



This is a repository copy of *Recent progress on combining geomorphological and geochronological data with ice sheet modelling, demonstrated using the last British-Irish Ice Sheet.*

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/143670/>

Version: Accepted Version

---

### Article:

Ely, J. [orcid.org/0000-0003-4007-1500](https://orcid.org/0000-0003-4007-1500), Clark, C., Hindmarsh, R. et al. (8 more authors) (2021) Recent progress on combining geomorphological and geochronological data with ice sheet modelling, demonstrated using the last British-Irish Ice Sheet. *Journal of Quaternary Science*, 36 (5). pp. 946-960. ISSN 0267-8179

<https://doi.org/10.1002/jqs.3098>

---

This is the peer reviewed version of the following article: Ely, J.C., Clark, C.D., Hindmarsh, R.C.A., Hughes, A.L.C., Greenwood, S.L., Bradley, S.L., Gasson, E., Gregoire, L., Gandy, N., Stokes, C.R. and Small, D. (2019), Recent progress on combining geomorphological and geochronological data with ice sheet modelling, demonstrated using the last British–Irish Ice Sheet. *J. Quaternary Sci.*, which has been published in final form at <https://doi.org/10.1002/jqs.3098>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

### Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

### Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

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

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

7  
8       1 – Department of Geography, The University of Sheffield, Sheffield, S10 2TN, UK

9       2 – British Antarctic Survey, High Cross, Madingley Road, Cambridge, CB3 0ET, UK

10      3 - Department of Earth Science, University of Bergen and Bjerknes Centre for Climate  
11      Research, Bergen, Norway

12      4 - Department of Geological Sciences, Stockholm University, Stockholm, Sweden

13      5 - Department of Geoscience and Remote Sensing, Delft University of Technology,  
14      Stevinweg 1, 2628 CN Delft, Netherlands

15      6 – School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8  
16      1SS, UK

17      7 - University of Leeds, School of Earth and Environment, Woodhouse Lane, Leeds, LS2  
18      9JT, UK

19      8 – Durham University, Department of Geography, Lower Mountjoy, South Road, Durham,  
20      DH1 3LE, UK

21  
22      **Abstract**

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

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

40

## 41 **1. Introduction**

42 Reconstructing the behaviour of palaeo-ice sheets enables a better understanding of the  
43 long-term (centennial to millennial) behaviour of ice sheets in the Earth system. The former  
44 extent and behaviour of ice sheets can be inferred principally from four main lines of  
45 evidence. First, relative sea-level (RSL) records (e.g. a raised beach or salt marsh) provide  
46 constraints on the loading history of an ice sheet. Through the application of a glacio-isostatic  
47 adjustment (GIA) model, RSL data can be used to infer palaeo-ice sheet thickness and extent  
48 (e.g. Lambeck and Chappell, 2001; Peltier, 2004; Bradley et al., 2011). Secondly, analysis of  
49 the properties and stratigraphic sequence of sediments transported and deposited by palaeo-  
50 ice sheets can be used to infer ice sheet history at a given location (e.g. Eyles and McCabe,  
51 1989; Piotrowski and Tulaczyk, 1999). The geomorphological record, composed of  
52 landforms such as drumlins and moraines, can be used to decipher former ice-flow directions  
53 and margin positions (e.g. Hughes et al., 2014; Clark et al., 2018). Finally, the timing of  
54 deposition of sediment and/or the time glacially transported or eroded bedrock has been  
55 exposed, and by inference the timing of formation of associated landforms, can be dated  
56 using laboratory-based techniques to produce the third line of evidence, geochronological  
57 data (e.g. Libby et al., 1949; Duller, 2006; Small et al., 2017a).

58 The body of empirical evidence related to palaeo-ice sheets is continually growing,  
59 producing an ever-expanding library of palaeo-ice sheet data (e.g. Dyke, 2004; Clark et al.,  
60 2012; Hughes et al., 2016; Stroeven et al., 2016). Producing a glaciologically-plausible  
61 *empirical reconstruction* of a palaeo-ice sheet is, however, a challenging process, with three  
62 main limitations. First, evidence is often temporally and spatially fragmented, thereby  
63 requiring some subjective inference to be made about ice sheet behaviour between the data-  
64 constraints (Clark et al., 2012; Hughes et al., 2016). Secondly, all sources of data have

65 inherent uncertainties due to factors such as preservation potential, inherent laboratory-based  
66 uncertainties and post-depositional modification (Hughes et al., 2016; Small et al., 2017a).  
67 Finally, a mathematically- and physically-based direct inversion from palaeo-glaciological  
68 information to infer past-ice sheet characteristics (e.g. former ice-flow velocities) has  
69 remained elusive owing to the complexity of processes involved, meaning that all  
70 reconstructions are subjective (albeit expert) inferences (Kleman and Borgström, 1996;  
71 Stokes et al., 2015). Despite these limitations, empirical reconstructions typically provide a  
72 spatially-coherent representation of ice sheet activity, often portrayed as a series of  
73 palaeogeographic maps showing ice extent, flow geometry, ice divides and their changes at  
74 any given time (or at several time-steps).

75 As an alternative to the data-driven approach of an empirical reconstruction, numerical  
76 ice sheet models can be used to reconstruct palaeo-ice sheet behaviour (e.g. Fisher et al.,  
77 1985; Tarasov and Peltier, 2004; Hubbard et al., 2009; Patton et al., 2017). The approach here  
78 is to apply a numerical model based on the understanding of ice sheet physics to produce a  
79 *modelled reconstruction* of a palaeo-ice sheet. Using this physics-based approach,  
80 information such as ice-thickness and velocity can be reconstructed across the entire model  
81 domain in a manner that is consistent with model physics. However, limitations with this  
82 approach mean that modelled reconstructions may struggle to replicate the information and  
83 detail provided by palaeo-data. Numerical ice sheet models require the specification of  
84 several input boundary conditions and parameters. One of the most uncertain of these is the  
85 climatic conditions used to determine the pattern of accumulation and ablation over the  
86 model domain through time (Stokes et al., 2015). Other factors relating to the nature of ice  
87 sheet flow, such as basal friction, subglacial hydrology and shear, may either rely upon  
88 poorly constrained model parameters (due to a lack of physical understanding), or simply be  
89 beyond the capabilities of the model (e.g. they operate at scales below the spatial resolution  
90 of the model). Compounding the problem, ice sheets exhibit instabilities, whereby small  
91 perturbations to boundary conditions are amplified by the instability and can affect the whole  
92 modelled ice sheet. Such instabilities may lead to highly non-linear responses that are  
93 difficult to predict. One example is the marine ice sheet instability (Hughes, 1973; Schoof,  
94 2007; 2012), which is an instability in the position of the grounding-line on a reverse bed  
95 slope that occurs as a consequence of ice-flux being proportional to ice-thickness at the  
96 grounding-line.

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

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

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

123           To account for the above limitations and uncertainties of modelled-reconstructions, two  
124 general approaches have been adopted which produce multiple ice sheet simulations. The  
125 first involves sensitivity analyses (e.g. Boulton and Hagdorn, 2006; Patton et al., 2016),  
126 whereby relevant model parameters and boundary conditions are perturbed to produce  
127 numerous simulations of the palaeo-ice sheet in question. Such tuning is conducted until a  
128 simulation is generated that is perceived to ‘best fit’ the empirical evidence, and is chosen as

129 the modelled reconstruction. The second adopts an ensemble approach (e.g. Tarasov and  
130 Peltier, 2004; Gregoire et al., 2012), whereby a wide set of plausible combinations of  
131 parameters are input into the ice sheet model to produce an array of model-outputs. Data-  
132 based constraints may then be used to rule out unrealistic simulations from the bank of  
133 ensemble simulations, leaving a combination of simulations that are yet to be ruled out (e.g.  
134 Gregoire et al., 2016). The second approach is to calibrate ensemble parameters against data  
135 constraints, ruling out simulations and their associated parameter sets based on acceptable fits  
136 to the data (e.g. Tarasov and Peltier, 2004). The remaining simulations are then supplemented  
137 by further simulations, which use the calibrated parameters. The final modelled  
138 reconstruction in this approach is a combination of calibrated model simulations, from which  
139 the distribution of plausible glaciological variables can be derived (e.g. mean ice velocity)  
140 (Tarasov et al., 2012).

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

150 Adopting this approach may create new opportunities for both empiricists and ice sheet  
151 modellers to drive the field forward. Empiricists could use models to help reduce data  
152 uncertainty and rule out physically-implausible interpretations. Modellers could use the data  
153 to score ensemble members and improve model formulation (as per Tarasov and Peltier,  
154 2004). Here, we extend some recent advances in this area to outline a procedure for  
155 comparing geochronological and geomorphological data with ice sheet model output. We  
156 illustrate this with example model output of the British-Irish Ice Sheet (BIIS). Given the  
157 expanding body of data constraining palaeo-ice sheet behaviour (e.g. Greenwood and Clark,  
158 2009; Hughes et al., 2014; Small et al., 2017a; Clark et al., 2018), it is one of the best ice  
159 sheets for model-data comparison. The primary purpose of the model runs presented here is  
160 not to simulate the intricacies of this palaeo-ice sheet or advance our understanding of the ice  
161 sheet, but simply to facilitate methodological comparisons between model output and

162 empirical data. Meaningful and more accurate simulations of the ice sheet are the subject of  
163 ongoing work as part of the BRITICE-CHRONO NERC consortium project (e.g. Gandy et  
164 al., 2018).

## 165 **2. Methods of model-data comparison**

166 Of the four sources of data that might be used to constrain palaeo-ice sheet simulations  
167 (RSL, sedimentology, geochronology, and geomorphology), it is perhaps not surprising that  
168 RSL has the longest tradition (Walcott, 1972; Peltier et al., 1978; Quinlan and Beaumont,  
169 1982). Sea-level index points provide a testable dataset with definable uncertainty (e.g.  
170 Engelhart and Horton, 2012). Furthermore, until recently, ice sheet models were run at a low-  
171 resolution of >20 km grid size. This meant that modelled reconstructions could be tested  
172 against relative-sea level data, which has a lack of abrupt spatial changes, through the use of a  
173 GIA model (e.g. Auriac et al., 2016). The advent of faster and parallel processing means that  
174 higher-resolution simulations of continental ice sheets are now achievable (~5 km),  
175 permitting comparison to other sources of information. However, these data need to be  
176 presented at a similar resolution to the model and will perhaps provide definitive and  
177 quantifiable characteristics that a model can predict. Ice sheet models are yet to have  
178 adequate sediment production, transportation, and deposition laws to make predictions to the  
179 same level of detail that might be observed in a sediment exposure. We here demonstrate how  
180 to make meaningful model-data comparisons to the remaining two classes of palaeo-ice sheet  
181 data, geomorphological (ice-margin position and ice-flow direction) and geochronological (in  
182 essence, the timing of ice-free conditions).

### 183 *2.1. Ice-Margin Position*

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

191 Napieralski et al. (2006) developed an Automated Proximity and Conformity Analysis  
192 (APCA) tool for comparing margin positions from mapped moraines and ice sheet model  
193 outputs (Table 1), later modified by Li et al. (2008). In this tool, mapped margins are first

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

## 211 *2.2. Ice-flow direction*

212 Subglacial bedforms record the ice-flow directions within a palaeo-ice sheet (e.g.  
213 Kleman, 1990; Clark, 1993; Kleman and Borgström, 1996; Stokes et al., 2009; Ely et al.,  
214 2016). Where cross-cutting subglacial bedforms are superimposed on each other, a sequence  
215 of flow directions is recorded (Clark, 1993). Neighbouring subglacial bedforms with a similar  
216 morphology and orientation can be grouped into flowsets – groups of subglacial bedforms  
217 interpreted to form in the same phase of ice-flow (e.g. Kleman and Borgström, 1996; Clark,  
218 1999). When grouped in this way, cross-cutting flowsets of subglacial bedforms can reveal  
219 major shifts in the flow patterns of an ice sheet, a consequence of shifting ice sheet geometry,  
220 ice-divide migration and ice-stream (de)activation (e.g. Boulton and Clark, 1990; Clark,  
221 1999; Greenwood and Clark, 2009). Whilst a single flowset provides a spatially limited  
222 constraint on ice-flow direction, the sequence and spatial patterning of flowsets across the  
223 former ice sheet bed can be used to reconstruct the ice-flow geometry of a palaeo-ice sheet  
224 and the evolution of that geometry through time (Boulton and Clark, 1990; Kleman et al.,  
225 1997; Greenwood and Clark, 2009; Hughes et al., 2014).



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

### 236 *2.3. Ice-free timing*

237 The timing of ice-free conditions can be derived from geochronological techniques.  
238 These have been applied most commonly to organic material in the case of radiocarbon  
239 dating (Libby et al., 1949; Arnold and Libby, 1951; Ó Cofaigh and Evans, 2007; McCabe et  
240 al., 2007), proglacial sands in the case of luminescence dating (e.g. Duller, 2006; Smedley et  
241 al., 2017; Bateman et al., 2018), and glacially transported boulders or glacially modified  
242 bedrock in the case of cosmogenic nuclide dating (Stone et al., 2003; Fabel et al., 2012; Small  
243 et al., 2017b). For some palaeo-ice sheets, compilations of thousands of dates recording ice-  
244 free conditions relevant to the timing of advance and retreat exist (Dyke, 2004; Hughes et al.,  
245 2016; Small et al., 2017a). However, dating the activity of an ice sheet is complex and, as  
246 such, not all dates are equally reliable constraints (Small et al., 2017a). To account for this, an  
247 assessment of data reliability, such as the traffic-light system proposed by Small et al.  
248 (2017a), should be conducted prior to model-data comparison. This involves initially filtering  
249 out ages irrelevant to the study period. The remaining ages are then assigned a quality rating  
250 based upon the stratigraphic and geomorphological context, supporting evidence and  
251 potential for significant and unquantifiable geological uncertainty (Small et al. 2017a).  
252 Depending on the stratigraphic setting of a dated sample (e.g. above or below glacial  
253 sediment), this timing constrains ice free conditions either prior to an advance of, or  
254 following the retreat of, an ice sheet (Hughes et al., 2011). Each site has an associated error,  
255 related to measurement uncertainties. Since geochronological techniques only record the  
256 timing of ice-free conditions prior to (advance) or after (retreat) the occupation of an area by  
257 an ice sheet, the associated error can be considered as one-sided (Figure 3; Briggs and  
258 Tarasov, 2013; Ely et al., in press).

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

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

#### 283 *3.1. Model-setup*

284 Our primary aim is to demonstrate various approaches to model-data comparison, and  
285 so we perform some simple experiments with the aim of creating a range of outputs. We  
286 therefore make numerous simplifications, especially regarding our climate input. It is  
287 unimportant for the model experiments to exactly replicate the detailed reconstructed history  
288 of the BIIS (e.g. Clark et al., 2012). However, the model output serves as a means for  
289 demonstrating how model-data comparison tools could work. We use the Parallel Ice Sheet  
290 Model (PISM; Winkelmann et al., 2011) to simulate the BIIS. PISM is a hybrid shallow-ice

291 shallow-shelf model which implements grounding line migration using a subgrid  
292 interpolation scheme. Ice movement is modelled as a combination of ice deformation and  
293 basal sliding. Internal deformation is determined by a flow law (Glen, 1952; Nye, 1953) with  
294 ice rheology altered by an enthalpy scheme (Aschwanden et al., 2012). Basal sliding occurs  
295 through a pseudo-plastic sliding law once basal shear stresses exceed yield stresses. Yield  
296 stress is determined to be a function of till friction, with till friction being a function of  
297 elevation and modelled basal effective pressure (Martin et al., 2011). Effective pressure is  
298 determined by a local subglacial hydrology model which relates overburden pressure to  
299 subglacial melt rates whilst ignoring horizontal water transport (Tulaczyk et al., 2000; Bueler  
300 and van Pelt, 2015). The model allows ice-shelves to form. Sub-shelf melt is determined  
301 using the parameterisation of Beckmann and Goosse (2003) perturbed by a melt factor  
302 (Martin et al., 2011), assuming that basal ice temperature is at pressure-melting point and  
303 ocean temperatures are at the freezing point at the depth of the ice-ocean interface (Martin et  
304 al., 2011). Calving rates are proportional to horizontal strain rates and are determined by a 2D  
305 parameterisation (Levermann et al., 2012; see also Supplementary Table 1 for key  
306 parameters).

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

316 To demonstrate differences between model simulations, we limit our analyses to the  
317 output from three model simulations. Parameters and boundary conditions are the same for all  
318 three simulations, with the exception that we vary the climate input. Climate is represented in  
319 our simulation as a spatially continuous field derived from multiple regression analysis of  
320 three sources of climate data; two modern day records and one from a palaeo-climate  
321 modelling experiment (Table 2; Braconnot et al. 2012). Prescribed temperatures are perturbed  
322 over time by a scalar offset derived from the Greenland ice core records (Seierstad et al.,  
323 2014) and fed into a positive degree day model to calculate surface mass balance (Calov and

324 Greve, 2005). Precipitation is also corrected with reference to the Greenland ice core record,  
325 with a 7.3% reduction in precipitation per degree Celsius decrease in temperature  
326 (Huybrechts, 2002). The model runs from 40 thousand years before present (ka BP) to the  
327 present day. Model output was recorded at 100-year intervals. The maximum extent of ice  
328 generated by each model simulation is shown in Figure 4. As expected, none of the model  
329 simulations perform well at replicating the reconstructed extent of the BIIS (e.g. Clark et al.,  
330 2012; Figure 4). The inability to reach these extents is most likely a consequence of the  
331 simplistic climate forcing and would therefore likely be ruled out by visual assessment alone.  
332 Such visual assessment is time consuming, especially considering that an ensemble is likely  
333 to produce thousands of model simulations. Furthermore, it may be that the parameters used  
334 in one simulation produce a closer fit to the data than others, guiding future models. It is  
335 therefore important to test model-data tools against these simulated ice sheets.

### 336 *3.2. Ice Margin Position*

337 We derived 189 ice margin positions from moraines reported in the BRITICE v.2  
338 database (Clark et al., 2018) and compared these using APCA (Li et al., 2008) against our  
339 modelled ice-margin positions (Figure 5). To determine reasonable thresholds of proximity  
340 and conformity beyond which model-data agreement can be declared, we conducted  
341 sensitivity analysis validated by visual inspection (Figure 5B). We found that a proximity  
342 threshold of 15 km and a conformity threshold of 3 km sufficiently identified modelled ice  
343 margin positions that visually agreed with the shape and location of each moraine (Figure 5B,  
344 5C). These thresholds could be used in similar experimental setups. A similar proximity  
345 measure (15 km) was reported by Napieralski et al. (2007). Figure 5B shows an example of a  
346 margin position where data-model agreement occurred. Data-model agreement occurred  
347 several times during the course of the simulation for this particular margin, as both measures  
348 of proximity and conformity fell below the agreement threshold on multiple occasions  
349 (Figure 5C). Marine based ice sheets, such as the BIIS, are prone to readvances (Schoof,  
350 2007; Kingslake et al., 2018). The potential to readvance means that we cannot make the  
351 simple assumption that moraines closer to the ice sheet centre are older, meaning that we do  
352 not consider time sequences of margin occupation as a test here.

353 Table 3 shows the percentage of margins matched by each model run. The most  
354 common reason for model-data mismatch was that margins were not reached by the  
355 simulated ice extent, meaning that they scored too low on the proximity score of APCA. This

356 is unsurprising given that 2 out of 3 of the models do not reach the extent of all considered  
357 margins (Figures 4, 5A). To test whether the model agrees with the observed shape and  
358 proximity of margins that are within modelled extent, we calculated a second statistic, which  
359 considered only those observed margins within the maximum extent of a given model  
360 simulation (Figures 4, 5A; Table 3). This shows that each simulation has model-data  
361 agreement with over 50% of the margins reached and their shape replicated by the model  
362 simulation (i.e. excluding mismatches for margins that are outside the maximum extent of the  
363 model simulation) (Table 3). However, direct comparisons between simulations become  
364 problematic when restricting the analysis to only moraines within the maximum extent, as  
365 this changes the number of data that are being compared (Table 3). We therefore created a  
366 third metric, the extent of margins matched within the extent Simulation C, the simulation  
367 which produced the smallest ice extent (Table 3; Figure 4).

368 3.3. *Ice-flow direction*

369 A total of 103 flowsets from Britain and Ireland were compared to our model  
370 simulations using AFDA (Li et al., 2007) (Figure 6A). These were assembled from  
371 Greenwood and Clark (2009) and Hughes et al. (2014) and include 32 cross cutting  
372 relationships. Combined, the datasets of Greenwood and Clark (2009 and Hughes et al.  
373 (2014) have over 150 flowsets. However, given the horizontal resolution of the models (5  
374 km), small (<20 km wide) flowsets were excluded from the analysis (n = 39). Flowsets  
375 identified as times-transgressive (i.e. formed asynchronously) were either divided into the  
376 stages of formation identified by Greenwood and Clark (2009) and Hughes et al. (2014), or  
377 excluded from the analysis (n = 20). Flow vectors were derived from the empirically-derived  
378 depiction of a flowset, rather than individual bedforms, because the orientation of these may  
379 vary on a sub-grid scale. For data-model agreement to occur, we applied a threshold of 10°  
380 mean residual vector, and 0.03 in mean variance. These values were initially derived by  
381 visually comparing the model and data and determining whether a modelled ice flow  
382 direction was sufficiently similar to a mapped flowset. These threshold values are consistent  
383 with those reported by Napieralski et al. (2007), and could be used to declare model-data  
384 agreement in similar experimental setups. To get a cross-cutting relationship registered to be  
385 in data-model agreement, the last occurrence of model conformity for the first flowset in a  
386 sequence needs to occur before the last occurrence of model conformity for the overprinted  
387 flowset.

388

389 Table 3 summarises the comparison between model output from the three Simulations  
390 and the assembled flowset database (Figure 6A). Overall, model-data agreement was low,  
391 with the majority of flowsets not replicated by the model simulations (Table 3). Similar to the  
392 margin comparison, this is partly a consequence of the models computed ice-covered area not  
393 replicating the full area covered by the BIIS (Figure 4). We therefore produced a second  
394 metric that restricted the analysis to those flowsets occurring within the modelled ice-extent.  
395 This was done to see if model-data mismatch was a consequence of ice-extent (in which a  
396 high number of ice-covered data points would be matched), or due to model-data mismatch  
397 even over the ice-covered area. However, note the caveat that this limits the ability to  
398 compare between simulations owing to the changing number of data in the model-data  
399 comparison. A third metric, the percentage of flowsets matched within the extent of

400 simulation C (the simulation with the smallest ice extent), allows for comparison between  
401 model runs. Even when this approach is adopted, the degree of model-data agreement for  
402 flowsets remains low, with simulation A being the best performing, matching 26% of  
403 flowsets within the extent of simulation C (Table 3). Furthermore, no models were able to  
404 replicate an observed cross-cutting relationship (Table 3). Figures 6B and C demonstrate an  
405 example of a matched flowset. Here, ice flow of sufficient coherence (a variance measure) in  
406 an agreed direction (vector orientation measure) is achieved toward the end of the model run  
407 (Figure 6C).

408

### 409 *3.4. Ice-free Timing*

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

423 Simulation B performs poorly in replicating the timing of ice-free conditions, with data-  
424 model conformity occurring for only 9% of the dates (Table 3). Simulations A and C have  
425 higher scores of this metric, with 41% and 89% of the dates agreeing with the modelled  
426 timing of ice-free conditions, respectively (Table 3). However, these model runs also have  
427 high wRMSE scores (Table 3), meaning that although ice-free conditions correctly occur,  
428 they are far from the mean age recorded by the geochronological data. For example, in  
429 Simulation C this indicates that although model-data agreement has occurred (i.e. the model  
430 has deglaciated an area before the empirical evidence indicates ice-free conditions), the

431 timing of modelled ice-free conditions is ~2000 years earlier on average than that recorded in  
432 the data. This pattern of premature deglaciation is apparent in Figure 7D.

## 433 **4. Discussion**

### 434 **4.1. Model-data fit**

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

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

461 All three model simulations do not replicate the maximum extent of the BIIS derived  
462 from observations. The maximum extent of an ice sheet is generally well known, and some of



463 these moraines record the maximum extent across different sectors of the BIIS (e.g. Bradwell  
464 et al., 2008; Clark et al., 2012). Therefore, future work may adopt a procedure of testing ice  
465 sheet models against only those margins derived from moraines which demark maximum  
466 palaeo-ice sheet extent and glaciated continental shelf-breaks (e.g. Seguinot et al., 2016;  
467 Patton et al., 2017), so as to identify simulations and glacio-climatic parameter combinations  
468 which achieve a reasonable ice sheet extent, before attempting to replicate margin positions  
469 occupied during ice retreat. A model which fits maximum ice extent margins in some places,  
470 may be able to interpolate between these constraints in a more consistent manner than  
471 empirical interpretations (e.g. Bowen et al., 1986; Clark et al., 2012; Seguinot et al., 2016;  
472 Patton et al., 2017).

473 All three simulations performed poorly at replicating the flow direction recorded by  
474 subglacial bedforms (Table 3). This is surprising given that the direction of many flowsets  
475 appears to be governed by the subglacial topography in Britain (Hughes et al., 2014), which  
476 is also likely to steer ice flow directions in numerical models that use that topography. One  
477 possibility is that this is due to the coarse (5 km) resolution of our model grid. Perhaps this  
478 model-data mismatch is also a consequence of the model being unable to fully replicate other  
479 conditions which determine ice flow direction such as basal thermal regime, subglacial  
480 hydrological conditions and the overall ice-sheet geometry (e.g. location of ice divides and  
481 domes). Areas with subglacial bedforms indicate warm-based ice, where basal  
482 sliding/subglacial till deformation is the dominant control upon ice-discharge. The most  
483 common reason for model-data mismatch in flow direction was the low mean residual  
484 variance scores. In other words, the model did not produce consistent flow-directions across  
485 the entire area of the flowset. Therefore, model-data mismatch is at least partially due to the  
486 model being unable to adequately simulate the dimensions of ice-streams and outlet glaciers,  
487 perhaps due to simplifications of physics (Hindmarsh, 2009; Stokes and Tarasov, 2010),  
488 poorly constrained patterns of basal sliding parameters (Bueler and Brown, 2009), or  
489 incomplete knowledge of basal sliding (Stearns and van der Veen, 2018). Climate  
490 uncertainties will also influence the ability of an ice sheet model to replicate empirically  
491 derived flow directions, as these impact the overall geometry of the modelled ice sheet. Since  
492 these factors are a large uncertainty in ice sheet modelling (Ritz et al., 2015; Gladstone et al.,  
493 2017), flowset direction is likely to be a robust test of ice sheet models. A question remains  
494 regarding how long flowing ice must occupy an area in order to produce lineated flow sets; if  
495 this time is decadal (e.g. Dowling et al., 2016) rather than centennial, it indicates that flowset

496 matching is not of the highest priority for ice sheet models which typically have a lower  
497 temporal resolution.

498         None of the three model simulations adequately replicated a cross-cutting relationship  
499 between flowsets. Such cross-cuts can be used to decipher the geometry of a palaeo-ice sheet  
500 and how it changes through time (Boulton and Clark, 1990), including factors such as ice-  
501 divide migration and margin position change (e.g. Greenwood and Clark, 2009; Hughes et al.,  
502 2014). This means, in addition to the problems of matching a single flow-set mentioned  
503 above, deglacial climate must be adequately simulated for a cross-cuts caused by climatically  
504 driven ice-divide migration to be matched. In addition to this, the model must also adequately  
505 represent the internal processes which cause ice-divide migration (e.g. flow piracy, ice stream  
506 initiation, saddle collapse). A further uncertainty is introduced by our ignorance of ice stream  
507 dynamics and how ice stream velocity and orientation can change over centennial and even  
508 decadal timescales. Given these potential difficulties at matching cross-cuts, they can be  
509 thought of as an even sterner test of an ice sheet model than the number of flowsets replicated  
510 alone.

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

#### 522 *4.2. An approach to measuring model-data fit*

523         As a consequence of the above complexity in model-data comparison, we suggest the  
524 following pragmatic approach to reconciling empirical reconstructions and model  
525 reconstructions, summarised in Figure 8. Here, the investigator starts with an ensemble of ice  
526 sheet model simulations; the number of simulations considered are progressively diminished  
527 in number by removing those which rank lowest against a particular metric (Figure 8). This

528 builds on the suggestion of Napieralski et al. (2007) who used APCA to rule out the majority  
529 of simulations, then AFDA to further evaluate model performance. Our order of rankings  
530 (Figure 8) is based upon what we ascertain from the above discussion to be progressively  
531 more stringent tests of a model simulation. Indeed, the order of these rankings is likely to  
532 change between users who are interested in specific aspects of a palaeo-ice sheet (e.g. more  
533 weighting may be given to flowset direction if studying ice-flow patterns). An alternative is  
534 to combine scores derived from the model-data comparison techniques for each simulation,  
535 and then rank simulations to either heavily weight the highest scoring simulations when  
536 producing a probabilistic output from an ensemble (e.g. Tarasov et al., 2012), or to rule out  
537 the lowest scoring simulations. In this case, the order that tests are applied is irrelevant.

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

554         After the application of these tests, the original ensemble of simulations will be much  
555 reduced, to a set which are yet to be ruled out (Figure 8). Given that it is unlikely that a  
556 perfect score will be found in these models, model-data mismatch between ‘best-fit’ models  
557 should be further investigated. It may be that certain areas of empirical evidence consistently  
558 produce model-data mismatch, and this may motivate further simulations if spatial or  
559 temporal patterns are clear. For example, a climate driver may under-represent a particular  
560 stadial, thereby producing a simulated timing which is disagrees with data. On the other hand,

561 if all surrounding empirical evidence is met, and a particular data point or subset of data  
562 cannot be replicated by the model, this may warrant re-evaluation of the data in question  
563 (Figure 8). In an analogous manner to climate modelling (Collins et al., 2017), it remains  
564 open as to whether all models which pass a threshold acceptance barrier should be  
565 incorporated into an acceptable set of reconstructions (i.e. a model democracy; Knutti, 2010)  
566 or whether a “best-fit” model which performs best against all constraints should be identified  
567 and used for further research. In either case, the procedure outlined above can help reduce  
568 model uncertainty and produce more robust palaeo-ice sheet reconstructions.

#### 569 *4.3. Suggestions for future developments*

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

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

591 For ice-flow direction comparison, our analysis shows that a key problem is replicating  
592 the synchronous flow directions recorded in some flowsets, and whether the model resolves

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

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

## 613 **5. Summary**

614 Progress toward an integration of empirically-based and numerical model-based  
615 reconstructions of palaeo-ice sheets have proven to be slow since being first suggested  
616 (Andrews, 1982; Stokes et al., 2015). Here, we have outlined a procedure of model-data  
617 comparison designed to score the degree of fit between ice sheet model simulations and  
618 palaeo-ice sheet data, which aims to further integrate these two approaches. We compared  
619 three ice sheet model simulations against the three data constraints of margin position (from  
620 moraines), flow direction (from subglacial bedforms) and timing of ice-free conditions (from  
621 geochronological data). In doing so, we highlighted the complexities of such model-data  
622 comparisons. As ice sheet models are unlikely to reproduce all the information provided at  
623 each constraint, we pragmatically suggest a hierarchical system for scoring ice sheet models,  
624 whereby successive tests are applied to the ice sheet model, progressively ruling out model

625 runs which perform the poorest against each constraint. This procedure could be used to  
626 ascertain best-fit models or used to calibrate models. Future work could consider in more  
627 depth the relative importance of the different data-based constraints. Furthermore, we argue  
628 that this approach could lead to models more frequently being used to test the plausibility of  
629 data-interpretations. In future work, this comparison should ideally be made in conjunction  
630 with other data-based constraints such as RSL data through GIA modelling and  
631 sedimentological observations. In this manner, an integration of empirical and model-based  
632 approaches to palaeo-ice sheet reconstruction can occur. The BIIS is a data rich environment  
633 for conducting such model-data integration.

634

### 635 **Acknowledgements**

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

### 642 **References**

- 643 Andrews, J.T., 1982. On the reconstruction of Pleistocene ice sheets: A review. *Quaternary*  
644 *Science Reviews*, 1(1), pp.1-30.
- 645 Arnold, J.R. and Libby, W.F., 1951. Radiocarbon dates. *Science*, 113(2927), pp.111-120.
- 646 Aschwanden, A., Bueller, E., Khroulev, C. and Blatter, H., 2012. An enthalpy formulation for  
647 glaciers and ice sheets. *Journal of Glaciology*, 58(209), pp.441-457.
- 648 Auriac, A., Whitehouse, P.L., Bentley, M.J., Patton, H., Lloyd, J.M. and Hubbard, A., 2016.  
649 Glacial isostatic adjustment associated with the Barents Sea ice sheet: a modelling  
650 inter-comparison. *Quaternary Science Reviews*, 147, pp.122-135.
- 651 Bateman, M.D., Evans, D.J., Roberts, D.H., Medialdea, A., Ely, J. and Clark, C.D., 2018. The  
652 timing and consequences of the blockage of the Humber Gap by the last British– Irish  
653 Ice Sheet. *Boreas*, 47(1), pp.41-61.
- 654 Bauer, P., Thorpe, A. and Brunet, G., 2015. The quiet revolution of numerical weather  
655 prediction. *Nature*, 525(7567), p.47-55.
- 656 Beckmann, A. and Goosse, H., 2003. A parameterization of ice shelf–ocean interaction for  
657 climate models. *Ocean modelling*, 5(2), pp.157-170.

- 658 Benetti, S., Dunlop, P. and Cofaigh, C.Ó., 2010. Glacial and glacially-related features on the  
659 continental margin of northwest Ireland mapped from marine geophysical data. *Journal*  
660 *of Maps*, 6(1), pp.14-29.
- 661 Boulton, G.S. and Clark, C.D., 1990. A highly mobile Laurentide ice sheet revealed by  
662 satellite images of glacial lineations. *Nature*, 346(6287), p.813.
- 663 Boulton, G. and Hagdorn, M., 2006. Glaciology of the British Isles Ice Sheet during the last  
664 glacial cycle: form, flow, streams and lobes. *Quaternary Science Reviews*, 25(23-24),  
665 pp.3359-3390.
- 666 Bowen, D.Q., Rose, J., McCabe, A.M., Sutherland, D.G., 1986. Correlation of Quaternary  
667 glaciations in England, Ireland, Scotland and Wales. *Quaternary Science Reviews* 5,  
668 pp. 299-340
- 669 Bradley, S.L., Milne, G.A., Shennan, I. and Edwards, R., 2011. An improved glacial isostatic  
670 adjustment model for the British Isles. *Journal of Quaternary Science*, 26(5), pp.541-  
671 552.
- 672 Bradwell, T., Stoker, M.S., Golledge, N.R., Wilson, C.K., Merritt, J.W., Long, D., Everest,  
673 J.D., Hestvik, O.B., Stevenson, A.G., Hubbard, A.L. Finlayson, A.G., and Mathers,  
674 H.E., 2008. The northern sector of the last British Ice Sheet: maximum extent and  
675 demise. *Earth-Science Reviews*, 88(3-4), pp.207-226.
- 676 Braconnot, P., Harrison, S.P., Kageyama, M., Bartlein, P.J., Masson-Delmotte, V., Abe-  
677 Ouchi, A., Otto-Bliesner, B. and Zhao, Y., 2012. Evaluation of climate models using  
678 palaeoclimatic data. *Nature Climate Change*, 2(6), p.417.
- 679 Briggs, R.D. and Tarasov, L., 2013. How to evaluate model-derived deglaciation  
680 chronologies: a case study using Antarctica. *Quaternary Science Reviews*, 63, pp.109-  
681 127.
- 682 Bueler, E.D., Lingle, C.S. and Brown, J., 2007. Fast computation of a viscoelastic deformable  
683 Earth model for ice sheet simulations. *Annals of Glaciology*, 46(1), pp.97-105.
- 684 Bueler, E. and Brown, J., 2009. Shallow shelf approximation as a “sliding law” in a  
685 thermomechanically coupled ice sheet model. *Journal of Geophysical Research: Earth*  
686 *Surface*, 114(F3).
- 687 Bueler, E. and Pelt, W.V., 2015. Mass-conserving subglacial hydrology in the Parallel Ice  
688 Sheet Model version 0.6. *Geoscientific Model Development*, 8(6), pp.1613-1635.
- 689 Calov, R. and Greve, R., 2005. A semi-analytical solution for the positive degree-day model  
690 with stochastic temperature variations. *Journal of Glaciology*, 51(172), pp.173-175.
- 691 Clark, C.D., 1993. Mega-scale glacial lineations and cross-cutting ice-flow landforms. *Earth*  
692 *surface processes and landforms*, 18(1), pp.1-29.
- 693 Clark, C.D., 1999. Glaciodynamic context of subglacial bedform generation and  
694 preservation. *Annals of Glaciology*, 28, pp.23-32.

- 695 Clark, C.D., Hughes, A.L., Greenwood, S.L., Jordan, C. and Sejrup, H.P., 2012. Pattern and  
696 timing of retreat of the last British-Irish Ice Sheet. *Quaternary Science Reviews*, 44,  
697 pp.112-146.
- 698 Clark, C.D., Ely, J.C., Greenwood, S.L., Hughes, A.L., Meehan, R., Barr, I.D., Bateman,  
699 M.D., Bradwell, T., Doole, J., Evans, D.J. Jordan, C.J., Moneys, X., Pellicer, M. and  
700 Sheehy, M., 2018. BRITICE Glacial Map, version 2: a map and GIS database of glacial  
701 landforms of the last British–Irish Ice Sheet. *Boreas*, 47(1), p.11.
- 702 Collins, M., 2017. Still weighting to break the model democracy. *Geophysical Research*  
703 *Letters*, 44(7), pp.3328-3329.
- 704 Dowling, T.P., Möller, P. and Spagnolo, M., 2016. Rapid subglacial streamlined bedform  
705 formation at a calving bay margin. *Journal of Quaternary Science*, 31(8), pp.879-892.
- 706 Duller, G.A.T., 2006. Single grain optical dating of glacial deposits. *Quaternary*  
707 *Geochronology*, 1(4), pp.296-304.
- 708 Dyke, A.S., 2004. An outline of North American deglaciation with emphasis on central and  
709 northern Canada. *Developments in Quaternary Sciences*, 2, pp.373-424
- 710 Ely, J.C., Clark, C.D., Spagnolo, M., Stokes, C.R., Greenwood, S.L., Hughes, A.L., Dunlop,  
711 P. and Hess, D., 2016. Do subglacial bedforms comprise a size and shape  
712 continuum?. *Geomorphology*, 257, pp.108-119.
- 713 Ely, J.C., Clark, C.D., Spagnolo, M., Hughes, A.L. and Stokes, C.R., 2018. Using the size  
714 and position of drumlins to understand how they grow, interact and evolve. *Earth*  
715 *Surface Processes and Landforms*, 43(5), pp.1073-1087.
- 716 Ely, J.C., Clark, C.D., Small, D., Hindmarsh, R.C.A, in press. ATAT 1.1, an Automated  
717 Timing Accordance Tool for comparing ice sheet model output with geochronological  
718 data. *Geoscientific Model Development Discussions*.
- 719 Engelhart, S.E. and Horton, B.P., 2012. Holocene sea level database for the Atlantic coast of  
720 the United States. *Quaternary Science Reviews*, 54, pp.12-25.
- 721 Eyles, N. and McCabe, A.M., 1989. The Late Devensian (< 22,000 BP) Irish Sea Basin: the  
722 sedimentary record of a collapsed ice sheet margin. *Quaternary Science Reviews*, 8(4),  
723 pp.307-351.
- 724 Fabel, D., Ballantyne, C.K. and Xu, S., 2012. Trimlines, blockfields, mountain-top erratics  
725 and the vertical dimensions of the last British–Irish Ice Sheet in NW  
726 Scotland. *Quaternary Science Reviews*, 55, pp.91-102.
- 727 Fisher, D., Reeh, N. and Langley, K., 1985. Objective reconstructions of the late Wisconsinan  
728 Laurentide Ice Sheet and the significance of deformable beds. *Géographie physique et*  
729 *Quaternaire*, 39(3), pp.229-238.
- 730 Gandy, N., Gregoire, L.J., Ely, J.C., Clark, C.D., Hodgson, D.M., Lee, V., Bradwell, T. and  
731 Ivanovic, R.F., 2018. Marine ice sheet instability and ice shelf buttressing of the Minch  
732 Ice Stream, northwest Scotland. *The Cryosphere*, 12(11), pp.3635-3651.



- 733 Gladstone, R.M., Warner, R.C., Galton-Fenzi, B.K., Gagliardini, O., Zwinger, T. and Greve,  
734 R., 2017. Marine ice sheet model performance depends on basal sliding physics and  
735 sub-shelf melting. *The Cryosphere*, 11(1), p.319.
- 736 Glen, J.W., 1952. Experiments on the deformation of ice. *Journal of Glaciology*, 2(12),  
737 pp.111-114.
- 738 Greenwood, S.L. and Clark, C.D., 2009. Reconstructing the last Irish Ice Sheet 2: a  
739 geomorphologically-driven model of ice sheet growth, retreat and  
740 dynamics. *Quaternary Science Reviews*, 28(27-28), pp.3101-3123.
- 741 Gregoire, L.J., Payne, A.J. and Valdes, P.J., 2012. Deglacial rapid sea level rises caused by  
742 ice sheet saddle collapses. *Nature*, 487(7406), p.219.
- 743 Gregoire, L.J., Otto-Bliesner, B., Valdes, P.J. and Ivanovic, R., 2016. Abrupt Bølling  
744 warming and ice saddle collapse contributions to the Meltwater Pulse 1a rapid sea level  
745 rise. *Geophysical research letters*, 43(17), pp.9130-9137.
- 746 Herbertson, A.J. (1908). *Outlines of Physiography: An Introduction to the Study of the Earth*.  
747 Edward Arnold, London. 3<sup>rd</sup> ed. Page 118.
- 748 Hindmarsh, R.C., 2009. Consistent generation of ice-streams via thermo-viscous instabilities  
749 modulated by membrane stresses. *Geophysical Research Letters*, 36(6).
- 750 Hubbard, A., Bradwell, T., Golledge, N., Hall, A., Patton, H., Sugden, D., Cooper, R. and  
751 Stoker, M., 2009. Dynamic cycles, ice streams and their impact on the extent,  
752 chronology and deglaciation of the British–Irish ice sheet. *Quaternary Science*  
753 *Reviews*, 28(7), pp.758-776.
- 754 Hughes, A.L., Greenwood, S.L. and Clark, C.D., 2011. Dating constraints on the last British-  
755 Irish Ice Sheet: a map and database. *Journal of Maps*, 7(1), pp.156-184.
- 756 Hughes, A.L., Clark, C.D. and Jordan, C.J., 2014. Flow-pattern evolution of the last British  
757 Ice Sheet. *Quaternary Science Reviews*, 89, pp.148-168.
- 758 Hughes, A.L., Gyllencreutz, R., Lohne, Ø.S., Mangerud, J. and Svendsen, J.I., 2016. The last  
759 Eurasian ice sheets—a chronological database and time-slice reconstruction, DATED-  
760 1. *Boreas*, 45(1), pp.1-45.
- 761 Hughes, T., 1973. Is the West Antarctic ice sheet disintegrating?. *Journal of Geophysical*  
762 *Research*, 78(33), pp.7884-7910.
- 763 Huybrechts, P., 2002. Sea-level changes at the LGM from ice-dynamic reconstructions of the  
764 Greenland and Antarctic ice sheets during the glacial cycles. *Quaternary Science*  
765 *Reviews*, 21(1-3), pp.203-231.
- 766 Imbrie, J., Hays, J.D., Martinson, D.G., McIntyre, A., Mix, A.C., Morley, J.J., Pisias, N.G.,  
767 Prell, W.L. and Shackleton, N.J., 1984. The orbital theory of Pleistocene climate:  
768 support from a revised chronology of the marine  $\delta^{18}\text{O}$  record. In A. L. Berger, J.  
769 Imbrie, J. Hays, G. Kukla, & B. Saltzman (Eds.), *Milankovitch and climate* (pp. 269–  
770 305). Dordrecht, Holland: Milankovitch and Climate.

- 771 Kingslake, J., Scherer, R.P., Albrecht, T., Coenen, J., Powell, R.D., Reese, R., Stansell, N.D.,  
772 Tulaczyk, S., Wearing, M.G. and Whitehouse, P.L., 2018. Extensive retreat and re-  
773 advance of the West Antarctic ice sheet during the Holocene. *Nature*, 558(7710), p.430.
- 774 Kleman, J., 1990. On the use of glacial striae for reconstruction of paleo-ice sheet flow  
775 patterns: with application to the Scandinavian ice sheet. *Geografiska Annaler: Series A,*  
776 *Physical Geography*, 72(3-4), pp.217-236.
- 777 Kleman, J. and Borgström, I., 1996. Reconstruction of palaeo-ice sheets: the use of  
778 geomorphological data. *Earth surface processes and landforms*, 21(10), pp.893-909.
- 779 Kleman, J., Hättestrand, C., Borgström, I. and Stroeven, A., 1997. Fennoscandian  
780 palaeoglaciology reconstructed using a glacial geological inversion model. *Journal of*  
781 *glaciology*, 43(144), pp.283-299.
- 782 Knutti, R., 2010. The end of model democracy? *Climatic Change*, 102, 395-404.
- 783 Kuchar, J., Milne, G., Hubbard, A., Patton, H., Bradley, S., Shennan, I. and Edwards, R.,  
784 2012. Evaluation of a numerical model of the British–Irish ice sheet using relative sea-  
785 level data: implications for the interpretation of trimline observations. *Journal of*  
786 *Quaternary Science*, 27(6), pp.597-605.
- 787 Lambeck, K. and Chappell, J., 2001. Sea level change through the last glacial  
788 cycle. *Science*, 292(5517), pp.679-686.
- 789 Levermann, A., Albrecht, T., Winkelmann, R., Martin, M.A., Haseloff, M. and Joughin, I.,  
790 2012. Kinematic first-order calving law implies potential for abrupt ice-shelf  
791 retreat. *The Cryosphere*, 6, pp.273-286.
- 792 Li, Y., Napieralski, J., Harbor, J. and Hubbard, A., 2007. Identifying patterns of  
793 correspondence between modeled flow directions and field evidence: an automated  
794 flow direction analysis. *Computers & geosciences*, 33(2), pp.141-150.
- 795 Li, Y., Napieralski, J. and Harbor, J., 2008. A revised automated proximity and conformity  
796 analysis method to compare predicted and observed spatial boundaries of geologic  
797 phenomena. *Computers & Geosciences*, 34(12), pp.1806-1814.
- 798 Libby, W.F., Anderson, E.C. and Arnold, J.R., 1949. Age determination by radiocarbon  
799 content: world-wide assay of natural radiocarbon. *Science*, 109(2827), pp.227-228.
- 800 Lorenz, E.N., 1963. Deterministic nonperiodic flow. *Journal of the Atmospheric*  
801 *Sciences*, 20(2), pp.130-141.
- 802 Margold, M., Stokes, C.R. and Clark, C.D., 2015. Ice streams in the Laurentide Ice Sheet:  
803 Identification, characteristics and comparison to modern ice sheets. *Earth-Science*  
804 *Reviews*, 143, pp.117-146.
- 805 Martin, M.A., Winkelmann, R., Haseloff, M., Albrecht, T., Bueller, E., Khroulev, C. and  
806 Levermann, A., 2011. The Potsdam Parallel Ice Sheet Model (PISM-PIK)—Part 2:  
807 Dynamic equilibrium simulation of the Antarctic ice sheet. *The Cryosphere*, 5(3),  
808 pp.727-740.

- 809 McCabe, A.M., Clark, P.U., Clark, J. and Dunlop, P., 2007. Radiocarbon constraints on  
810 readvances of the British–Irish Ice Sheet in the northern Irish Sea Basin during the last  
811 deglaciation. *Quaternary Science Reviews*, 26(9-10), pp.1204-1211
- 812 Napieralski, J., Li, Y. and Harbor, J., 2006. Comparing predicted and observed spatial  
813 boundaries of geologic phenomena: Automated Proximity and Conformity Analysis  
814 applied to ice sheet reconstructions. *Computers & geosciences*, 32(1), pp.124-134.
- 815 Napieralski, J., Hubbard, A., Li, Y., Harbor, J., Stroeven, A.P., Kleman, J., Alm, G. and  
816 Jansson, K.N., 2007. Towards a GIS assessment of numerical ice sheet model  
817 performance using geomorphological data. *Journal of Glaciology*, 53(180), pp.71-83.
- 818 Nye, J.F., 1953. The flow law of ice from measurements in glacier tunnels, laboratory  
819 experiments and the Jungfraufirn borehole experiment. *Proc. R. Soc. Lond.*  
820 *A*, 219(1139), pp.477-489.
- 821 Ó Cofaigh, C. and Evans, D.J., 2007. Radiocarbon constraints on the age of the maximum  
822 advance of the British–Irish Ice Sheet in the Celtic Sea. *Quaternary Science*  
823 *Reviews*, 26(9-10), pp.1197-1203.
- 824 Patton, H., Hubbard, A., Andreassen, K., Winsborrow, M. and Stroeven, A.P., 2016. The  
825 build-up, configuration, and dynamical sensitivity of the Eurasian ice sheet complex to  
826 Late Weichselian climatic and oceanic forcing. *Quaternary Science Reviews*, 153,  
827 pp.97-121.
- 828 Patton, H., Hubbard, A., Andreassen, K., Auriac, A., Whitehouse, P.L., Stroeven, A.P.,  
829 Shackleton, C., Winsborrow, M., Heyman, J. and Hall, A.M., 2017. Deglaciation of the  
830 Eurasian ice sheet complex. *Quaternary Science Reviews*, 169, pp.148-172.
- 831 Peltier, W.R., 2004. Global glacial isostasy and the surface of the ice-age Earth: the ICE-5G  
832 (VM2) model and GRACE. *Annu. Rev. Earth Planet. Sci.*, 32, pp.111-149.
- 833 Peltier, W.R., Farrell, W.E. and Clark, J.A., 1978. Glacial isostasy and relative sea level: a  
834 global finite element model. *Tectonophysics*, 50(2-3), pp.81-110.
- 835 Piotrowski, J.A. and Tulaczyk, S., 1999. Subglacial conditions under the last ice sheet in  
836 northwest Germany: ice-bed separation and enhanced basal sliding?. *Quaternary*  
837 *Science Reviews*, 18(6), pp.737-751.
- 838 Quinlan, G. and Beaumont, C., 1982. The deglaciation of Atlantic Canada as reconstructed  
839 from the postglacial relative sea-level record. *Canadian Journal of Earth*  
840 *Sciences*, 19(12), pp.2232-2246.
- 841 Retzlaff, R. and Bentley, C.R., 1993. Timing of stagnation of Ice Stream C, West Antarctica,  
842 from short-pulse radar studies of buried surface crevasses. *Journal of*  
843 *Glaciology*, 39(133), pp.553-561.
- 844 Ritz, C., Edwards, T.L., Durand, G., Payne, A.J., Peyaud, V. and Hindmarsh, R.C., 2015.  
845 Potential sea-level rise from Antarctic ice sheet instability constrained by  
846 observations. *Nature*, 528(7580), pp.115-118.

- 847 Schoof, C., 2007. Ice sheet grounding line dynamics: Steady states, stability, and  
848 hysteresis. *Journal of Geophysical Research: Earth Surface*, 112(F3).
- 849 Schoof, C., 2012. Marine ice sheet stability. *Journal of Fluid Mechanics*, 698, pp.62-72.
- 850 Seierstad, I.K., Abbott, P.M., Bigler, M., Blunier, T., Bourne, A.J., Brook, E., Buchardt, S.L.,  
851 Buizert, C., Clausen, H.B., Cook, E. Dahl-Jensen, D., Davies, S.M., Guillevic, M.,  
852 Johnsen, S.J., Pedersen, D.S., Popp, T.J., Rasmussen, S.O., Severinghaus, J.P.,  
853 Svensson, A. and Vinther, B.M., 2014. Consistently dated records from the Greenland  
854 GRIP, GISP2 and NGRIP ice cores for the past 104 ka reveal regional millennial-scale  
855  $\delta^{18}\text{O}$  gradients with possible Heinrich event imprint. *Quaternary Science*  
856 *Reviews*, 106, pp.29-46.
- 857 Seguinot, J., Khroulev, C., Rogozhina, I., Stroeven, A.P. and Zhang, Q., 2014. The effect of  
858 climate forcing on numerical simulations of the Cordilleran ice sheet at the Last Glacial  
859 Maximum. *The Cryosphere*, 8(3), pp.1087-1103.
- 860 Seguinot, J., Rogozhina, I., Stroeven, A.P., Margold, M., Kleman, J., 2016. Numerical  
861 simulations of the Cordilleran ice sheet through the last glacial cycle. *The Cryosphere*  
862 10, pp. 639-664.
- 863 Seguinot, J., Ivy-Ochs, S., Jouviet, G., Huss, M., Funk, M. and Preusser, F., 2018. Modelling  
864 last glacial cycle ice dynamics in the Alps. *The Cryosphere*, 12(10), pp.3265-3285.
- 865 Simpson, M.J., Milne, G.A., Huybrechts, P. and Long, A.J., 2009. Calibrating a glaciological  
866 model of the Greenland ice sheet from the Last Glacial Maximum to present-day using  
867 field observations of relative sea level and ice extent. *Quaternary Science*  
868 *Reviews*, 28(17-18), pp.1631-1657.
- 869 Small, D., Clark, C.D., Chiverrell, R.C., Smedley, R.K., Bateman, M.D., Duller, G.A., Ely,  
870 J.C., Fabel, D., Medialdea, A. and Moreton, S.G., 2017a. Devising quality assurance  
871 procedures for assessment of legacy geochronological data relating to deglaciation of  
872 the last British-Irish Ice Sheet. *Earth-science reviews*, 164, pp.232-250.
- 873 Small, D., Benetti, S., Dove, D., Ballantyne, C.K., Fabel, D., Clark, C.D., Gheorghiu, D.M.,  
874 Newall, J. and Xu, S., 2017b. Cosmogenic exposure age constraints on deglaciation and  
875 flow behaviour of a marine-based ice stream in western Scotland, 21–16  
876 ka. *Quaternary Science Reviews*, 167, pp.30-46.
- 877 Smedley, R.K., Scourse, J.D., Small, D., Hiemstra, J.F., Duller, G.A.T., Bateman, M.D.,  
878 Burke, M.J., Chiverrell, R.C., Clark, C.D., Davies, S.M. Fabel, D., Gheorghiu, D.M.,  
879 McCarroll, D., Medialdea, A., and Xu S., 2017. New age constraints for the limit of the  
880 British–Irish Ice Sheet on the Isles of Scilly. *Journal of Quaternary Science*, 32(1),  
881 pp.48-62
- 882 Stearns, L.A. and van der Veen, C.J., 2018. Friction at the bed does not control fast glacier  
883 flow. *Science*, p.eaat2217.
- 884 Stokes, C.R. and Clark, C.D., 1999. Geomorphological criteria for identifying Pleistocene ice  
885 streams. *Annals of glaciology*, 28, pp.67-74.

- 886 Stokes, C.R. and Tarasov, L., 2010. Ice streaming in the Laurentide Ice Sheet: A first  
887 comparison between data-calibrated numerical model output and geological  
888 evidence. *Geophysical Research Letters*, 37(L01501).
- 889 Stokes, C.R., Clark, C.D. and Storrar, R., 2009. Major changes in ice stream dynamics during  
890 deglaciation of the north-western margin of the Laurentide Ice Sheet. *Quaternary*  
891 *Science Reviews*, 28(7-8), pp.721-738.
- 892 Stokes, C.R., Tarasov, L., Blomdin, R., Cronin, T.M., Fisher, T.G., Gyllencreutz, R.,  
893 Hättestrand, C., Heyman, J., Hindmarsh, R.C., Hughes, A.L., Jakobsson, M., Kirchner,  
894 N., Livingstone, S.J., Margold, M., Murton, J.B., Noormets, R., Peltier, R.W., Peteet,  
895 D.M., Piper, D.J.W., Preusser, F., Renssen, H., Roberts, D.H., Roche, D.M., Saint-  
896 Ange, F., Stroeven, A.P. and Teller, J.T., 2015. On the reconstruction of palaeo-ice  
897 sheets: recent advances and future challenges. *Quaternary Science Reviews*, 125, pp.15-  
898 49.
- 899 Stone, J.O., Balco, G.A., Sugden, D.E., Caffee, M.W., Sass, L.C., Cowderly, S.G. and  
900 Siddoway, C., 2003. Holocene deglaciation of Marie Byrd land, west  
901 Antarctica. *Science*, 299(5603), pp.99-102.
- 902 Stroeven, A.P., Hättestrand, C., Kleman, J., Heyman, J., Fabel, D., Fredin, O., Goodfellow,  
903 B.W., Harbor, J.M., Jansen, J.D., Olsen, L., Caffee, M.W., Fink, D., Lundqvist, J.,  
904 Rosqvist, G.C., Strömberg, B. and Jansson, K.N., 2016. Deglaciation of  
905 Fennoscandia. *Quaternary Science Reviews*, 147, pp.91-121.
- 906 Tarasov, L. and Peltier, W.R., 2004. A geophysically constrained large ensemble analysis of  
907 the deglacial history of the North American ice sheet complex. *Quaternary Science*  
908 *Reviews*, 23(3-4), pp.359-388.
- 909 Tarasov, L., Dyke, A.S., Neal, R.M. and Peltier, W.R., 2012. A data-calibrated distribution of  
910 deglacial chronologies for the North American ice complex from glaciological  
911 modeling. *Earth and Planetary Science Letters*, 315, pp.30-40.
- 912 Tulaczyk, S., Kamb, W.B. and Engelhardt, H.F., 2000. Basal mechanics of ice stream B,  
913 West Antarctica: 1. Till mechanics. *Journal of Geophysical Research: Solid*  
914 *Earth*, 105(B1), pp.463-481.
- 915 Walcott, R.I., 1972. Late Quaternary vertical movements in eastern North America:  
916 Quantitative evidence of glacio- isostatic rebound. *Reviews of Geophysics*, 10(4),  
917 pp.849-884.
- 918 Weatherall, P., Marks, K.M., Jakobsson, M., Schmitt, T., Tani, S., Arndt, J.E., Rovere, M.,  
919 Chayes, D., Ferrini, V. and Wigley, R., 2015. A new digital bathymetric model of the  
920 world's oceans. *Earth and Space Science*, 2(8), pp.331-345.
- 921 Winkelmann, R., Martin, M.A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C. and  
922 Levermann, A., 2011. The Potsdam parallel ice sheet model (PISM-PIK)-Part 1: Model  
923 description. *The Cryosphere*, 5(3), p.715.

924 **Tables:**

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

<b>Glaciological characteristic</b>	<b>Model representation</b>	<b>Empirical data basis</b>	<b>BIIS data used in this study</b>	<b>Data-model comparison tool</b>
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, <sup>14</sup> 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).

926

927

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

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

930

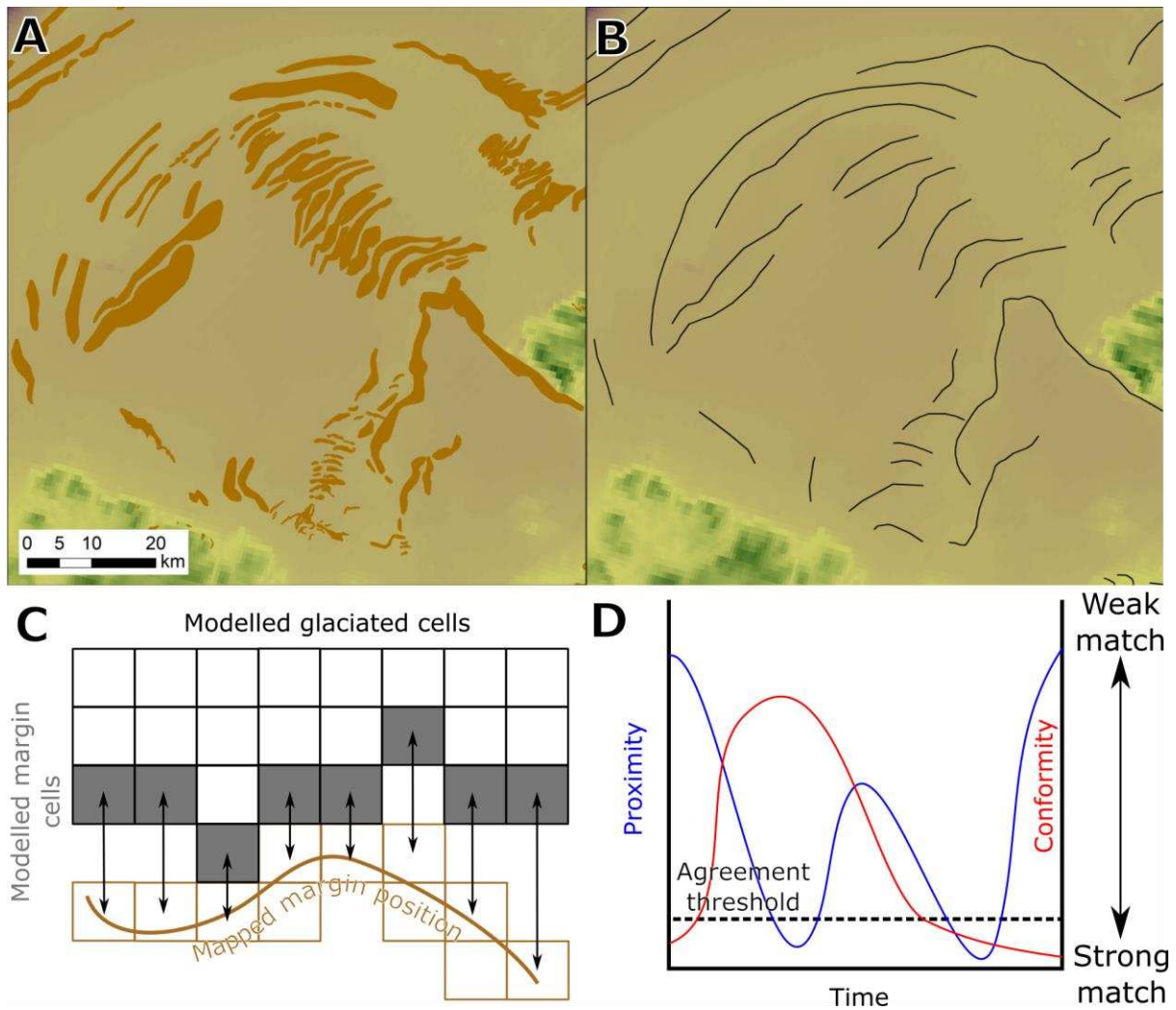
931

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

	Simulation																		
A	60	% of margins matched (n = 189)	76 (n = 151)	% of margins matched within maximum modelled extent	61	% of margins matched within extent of simulation C	9	% of flowsets matched (n = 103)	21 (n = 41)	% of flowsets matched within maximum modelled extent	26	% of flowsets matched within extent of simulation C	0	% of cross-cuts matched	41	% of dates where model-data agreement occurs (n = 108)	1898	wRMSE of model-data difference for ice covered dates where model-data agreement occurs (years)	
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			

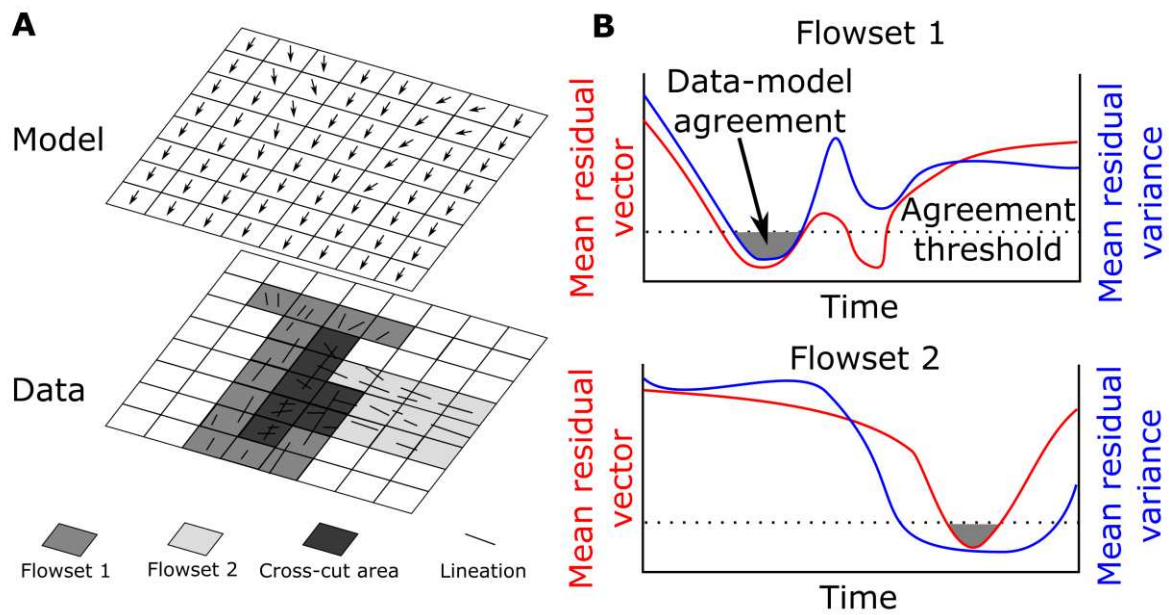
934  
 935





937

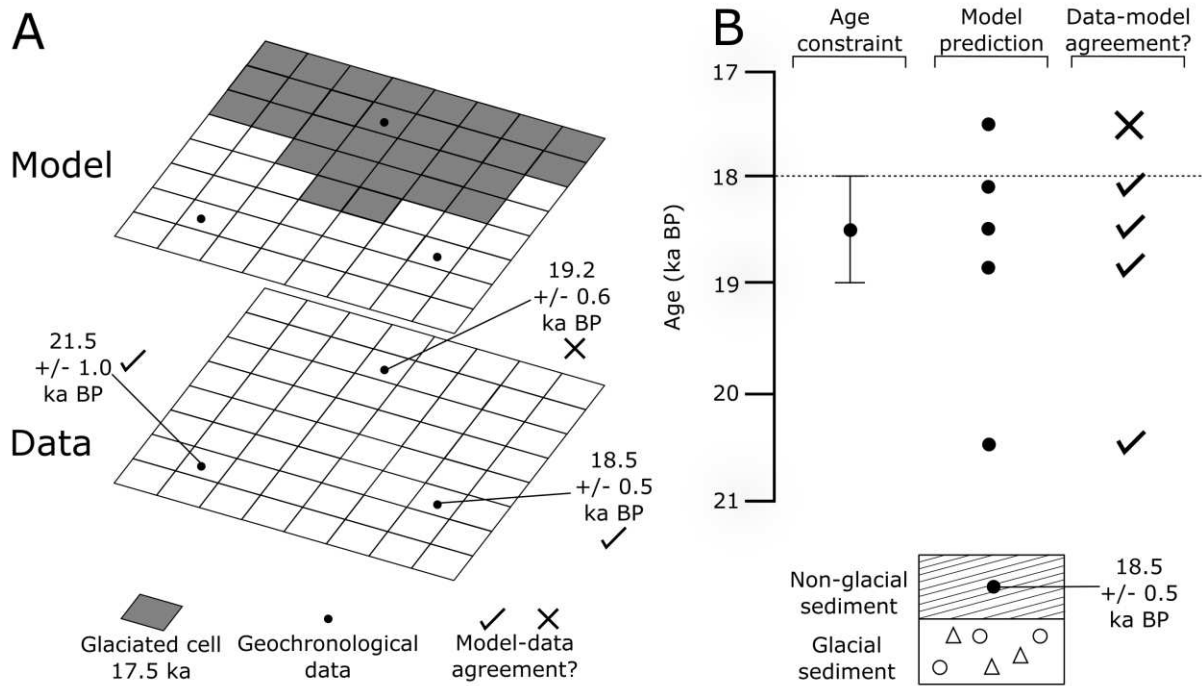
938 Figure 1. A) Mapped offshore moraines, Donegal Bay, Ireland, from Benetti et al. (2010). B)  
 939 Interpreted margin positions from A. C) Schematic representation of the Automated  
 940 Proximity and Conformity Analysis (APCA), whereby the distance between modelled and  
 941 mapped margin position is measured. Proximity is defined as the mean of these  
 942 measurements and conformity as the standard deviation (Napieralski et al., 2006; Li et al.,  
 943 2008). D) Schematic output from APCA. Here, a model-data agreement is only declared  
 944 when both proximity and conformity are below a defined threshold. Such instances are  
 945 shaded in grey.



946

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

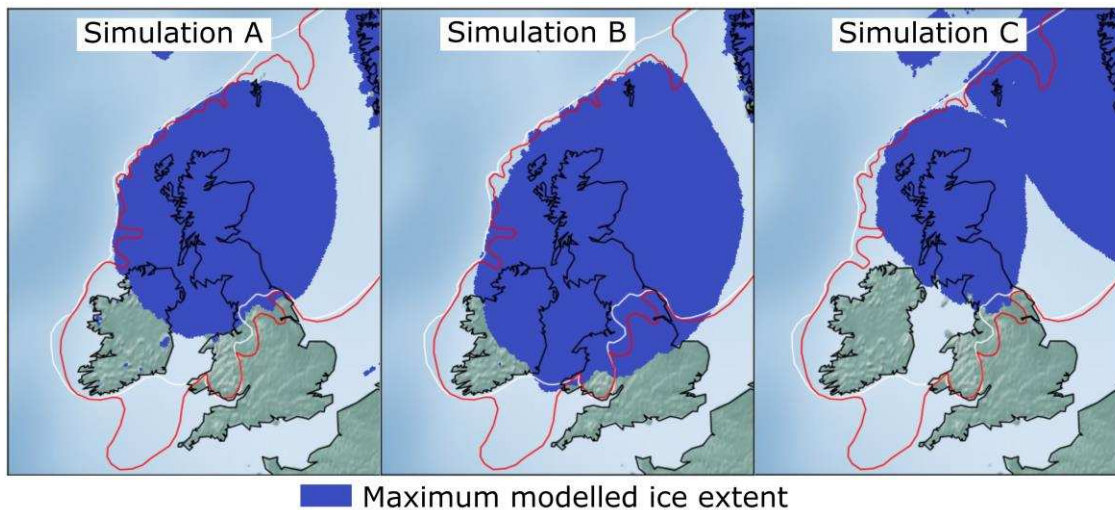
955



957

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

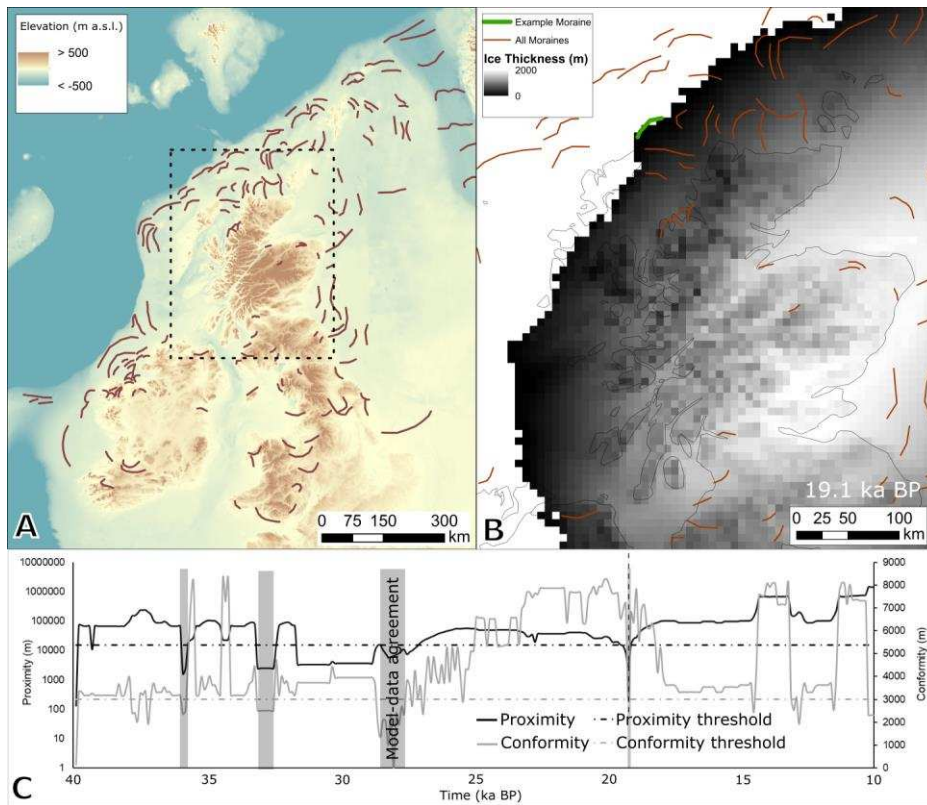
966



967

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

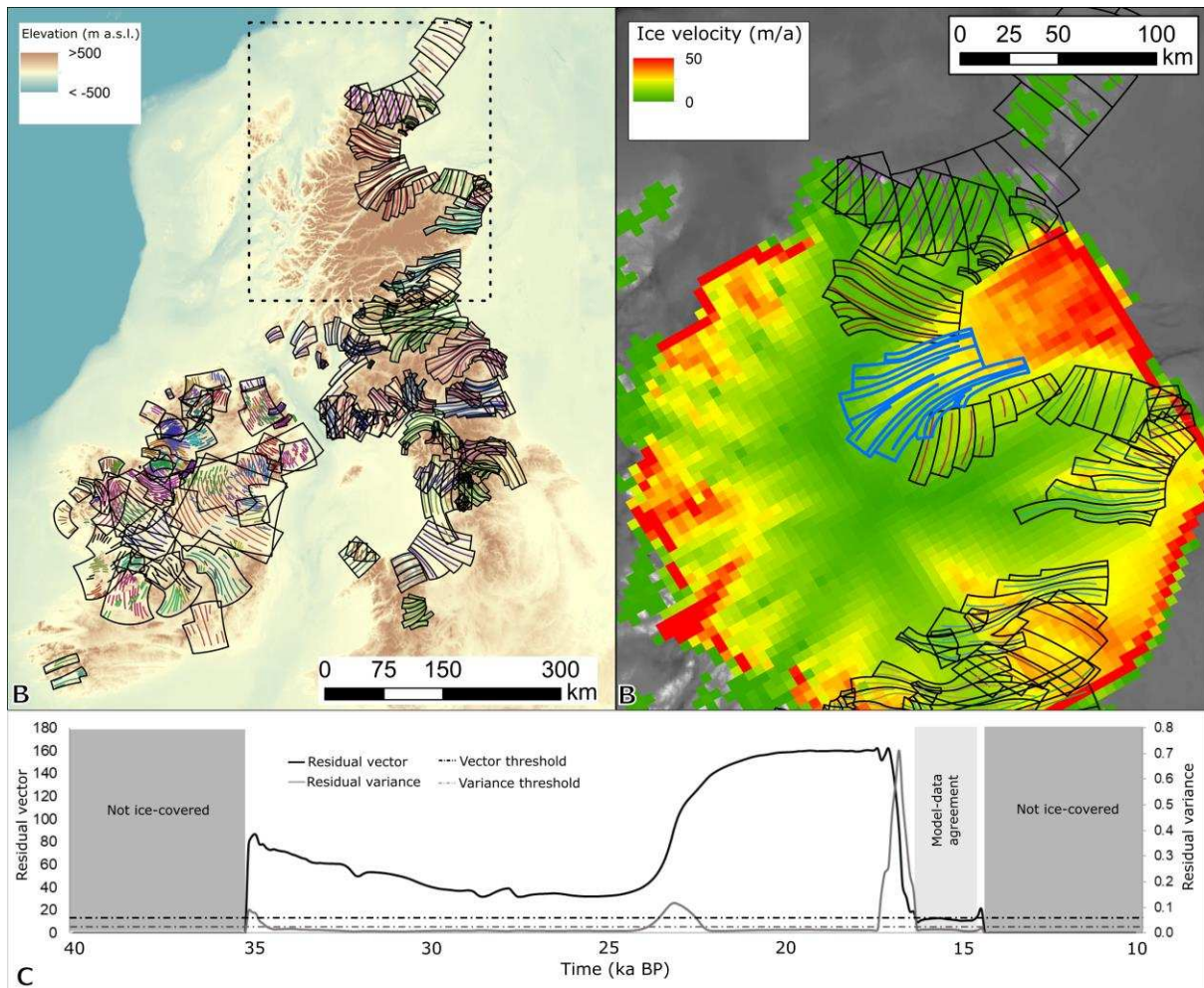
973



974

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

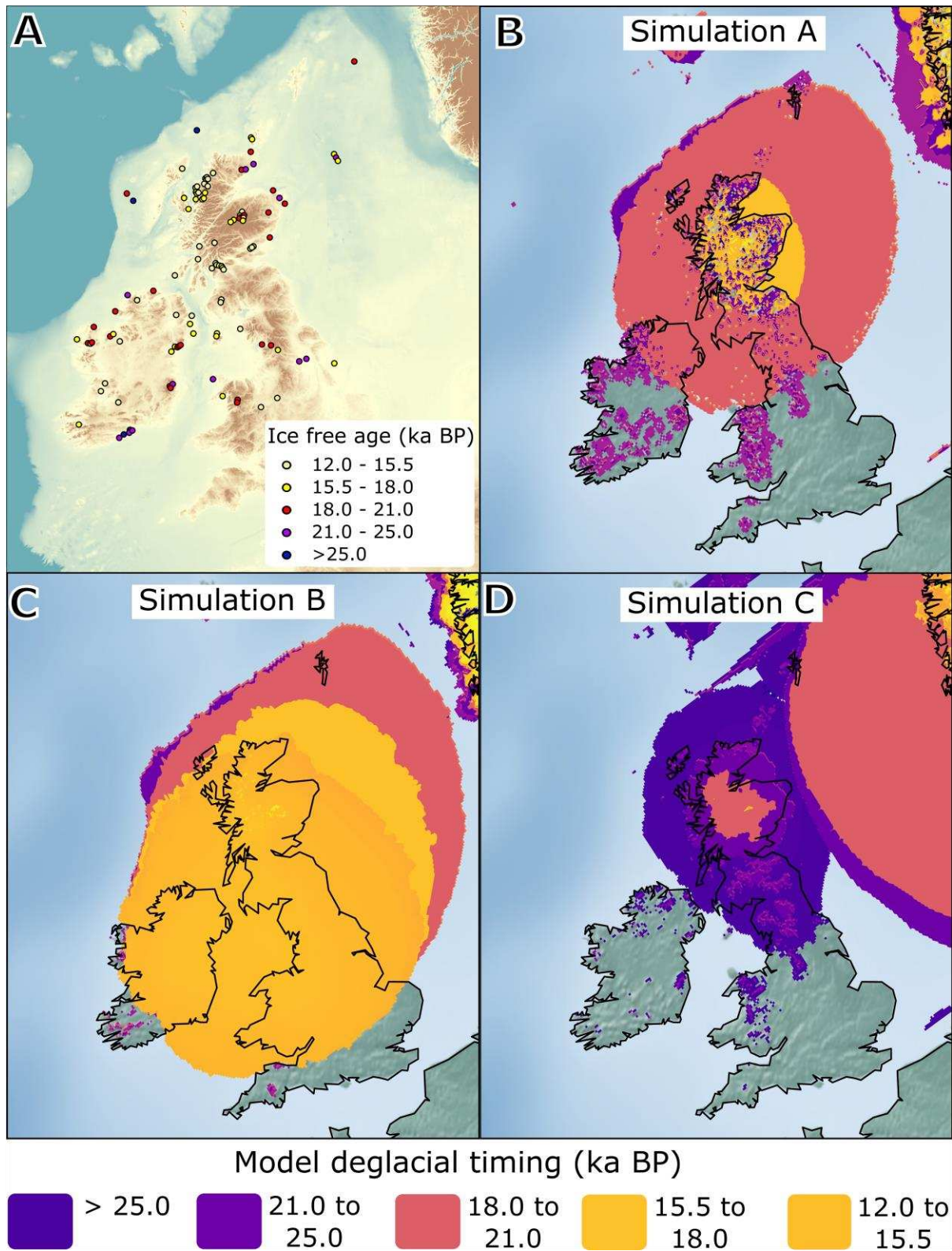
985



986

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

995



996

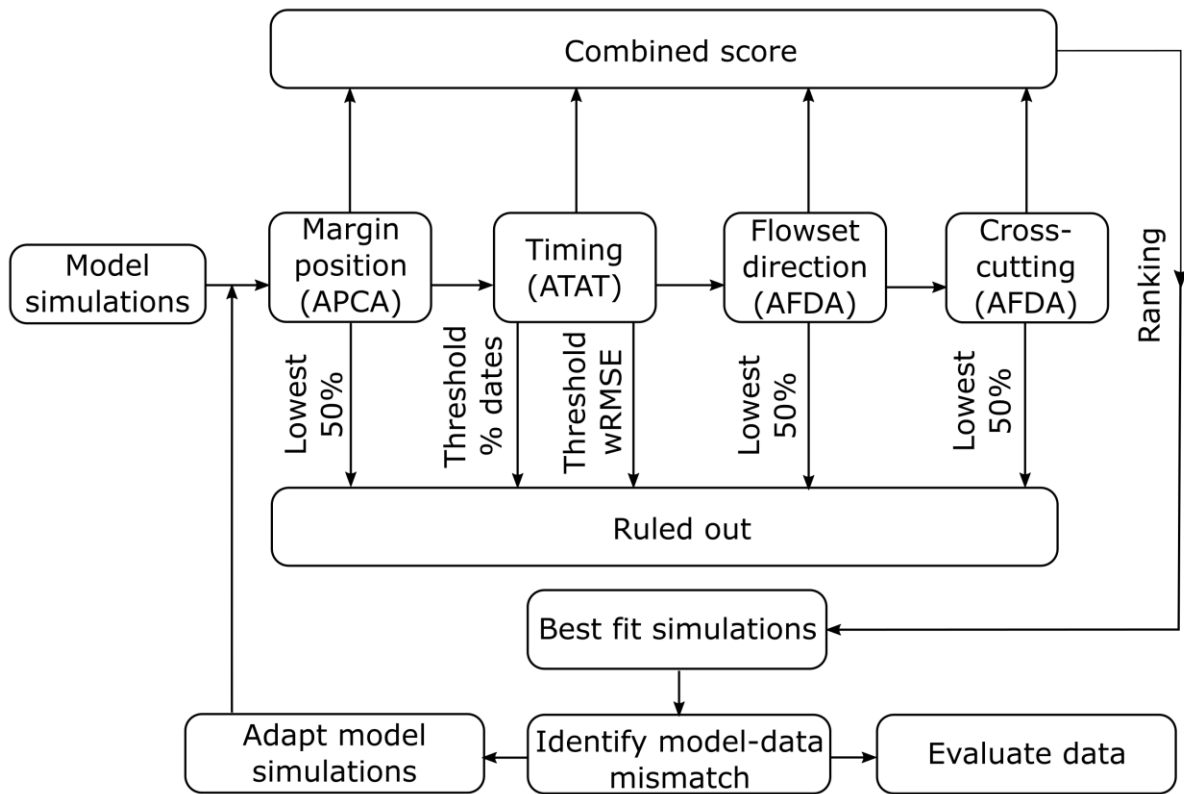
997

998

999

1000

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.



1001

1002 Figure 8. Proposed procedure for comparing multiple model-runs to geochronological data.