Review of : A statistical method to validate reconstructions of late−glacial relative sea level − Application to shallow water shells rated as low−grade sea−level indicators

It is great to see a study that considers how to rigorously define conditional probability distributions for RSL for paleo contexts. However, the current submission has a major flaw. The SLI residuals are not independent and this must be explicitly accounted for. The current formulation explicitly assumes independence but then contradicts this with a 1/N normalization. The consequence of SLI dependence is clear, for instance, when considering the whole Dyke RSL database for North America. The spatial−temporal density of RSL datapoints varies greatly with resultant variations in datapoint redundancy. Without taking this density variation explicitly into account, use of your scoring scheme for say deglacial ice sheet model calibration will give results with model−data fits biased to where datapoints density is highest, even if the sectors where this occurs represent just a small area fraction of the LGM North American ice complex. Until this is addressed, the statistical method is invalid.

I should also note that this flaw might have been avoided with a more careful consideration of the existing litterature (which is not evident in the reference list), eg Briggs and Tarasov, 2013 and Love et al, 2016.

I do not understand the choice of journal. This submission would seem to me much more appropriate in GMD especially since the novelty here isn’t the theory (this is standard Bayesian and probability theory) but the actual implementation. The first line in the abstract also delineates this as a methodology paper: "In this study, we propose a statistical method to validate sea−level reconstructions using geological records known as sea−level indicators (SLIs)." Futhermore, the paper focus is on the method with the viscosity results only provided as an example : "findings are only meant to explain the method and not actually to constrain models."

The paper would also strongly benefit from more concrete details on implementation (probably best included in the supplement) to enable others to do so (especially since the software toolbox is not being made available).

Submission to GMD though requires provision of necessary code/software. This then raises an inequity between the two journals, submit to CPD and avoid the need to provide required code.... I’ll defer the appropriate journal choice to the Chief Editor who should have a clearer sense of journal scope. I would like to see a statement from the editor clarifying how to resolve the scope intersection between GMD and CP with respect to software availability.

I would also like to see explicit consideration of tidal range and wave impacts, especially given the significant tides in Hudson Bay along with the well−known "storm−beach" displacement of SLIs.

Once these issues (and the points below) are addressed, I would see this submission as worthy of publication in GMD (or CP if justified by the chief editor).

specific comments

For this conceptional study, we restrict ourselves to one type of indicators, shallow10 water shells, which are usually considered as low−grade samples giving only a lower limit of former sea level, as the depth range in which they live spreads over several tens of meters, and does not follow a normal distribution
This statement is too sweeping. Eg Dick Peltier and myself treat
certain inter−tidal species (eg Myt. Ed.) as providing more than
just 1−way bounding.

The shells' depth range is derived from the OBIS database,
You need to make clear whether the database only includes shells
that were found in living position as well as whether the shells
were living or not.

In addition to the indicative meaning, each sample's depth is attributed to additional measurement errors which we have to account for. We assume them to be normally distributed, i.e.,
It should be stated whether all considered SLIs were found in a living positio
n.
If not, how are the additional uncertainties addressed?

sums up the uncertainties derived from the leveling 5 of the OBIS data, sigma_OBIS, and those of the SLI, sigma_SLI.
I'm confused. Doesn't the gamma function account for sigma_SLI?

For the time range of considered SLIs, the IntCal13 curve
Why wasn't the marine calibration curve used? Furthermore there
needs to be accounting of Reservoir age uncertainties (and reservoir
age itself if you are using the IntCal curve). The text should also
briefly describe reservoir ages uncertainties (given their
non−trivial space/time variations).

Fig 3. LGM RSL is kind of meaningless since all the SLIs are only present
after local deglaciation. Better to show eg 8 ka RSL around when most of
the critical Hudson Bay dates are available.

eq 9: here am and bm are predicted height and uplift velocity
a_m is the predicted height at t=0 only

lithospehere −> lithosphere

Fig 9
please use a higher contrast colour scheme to make this easier to read

Notation: equations 10−12
Use consistent notation. Eg eq 11 use P_i for conditional probability
but eq 10 uses F_{h,t/m}. Best would be to use standard statistical
notation for conditional probability, eg (h|x) for h conditioned on x.

eq 10

```
# How is this implemented? And how is pa(t) retrieved from oxcal?
# oxcal is a complex enough application that a bit of guidance here
# would help others with their own implementation.


Assuming that the conditional probabilities of the individual SLIs,
P_i, are independent, the joint probalility
eq 12

# the 1/N_data normalization in eq 12 breaks the stated assumption of
# independent conditional probabilities.  The likelihood is the joint
# conditional probability given by P in eq 11.  ln(L) would just be
# SUM(ln(P_sli_i) if the residuals were truly independent. Anyway,
# there is no basis to assume all the SLI residuals are
# independent.
```