Recent progress on combining geomorphological and geochronological data with ice sheet modelling, demonstrated using the last British-Irish Ice Sheet

Jeremy C. Ely[1], Chris D. Clark[1], Richard C.A. Hindmarsh[2], Anna L.C. Hughes[3], Sarah L. Greenwood[4], Sarah L. Bradley[5], Edward Gasson[6], Lauren Gregoire[7], Niall Gandy[7], Chris R. Stokes[8] and David Small[8].

1 – Department of Geography, The University of Sheffield, Sheffield, S10 2TN, UK
2 – British Antarctic Survey, High Cross, Madingley Road, Cambridge, CB3 0ET, UK
3 - Department of Earth Science, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway
4 - Department of Geological Sciences, Stockholm University, Stockholm, Sweden
5 - Department of Geoscience and Remote Sensing, Delft University of Technology, Stevinweg 1, 2628 CN Delft, Netherlands
6 – School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8 1SS, UK
7 - University of Leeds, School of Earth and Environment, Woodhouse Lane, Leeds, LS2 9JT, UK
8 – Durham University, Department of Geography, Lower Mountjoy, South Road, Durham, DH1 3LE, UK

Abstract

Palaeo-ice sheets are important analogues for understanding contemporary ice sheets, offering a record of ice sheet behaviour that spans millennia. There are two main approaches to reconstructing palaeo-ice sheets. Empirical reconstructions use the available glacial geological and chronological evidence to estimate ice sheet extent and dynamics but lack direct consideration of ice physics. In contrast, numerically-modelled simulations implement ice physics, but often lack direct quantitative comparison to empirical evidence. Despite being long-identified as a fruitful scientific endeavour, few ice sheet reconstructions attempt to reconcile the empirical and model-based approaches. To achieve this goal, model-data comparison procedures are required. Here, we compare three numerically-modelled simulations of the former British-Irish Ice Sheet with the following lines of evidence: (i)

position and shape of former margin positions, recorded by moraines; (ii) former ice-flow direction and flow-switching, recorded by flowsets of subglacial bedforms; and (iii), the timing of ice-free conditions, recorded by geochronological data. These model-data comparisons provide a useful framework for quantifying the degree of fit between numerical model simulations and empirical constraints. Such tools are vital for reconciling numerical modelling and empirical evidence, the combination of which will lead to more robust palaeo-ice sheet reconstructions with greater explicative and ultimately predictive power.

1. Introduction

Reconstructing the behaviour of palaeo-ice sheets enables a better understanding of the long-term (centennial to millennial) behaviour of ice sheets in the Earth system. The former extent and behaviour of ice sheets can be inferred principally from four main lines of evidence. First, relative sea-level (RSL) records (e.g. a raised beach or salt marsh) provide constraints on the loading history of an ice sheet. Through the application of a glacio-isostatic adjustment (GIA) model, RSL data can be used to infer palaeo-ice sheet thickness and extent (e.g. Lambeck and Chappell, 2001; Peltier, 2004; Bradley et al., 2011). Secondly, analysis of the properties and stratigraphic sequence of sediments transported and deposited by palaeo-ice sheets can be used to infer ice sheet history at a given location (e.g. Eyles and McCabe, 1989; Piotrowski and Tulaczyk, 1999). The geomorphological record, composed of landforms such as drumlins and moraines, can be used to decipher former ice-flow directions and margin positions (e.g. Hughes et al., 2014; Clark et al., 2018). Finally, the timing of deposition of sediment and/or the time glacially transported or eroded bedrock has been exposed, and by inference the timing of formation of associated landforms, can be dated using laboratory-based techniques to produce the third line of evidence, geochronological data (e.g. Libby et al., 1949; Duller, 2006; Small et al., 2017a).

The body of empirical evidence related to palaeo-ice sheets is continually growing, producing an ever-expanding library of palaeo-ice sheet data (e.g. Dyke, 2004; Clark et al., 2012; Hughes et al., 2016; Stroeven et al., 2016). Producing a glaciologically-plausible empirical reconstruction of a palaeo-ice sheet is, however, a challenging process, with three main limitations. First, evidence is often temporally and spatially fragmented, thereby requiring some subjective inference to be made about ice sheet behaviour between the data-constraints (Clark et al., 2012; Hughes et al., 2016). Secondly, all sources of data have inherent uncertainties due to factors such as preservation potential, inherent laboratory-based uncertainties and post-depositional modification (Hughes et al., 2016; Small et al., 2017a). Finally, a mathematically- and physically-based direct inversion from palaeo-glaciological information to infer past-ice sheet characteristics (e.g. former ice-flow velocities) has remained elusive owing to the complexity of processes involved, meaning that all reconstructions are subjective (albeit expert) inferences (Kleman and Borgström, 1996; Stokes et al., 2015). Despite these limitations, empirical reconstructions typically provide a spatially-coherent representation of ice sheet activity, often portrayed as a series of palaeogeographic maps showing ice extent, flow geometry, ice divides and their changes at any given time (or at several time-steps).

As an alternative to the data-driven approach of an empirical reconstruction, numerical ice sheet models can be used to reconstruct palaeo-ice sheet behaviour (e.g. Fisher et al., 1985; Tarasov and Peltier, 2004; Hubbard et al., 2009; Patton et al., 2017). The approach here is to apply a numerical model based on the understanding of ice sheet physics to produce a modelled reconstruction of a palaeo-ice sheet. Using this physics-based approach, information such as ice-thickness and velocity can be reconstructed across the entire model domain in a manner that is consistent with model physics. However, limitations with this approach mean that modelled reconstructions may struggle to replicate the information and detail provided by palaeo-data. Numerical ice sheet models require the specification of several input boundary conditions and parameters. One of the most uncertain of these is the climatic conditions used to determine the pattern of accumulation and ablation over the model domain through time (Stokes et al., 2015). Other factors relating to the nature of ice sheet flow, such as basal friction, subglacial hydrology and shear, may either rely upon poorly constrained model parameters (due to a lack of physical understanding), or simply be beyond the capabilities of the model (e.g. they operate at scales below the spatial resolution of the model). Compounding the problem, ice sheets exhibit instabilities, whereby small perturbations to boundary conditions are amplified by the instability and can affect the whole modelled ice sheet. Such instabilities may lead to highly non-linear responses that are difficult to predict. One example is the marine ice sheet instability (Hughes, 1973; Schoof, 2007; 2012), which is an instability in the position of the grounding-line on a reverse bed slope that occurs as a consequence of ice-flux being proportional to ice-thickness at the grounding-line.

A complementary approach to the above is to view ice sheet behaviour as an expression of the weather/climate duality; “climate is what on an average we expect, weather is what we actually get” (Herbertson, 1908, p. 118). Restricting our attention to NW Europe, over diurnal periods weather is quite predictable, but this statement is false over time periods of a few days. On the other hand, it is true to say that winter months will be colder than summer months. The loss of predictability on a weekly time-scale arises from physical instabilities in the atmospheric circulation (Lorenz, 1963), and decades of observations have allowed scientists to make general statements about the temporal and spatial scales associated with these instabilities, improving predictability (Bauer et al., 2015).

Unfortunately, we do not have enough observations of ice sheet behaviour to make similar statements about the spatial and temporal scales associated with glaciological variability. Ice streams are a good example; the Kamb Ice Stream shut down in the past two centuries (Retzlaff and Bentley, 1993), and a myriad of ice streams with similar potential behaviour have been identified from the geological record in North America and Europe (Stokes and Clark, 1999; Margold et al., 2015). Modelling has shown that ice streams can be generated by, for example, instabilities in thermo-mechanical coupling (Hindmarsh, 2009), but none of these models have been used to match the extent of specific ice streams, due in part or largely to lack of data. Another example, most likely with greater spatial extent, is the marine ice sheet instability (MISI, Schoof, 2007), which acts on marine ice sheets with grounding lines on reverse slopes. Both ice streams and the MISI can be viewed as examples of ice sheet ‘weather’ – lack of predictability caused by instabilities, in exactly the same way as atmospheric weather is generated by instabilities.

This leads to a conundrum increasingly faced by geologists and geomorphologists; is the unusual behaviour frequently observed a signal from the whole ice sheet, or is it a signal of local variability? This is where modellers can inform field scientists, since modelling can give physically-based estimates of the spatial and temporal scale of unstable behaviour.

To account for the above limitations and uncertainties of modelled-reconstructions, two general approaches have been adopted which produce multiple ice sheet simulations. The first involves sensitivity analyses (e.g. Boulton and Hagdorn, 2006; Patton et al., 2016), whereby relevant model parameters and boundary conditions are perturbed to produce numerous simulations of the palaeo-ice sheet in question. Such tuning is conducted until a simulation is generated that is perceived to ‘best fit’ the empirical evidence, and is chosen as the modelled reconstruction. The second adopts an ensemble approach (e.g. Tarasov and Peltier, 2004; Gregoire et al., 2012), whereby a wide set of plausible combinations of parameters are input into the ice sheet model to produce an array of model-outputs. Data-based constraints may then be used to rule out unrealistic simulations from the bank of ensemble simulations, leaving a combination of simulations that are yet to be ruled out (e.g. Gregoire et al., 2016). The second approach is to calibrate ensemble parameters against data constraints, ruling out simulations and their associated parameter sets based on acceptable fits to the data (e.g. Tarasov and Peltier, 2004). The remaining simulations are then supplemented by further simulations, which use the calibrated parameters. The final modelled reconstruction in this approach is a combination of calibrated model simulations, from which the distribution of plausible glaciological variables can be derived (e.g. mean ice velocity) (Tarasov et al., 2012).

Ideally, palaeo-ice sheet reconstructions should combine the data-rich empirical approach with physically-based modelled reconstructions. Indeed, this suggestion was put forward in a landmark paper by Andrews (1982), when numerical modelling was very much in its infancy, and yet it has been very difficult to achieve. Ice sheet model outputs are often compared to RSL data through GIA modelling (e.g. Simpson et al., 2009; Kuchar et al., 2012; Auriac et al., 2016; Patton et al., 2017), but quantitative model-data comparison using other forms of palaeo-ice sheet data has remained rare (although see Briggs and Tarasov, 2013; Patton et al., 2016). This is despite the development (Napieralski et al., 2006; Li et al., 2007) and demonstration (Napieralski et al., 2007) of tools for data-model comparison.

Adopting this approach may create new opportunities for both empiricists and ice sheet modellers to drive the field forward. Empiricists could use models to help reduce data uncertainty and rule out physically-implausible interpretations. Modellers could use the data to score ensemble members and improve model formulation (as per Tarasov and Peltier, 2004). Here, we extend some recent advances in this area to outline a procedure for comparing geochronological and geomorphological data with ice sheet model output. We illustrate this with example model output of the British-Irish Ice Sheet (BIIS). Given the expanding body of data constraining palaeo-ice sheet behaviour (e.g. Greenwood and Clark, 2009; Hughes et al., 2014; Small et al., 2017a; Clark et al., 2018), it is one of the best ice sheets for model-data comparison. The primary purpose of the model runs presented here is not to simulate the intricacies of this palaeo-ice sheet or advance our understanding of the ice sheet, but simply to facilitate methodological comparisons between model output and empirical data. Meaningful and more accurate simulations of the ice sheet are the subject of ongoing work as part of the BRITICE-CHRONO NERC consortium project (e.g. Gandy et al., 2018).

2. Methods of model-data comparison

Of the four sources of data that might be used to constrain palaeo-ice sheet simulations (RSL, sedimentology, geochronology, and geomorphology), it is perhaps not surprising that RSL has the longest tradition (Walcott, 1972; Peltier et al., 1978; Quinlan and Beaumont, 1982). Sea-level index points provide a testable dataset with definable uncertainty (e.g. Engelhart and Horton, 2012). Furthermore, until recently, ice sheet models were run at a low-resolution of >20 km grid size. This meant that modelled reconstructions could be tested against relative-sea level data, which has a lack of abrupt spatial changes, through the use of a GIA model (e.g. Auriac et al., 2016). The advent of faster and parallel processing means that higher-resolution simulations of continental ice sheets are now achievable (∼5 km), permitting comparison to other sources of information. However, these data need to be presented at a similar resolution to the model and will perhaps provide definitive and quantifiable characteristics that a model can predict. Ice sheet models are yet to have adequate sediment production, transportation, and deposition laws to make predictions to the same level of detail that might be observed in a sediment exposure. We here demonstrate how to make meaningful model-data comparisons to the remaining two classes of palaeo-ice sheet data, geomorphological (ice-margin position and ice-flow direction) and geochronological (in essence, the timing of ice-free conditions).

2.1. Ice-Margin Position

Mapping of moraines underpins empirical palaeo-glaciology, providing information on former ice margin position, the direction of ice sheet retreat, and the shape of the margin (Figure 1A; Clark et al., 2012). Palaeo-ice sheet models can also predict these characteristics of a margin through time. However, only the largest moraines are likely to be of a sufficient scale to permit meaningful comparison with ice sheet model output. To compensate for this, neighbouring morainic ridges are often grouped/interpreted into larger composite margin positions, which collectively delineate ice margin retreat patterns (e.g. Figure 1B).

Napieralski et al. (2006) developed an Automated Proximity and Conformity Analysis (APCA) tool for comparing margin positions from mapped moraines and ice sheet model outputs (Table 1), later modified by Li et al. (2008). In this tool, mapped margins are first coarsened to conform to the ice sheet model grid size. Then, for each model-output time-slice, APCA measures the distance of an ice-margin determined based on mapped moraines to the modelled ice-margin (Figure 1C). The conformity of shape between margin positions determined from moraines and the model output is defined as the standard deviation of proximity for each cell occupied by a mapped margin position (Li et al., 2008; Figure 1C). An ideal simulation of a palaeo-ice sheet would match the location and shape of each moraine, which would be quantified by APCA as simultaneous zero proximity and perfect conformity at some point during the model run. However, model resolution limitations mean that a perfect score is unlikely to occur. Consequently, a more pragmatic approach would be to apply a proximity and conformity threshold, below which an acceptable level of model-data agreement occurs (Figure 1D). Only when both measures are below this predetermined acceptance threshold will model-data agreement be declared, i.e. the model matches the location and shape of the mapped margin derived from mapped moraines sufficiently. Where the relative sequence of moraine formation is known (e.g. in a retreat sequence of concentric moraines), the timing of margin matching could be considered. However, caution should be taken if relative timing of moraine formation criteria are utilised, in order that simulations which produce readvances that reoccupy margin positions are not excluded.

2.2. Ice-flow direction

Subglacial bedforms record the ice-flow directions within a palaeo-ice sheet (e.g. Kleman, 1990; Clark, 1993; Kleman and Borgström, 1996; Stokes et al., 2009; Ely et al., 2016). Where cross-cutting subglacial bedforms are superimposed on each other, a sequence of flow directions is recorded (Clark, 1993). Neighbouring subglacial bedforms with a similar morphology and orientation can be grouped into flowsets – groups of subglacial bedforms interpreted to form in the same phase of ice-flow (e.g. Kleman and Borgström, 1996; Clark, 1999). When grouped in this way, cross-cutting flowsets of subglacial bedforms can reveal major shifts in the flow patterns of an ice sheet, a consequence of shifting ice sheet geometry, ice-divide migration and ice-stream (de)activation (e.g. Boulton and Clark, 1990; Clark, 1999; Greenwood and Clark, 2009). Whilst a single flowset provides a spatially limited constraint on ice-flow direction, the sequence and spatial patterning of flowsets across the former ice sheet bed can be used to reconstruct the ice-flow geometry of a palaeo-ice sheet and the evolution of that geometry through time (Boulton and Clark, 1990; Kleman et al., 1997; Greenwood and Clark, 2009; Hughes et al., 2014).

Li et al. (2007) developed an Automated Flow Direction Analysis (AFDA) tool for comparing modelled and empirically derived ice sheet flow directions. To measure flow correspondence, AFDA calculates the mean residual angle and variance of offset between modelled and empirically derived ice-flow directions (Figure 2). Where detailed flowset reconstructions exist (e.g. for the BIIS; Greenwood and Clark, 2009; Hughes et al., 2014), the relative age of cross-cutting flowsets can be used as a further constraint by evaluating whether a model run recreates a cross-cutting sequence of flow directions in the inferred order of time (Figure 2). To do this, flow-direction model agreement would need to have occurred in the specified order, beneath a predetermined (user-specified) threshold which corresponds to an acceptable level of model-data agreement (Figure 2B).

2.3. Ice-free timing

The timing of ice-free conditions can be derived from geochronological techniques. These have been applied most commonly to organic material in the case of radiocarbon dating (Libby et al., 1949; Arnold and Libby, 1951; Ó Cofaigh and Evans, 2007; McCabe et al., 2007), proglacial sands in the case of luminescence dating (e.g. Duller, 2006; Smedley et al., 2017; Bateman et al., 2018), and glacially transported boulders or glacially modified bedrock in the case of cosmogenic nuclide dating (Stone et al., 2003; Fabel et al., 2012; Small et al., 2017b). For some palaeo-ice sheets, compilations of thousands of dates recording ice-free conditions relevant to the timing of advance and retreat exist (Dyke, 2004; Hughes et al., 2016; Small et al., 2017a). However, dating the activity of an ice sheet is complex and, as such, not all dates are equally reliable constraints (Small et al., 2017a). To account for this, an assessment of data reliability, such as the traffic-light system proposed by Small et al. (2017a), should be conducted prior to model-data comparison. This involves initially filtering out ages irrelevant to the study period. The remaining ages are then assigned a quality rating based upon the stratigraphic and geomorphological context, supporting evidence and potential for significant and unquantifiable geological uncertainty (Small et al. 2017a). Depending on the stratigraphic setting of a dated sample (e.g. above or below glacial sediment), this timing constrains ice free conditions either prior to an advance of, or following the retreat of, an ice sheet (Hughes et al., 2011). Each site has an associated error, related to measurement uncertainties. Since geochronological techniques only record the timing of ice-free conditions prior to (advance) or after (retreat) the occupation of an area by an ice sheet, the associated error can be considered as one-sided (Figure 3; Briggs and Tarasov, 2013; Ely et al., in press).

Ely et al. (in press) developed an Automated Timing Accordance Tool (ATAT) for comparing ice sheet model output with geochronological data. Ice-free dates must first be grouped as constraints on the retreat or advance of the ice sheet and then gridded (rasterised) to the resolution of the ice sheet model (Figure 3). Loose constraints, for example ice-free dates that are thousands of years younger or older than those indicated by the regional advance or retreat chronologies, can be ignored when creating the geochronological grid because they provide a poor test of the ice sheet model. ATAT produces several statistics based on the agreement between ice-free ages and modelled deglacial chronologies. It categorises dates as to whether there is agreement within both model and data uncertainty, including a procedure that considers whether a dated site could have become ice-free due to thinning of the ice sheet surface (i.e. nunataks or emergent hills close to margins). After classifying dates, ATAT calculates the route-mean square error (RMSE) between measured and modelled ice-free timings, with an additional weighted statistic which accounts for the uneven spatial distribution of dates (wRMSE). ATAT therefore measures both the number of dates that agree with a simulation (% of dates that agree), and how close the simulation gets to replicating the dates (wRMSE). Ideally, the ice sheet model would simulate ice-free conditions within the error of each geochronological constraint. Given the limitations of models, and the uncertainty associated with geochronological dates, the statistics generated by ATAT can be used more pragmatically to distinguish which model-runs better conform to the available geochronological archive (Ely et al., in press). For example, Ely et al. (in press) suggest that the measure “number of ice-free dates agreed with within error” is a good indictor from which to initially sift model simulations. A further application of ATAT is demonstrated in this paper.

3. Demonstration of approach using the British-Irish Ice Sheet.

3.1. Model-setup

Our primary aim is to demonstrate various approaches to model-data comparison, and so we perform some simple experiments with the aim of creating a range of outputs. We therefore make numerous simplifications, especially regarding our climate input. It is unimportant for the model experiments to exactly replicate the detailed reconstructed history of the BIIS (e.g. Clark et al., 2012). However, the model output serves as a means for demonstrating how model-data comparison tools could work. We use the Parallel Ice Sheet Model (PISM; Winkelmann et al., 2011) to simulate the BIIS. PISM is a hybrid shallow-ice shallow-shelf model which implements grounding line migration using a subgrid interpolation scheme. Ice movement is modelled as a combination of ice deformation and basal sliding. Internal deformation is determined by a flow law (Glen, 1952; Nye, 1953) with ice rheology altered by an enthalpy scheme (Aschwanden et al., 2012). Basal sliding occurs through a pseudo-plastic sliding law once basal shear stresses exceed yield stresses. Yield stress is determined to be a function of till friction, with till friction being a function of elevation and modelled basal effective pressure (Martin et al., 2011). Effective pressure is determined by a local subglacial hydrology model which relates overburden pressure to subglacial melt rates whilst ignoring horizontal water transport (Tulaczyk et al., 2000; Bueler and van Pelt, 2015). The model allows ice-shelves to form. Sub-shelf melt is determined using the parameterisation of Beckmann and Goosse (2003) perturbed by a melt factor (Martin et al., 2011), assuming that basal ice temperature is at pressure-melting point and ocean temperatures are at the freezing point at the depth of the ice-ocean interface (Martin et al., 2011). Calving rates are proportional to horizontal strain rates and are determined by a 2D parameterisation (Levermann et al., 2012; see also Supplementary Table 1 for key parameters).

We run the model at 5 km resolution, using bed topography gridded from the General Bathymetric Chart of the Oceans (www.gebco.net; Weatherall et al., 2015). Though higher resolution simulations of palaeo-ice sheets are possible (e.g. Seguinot et al., 2018), they are computationally expensive, limiting the ability to run ensembles or sensitivity analyses. Furthermore, larger palaeo-ice sheets (e.g. the Laurentide), where similar approaches could be conducted, require similar or coarser resolutions. Topography is updated to account for isostasy using a parameterisation of viscoelastic Earth deformation in response to loading (Bueler et al., 2007). Eustatic sea level change is accounted for by applying a scalar offset from the SPECMAP data (Imbrie et al., 1984).

To demonstrate differences between model simulations, we limit our analyses to the output from three model simulations. Parameters and boundary conditions are the same for all three simulations, with the exception that we vary the climate input. Climate is represented in our simulation as a spatially continuous field derived from multiple regression analysis of three sources of climate data; two modern day records and one from a palaeo-climate modelling experiment (Table 2; Braconnot et al. 2012). Prescribed temperatures are perturbed over time by a scalar offset derived from the Greenland ice core records (Seierstad et al., 2014) and fed into a positive degree day model to calculate surface mass balance (Calov and

Greve, 2005). Precipitation is also corrected with reference to the Greenland ice core record, with a 7.3% reduction in precipitation per degree Celsius decrease in temperature (Huybrechts, 2002). The model runs from 40 thousand years before present (ka BP) to the present day. Model output was recorded at 100-year intervals. The maximum extent of ice generated by each model simulation is shown in Figure 4. As expected, none of the model simulations perform well at replicating the reconstructed extent of the BIIS (e.g. Clark et al., 2012; Figure 4). The inability to reach these extents is most likely a consequence of the simplistic climate forcing and would therefore likely be ruled out by visual assessment alone. Such visual assessment is time consuming, especially considering that an ensemble is likely to produce thousands of model simulations. Furthermore, it may be that the parameters used in one simulation produce a closer fit to the data than others, guiding future models. It is therefore important to test model-data tools against these simulated ice sheets.

3.2. Ice Margin Position

We derived 189 ice margin positions from moraines reported in the BRITICE v.2 database (Clark et al., 2018) and compared these using APCA (Li et al., 2008) against our modelled ice-margin positions (Figure 5). To determine reasonable thresholds of proximity and conformity beyond which model-data agreement can be declared, we conducted sensitivity analysis validated by visual inspection (Figure 5B). We found that a proximity threshold of 15 km and a conformity threshold of 3 km sufficiently identified modelled ice margin positions that visually agreed with the shape and location of each moraine (Figure 5B, 5C). These thresholds could be used in similar experimental setups. A similar proximity measure (15 km) was reported by Napieralski et al. (2007). Figure 5B shows an example of a margin position where data-model agreement occurred. Data-model agreement occurred several times during the course of the simulation for this particular margin, as both measures of proximity and conformity fell below the agreement threshold on multiple occasions (Figure 5C). Marine based ice sheets, such as the BIIS, are prone to readvances (Schoof, 2007; Kingslake et al., 2018). The potential to readvance means that we cannot make the simple assumption that moraines closer to the ice sheet centre are older, meaning that we do not consider time sequences of margin occupation as a test here.

Table 3 shows the percentage of margins matched by each model run. The most common reason for model-data mismatch was that margins were not reached by the simulated ice extent, meaning that they scored too low on the proximity score of APCA. This is unsurprising given that 2 out of 3 of the models do not reach the extent of all considered margins (Figures 4, 5A). To test whether the model agrees with the observed shape and proximity of margins that are within modelled extent, we calculated a second statistic, which considered only those observed margins within the maximum extent of a given model simulation (Figures 4, 5A; Table 3). This shows that each simulation has model-data agreement with over 50% of the margins reached and their shape replicated by the model simulation (i.e. excluding mismatches for margins that are outside the maximum extent of the model simulation) (Table 3). However, direct comparisons between simulations become problematic when restricting the analysis to only moraines within the maximum extent, as this changes the number of data that are being compared (Table 3). We therefore created a third metric, the extent of margins matched within the extent Simulation C, the simulation which produced the smallest ice extent (Table 3; Figure 4).

3.3. Ice-flow direction

A total of 103 flowsets from Britain and Ireland were compared to our model simulations using AFDA (Li et al., 2007) (Figure 6A). These were assembled from Greenwood and Clark (2009) and Hughes et al. (2014) and include 32 cross cutting relationships. Combined, the datasets of Greenwood and Clark (2009 and Hughes et al. (2014) have over 150 flowsets. However, given the horizontal resolution of the models (5 km), small (<20 km wide) flowsets were excluded from the analysis (n = 39). Flowsets identified as times-transgressive (i.e. formed asynchronously) were either divided into the stages of formation identified by Greenwood and Clark (2009) and Hughes et al. (2014), or excluded from the analysis (n = 20). Flow vectors were derived from the empirically-derived depiction of a flowset, rather than individual bedforms, because the orientation of these may vary on a sub-grid scale. For data-model agreement to occur, we applied a threshold of 10° mean residual vector, and 0.03 in mean variance. These values were initially derived by visually comparing the model and data and determining whether a modelled ice flow direction was sufficiently similar to a mapped flowset. These threshold values are consistent with those reported by Napieralski et al. (2007), and could be used to declare model-data agreement in similar experimental setups. To get a cross-cutting relationship registered to be in data-model agreement, the last occurrence of model conformity for the first flowset in a sequence needs to occur before the last occurrence of model conformity for the overprinted flowset.

Table 3 summarises the comparison between model output from the three Simulations and the assembled flowset database (Figure 6A). Overall, model-data agreement was low, with the majority of flowsets not replicated by the model simulations (Table 3). Similar to the margin comparison, this is partly a consequence of the models computed ice-covered area not replicating the full area covered by the BIIS (Figure 4). We therefore produced a second metric that restricted the analysis to those flowsets occurring within the modelled ice-extent. This was done to see if model-data mismatch was a consequence of ice-extent (in which a high number of ice-covered data points would be matched), or due to model-data mismatch even over the ice-covered area. However, note the caveat that this limits the ability to compare between simulations owing to the changing number of data in the model-data comparison. A third metric, the percentage of flowsets matched within the extent of simulation C (the simulation with the smallest ice extent), allows for comparison between model runs. Even when this approach is adopted, the degree of model-data agreement for flowsets remains low, with simulation A being the best performing, matching 26% of flowsets within the extent of simulation C (Table 3). Furthermore, no models were able to replicate an observed cross-cutting relationship (Table 3). Figures 6B and C demonstrate an example of a matched flowset. Here, ice flow of sufficient coherence (a variance measure) in an agreed direction (vector orientation measure) is achieved toward the end of the model run (Figure 6C).

3.4. Ice-free Timing

Simulated ice sheet retreat timing from the model was compared to 108 published dated sites of ice sheet retreat using the ATAT (Ely et al., in press). Only sites with a green or amber quality rating from the traffic light system of Small et al. (2017a) were used. This means that the quality control considerations of dating techniques and stratigraphic contexts were deemed to be high-quality (green) or acceptable (amber). Sites flagged with ‘caution when interpreting (red)’, due to specific site or technique uncertainty, were not considered here (see Figure 7A for the location of sites used). For each model run, we report the percentage of dates where model-data agreement occurs (i.e. when a model recreates the ice-free timing recorded by geochronological data) and a spatially weighted root-mean square error (wRMSE) between data-based and model-based deglaciation timing (Ely et al., in press; Table 3). These measures consider the uncertainty in model-margin timing and the vertical uncertainty introduced when comparing low resolution modelled ice-surface topography to geochronological data collected at a point location (Ely et al., in press).

Simulation B performs poorly in replicating the timing of ice-free conditions, with data-model conformity occurring for only 9% of the dates (Table 3). Simulations A and C have higher scores of this metric, with 41% and 89% of the dates agreeing with the modelled timing of ice-free conditions, respectively (Table 3). However, these model runs also have high wRMSE scores (Table 3), meaning that although ice-free conditions correctly occur, they are far from the mean age recorded by the geochronological data. For example, in Simulation C this indicates that although model-data agreement has occurred (i.e. the model has deglaciated an area before the empirical evidence indicates ice-free conditions), the timing of modelled ice-free conditions is ~2000 years earlier on average than that recorded in the data. This pattern of premature deglaciation is apparent in Figure 7D.

4. Discussion

4.1. Model-data fit

Integration of the empirically-based and model-based approaches of ice sheet reconstruction requires tools for quantifying the degree of fit between models and data. Comparisons between the varied constraints of margin position, flow direction and timing, such as those conducted in Section 3, are a step towards achieving this goal. A model-based reconstruction is likely to be more robust if it involves multiple (100s-1000s) model simulations, rather than just the three illustrated here. However, given that none of these individual simulations is likely to match every piece of available evidence the question "which simulations adequately recreate the available geological data?" must be addressed. By addressing this question, an investigator may be able to find the optimum model reconstruction (e.g. Napieralski et al., 2007; Patton et al., 2016; Seguinot et al., 2016). Alternatively, these model-data tests could be incorporated as additional calibration criteria for ensemble simulations (e.g. Tarasov et al., 2012), which could potentially reduce the produced uncertainty of an ensemble model reconstruction.

Despite only using three model runs, our comparison highlights some difficulties in answering the above question. For margin positions, all models performed reasonably well, matching over 50% of the margins within the modelled ice sheet extent and, in the case of Simulation A, 75% (Table 3). Therefore, if looking at margin position in isolation from other metrics, Simulation A would be considered the best performing model-run. Since all models perform well at replicating ice-marginal positions, our results, albeit limited to a small sample of three simulations, suggest that the margin metric is the least stringent test of the ice sheet simulations (Table 3). One possible reason for this is that models are better at replicating margin shapes and positions than other data-based characteristics. However, a second interpretation is that that the generalisation of margin shape to a 5 km grid removes any complexity in margin shape, thus promoting conformity between model and data. Future work, which considers ice sheet models and margin data at different resolutions should be undertaken to examine this in more detail.

All three model simulations do not replicate the maximum extent of the BIIS derived from observations. The maximum extent of an ice sheet is generally well known, and some of these moraines record the maximum extent across different sectors of the BIIS (e.g. Bradwell et al., 2008; Clark et al., 2012). Therefore, future work may adopt a procedure of testing ice sheet models against only those margins derived from moraines which demark maximum palaeo-ice sheet extent and glaciated continental shelf-breaks (e.g. Seguinot et al., 2016; Patton et al., 2017), so as to identify simulations and glacio-climatic parameter combinations which achieve a reasonable ice sheet extent, before attempting to replicate margin positions occupied during ice retreat. A model which fits maximum ice extent margins in some places, may be able to interpolate between these constraints in a more consistent manner than empirical interpretations (e.g. Bowen et al., 1986; Clark et al., 2012; Seguinot et al., 2016; Patton et al., 2017).

All three simulations performed poorly at replicating the flow direction recorded by subglacial bedforms (Table 3). This is surprising given that the direction of many flowsets appears to be governed by the subglacial topography in Britain (Hughes et al., 2014), which is also likely to steer ice flow directions in numerical models that use that topography. One possibility is that this is due to the coarse (5 km) resolution of our model grid. Perhaps this model-data mismatch is also a consequence of the model being unable to fully replicate other conditions which determine ice flow direction such as basal thermal regime, subglacial hydrological conditions and the overall ice-sheet geometry (e.g. location of ice divides and domes). Areas with subglacial bedforms indicate warm-based ice, where basal sliding/subglacial till deformation is the dominant control upon ice-discharge. The most common reason for model-data mismatch in flow direction was the low mean residual variance scores. In other words, the model did not produce consistent flow-directions across the entire area of the flowset. Therefore, model-data mismatch is at least partially due to the model being unable to adequately simulate the dimensions of ice-streams and outlet glaciers, perhaps due to simplifications of physics (Hindmarsh, 2009; Stokes and Tarasov, 2010), poorly constrained patterns of basal sliding parameters (Bueler and Brown, 2009), or incomplete knowledge of basal sliding (Stearns and van der Veen, 2018). Climate uncertainties will also influence the ability of an ice sheet model to replicate empirically derived flow directions, as these impact the overall geometry of the modelled ice sheet. Since these factors are a large uncertainty in ice sheet modelling (Ritz et al., 2015; Gladstone et al., 2017), flowset direction is likely to be a robust test of ice sheet models. A question remains regarding how long flowing ice must occupy an area in order to produce lineated flow sets; if this time is decadal (e.g. Dowling et al., 2016) rather than centennial, it indicates that flowset matching is not of the highest priority for ice sheet models which typically have a lower temporal resolution.

None of the three model simulations adequately replicated a cross-cutting relationship between flowsets. Such cross-cuts can be used to decipher the geometry of a palaeo-ice sheet and how it changes through time (Boulton and Clark, 1990), including factors such as ice-divide migration and margin position change (e.g. Greenwood and Clark, 2009; Hughes et al., 2014). This means, in addition to the problems of matching a single flow-set mentioned above, deglacial climate must be adequately simulated for a cross-cuts caused by climatically driven ice-divide migration to be matched. In addition to this, the model must also adequately represent the internal processes which cause ice-divide migration (e.g. flow piracy, ice stream initiation, saddle collapse). A further uncertainty is introduced by our ignorance of ice stream dynamics and how ice stream velocity and orientation can change over centennial and even decadal timescales. Given these potential difficulties at matching cross-cuts, they can be thought of as an even sterner test of an ice sheet model than the number of flowsets replicated alone.

None of the three model simulations performed well when compared to the assembled database of ice-free dates (Table 3). Simulation C has agreements with many sites (Table 3), but simulated deglaciation occurs thousands of years prior to the age indicated by the geochronological record at many sites, suggesting that retreat occurs too early and rapidly. Other modelling simulations have qualitatively demonstrated a better fit to deglacial chronologies by visually comparing the pattern and timing of modelled reconstructions to empirically-based reconstructions (e.g. Patton et al., 2017). However, replicating the timing of ice-free conditions across an ice sheet requires adequately constraining all internal and external forcings through time, as well as the interactions between the two. Therefore, our approach of site-by-site comparison to modelled deglacial timing provides a more stringent test of model-data fit than qualitative comparisons.

4.2. An approach to measuring model-data fit

As a consequence of the above complexity in model-data comparison, we suggest the following pragmatic approach to reconciling empirical reconstructions and model reconstructions, summarised in Figure 8. Here, the investigator starts with an ensemble of ice sheet model simulations; the number of simulations considered are progressively diminished in number by removing those which rank lowest against a particular metric (Figure 8). This builds on the suggestion of Napieralski et al. (2007) who used APCA to rule out the majority of simulations, then AFDA to further evaluate model performance. Our order of rankings (Figure 8) is based upon what we ascertain from the above discussion to be progressively more stringent tests of a model simulation. Indeed, the order of these rankings is likely to change between users who are interested in specific aspects of a palaeo-ice sheet (e.g. more weighting may be given to flowset direction if studying ice-flow patterns). An alternative is to combine scores derived from the model-data comparison techniques for each simulation, and then rank simulations to either heavily weight the highest scoring simulations when producing a probabilistic output from an ensemble (e.g. Tarasov et al., 2012), or to rule out the lowest scoring simulations. In this case, the order that tests are applied is irrelevant.

The original ensemble of simulations is likely to contain hundreds of members and may have involved some prior tuning of parameters to broadly replicate ice sheet extent (e.g. Boulton and Hagdorn, 2006). Since margin position seems a comparatively simple metric with which an ice sheet model result must conform, we suggest that the first set of models to be ruled out are those that perform lowest in the APCA tests against margins (Napieralski et al., 2006; Li et al., 2008; Figure 8). The top-performing simulations are then compared to timing through the ATAT tool (Ely et al., in press; Figure 8). ATAT will produce statistics on the number of dated positions matched, and how close overall the simulation gets to replicating the timing of ice-free conditions (wRMSE). Thresholds of acceptance should be applied for each, so that only simulations that replicate an adequate number of dates within a reasonable time window from the data will remain in the 'not-ruled out' category of simulations (Figure 8).This will rule out simulations which perform badly at replicating the timing and rate of palaeo-ice sheet retreat recorded in geochronological data. Since flowset conformity is likely to be a demanding test of ice sheet models, with the ability to produce cross cutting flow even more demanding, we suggest remaining simulations should then be ranked according to their performance according to the AFDA (Li et al., 2007; Figure 8).

After the application of these tests, the original ensemble of simulations will be much reduced, to a set which are yet to be ruled out (Figure 8). Given that it is unlikely that a perfect score will be found in these models, model-data mismatch between 'best-fit' models should be further investigated. It may be that certain areas of empirical evidence consistently produce model-data mismatch, and this may motivate further simulations if spatial or temporal patterns are clear. For example, a climate driver may under-represent a particular stadial, thereby producing a simulated timing which is disagrees with data. On the other hand, if all surrounding empirical evidence is met, and a particular data point or subset of data cannot be replicated by the model, this may warrant re-evaluation of the data in question (Figure 8). In an analogous manner to climate modelling (Collins et al., 2017), it remains open as to whether all models which pass a threshold acceptance barrier should be incorporated into an acceptable set of reconstructions (i.e. a model democracy; Knutti, 2010) or whether a “best-fit” model which performs best against all constraints should be identified and used for further research. In either case, the procedure outlined above can help reduce model uncertainty and produce more robust palaeo-ice sheet reconstructions.

4.3. Suggestions for future developments

The model-data comparison conducted here has highlighted some areas where comparison tools and procedures require further development. Some required developments are listed below and may aid in the reduction of both model and data uncertainty.

When comparing modelled and empirically derived margins using APCA, the occupied side of a moraine is not considered. In situations where ice-flow geometry is likely to be simple, for example in a deep trough or at the continental shelf break, this is unlikely to matter. However, in more complex settings, for example where two ice sheets converge such as in the North Sea, this may introduce false positives whereby a mapped margin is recorded to be matched by ice flowing from the wrong direction. Our margin comparison was also conducted throughout both the advance and retreat of the ice sheet. Again, this may introduce false positives, as moraines known to have formed in retreat may be matched during ice advance. We therefore suggest that future adaptations of APCA should consider ice flow direction and the trajectory of the modelled ice margin (advance or retreat). For the latter, this is unlikely to be as simple as restricting analysis to a certain time period from which deglaciation commences, as maximum extents may be asynchronous (e.g. Patton et al., 2016; Seguinot et al., 2018) and readvances may occur (e.g. Kingslake et al., 2018). Future work should also consider penalising a model for extending beyond a well-known limit of ice-extent (i.e. producing an ice-sheet that is too large). Furthermore, given the uncertainty of data, it is worth considering how certain the origin of each moraine system is when applying these tools. For example, could a moraine have formed during ice advance, and been preserved beneath cold-based ice?

For ice-flow direction comparison, our analysis shows that a key problem is replicating the synchronous flow directions recorded in some flowsets, and whether the model resolves the timescales involved in bedform formation. Given that there is some evidence that drumlins can form rapidly (Dowling et al., 2016) and the pattern of drumlins within a flowset can evolve with time (Ely et al., 2018), another way of extracting more information from a model-data comparison would be to compare the direction of individual bedforms to modelled-flow directions. If neighbouring bedforms match within a reasonable time difference, then the model could be used to classify bedforms into flowsets that could then be compared to those which are empirically derived (e.g. Greenwood and Clark, 2009; Hughes et al., 2014). Interpolating directions between modelled time-slices may also help improve model-data comparison of flow direction, potentially capturing the flow direction of some bedforms which form between model output timesteps.

Although influenced by overall ice sheet geometry, both margin and flow-direction are predominantly constraints upon the horizontal dimension of an ice sheet. Given that the thickness of ice is a vital variable for determining sea-level contribution and impacts upon the landscape, vertical constraints are also important. As stated above, our comparison would ideally be conducted alongside the use of a GIA model which compares to RSL data (e.g. Kuchar et al., 2012; Auriac et al., 2016; Patton et al., 2017). ATAT also has a procedure for identifying whether an ice-free date is positioned higher than the modelled ice-elevation (Ely et al., in press), for example if a nunatak is predicted. Given the importance of these vertical constraints on ice-sheet geometry, perhaps future comparisons should isolate these data as a separate test of model performance.

5. Summary

Progress toward an integration of empirically-based and numerical model-based reconstructions of palaeo-ice sheets have proven to be slow since being first suggested (Andrews, 1982; Stokes et al., 2015). Here, we have outlined a procedure of model-data comparison designed to score the degree of fit between ice sheet model simulations and palaeo-ice sheet data, which aims to further integrate these two approaches. We compared three ice sheet model simulations against the three data constraints of margin position (from moraines), flow direction (from subglacial bedforms) and timing of ice-free conditions (from geochronological data). In doing so, we highlighted the complexities of such model-data comparisons. As ice sheet models are unlikely to reproduce all the information provided at each constraint, we pragmatically suggest a hierarchical system for scoring ice sheet models, whereby successive tests are applied to the ice sheet model, progressively ruling out model runs which perform the poorest against each constraint. This procedure could be used to ascertain best-fit models or used to calibrate models. Future work could consider in more depth the relative importance of the different data-based constraints. Furthermore, we argue that this approach could lead to models more frequently being used to test the plausibility of data-interpretations. In future work, this comparison should ideally be made in conjunction with other data-based constraints such as RSL data through GIA modelling and sedimentological observations. In this manner, an integration of empirical and model-based approaches to palaeo-ice sheet reconstruction can occur. The BIIS is a data rich environment for conducting such model-data integration.

Acknowledgements

This work was supported by the Natural Environment Research Council (NERC) consortium grant; BRITICE-CHRONO NE/J009768/1. J.C.E. acknowledges support from a NERC independent research fellowship (NE/R014574/1). Development of PISM is supported by NASA grant NNX17AG65G and NSF grants PLR-1603799 and PLR-1644277. We thank Arjen Stroeven for editing this manuscript, as well as Irina Rogozhina and an anonymous reviewer for their insightful comments.

References

Andrews, J.T., 1982. On the reconstruction of Pleistocene ice sheets: A review. Quaternary Science Reviews, 1(1), pp.1-30.

Arnold, J.R. and Libby, W.F., 1951. Radiocarbon dates. Science, 113(2927), pp.111-120.

Aschwanden, A., Bueler, E., Khroulev, C. and Blatter, H., 2012. An enthalpy formulation for glaciers and ice sheets. Journal of Glaciology, 58(209), pp.441-457.

Auriac, A., Whitehouse, P.L., Bentley, M.J., Patton, H., Lloyd, J.M. and Hubbard, A., 2016. Glacial isostatic adjustment associated with the Barents Sea ice sheet: a modelling inter-comparison. Quaternary Science Reviews, 147, pp.122-135.

Bateman, M.D., Evans, D.J., Roberts, D.H., Medialdea, A., Ely, J. and Clark, C.D., 2018. The timing and consequences of the blockage of the Humber Gap by the last British− Irish Ice Sheet. Boreas, 47(1), pp.41-61.

Bauer, P., Thorpe, A. and Brunet, G., 2015. The quiet revolution of numerical weather prediction. Nature, 525(7567), p.47-55.

Beckmann, A. and Goosse, H., 2003. A parameterization of ice shelf–ocean interaction for climate models. Ocean modelling, 5(2), pp.157-170.

Benetti, S., Dunlop, P. and Cofaigh, C.Ó., 2010. Glacial and glacially-related features on the continental margin of northwest Ireland mapped from marine geophysical data. Journal of Maps, 6(1), pp.14-29.

Boulton, G.S. and Clark, C.D., 1990. A highly mobile Laurentide ice sheet revealed by satellite images of glacial lineations. Nature, 346(6287), p.813.

Boulton, G. and Hagdorn, M., 2006. Glaciology of the British Isles Ice Sheet during the last glacial cycle: form, flow, streams and lobes. Quaternary Science Reviews, 25(23-24), pp.3359-3390.

Bowen, D.Q., Rose, J., McCabe, A.M., Sutherland, D.G., 1986. Correlation of Quaternary glaciations in England, Ireland, Scotland and Wales. Quaternary Science Reviews 5, pp. 299-340

Bradley, S.L., Milne, G.A., Shennan, I. and Edwards, R., 2011. An improved glacial isostatic adjustment model for the British Isles. Journal of Quaternary Science, 26(5), pp.541-552.

Bradwell, T., Stoker, M.S., Golledge, N.R., Wilson, C.K., Merritt, J.W., Long, D., Everest, J.D., Hestvik, O.B., Stevenson, A.G., Hubbard, A.L. Finlayson, A.G., and Mathers, H.E., 2008. The northern sector of the last British Ice Sheet: maximum extent and demise. Earth-Science Reviews, 88(3-4), pp.207-226.

Braconnot, P., Harrison, S.P., Kageyama, M., Bartlein, P.J., Masson-Delmotte, V., Abe-Ouchi, A., Otto-Bliesner, B. and Zhao, Y., 2012. Evaluation of climate models using palaeoclimatic data. Nature Climate Change, 2(6), p.417.

Briggs, R.D. and Tarasov, L., 2013. How to evaluate model-derived deglaciation chronologies: a case study using Antarctica. Quaternary Science Reviews, 63, pp.109-127.

Bueler, E.D., Lingle, C.S. and Brown, J., 2007. Fast computation of a viscoelastic deformable Earth model for ice sheet simulations. Annals of Glaciology, 46(1), pp.97-105.

Bueler, E. and Brown, J., 2009. Shallow shelf approximation as a “sliding law” in a thermomechanically coupled ice sheet model. Journal of Geophysical Research: Earth Surface, 114(F3).

Bueler, E. and Pelt, W.V., 2015. Mass-conserving subglacial hydrology in the Parallel Ice Sheet Model version 0.6. Geoscientific Model Development, 8(6), pp.1613-1635.

Calov, R. and Greve, R., 2005. A semi-analytical solution for the positive degree-day model with stochastic temperature variations. Journal of Glaciology, 51(172), pp.173-175.

Clark, C.D., 1993. Mega-scale glacial lineations and cross-cutting ice-flow landforms. Earth surface processes and landforms, 18(1), pp.1-29.

Clark, C.D., 1999. Glaciodynamic context of subglacial bedform generation and preservation. Annals of Glaciology, 28, pp.23-32.

Clark, C.D., Hughes, A.L., Greenwood, S.L., Jordan, C. and Sejrup, H.P., 2012. Pattern and timing of retreat of the last British-Irish Ice Sheet. Quaternary Science Reviews, 44, pp.112-146.

Clark, C.D., Ely, J.C., Greenwood, S.L., Hughes, A.L., Meehan, R., Barr, I.D., Bateman, M.D., Bradwell, T., Doole, J., Evans, D.J. Jordan, C.J., Moneys, X., Pellicer, M. and Sheehy, M., 2018. BRITICE Glacial Map, version 2: a map and GIS database of glacial landforms of the last British–Irish Ice Sheet. Boreas, 47(1), p.11.

Collins, M., 2017. Still weighting to break the model democracy. Geophysical Research Letters, 44(7), pp.3328-3329.

Dowling, T.P., Möller, P. and Spagnolo, M., 2016. Rapid subglacial streamlined bedform formation at a calving bay margin. Journal of Quaternary Science, 31(8), pp.879-892.

Duller, G.A.T., 2006. Single grain optical dating of glacigenic deposits. Quaternary Geochronology, 1(4), pp.296-304.

Dyke, A.S., 2004. An outline of North American deglaciation with emphasis on central and northern Canada. Developments in Quaternary Sciences, 2, pp.373-424

Ely, J.C., Clark, C.D., Spagnolo, M., Stokes, C.R., Greenwood, S.L., Hughes, A.L., Dunlop, P. and Hess, D., 2016. Do subglacial bedforms comprise a size and shape continuum?. Geomorphology, 257, pp.108-119.

Ely, J.C., Clark, C.D., Spagnolo, M., Hughes, A.L. and Stokes, C.R., 2018. Using the size and position of drumlins to understand how they grow, interact and evolve. Earth Surface Processes and Landforms, 43(5), pp.1073-1087.

Ely, J.C., Clark, C.D., Small, D., Hindmarsh, R.C.A, in press. ATAT 1.1, an Automated Timing Accordance Tool for comparing ice sheet model output with geochronological data. Geoscientific Model Development Discussions.

Engelhart, S.E. and Horton, B.P., 2012. Holocene sea level database for the Atlantic coast of the United States. Quaternary Science Reviews, 54, pp.12-25.

Eyles, N. and McCabe, A.M., 1989. The Late Devensian (< 22,000 BP) Irish Sea Basin: the sedimentary record of a collapsed ice sheet margin. Quaternary Science Reviews, 8(4), pp.307-351.

Fabel, D., Ballantyne, C.K. and Xu, S., 2012. Trimlines, blockfields, mountain-top erratics and the vertical dimensions of the last British–Irish Ice Sheet in NW Scotland. Quaternary Science Reviews, 55, pp.91-102.

Fisher, D., Reeh, N. and Langley, K., 1985. Objective reconstructions of the late Wisconsinan Laurentide Ice Sheet and the significance of deformable beds. Géographie physique et Quaternaire, 39(3), pp.229-238.

Gandy, N., Gregoire, L.J., Ely, J.C., Clark, C.D., Hodgson, D.M., Lee, V., Bradwell, T. and Ivanovic, R.F., 2018. Marine ice sheet instability and ice shelf buttressing of the Minch Ice Stream, northwest Scotland. The Cryosphere, 12(11), pp.3635-3651.

Gladstone, R.M., Warner, R.C., Galton-Fenzi, B.K., Gagliardini, O., Zwinger, T. and Greve, R., 2017. Marine ice sheet model performance depends on basal sliding physics and sub-shelf melting. The Cryosphere, 11(1), p.319.

Glen, J.W., 1952. Experiments on the deformation of ice. Journal of Glaciology, 2(12), pp.111-114.

Greenwood, S.L. and Clark, C.D., 2009. Reconstructing the last Irish Ice Sheet 2: a geomorphologically-driven model of ice sheet growth, retreat and dynamics. Quaternary Science Reviews, 28(27-28), pp.3101-3123.

Gregoire, L.J., Payne, A.J. and Valdes, P.J., 2012. Deglacial rapid sea level rises caused by ice sheet saddle collapses. Nature, 487(7406), p.219.

Gregoire, L.J., Otto-Bliesner, B., Valdes, P.J. and Ivanovic, R., 2016. Abrupt Bølling warming and ice saddle collapse contributions to the Meltwater Pulse 1a rapid sea level rise. Geophysical research letters, 43(17), pp.9130-9137.

Herbertson, A.J. (1908). Outlines of Physiography: An Introduction to the Study of the Earth. Edward Arnold, London. 3rd ed. Page 118.

Hindmarsh, R.C., 2009. Consistent generation of ice-streams via thermo-viscous instabilities modulated by membrane stresses. Geophysical Research Letters, 36(6).

Hubbard, A., Bradwell, T., Golledge, N., Hall, A., Patton, H., Sugden, D., Cooper, R. and Stoker, M., 2009. Dynamic cycles, ice streams and their impact on the extent, chronology and deglaciation of the British–Irish ice sheet. Quaternary Science Reviews, 28(7), pp.758-776.

Hughes, A.L., Greenwood, S.L. and Clark, C.D., 2011. Dating constraints on the last British-Irish Ice Sheet: a map and database. Journal of Maps, 7(1), pp.156-184.

Hughes, A.L., Clark, C.D. and Jordan, C.J., 2014. Flow-pattern evolution of the last British Ice Sheet. Quaternary Science Reviews, 89, pp.148-168.

Hughes, A.L., Gyllencreutz, R., Lohne, Ø.S., Mangerud, J. and Svendsen, J.I., 2016. The last Eurasian ice sheets–a chronological database and time-slice reconstruction, DATED-1. Boreas, 45(1), pp.1-45.

Hughes, T., 1973. Is the West Antarctic ice sheet disintegrating?. Journal of Geophysical Research, 78(33), pp.7884-7910.

Huybrechts, P., 2002. Sea-level changes at the LGM from ice-dynamic reconstructions of the Greenland and Antarctic ice sheets during the glacial cycles. Quaternary Science Reviews, 21(1-3), pp.203-231.

Imbrie, J., Hays, J.D., Martinson, D.G., McIntyre, A., Mix, A.C., Morley, J.J., Pisias, N.G., Prell, W.L. and Shackleton, N.J., 1984. The orbital theory of Pleistocene climate: support from a revised chronology of the marine d18O record. In A. L. Berger, J. Imbrie, J. Hays, G. Kukla, & B. Saltzman (Eds.), Milankovitch and climate (pp. 269–305). Dordrecht, Holland: Milankovitch and Climate.

Kingslake, J., Scherer, R.P., Albrecht, T., Coenen, J., Powell, R.D., Reese, R., Stansell, N.D., Tulaczyk, S., Wearing, M.G. and Whitehouse, P.L., 2018. Extensive retreat and re-advance of the West Antarctic ice sheet during the Holocene. Nature, 558(7710), p.430.

Kleman, J., 1990. On the use of glacial striae for reconstruction of paleo-ice sheet flow patterns: with application to the Scandinavian ice sheet. Geografiska Annaler: Series A, Physical Geography, 72(3-4), pp.217-236.

Kleman, J. and Borgström, I., 1996. Reconstruction of palaeo-ice sheets: the use of geomorphological data. Earth surface processes and landforms, 21(10), pp.893-909.

Kleman, J., Hättestrand, C., Borgström, I. and Stroeven, A., 1997. Fennoscandian palaeoglaciology reconstructed using a glacial geological inversion model. Journal of glaciology, 43(144), pp.283-299.

Knutti, R., 2010. The end of model democracy? Climatic Change, 102, 395-404.

Kuchar, J., Milne, G., Hubbard, A., Patton, H., Bradley, S., Shennan, I. and Edwards, R., 2012. Evaluation of a numerical model of the British–Irish ice sheet using relative sea-level data: implications for the interpretation of trimline observations. Journal of Quaternary Science, 27(6), pp.597-605.

Lambeck, K. and Chappell, J., 2001. Sea level change through the last glacial cycle. Science, 292(5517), pp.679-686.

Levermann, A., Albrecht, T., Winkelmann, R., Martin, M.A., Haseloff, M. and Joughin, I., 2012. Kinematic first-order calving law implies potential for abrupt ice-shelf retreat. The Cryosphere, 6, pp.273-286.

Li, Y., Napieralski, J., Harbor, J. and Hubbard, A., 2007. Identifying patterns of correspondence between modeled flow directions and field evidence: an automated flow direction analysis. Computers & geosciences, 33(2), pp.141-150.

Li, Y., Napieralski, J. and Harbor, J., 2008. A revised automated proximity and conformity analysis method to compare predicted and observed spatial boundaries of geologic phenomena. Computers & Geosciences, 34(12), pp.1806-1814.

Libby, W.F., Anderson, E.C. and Arnold, J.R., 1949. Age determination by radiocarbon content: world-wide assay of natural radiocarbon. Science, 109(2827), pp.227-228.

Lorenz, E.N., 1963. Deterministic nonperiodic flow. Journal of the Atmospheric Sciences, 20(2), pp.130-141.

Margold, M., Stokes, C.R. and Clark, C.D., 2015. Ice streams in the Laurentide Ice Sheet: Identification, characteristics and comparison to modern ice sheets. Earth-Science Reviews, 143, pp.117-146.

Martin, M.A., Winkelmann, R., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C. and Levermann, A., 2011. The Potsdam Parallel Ice Sheet Model (PISM-PIK)–Part 2: Dynamic equilibrium simulation of the Antarctic ice sheet. The Cryosphere, 5(3), pp.727-740.

McCabe, A.M., Clark, P.U., Clark, J. and Dunlop, P., 2007. Radiocarbon constraints on readvances of the British–Irish Ice Sheet in the northern Irish Sea Basin during the last deglaciation. Quaternary Science Reviews, 26(9-10), pp.1204-1211

Napieralski, J., Li, Y. and Harbor, J., 2006. Comparing predicted and observed spatial boundaries of geologic phenomena: Automated Proximity and Conformity Analysis applied to ice sheet reconstructions. Computers & geosciences, 32(1), pp.124-134.

Napieralski, J., Hubbard, A., Li, Y., Harbor, J., Stroeven, A.P., Kleman, J., Alm, G. and Jansson, K.N., 2007. Towards a GIS assessment of numerical ice sheet model performance using geomorphological data. Journal of Glaciology, 53(180), pp.71-83.

Nye, J.F., 1953. The flow law of ice from measurements in glacier tunnels, laboratory experiments and the Jungfraufirn borehole experiment. Proc. R. Soc. Lond. A, 219(1139), pp.477-489.

Ó Cofaigh, C. and Evans, D.J., 2007. Radiocarbon constraints on the age of the maximum advance of the British–Irish Ice Sheet in the Celtic Sea. Quaternary Science Reviews, 26(9-10), pp.1197-1203.

Patton, H., Hubbard, A., Andreassen, K., Winsborrow, M. and Stroeven, A.P., 2016. The build-up, configuration, and dynamical sensitivity of the Eurasian ice sheet complex to Late Weichselian climatic and oceanic forcing. Quaternary Science Reviews, 153, pp.97-121.

Patton, H., Hubbard, A., Andreassen, K., Auriac, A., Whitehouse, P.L., Stroeven, A.P., Shackleton, C., Winsborrow, M., Heyman, J. and Hall, A.M., 2017. Deglaciation of the Eurasian ice sheet complex. Quaternary Science Reviews, 169, pp.148-172.

Peltier, W.R., 2004. Global glacial isostasy and the surface of the ice-age Earth: the ICE-5G (VM2) model and GRACE. Annu. Rev. Earth Planet. Sci., 32, pp.111-149.

Peltier, W.R., Farrell, W.E. and Clark, J.A., 1978. Glacial isostasy and relative sea level: a global finite element model. Tectonophysics, 50(2-3), pp.81-110.

Piotrowski, J.A. and Tulaczyk, S., 1999. Subglacial conditions under the last ice sheet in northwest Germany: ice-bed separation and enhanced basal sliding?. Quaternary Science Reviews, 18(6), pp.737-751.

Quinlan, G. and Beaumont, C., 1982. The deglaciation of Atlantic Canada as reconstructed from the postglacial relative sea-level record. Canadian Journal of Earth Sciences, 19(12), pp.2232-2246.

Retzlaff, R. and Bentley, C.R., 1993. Timing of stagnation of Ice Stream C, West Antarctica, from short-pulse radar studies of buried surface crevasses. Journal of Glaciology, 39(133), pp.553-561.

Ritz, C., Edwards, T.L., Durand, G., Payne, A.J., Peyaud, V. and Hindmarsh, R.C., 2015. Potential sea-level rise from Antarctic ice sheet instability constrained by observations. Nature, 528(7580), pp.115-118.

Schoof, C., 2007. Ice sheet grounding line dynamics: Steady states, stability, and hysteresis. Journal of Geophysical Research: Earth Surface, 112(F3).

Schoof, C., 2012. Marine ice sheet stability. Journal of Fluid Mechanics, 698, pp.62-72.

Seierstad, I.K., Abbott, P.M., Bigler, M., Blunier, T., Bourne, A.J., Brook, E., Buchardt, S.L., Buizert, C., Clausen, H.B., Cook, E. Dahl-Jensen, D., Davies, S.M., Guillevic, M., Johnsen, S.J., Pedersen, D.S., Popp, T.J., Rasmussen, S.O., Severinghaus, J.P., Svensson, A. and Vinther, B.M., 2014. Consistently dated records from the Greenland GRIP, GISP2 and NGRIP ice cores for the past 104 ka reveal regional millennial-scale δ18O gradients with possible Heinrich event imprint. Quaternary Science Reviews, 106, pp.29-46.

Seguinot, J., Khroulev, C., Rogozhina, I., Stroeven, A.P. and Zhang, Q., 2014. The effect of climate forcing on numerical simulations of the Cordilleran ice sheet at the Last Glacial Maximum. The Cryosphere, 8(3), pp.1087-1103.

Seguinot, J., Rogozhina, I., Stroeven, A.P., Margold, M., Kleman, J., 2016. Numerical simulations of the Cordilleran ice sheet through the last glacial cycle. The Cryosphere 10, pp. 639-664.

Seguinot, J., Ivy-Ochs, S., Jouvet, G., Huss, M., Funk, M. and Preusser, F., 2018. Modelling last glacial cycle ice dynamics in the Alps. The Cryosphere, 12(10), pp.3265-3285.

Simpson, M.J., Milne, G.A., Huybrechts, P. and Long, A.J., 2009. Calibrating a glaciological model of the Greenland ice sheet from the Last Glacial Maximum to present-day using field observations of relative sea level and ice extent. Quaternary Science Reviews, 28(17-18), pp.1631-1657.

Small, D., Clark, C.D., Chiverrell, R.C., Smedley, R.K., Bateman, M.D., Duller, G.A., Ely, J.C., Fabel, D., Medialdea, A. and Moreton, S.G., 2017a. Devising quality assurance procedures for assessment of legacy geochronological data relating to deglaciation of the last British-Irish Ice Sheet. Earth-science reviews, 164, pp.232-250.

Small, D., Benetti, S., Dove, D., Ballantyne, C.K., Fabel, D., Clark, C.D., Gheorghiu, D.M., Newall, J. and Xu, S., 2017b. Cosmogenic exposure age constraints on deglaciation and flow behaviour of a marine-based ice stream in western Scotland, 21–16 ka. Quaternary Science Reviews, 167, pp.30-46.

Smedley, R.K., Scourse, J.D., Small, D., Hiemstra, J.F., Duller, G.A.T., Bateman, M.D., Burke, M.J., Chiverrell, R.C., Clark, C.D., Davies, S.M. Fabel, D., Gheorghiu, D.M., McCarroll, D., Medialdea, A., and Xu S., 2017. New age constraints for the limit of the British–Irish Ice Sheet on the Isles of Scilly. Journal of Quaternary Science, 32(1), pp.48-62

Stearns, L.A. and van der Veen, C.J., 2018. Friction at the bed does not control fast glacier flow. Science, p.eaat2217.

Stokes, C.R. and Clark, C.D., 1999. Geomorphological criteria for identifying Pleistocene ice streams. Annals of glaciology, 28, pp.67-74.

Stokes, C.R. and Tarasov, L., 2010. Ice streaming in the Laurentide Ice Sheet: A first comparison between data-calibrated numerical model output and geological evidence. Geophysical Research Letters, 37(L01501).

Stokes, C.R., Clark, C.D. and Storrar, R., 2009. Major changes in ice stream dynamics during deglaciation of the north-western margin of the Laurentide Ice Sheet. Quaternary Science Reviews, 28(7-8), pp.721-738.

Stokes, C.R., Tarasov, L., Blomdin, R., Cronin, T.M., Fisher, T.G., Gyllencreutz, R., Hättestrand, C., Heyman, J., Hindmarsh, R.C., Hughes, A.L., Jakobsson, M., Kirchner, N., Livingstone, S.J., Margold, M., Murton, J.B., Noormets, R., Peltier, R.W., Peteet, D.M., Piper, D.J.W., Preusser, F., Renssen, H., Roberts, D.H., Roche, D.M., Saint-Ange, F., Stroeven, A.P. and Teller, J.T., 2015. On the reconstruction of palaeo-ice sheets: recent advances and future challenges. Quaternary Science Reviews, 125, pp.15-49.

Stone, J.O., Balco, G.A., Sugden, D.E., Caffee, M.W., Sass, L.C., Cowdery, S.G. and Siddoway, C., 2003. Holocene deglaciation of Marie Byrd land, west Antarctica. Science, 299(5603), pp.99-102.

Stroeven, A.P., Hättestrand, C., Kleman, J., Heyman, J., Fabel, D., Fredin, O., Goodfellow, B.W., Harbor, J.M., Jansen, J.D., Olsen, L., Caffee, M.W., Fink, D., Lundqvist, J., Rosqvist, G.C., Strömberg, B. and Jansson, K.N., 2016. Deglaciation of Fennoscandia. Quaternary Science Reviews, 147, pp.91-121.

Tarasov, L. and Peltier, W.R., 2004. A geophysically constrained large ensemble analysis of the deglacial history of the North American ice sheet complex. Quaternary Science Reviews, 23(3-4), pp.359-388.

Tarasov, L., Dyke, A.S., Neal, R.M. and Peltier, W.R., 2012. A data-calibrated distribution of deglacial chronologies for the North American ice complex from glaciological modeling. Earth and Planetary Science Letters, 315, pp.30-40.

Tulaczyk, S., Kamb, W.B. and Engelhardt, H.F., 2000. Basal mechanics of ice stream B, West Antarctica: 1. Till mechanics. Journal of Geophysical Research: Solid Earth, 105(B1), pp.463-481.

Walcott, R.I., 1972. Late Quaternary vertical movements in eastern North America: Quantitative evidence of glacio- isostatic rebound. Reviews of Geophysics, 10(4), pp.849-884.

Weatherall, P., Marks, K.M., Jakobsson, M., Schmitt, T., Tani, S., Arndt, J.E., Rovere, M., Chayes, D., Ferrini, V. and Wigley, R., 2015. A new digital bathymetric model of the world's oceans. Earth and Space Science, 2(8), pp.331-345.

Winkelmann, R., Martin, M.A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C. and Levermann, A., 2011. The Potsdam parallel ice sheet model (PISM-PIK)-Part 1: Model description. The Cryosphere, 5(3), p.715.

Tables:

Table 1 – Summary of sources of data and comparison tools discussed in this paper.

Glaciological characteristic	**Model representation**	**Empirical data basis**	**BIIS data used in this study**	**Data-model comparison tool**
Margin-position.	Extent mask or determined from ice thickness.	Moraines (or other ice-contract/marginal landforms).	189 margin positions derived from the BRITICE v.2 compilation (Clark et al., 2018; Figure 5).	Automated Proximity and Conformity Analysis (APCA) (Napieralski et al., 2006; Li et al., 2008).
Ice flow direction.	Continuous field produced by model.	Subglacial bedforms, often grouped into flowsets (distinct flow events).	103 flowsets with 32 cross-cutting relationships (Greenwood and Clark, 2009; Hughes et al., 2014; Figure 6).	Automated Flow Direction Analysis (AFDA) (Li et al., 2007)
Timing of ice-free conditions.	Change in ice sheet extent mask, or ice thickness grid to 0 metres.	Geochronological data (mainly from Terrestrial Cosmogenic Nuclide, ^{14}C and Optically Stimulated Luminescence dating)	108 dated sites derived from previous literature (Small et al., 2017a; Figure 7). Only sites with Green or Amber quality rating are used (see Section 2.3).	Automated Timing Accordance Tool (ATAT) (Ely et al., in press).

Table 2. Multiple regression fields for climate. lat = latitude, lon = longitude, topg = surface topography (i.e. elevation in metres above present-day sea-level).

Simulation	Precipitation (mm/a)	Mean Annual temperature (°C)	July temperature (°C)	Source of climate data
A	374.6 + 10.1 x lat – 26.0 x lon	25.3 – 0.004 x topg – 0.294 x lat – 0.035 x lon	32.2 – 0.004 x topg – 0.316 x lat – 0.009 x lon	www.worldclim.org/
B	81.1 + 0.116 x lat – 1.502 x lon	35.8 – 0.005 x topg – 4.97 x lat – 0.07 x lon	34.2 – 0.004 x topg - 0.343 x lat + 0.112 x lon	www.cru.uea.ac.uk/data
C	159.8 – 16.545 x lat – 12.342 x lon	33.7 – 0.007 x topg – 0.674 x lat - 0.218 x lon	39.358 – 0.007 x topg – 0.621 x lat + 0.18 x lon	pmip3.lsce.ipsl.fr/

Table 3. Summary of results from model-data comparisons. Note that when measures are restricted to the modelled ice extent, the number of comparisons change, limiting the ability to compare between simulations.

Simulation	% of margins matched (n = 189)	% of margins matched within maximum modelled extent	% of margins matched within extent of simulation C	% of flowsets matched (n = 103)	% of flowsets matched within maximum modelled extent	% of flowsets matched within extent of simulation C	% of cross-cuts matched	% of dates where model-data agreement occurs (n = 108)	wRMSE of model-data difference for ice covered dates where model-data agreement occurs (years)
A	60	76 (n = 151)	61	9	21 (n = 41)	26	0	41	1898
B	36	54 (n = 125)	43	16	19 (n = 88)	21	0	9	1182
C	43	66 (n = 124)	66	3	8 (n = 39)	8	0	89	2057

Figures:

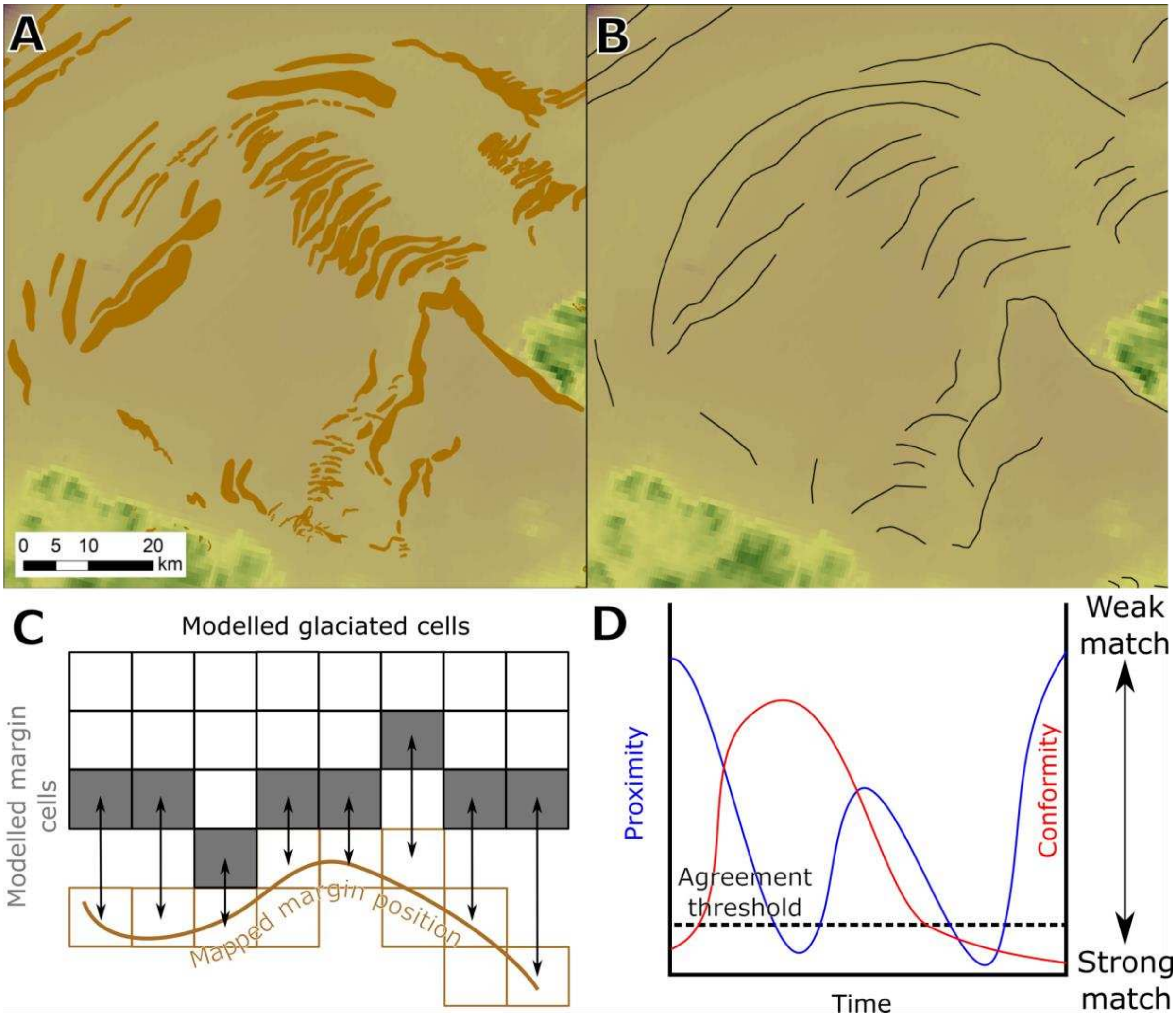

Figure 1. A) Mapped offshore moraines, Donegal Bay, Ireland, from Benetti et al. (2010). B) Interpreted margin positions from A. C) Schematic representation of the Automated Proximity and Conformity Analysis (APCA), whereby the distance between modelled and mapped margin position is measured. Proximity is defined as the mean of these measurements and conformity as the standard deviation (Napieralski et al., 2006; Li et al., 2008). D) Schematic output from APCA. Here, a model-data agreement is only declared when both proximity and conformity are below a defined threshold. Such instances are shaded in grey.

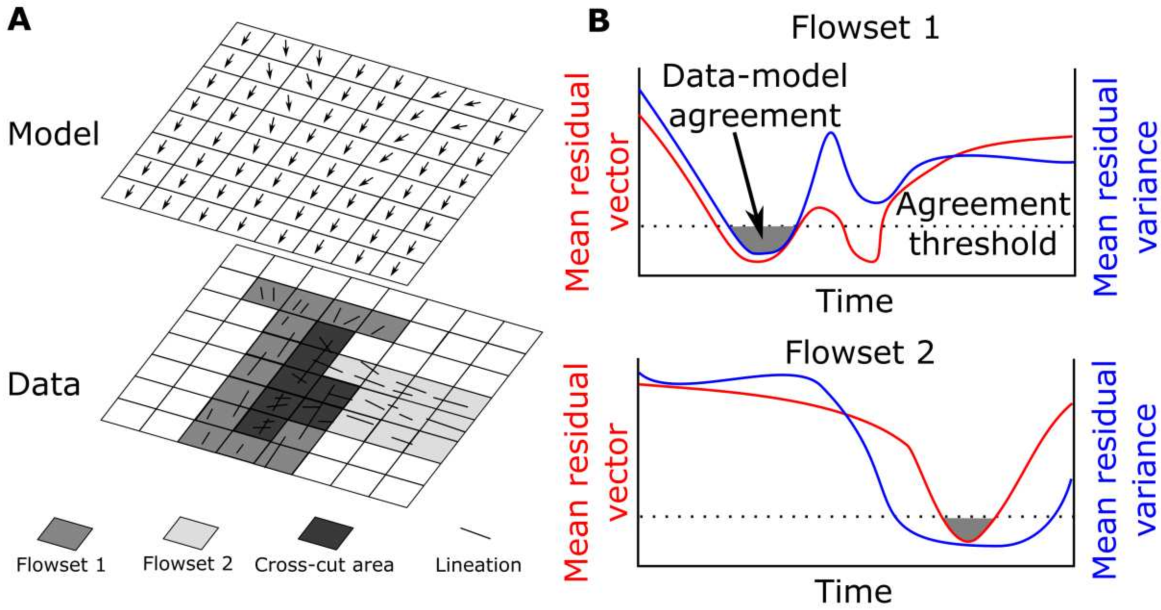

Figure 2. A) Schematic of Automated Flow Direction Analysis comparison technique (after Li et al., 2007). At this point in time, the model agrees well with Flowset 1, but is flowing at right angles to the superimposed Flowset 2. For complete model-data agreement to occur, the model must replicate the flow direction of flowset 2 at a later stage. B) Schematic output from AFDA for Flowsets 1 and 2 depicted in A. In this case, data-model agreement occurs when both mean residual variance and the mean residual vector are below an applied threshold. As this occurs in the observed sequence (Flowset 1 then Flowset 2), model data-agreement of this cross-cutting relationship can be said to occur.

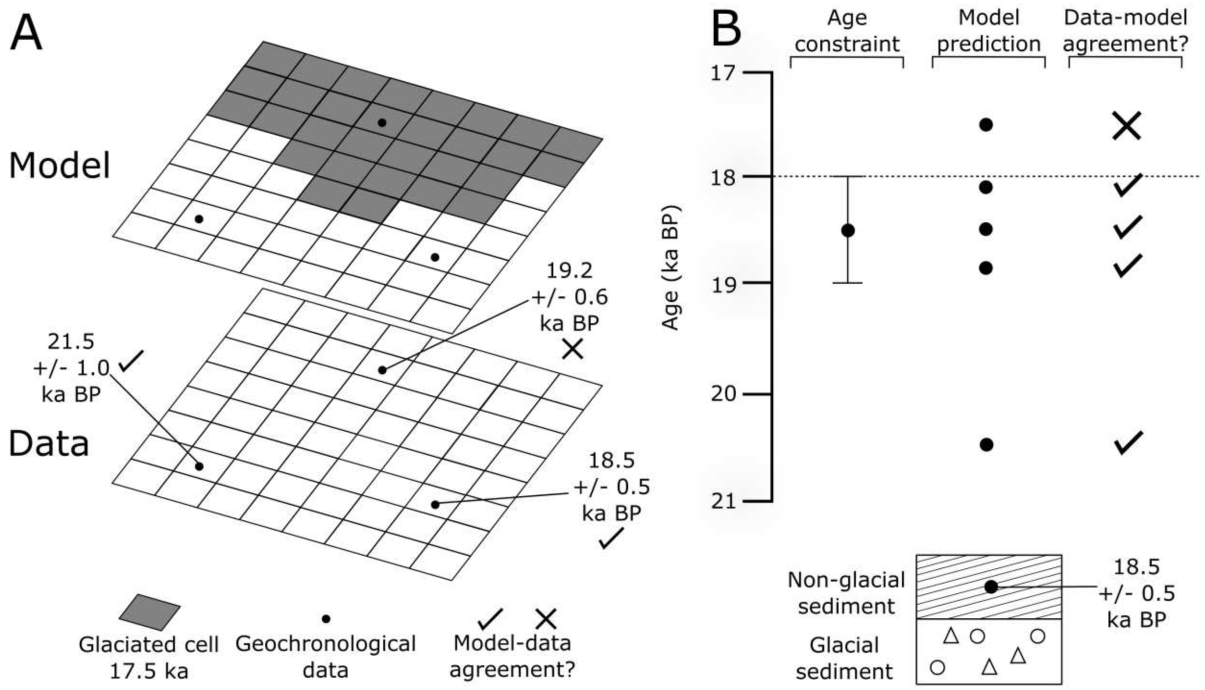

Figure 3. A) Schematic of the comparison between model and data made by ATAT (Ely et al., in press). Example shows a deglaciating ice sheet model output at 17.5 ka BP. The model replicates the ice-free conditions recorded by the lower two sites and thus there is model-data agreement. However, the model still produces ice cover at this time within the range of the date of 19.2 ± 0.6 ka BP. In this case, there is model-data disagreement. B) Example of comparison procedure for one site, dated to 18.5 ± 0.5 ka BP. Model predictions that occur before an ice-free age, or during the associated error, are considered to agree with the data. Adapted from Ely et al. (in press).

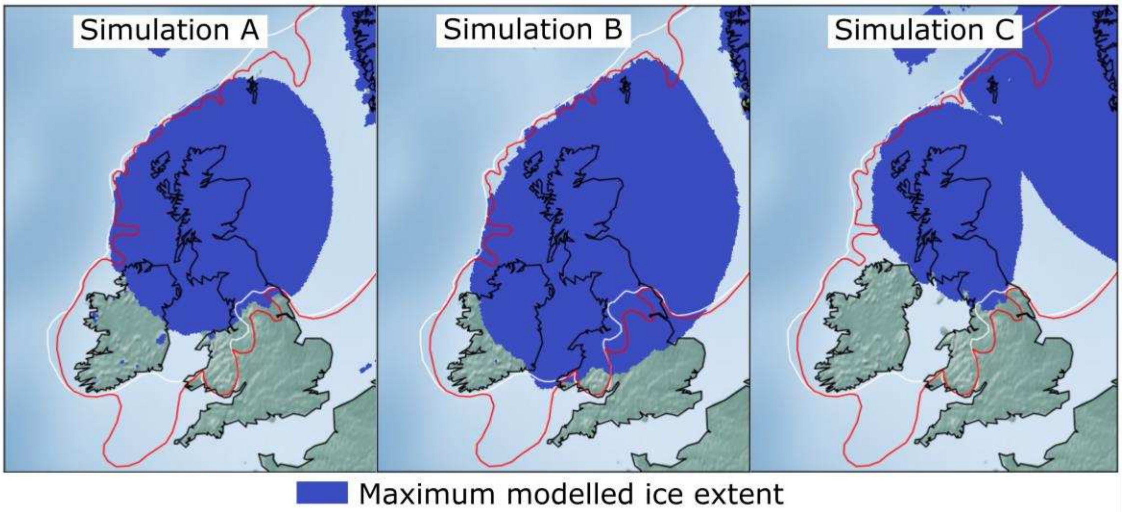

Figure 4. The maximum extent of the three model simulations. Note that these simulations are only driven by climate and are not calibrated to any empirical evidence of the ice sheet. Thus, they do not achieve a state which resembles the empirically reconstructed ice sheet. Reconstructed extents at 27 ka BP (white line) and 23 ka BP from Clark et al. (2012) are shown for comparison.

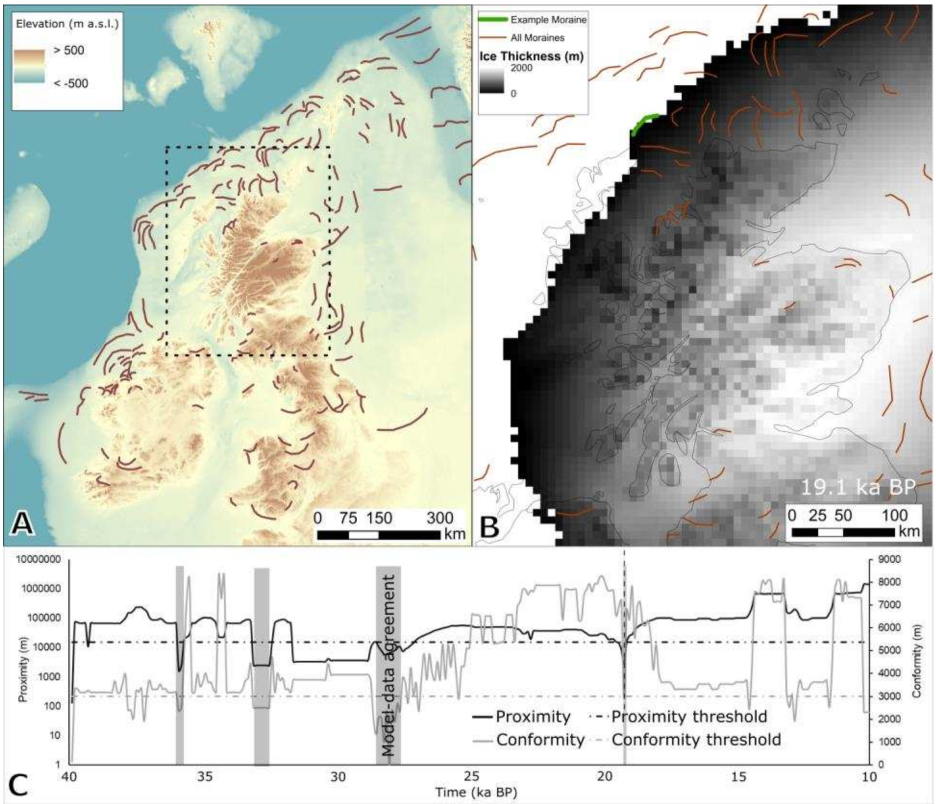

Figure 5. A) Generalised margin positions tested, derived from moraines reported in Clark et al. (2018). Merged bathymetry and topography from the General Bathymetric Chart of the Oceans 2014 grid (GEBCO; Weatherall et al., 2015). B) Modelled ice sheet thickness at 19.1 ka BP from Simulation A, centred on north-west Scotland with ice margin positions plotted on top. The example moraine considered in panel C is highlighted in green. Location of this panel is the dashed box on panel A. C) Output of proximity and conformity analysis for the example moraine shown in B for the duration of simulation A (40 ka BP to 10 ka BP). Note there are several periods of time when both proximity and conformity indicate model-data agreement, the most recent being at 19.1 ka BP. Note that the axis for "Proximity" is logarithmic.

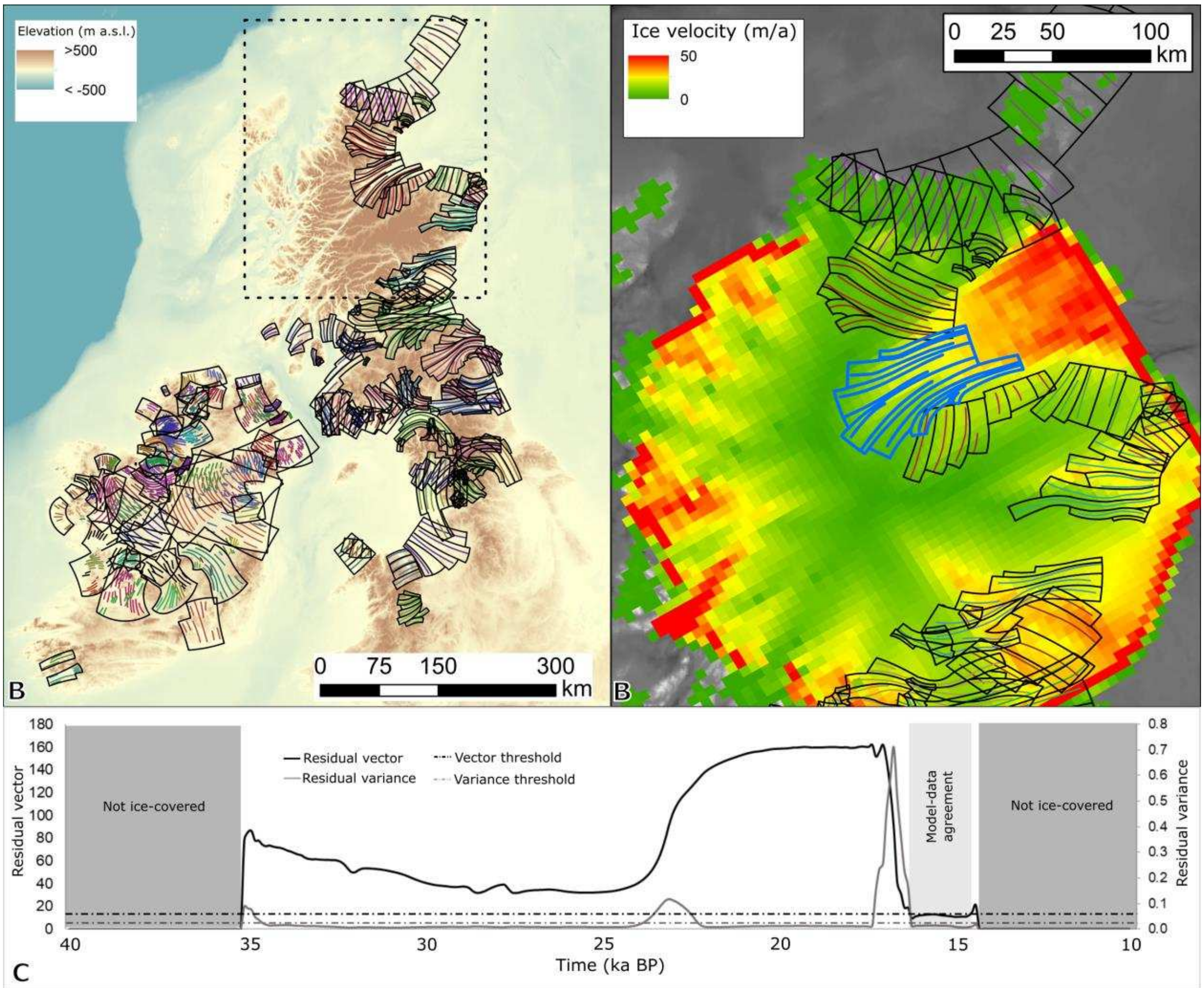

Figure 6. A) Flowsets used to compare to model simulations, with colours indicating different flowsets. Background from GEBCO (Weatherall et al., 2015). Overlapping regions are regions of cross-cutting (from Greenwood and Clark (2009a) and Hughes et al. (2014)). B) An example of a matched flowset, highlighted in blue, from simulation B at 17.1 ka BP. Other flowsets are indicated by coloured lines encompassed by black boxes. This panel is located by the dashed box on panel A. C) Output from AFDA for model simulation B (40 ka BP to 10 ka BP), showing periods of model-data agreement over time for the flowset shown in (B).

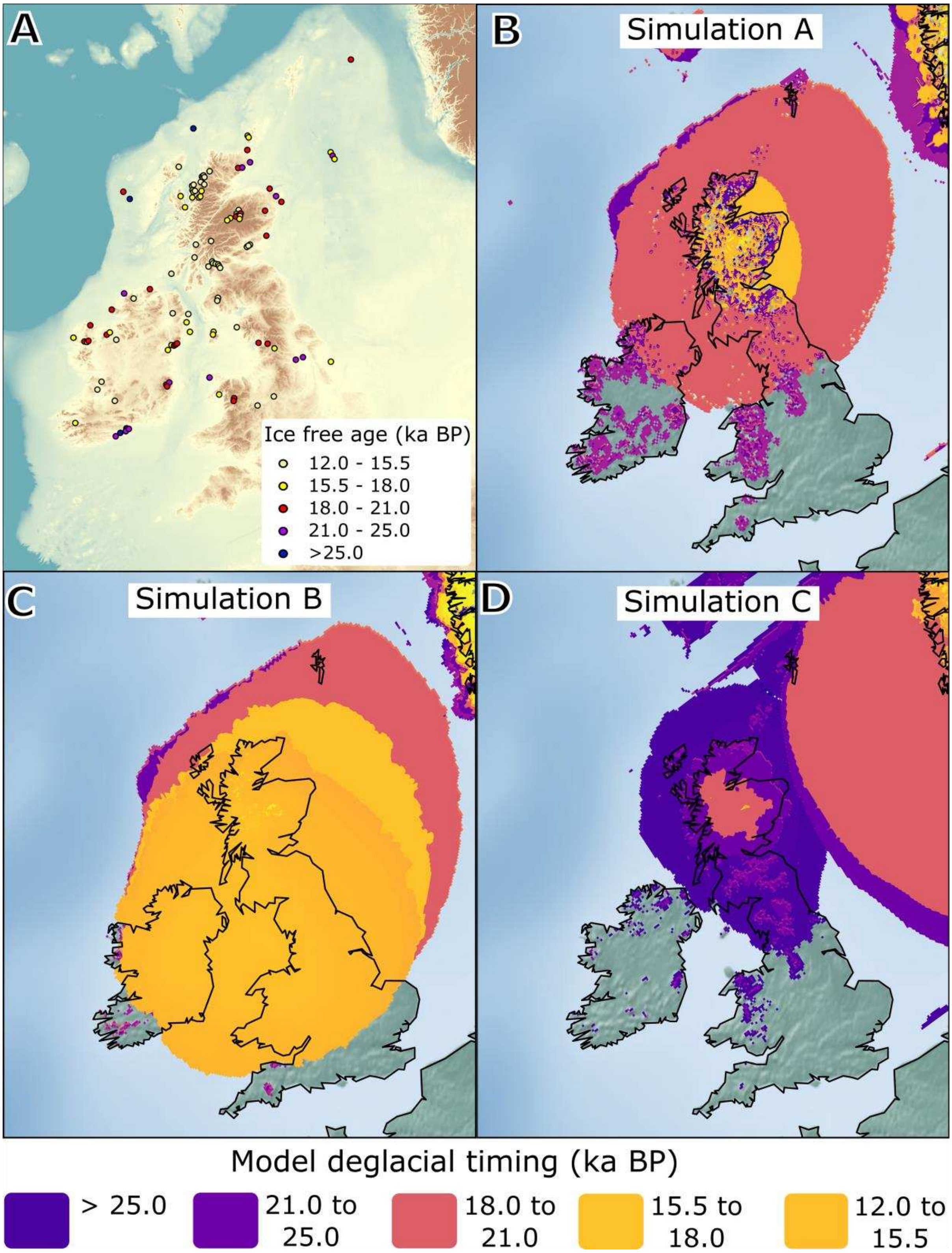

Figure 7. A) Dated locations assembled from Small et al. (2017a) that have a quality rating of green or amber. B to D) Simulated timing of ice-free conditions from model simulations A to C. Note that these simulations are uncalibrated to any empirical evidence, and a better fit may be achieved by tuning parameters and boundary conditions.

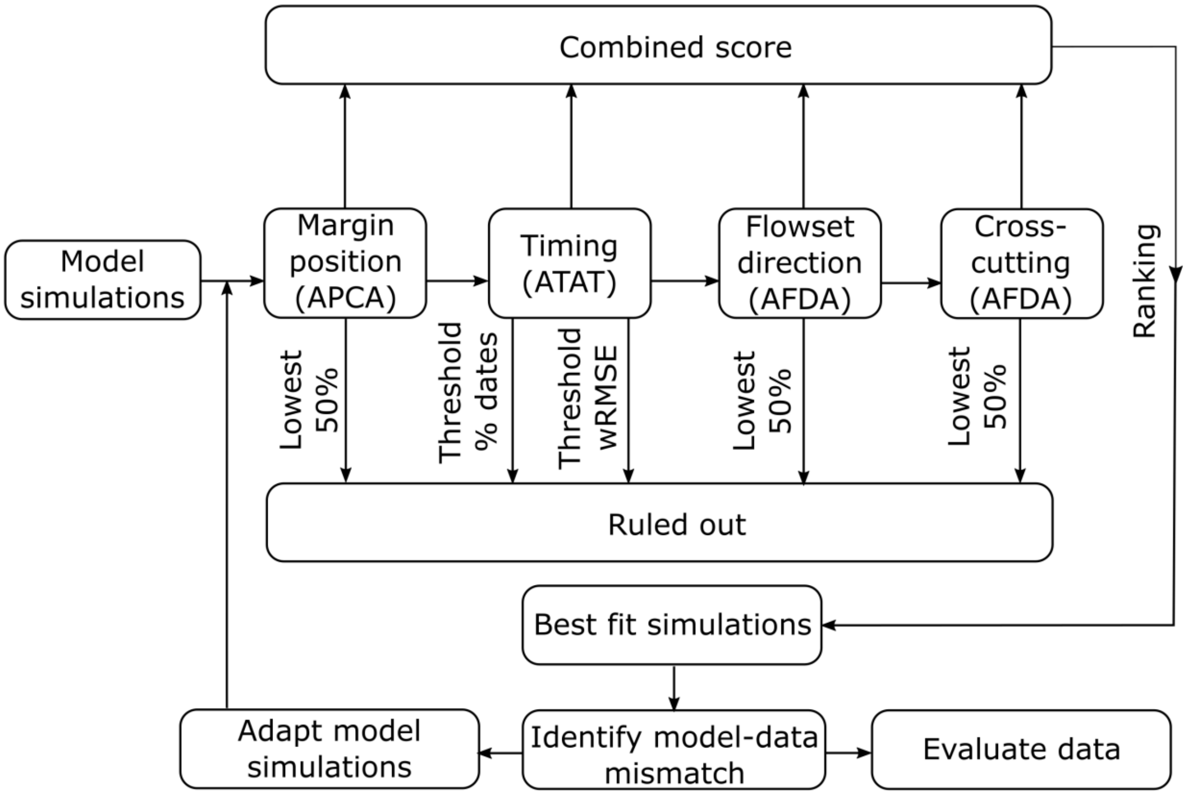

Figure 8. Proposed procedure for comparing multiple model-runs to geochronological data.