Reply to the comments by Referee # 1

We would like to thank the referee for the helpful comments. In the following, the comments by the referee are listed in Italic, and our reply is provided for each comment in Roman.

Comment:
. . . page 5, line 135 and below. As in my previous review, I point out that there is nothing Bayesian about a state space model. The Bayesian aspect comes in later, when priors are put on parameters theta.

Reply: We have modified the expressions (L. 180), and we have also modified the title of Section 3. We appreciate the referee for the comment.

Comment:
Likewise, the distribution $p(x_{0:z}|y_{1:z}, \theta)$ is referred to as the posterior distribution (e.g. page 7 line 188), but it is customary to call it the filtering distribution (on paths), or the smoothing distribution. The term "conditional distribution", used page 8 line 224, is neutral and might be preferred.

Reply: The distribution $p(x_{0:z}|y_{1:z}, \theta)$ is referred to as the 'posterior distribution' in the original paper of PMCMC (Andrieu et al., 2010), although the notation is a little different. We use the term 'posterior distribution' according to this original paper. As noted by Doucet et al. (2001), the sequential Monte Carlo method is derived from the Bayes' theorem for $p(x_{0:z}|y_{1:z}, \theta)$:

$$p(x_{0:z}|y_{1:z}, \theta) = \frac{p(y_{1:z}|x_{0:z})p(x_{0:z})}{\int p(y_{1:z}|x_{0:z})p(x_{0:z})dx_{0:z}}.$$

Therefore, we think it is reasonable to refer to it as a 'posterior' distribution.

We usually refer to $p(x_z|y_{1:z}, \theta)$ as the filtered (filtering) distribution and refer to $p(x_{0:z-1}|y_{1:z}, \theta)$ as the smoothed (smoothing) distribution. Thus, in our terminology, $p(x_{0:z}|y_{1:z}, \theta)$ contains both the filtered and smoothed distributions. We think the term 'posterior distribution' is a more neutral term than the filtered and smoothed distribution.

Comment:
Likewise, "sequential Bayesian model" are usually called state space models (e.g. around Eq 23).

Reply: We agree the term 'sequential Bayesian model' is not common. Since that sentence is not essential, we have removed it to avoid confusion.

Comment:
page 7, line 210. I believe you should indicate that you consider the uniform distribution on the real line for each parameter, also referred to as the flat prior (as opposed to the uniform distribution on an interval). Thus you use what is called an improper prior.

Reply: We thank the referee for the comment. We have modified the expression according to this comment. As we omitted to describe that the parameters except a and b were assumed to be non-negative, we now explicitly say that the uniform distribution on the non-negative real line is considered for the non-negative parameters (P. 8, L. 232–236).

Comment:
page 10 line 284 "can accordingly obtained" + be

Reply: Corrected (P. 11, L. 308). We appreciate the referee for the correction.

Comment:
page 11 line 320: note that if you needed to store the whole trajectories, there would be efficient ways to do so (Jacob, P. E., Murray, L. M., & Rubenthaler, S. (2013). Path storage in the particle filter. Statistics and Computing, 25(2), 487-496.).

Reply: We thank the referee for this information. We was not aware of that paper.
Our main purpose of discarding $x_{1:z-1}$ is to reduce the computational time for SMC at each iteration of MCMC. However, it is true that the memory cost can be reduced by the algorithm by Jacob et al. We have added a mention on the paper by Jacob et al. (P. 12, L. 346).

Comment:
page 13 line 363: "is sufficiently converged" should be rewritten... perhaps "is considered reliable" ?

Reply: The expression has been edited according to the suggestion (P. 13, L. 395).

Comment:
I believe the language is very understandable, I didn't see many spelling mistakes, but some sentences are a bit clumsy. For instance, I believe that most of the articles "the" could be safely deleted ("the MCMC", "the SMC", ...).

Reply: We have tried to improve the expressions. Grammatical errors have also been fixed as far as we found.

Reference

Doucet, A., de Freitas, N., and Gordon, N.: An introduction to sequential Monte Carlo methods, in: Sequential Monte Carlo methods in practice, edited by Doucet, A., de Freitas, N., and Gordon, N., pp. 3–14, Springer-Verlag, New York, 2001.

Reply to the comments by Referee # 2

We greatly appreciate the referee for many constructive comments and suggestions. In the following, the comments by the referee are listed in Italic, and our reply is provided for each comment in Roman.

Comment:
. . . The manuscript would benefit from being written more to-the-point. Several places ? especially in the newly added sections ? the text contain redundancies. I have noted some of these under "technical comments", but they exist throughout the text. I encourage the authors to carefully go through the manuscript and remove such redundancies.
Additionally, I strongly suggest the authors to improve the readability of the paper by starting each section with an introductory paragraph explaining "what is the underlying purpose of the equations that we're going to develop now". While such explanations do exist in the paper, many of these are currently located towards the end of the section, after the equations have been developed. E.g. the idea behind introducing error functions is described almost parenthetically in line 130-133, after the ice dynamic equations have been developed. Another example of this is e.g. on page 8, line 245. On a similar note, it would improve readability if variables were introduced in the text before, rather than after, the equation in which they are used.

Reply: We are grateful for the suggestion. We have tried to remove the redundancies as far as we found. We have also tried to improve the readability. We now provide short introductions at the beginnings of Sections 2 and 3. (We think an introduction for Section 4 is already provided at the beginnings of that sections. But, it contained an inconsistency in the expression. Thus it has been resolved.) The outline described in the end of Section 1 has also been revised. If we can, the variables are introduced before the equation.

The discussion in Lines 130–133 (in the previous version) was about the inconsistency between the assumption for the accumulation rate and that for the thinning factor. As described below, however, we have found that the assumption for the thinning factor is not necessarily inconsistent with the variable accumulation rate. Thus, it has been removed in this revision.

Comment:
To clarify how the PMCMC method is used, I further suggest the authors to explicitly provide the sequence of steps taken in the algorithm, either in the text as a series of numbered steps, or in a figure as e.g. a flow diagram.

Reply: Section 4.2 is intended to provide a summary of a sequence of PMCMC. We have no further idea to show the sequence of PMCMC. However, in the original paper by Andrieu et al. (2010), a pseudo-code of PMCMC is provided. We think this pseudo-code might be helpful for the readers who want to know the algorithm. We have added the mention on the pseudo-code (P. 12, L. 335).

Comment:
The paper lacks a paragraph stating explicitly which glaciological parameters are assumed to be (approximately) constant in time, and which parameters are allowed to vary. Currently, this important information is not obvious from reading the manuscript. This is especially important to clarify since the age model equations are developed under the assumption of steady state, an assumption which is subsequently relaxed with e.g. the accumulation rate allowed to change over time. It appears that the authors assume the horizontal velocity profile (described by variables p, s) to stay constant over time, as well as the thickness of the ice sheet, H. Also μ is held constant in time, and with that also the basal melt rates (please correct me if I'm wrong). Consequently, the entire thinning profile is held constant in time, although this is not consistent with changes in

accumulation. I agree that these are good starting approximations, but all this should be clear from the text. I encourage the authors to explicitly state, in a single paragraph, all the assumptions that go into the development of the age-model equations.

Reply: We agree it is important to describe about the assumptions in Section 2. In this revision, we have improved the descriptions on the assumptions used in our method.

After checking the formulations and assumptions in the model by Parrenin et al., it has been found that it assumes the 'pseudo-steady' state, not the steady state. Accordingly, our method, which is based on the formulation by Parrenin et al., also assumes the 'pseudo-steady' state. We have corrected the explanation on the assumption.

As the referee points out, it is assumed that the velocity profile and the thickness of the ice sheet stay constant over time. It is also assumed that μ is constant. However, the constant μ means that the ratio m/A is constant. Since the basal melt rate m is typically much smaller than A, μ can be maintained to be approximately constant even if A varies.

In this paper, variables which depends on depth (or time) are indicated by the subscript z. The unknown constant parameters are collected into one vector $\boldsymbol{\theta}$. Thus, Section 3 describes what quantities are constant and what quantities depend on time. However, we agree it was not clear which quantities are constant in Section 2. We try to add descriptions of what are constant parameters in Section 2 in the revised version.

Comment:
A second issue concerns the employed tie points. The authors have now included some information about the tie points, but I'm still missing some discussion about the nature of these, and, not least, their associated uncertainties.
Most of the employed tie points are age estimates derived from O2/N2 cycles. These tie points have an absolute uncertainty. However, many tie points commonly used for providing age constraints for ice core timescales are not based on N2O2 measurements. Indeed, probably the most common age constraint for ice cores is based on volcanic tie points; volcanic acidity horizons that can be correlated to sulfate horizons in other ice cores that are more accurately dated. Such more accurate ice core timescales may e.g. have been developed by counting annual layers in the ice core data. In this case, however, the uncertainty of the age of a tie point will no longer be absolute. Instead, one would have an estimate of the uncertainty of the time interval (i.e. the number of layers wrongly counted/not counted) between two age markers. Would it be possible to include such age markers in the current version of the proposed model? I don't think so, since as they would imply a relations between the ages of the ice core at different depths, in contradiction with the discussion at P. 9, line 270-272. I hence suggest the authors to include in the paper a discussion on which type of tie points can be used in the model (and to mention the origin of the uppermost two tie points).

Reply: The uncertainties in a time interval between two age markers can easily be considered by augmenting the state vector $\boldsymbol{x}_z$ with a new variable to represent the time interval. Or, such uncertainties can also be considered by modifying Eq. (23) in the previous version.)

In Lines 270–272 on Page 9 in the previous version, we discuss a technical issue about how to enhance the efficiency of the SMC procedure at each iteration of the MCMC. This discussion is not affected by any glaciological assumptions.

Comment:
P.1, line 7: Explain what the d18O data is used for (this is not trivial information)

Reply: In the abstract in the revised version, we describe how $\delta^{18}O$ is used (L. 8–9).

Comment:
P. 1, line 4, and other places: The accumulation process (i.e. "how does the snow fall") is not important here - what we don't know is how the accumulation rate has changed over time. This should be made clear, here as well as in the rest of the paper.

Reply: It is true that the accumulation process is not important here. We have improved the expression (L. 4). We appreciate for the comment.

Comment:
P. 3, line 66-68, and other places: While I understand what the authors mean by A(z), this should be more carefully described: A(z) is the accumulation rate at the time when the ice now at depth z was deposited. Similarly, in the rest of the paper, the authors still on several occasions mention "accumulation at the surface, A0" (e.g. P. 4, line 90, P. 2, line 337). Accumulation always take place at the surface! Please use correct wording; A0 is present accumulation rate, and A(z) could e.g. be described as "accumulation rate corresponding to the ice at depth z". Also, note that the proper term for A is not accumulation, but accumulation rate.

Reply: According to the comment, the descriptions about $A(z)$ and A_0 have been corrected in the revised manuscript. We are grateful for the comment.

Comment:
P. 3, line 88: The Lliboutry equation is put forward without any justification

Reply: We have added a justification for using the Lliboutry equation (P. 4, L. 99–104).
There is a study which suggested that the Lliboutry equation is not appropriate for a steady dome. However, we can expect that the domes in central Antarctica are non-steady because the Raymond bumps have never been observed. We thus assume that the Lliboutry equation can be used.

Comment:
P. 4, line 107: What kind of age markers are the two upper tiepoints? This should be mentioned, as it is important for the way their uncertainties are incorporated in the model. Additionally, this information could be included (for all tie points) in table 2.

Reply: We have added the information on the two upper tiepoints. As described in the paper by Parrenin et al. (2007), one is ACR-Holocene transition and the other is Beryllium 10 peak (P. 5, L. 122–124).

Comment:
P. 8, line 220: I do not understand the sentence "the posterior .. d18O variation". The sentence seems to imply that the discrepancy between actual accumulation and that predicted from isotopes is caused by something that has to do with the isotope data. But the differences are because the relationship between isotopes and accumulation is only approximate. Please revise the sentence to reflect this.

Reply: We agree that it is more appropriate to say σ_w represents the discrepancy the estimated accumulation rate and the isotope data. We have revised the expression (L. 8, P. 245–247).

Comment:
P. 10, line 296-289: Why is the value of lambda chosen as 1/5? Are there 5 times as many isotope data points to age markers? How does the resulting age model depend on the chosen value of lambda? Is it robust?

Reply: It does not mean the number of isotope data points is 5 times as large as that of age markers. The value of λ was chosen so that the information of age markers becomes effective enough. This is written in the revised version (P. 11, L. 312).

Even if λ was taken to be different values around $1/5$, visible changes were not observed in the posterior distribution Thus, we think the result is robust to λ.

Comment:
P. 10, line 290-292: This is the kind of information that should go into an introductory paragraph of the section.

Reply: This information is already provided in the first paragraph of this section, which says, "the likelihood of the parameters is estimated using the SMC method".

We think the inconsistency in our terminology might cause the confusion. We sometimes referred to $\hat{p}(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})$ as an 'estimate' of the likelihood, but we sometimes referred to it as an 'approximation' of the likelihood. In this revision, we have resolved this inconsistency and we now refer to it as an 'approximation' of the likelihood.

Comment:
P. 11, line 299: Are also the variance parameters assumed to be belong to zero-mean Gaussian distributions? If so, the variances may become negative.

Reply: We omitted to say that the parameters except a and b were assumed to be non-negative. This assumption was realized by setting the prior distribution to be the uniform distribution on the non-negative real line. We now describe it in the revised version (P. 8, L. 232–236).

Even if a negative value is drawn from the proposal distribution for a non-negative parameter, it will be rejected because the prior distribution does not allow negative values. Therefore, we can use a zero-mean Gaussian distribution for the proposal distribution even for a non-negative parameter.

Comment:
P. 11, line 303: I assume that these variances were selected to produce reasonable acceptance rates during the MC sampling. If so, this should be stated. In that case, the way that these are selected should not influence the results (but would affect the burn-in period and the rapidity that the algorithm is moving around in parameter space). If absolutely required, I suggest these numbers to be placed in a table instead.

Reply: This comment is right. We now describe that the variances were selected in order that the Markov chain rapidly moves around in the parameter space (P. 11, L. 329).

Comment:
P. 12, line 333: It does not seem likely that every 5th sample is independent. Later, the authors (P. 15, line 433) quote an acceptance rate of 26%, meaning that on average every 5th sample might be different - but only slightly - and they will definitely not be independent. As the Metropolis-Hastings sampling requires approximately independent samples, it would be preferable for obtaining proper statistics that samples were selected further apart. This is even more important when the authors test the algorithm when using fewer particles (P. 14-15), in which case they quote an acceptance rate of only 7.6%, making every 5th sample highly dependent (indeed, often every 5th sample would be the same as the previous).

Reply: This study does not aim at obtaining a set of independent samples from the posterior distribution, but it simply aims at estimating the posterior distribution. Thus, it would not be necessary that every 5th sample is independent.

In this study, since most of five successive samples would take the same value, the estimate would hardly be affected by discarding four of five samples. That was the reason why we discarded four of five samples. If too many of samples are discarded, it becomes necessary to run much more iterations of the MCMC. Thus, we allow duplicates of some of the samples to be retained.

Comment:
P. 12, line 340: Provide the physical interpretation of the results for these parameters. For instance, with μ being approximately 0, this implies a most likely basal melt rate of zero. This is important glaciological information and should be noted.

Reply: We agree that it is important to remark that the basal melt is near zero. In this revision, some discussions on the results have been added (P. 13, L. 371–375).

Comment:
P. 12, line 346-348, P. 13, line 384-386: Collect these sentences in a single paragraph about the differences to the model used in Parrenin et al (2007).

Reply: Lines 346–348 in the previous version just mentions the thinning factor as a possible reason of the difference in the estimate of the parameters. Lines 384–348 in the previous version provides a main discussion on the estimates of the thinning factor. Thus, we think it is not appropriate to collect these senteces. However, a mention on the latter sentece has been added to the former sentence (L. 380 in the revised version).

Comment:
P. 14, line 408-410: These differences might also have to do with the large covariance between exactly these three variables.

Reply: Right. The marginal posterior distribution for a and b can be modified because of the large covariance with σ_η shown in Figure 3. This is now mentioned (P. 15, L. 441–443).

Comment:
P. 14, line 426-437: Since this run with fewer particles has not converged, I'm not convinced how much information these results contain. I suggest to significantly reduce this paragraph, as well as remove figure 12.

Reply: We agree Figure 12 in the previous version did not contain much information. It has been removed in this revision. The paragraph discussing about the case with fewer particles was also removed accordingly.

Comment:
P. 15, line 464: A better proposal distribution ? for what?

Reply: A better proposal distribution for SMC. We now state it explicitly (P. 15, L. 464).

Comment:
P. 15: Part of the text in the conclusion should be moved to the discussion

Reply: Some paragraphs have been moved to the discussion section (P. 15, L. 460–P. 16, L. 480).

Comment:
P. 1, line 18-20: Write sentence more to-the-point, remove redundancy
P. 2, line 24-26: same

Reply: We have edited them to remove the redundancies (P. 1, L. 20–P. 2, L. 26).

Comment:
P. 2, line 33: Klauenberg et al (2011) also produces a timescale - not just "accumulation and some parameters from d18O". Which are these "other parameters"?

Reply: In this paragraph, some existing Baysian methods for estimating the age–depth relationship are mentioned. We do not intend to ignore their result on the timescale. We have edited this sentence in order to avoid the misunderstandings (P. 2., L. 32–36).

Comment:
P. 2, line 57: section 6 -¿ Section 6 and 7

Reply: We are grateful for pointing out this error. The last paragraph of Section 1 has been revised.

Comment:
P. 3, line 59: Remove "It is thought that". It is unquestionable that the age-depth relationship is determined by accumulation and thinning.
P. 3, line 54: Remove "in year"

Reply: Removed.

Comment:
P. 3, line 64: z is later referred to in units of m, not cm. Which unit is used?

Reply: We are grateful for pointing out this inconsistency. In the revised version, we use m. The unit of accumulation rate shown in the figures has been changed accordingly.

Comment:
P. 3, line 65: The word "past" refers to time, not depth, hence z cannot be referred to as positive towards the past.

Reply: We think it would be our grammatical problem. We meant ξ is defined so that past is postive. But, the expression was confusing. It has been removed to avoid the confusion.

Comment:
P. 3, line 81, P. 4, line 115: The values used for H and Z are appropriate only to Dome Fuji, and should be mentioned in section about Dome Fuji model results, not here.

Reply: They have been moved to the beginning of Section 5.

Comment:
P. 4, line 98-104: This paragraph should be written more to-the-point. Further, it should be made abundantly clear that the accumulation here is allowed to change over time, and that this is in contrary to the assumption of steady state, based on which the thinning rate equation is developed.

Reply: We have tried to improve that paragraph (P. 4, L. 113–117).

Comment:
P.4, line 105: Remove "more reliable age information". Only the age markers provide age information (d18O provides accumulation information).

Reply: That sentence has been edited (P. 4, L. 120).

Comment:
P. 4, line 118, line 128: Provide information here on the distributions of ν_z, η_z

Reply: The information on the distributions of ν_z and η_z has been moved from the sentences before Eq. (13) (in the revised version). We appreciate for the comment.

Comment:
P. 5, line 121-123: Correct the grammar, remove redundancies

Reply: The expression has been edited to remove redundancies (P. 5, L. 146–148).

Comment:
P. 6, line 156-162: Write more to-the-point, remove redundancies as well as the apparent contradiction ("accumulation and thinning might be affected $\cdots$ accumulation and thinning would not be sensitive"). Indeed, I agree that any depth uncertainty on the tie points could to large extent be compensated by increasing the age uncertainties. The effect of such depth uncertainties would, however, depend on the proximity of the age markers.

Reply: In this revision, we tried to remove the redundancy.

As the referee points out, it is not necessarily correct to say, "the uncertainty in age would compensate the possible effect of the depth uncertainty on the estimates of accumulation and thinning". We consider the uncertainty in the increment of age by ν_z. This compensate the potential effect of the depth uncertainty on the estimates of accumulation rate and thinning factor. We have revised the descriptions (P. 6, L. 173–180).

Comment:
P. 6, line 175-181: Remove redundancies

Reply: We think this paragraph does not contain redundancy. However, we have found that the last sentence of this paragraph did not say any essential thing. Thus, it has been removed (P. 7, L. 194).

Comment:
P. 8, line 225: Spell out the name of PMCMC

Reply: The PMCMC is already spelling out in Section 1.

Comment:
P. 9, line 255-256: Redundancy

Reply: We agree the explanation on the likelihood would not be necessary for the readers of this paper. It has been revised (P. 10, L. 180).

Comment:
P. 10, approx. line 277: No need to include this middle step of the equation, since it is not used in the following.

Reply: Removed (Eq. (32) in the revised version).

Comment:
P. 11, line 301: What is the difference between theta' and theta?*

Reply: We gave the definition of $\boldsymbol{\theta}^*$ just before Eq. (37) in the previous version, but it contained an error. We denote a proposal sample drawn from $q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k-1)})$ by $\boldsymbol{\theta}^*$.

The vector $\boldsymbol{\theta}'$ is an arbitrary vector having the same dimension as $\boldsymbol{\theta}$, as indicated just after Eq. (38) in the previous version, which is now Eq. (37).

Maybe, the description on the proposal distribution just before Eq. (37) (in the revised version) version was confusing. It has been modified in this revision.

Comment:
P. 12, line 337, 239: Provide mass balance in the same unit in the two places.

Reply: We now compare them in the same unit (P. 13, L. 367). We thank the referee for the suggestion.

Comment:
P. 13, line 366-369, figure 4: Since all four lines are on top of each other, it makes no sense to include them all on the figure. Their differences are shown well in figure 5.

Reply: We can see small differences among the four lines. We believe the four lines in Figure 4 are helpful for comparing the scale of uncertainty with the entire age range of the ice core.

Comment:
Table 1: This table should include a description for all the employed variables, also a, b, σ_ν etc.

Reply: We thank for the suggestion. We have revised the table.

Comment:
Figure 7: Units for the accumulation data (top panel) are missing. Blue line is not mentioned in figure text. Label on the axis is likely wrong (according to the text, it shows accumulation differences).
Figure 8: Same

Reply: We appreciate the referee for the corrections. Units have been added. Label on the axis has been corrected. We think blue line is mentioned.

Comment:
Figure 10: Does not contain any new information, figure can be removed. Instead it would be interesting to see a figure similar to figure 5 for the case where the five lowermost age markers are removed.

Reply: That figure was replaced by a figure showing the difference from the median of the posterior like Figure 5. We thank the referee for the suggestion.