Ice Sheet Speed-dating: Using expert judgement to identify "good" simulations of the Last Glacial Maximum North American Ice Sheets

GANDY, Niall, ASTFALCK, Lachlan C., IVES, Gemma L. and RIVERS, Gwyneth

Quaternary Science Reviews 333 (2024) 108690

Contents lists available at ScienceDirect

Quaternary Science Reviews

journal homepage: www.elsevier.com/locate/quascirev

Short communication

Ice sheet speed-dating: Using expert judgement to identify "good" simulations of the last glacial maximum North American ice sheets

Niall Gandy[a,*], Lachlan C. Astfalck[b,c], Gemma L. Ives[d], Gwyneth E. Rivers[a]

[a] Department of Natural and Built Environment, Sheffield Hallam University, UK
[b] Oceans' Graduate School, The University of Western Australia, Australia
[c] School of Physics, Mathematics and Computing, The University of Western Australia, Australia
[d] Computing Services, The University of Sheffield, UK

ARTICLE INFO

Handling Editor: Qiuzhen Yin

Keywords:
Glaciology
Data analysis
Expert judgement

ABSTRACT

After generating a large ensemble of palaeo ice sheet model runs, it is common to either rank the simulations, or classify each simulation as an acceptable match to observations or not. Both tasks require implicit human judgement, usually left to the discretion of the research authors. For instance, even numerical comparisons to reconstructions require human input on values for match thresholds and allowances for model-mismatch. We embrace the subjectivity of human judgement and calibrate an ice-sheet model by explicitly asking ~100 experts to identify simulations that are good enough. Expert judgement is made based on a set of features of the model output that is of interest (for example, margin shapes and regional ice volumes); where possible we also record such knowledge. By seeking the input of many experts, we obtain a community consensus that can be used to develop guidance to determine the quality of future simulations. This short communication describes our exercise in seeking expert classifications of simulations of the Last Glacial Maximum (LGM) North American Ice Sheets, discusses the lessons learnt, and suggests future analysis of the responses.

1. Introduction

Palaeo ice sheet modelling studies often produce large ensembles of simulations to properly explore the uncertain parameter space, in search of simulations that are a reasonable match to empirical observations and to better understand the potential range of ice sheet behaviour (DeConto and Pollard, 2016; Gregoire et al., 2016; Gandy et al., 2021). This ensemble approach produces many simulations which are a poor match to observations, and some which are a better match (Fig. 1). To determine which simulations are sufficiently accurate (be that for further analysis, or for feeding into another model experiment or model) simulations can be compared against empirical data constraining the size, shape, and behaviour of past ice sheets. However, it is a challenge to compare multiple simulations to the wealth of empirical data, including deglaciation ages (Small et al., 2017; Dalton et al., 2020), flowsets (Boulton and Clark, 1990; Hughes et al., 2014), margin reconstructions (Dalton et al., 2020; Clark et al., 2022), and relative sea level curves (Dyke et al., 2005; Argus and Peltier, 2010; Vacchi et al., 2018). All these reconstruction methods include their own uncertainties, a variety of sub-methods, and in some cases contested interpretation. Further complicating matters, guidance on quantifying acceptable model-mismatch for each reconstruction method is sparse, and highly uncertain. In a pioneering paper, Andrews (1982) proposed that ice sheet reconstructions would be strengthened by combining and reconciling model and empirical evidence. Numerical ice sheet modelling was in its infancy, but despite four decades of progress, the full integration of model and empirical data has proved difficult to implement.

A simple approach to model-data comparison is to compare model features, such as ice sheet volume and area, to data-based reconstructions; for example, in Gregoire et al. (2016) and Gandy et al. (2023). It is also not certain that the smallest and largest existing reconstructions are the actual bounds of physical plausibility, so the ranges need to be extended accordingly. This problem is more acute when comparing palaeo ice sheet simulations to a reconstructed area. For example, the Last Glacial Maximum (LGM) footprint of the North American Ice Sheets were reconstructed with only limited uncertainty by Dyke et al. (2002), and further refined by Dalton et al. (2020). The precision of this reconstruction is difficult for a freely-evolving ice sheet model to reproduce, so we need a judgement of where and how much deviation from the reconstruction is tolerable. As an extreme example, a

* Corresponding author. Department of Natural and Built Environment, Sheffield Hallam University, Sheffield, S1 1WB, UK
E-mail address: n.gandy@shu.ac.uk (N. Gandy).

https://doi.org/10.1016/j.quascirev.2024.108690
Received 21 February 2024; Received in revised form 26 April 2024; Accepted 27 April 2024
Available online 3 May 2024

simulation that is perfectly matched to the southern margin of the LGM Laurentide Ice Sheet, but with a doughnut-like hole in the ice sheet, would be less plausible than a poorly matched southern margin with contiguous ice in the interior, even if the absolute error compared to the reconstruction was comparable. Recent reconstructions of the Eurasian (Hughes et al., 2016) and North American (Dalton et al., 2023) ice sheets include a maximum and minimum plausible extent, which partly helps address comparison uncertainty. In Gregoire et al. (2016) and Gandy et al. (2023), the error between each simulation and the reconstruction area was calculated, and used to rank the ensemble. An error cut-off between implausible and plausible simulations was decided by the authors.

There are more sophisticated methods to compare models to empirical data. For example, reconstructed and expected relative sea level curves can be compared to constrain the spatial distribution of ice mass. This is a developed route of model-data intercomparison e.g. (Walcott, 1972; Peltier et al., 1978; Quinlan and Beaumont, 1982), and there is a wealth of data to compare to. To assist with model-data comparison, tools have also been developed to compare model outputs to palaeo ice sheet geomorphology. The Automated Proximity and Conformity Analysis (APCA) tool has been developed to compare mapped moraine positions to modelled margins (Napieralski et al., 2006; Li et al., 2008). This tool computes the similarity in shape of modelled ice margins and moraines, identifying occurrences for each moraine when both the proximity and conformity is below some acceptance threshold. The Automated Flow Direction Analysis (AFDA) tool follows a similar principle, computing the mismatch in flow direction between modelled and empirically mapped flowsets (Li et al., 2007), and the recently developed Likelihood of Accordant Lineations Analysis (LALA) tool uses a statistical methodology to consider variability in empirical mapping techniques (Archer et al., 2023).

Finally, geochronology has been compared to ice sheet models using the Automated Timing Accordance Tool (ATAT) (Ely et al., 2019). The tool considers an empirical age constraint, and identifies all instances that an age is not conflicted by the model outputs. Unlike other tools, this does not require any set acceptance thresholds, instead calculating the percentage of ages which do not conflict with the model output. APCA, AFDA, and ATAT have all been used to rank ensemble members of palaeo ice sheet simulations (Ely et al., 2021; Gandy et al., 2021). However, there are some recurring shortcomings of these tools, such as the need to set acceptance thresholds. This returns us to the human-judgment element of "how good is good enough?".

1.1. Our approach

Here, we explore the use of expert judgement to identify "good" simulations of palaeo ice sheets in a large ensemble. Expert elicitation is a well cited and often used method that quantifies probabilistic beliefs from subject matter experts (Astfalck et al., 2018; O'Hagan, 2019). However, rigorous expert elicitation for even the simplest quantities is exacting. We instead treat experts as a classifier, or data-generating source, and capture expert beliefs through modelling. Expert opinion

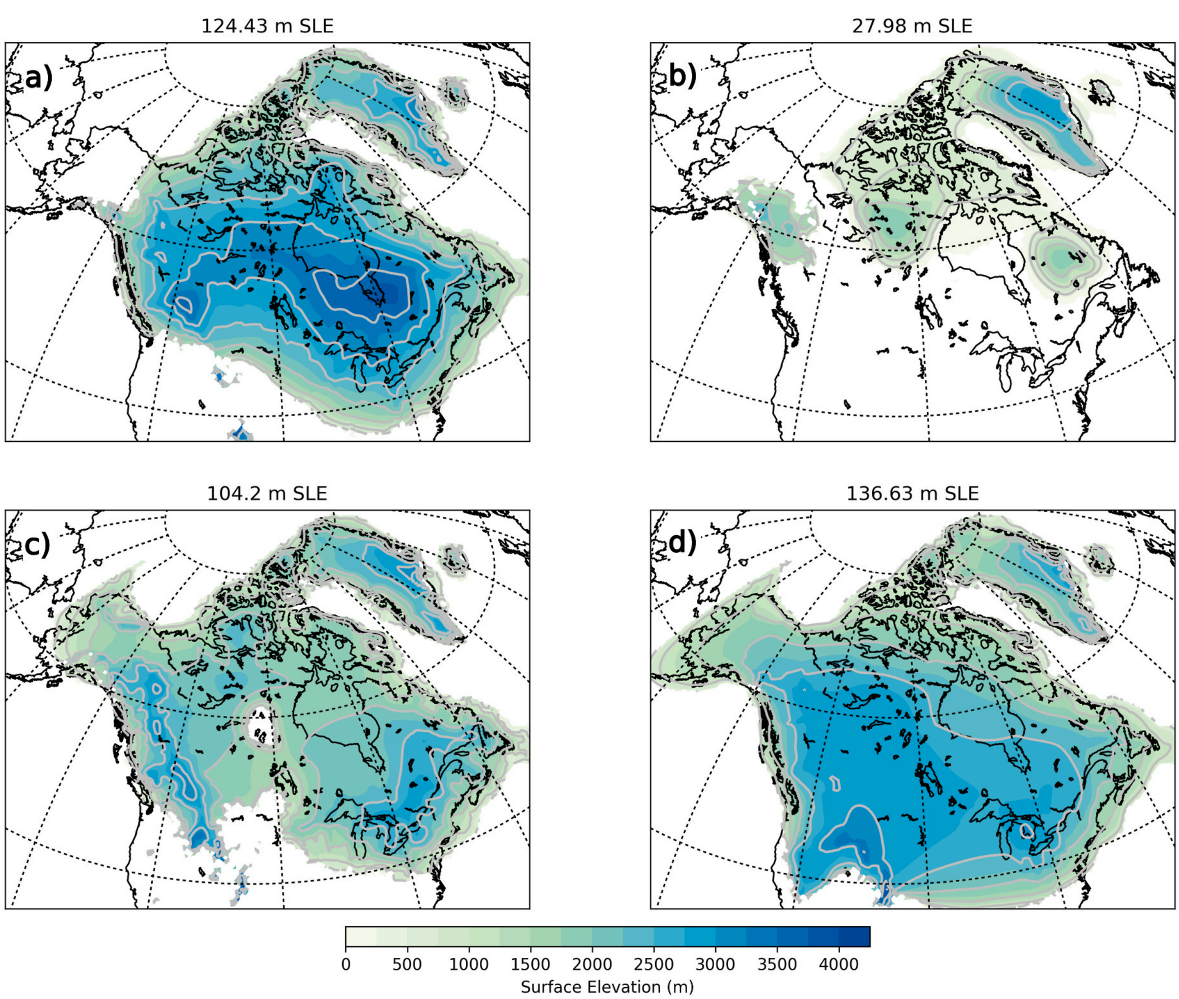

Fig. 1. The good, the bad, and the ugly. A range of simulations of the LGM North American Ice Sheets, from Gregoire et al. (2016) and Gandy et al. (2023), including a potentially reasonable simulation (a), a simulation with restricted ice (b), a glaciologically or climatologically implausible simulations (c), and a simulation with extensive ice (d).

has been used within glaciology to identify slush from remote sensing imagery (Dell et al., 2022), to identify plausible future ice sheet projections (Bamber and Aspinall, 2013), and to map fields of subglacial linneations (Archer et al., 2023). We argue that this method can be added to the existing arsenal of tools described above.

2. Methods

We sought to acquire a database of categorisations, classified as either “good enough” or “not good enough” for an existing ensemble of simulations of the LGM North American Ice Sheets (the combined Laurentide, Innuitian, and Cordilleran Ice Sheet complex) from Gregoire et al. (2016) and Gandy et al. (2023). The LGM North American Ice Sheets are well-constrained by empirical data (e.g. Pico et al., 2017; Gowan et al., 2021), and has been extensively modelled (e.g. Tarasov and Peltier, 2007; Gowan et al., 2016), compared to the Eurasian Ice Sheet or other time periods. The definition of “good enough” will vary between participants, depending on individual requirements for ice sheet simulation usage. This diversity of opinion helps us understand two things: first, the variability in community opinion for a simulation to be classified as “good enough”; and second, what the standard communitywide “good enough” simulation looks like. We received enough responses to categorise each ensemble member multiple times, allowing for some indication of the response variability. Although, at first, this may seem a crude method of calibration, we show in the responses that the community recovered many known important features.

All simulations are of the LGM ice sheet, with 567 simulations from Gregoire et al. (2016) which are forced with climate simulations of 21 ka BP using climate model output. There are also 347 simulations from Gandy et al. (2023), which are from coupled climate-ice sheet simulations using the Glimmer ice sheet model and the FAMOUS General Circulation Model running with PMIP4 LGM boundary conditions and a statistical reconstruction of ocean conditions (Astfalck et al., 2024). Each simulation was displayed to the user showing the simulated ice sheet shape, and the total ice sheet volume in sea level equivalent for the whole domain. An example of how the simulations are displayed is shown in Fig. 2. The plots do not include any information about the empirical constraints, for example by showing the Dyke et al. (2002) margin. This is to avoid bias in the results towards any specific information displayed on the plots.

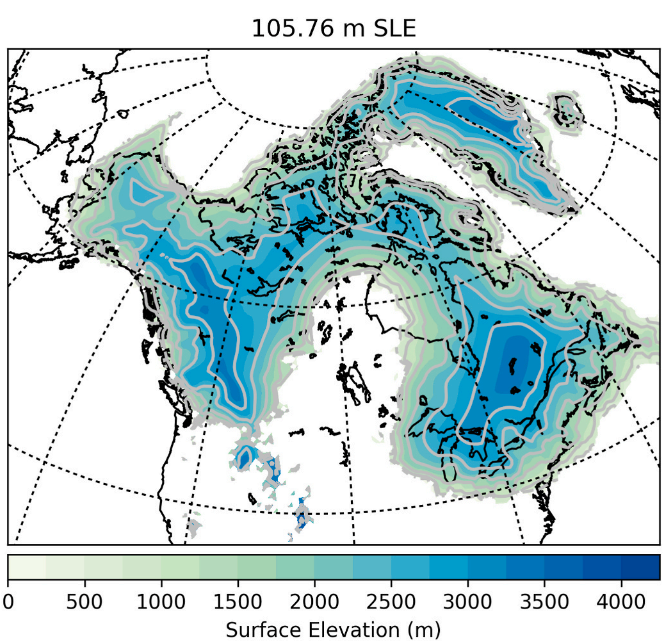

Fig. 2. An example LGM ice sheet simulation which could pass some simple metrics, with a reasonable ice volume and area, but a glaciologically and climatologically implausible southern margin.

We use the citizen science web portal Zooniverse to host the categorisation workflow. Zooniverse has previously been used to categorise galaxies (e.g. Weisz et al., 2015; Lingard et al., 2021), cyclone centres (Knapp et al., 2016), and to annotate historical documents e.g. (Williams et al., 2014; Grayson, 2016; Blickhan et al., 2019). Zooniverse displays a random simulation plot (such as Fig. 2), and users can click if the simulation is “good enough” or “not good enough”. This response is saved, and a new randomised plot is shown. Each new plot is chosen at random from the full sample of 914, so can be categorised multiple times. We collected the majority of the data in person during conference poster sessions, allowing conversations which collected soft-knowledge from the experts on what features of the ice sheets they were explicitly looking at. No other user information is retained and users are not required to log on or provide any personal information. A user can continue categorising in perpetuity; there is no set number of simulations to categorise. We encouraged users to spend 5–10 min categorising, which would equate to roughly 60–120 individual categorisations.

Participation in the classification was advertised on social media and through the subject-related mailing list “Cryolist”, including a link to participate. We also advertised the project through a poster presentation at both EGU and INQUA 2023, where a tablet was set up for participation during the poster presentation, and a link was provided so attendees could contribute in their own time from their own device. Just less than half of all the ice sheet classifications were made during the weeks of EGU and INQUA.

3. Results and discussions

From February to August 2023, we received 3396 responses, with 1220 during the week of EGU and 195 during the week of INQUA. The classifications were made across 107 sessions, suggesting the participation of ~100 people. Of the 3393 responses, 94.8% of the results were negative, which is consistent with most ensemble simulations being ruled out when they were previously assessed with more traditional methods in Gregoire et al. (2016) and Gandy et al. (2023). As the displayed simulations were selected randomly, 1 simulation was never categorised, and 60 only had one categorisation. Each simulation had an average of 3.7 responses.

By comparing the set of positive responses to the entire set of simulations, it is possible to determine some features of classification that lead to positive responses. Fig. 3 shows, for each gridbox, the percentage of simulations with ice cover for the entire simulation set, compared to just the positive responses. This shows that the positive response set is much more extensive, on average, with ice extending down towards the expected southern margin (Dyke et al., 2002; Dalton et al., 2020). There is also no increase in Alaskan ice, despite larger simulations in the ensembles typically having extensive Alaskan ice cover, showing that responses favoured simulations without extensive Alaskan ice.

3.1. Analysis of the classifications

The collected categorisations only contain responses for ice sheet simulations that have been presented; without further analysis they do not immediately help us determine if any future simulation of the LGM North American Ice Sheets would be acceptable. Accordingly, we seek to find trends in the responses from which we can construct generalised rules to replicate the results of the community categorisation. As we have laboured, we do not recommend this to replace any of the existing calibration methodologies, but to be used in conjunction.

First, we look at the results as a function of the total whole-domain volume and footprint area of each simulation. This is shown in the top plots of Fig. 4; a “good enough” response is marked by blue and “not good enough” by red. This whole domain approach is a simple metric, previously used by Gandy et al. (2023). It shows that simulations with a

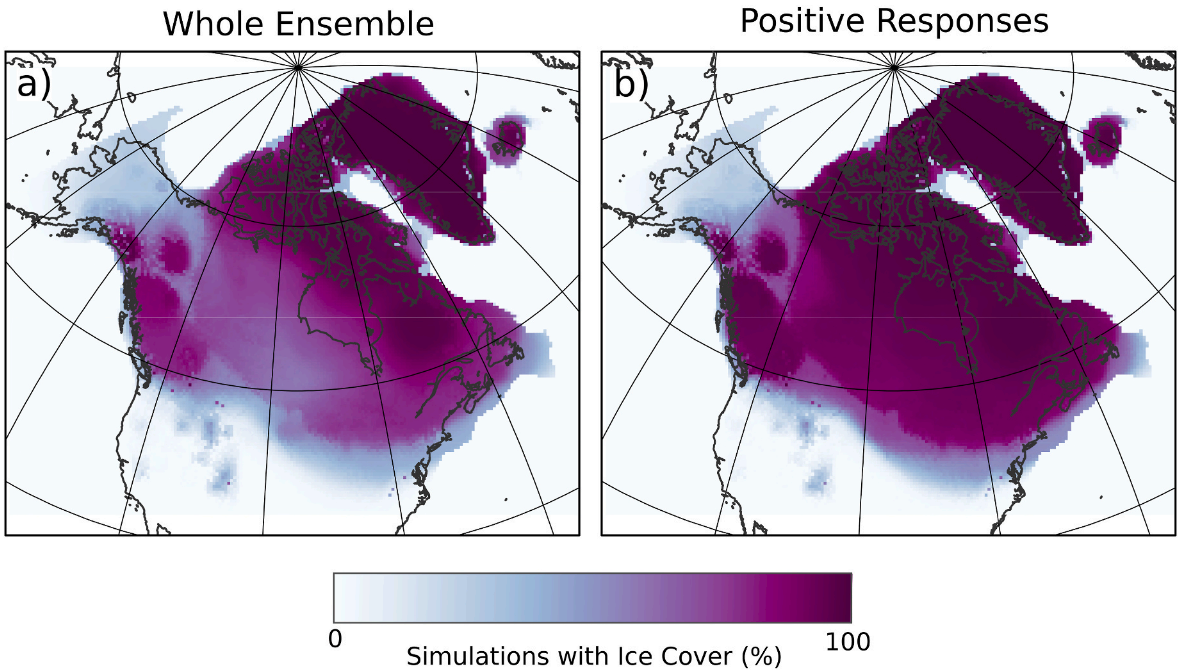

Fig. 3. The percentage of simulations with ice cover for each grid box, for all simulations (a) and all positive responses (b).

higher volume and area are more likely to be categorised as "good enough", with many simulations categorised this way once the volume is above 3 × 10^7km^3 and the area is above 1.5 × 10^7km^2. This region contains 72.1% of the positive responses and rejects 61.6% of the negative ones, but still is only 30% positive. This is likely the result of the fact that ice sheet simulations with a very similar total volume and area can have very different configurations and such a simple rule is unable to capture this sophistication. An example of an unreasonable simulation which has a similar volume and area to a reasonable simulation is shown in Fig. 2. The variability will also reflect the inherent variability of the community engagement, with individual contributors having different approaches to identifying good simulations. Overall, this whole domain volume and area approach, alone, has limited ability in helping to identity if a simulation is likely to be favourably categorised by the community, but it can certainly identify when a simulation is very unlikely to be categorised favourably.

We asked contributors what explicit features they were looking for to help identify a good simulation. Two main responses were the extent of the southern margin by itself, and the extent of the southern margin without overextending onto other regions such as Alaska. The distribution of good simulations of the southern margin's volume and area is shown in the middle plots of Fig. 4. If we just examine simulations with a Southern Area above 1.5 × 10^4km^2 and a Southern Volume above 0.5 × 10^7km^3, we would include 73.2% of all positive responses, and exclude 52.2% of all negative response. This is not a significant improvement on using the entire ice sheet volume and area. Next, the southern volume is compared to the Alaskan volume in the bottom plots of Fig. 4. The extent of the Alaskan sector (Fig. 4e) is kept broad to test for simulations where a limited Alaskan ice extent is coincident with a deglaciated northern Cordilleran Ice Sheet. Interestingly, when examining the interplay of the southern margin and Alaska, three distinct behaviour patterns are apparent. The top left cluster has very few positive responses. The middle cluster simulations, where the southern ice volume is above 1 × 10^7km^3 and Alaskan volume is between 0.4 and 0.8 × 10^7km^3 have a high chance of being categorised as "good enough". Finally, the bottom cluster shows a similar pattern as in the domain-wide analysis, where ice-sheet volumes above 0.5 × 10^7km^3 result in a mixed response.

If we apply the constraint of all three metrics together we capture 46.6% of the positive responses, and reject 69.5% of the negative ones. Note, this does not perfectly calibrate our model to a user categorisation, but rather provides some notion similar to not-ruled-out-yet space in history matching (Williamson et al., 2013). Our analysis of the responses is a relatively rudimentary first attempt at presenting these results. As demonstrated by looking at the interplay between the southern volume and the Alaskan volume, more complicated studies of model features could provide tighter calibration. We expect as we continue to explore increasingly complex metrics of ice sheet configuration, informed by statistical and machine learning feature selection methods, an even stronger classification signal will emerge.

3.2. Considerations

A shortcoming of this approach is that it does not directly consider the wealth of empirical data available on palaeo ice sheets (e.g. Dalton et al., 2020). This empirical data has been hard-won as part of large innovative projects (for example, Clark et al., 2022), and there are examples of model ensembles being robustly compared to this data. Where comparisons to empirical data may fail is when it is difficult to judge and compare the complex geometry of the ice sheets; a total volume and extent may be reasonable, with an implausible mass distribution, for example. Of course, the categorisations of contributors are made considering their implicit knowledge of the empirical data, synthesised into a general view and assessment, but this generalisation can exclude some nuance in the data (such as the non-synchronous maximum extent of the North American Ice Sheets), contain a lack of information, or could include personal biases. Applicability of this method to other ice sheets would need to consider the higher uncertainty in the empirical data, and potentially a resulting greater spread of "good enough" responses.

The most common question we were asked at the EGU and INQUA poster sessions was a derivative of "am I enough of an expert to help?" Our approach was to encourage anyone who was interested in the project and the LGM North American Ice Sheets to contribute, in contrast to usual expert elicitation studies where the expert group is selected and relatively small. This open approach was motivated by the wish for a range of contributions, reflecting the wide range of researchers who may use the output of ice sheet models. We also wanted to avoid acting as an arbiter of who qualifies as an expert in the field. A

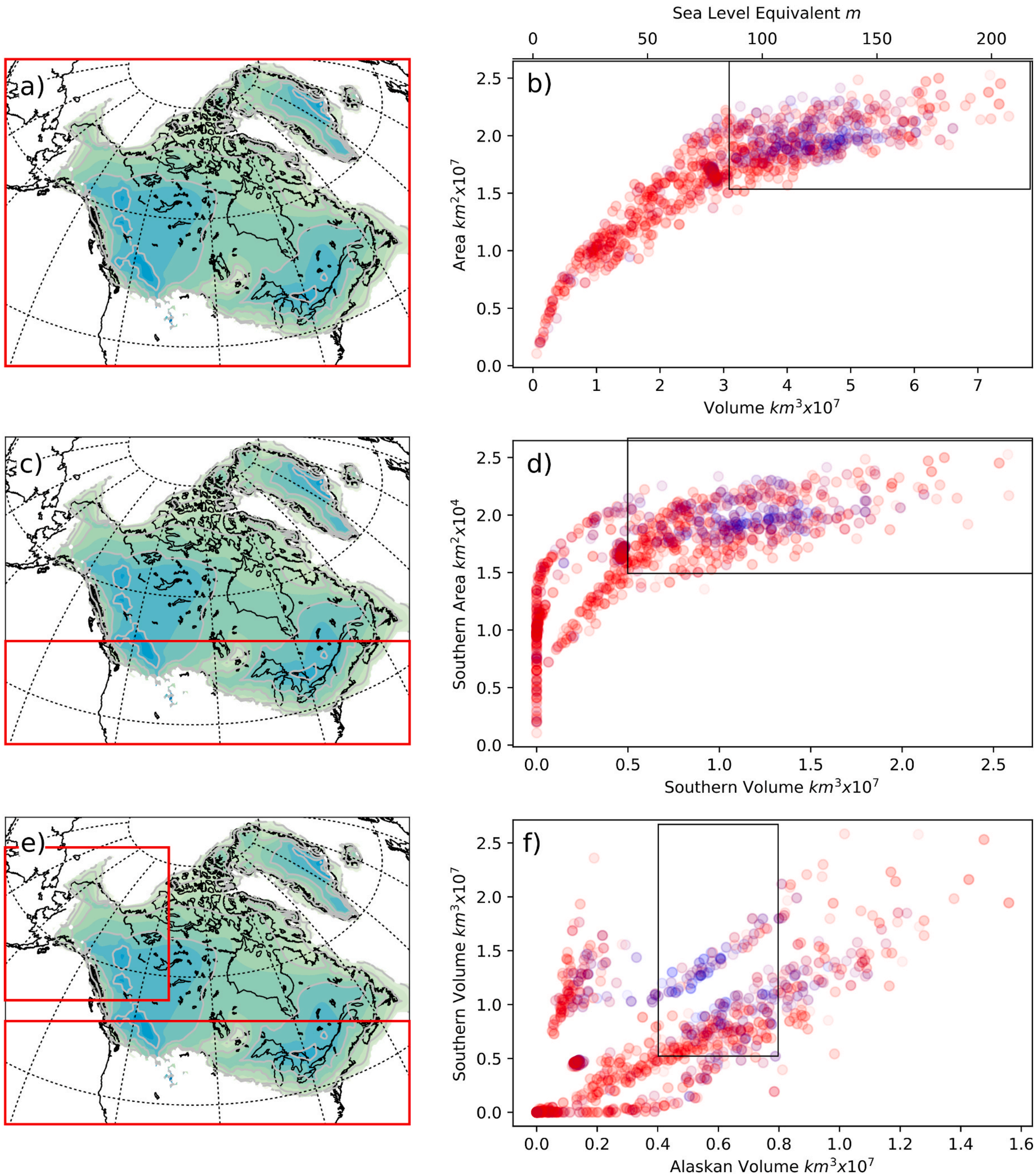

Fig. 4. An attempt to identify simulations likely to be categorised as "good enough" using simple volume and footprint area categorisations. The extents used to calculate volume and area are shown in panels a, c, and e, with the same elevation colormap as Figs. 1 and 2 b) The volume and area of the whole domain, d) The volume and area of the southern margin, f) The volume of the southern margin and the Alaskan sector. Red points categorised as "not good enough", blue points categorised as "good enough". The domains used to compute volume and footprint area are shown in red on the corresponding maps.

challenge is that there were certainly contributors who had limited knowledge of the LGM North American Ice Sheets. Our view is that each simulation was, on average, classified multiple (3.7) times, protecting against this issue. We also note that contributors who were highly engaged in the research spent much more time classifying simulations than contributors with a passing interest. Still, this range of expertise in the contributors will likely have contributed to some variability seen in the results in Fig. 4.

3.3. Future developments

The work presented in this manuscript is primarily a proof of concept for a much more detailed and robust statistical analysis that is being designed. We have three focuses for this analysis, that are expected to make this use of expert judgement a more powerful calibration tool. First, we will examine how the binary expert responses can be translated into the underlying belief distribution of the experts, both in the model space and in the parameter space. Next, we will look to feature selection techniques to dial in the use of this data as a classifier for future model results. Modern and sophisticated machine learning techniques increasingly have the capability to describe very sophisticated, and non-linear, features of model output, and we expect such techniques to work well in this application. Finally, we will develop theory for adaptive sequential design, to maximise the utility of expert response. As we have mentioned, 94.8% of responses were negative, and so only $\sim 1/20^{th}$ of the experts' decisions were used to positively classify an ice sheet. Rather than naively randomly sampling simulations, we can use adaptive sequential design methods to inform what models we present to the experts to maximise the information gain for a fixed number of responses. Additional surveys could also be completed with a non-specialist audience to help determine the level of expertise required, and the required survey size.

4. Conclusions

This study presents an innovative approach to evaluating palaeo ice sheet simulations, employing expert judgement through a citizen science framework to assess the quality of numerous model runs of the LGM North American Ice Sheets. Whilst traditional methods of model-data comparison often rely on subjective thresholds of a small number of research authors, here we compile the subjective decision of many experts. Our exploration of the experts' responses yielded valuable insights into perceptions on what requires a simulation to be "good enough". Our findings reveal the complexity in pinpointing a definitive metric for simulation quality. While certain characteristics such as total ice sheet volume and regional ice configurations offer predictive potential, the variability in assessments underscores the limitations of simple metrics in predicting community-consensus on model quality. We have offered some avenues for future work that will lead to better predictive power of a model classifier and better utilise expert resources. Future work could address the incorporation of empirical data and refine the elicitation process, aiming to establish a more nuanced, comprehensive methodology for assessing ice sheet simulations in palaeo-climate studies.

CRediT authorship contribution statement

Niall Gandy: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Funding acquisition. **Lachlan C. Astfalck:** Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing. **Gemma L. Ives:** Methodology, Software, Writing – review & editing. **Gwyneth E. Rivers:** Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The recorded responses and the plot image for each simulation are archived online (https://doi.org/10.17632/5tm6jgxhg2.1).

Acknowledgements

This research was funded with a Sheffield Hallam University Department of Natural and Built Environment Research Support Grant. Lachlan Astfalck was supported by the ARC ITRH for Transforming energy Infrastructure through Digital Engineering (TIDE), Grant No. IH200100009. Gwyneth E. Rivers is funded by a PhD GTA studentship awarded by Sheffield Hallam University. We thank Christine Batchelor and two anonymous reviewers for their constructive reviews,

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version of this paper arising from this submission.

References

Argus, Donald F., Peltier, W Richard, 2010. Constraining models of postglacial rebound using space geodesy: a detailed assessment of model ice-5g (vm2) and its relatives. Geophys. J. Int. 181 (2), 697–723.

Andrews, John T., 1982. On the reconstruction of pleistocene ice sheets: a review. Quat. Sci. Rev. 1 (1), 1–30.

Archer, R.E., Ely, Jeremy, C., Heaton, T.J., Butcher, Frances, EG., Hughes, Anna, LC., Clark, Chris, D., 2023. Assessing ice sheet models against the landform record: the likelihood of accordant lineations analysis (lala) tool. Earth Surf. Process. 48 (14), 2754–2771.

Astfalck, L.C., Cripps, E.J., Gosling, J.P., Hodkiewicz, M.R., Milne, I.A., 2018. Expert elicitation of directional metocean parameters. Ocean Eng. 161, 268–276.

Astfalck, L.C., Williamson, D., Gandy, N., Gregoire, L., Ivanovic, R., 2024. Coexchangeable process modeling for uncertainty quantification in joint climate reconstruction. J. Am. Stat. Assoc. 1–14.

Bamber, Jonathan L., Aspinall, W.P., 2013. An expert judgement assessment of future sea level rise from the ice sheets. Nat. Clim. Change 3 (4), 424–427.

Blickhan, Samantha, Coleman, Krawczyk, Hanson, Daniel, Boyer, Amy, Simenstad, Andrea, Van Hyning, Victoria, 2019. Individual vs. collaborative methods of crowdsourced transcription. J. Data Min. Digital Humanit. 2416-5999.

Boulton, G.S., Clark, C.D., 1990. The laurentide ice sheet through the last glacial cycle: the topology of drift lineations as a key to the dynamic behaviour of former ice sheets. Earth and Environmental Science Transactions of The Royal Society of Edinburgh 81 (4), 327–347.

Clark, Chris D., Ely, Jeremy C., Hindmarsh, Richard CA., Bradley, Sarah, Ignéczi, Adam, Fabel, Derek, Cofaigh, Colm Ó., Chiverrell, Richard C., Scourse, James, Benetti, Sara, et al., 2022. Growth and retreat of the last british–Irish ice sheet, 31 000 to 15 000 years ago: the britice-chrono reconstruction. Boreas 51 (4), 699–758.

Dalton, April S., Margold, Martin, Stokes, Chris R., Tarasov, Lev, Dyke, Arthur S., Adams, Roberta S., Allard, Serge, Arends, Heather E., Atkinson, Nigel, Attig, John W., et al., 2020. An updated radiocarbon-based ice margin chronology for the last deglaciation of the north american ice sheet complex. Quat. Sci. Rev. 234, 106223.

Dalton, April S., Dulfer, Helen E., Margold, Martin, Heyman, Jakob, Clague, John J., Froese, Duane G., Gauthier, Michelle S., Hughes, Anna LC., Jennings, Carrie E., Norris, Sophie L., et al., 2023. Deglaciation of the north american ice sheet complex in calendar years based on a comprehensive database of chronological data: Nadi-1. Quat. Sci. Rev. 321, 108345.

DeConto, Robert M., Pollard, David, 2016. Contribution of Antarctica to past and future sea-level rise. Nature 531 (7596), 591–597.

Dell, Rebecca L., Banwell, Alison F., Willis, Ian C., Arnold, Neil S., Halberstadt, Anna Ruth W., Chudley, Thomas R., Pritchard, Hamish D., 2022. Supervised classification of slush and ponded water on antarctic ice shelves using landsat 8 imagery. J. Glaciol. 68 (268), 401–414.

Dyke, A.S., Andrews, J.T., Clark, P.U., England, J.H., Miller, G.H., Shaw, J., Veillette, J. J., 2002. The laurentide and innuitian ice sheets during the last glacial maximum. Quat. Sci. Rev. 21 (1–3), 9–31.

Dyke, Arthur S., Dredge, Lynda A., Hodgson, Douglas A., 2005. North american deglacial marine-and lake-limit surfaces. Géographie physique et Quaternaire 59 (2), 155–185.

Ely, Jeremy C., Clark, Chris D., Small, David, Hindmarsh, Richard CA., 2019. Atat 1.1, the automated timing accordance tool for comparing ice-sheet model output with geochronological data. Geosci. Model Dev. (GMD) 12 (3), 933–953.

Ely, Jeremy C., Clark, Chris D., Hindmarsh, Richard CA., Hughes, Anna LC., Greenwood, Sarah L., Bradley, Sarah L., Gasson, Edward, Gregoire, Lauren, Gandy, Niall, Stokes, Chris R., et al., 2021. Recent progress on combining geomorphological and geochronological data with ice sheet modelling, demonstrated using the last british–Irish ice sheet. J. Quat. Sci. 36 (5), 946–960.

Gandy, Niall, Gregoire, Lauren J., Ely, Jeremy C., Cornford, Stephen L., Clark, Christopher D., Hodgson, David M., 2021. Collapse of the last eurasian ice sheet in the north sea modulated by combined processes of ice flow, surface melt, and marine ice sheet instabilities. J. Geophys. Res.: Earth Surf. 126 (4), e2020JF005755.

Gandy, Niall, Astfalck, L.C., Gregoire, L.J., Ivanovic, R.F., Patterson, V.L., Sherriff-Tadano, Sam, Smith, Robin S., Williamson, Daniel, Rigby, Richard, 2023. De-tuning albedo parameters in a coupled climate ice sheet model to simulate the north american ice sheet at the last glacial maximum. J. Geophys. Res.: Earth Surf., e2023JF007250

Gowan, Evan J., Zhang, Xu, Khosravi, Sara, Rovere, Alessio, Stocchi, Paolo, Hughes, Anna LC., Gyllencreutz, Richard, Mangerud, Jan, Svendsen, John-Inge, Lohmann, Gerrit, 2021. A new global ice sheet reconstruction for the past 80000 years. Nat. Commun. 12 (1), 1199.

Gowan, Evan J., Paul, Tregoning, Purcell, Anthony, Montillet, Jean-Philippe, McClusky, Simon, 2016. A model of the western laurentide ice sheet, using observations of glacial isostatic adjustment. Quat. Sci. Rev. 139, 1–16.

Grayson, Richard, 2016. A life in the trenches? the use of operation war diary and crowdsourcing methods to provide an understanding of the british army's day-to-day life on the western front. British Journal for Military History 2 (2).

Gregoire, Lauren J., Otto-Bliesner, Bette, Valdes, Paul J., Ivanovic, Ruza, 2016. Abrupt bølling warming and ice saddle collapse contributions to the meltwater pulse 1a rapid sea level rise. Geophys. Res. Lett. 43 (17), 9130–9137.

Hughes, Anna LC., Clark, Chris D., Jordan, Colm J., 2014. Flow-pattern evolution of the last british ice sheet. Quat. Sci. Rev. 89, 148–168.

Hughes, Anna LC., Gyllencreutz, Richard, Lohne, Øystein S., Mangerud, Jan, Svendsen, John Inge, 2016. The last eurasian ice sheets–a chronological database and time-slice reconstruction, dated-1. Boreas 45 (1), 1–45.

Knapp, Kenneth R., Matthews, Jessica L., Kossin, James P., Hennon, Christopher C., 2016. Identification of tropical cyclone storm types using crowdsourcing. Mon. Weather Rev. 144 (10), 3783–3798.
Li, Yingkui, Jacob, Napieralski, Harbor, Jon, Hubbard, Alun, 2007. Identifying patterns of correspondence between modeled flow directions and field evidence: an automated flow direction analysis. Comput. Geosci. 33 (2), 141–150.
Li, Yingkui, Jacob, Napieralski, Harbor, Jon, 2008. A revised automated proximity and conformity analysis method to compare predicted and observed spatial boundaries of geologic phenomena. Comput. Geosci. 34 (12), 1806–1814.
Lingard, Timothy, Masters, Karen L., Krawczyk, Coleman, Lintott, Chris, Kruk, Sandor, Simmons, Brooke, Keel, William, Nichol, Robert C., Baeten, Elisabeth, 2021. Galaxy zoo builder: morphological dependence of spiral galaxy pitch angle. Mon. Not. Roy. Astron. Soc. 504 (3), 3364–3374.
Napieralski, Jacob, Li, Yingkui, Harbor, Jon, 2006. Comparing predicted and observed spatial boundaries of geologic phenomena: automated proximity and conformity analysis applied to ice sheet reconstructions. Comput. Geosci. 32 (1), 124–134.
O'Hagan, Anthony, 2019. Expert knowledge elicitation: subjective but scientific. Am. Statistician 73 (Suppl. 1), 69–81.
Peltier, W.R., Farrell, W.E., Clark, J.A., 1978. Glacial isostasy and relative sea level: a global finite element model. Tectonophysics 50 (2–3), 81–110.
Pico, T., Creveling, J.R., Mitrovica, J.X., 2017. Sea-level records from the us mid-atlantic constrain laurentide ice sheet extent during marine isotope stage 3. Nat. Commun. 8 (1), 15612.
Quinlan, Garry, Beaumont, Christopher, 1982. The deglaciation of atlantic Canada as reconstructed from the postglacial relative sea-level record. Can. J. Earth Sci. 19 (12), 2232–2246.
Small, David, Clark, Chris D., Chiverrell, Richard C., Smedley, Rachel K., Bateman, Mark D., Duller, Geoff AT., Ely, Jeremy C., Fabel, Derek, Medialdea, Alicia, Moreton, Steven G., 2017. Devising quality assurance procedures for assessment of legacy geochronological data relating to deglaciation of the last british-Irish ice sheet. Earth Sci. Rev. 164, 232–250.
Tarasov, Lev, Peltier, W.R., 2007. Coevolution of continental ice cover and permafrost extent over the last glacial-interglacial cycle in North America. J. Geophys. Res.: Earth Surf. 112 (F2).
Vacchi, Matteo, Engelhart, Simon E., Nikitina, Daria, Ashe, Erica L., Peltier, W Richard, Roy, Keven, Kopp, Robert E., Horton, Benjamin P., 2018. Postglacial relative sea-level histories along the eastern canadian coastline. Quat. Sci. Rev. 201, 124–146.
Walcott, R.I., 1972. Late quaternary vertical movements in eastern north America: quantitative evidence of glacio-isostatic rebound. Rev. Geophys. 10 (4), 849–884.
Weisz, Daniel R., Johnson, L Clifton, Foreman-Mackey, Daniel, Dolphin, Andrew E., Beerman, Lori C., Williams, Benjamin F., Dalcanton, Julianne J., Rix, Hans-Walter, Hogg, David W., Morgan, Fouesneau, et al., 2015. The high-mass stellar initial mass function in m31 clusters. Astrophys. J. 806 (2), 198.
Williams, Alex C., Wallin, John F., Yu, Haoyu, Perale, Marco, Carroll, Hyrum D., Lamblin, Anne-Francoise, Fortson, Lucy, Obbink, Dirk, Lintott, Chris J., Brusuelas, James H., 2014. A computational pipeline for crowdsourced transcriptions of ancient Greek papyrus fragments. In: 2014 IEEE International Conference on Big Data (Big Data). IEEE, pp. 100–105.
Williamson, Daniel, Goldstein, Michael, Allison, Lesley, Blaker, Adam, Challenor, Peter, Jackson, Laura, Yamazaki, Kuniko, 2013. History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble. Clim. Dynam. 41, 1703–1729.