REFEREE REPORT FOR THE PAPER "INFERRING PALEO-ACCUMULATION RECORDS FROM ICE-CORE DATA BY AN ADJOINT METHOD: APPLICATION TO JAMES ROSS ISLAND'S ICE CORE" BY N. MARTIN, R. MULVANEY, G. .H. GUDMUNDSSON, AND H. CORR

NOEMI PETRA

The paper applies an adjoin-based inversion method to an inverse problem governed by a one-dimensional transport equation. The inversion is for the accumulation, melting and the initial distribution of age depth parameters from ice age measurements. The inverse problem is formulated as a nonlinear least-squares optimization problem whose cost functional is the misfit between ice age observations and model predictions. Tikhonov regularization is added to the cost functional to render the problem well-posed and account for observational error. The authors show that the reconstructions of the accumulation parameter field deteriorates when the noise added to synthetic observations increases, and the quality of the reconstruction depends on the available observations. The method is also applied with real observations extracted from the James Ross Island's ice cores.

I find the topic of the paper interesting. To the best of my knowledge, this paper is the first to report on an adjoint-based solution method applied to inverse problems using vertical ice information. However, I would like to point out that the inversion method itself is not novel, as claimed in the abstract and conclusions.

Below I list my detailed comments:

1. **Title and Abstract:**
   (a) I sugest replacing "adjoin method" with "adjoint-based method" (the method is not adjoint, it is based on adjoint to compute gradient information).
   (b) It is not easy to understand from the abstract what exactly is the inverse problem, what are the inversion parameters, and what are the observations.
   (c) It is not clear what the authors mean by "gaps in the data".

2. **Sections:**
   Section 1.
   (a) In general I think the authors can improve on the literature review and making clear what exactly is novel in this paper.
   (b) What specific ice-flow parameters are the authors refering to in line 20?
   (c) Please elaborate a little on the "direct search" method applied by Morgan et al., etc. (line 22).
   (d) Line 25: I suggest the authors use the same terminology for the method applied (as in the abstract).

Section 2.

(a) What is the equation in the forward problem (modeling the age of ice)? It seems to be a transport equation (in one spatial dimension), right? It would be helpful to mention this and to list some references for this equation/model.

(b) Why is the right hand side 1?

(c) Is A_d finite-dimensional and A infinite-dimensional? If so care must be taken in first line in eq. 3. For instance one can use an observation operator B that extracts the values of A at the measurement points.

(d) The notation $A(t = T)$ for instance is confusing and looks weird. I recommend using the notation $A(t, z)$, and keep it this way for consistency. For instance when $t = T$, this will translate to $A(T, z)$, or when $z = H$, then one would have $A(t, H)$. Same comment for λ.

(e) After eq. 3, need to make clear that A solves the transport problem (1).

(f) The language after equation (4) talking about the "formal Lagrange approach" is a bit vague (especially for the cryosphere community).

(g) After eq. (5), I recommend explaining that the gradient of the cost J can be found by requiring that variations of the Lagrangian with respect to the forward and adjoint variables vanish.

(h) I recommend adding the details for the derivation of the adjoint into an appendix section. I am surprised to see the positive sign for $\partial_t \lambda$ and w under the partial derivative. Was the gradient checked via finite differences? It would be great to show this verification in the numerics section.

(i) Eq. (6): the boundary condition should hold for all $t \in [0, T]$, and the *final* condition should hold for all $z \in [0, H]$.

(j) I recommend calling the "control equations" gradients and not setting these to zero, since this is true (close to zero) only at the (optimal) solution. (This will also avoid the need to repeat equations in (9).) Same comment as for the adjoint, please add the derivation of the gradient in the appendix.

(k) I suggest splitting Section 2.2 into two sections, namely 2.2. inverse model/problem section and 2.3. the solution method which would show the adjoin-based approach, including everything from section 2.3. (The title of Sec. 2.3. is misleading, it suggests that this section covers the numerical solution/results.)

(l) Why is x either of the three parameters? Is the inverse problem solved for one inversion parameter at a time?

(m) I may have missed it, how are the forward and adjoint problems solved?

(n) Typical value for c in the context of the Armijo condition is 10^{-4} (see Nocedal and Wright, Numerical Optimization, pag. 33). Is there a reason for using $c = 10^{-2}$?

(o) It would be useful for the readers to see the L-curve plot.

Section 3.

(a) How much noise was added to the data? (line 5)

(b) It is not clear (from the text) which colors (lines) show the reference accumulation and age depth.

(c) Eq. 13 zk versus z_k. I also recommend showing the dependence of η on

z, by defining $\eta(z)$.

(d) The steepest descent is known to perform poorly compared to not only Hessian-based methods but also compared to more sophisticated gradient-based methods (e.g., nonlinear CG). I am very surprised to see the very small number of iterations (for the steepest descent method). What do the authors mean by "numerical details" in line 19 / pag 3831? What was the tolerance for the optimization problem? Also, what was the stopping criterion?

(e) Please describe what we see in Figure 1 more specifically (in the text). Also, the blue line is not visible. This is a general comment for all Figures, it is difficult to navigate through colors and lines and match the quantity that is shown and referred to.

(f) It is not clear how the authors "measure the sensitivity". Also, what do the authors mean by "uncertainty in a parameter"?

(g) The numerical results sections (3.1 and 3.2) show inversion results only for the accumulation parameter. It is not clear why the gradient is shown for accumulation, melting and the initial distribution of age depth parameters. Adding more inversion parameters makes the presentation more difficult (especially for readers not expert in inverse problems). If the authors insist to keep the expressions for the more general case, then please make it clear at the beginning of the paper that the inversion results are only for the accumulation.

Sections 4 and 5. Overall these two sections need more serious work. There are several terminology inconsistencies, language tends to be casual, and grammar issues (e.g., line 26). Please try to reformulate. Below I list a few concerns:

(a) Since it is a relatively short paper, and there is some repetition in these two sections, I would suggest to combine these two sections into one.

(b) The first sentence in Section 4 is a strong claim. It is not clear (to me) how the quality of the reconstruction (i.e., the inverse problem solution) depends on the "assumptions made in the forward model". Also, what do the authors mean by: "information recovered"? or "model inversion"?

(c) It is not clear how the "propagation of the uncertainty" was done.

(d) What are "'gaps" (line 5), "perfectly sampled data" (line 11), and "retrieved profile" (line 14). Please be consistent with the terminology.

(e) The sentence (starting in line 16) is confusing. Is the intention to compare with results using other methods?

(f) Is the "uncertainty" in line 21 referring to the noise level?

(g) It is unusual to call the adjoin-based solution method (for inverse problems) applied in this paper the "Lagrange method for constrained optimization". Perhaps the authors meant "Lagrangian-based method", or "Lagrange multiplier method"? See comment in 2c.