. Both parameters have a direct influence on the trade-off between accuracy and smoothness, with high thickness standard deviations introducing stronger variability in the predictions. The SHAP-analysis confirms that these

parameters exert both positive and negative influences depending on their magnitude, whereas factors like the anisotropy parameter and initial sliding coefficient play a secondary role. IGM's instability in some partial dependence plots could point to some numerical instability. The opti-smooth-anisotropy-factor, while deemed highly influential in theory, had to be fixed due to the emergence of a chessboard artifact. This highlights a limitation of the current IGM implementation and points to a potential avenue for improving the robustness of anisotropy handling in future work.

Limitations

All interpolations are largely influenced by the parameter values they are given. As the sensitivity analysis of the different methods showed, especially the accuracy of the input data is very influential.

It is important to note that the sensitivity analysis was performed 'on the side': all used values result from all intermediate steps during the optimisation algorithm. Since the optimisation algorithm searches for the best performing combination of parameters, certain combinations of parameters are overrepresented, while other combinations might be neglected. A better approach would be to make a selection of values for each parameter and then check all combinations. The choice for this approach was made because of its time- and computing efficiency and because the sensitivity analysis is not the core of the research.

The residual analysis (figures 3.2 to 3.7) shows that IGM generally produces smoother residual patterns with lower spatial autocorrelation (Moran's I) compared to kriging and ANUDEM, especially in data-sparse regions. This smoothness is beneficial for avoiding local thickness extremes, but may also under-represent small-scale bed irregularities that can influence local velocity gradients. Subset performance differences also reveal that central-flowline measurements alone provide less reliable interpolation than random distributions, highlighting the importance of cross-glacier coverage for constraining the inversion. These interpolation characteristics directly influence the initial geometry used for forward modelling and therefore partially explain differences in subsequent glacier evolution.

4.2 Glacier evolution

Overall evolution and regional context

Batysh-Sook Glacier is projected to continue shrinking throughout the 21st century. Under SSP5-8.5, both HOM and IGM simulations predict complete disappearance by 2100, and under SSP3-7.0 this occurs for HOM as well. Even in the most optimistic scenario (lower bound of SSP1-1.9), around two-thirds of the 1850 glacier area and over 99 % of the original volume are lost by the end of the century. Volume loss accelerates faster than area reduction, indicating that thinning dominates over lateral retreat.

A comparison with the glaciers modelled by Van Tricht and Huybrechts (2023) is required, since these are similar glaciers (and one ice cap, Grigoriev) in the same area. Table 4.1 contains the percentage of area and volume for different glaciers compared to the present day situation.

For the glaciers by Van Tricht and Huybrechts (2023), the present day situation is 2021. For Batysh-Sook, the present day situation refers to 2024.

Table 4.1: Projected percentage of glacier area and volume remaining relative to the current situation (2021, except for Batysh-Sook) values under different SSPs by 2100 concerning Van Tricht and Huybrechts (2023)

Glacier	Area Remaining (%)			Volume Remaining (%)		
	SSP1-1.9	SSP2-4.5	SSP5-8.5	SSP1-1.9	SSP2-4.5	SSP5-8.5
Ashu-Tor	52	14	0	25	4	0
Bordu	49	24	0	30	11	0
Golubin	70	40	0	46	13	0
Grigoriev	69	39	0	12	2	0
Kara-Batkak	80	60	19	88	33	5
Sary-Tor	41	11	0	19	4	0
Batysh-Sook (HOM)	25	6	0	10	2	0
Batysh-Sook (IGM)	34	8	0	21	5	0

Compared to other glaciers in the region, Batysh-Sook exhibits a markedly higher sensitivity to projected climate scenarios. Both modelling approaches indicate significant area and volume loss by 2100, even under the most optimistic scenario (SSP1-1.9), where less than 35% of the glacier area and only 10–21% of its volume are projected to remain. Most other studied glaciers are significantly larger and have at least double or triple the ice volume. Because Batysh-Sook is small, thin, and low in altitude, it warms more quickly and cannot buffer climatic impacts as effectively as bigger, colder glaciers, leading to earlier disappearance.

Altitude plays a minor role in explaining this sensitivity. Batysh-Sook has a mean elevation of just 4200 meters, an intermediate value for the glaciers in the dataset. Lower-elevation glaciers are more exposed to warm air temperatures and experience longer melt seasons. For comparison, Kara-Batkak, which retains a relatively high portion of ice under moderate scenarios, is located at a mean altitude of 3950 meters. While orientation can influence energy balance, its role here appears secondary. Batysh-Sook faces north, which typically reduces direct solar radiation and can help preserve ice. However, this advantage seems not to play a crucial role. Interestingly, Kara-Batkak also has a northern orientation, but in its case, the combination of a lower altitude and small size results in better retention of both area and volume in all scenarios.

In terms of ice thickness, Batysh-Sook reaches a maximum of 185 meters, which is relatively high considering its small size. This suggests it may have a steep or confined valley profile. Nevertheless, its overall volume remains very low, meaning the glacier lacks the storage capacity to withstand sustained melting. This emphasizes the importance of volume–altitude distribution: small, low glaciers have most of their ice mass near the melt threshold, making them disproportionately vulnerable to warming.

Model response to climate forcing

Despite geometric differences, both models predict nearly identical peak runoff timing (2052 for HOM, 2054 for IGM) suggesting that basin-integrated meltwater release is controlled primarily by mass balance forcing rather than ice dynamics. This outcome reflects the shared SMB model input, which removes temperature–elevation feedback from IGM’s runs. The implication is that for small glaciers such as Batysh-Sook, climatic forcing exerts stronger control over hydrological response than dynamic complexity.

Table 4.2: Projected peak runoff years for the glaciers studied by Van Tricht and Huybrechts (2023) and Batysh-Sook).

Glacier	Peak Runoff Year
Ashu-Tor	2054
Bordu	2030
Golubin	2054
Grigoriev	2037
Kara-Batkak	2037
Sary-Tor	2032
Batysh-Sook (HOM)	2052
Batysh-Sook (IGM)	2054

The enhancement factor E increases the deformation of the ice with the temperature. However, in colder parts of the glacier, the ice is already so stiff that changes in E have little effect. In contrast, ASL controls how much the glacier slides at the base, depending on the shape and slope of the ice. But if the base is too cold or flat, sliding will not happen much, even if ASL is high.

Comparison IGM and HOM

Both models produce broadly similar trends in glacier length and area, but systematic differences emerge. IGM consistently predicts thicker ice and slightly larger areas, while HOM tends to better match observed length. In the historic period, IGM slightly overestimates 2023 glacier area and strongly overestimates volume, while HOM slightly underestimates both. For glacier length, HOM aligns well with observations since the 1970s, whereas IGM underestimates length. In 2023, IGM yields a glacier that is smaller in length but thicker and wider than observed, whereas HOM more closely reproduces cross-sectional shapes from RES measurements. HOM performs therefore better over the historic period, but that does not necessarily means HOM also performs better for the future. Given the difference between both models in area and volume does not increases much anymore in the future period, this could indicate the difference in 2023 is the cause for the difference in the future period.

The most important difference between HOM and IGM lies in their treatment of ice rheology and thermodynamics. In HOM, Glen’s flow law links deviatoric stress and strain rate, with a temperature-dependent flow factor $A(T)$ derived from an Arrhenius-like formulation. This enables HOM to capture nonlinear stress–strain behaviour and the influence of temperature on ice deformation. Conversely, IGM as implemented in this thesis assumes a constant

$A(T)$ and therefore does not allow thermodynamics to vary over time and space, potentially underestimating the role of spatial variations in temperature and basal conditions.

While HOM explicitly resolves stress transfer between different parts of the glacier (to the extent allowed by the Blatter–Pattyn approximation), IGM simulates bridging effects via its emulator. The difference might contribute to the differences in steep or narrow sections of the glacier, where HOM predicted higher velocity gradients and localized acceleration, whereas IGM's flow field was smoother. In these places, where higher-order stress terms are significant, the emulated ice dynamics result in slightly different results.

The maps of the surface velocities, points to this difference as the main reason for the differences between the results of both models. In 1850, IGM predicts much faster flow almost everywhere except in the glacier tongue. This faster flow in the accumulation area tends to export mass more quickly downstream (unless it would be compensated by higher accumulation), which can lead to thinning in the upstream parts and thickening downstream. Since IGM has a constant flow rate factor A , it may overestimate ice softness in cold regions early on, which would make the ice flow faster than in HOM's temperature-dependent rheology. This can explain why IGM produces a glacier that is thicker. From 1950 onwards, HOM's surface velocity is higher than IGM's almost everywhere. These higher velocities allow HOM to move ice more efficiently towards the tongue, maintaining glacier length closer to the observed values. The lower velocities modelled by IGM cause less supply of ice in the snout. Therefore, the retreat is faster and the length is slightly shorter than the one predicted by HOM. Cross-section analysis (figures 3.34 to 3.36) confirms these broader patterns: in 2023, IGM overestimates thickness and width in the upper and mid-glacier compared to RES measurements, whereas HOM aligns more closely with observed cross-sectional shapes. Over time, HOM's cross-sections thin more evenly, while IGM's maintain greater central thickness but lose width more rapidly. These structural differences in glacier shape contribute to the contrasting length evolution and local velocity fields.

The annual accumulation/ablation area evolution (figure 3.37) shows that IGM consistently maintains a slightly larger accumulation area fraction in the early historic period, consistent with its thicker ice in the upper parts and initially higher velocities there. HOM, with smaller accumulation area but higher late-century velocities, shifts mass more efficiently to the ablation zone, sustaining snout position longer.

Another key difference lies in the response to climate forcing. The HOM's dynamic calibration was sensitive to variations in precipitation and temperature, and the retreat pattern aligned with satellite-derived glacier outlines from Landsat images. IGM, by using the same mass balance model but different internal flow dynamics, responded slightly differently to the same forcings. While both models predicted overall retreat, the timing of the onset of rapid thinning and length reduction was offset by several years. This divergence underlines the importance of correctly parameterizing basal sliding and rheological parameters, which are directly tuned in HOM but implicitly encoded in IGM's CNN.

Limitations and further research

The biggest limitation of this experiment is the disabling of the mass balance model in IGM. Because the goal of the research was to compare the flow model and its emulated counterpart

in IGM, the choice fell to implement the mass balance model from HOM directly into IGM. Therefore, the 'enthalpy' module of IGM, which contains the thermodynamic model, was turned off. As discussed before, many of the observed differences between the results of both models can potentially be assigned to the constant A in IGM.

The IGM model runs are performed by using the LIA-geometry as created by the spin-up experiment of the HOM. The glacier could also be spinned up via IGM itself. For the purpose of this thesis, only the evolution of the glacier is compared. A further research could investigate the spinning up. However, a different spin-up might cause a slightly different evolution.

The output of the mass balance model from the HOM was applied on IGM as well. It might be interesting to do the inverse as well, and applying the mass balance model of IGM to HOM. The drawback of not using the mass balance model of IGM is the loss of the temperature-elevation feedback. However, Zekollari et al. (2014) showed that the difference between a fixed and a dynamic geometry is relatively small for a warming of 4 °C at Morteratsch glacier (Switzerland), which in 2010 had an area of 16 km^2 and ranges from 2050 to 4000 m. Since Batysh-Sook is much smaller and stretched over a smaller elevation range, the effect of the temperature-elevation feedback is assumed to be negligible.

This thesis shows that IGM exhibits similar results to HOM, but an examination of the thermodynamic model is required to have a better view on the performance of the emulator. A possible follow-up research could be to implement the mass balance model of HOM in IGM. This would cancel out the influence of the mass balance model and would enable to use the 'enthalpy' module of IGM. Another promising direction would be to implement ensemble approaches, where HOM is used for calibration and parameter tuning, and IGM is used for fast projection runs.

Practical implementation

IGM is coded in Python, currently the most popular coding language. This makes it more user-friendly than HOM, which is coded in Fortran code. To make use of IGM, it is usually not needed to make changes in the code. However, it is easy to upload your own script for specific experiments. Only the flowmodel cannot be changed, since that is included in the emulator. The numerous Fortran scripts of the HOM on the other hand are more difficult to adapt, not only because of the language but also because of the interdependence of the different scripts on each other.

IGM is well documented. A GitHub repository contains all codes and examples of implementations. Since the model is open-source, everyone is allowed to perform improvements and additions. Also a technical paper is provided. Despite its relatively recent origin, a relatively large community of users and contributors is working on the development of the model. A active Discord server exists where users can ask questions and provide suggestions. Both the makers of the model, as well as other users, quickly reply and are helpful for suggestions and fixes. HOM on the contrary is not free available on the internet. Its documentation is rather poor compared to that of IGM. Since it longer existence, more VUB-founded scientists are familiar with the model, making it easier to ask assistance.

IGM has the main advantage to be faster. Where HOM requires a number of hours to

compute the entire evolution of the glacier, running IGM takes only a few minutes. It is important to note that during this time, HOM has to calculate the mass balance as well, while this was provided as input to IGM. Remind IGM has a mass balance model as well that could have been used. For HOM, a calibration experiment is required where E and ASL are calibrated by running the model (or a simplified version of the model like SIA) for a large number of combinations of E and ASL . IGM computes optimizes its own parameters.

To run IGM, a computer with GPUs is useful. Theoretically, the model can be ran on CPUs, but this removes the computational benefits of the model. To run HOM, a high performance computer (HPC) is required. In this case, the model was submitted to the Hydra cluster of the Vlaamse Supercomputer Centre (VSC). Getting access to such hpc's reduces the ease of use. IGM is driven by a short JPEG-file containing the required model settings and parameters. This provides a relatively easy use. To change parameters or settings in HOM, one has to make changes in the Fortran files itself. Although most parameters are included in a parameterlist file, it can be hard to find the right parameter and its place.

From a regional modelling perspective, these results highlight that even when using the same SMB model, differences in ice dynamics parameterisation can lead to persistent discrepancies in glacier evolution. For small, steep, or complexly shaped glaciers, correctly representing higher-order stress transfer and temperature-dependent rheology is essential for matching observed historic changes. At the same time, initial thickness uncertainty can imprint on future projections for decades, underscoring the need to integrate high-quality thickness data into regional projections.

Regional implications

From a regional modelling perspective, these results demonstrate that even with identical SMB forcing, differences in ice dynamics parameterisation create persistent discrepancies in glacier evolution. For small, steep, or low-altitude glaciers, representing higher-order stress transfer and temperature-dependent rheology is crucial to reproduce observed changes. At the same time, biases in initial thickness propagate into future projections, highlighting the importance of accurate thickness measurements.

For water resource management, the finding that runoff timing is largely insensitive to dynamics but strongly governed by SMB forcing underscores the dominant role of climate in small glacier hydrology. For regional projections, combining models like HOM and IGM may therefore be a powerful approach: HOM constrains dynamics, while IGM enables efficient large-scale scenario exploration. Together, they provide complementary insights into the fate of vulnerable glaciers such as Batysh-Sook.

Chapter 5

Conclusions

5.1 Batysh Sook evolution

The Batysh-Sook glacier, a small, middle-elevation glacier in Kyrgyzstan's Tien Shan mountains, exhibits high vulnerability to climate change. Under all SSPs, both the HOM and IGM project significant retreat. High-emission scenarios predict a complete glacier loss by 2069 (HOM) and by 2085 (IGM). Low-emission scenarios (SSP1-1.9) preserve 25–34% of area and 10–21% of volume by 2100. The glacier's rapid decline is attributed to its small size (0.8 km²). Volume loss outpaces area reduction, indicating dominant thinning over lateral retreat. Peak runoff is projected around 2052–2054 for both models, indicating critical water supply challenges for downstream agriculture in Central Asia.

5.2 Performance of IGM for ice thickness interpolations

IGM's inverse modelling module demonstrated superior performance in interpolating sparse ice thickness measurements compared to traditional methods. IGM shows robustness in data-sparse regions, since it produces smoother, physically plausible thickness fields with lower spatial autocorrelation (Moran's I) than kriging or ANUDEM, especially in the accumulation zone where radar data were scarce. Residuals from IGM interpolations were closer to a normal distribution (lower KS statistics) and less skewed than those from spatial methods, reducing systematic bias. For subsets like the top/bottom 50% of measurements, IGM achieved lower RMSE than spatial methods. However, for randomly distributed data (90%), kriging/ANUDEM had marginally better RMSE, though IGM retained superior generalizability. By integrating surface elevation data and ice flow physics via a neural network emulator, IGM avoided unrealistic extrapolations, outperforming the yield stress method in global smoothness.

5.3 Performance of IGM for modelling glacier evolution

For forward modelling, IGM showed strong alignment with HOM in projecting long-term trends but exhibited key differences. IGM consistently simulated thicker ice (+20–30% volume) and slower retreat than HOM, leading to delayed disappearance dates (e.g., 2085 vs. 2069 for SSP5-8.5). It is important to note that the experimental setup might be the reason for this difference, since the thermodynamic model of IGM was not used. Both models agreed

on peak runoff timing (2052–2054), glacier length evolution, and accumulation-ablation area shifts, confirming that mass balance forcing dominated over ice dynamics. GPU enables faster simulations while capturing stress coupling and flow dynamics via its physics-informed neural network. IGM overestimated modern ice thickness due to under-resolved accumulation processes. Enabling its enthalpy module and refining initial conditions could improve accuracy.

Bibliography

- Aizen, V. B., Aizen, E. M., Melack, J. M., and Dozier, J.: Climatic and Hydrologic Changes in the Tien Shan, Central Asia, *Journal of Climate*, 10, 1393 – 1404, [https://doi.org/10.1175/1520-0442\(1997\)010<1393:CAHCIT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<1393:CAHCIT>2.0.CO;2), 1997.
- Bernauer, T. and Siegfried, T.: Climate change and international water conflict in Central Asia, *Journal of Peace Research*, 49, 227–239, 2012.
- Blatter, H.: Velocity and stress fields in grounded glaciers: a simple algorithm for including deviatoric stress gradients, *Journal of Glaciology*, 41, 333–344, 1995.
- Bolch, T., Duethmann, D., Wortmann, M., Liu, S., and Disse, M.: Declining glaciers endanger sustainable development of the oases along the Aksu-Tarim River (Central Asia), *International journal of sustainable development and world ecology*, 29, 209–218, 2022.
- Burt, J. E., Barber, G. M., and Rigby, D. L.: *Elementary statistics for geographers*, Guilford Press, 2009.
- Chen, H., Chen, Y., Li, W., and Li, Z.: Quantifying the contributions of snow/glacier meltwater to river runoff in the Tianshan Mountains, Central Asia, *Global and Planetary Change*, 174, 47–57, 2019.
- Climate-Data.org: Climate Data for Cities Worldwide, <https://en.climate-data.org/>, accessed: 2024-10-10, 2024.
- Cogley, J. G., Arendt, A. A., Bauder, A., Braithwaite, R. J., Hock, R., Jansson, P., Kaser, G., Moller, M., Nicholson, L., Rasmussen, L. A., and Zemp, M.: Glossary of glacier mass balance and related terms, Technical report, International Hydrological Programme, 2010.
- Cuffey, K. M. and Paterson, W. S. B.: *The physics of glaciers*, Academic Press, 2010.
- Emmer, A. et al.: Progress and challenges in glacial lake outburst flood research (2017-2021): a research community perspective, *Natural Hazards and Earth System Sciences*, 22, 3041–3061, 2022.
- ESRI: How Topo To Raster Works, URL <https://pro.arcgis.com/en/pro-app/latest/tool-reference/3d-analyst/how-topo-to-raster-works.htm>, accessed: 2025-05-30, 2023a.
- ESRI: How Kriging Works, URL <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/how-kriging-works.htm>, accessed: 2025-04-16, 2023b.

- Farinotti, D., Huss, M., Bauder, A., Funk, M., and Truffer, M.: A method to estimate the ice volume and ice-thickness distribution of alpine glaciers, *Journal of Glaciology*, 55, 422–430, <https://doi.org/10.3189/002214309788816759>, 2009.
- Farinotti, D., Longuevergne, L., Moholdt, G., Duethmann, D., Mölg, T., Bolch, T., Vorogushyn, S., and Güntner, A.: Substantial glacier mass loss in the Tien Shan over the past 50 years, *Nature Geoscience*, 8, 716–722, 2015.
- Farinotti, D., Brinkerhoff, D. J., Clarke, G. K. C., Fürst, J. J., Frey, H., Gantayat, P., Gillet-Chaulet, F., Girard, C., Huss, M., Leclercq, P. W., Linsbauer, A., Machguth, H., Martin, C., Maussion, F., Morlighem, M., Mosbeux, C., Pandit, A., Portmann, A., Rabatel, A., Ramsankaran, R., Reerink, T. J., Sanchez, O., Stentoft, P. A., Singh Kumari, S., van Pelt, W. J. J., Anderson, B., Benham, T., Binder, D., Dowdeswell, J. A., Fischer, A., Helfricht, K., Kutuzov, S., Lavrentiev, I., McNabb, R., Gudmundsson, G. H., Li, H., and Andreassen, L. M.: How accurate are estimates of glacier ice thickness? Results from ITMIX, the Ice Thickness Models Intercomparison eXperiment, *The Cryosphere*, 11, 949–970, <https://doi.org/10.5194/tc-11-949-2017>, 2017.
- Frank, T. and van Pelt, W. J. J.: Ice volume and thickness of all Scandinavian glaciers and ice caps, *Journal of Glaciology*, pp. 1–14, 2024.
- Frey, H., Machguth, H., Huss, M., Huggel, C., Bajracharya, S., Bolch, T., Kulkarni, A., Linsbauer, A., Salzmann, N., and Stoffel, M.: Estimating the volume of glaciers in the Himalayan-Karakoram region using different methods, *The Cryosphere*, 8, 2313–2333, <https://doi.org/10.5194/tc-8-2313-2014>, 2014.
- Fürst, J. J., Rybak, O., Goelzer, H., De Smedt, B., de Groen, P., and Huybrechts, P.: Improved convergence and stability properties in a three-dimensional higher-order ice sheet model, *Geoscientific Model Development*, 4, 1133–1149, <https://doi.org/10.5194/gmd-4-1133-2011>, 2011.
- Fürst, J. J., Gillet-Chaulet, F., Benham, T. J., Dowdeswell, J. A., Grabiec, M., Navarro, F., Pettersson, R., Moholdt, G., Nuth, C., Sass, B., Aas, K., Fettweis, X., Lang, C., Seehaus, T., and Braun, M.: Application of a two-step approach for mapping ice thickness to various glacier types on Svalbard, *The Cryosphere*, 11, 2003–2032, <https://doi.org/10.5194/tc-11-2003-2017>, 2017.
- Glen, J.: Rate of flow of polycrystalline ice, *Nature*, 172, 721–722, 1953.
- Glen, J. W. and Paren, J. G.: The Electrical Properties of Snow and Ice, *Journal of Glaciology*, 15, 15–38, <https://doi.org/10.3189/S0022143000034249>, 1975.
- Hagg, W., Mayer, C., Lambrecht, A., Kriegel, D., and Azizov, E.: Glacier changes in the Big Naryn basin, Central Tian Shan, *Global and Planetary Change*, 110, 40–50, URL <https://www.sciencedirect.com/science/article/pii/S0921818112001580>, 2013a.
- Hagg, W., Mayer, C., Lambrecht, A., Kriegel, D., and Azizov, E.: Glacier changes in the Big Naryn basin, Central Tian Shan, *Global and Planetary Change*, 110, 40–50, 2013b.

- Hoelzle, M., Azisov, E., Barandun, M., Huss, M., Farinotti, D., Gafurov, A., Hagg, W., Kenzhebaev, R., Kronenberg, M., Machguth, H., Merkushev, A., Moldobekov, B., Petrov, M., Saks, T., Salzmann, N., Schöne, T., Tarasov, Y., Usabaliev, R., Vorogushyn, S., Yakovlev, A., and Zemp, M.: Re-establishing glacier monitoring in Kyrgyzstan and Uzbekistan, Central Asia, *Geoscientific Instrumentation, Methods and Data Systems*, 6, 397–418, <https://doi.org/10.5194/gi-6-397-2017>, 2017a.
- Hoelzle, M., Azisov, E., Barandun, M., Huss, M., Farinotti, D., Gafurov, A., Hagg, W., Kenzhebaev, R., Kronenberg, M., Machguth, H., et al.: Re-establishing glacier monitoring in Kyrgyzstan and Uzbekistan, Central Asia, *Geoscientific Instrumentation, Methods and Data Systems*, 6, 397–418, 2017b.
- Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., et al.: Accelerated global glacier mass loss in the early twenty-first century, *Nature*, 592, 726–731, 2021.
- Huss, M. and Hock, R.: Global-scale hydrological response to future glacier mass loss, *Nature climate change*, 8, 135–140, 2018.
- Hutchinson, M.: ANUDEM version 5.3, user guide, Canberra: Fenner School of Environment and Society, Australian National University, 2011.
- Hutchinson, M. F.: Calculation of hydrologically sound digital elevation models, in: *Proceedings of Third International Symposium on Spatial Data Handling*, pp. 17–19, International Geography Union, Columbus, OH, 1988.
- Immerzeel, W. W. et al.: Importance and vulnerability of the world’s water towers, *Nature*, 577, 364–369, 2020.
- IPCC: Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, <https://doi.org/10.59327/IPCC/AR6-9789291691647>, 2023.
- Johnson, R.: Thermal low, *Encyclopedia of atmospheric science*, pp. 2269–2273, 2003.
- Jouvet, G.: Inversion of a Stokes glacier flow model emulated by deep learning, *Journal of Glaciology*, 69, 13–26, 2023.
- 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, 2022.
- Jouvet, G., Cook, S., Cordonnier, G., Finley, B., Henz, A., Herrmann, O., Maussion, F., Mey, J., Scherler, D., and Welty, E.: Concepts and capabilities of the Instructed Glacier Model, 2024.
- Kenzhebaev, R., Barandun, M., Kronenberg, M., Chen, Y., Usabaliev, R., and Hoelzle, M.: Mass balance observations and reconstruction for Batysh Sook Glacier, Tien Shan, from 2004 to 2016, *Cold Regions Science and Technology*, 135, 76–89, URL <https://www.sciencedirect.com/science/article/pii/S0165232X16304001>, 2017.

- Kick, W.: The first glaciologists in Central Asia, *Journal of Glaciology*, 3, 687–692, 1960.
- Kraaijenbrink, P. D., Bierkens, M. F., Lutz, A. F., and Immerzeel, W.: Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers, *Nature*, 549, 257–260, 2017.
- Kutuzov, S. and Shahgedanova, M.: Glacier retreat and climatic variability in the eastern Terskey–Alatoo, inner Tien Shan between the middle of the 19th century and beginning of the 21st century, *Global and Planetary Change*, 69, 59–70, 2009.
- Lapazaran, J., Otero, J., Martín-Español, A., and Navarro, F.: On the errors involved in ice-thickness estimates I: ground-penetrating radar measurement errors, *Journal of Glaciology*, 62, 1008–1020, 2016a.
- Lapazaran, J., Otero, J., Martín-Español, A., and Navarro, F.: On the errors involved in ice-thickness estimates II: errors in digital elevation models of ice thickness, *Journal of Glaciology*, 62, 1021–1029, 2016b.
- Levintanus, A.: Saving the Aral Sea, *International journal of water resources development*, 8, 60–64, 1992.
- Li, H., Ng, F., Li, Z., Qin, D., and Cheng, G.: An extended “perfect-plasticity” method for estimating ice thickness along the flow line of mountain glaciers, *Journal of Geophysical Research: Earth Surface*, 117, 2012.
- Linsbauer, A., Paul, F., and Haeberli, W.: Modeling glacier thickness distribution and bed topography over entire mountain ranges with GlabTop: Application of a fast and robust approach, *Journal of Geophysical Research: Earth Surface*, 117, 2012.
- Millan, R., Mouginot, J., Rabatel, A., and Morlighem, M.: Ice velocity and thickness of the world's glaciers, *Nature Geoscience*, 15, 124–129, 2022.
- Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Ben Dhia, H., and Aubry, D.: A mass conservation approach for mapping glacier ice thickness, *Geophysical Research Letters*, 38, 2011.
- Mountain Research Initiative EDW Working Group: Elevation-dependent warming in mountain regions of the world, *Nature climate change*, 5, 424–430, 2015.
- Neto, F. D. M. and da Silva Neto, A. J.: An introduction to inverse problems with applications, Springer Science & Business Media, 2012.
- Oerlemans, J.: The atmospheric boundary layer over melting glaciers, *VNE*, 48, 129–153, 1998.
- Oliver, M. and Webster, R.: A tutorial guide to geostatistics: Computing and modelling variograms and kriging, *Catena*, 113, 56–69, 2014.
- Osborn, T., Jones, P., Lister, D., Morice, C., Simpson, I., Winn, J., Hogan, E., and Harris, I.: CRUTEM5.0.0.0: Climatic Research Unit (CRU) gridded near-surface air temperature anomalies over land, Centre for Environmental Data Analysis, URL <https://catalogue.ceda.ac.uk/uuid/901f576daca4e049630ab879d6fb476>, 2020.

- Pattyn, F.: A new three-dimensional higher-order thermomechanical ice sheet model: Basic sensitivity, ice stream development, and ice flow across subglacial lakes, *Journal of Geophysical Research: Solid Earth*, 108, 2003.
- Paul, F. and Linsbauer, A.: Modeling of glacier bed topography from glacier outlines, central branch lines, and a DEM, *International Journal of Geographical Information Science*, 26, 1173–1190, <https://doi.org/10.1080/13658816.2011.627859>, 2012.
- Plewes, L. A. and Hubbard, B.: A review of the use of radio-echo sounding in glaciology, *Progress in Physical Geography: Earth and Environment*, 25, 203–236, <https://doi.org/10.1177/030913330102500203>, 2001.
- Pritchard, H. D.: Asia's shrinking glaciers protect large populations from drought stress, *Nature*, 569, 649–654, 2019.
- RGI 7.0 Consortium: Randolph Glacier Inventory - A Dataset of Global Glacier Outlines, Version 7.0, <https://doi.org/10.5067/f6jmovy5navz>, <https://doi.org/10.5067/f6jmovy5navz>, 2023.
- Seguinot, J., Ivy-Ochs, S., Jouvet, G., Huss, M., Funk, M., and Preusser, F.: Modelling last glacial cycle ice dynamics in the Alps, *The Cryosphere*, 12, 3265–3285, <https://doi.org/10.5194/tc-12-3265-2018>, 2018.
- Sehring, J.: Irrigation reform in Kyrgyzstan and Tajikistan, *Irrigation and Drainage Systems*, 21, 277–290, 2007.
- Solomina, O., Barry, R., and Bodnya, M.: The retreat of Tien Shan glaciers (Kyrgyzstan) since the Little Ice Age estimated from aerial photographs, lichenometric and historical data, *Geografiska Annaler: Series A, Physical Geography*, 86, 205–215, 2004.
- TensorFlow: TensorFlow: An end-to-end platform for machine learning, <https://www.tensorflow.org/>, accessed: (28/12/2024), 2024.
- The Observatory of Economic Complexity: The Observatory of Economic Complexity (OEC): Kyrgyzstan Profile, URL <https://oec.world/en/profile/country/kgz>, accessed on March 24, 2024, Accessed 2024.
- van Pelt, W. and Frank, T.: New glacier thickness and bed topography maps for Svalbard, *The Cryosphere*, 19, 1–17, 2025.
- Van Tricht, L. and Huybrechts, P.: Modelling the historical and future evolution of six ice masses in the Tien Shan, Central Asia, using a 3D ice-flow model, *The Cryosphere*, 17, 4463–4485, 2023.
- Van Tricht, L., Huybrechts, P., Van Breedam, J., Fürst, J. J., Rybak, O., Satylkanov, R., Ermenbaiev, B., Popovnin, V., Neyns, R., Paice, C. M., and et al.: Measuring and inferring the ice thickness distribution of four glaciers in the Tien Shan, Kyrgyzstan, *Journal of Glaciology*, 67, 269–286, <https://doi.org/10.1017/jog.2020.104>, 2021a.

- Van Tricht, L., Paice, C. M., Rybak, O., Satylkanov, R., Popovnin, V., Solomina, O., and Huybrechts, P.: Reconstruction of the historical (1750–2020) mass balance of Bordu, Kara-Batkak and Sary-Tor Glaciers in the inner Tien Shan, Kyrgyzstan, *Frontiers in Earth Science*, 9, 734802, 2021b.
- Van Tricht, L., Paice, C. M., Rybak, O., Popovnin, V., Satylkanov, R., and Huybrechts, P.: Monitoring the annual geodetic mass balance of Bordu and Sary-Tor glaciers using UAV data, *Annals of Glaciology*, pp. 1–10, 2023a.
- Van Tricht, L., Zekollari, H., Huss, M., Farinotti, D., and Huybrechts, P.: Global vs local glacier modelling: a comparison in the Tien Shan, *The Cryosphere Discussions*, 2023, 1–25, 2023b.
- Zekollari, H., Huybrechts, P., Fürst, J., Rybak, O., and Eisen, O.: Calibration of a higher-order 3-D ice-flow model of the Morteratsch glacier complex, Engadin, Switzerland, *Annals of Glaciology*, 54, 343–351, <https://doi.org/10.3189/2013AoG63A434>, 2013.
- Zekollari, H., Fürst, J. J., and Huybrechts, P.: Modelling the evolution of Vadret da Morteratsch, Switzerland, since the Little Ice Age and into the future, *Journal of Glaciology*, 60, 1155–1168, <https://doi.org/10.3189/2014JoG14J053>, 2014.
- Zekollari, H., Huss, M., Farinotti, D., and Lhermitte, S.: Ice-dynamical glacier evolution modeling—a review, *Reviews of Geophysics*, 60, e2021RG000754, 2022.

Appendices

Appendix A

Validation interpolation methods

Table A.1: Performance metrics for different interpolation methods and training sets: the RMSE, the Moran's I, the KS test statistic and the p-value of that KS. The p-values tested to be significant at a significance level of 0.05 are coloured.

Interpolation Method	Training Set	RMSE	Moran's I	Kolmogorov–Smirnov Test Statistic	Kolmogorov–Smirnov p-value
Kriging	Top 50%	63,87	0,81	0,16	0,01
Kriging	Bottom 50%	74,22	1,94	0,20	0,00
Kriging	Random 50%	13,54	0,07	0,11	0,11
Kriging	Random 90%	14,46	0,01	0,08	0,41
Kriging	Central Flowline	41,04	0,40	0,14	0,03
Kriging	All points	15,52	0,02	0,09	0,28
ANUDEM_ice_thickness	Top 50%	64,63	1,49	0,19	0,00
ANUDEM_ice_thickness	Bottom 50%	72,09	1,60	0,19	0,00
ANUDEM_ice_thickness	Random 50%	15,81	0,34	0,13	0,05
ANUDEM_ice_thickness	Random 90%	11,73	0,16	0,14	0,03
ANUDEM_ice_thickness	Central Flowline	40,93	0,04	0,16	0,01
ANUDEM_ice_thickness	All points	11,03	0,17	0,12	0,08
ANUDEM_bedrock	Top 50%	64,75	1,52	0,19	0,00
ANUDEM_bedrock	Bottom 50%	72,92	1,57	0,19	0,00
ANUDEM_bedrock	Random 50%	15,56	0,31	0,13	0,06
ANUDEM_bedrock	Random 90%	11,74	0,16	0,14	0,03
ANUDEM_bedrock	Central Flowline	40,72	0,04	0,16	0,01
ANUDEM_bedrock	All points	11,23	0,16	0,13	0,06
IGM_optimize	Top 50%	24,06	0,52	0,06	0,84
IGM_optimize	Bottom 50%	34,64	1,72	0,10	0,23
IGM_optimize	Random 50%	25,49	0,55	0,06	0,88
IGM_optimize	Random 90%	26,22	0,77	0,07	0,73
IGM_optimize	Central Flowline	30,76	1,60	0,11	0,15
IGM_optimize	All points	29,73	1,02	0,07	0,70
IGM_optimize_velocity	Top 50%	26,90	0,88	0,05	0,95
IGM_optimize_velocity	Bottom 50%	34,97	1,68	0,09	0,30
IGM_optimize_velocity	Random 50%	28,33	0,93	0,05	0,90
IGM_optimize_velocity	Random 90%	28,99	1,13	0,07	0,67
IGM_optimize_velocity	Central Flowline	32,12	1,61	0,09	0,29
IGM_optimize_velocity	All points	31,81	1,25	0,07	0,65
Yieldstress	Top 50%	49,04	0,68	0,11	0,11
Yieldstress	Bottom 50%	38,10	0,12	0,13	0,04
Yieldstress	Random 50%	22,54	0,13	0,10	0,26
Yieldstress	Random 90%	17,28	0,00	0,09	0,35
Yieldstress	Central Flowline	34,46	0,04	0,10	0,19
Yieldstress	All points	13,95	0,03	0,05	0,96

Appendix B

Sensitivity analysis spatial interpolation methods

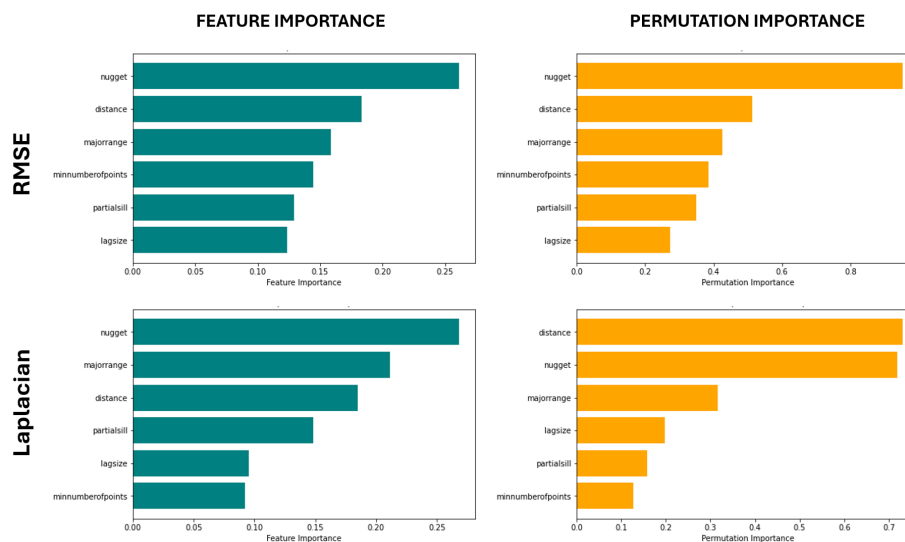


Figure B.1: Parameter importance of the kriging parameters. The green barplots represent the feature importance for the RMSE (top) and the Laplacian (bottom). The yellow barplots represent the permutation importance for the same two indicators.

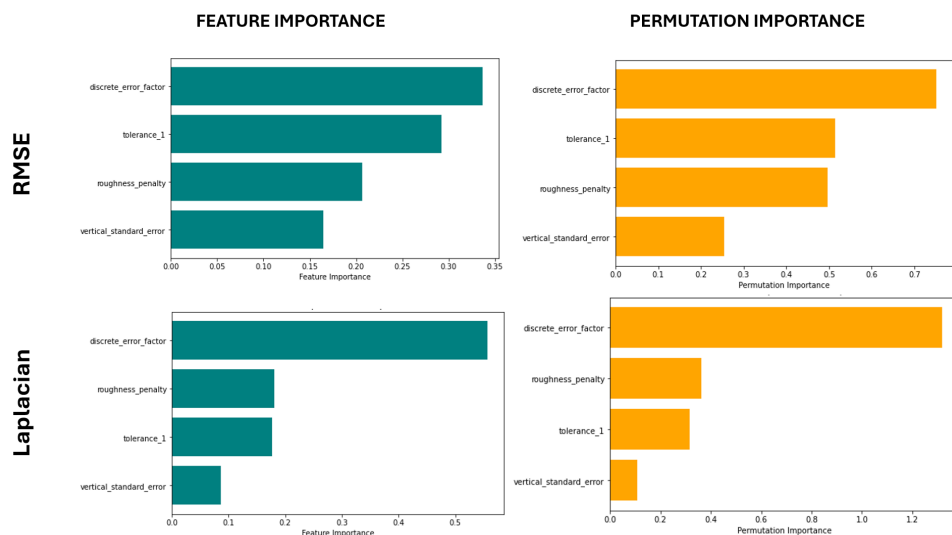


Figure B.2: Parameter importance of the ANUDEM parameters. The green barplots represent the feature importance for the RMSE (top) and the Laplacian (bottom). The yellow barplots represent the permutation importance for the same two indicators.

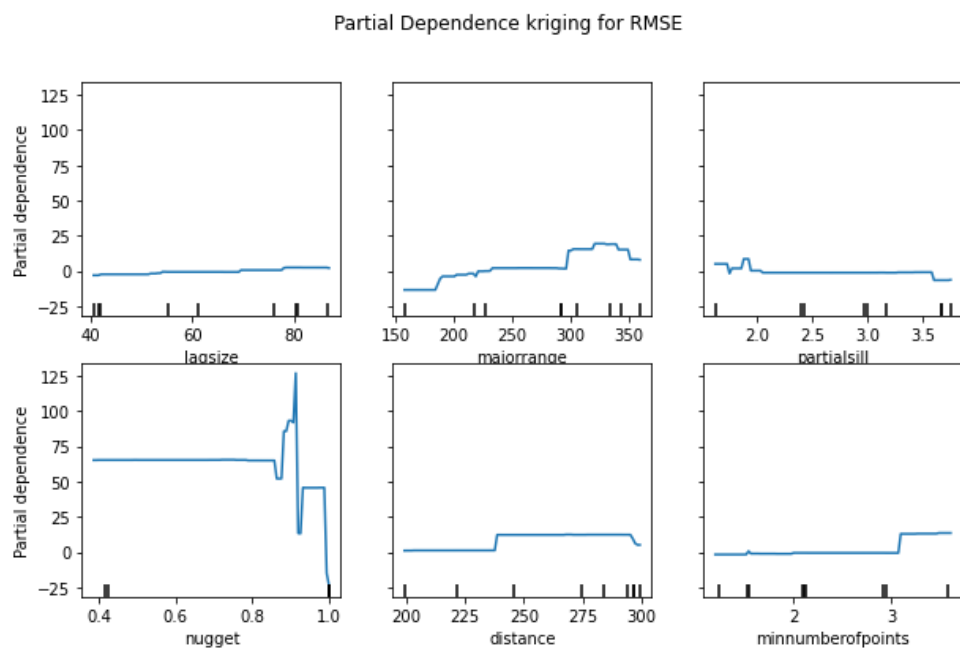


Figure B.3: Partial dependence plots of the kriging parameters for the RMSE. Lower values correspond to better model performance (see methods section).

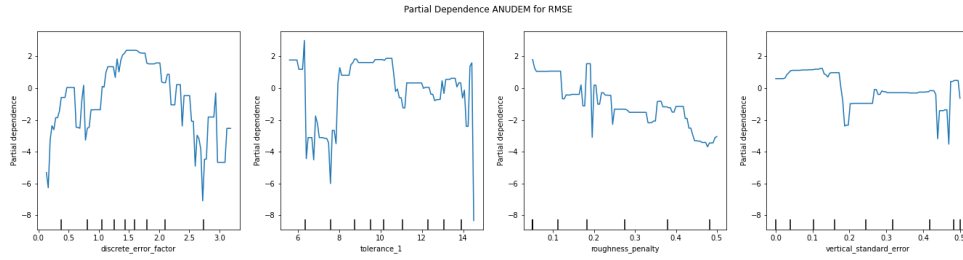


Figure B.4: Partial dependence plots of the ANUDEM parameters for the RMSE. Lower values correspond to better model performance (see methods section).

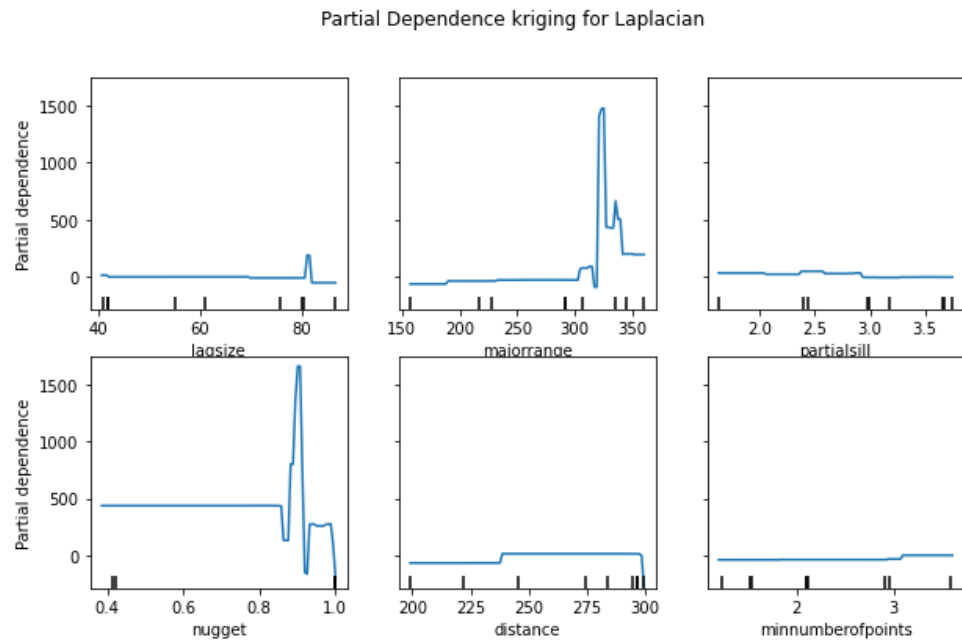


Figure B.5: Partial dependence plots of the kriging parameters for the Laplacian. Lower values correspond to better model performance (see methods section).

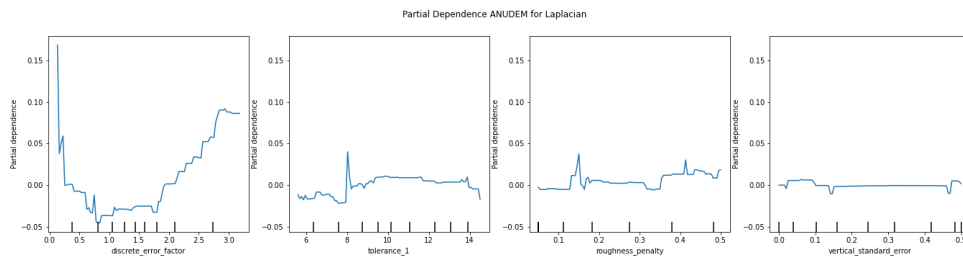


Figure B.6: Partial dependence plots of the ANUDEM parameters for the Laplacian. Lower values correspond to better model performance (see methods section).

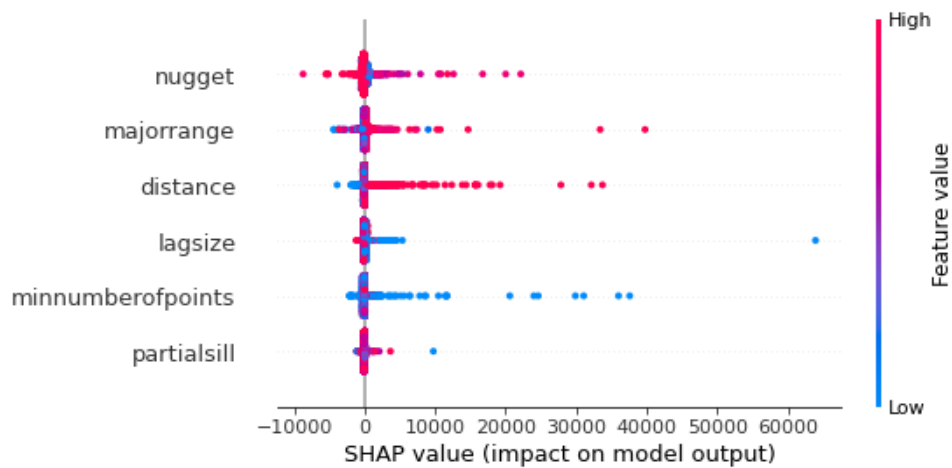


Figure B.7: Summary plot SHAP values for the kriging interpolations. The vertical axis ranks the parameters by importance, from most to least influential. In the plot, low parameter values are coloured blue, and high values are coloured red. A negative SHAP value indicates that the parameter value pushes the prediction lower; a positive SHAP value means it pushes the prediction higher.

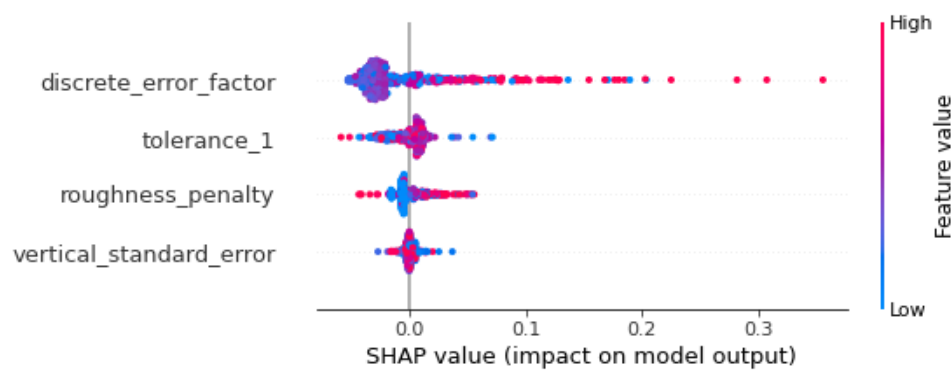


Figure B.8: Summary plot SHAP values for the ANUDEM interpolations. The vertical axis ranks the parameters by importance, from most to least influential. In the plot, low parameter values are coloured blue, and high values are coloured red. A negative SHAP value indicates that the parameter value pushes the prediction lower; a positive SHAP value means it pushes the prediction higher.

Afdeling geografie
Celestijnenlaan 200E
3001 Heverlee (Leuven), BELGIË
tel. + 32 16 00 00 00
fax + 32 16 00 00 00
www.kuleuven.be

