

Responses to referee comments on “Reconstruction of North American drainage basins and river discharge since the Last Glacial Maximum” by A. D. Wickert

I would like to thank both reviewers for their thorough comments. Having a geologist and a geophysicist/ice-sheet modeler review this work covers much of the breadth of its projected readership, and both of your thorough questions and comments have lead to an improved revised version of the manuscript.

– Andy Wickert

Responses to Reviewer 1 (Breckenridge)

General Comments

This is an outstanding first attempt to model North American late glacial freshwater discharge. I appreciate the side by side comparison that results from different ice sheet models, and a synthesis of the extensive literature concerned with estimating meltwater routing and discharge from North America. Licciardi et al (1999) is the oft-cited study to which this effort closely compares – but Licciardi et al is based on data interpretations versus the models used here. It could be useful to compare the Licciardi data to the output here and discuss the differences and similarities directly (perhaps in Fig. 15)?

The Licciardi et al. (1999) estimates are now included in Fig. 15. You are right – it is instructive to have these side-by-side. I also added reference to Marshall and Clarke (1999) to invoke another former integrated ice-sheet—meltwater runoff model.

I have also addressed this comparison, and some general features of the calculations featured here, in a new paragraph as follows:

“A comparison between these model results and data-driven drainage histories of \cite{Licciardi1999} (Figure \ref{fig:Q_all_rivers}: Mississippi, Hudson, Saint Lawrence, Hudson Strait) includes both similarities and differences when compared to the river system reconstructions presented here. For the Mississippi River System, \cite{Licciardi1999} expect enhanced discharge between 18.4 and 16.3 ka and prior to 19.9 ka -- that is, at times when all Great Lakes flow was routed down the Mississippi River. This is not borne out in the models analyzed here, which show enhanced discharge but not to the degree predicted by \cite{Licciardi1999}. This likewise is not seen in the $\delta^{18}O$ record of the Gulf of Mexico, though this is more sensitive to the very negative $\delta^{18}O$ of meltwater discharge than to the overall freshwater discharge. The modeled Hudson River discharge does not exhibit the predicted Erie or Mackinaw Interstade periods of enhanced discharge. The latter is greatly diminished in the chronology presented here, but the former includes compelling geological evidence that meltwater flowed through the Mohawk Valley into the Hudson River \cite{Ridge1997}, making this lack of an early Hudson River discharge in these models troubling. The Saint Lawrence River constraints from \cite{Licciardi1999} remain consistent in this work with the flow-routed G12 model \cite{Gregoire2012} and ICE-6G \cite{Argus2014, Peltier2015} providing good fits to these constraints. The Hudson Strait Record is likewise largely consistent, with a difference in the timing of the major meltwater release due to the reconstruction presented here having a faster collapse following the 8.2 ka event \cite{Barber1999} than did \cite{Licciardi1999}.”

The synthesis of the freshwater routing provided by this study is excellent, but there is a lot of uncertainty in the chronology of meltwater routing— especially with respect to dating. Agassiz routing is particularly contentious – but it’s also the most intensively studied. Part of this challenge is attributed to the size of the basin, but I also think the more lines of evidence that are uncovered, the greater the potential for uncertainty (at least at first).

Yes: with one data point, the result seems clear. With maybe 10, not so obvious; with maybe 100, trends appear and we are able to piece together the answer. So hopefully we can move towards having enough data to put together better reconstructions.

I added this sentence prefacing the data setting, “Current geologic evidence, while incomplete and subject to revision, provides a good picture of the broad changes in drainage across North America since the Last Glacial Maximum.”

Another challenge I see when comparing model output and data is that the geologic record is often highlighted by short term (<100 yrs) flood events (e.g. Kankakee Torrent, Champlain freshening event, Murton flood, 9.3 event, etc) – and this modeled output has to average discharge over longer periods of time. Some of the flood events must have resulted from short term release of meltwater stored in lakes, which this modeling attempt cannot resolve (perhaps in the future?) But I also wonder if melt rates may have changed suddenly at times due to climate (cf. pg. 11, 315-323). The varve records from Lake Agassiz, Hitchcock, Superior, and Ojibway clearly illustrate that varve thickness (which must correlate with discharge) could rapidly fluctuate on decadal scales – and the Hitchcock records shows this was at times the result of climate change. I would think that combined, these two factors (storage of meltwater in lakes, and climate change on 10-100 yr scales), may have had dramatic effects overall freshwater flux.

The model time-steps are 100 to ~1000 years, meaning that the model results will average over these events. In my experience, those events that create a strong geomorphic record are related to large past events (e.g., flooding) but do not persist over enough time to be important on the scale of total meltwater release. The upside is that the increase in application of numerical modeling means that time steps will become shorter in the future, with the possibility to investigate short-term patterns of change... though my prediction is that for the time being, we will be left with evaluating these events on a case-by-case basis.

I have added a sentence to a new paragraph on time steps that reads, “This means that the model outputs will generally average over short-term geologic events that are responsible for creating some of the major spillways and deposits \citep[e.g.,][Kehew1994, Breckenridge2007, Murton2010}.”

And this one to a new paragraph on numerical modeling, “... the potential to produce outputs at a high enough temporal resolution to record short-term events with that left a strong mark in the geologic record \citep[e.g.,][Kehew1994, Breckenridge2007, Murton2010}”

It'd love to see a future attempt to model d18 O output for the modeled discharge. Can these ice sheet models be combined with a dynamic oxygen isotope model for the LIS - as well as a model for d18O precipitation? I wonder how the d18 O composition of meltwater changed through time – prior attempts to estimate glacial meltwater discharge assume a static d18 O composition for meltwater (or range of values). We see decreasing d18 O water values for Lake Agassiz and Minong, and these have been interpreted as evidence for increasing meltwater discharge – but is this valid? Perhaps meltwater was becoming progressively enriched in O-16 as the ice sheet decayed? Long term, the best records to reconstruct freshwater discharge continue to be d18 O records (both from the interior, and offshore), and this seems to be an important unknown to interpreting these records. Circulation dynamics and the influence of climate on the overflow of water from the large pro-glacial lakes could be another important complicating factor.

Well, you've anticipated the general track! I wrote a paper in 2013 in which I do this in a simplified way for only the Mississippi drainage. A few colleagues and I have a proposal in review to integrate isotope-enabled ice-sheet models and the isotopic records for all of North America from the LGM to present. Not sure if we'll get the funding, but in general – yes, agreed!

On the track of isotopes in lakes, I have now seen a new paper on Lake Ontario d18O, and have included that reference, as well as your review Great Lakes oxygen isotope papers. (I usually cite original sources, but the list was getting very long for the end of a discussion section.)

From a more general standpoint, I have added a new paragraph about future directions in the ice-sheet reconstruction realm:

“Pursuit of these targets does not preclude the search for a more representative ice-sheet reconstruction than ICE-6G, the best-fitting ice model of those selected, and better ways to integrate models and data. New reduced-complexity modeling rooted in ice-margin mapping and ice physics \citep{Gowan2016, Gowan2016gmd} presents a strong case to span the gap between the complete empiricism and modeling approaches highlighted here. Surface energy balance calculations \citep{Ullman2015} also elucidate which part of the ice sheet was melting, and when -- providing a link to record of glacial meltwater input. Continued development of numerical ice-sheet models \citep[e.g.,][{}]{Cornford2013, LeMorzadec2015}, with the latter continuing earlier work \citep{Tarasov2006, Tarasov2012} in developing deglacial drainage chronologies, at a coarser spatial resolution than the present study but in a way that is fully coupled with its numerical simulation and the potential to produce outputs at a high enough temporal resolution to recreate short-term events with that left a strong mark in the geologic record \citep[e.g.,][{}]{Kehew1994, Breckenridge2007, Murton2010}. Isotope-enabled climate and ice-sheet models can be combined with isotopic data from geological archives to more closely tie modeled processes and the chemostratigraphic record \citep{Ferguson2015, Gasson2016}. Each of these approaches presents new opportunities to test deglacial drainage patterns and converge towards a solution for the shape and evolution of the North American ice sheet complex that integrates data from glacial, fluvial, and marine systems.”

Specific and Technical comments

Abstract/Conclusions – for me examining the differences between the ice sheet models is an important contribution - but there's no mention in the abstract or conclusions regarding how well the models compare with the synthesis provided. Your synthesis suggests the 6G model performs better than

predecessors – and I wonder if this is worth drawing attention to?

Now added to the conclusions:

“Of the models tested, ICE-6G/VM5a compares the best against the suite of geological data employed in this study, indicating that it is a clear step forward from its predecessors. G12, while only loosely calibrated to match the outlines of the North American Ice Sheet complex, nonetheless performed acceptably well and produces paleodischarge reconstructions with the highest temporal resolution, indicating the value of ice-physics-based numerical modeling in paleo-ice-sheet reconstructions.”

In reading over my abstract, I realized that I did not state which drainage basins were analyzed, so I have added that there.

Pg. 2(43) – d18 O records can provide estimates of meltwater volume, e.g. Moore et al., 2000 (YD interval & outflow from the LIS, *Paleoceanography* 15(1), 4-18.) I think the treatment of the d18 O composition of Lake Superior overflow and LIS is treated too simply in this model – but it’s an excellent attempt to get at meltwater volume from geologic data.

True. I think I was thinking of just viewing the raw data when I said that they didn't give volume, because they certainly can! Added:

“Geologic data provide a timeline of meltwater delivery to one basin or another \citep[e.g.,][{}Ridge1997, Knox2007, Rittenour2007, Breckenridge2009, Murton2010, Williams2012, Breckenridge2015}, with some types of data, such as oxygen isotopes, being able to be analyzed to produce reconstructions of meltwater discharge \citep{Moore2000, Carlson2007, Carlson2009, Wickert2013nature}.”

Pg 2, line 48. “connective tissue” is an odd phrase and ensure is repeated. The sentence could be simplified.

That did look strange. I removed everything after the comma, which solves both issues.

Pg 3, lines 67-74, fig. 4b – it’s odd to see native names given for rivers. Later on in the text, all the rivers are simply referred to by their English name. Perhaps just use the official US, Mexican, and Canadian names? - otherwise you are making errors of omission. For example, the Inuit name for the Mackenzie is omitted, as are native names for the Colorado and Columbia. My suspicions is that there were other native names for the Hudson and St. Lawrence.

It is true that these are errors of omission, but there is something that doesn't sit properly with me when we use only the European names for places that people named long before.

I have decided to remove the other names from the text (using only the primary English-language names), but keep them and augment them on Figure 4b, following your note that they are incomplete. I have retained the note that I wrote for the Figure 4b caption that indicates that the list of languages is incomplete.

The short version, I suppose, is to say that not perfectly including all non-European-American names for the rivers is better than discounting the names that were used before Euro-Americans arrived.

Pg. 4 (102) – missing parenthesis

Thanks!

Pg. 5 (144-145) I'm confused by the idea that the St. Louis River is progressively capturing the Upper Mississippi R - I don't see this in Dean and Philips (a challenging field guide to decipher), and haven't looked at van Hise and Leith 1911. Presumably this is at the divide near Floodwood? It seems like that in general southern basins would capture northern basins due to rebound?

I'm glad you mentioned to take another look at this – since the study was one just looking at maps in 1911, and hasn't been followed up (maybe something for an undergraduate student project?), I have added some weasel words, “...is believed to have progressively captured Upper Mississippi drainage near Savanna Portage...”.

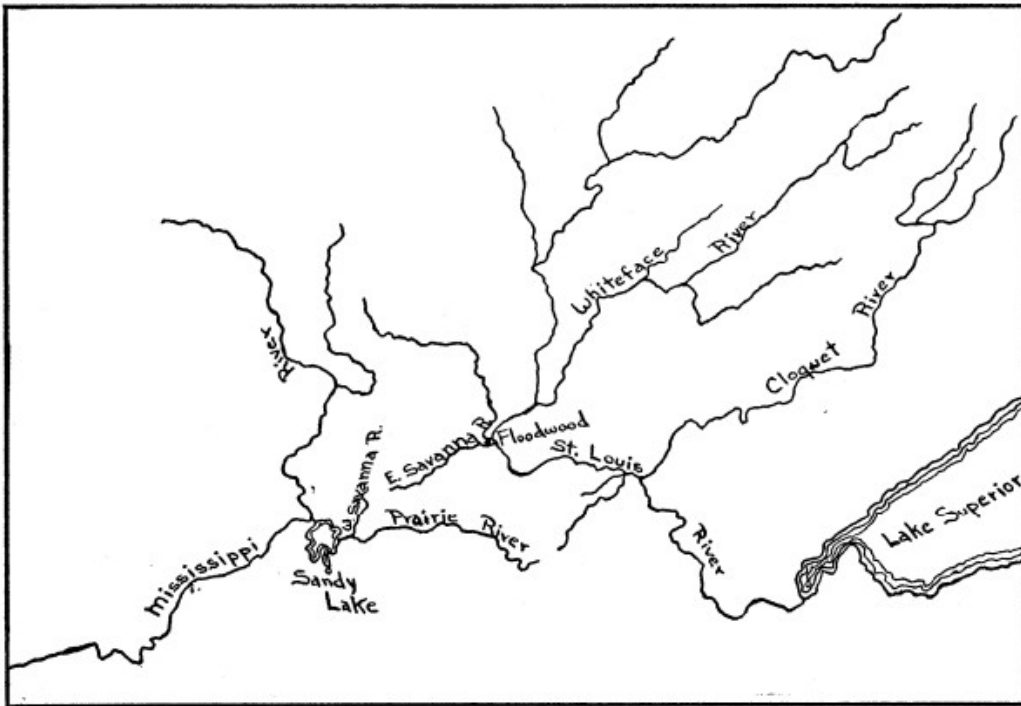
I haven't gotten my hands on an original copy of Dean and Philips either, but the salient parts are reproduced in a Minnesota Historical Society article from 1932: http://www.jstor.org/stable/20161011?seq=1#page_scan_tab_contents

In case you cannot access the article, I've copied the text excerpt and map below:

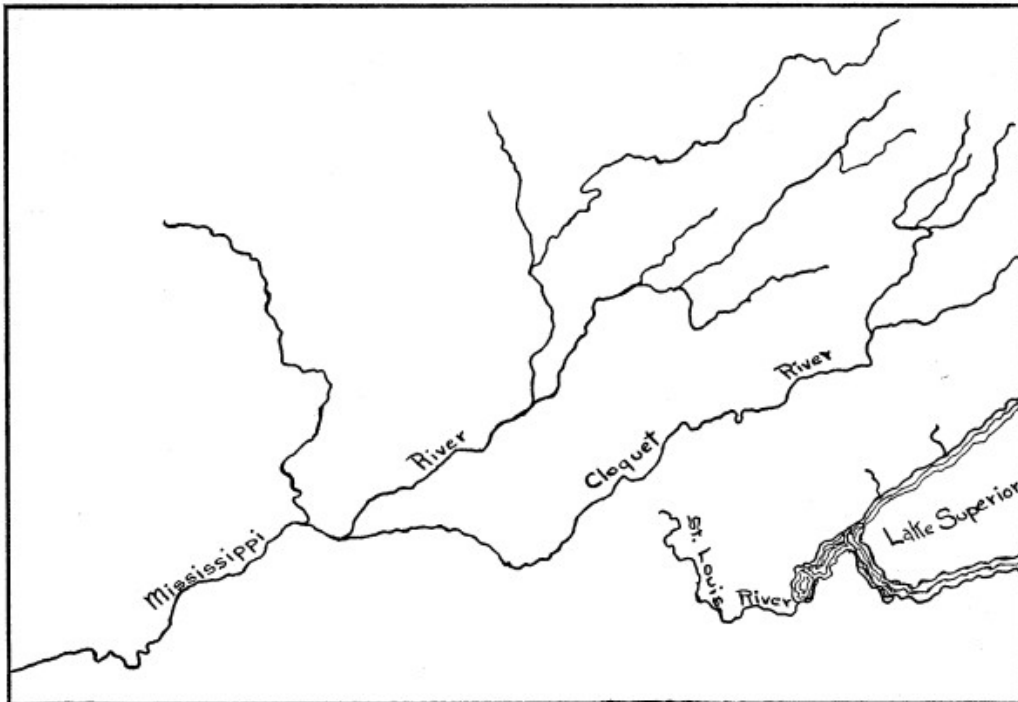
The streams of the Duluth escarpment descend very steeply to Lake Superior; few of them head more than 4 or 5 miles from Lake Superior . . . the greatest distance being 12 to 14 miles, in contrast with lengths of 30 to 75 miles on the north and northeast shores of Lake Superior. Many of them have as steep an average grade as 150 to 250 feet to the mile . . . the general average being 80 to 160 feet to the mile. No one of these rather tumultuous streams has cut a significantly deep valley in the face of the escarpment and most of them have only cut short gorges with small rapids and waterfalls.

Quite in contrast with these steep-graded, rapidly falling streams of the escarpment are the leisurely flowing streams of the plateau surface above. The Cloquet, the upper St. Louis, and various other rivers have an average slope of about 8 or 10 feet to the mile. It is well established that a rapidly flowing stream with a steep grade is able to deepen its valley rapidly and to extend its headwater area so that it encroaches upon the area drained by an adjacent leisurely flowing stream . . . capturing and diverting the latter or some portion of its headwaters. Stream captures or piracies, as they are called, of this kind are common. We should expect, then, that in the course of stream development for a great length of time several of the swiftly flowing streams of the escarpment would have extended their headwaters back to the region drained by the leisurely flowing streams of the plateau surface and captured part or all of these drainage systems. The fact that many of the large streams have not done so is evidence of their youth.

The largest stream in the region, however, seems to have already done just what would be expected . . . and it is natural that the largest stream should be able to do this first. St. Louis River, cutting back at a point near the end of the escarpment where it is rather low, has been able to extend its headwater region northwestward until it has captured the southwestward-flowing Cloquet and the southwestward-flowing stream that forms the present headwaters of the St. Louis itself. These captured streams had been a part of the leisurely drainage system of the plateau surface, and, it seems certain, were within the Mississippi basin. . . . Indeed, a large valley extending southwestward from the town of Floodwood, where the St. Louis now turns abruptly to the southeast, indicates that this is probably the latest elbow of capture at which the piratical St. Louis



PRESENT DRAINAGE SYSTEMS OF THE ST. LOUIS AND MISSISSIPPI RIVER HEADWATERS



ANCIENT DRAINAGE SYSTEMS OF THE ST. LOUIS AND MISSISSIPPI RIVER HEADWATERS

[Maps redrawn from Van Hise and Leith, *Geology of the*

Pg. 6 (182) – perhaps which recent models include a more massive LIS

“it provides a counterpoint to the more recent ice-sheet reconstructions, all of which include a much more massive Laurentide ice sheet.”

(pg. 7, sec. 2.1.3) I wonder about the accuracy of the GIA models – the ice sheet models are visually easy to compare to the Dyke ice sheet reconstructions – but this isn't true for GIA, but there are clear flaws in GIA models for the mid-continent after these data are compared to strandline reconstructions (e.g. Lewis et al, GPQ, 2005). I don't have a suggestion, I just wonder how we can go about assessing and describing the spatio-temporal accuracy of the various GIA models.

There are also inconsistencies (and consistencies) with GPS data, GRACE, paleoshorelines, ...

*The GIA response computed by the combination of the ice sheet reconstructions and their respective solid Earth models is generally not calibrated to lake data from the mid-continent. This is currently on my mind and to-do this, and has been for some time. It is beyond the scope of this current work, but one study that does incorporate them is brand-new: Gowan et al. (QSR: 2016)
<http://www.sciencedirect.com/science/article/pii/S0277379116300646>*

I have been talking with Evan Gowan about trying to use his model for all of North America, and he considers this to be a priority as well; I think his approach of applying a bit of the physics but keeping it fairly simple is a good way to take a first stab.

Of course, Tarasov (the other reviewer) has considered lake records as well for his ice sheet and GIA model.

Pg. 8 (255), add “in” after changes

Oops! Thanks!

Pg. 9 (263-265) I can imagine the 30 arcsecond grid may create problems for areas near drainage divides, where the potential sills have poorly approximated elevations? It seems like it would be a good idea to compare elevations in the DEM used to higher resolution topographic data – and modify the 30 arcsecond DEM if necessary. I'm thinking about headwater areas – such as west of Lake Nipigon, or across potential NW outlet divides of Lake Agassiz.

For this work, at least, I think that the relief around drainage divides has not been so important – in a fairly low-relief landscape in which drainage is rearranged by km-thick ice masses and ~100's of meters of GIA, topographic variability is relatively small. Further, modifying a 1-km-scale continent-wide DEM by hand would be incredibly time-consuming and somewhat subjective, so I would prefer to find an automated way to do this, if possible.

Developing a method for modifying the continental-scale DEM to be hydrologically correct as it were

modified over geologic time, on the other hand, could be an approach to take in the future.

Another option for future work would be to just do these calculations at 3 arcsecond resolution; the whole deglacial cycle currently runs overnight on a desktop computer, so if this could be parallelized onto 100 processors, it would be the same runtime.

And finally, just to add some context – most such drainage calculations are run at 0.5-1 degree resolution. So this is already a lot better than the old standard.

Pg. 12 (351) consider using Caribou “sub-basin” rather than basin, as this is a sub-basin of the great Lake Superior basin.

That makes sense! Changed.

Pg. 12 (359), “Our” is used? – consider “the”

Done.

Pg 15 (407) missing parenthesis after Figure 5

Fixed.

Pg. 15 (426) missing period before For

Thanks, fixed.

Pg. 18 (448) seems like the eskers and ice streams should only provide good evidence for ice divides for time periods when each were active.

Yes; hence the following admission-of-approximation sentence, “As there are few chronological controls on when these ice-streams were active, ice divides are drawn between diverging ice-streams.”

I have also added “While this is a potential source of error...” immediately after this.

Pg. 19 (481) perhaps cite the other discharge studies driven by direct interpretations of data

Added four citations

Pg. 30 (521) perhaps add Levac et al., 2015 (GPC 130: 47-65)

I hadn't seen that paper – added a reference to it.

Pg. 30 (522) I'd remove the reference to Breckenridge, 2015. It looks like the NW outlet didn't open until water dropped from the Tintah, and the model I used for rebound would place after the onset of the YD. I'd argue the main Moorhead low did not occur until long after the onset of the YD.

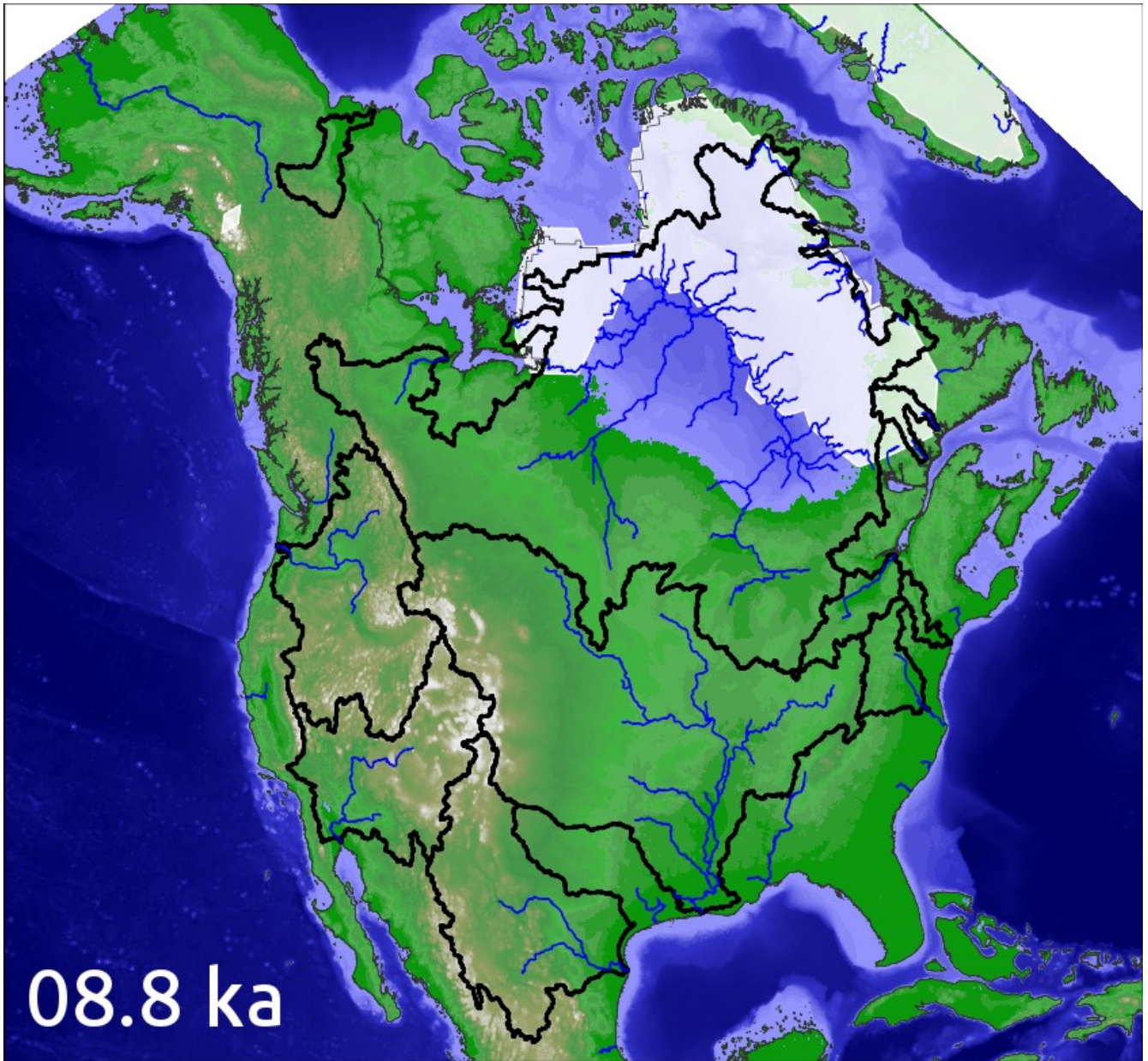
Glad to have the update and the reminder about the age uncertainty in Lake Agassiz' drop, though I should have gotten this right. I have removed the reference, and added in a sentence after my critiques of the drying hypothesis, as follows:

“The drying hypothesis has received further critiques from \cite{Teller2013} and \cite{Breckenridge2015} based on field evidence, with strandline mapping by \cite{Breckenridge2015} indicating that the level of Lake Agassiz dropped when it flowed to the northwest, though current evidence suggests that Lake Agassiz drawdown postdates the onset of the Younger Dryas.”

Pg. 30 (528) Roy et al (2011) QSR 30: 682-692 would be a key publication for the timing of Ojibway drainage. As you note in line 539, our work shows there were two drainages – an initial partial drainage, and a second complete drainage sometime later. The first drainage is not accounted for in current models/hypotheses. I would guess that Lake Agassiz drained first in Manitoba, then either completely re-filled from closer/blockage of that drainage pathway, or the divide between Ojibway and Manitoba was covered by ice advance, re-filling Lake Ojibway. I'm curious if your models suggest other drainage possibilities (water routed from Agassiz to the Mackenzie via Wollaston-Athabasca?)

It would be a key publication! I've added it now.

That is a good question – where did the first drainage go? I've attached a few figures from right before this time (they appear on the next pages). From these, it would seem that a drainage towards the NW (but not the Mackenzie River), via the Manitoba–Saskatchewan northern border region, could be plausible along the very subsided portion of the ice-sheet margin. I don't want to steal your thunder, especially as I can only speculate on this, so I'm not writing anything about it in the paper. Both G12 and ICE-6G show a tiny Mackenzie River drainage basin because of extensive isostatic subsidence; I would be curious if this can be confirmed/denied from stratigraphic / isotopic data from the Beaufort Sea.



G12



ICE-6G

Pg. 31 (557) I'd change or add the reference, or modify the text. Fisher and Lowell (2012) specifically state that Agassiz could not have gone NW "until 10.6-10.1 ka cal BP)". IMO they are ignoring contrary evidence in Fisher et al. 2002 and Fisher and Souch 1998, and place too much confidence on the utility of basal radiocarbon ages to establish ice margins – but they don't suggest a 10.8 drainage in Fisher and Lowell (2012).

I tried to find why I had written this, and I see that this was a mistake made while writing my dissertation that propagated. I have now removed this, and instead mentioned the age range of Agassiz rerouting from your 2015 work.

Pg. 34 (624) most favored hypothesis for NW routing is too generous. As noted earlier, until the Tintah can be directly dated, the strandlines suggest Agassiz did not drain NW until after the YD onset. I have unpublished sediment core data that suggests Agassiz drained east prior to routing NW – as others have argued, there was likely an early Moorhead low – caused by the opening of eastern routes – which was followed by a major Moorhead low when a NW path opened.

Yes, I was sloppy here. Revised to, "Currently, evidence indicates that Lake Agassiz first drained east at the onset of the Younger Dryas \citep{Carlson2007, Rayburn2011, CarlsonClark2012} and then drained north at 12.18 \pm 0.48 ka \citep{Breckenridge2015}."

Pg. 34 (626) – correct the phrase "as the model simulate it"

"as the model to simulate it"

Pg 34. (630) missing parenthesis around Anderson et al., 2014

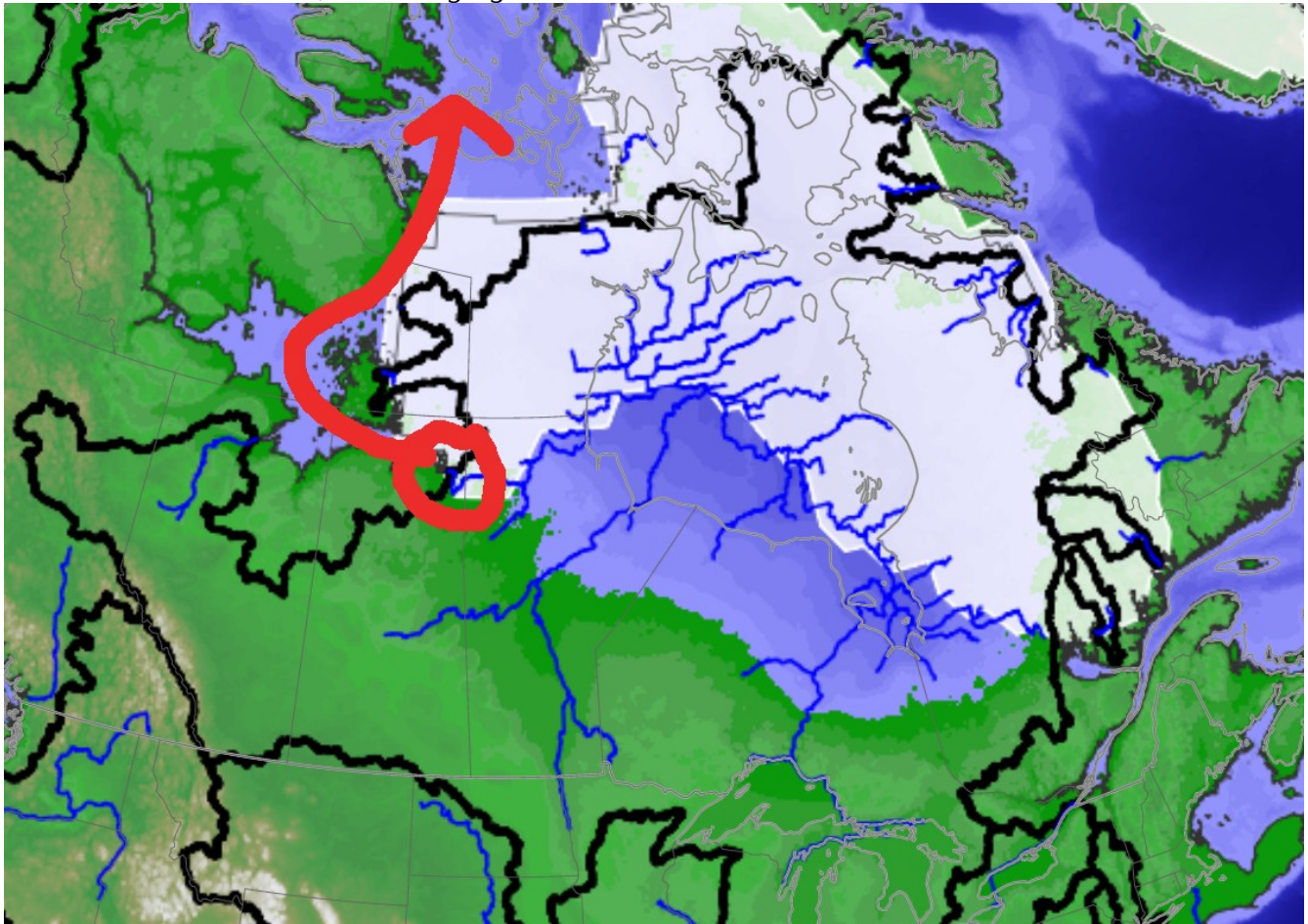
Fixed.

Pg 35 (653) where is this gap to the southwest exactly? This could be an explanation for early (pre-8.2k) partial drainage of Lake Ojibway?

Good question! I have rewritten some awkward phrasing and answered it in the paragraph in the paper:

"... ICE-5G, ICE-6G, and ANU have a drainage divide that splits the east--west Saint Lawrence drainage basin, which at this time extended west to the Canadian Rockies \citep{Teller2004, Breckenridge2012, Stroup2013}. G12 has this continuous drainage basin, but instead of flowing into the Saint Lawrence River, water in its drainage reconstruction flows to the Arctica ocean via a gap at the Manitoba--Saskatchewan border along the southwest corner of the retreating ice-sheet (Fig. \ref{fig:mapsEH}). It is notable that ICE-5G, ICE-6G, and G12 have a tiny early Holocene Mackenzie River basin because the area to the east of the Mackenzie River slopes away from it due to recent ice retreat that exposed the isostatically-subsided land surface; evaluating the plausibility of this drastic Holocene change in the Mackenzie River could be a goal for future field work."

To make this a bit clearer, here is a gorgeous annotated screenshot:



Thinking about this in light of your earlier question about the drainage of Lake Ojibway/Agassiz ~8.6 ka, this model would agree with your intuition.

Figures

I might keep the scale of the axes consistent between the different model runs.

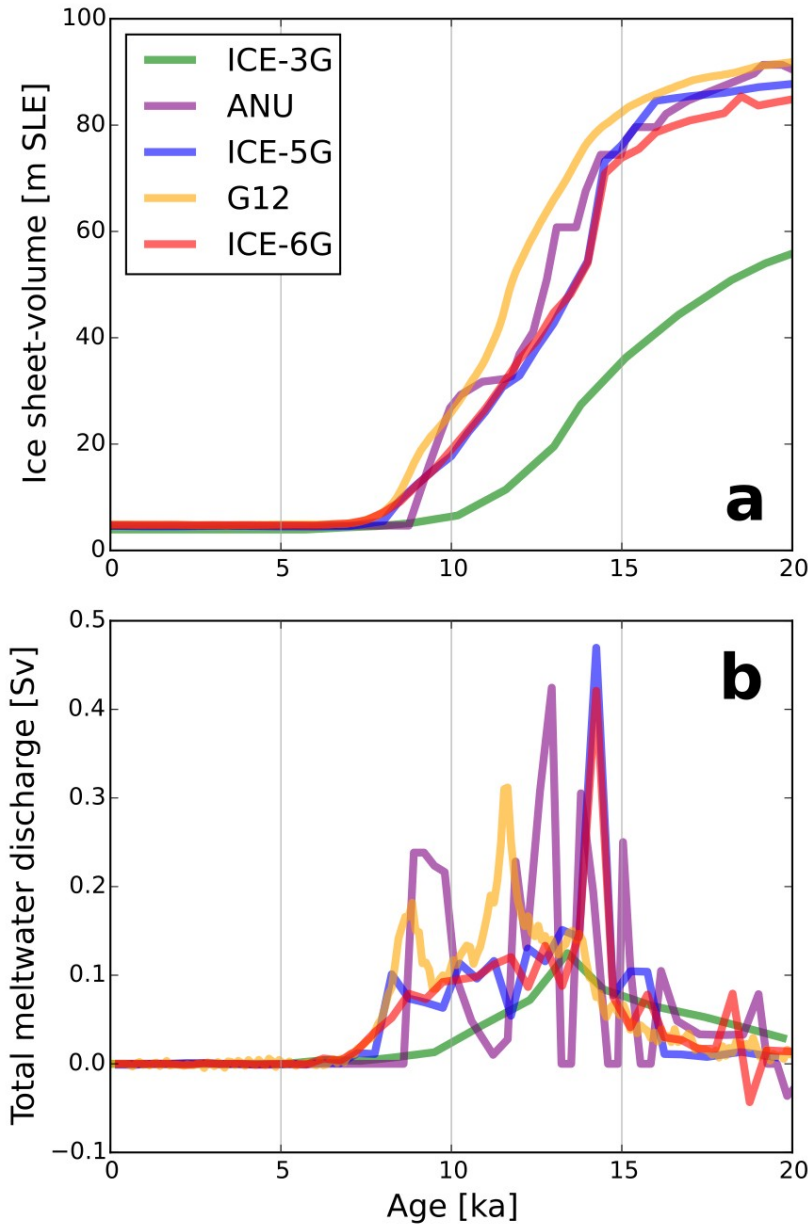
Good point! I had meant to do this, but had not copied over the lines of code that do this to the proper section of the script. Done now.

On the paleogeographic maps, only the “from data” figures include political boundaries – I find these useful when looking at the details between each map.

These are now on all maps.

It might be useful to see total discharge to the oceans through time compared for all the models.

Done – the new figure and its caption are below.



Ice volume histories for each of the ice-sheet reconstructions. **a.** Cumulative change in ice volume within the computational domain, which here contains the entire North American ice-sheet complex and part of Greenland. These ice volumes are represented in meters sea level equivalent, assuming a constant ocean surface area of $3.62 \times 10^8 \text{ km}^2$ [MarshallClarke1999]. **b.** Total meltwater discharge over the computational domain for each ice-sheet reconstruction. ICE-5G and ICE-6G have a noticeable peak during Meltwater Pulse 1A (MWP-1A), 14.65--14.31 ka [Deschamps2012]. Though the time-scale of their numerical model is shifted, [Gregoire2012] also consider their meltwater peak, due to collapse of the ice saddle between the Cordilleran and Laurentide ice sheets, to represent MWP-1A. The vertical axis is in Sverdrups, where 1 Sv is $10^6 \text{ m}^3 \text{ s}^{-1}$.

Responses to Reviewer 2 (Tarasov)

This is a worthwhile study comparing downslope drainage of GIA-based reconstructions (and one chronology from a glaciological model) against geological inferences. It is interesting to see how consistent or not GIA reconstructions are with inferred deglacial meltwater drainage.

The inclusion of only one glaciological model that lacks any data documented calibration and that is not self-consistent with the GIA model used to infer deglacial drainage in this study is a bit unfortunate.

It is – and I had considered asking you for your models as well, but you already computed drainage in your work. In hindsight, I should have asked you if I could check yours with the method developed here.

Technically, the drainage computation is sound. However, as detailed below, more consideration needs to be given to uncertainties in the inferred "data-driven" reconstruction. This is especially the case wrt temporal uncertainties in the inferred high/low drainage. "~15.6 ka" has no clear interpretation. I've also noted some errors in the inferred chronologies stemming from out of date references with old age models. I am not on top of the current literature and as such there may be more such errors.

There were some out-of-date references – thank you for the notes on these. I have checked through many of the others as well. Largely, the poor age control and lack of quantitative error is simply the state of the data and available literature.

There are also some inaccurate statements and some poor referencing as detailed below.

I am curious why proglacial lakes were ignored. The GRASS solver could easily have computed these and would have enabled some quantitative data comparisons as opposed to the purely qualitative comparison currently.

The GRASS solver cannot compute proglacial lakes; see further notes below. Agreed that the lakes are a great target to hit, and one that I am keen to see done in a more quantitative comparison in the future.

Anyway, once comments below are addressed, I would support publication in ESurf. The topic is appropriate for ESurf, figures and abstract are appropriate,...

** detailed comments

generating up to $\sim 1/44$ km of high-albedo ice-surface topography (Kutzbach and Wright, 1985; Ullman et al., 2014)

Neither of those references are appropriate for a claim of " ~ 4 km" elevation
(and high-albedo is obvious). Cite an appropriate primary source.

OK; changed to primary sources and an in-review integrative paper (EGU discussions).

but are evaluated using limited (Tarasov and Peltier, 2006) to no (all others) geologic evidence for past drainage patterns.

Incorrect as worded, Tarasov et al, 2012 used the same geological strandline
data from 2006 as constraints.

My mistake. Fixed.

Reconstructions of past ice-sheet thickness have proliferated (e.g., Tushingham and Peltier, 1991; Lambeck et al., 2002; Peltier, 2004; Tarasov and Peltier, 2006; Tarasov et al., 2012; Gregoire et al., 2012; Argus et al., 2014; Peltier et al., 2015), ... Instead, they are tested against terminal moraine positions and glacial isostatic adjustment,

Incorrect. The models are either tuned or calibrated to these
constraints. This is different than "testing".

Reworded.

"glacial isostatic adjustment" what? That is a physical
process. Do you mean "records of"?

Yes; changed.

Also, Gregoire et al, 2012 was not tested nor tuned against
terminal moraine positions nor glacial isostatic adjustment records
contrary to what I'm inferring you are currently stating.

Gregoire et al. (2012) was not formally tuned to terminal moraine positions, but was built to match the ice outlines. I have separated this as follows:

All but the G12 model \cite{Gregoire2012} were tuned or calibrated terminal moraine positions and records of glacial isostatic adjustment, with the latter being a response that is spread across hundreds of to ~ 1000 kilometers due to the high flexural rigidity of the lithosphere, especially in the Canadian Shield province that was covered by the ice-sheet \cite{Kirby2009, Tesauro2012}. G12 \cite{Gregoire2012} was built to roughly match ice-sheet outlines and evolve more freely.

with the latter being a response that is spread across hundreds of to
~1000 kilometers due to the high flexural rigidity of the lithosphere
Incorrect scales. Yup, there is some smoothing to order 4*
Lithospheric thickness (so order 400 km for a start), but I can also
get significantly different RSL response between 2 sites that are
less than 50 km apart or you can also look at RSL records to also
note such resolution sensitivity

Correct scales, but not precise enough wording. Updated to:

“...with the latter being a response with a characteristic two-dimensional flexural half-wavelength of 500--850 km over the high-flexural-rigidity Canadian Shield province that was covered by the ice-sheet \citep{VeningMeinesz1931, Kirby2009, Tesauro2012}”

many studies do not include any well defined
picture of drainage basin evolution. This....
With putting out this partial straw-man, why is there
no mention of the Tarasov and Peltier, 2006 that does
actively compute the self-consistent drainage chronology,
and that documents the change in drainage basins?

Tarasov and Peltier (2006) was included in the list here of those that “did it right”. I think that the problem is that your figures are colored where meltwater flowed, but do not give full drainage basin boundaries. This removes some of the information that a geologist might be interested in (where is the ice divide, on the ice?) and therefore is a need for a more illustrative work (i.e. this).

As an aside, I think that too few geologists have made note of and/or read your work, so hopefully this paper will provide another portal to approach paleodrainage work.

In general, there exists a lack of recognition of the importance of
Pleistocene drainage rearrangement on river systems.
You are creating another unsupported strawman, especially since you
subsequently cite more references that do recognize the changes
compared to those that don't. Yes there has been a problem to date
with most (but not all) GCMs using fixed present-day routing, but
there is a lot more to paleo science than GCM modelling.

There were more references for the “we did it right” side. That's because I didn't want to miss anyone in that crowd. As you are in that crowd, it might surprise you that there is a broad blank spot in glacial-stage drainage for many geologists. (I'm happy as well about the GCM's coming through, though.)

My point here is that, if the papers that did it right presented the data in a way that were easier for

geologists in related fields to digest, that these continued mistakes wouldn't happen. So once again, the goal of this paper is in part communicating this to a broader community.

I have added some more to the “not getting it, or at least not entirely” category such that there are now more papers on the “not getting it” side.

This ice-sheet surface is chosen to drive the flow-routing calculations because this is what drives ice flow (cf. Cuffey and Paterson, 2010),

"drive" makes no sense. The ice sheet surface is used for flow routing because you are extracting surface drainage. This is independent of the physics of ice flow.

Rephrased: This ice-sheet surface is chosen for the flow-routing calculations because this is what generates the driving stress for ice flow \citep[cf.][]{}Cuffey2010}.

I am extracting surface drainage, yes. But this is not surface melt (i.e. flow). This is ice flow direction and water flow direction, calculated at once. This only gets a bit messy with subglacial water near the ice margins, but at this point, the drainage direction is already determined, and this error is only a factor of ($\rho_{\text{water}} / \rho_{\text{ice}}$)

Greenland Ice Sheet, where the subglacial topography is interpolated from the etopo1 1-arcminute global topographic data set (Amante and Eakins, 2009).

I suggest you obtain a more current DEM for Greenland for any future work where accuracy is critical

I agree; I kept the Amante and Eakins map to match my 2013 work and because Greenland just happens to be inside the rectangle I clipped around North America (not actually discussing it).

H_i , ice-sheet thickness,...therefore are interpolated using an iterative nearest-neighbor approach to remove stepwise discontinuities that would otherwise introduce artifacts in the flow-routing calculations.

Please make the description clear and precise. How do you decide where the ice margin is when downscaling from order 1 degree resolution to 30 arc-second resolution? This would be a critical issue for getting drainage correct, especially when considering switching between Mississippi, St. Lawrence and Arctic drainage for Lake Agassiz region.

Addressed and updated to:

H_i , ice-sheet thickness, is provided by pre-generated ice-sheet reconstructions. These are produced at a coarser resolution than the 30-arcsecond flow-routing grid, and therefore are interpolated using

an iterative nearest-neighbor approach to remove stepwise discontinuities. This smoothing works successively from 1/3 the native resolution to 30 arcseconds, and is necessary because some coarse ice-sheet reconstructions contain multi-cell interior plateaus and several-hundred-meter cliffs at their termini that would otherwise introduce artifacts in the flow-routing calculations. For a 1-degree grid, this requires 5 iterations and leads to smoothing over a half-degree distance ($\sim 40\text{--}50$ km) outside of a given cell. While this somewhat smears the ice-sheet boundary (see, e.g., "ICE-6G" vs. "From data" in Figures \ref{fig:mapsLGM} and \ref{fig:mapsBA}), it does not significantly change its fit to the published ice-margin chronology \citep{Dyke2003, Dyke2004}.

Ice physics, on the other hand, are sensitive to the thermal evolution of the ice-sheet, related thermomechanical effects on ice rheology, and the ice basal conditions
Thermal evolution is a derivative uncertainty as the physics of heat
transfer is well represented in current ice sheet models. The
primary uncertainty for paleo glaciological ice sheet modelling is
the climate forcing (which drives the thermal evolution). For
regions with present-day ice cover, another primary uncertainty is
the basal conditions (topography, roughness, sediment cover,
geothermal heat flux).

Updated to, "Ice physics, on the other hand, are most sensitive to the paleoclimate forcing and basal conditions (e.g., topography, roughness, sedimentary cover, geothermal heat flux) \citep{Tarasov2006, Cuffey2010}"

I investigate both types of models.
Hardly. You only examine one glaciological model and one that was
subject to few constraints.

Clarified to, "Here, I investigate four geophysically-based ice-sheet reconstruction \citep{Tushingham1991, Lambeck2002, Peltier2004, Argus2014, Peltier2015} and one ice-physics based reconstruction \citep{Gregoire2012}, with the latter being allowed to evolve more freely than the tuned model reconstructions of, e.g., \cite{Tarasov2006} and \cite{Tarasov2012}."

and matches geological evidence for the locations of the ice domes and ice-sheet outlines
Misleading: Geological evidence (ie the Dyke 2004 ice margin
chronology) for ice-sheet outlines is IMPOSED on ICE6G.

Good point! Changed to, "and was built to match geologic evidence..."

GMSL was ~ 135 m lower than present
Given different interpretations of what GMSL means (relative to center of
mass of earth, relative to present-day shorelines,..), you need to define

exactly what you mean (Note, Lambeck 2014 eustatic sealevel is neither of
the two definitions I provided above in the brackets)

Lambeck et al. (2014) use ice-volume-equivalent sea-level; I have clarified this point at all places where it appeared. "GMSL" is no longer used because it became unwieldy due to the extra explanation. And, following a closer look at Lambeck's sea-level chronology, I have updated this to 130-135 m, which is the range of LGM values in the supplement provided with the paper

This gravitationally self-consistent sea level theory and its numerical implementation are described by Mitrovica and Milne (2003) and Kendall et al. (2005)

Keen to stay out of the politics of the GIA community, but it is
scientific convention to also cite the originators of a theory and
not just a recent description of it.

My aim here was not to ignore the other work on GIA (Farrell and Clark, Peltier, ...) but to say that it is this particular theory and its implementation that I use here. To clarify this, I have written, "The gravitationally self-consistent sea level theory applied here and its numerical implementation..."

truncated at spherical harmonic degree and order 256. This satisfies the Nyquist flexural wavenumber for the solid Earth models employed, thereby allowing the solutions to be interpolated to an arbitrarily high resolution

Provide a reference for "Nyquist flexural wavenumber" (couldn't find
one on google, web of science,..). Most readers will not understand
what this means and even more will not know it's value is
determined...

I didn't realize that I was the first to describe it in this way – I am simply using the Nyquist frequency, except as a wavenumber. I have clarified as follows: "This satisfies the Nyquist flexural wavenumber for the solid Earth models employed, meaning that the wavelength of the spectral representation is always less than half of the flexural wavelength of the lithosphere in the solid Earth model \citep{VeningMeinesz1931}. The satisfaction of this condition..."

VM5a is largely a simplification of VM2,
Not really. Better to describe it as a modification of VM2 (yes it
has a simplified structure but it also has 10^{21} Pa s viscosity
layer between the Lithosphere and mantle that is not present in VM2.

Good catch – I was only looking deeper. In fact, I don't know what I was thinking when I wrote this section – I completely left out VM1. Now the paragraph is much longer and more complete.

I initiate the GIA modeling of G12 in equilibrium at the LGM

Should mention that is a source of error, as there is no evidence to
support isostatic equilibrium for North America at LGM

True. I added the following, “This is a potential source of error: while the sea-level records of \cite{Lambeck2014} indicate that global ice mass was grew only slightly between 30 and 20 ka, providing enough time to approach isostatic equilibrium, this is only ~2 τ -folding time-scales based on data from the \AA ngerman River and Hudson Bay \cite{Mitrovica2004, Nordman2015}. Therefore, it is very likely that the North American Ice Sheet complex was not yet in isostatic equilibrium at the LGM.”

Our flow–routing calculations neglect geomorphic change for three major reasons. First, significant changes Lake Agassiz outflow directions occurred as a precursor to spillway incision
This approximation is reasonable since you are not modelling lake
levels and are therefore not making direct comparisons to
strandlines.

Sounds good.

Precipitation and evaporation inputs changed through time (Fig. 2). These changes were calculated by Liu et al. (2009) and He (2011) using a continuous LGM (22 ka) to present run of the National Center for Atmospheric Research (NCAR) Community Climate System Model version 3 (CCSM3) (Collins et al., 2006).
Should mention that the TraCE–21K simulation used ICE–5G as a
boundary condition, and so will not be self–consistent with other
deglacial ice thickness reconstructions you are using.

Important – yes. Added, “Therefore, the only fully self-consistent calculations presented here are from the ICE-5G/VM2 reconstruction.”

Flow routing and drainage basins can be calculated only on time steps when the ice–sheet and GIA models provide surface topography (Section 2.3, below), but the ice–sheet contribution to runoff re–quires ice–sheet thickness to be differenced, and thus should be most valid for the times halfway between the ice thickness time–steps.
There is no basis for the last claim when you have 500 year
timesteps (eg ICE5G). On that note, you need to explicitly provide
what timesteps were used

To the first point, I think this is a simple miscommunication. I have removed the awkward wording that you object to; all I was saying is that a central differencing approximation is better than differencing and placing that difference at some other point between the new time steps.

To the second point, I have provided a new paragraph that explains this:

“Time steps vary by model. ICE-3G \cite{Tushingham1991} and ANU \cite{Lambeck2002} use irregular time steps separated by a ~ 1000 years (ICE-3G) and ~ 500 years (ANU). Time steps for ICE-5G are 1000 years from 32 to 17 ka, and 500 years thereafter \cite{Peltier2004}. Model outputs from \cite{Gregoire2012} were provided at 10-year intervals, with GIA computed at uniform 100-year time-steps \cite{Wickert2013nature}, which are the same as those used here for the flow routing calculations.”

This means that these models implicitly include the lakewater volume within the ice mass.... as ice must be used in the models to represent the surface loads of these large lakes in the geologic past....

Incorrect and no easy choices here. Since most of the proglacial lake sites lack proximal RSL constraints, there is little basis to assume that the GIA models are implicitly taking pro-glacial lake loads into account (except wrt global ice volume required to match far-field constraints). Most proglacial lake regions are more constrained in current GIA models by present-day vertical velocities, but this has much poorer time resolution. Furthermore, models such ICE-5G have their ice load spatial extent set to (though with no clear documentation whether any discrepancies are allowed) the Dyke et al 2004 ice margin chronology. This means they have no load where there are proglacial lakes.
It would be best to repeat at least one model calculation with inclusion of pro-glacial lake loads to quantify the uncertainty.

I disagree in the specifics but agree in spirit. As you note, the lakes must take global ice-volume equivalent sea level. Therefore, that lake water volume must be in the ice sheets. To include the lakes would therefore be to ignore mass balance. As my goal here is to evaluate existing models, I will not add the lakes for this work. However, I am 100% on board with the spirit in which this comment is written, and in fact, have plans to include proglacial lakes in future analyses. So stay tuned...

models to match the observed ice-sheet geometry,

"observed"? Sure wish we had observations of paleo ice-sheet geometry...

Wishful thinking! Changed to, “match the observed record of ice-marginal geometry”.

While it is essential to ignore local depressions in the DEM

Why? "Local depressions" = lakes = more data for comparison.

A perfect question for an age-old problem! I have added the following paragraph for clarification and changed the one after it as follows:

“The ability of the \cite{Metz2011} algorithm to route water across local depressions also prevents both natural and artificial (human-built or DEM artifact) highs and lows in the landscape from

stopping the flow routing. This is an advantage because flow-routing calculations without this feature fail unless they are performed on a hydrologically-corrected DEM (i.e. one in which the elevation is altered to produce the known drainage basins). Using a hydrologically-corrected DEM is a potential source of error in reconstructions of the past because its actual terrain has been artificially changed \citep[e.g.,][Grimaldi2007], with difficult-to-discern effects on error.

In spite of these advantages, the \cite{Metz2011} method causes internally drained...

This is an improvement over hosing experiments (e.g., Condron and Winsor, 2012)

This citation does not make sense for this context. Better to cite
PMIP III. Unlike PMIP III hosing that distributed meltwater fluxes
across a band of the North Atlantic, Condron and Winsor for the
first time resolved what would happen (all be it for only a few
years) to discharge from major river outlets. If you are citing
Condron and Winsor as an example of why PMIP style hosing has no
geophysical basis, then you need to rewrite the sentence to make
this clear

I'm not sure where I miscommunicated, but I have reworded the paragraph. To further elaborate. Condron and Winsor (2012) changed a patch of the polar ocean and watched what would happen – an excellent intuition-builder! What I am writing about is computing all discharges to the ocean and then sending these to the ocean – a different task.

The fully-distributed GCM meltwater inputs create a more realistic ice-age ocean

A bit simplistic. It is unclear whether fully-distributed GCM
meltwater inputs can create a more realistic ice-age ocean in
current paleoclimate modelling AOGCMs since these models lack the
resolution to accurately resolve meltwater flux transports and
turbulent mixing

Qualified to, “when coupled with models that can appropriately handle point sources of fresh water, create a more realistic ice-age ocean”

I developed a set of data-driven drainage basin boundaries for each of the study basins ... The second is the ice-sheet margin chronology of Dyke et al...

There are many significant uncertainties in this approach, that need
some tabulation For instance, the Dyke et al chronology has
significant temporal uncertainty (cf Tarasov et al, 2012), that
could significantly affect deglacial drainage chronologies.

Added, “\cite{Tarasov2012} and \cite{Gowan2013} note that uncertainties exist in this deglaciation

chronology”

The second is the ice-sheet margin chronology of Dyke et al. (2003); Dyke (2004), which, when lacking independent information, I use as an approximate set of contours of ice-sheet thickness
Does not make sense. How do you get from an ice margin chronology to
ice sheet thickness contours?

The last part of the ice-sheet to melt is probably close to the zone of maximum thickness of ice. Of course, the highest portion of the ice sheet can migrate, but it is not a bad first-order approximation, and in any case, is corroborated by the records of ice-streaming by Margold et al.

As such, disagreements between these data-derived basins and those derived from models, especially where the data-driven drainage basins are tightly-constrained
For such interpretation, you then need to provide uncertainty
estimates for the data driven approach

I have changed it and the whole paragraph. Uncertainties in the data-driven approach are not quantitatively constrained, but provide a rules-based starting point for thinking about how to look at the data and form simple conclusions. Drainage basin areas are robust based on the terrestrial record, but discharges are less so. I think that the recent work by Gowan et al. (2015, 2016) is the way forward in this regard.

The paragraph now reads:

“The counterpart to these computed histories are those that are driven by more direct interpretations of the data. I write, “counterpart” as opposed to “ground-truth” because these are also to some extent interpretations, based on limited data. Their key utility is that they offer a much simpler path from the raw data to the interpreted drainage basin extents and periods of high discharge, and one that can be more closely tied to geological fact. No age uncertainty is provided for these data-driven reconstructions: The maps display multiple events that coincide, assembled from disparate data sources, of which many do not have well-quantified error. The periods of high and low discharge are marked with sharp lines even where chronological constraints are sparse, and should be viewed as a current geological best estimate. Nevertheless, the large-scale drainage basins noted on Figures \ref{fig:mapsLGM}--\ref{fig:mapsEH}, each of which displays an important time in deglacial history, display the current state of knowledge for that point in the geologic time scale, regardless of chronological error. Therefore, large-scale disagreements between the data-constrained drainage basin outlines and the model results at these key points in the geological time scale generally indicate a need to improve the model-based reconstructions. Mismatch between data-driven reconstructions and models in the discharge histories is less straightforward as it involves rates of ice-sheet melt, while the terrestrial geologic record is more closely tied to drainage basin area, and these potential mismatches are discussed below.”

but by 19.9 ka,

I'm assuming calendar before present? Needs to be clarified

All calendar, especially because not all chronological control is 14C-based. I have now included, “All radiocarbon ages were recalibrated to calendar years BP using IntCal13 \citep{Reimer2013}.” early in the paper.

Figures 6–9: "from data" panel: what is the age uncertainty?

The “from data” panel includes all events occurring during this time, each of which has its own specific age uncertainty, as does the ice margin. These figures are a reconstruction in a very geological sense: no quantified errors, lots of correlation. Why? Lots of unreported error, lots of different data types, data sources for which error is difficult to quantify (how long did the ice stay at these hills based on a handful of dates that may each have their own systematic errors), ...

Therefore, I have made the changes noted two comments above.

Figures 10–15 the high/low/none inferred discharge shading needs
a visually representation of temporal uncertainties in the transitions
eg, with hatching or someother visual texturing

As before, this would be possible in a perfect world, but with the limited age control, we often just know that “X happened between A and B”, and more likely between [subset of ages]. I think that it is better to leave these as they are. However, I do agree that your point is valid, and we do know that “X happened during the Younger Dryas”. So, as noted above, I have changed the respective paragraph to account for this.

18.2±0.5 ka (Rashid et al., 2003)

You need update your citations, this is not a consensus estimate

I can see two reasons for this comment:

1. That “Heinrich Event” is canonically used for the timing of IRD release, typically shorter than the climate impacts (which also use the “Heinrich Event” terminology).

2. That the reference is, in fact, old!

(1) makes it hard to find a “consensus” estimate (it bounces around based on field), but led me to check that the IRD definition of Heinrich Events is in the text (it is).

(2) Pushed me to find a newer reference (Gil et al., 2015). While Carlson comments on the Younger Dryas part, Gil et al. (2015) show that their two-phased IRD pulse from H1 on the Laurentian Fan is consistent with near-European records. As such, I have updated the dates of H1 to:

H1A: 16.8-16.3 ka

H1B: 15.9-15.7 ka

(these are calibrated with IntCal13)

Heinrich Event 0, corresponding to the Younger Dryas,

was 12.8 ka (Clark et al., 2001)
Again, this is an out of date termination estimate.
Current chronologies end it around 11.6 ka
Your age extraction from old references also ignores recent
refinements in ice core age chronologies and C14 calibration

To the first two lines: this is great to be aware of. I was never satisfied with the Hudson Bay chronology from the Younger Dryas through the early Holocene, and now see that numerous papers on it came out in 2015, after I had ended my literature review.

The points about 14C calibration are incorrect: I am using IntCal13 and recalibrating all of the ages from old papers accordingly. To make this clear, I have added the line, "All radiocarbon ages were recalibrated to calendar years BP using IntCal13 \cite{Reimer2013}."

started ~15.6 ka
what does "~" mean? +/- 0.1 ka, +/- 1 ka or ?

"~" means that age control is weak, old, and from a data-poor region. The age is 14C without reported error, because it is an amalgamation of data sources. I have added a comment to this effect.

At 13.1 ka, ice retreat rerouted Glacial Lake Peace
What dating scheme is going to give 100 year accuracy at 13.1 ka???

Revised to ~13.1. I know it is unsatisfactory, but no good error estimates exist for these ages.

Though to answer the question you pose in a more general sense: varved sediments with absolute age control in nearby layers. For example, the Laacher See Tephra.

I then generated histograms for the high-flow and low-flow segments of the model results, and performed Kolmogorov-Smirnov tests between the pairs of distributions for each ice model and river system to test how different they are from one another
again how was temporal uncertainty taken into account?

Temporal uncertainty was not taken into account, but would provide a data-model mismatch if it were on either side. Yes, the geological constraints are also imperfect. This is not a formal data-model integration exercise, but rather is intended to highlight where the two are more or less in agreement. To that extent, I think that it successfully separated the more successful and less successful models.

The ice-physics-based G12 performed the best in generating discharges during high-flow periods that were much higher than those during low-flow periods.
Again need to make clear what timesteps were used. Was the above due

to G12 being provided at higher temporal resolution or not?

Included the time steps as noted above. It is not clear whether finer time steps help or hurt: these values are means across larger swaths of the post-glacial time. One indication that finer time steps can hurt is the lower match with data for the time-step sparse ICE-3G and the time-step rich G12 for the non-glacial rivers, which should be the same in both.

I have added the following sentence to address this,

“For both of these metrics, the inconsistent time steps between ice-sheet reconstructions can affect the fits, as can be seen especially for the Rio Grande and Columbia River, which should be the same for all runs as the same climate reconstruction was employed in all cases. Therefore, these catchments were removed from the above analysis, the broad strokes of which provide a useful template for comparison.”

(1) evaporative loss that is generally >1.5 times today’s observations, in spite of cooler temperatures and a longer lake-ice season

you ignore the likelihood of much higher mean wind speeds near an ice margin and their resultant impact on evaporative loss. This needs to be mentioned as a potentially offsetting factor to your one-sided critique.

Katabatic winds are certain, but there would be no topographic funneling of such winds. I have added text that the winds would be present. The critique was one-sided because Lowell et al. (2013) showed that Lake Agassiz could be a closed basin if they picked parameters outside reasonable limits, and concluded that their modeling supported their hypothesis that lake-level lowered due to evaporation. I have trouble understanding this sort of conclusion. The two parameters that I mention are not the only two that are problematic: they also had to choose a low PDD factor, low runoff, low precipitation, large lake area (contributing to evaporation)... more or less, every knob that could be turned to make Agassiz a closed basin, was turned. So plausible? Yes. Likely? I think not.