Comment on gmd-2022-63

Anonymous Referee #2

Referee comment on "ICLASS 1.1, a variational Inverse modelling framework for the Chemistry Land-surface Atmosphere Soil Slab model: description, validation and application" by Peter J. M. Bosman and Maarten C. Krol, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-63-RC2, 2022

The authors present the implementation of an inverse modelling framework for the CLASS model. This framework includes the variational scheme only since the model is non-linear but it is associated to a Monte-Carlo scheme to retrieve estimates of the posterior uncertainties. Tests for validating the code are shown and finally, an example of a "real data" application.

General comments

The developments described in this paper are relevant since the CLASS model is simple and the inverse framework has been carefully implemented and tested. It may therefore provide a robust and easily manageable tool for field studies such as the one used as an example application. Nevertheless, some major revisions are required before this is made possible, both in the way the paper is written and in the validation of the inverse.

The introduction to the paper is off the mark. It does not explain the links between ICLASS and the efforts of other models but contains a lot of more or less technical information (e.g. on the tangent-linear and adjoint). I think that readers interested in a variational inverse modelling framework may already know about the TL and adjoint. If the aim is to teach users of CLASS what is an inversion and how they can use it, it may not be best done with a paper in GMD.

The order for presenting the variables and various definitions is not always very logical or at least, easy to follow for the reader, particularly in Section 3. The whole of Section 4 and most of Section 8 are not relevant, as well as some theoretical paragraphs in Sections 3 and 5 (see Specific comments for more details).

The validation (Section 9) must deal with more relevant tests and show the uncertainties. The same remark applies to the application example (Section 10): no posterior uncertainties are shown even though ICLASS can estimate them with its Monte-Carlo scheme.

Finally, some very practical information is missing, e.g. about the computation costs.

Specific comments

Introduction

The introduction should be rewritten to include more of the general context surrounding ICLASS e.g. how is it linked to the efforts around other models. Nevertheless, in case they are useful, here are some remarks on specific points:
-p.2 l.31-34: what is the typical frequency of the "golden days" in a year? How are they distributed? At least in the area where the example application is located.
-p.2 l.39-40: this is not true: neural networks or statistical models have no physics at all and their results can be consistent with measurements...
-p.2 l.49 "capable of correcting observations for biases": this is a bit misleading as to what is done by ICLASS. Any inversion set-up can "correct observations for biases" if a control variable is created for it. The issue is whether the resulting corrections have any physical meaning.
-p.3 l.64-65: beware, non-linear is not random (which I assume to be the meaning of "erratically" here).

Forward model
Please check which pieces of information are actually relevant for the inversion framework. If an option is not used in the tests or example application, it may not be explained here.
-p.4 l.96: how is the cloud mass flux included? Or is it not relevant here?
-p.4 l.98: how are cloud effects on the BLH accounted for? Or is it not relevant here?
-p.4 l.101: do you use the option for the Monin-Obukhov similarity?
-p.4 l.102-105: this very long sentence is not clear, please rephrase.
-p.4 l.107: do you use this option?
-p.4 l.107-108: "a-gs" module and big-leaf method are not defined/referenced anywhere. Is it supposed to be commonly known methods?
-p.4 l.111: from which data does the model dynamically compute the long and short wave radiations?
-p.4 l.114: where do the surface temperatures come from?

Inverse modelling framework
-p.4 l.122-123: please clearly list the inputs and/or put them in Fig.1
-p.4 l.123: "[y]our bias correction scheme" has not yet been described. Moreover, the remaining parts of this subsection deals only with xm: please try to make the layout

easier to follow for the reader.

-p.5 l.125-126: what are the "model parameters that are not part of the state"? If they don't, why are they in the model at all?

-p.5 l.126seq: your notations are not conventional - at least, not from the atmospheric inversion conventions. We use R and B for the covariance matrices, for example.

-p.5 l.132-p.6 l.148: all this is part of the general theory of the inversion, it is not particular to ICLASS so I think it must be omitted. Only the information that the observation errors are uncorrelated is relevant.

-p.6 l.154: how can these factors be optimised?

-p.6 l.158-165: this is again part of the general theory of the inversion.

-p.7 3.2: put the definitions of xm before l.125. Maybe xb also.

-p.7 l.179: where do FracH appear in J? This is only indicated in Eq.11.

-p.7 l.180: "this is the topic of the next section": this is not a valid transition between sections. It is useless or may indicate that the sectionning and order of the sections is not logical enough.

-p.7 l.186-197: why may the user desire to specify their own observational energy balance closure residual?

-p.8 l.193: can you conclude on the advantages and limitations of this bias correction?

-p.8 l.195-211: this is the general theory of the adjoint, it is not particular to ICLASS.

-p.8 l.214: what are "forcing vectors"?

-p.8 l.215-217: this is not clear: what is the link between FracH, FH, the observation scaling factors? Please clarify the vocabulary.

-p.9 l.226: what are the advantages and limitations of the numerical derivative compared to the analytical gradient?

-p.9 l.228-230: general theory, remove.

-p.9 l.230-232: if the forward model crashes, aren't there any other issues than the inversion?

-p.9 l.233: "on which state vectors are tested": a missing word?

-p.9 l.236-239: general explanation on the Monte-Carlo principle, not particular to this work.

Figure 1: please indicate also the inputs and outputs.

Adjoint model

I appreciate the very pedagogical drive regarding the adjoint but I think that this section must be removed altogether since I don't think the reader of such a paper expects a lecture on the adjoint.

Error statistics

-p.14 l.381-383: does it invalidate the approach not to keep in the normality assumption? Why?

-p.14 l.384-392: make a graph? Also please check that you don't need to repeat information already given previously or to anticipate.

-p.14 l.395-p.15 l.419: "it will be shortly explained here": not necessary if it is the same as Chevallier et al. (2007), only detail the differences if any.

Output

-p.15 l.422: "in ICLASS": what is the difference with the general definition of the chi-

square?
-p.15 l.426: what does "default" mean? That the user can choose otherwise?
-p.16 l.412-452: a lot of this is generally known and used. Please keep to what is particular to ICLASS. Maybe also use tables.
-p.16 l.453- p.17 l.464: please use a graph or a list of a table.

Technical details of the code
-p.17 l.467seq: here again, please use a graph or a list or a table
-p.17 l.477: "can easily be adapted": wouldn't netcdf be easier to use than pickle?

Adjoint model validation
-p.17 l.480 - p.18 l.506: this is the general theory and must be removed.
-p.18 l.509 - p.19 l.519: same remark.
-p.19 l.522: how many is "the vast majority"? What about those that don't pass? What does "executed in this file" mean? How could you deal with numerical noise?

Invese modelling validation: OSSE
The tests described in this section are useful but they are only very basic tests since, for example, I understand that four out of five are set-up without any perturbations of the observations. The error statistics are not described: are they the "true" ones or are they mis-specified in some tests? The convergence criteria are not discussed, which makes it difficult to compare the tests. Moreover, without the posterior uncertainties, the results are not complete nor comparable.

-p.19 l.527: what does "constructed adjoint" mean?
-p.19 l.530: 5 experiments is a bit too small a number for actual validation of a code.
-p.19 l.535-536: keeping out the background makes them very basic tests.
-p.19 l.547-548: you can quantify the influence of a state parameter with the adjoint.
-p.20 l.549: what is "a very good fit"? How can it be defined without the uncertainties?
-p.20 l.552: "a more complex problem": the problem is not well defined but is it complex?
-p.20 l.553-554: if the parameters have no influence on the cost function (which can be checked with the adjoint), then the inversion is useless.
-p.20 l.554-55: the parameter interdependency issues are not the only ones which may arise in this case.
-p.21 l.567-568: why does the analytical gradient perform better than the numerical calculation?
-p.21 l.573: the framework finds a minimum, not the minimum of the cost function.
-p.21 l.575: what is "a good fit"?

Figure 3: what about the uncertainties?

Application example
-p.21 l.584-586: this is strangely put: observations are "derived" from other observations
- it looks like you use the same word for actual observations i.e physical variables that are

measured and "observations" in the modelling framework i.e. variables of which the model computes an equivalent for comparison.

-p.21 l.587: what are the "non-state parameters"? Put them in the table?

-p.22 l.589: "the detailed settings on chosen model errors, etc" are crucial information, I think they should be put in the main text or at least in an appendix.

-p.22 l.591: 591-592: what about the uncertainties of the prior and posterior? Without them, "a much better fit" cannot be defined. Moreover, fitting the observations is not the reason why inversions are run. The aim is to reduce the uncertainty on the optimised parameters, which is not shown in the figures.

-p.22 l.594: what could be done about the non-optimal error specifications? Lacking information on the computing cost of the inversions, it is not possible to assess whether a number of error set-ups could be tested.

-p.22 l.595 - p.23 l.597: why are some observation streams different from the others with respects to the variance?

-p.23 l.599-600: I don't understand the link between "the model also has a closed energy balance" and the "good fit".

-p.23 l.601-602: this sentence is not clear.

-p.23 l.603 - p.24 l.604: aren't there any data available to check the cumulus clouds or the drop in net radiation?

-p.24 l.608-609: is this assumption very limiting?

-p.25 l.614: "we shortly return to this later in this section": avoid this with a more explicit division in subsections?

-p.26 l.627-628: this sentence calls for a discussion on the impacts of the mis-specification of prior errors.

-p.26 l.630: what does "relatively strongly" mean?

-p.26 l.632: what are these differences?

-p.27 l.637: is 0.05 the average?

-p.27 l.638: what does "reasonably robust" mean?

-p.27 l.642: is there is "no clear reduction in uncertainty", then the inversion was useless. It may not have failed mathematically but its results are not interesting as such. (The fail may be interesting to ask for more observations.)

-p.27 l.642 - p.28 l.643: this is not clear to me.

-p.27 l.647-654: this should come sooner in the text.

Tables 1 and 3: what about the convergence criteria? What about the uncertainties (prior and posterior)?

Table 2: how much are the sensible and latent heat flux observations corrected?

Figures 4,5 and 6,7: what about the uncertainties?

Figure 9: what about the Gaussian assumption?

Concluding discussion

-p.28 l.657-658: general theory of inversions.

-p.28 l.659: what could the more advanced error estimation methods be?

-p.30 l.679-680: the correction of biases is a very complex topic. It is often done outside the inversion framework. a bias correction scheme such as tested here probably cannot be expected to deal completely with the issue.

- please add information on the computation costs.

Technical corrections

-p.3 l.76 and others: why is the term "adjoint" in italics?
-p.6 l.156: what are the (-)? Also found elsewhere.
-p.16 l.436: "similar to" instead of "similar as"?