Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2019-1-AC2, 2019

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

Interactive comment on "Precision Annealing Monte Carlo Methods for Statistical Data Assimilation: Metropolis-Hastings Procedures" *by* Adrian S. et al.

Adrian S. et al.

asw012@ucsd.edu

Received and published: 19 April 2019

Thank you for your detailed review. We were, unfortunately, unable to respond to all your (11-pages worth of) comments before the end of the discussion period. All the comments will be addressed and incorporated into the upcoming draft.

Printer-friendly version

Discussion paper

General Comments:

We agree with your suggestion of the computational cost analysis. But we are concerned that the study of cycling dynamics, performing more experiments, and comparing against more traditional methods, would be too ambitious for our one paper.

L22-23:

Some may also have interest in the probability distribution or higher moments of the distribution.

G(X) is more general than just the (scalar) moments; it may even be a vector valued function. By a proper choice of G(X), say $(X - \mu)^n$, we can represent any n^{th} moment to evaluate and much more.

L33:

It might be worth clarifying: "... information from [measured] data ..."

Good point; thanks.

L 41:

We also want to estimate the measured states which are inaccurately known due to errors in the measurement devices.

We will include the motivation for wanting the measure states as well, as per your suggestion.

L 45:

It might be worth mentioning that in prediction applications, this process is repeated by redefining the time window to start at t_F.

Though it might be true, this paper does not use cycling.

L 52:

Could the bold formatting vector notation from the text for functions F and f be used consistently in equations (1) and (2).

We believe that using the subscript may be clearer for equations (7) and (16), which are of utmost importance. Equations (1) and (2) have, as the input, the full D-dimensional state of the system.

L 57:

Perhaps since ‘F’ is already used as a function name, another symbol could be used for the largest ‘k’ value index.

We think that it would be clearer to have F represent final, and there should be little confusion that the dynamical system F. But we can replace t_F in Line 57 with t_{Final} for clarity perhaps?

L 66:

It should maybe be mentioned that this assumes the model parameters are global, i.e. they are not time-dependent or spatially dependent. Though you could also point out that if you were to cycle the analysis window, without any other changes, it would support a slow time varying global forcing parameter as is.

It has been previously stated multiple times, eg L44, L54.

Equation (3):

This may be an item for the editors to address, but the equation is difficult to read and would benefit from improved formatting. I assume the large ‘dot’ is multiplication. It seems that the second large ‘dot’ (out of three) does not belong. I like the use of color in the equation, but it might help to explain what it’s purpose is.

We are not sure how to improve the formatting for clarity, and we welcome specific recommendations. Thank you for pointing out the mistake. The 'dot' is replaced with an 'x' and the out-of-place 'dot' has been removed. The color is to emphasize the recursion relation, and will be added as a comment.

L 101:

I think there's an error in the denominator, please revise.

Yes, thank you. One should be a and the other should be b.

Equation (6):

I'm assuming the summation with index "k" should be inside the left bracket, since the second term doesn't have any "k" indices. Perhaps also double check the placement of the negative signs once the terms are adjusted.

Seem like the notation is a little off here, thanks for catching it. It has been fixed.

L 132:

Maybe change "perform" to "solve" or "evaluate" the integral.

Okay, evaluate.

L 133:

“[occurs at] the [value of X that produces the] maxima of. . .”

Typo, thanks.

L 138:

“methods to [solve] the integral. . .”

Typo, thanks.

L 148:

change “yielding the smallest value of” to “minimizing”

Typo, thanks.

L 173:

“In practice, [the closer] one can choose X[init] to [correspond to] the maximum of P(X|Y), . . .”

Thank you.

L 176:

“[was] located (Shirman 2018), we [would]. . .”

Typo, thanks.

Equation (8):

This definition should probably appear where X first appears in the manuscript.

A similar form is on L67, but maybe it should be reiterated here.

L 214:

There’s an extra space before the period in “Eq. (9) .”

Yes, thanks.

L 227-229:

It would be helpful to explain the fixed iterations of M-H, what a ‘burn-in’ step is, etc. The ensemble generation stem is important and at the moment it seems a bit ambiguous, which makes analyzing the following steps more difficult as a reader.

We will add some slightl explanation for readers who are not familiar with Monte Carlo methods. However, Metropolis-Hasting is a commonly and widely used algorithm, so we will probably reference a notable paper here. A full tutorial here will be too lengthy. There are many hyper-parameters to tweak, and this is dependent on available computational resources, problem statements, systems, etc.

L 231:

What does the second argument in $N_A(q, 0)$ stand for?

$N_A(q, \beta)$ means the β^{th} annealing step, to be specific, $R_f = R_{f0}\alpha^{\beta}$. We should explain this clearer before hand, probably with a table, but was previously addressed in Eq. (13).

L 238:

"All these paths have zero action."
Just to make sure I understand - the paths all have zero action because they all intersect at the observed points, and at this point the Action term only includes the first term that applies a penalty as the state deviates from the observed values. Is this interpretation correct?

Yes, the interpretation is correct.

L 249:

missing space: “X[0]for”

Thanks.

L 251:

Is there a guarantee that this procedure finds the true minimum and does not get stuck in one of the local minima?

The Metropolis-Hastings algorithm guarantees a probability to climb uphill and thus will be likely to get out of local minima. The annealing heuristic has also been shown, empirically, to find a "superior minimum’. With better tuned hyper-parameters, there is a higher chance to arrive at the global minimum, but never guaranteed.

L 246-264:

In this case, I think this is more explanation than is needed and this iterative processes can probably be consolidated by just defining the first step and the induction step.

We think that it is an appropriate amount of explanation.

Equation (14):

I wonder if there is something that can be learned from computing and examining the error covariance matrix produced by this resulting ensemble. It would be interesting to see how the error covariance matrix changes/converges over each of these iterations from 0 to beta.

Indeed we think that it is interesting. I am not sure if we still have room in this paper to add it here, but we could try it and see if anything interesting happens.

L 301:

I assume the choices of L=5 to 12 is only applied for the D=20 case. But it mentions above (L290) that there are cases with D=5, does this experiment case also have a similar parameter list?

Yes, the parameters are the same. We should clarify.

Equation (16):

I asked before, but I'd like clarification about which 'x' terms in equation (16) are the control variables (i.e. the ones corresponding to X). Based on equation (8), I assume it is the x_l and x_a terms? But I know you are also trying to estimate p. My main point is that it is unclear and needs to be laid out in more detail when first introduced. It's not really consequential, but the order of (y_l - x_l) changed from the previous action

equation (7). It might be better to keep it consistent.

X is defined previously in line 65, but we could remind the readers here by restating things for clarity. Fixed the changed order of x and y.

Figure (3):

“This leads the action to become effectively equal to the action itself” I don’t understand this statement.

Typographical mistake. "...equal to the measurement term itself"

Figure (4):

It seems like most of the members are not finding the correct value of the forcing parameter. Am I interpreting this correctly?

We will likely remove figure (4).

Figure (5):

Are these results aggregating the cases in Figure (4)? They don’t appear to correspond. Or is Figure (4) before convergence?

We will likely remove figure (4).

L 423:

Put the actual url - a reader of a printed copy of the text will not be able to see the link.

Good point, thanks.

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2019-1, 2019.

Printer-friendly version

Discussion paper