Anonymous Referee 4

First of all thank you very much for your constructive review, which greatly helped improve the relevance of the manuscript.

The manuscript describes a new implementation of the Datice model, which most notably was used to create the AICC2012 ice-core time chronological framework. The new implementation, IceChrono, is only marginally conceptually different from the Datice model, but is based on a Python platform that makes the model more accessible to the wider community. I believe IceChrono is appropriate for release/publication in GMDD. However, a number of open questions about the function of the model remain, and the presentation is not very clear. Without being an expert on the mathematical formulation of Bayesian models, I agree on the comments provided by Tim Heaton and the other anonymous reviewer:

- The model description needs to be more detailed, especially because none of the previous papers on Datice have clearly presented the modelling framework in an accessible format discussing in depth how to address problems with meaningful uncertainty assignment etc.. Some concrete examples are given below.

We tried to clarify the description of the model.

- The glaciological validity of the results needs to explored further,

The glaciological validity is taken into account using the prior information.

- The section on the Berkner Island dating needs massive improvements or could be removed to free space for a more thorough discussion of the model. I provide comments about the Berkner Island section separately below.

The section on Berkner Island has been removed.

- The manuscript is in serious need of significant improvements in language and clarity. Just to add to the list provided by the other reviewers: use dash in “ice-core records” and similar expressions. The manuscript should pass a thorough grammar check before resubmission.

The language of the manuscript has been checked independently by several people.

With regard to the first point above, I find the best approach to be to describe the workings of the model conceptually first and then, in a separate section that can be skipped by readers with limited technical interest and/or skill in Bayesian modelling, go into the level of detail asked for by reviewer Tim Heaton.

We now have a initial “method summary” sub-section to give the reader a rough idea on how the model works. We did not adopt entirely the Bayesian framework as described by T. Heaton since we could not come up with a clear description using this framework. So we kept our framework based on the books by A. Tarantola:

The true chronology of an ice core, i.e. the ice and air ages at any depth, is a function of three variables (also functions of the depth): the initial accumulation rate (the accumulation rate when and where the particle was at surface), the lock-in depth of the air and the vertical thinning function (the ratio between the thickness of a layer in the ice core to the initial accumulation rate). This is what we call the forward model. These variables are unknown, and to find our optimal chronology we estimate them based on:

- *Prior information about their values on each ice core;*
- *Chronological observations, such as (see Figure Erreur : source de la référence non trouvée): the ice or air age at a certain depth, the time elapsed between two depths, the synchroneity between two ice or air depths within two different ice cores or the depth shift between synchronous ice and air depths within the same ice core.*

All these different types of information, mathematically described as probability density functions (PDF), are assumed to be independent and are combined using a Bayesian framework to obtain posterior estimates of the three input variables (accumulation, LID and thinning) and of the resulting chronologies. Uncertainties on the prior estimates and on the observations are further assumed to be Gaussian and the forward model is linearized, which allow to use the Levenberg-Marquardt (hereafter LM) algorithm (Levenberg, 1944; Marquardt, 1963) to solve this least-squares optimization problem. The philosophy of the method is similar to that of the Datice method (Lemieux-Dudon et al., 2010a, 2010b).

In addition, I would like to raise three other central points, first one regarding the presentation of the IceChrono (and Datice) results, and two of more technical nature. Firstly, I find it unjustified to claim that the Datice or IceChrono models produce “optimal” chronologies. The models use optimization techniques, and therefore, the use of the word “optimization” is acceptable. However, optimization only produces optimal results if the model underlying assumptions are justified, the simplifications insignificant, and the data basis is correct and with correct representation of uncertainty (which is particularly problematic in this context, see comment below on volcanic matching). I know that “optimal” is used in the title of Lemieux-Dudon et al 2010b, but I still think that this use of “optimal” should be discontinued. I think that the word “consistent” in the title of the original Datice paper of Lemieux-Dudon et al. 2010a is the most appropriate description of the approach.

Consistent was used to mean that the chronologies for the different ice cores are consistent (following the stratigraphic information). “Optimal” means that the chronology is optimal with respect to the different chronological informations which are used. This term is very often used in the field of data assimilation and we see no reason not to use it here. Of course we do not claim we produce “perfect” chronologies, but optimal chronologies in a certain sense.

Secondly, a more technical point is the balance between data and model inputs. The Datice and IceChrono models essentially make trade-offs between background scenarios (that are known to be wrong as there would otherwise not be a reason to apply the model) and data-based constraints. The trade-off is made in the models' costs functions, which gets contributions from the misfit between model and data constraints and the deviation of the model from the background scenarios. From eq. 9-14, it appears to me that all contributions are added weighted only by their uncertainties. There may be no "optional" way to do this, but at least the question of how to obtain a good balance between the data constraints and the backgrounds should be thoroughly discussed. In judging whether a fair balance has been obtained, I encourage the author to consider if a measure and/or figures showing the contribution to the cost function from the different terms in the cost function could be useful.

Following Tarantola's framework, we combine different information (prior or observations). Each prior and observation has a weight which is related to its variance-covariance matrix.

Some specific questions to address:

- It seems clear that the resolution of the background scenarios influences the balance between data constraints and backgrounds. The experiment on page 6823 explores the effect of doubling the resolution of the correction functions of the background scenarios, which is a good test to make. Does the resolution of the background scenarios themselves (and not only their correction functions) influence the results? In other words: Is the relative contribution of the deviation of the model from the background scenario dependent on the grid resolution for the background, the grid resolution of the correction function, or both?

The prior scenario are given on the age equation grid. This age equation grid is used to solve Eqs (1-3). Eq (2) is simply solved using a linear interpolation, so it is certainly robust to a change of resolution. Eq (1) and Eq (3) are integrals that clearly converge when the depth step decreases to zero. We made this point clear in the revised manuscript:

It is important to assess whether the formulation of IceChrono1 is robust to a change of resolution: when the resolution increases, the simulations should converge toward a meaningful result. IceChrono1 uses different two different types of grids to optimize the ice cores age scales: the age equation grids and the correction function grids.
The age equation grids are used to solve Eqs (1), (2) and (3). Eq (2) is the value of the ice age function at a given depth, so it is clearly robust to a change of resolution. Eqs (1) and (3) are integrals and are therefore also robust to a change of resolution.
Concerning the correction functions grids, we made two test experiments with different resolutions: AICC2012-VHR and AICC2012-V2HR. The fact that the AICC2012-VHR and AICC2012-V2HR experiments agree well indicates that the formulation of the optimization problem in IceChrono1 is robust to a change of resolution of the correction functions.

- Especially if the only weighting factor of the background scenarios is determined by the width of their confidence intervals, the assignment of confidence intervals/uncertainty is of central importance. What measures have been taken to ensure consistent assignment of background scenario uncertainty between different cores and through time?

The uncertainties on the prior and observations should be determined by an in-depth analysis of the methods which were used to determine them. There is no common solution to this problem and it is up to the user of IceChrono to ensure consistent error assignment to the different chronological information.

- Given that the models include background scenarios for cores that cover from 1 to 8 glacial cycles, and given that the flow regime and accumulation reconstructions are likely much better constrained for some cores and some time intervals than others, is it a reasonable assumption to use the same resolution of background scenarios for different cores and times? For example, by using a temporal resolution of 1 kyr for the accumulation background correction functions, I guess there will be 6-8 times more points that relate to the EDC accumulation reconstruction than the NGRIP reconstruction? If so, could/should the gridding/background scenario resolution be made variable to adapt to this?

It is normal that a 800 kyr long ice core has more weight than the first 100 kyr of the same ice core, since it brings more information to the optimization problem. But the weight also depends on the uncertainties prescribed for the prior: the smaller the uncertainty, the larger the weight. Concerning the resolution of the correction function, we generally assign a correlation length to the priors. This way, if one increases the resolution below this correlation length, we indeed increase the number of points of the prior but these points are not independent anymore, so there is no real new information.

- A specific question along the same lines: Assume that two cores have been linked stratigraphically using 1000 volcanic fix points and 50 methane horizons with the same uncertainty. Does each horizon enter the cost function with the same weight? If so: Is this reasonable? If not: How is the weighting determined?

As explained above, we use the Tarantola framework and we assume that each independent observation is treated the same way. Of course, the weight of an observation is related to its uncertainty. This is even more complicated when there is a correlation between the various observations.

Thirdly, the author is encouraged to discuss how to better represent the uncertainty of volcanic matches in the model. I understand that no new such data are introduced here, but are adopted from the AICC2012 data basis. However, this is an obvious place to improve Datice to increase the confidence of the results. The issue is that volcanic matches come in different categories. In rare cases, tephras have been found in several cores and geochemistry has confirmed that the tephra are indeed coming from the same volcano. However, more commonly, a series of acidity/sulphate peaks representing a characteristic pattern are matched. Assuming that the pattern match is correct, the uncertainties of the individual horizons are on the order of centimeters (i.e. years to decades) and largely uncorrelated between horizons. However, there is a small risk that the pattern match is incorrect, in which case the error can be several meters or more (centuries or millenia), while being highly correlated between the horizons that belong to the same matched pattern. In contrast to this, as far as I know, the uncertainty of volcanic horizons is represented in Datice and IceChrono as Gaussian errors of typically 20-200 years. It would be great to hear if the author has ideas about how to implement a more realistic uncertainty estimate, and to respect the different types of volcanic ties.

The volcanic matching is, as you described, an ambiguous and complex problem. IceChrono v1 assumes gaussian uncertainties, so in this case we assume the volcanic matching has been done correctly. We added this point in the discussion. Note that IceChrono v1 could be used to test different plausible volcanic matching scenarios, but it is not able to synthesize the different results at the moment. The automatic synchronization of records is outside the scope of the present study and will be done later one:

1) All uncertainties are assumed Gaussian. While the Gaussian probability density functions are the most often encountered in Science, it is not always appropriate. For example, the volcanic synchronization of ice cores (e.g., Parrenin et al., 2012b) is ambiguous: a volcanic peak in one ice core can sometimes be synchronized to several other volcanic peaks in the other ice core. IceChrono1 therefore assumes that the features recognition has been performed without ambiguity and that the only uncertainty remaining is the exact match of the two features within their durations.

Concrete comments to the manuscript:

6812

7: “use of dated depth intervals” is slightly misleading as the depth intervals are not dated, but represent a certain duration in the record. Please use another word, e.g. “use of intervals with known duration” or similar. This change should be applied consistently throughout the manuscript.

Changed to “intervals with known durations”.

6813

3: . . . field STRENGTH.

The introduction has been entirely rewritten:

Polar ice cores provide continuous records of key past features of the climate and the environment, with local, regional and global relevance (e.g., Ahmed et al., 2013; EPICA community members, 2004; NorthGRIP project members, 2004; WAIS Divide Project Members, 2013). Tracers of polar climate (e.g., Jouzel et al., 2007), ice sheet topography (NEEM community Members, 2013), water cycle (Schoenemann et al., 2014; e.g., Stenni et al., 2010; Winkler et al., 2012), aerosol deposition (e.g., Lambert et al., 2012; Wolff et al., 2006) and global atmospheric composition (e.g., Ahn and Brook, 2014; Loulergue et al., 2008; Marcott et al., 2014) measured in ice cores unveil sequences of events on seasonal to glacial-interglacial time scales.
However, the prior to interpretation of polar ice core records is the complex task of building two robust time-depth relationships, one for the tracers measured in the ice phase (e.g. water isotopes, particulates and chemical impurities) and one for those measured in the air phase (e.g. greenhouse gas concentration, isotopic composition of gases). The firn, where snow is gradually compacted into ice, constitutes the upper 50-120 m part of ice sheets. The firn is permeable and air is only locked-in at its base, at a depth level called the Lock-In Depth (LID). As a result, the entrapped air is always younger than the surrounding ice at any depth level. Through gravitational fractionation processes, LID is closely related to the isotopic composition of δ15N of N2 in air bubbles data (e.g., Buizert et al., 2015; Goujon et al., 2003; Parrenin et al., 2012a; Schwander et al., 1993). The temporal evolution of the age difference between ice and air at a given depth must therefore be estimated using firn densification modeling and air δ15N. This age difference is essential for clarifying the exact timing between changes in atmospheric CO2 concentration and Antarctic surface temperature during deglaciations (Caillon et al., 2003; e.g., Landais et al., 2013; Monnin et al., 2001; Parrenin et al., 2013; Pedro et al., 2011). However, glacial-interglacial Antarctic firn

changes remain poorly understood (e.g., Capron et al., 2013).
Several strategies have been developed to build ice and gas chronologies. We briefly describe these methods, their strengths and caveats hereafter:
1) Annual layer counting (e.g., Rasmussen et al., 2006; WAIS Divide Project Members, 2013; Winstrup et al., 2012). Only applicable when accumulation rates are sufficiently high to make this annual layer identification possible, this method provides accurate estimates of event durations and small uncertainties on the absolute age of the upper ice sections. However, the cumulative nature of the errors, associated with the increasing number of counted layers, leads to a decrease of the accuracy of absolute age with depth. For instance, the GICC05 (Greenland Ice Core Chronology 2005) composite timescale for Greenland ice cores (Rasmussen et al., 2006; Seierstad et al., 2014; Svensson et al., 2008), is associated with a maximum counting error of only 45 years at ~8.2 ka B1950 (Before 1950 C.E.). This error increases progressively with depth, reaching more than 2500 years at ~60 ka B1950. Annual layer counting techniques cannot be applied when the annual layer thickness is too small to be resolved visually, e.g. in ice cores from central East Antarctica.
2) Use of absolute age markers in ice cores. Well-dated tephra layers identified in ice cores during the last millennia provide precious constraints (e.g., Sigl et al., 2014). Beyond that period, absolute age markers are very scarce. The links between 10Be peaks and well-dated magnetic events (Raisbeck et al., 2007) have provided an age marker for the Laschamp event (Singer et al., 2009). Promising results have recently been obtained using radiochronologic dating tools (Aciego et al., 2011; Buizert et al., 2014; Dunbar et al., 2008).
3) Orbital dating in ice cores. Because there are few absolute constraints in ice cores beyond 60 ka B1950 (limit for the layer counting in the NGRIP ice core), orbital tuning is the most effective method to provide chronological constraints on ice core deepest sections. In the first orbital dating exercises, tie points were determined from the tuning of water isotopic records on insolation curves (e.g., Parrenin et al., 2004), which limits further investigations of polar climate relationships with orbital forcing. More recent chronologies tried to circumvent this assumption and focused on non-climatic orbital markers. Three complementary tracers are currently used: the δ18O of atmospheric O2 (δ18Oatm) (e.g., Bender et al., 1994; Dreyfus et al., 2007), δO2/N2 (e.g., Bender, 2002; Kawamura et al., 2007; Suwa and Bender, 2008) and the total air content (e.g., Raynaud et al., 2007). While the link between δ18Oatm and precession is explained by variations in the water cycle of the low latitudes, relationships between δO2/N2, air content and local summer insolation are understood to arise from changes in the surface snow energy budget influencing its metamorphism. Without a precise understanding of mechanisms linking these tracers to their respective orbital targets, the associated uncertainties remain large, 6 ka for δ18Oatm and 3 to 4 ka for δO2/N2 and air content (Bazin et al., 2013, 2014; Landais et al., 2012).
4) Ice core record synchronization. Inter-ice core matching exercises are undertaken to transfer absolute or orbital dating information from one ice core to another one. It generally relies on the global synchroneity of changes in atmospheric composition (CO2, CH4 concentration, and δ18Oatm) (Bender et al., 1994; Blunier and Brook, 2001; Monnin et al., 2004), the identification of volcanic sulfate spikes within a given area (Parrenin et al., 2012b; Severi et al., 2007) or the hypothesis of synchronous regional deposition of aerosols recorded as ice impurities (Seierstad et al., 2014). In the first case, limitations are associated with the smooting of atmospheric composition changes through firn air diffusion. In the second case, mismatches may arise through incorrect identification of events in different ice cores.
5) Correlation with other well-dated climatic records. In some cases, high-resolution calcite δ18O records and precise U/Th dates on speleothems have been used to adjust ice core chronologies (Barker et al., 2011; Buizert et al., 2015; Parrenin et al., 2007a). Pinning ice core and speleothem records is attractive to reduce absolute age uncertainties especially during past abrupt climatic events of glacial periods. However, these exercises rely on the assumption of simultaneous abrupt climatic changes recorded in ice core (e.g. water isotopes, CH4) and low latitudes speleothem δ18O records (mostly reflecting changes in regional atmospheric water cycle). A main limitation of this method lies in the validity of this assumption.

6) Modeling of the sedimentation process: snow accumulation, snow densification into ice, air bubbles trapping and ice flow (Goujon et al., 2003; Huybrechts et al., 2007; Johnsen et al., 2001). Glaciological modeling provides a chronology derived from the estimate of the annual layer thickness, and therefore, leads to more realistic event durations when the accumulation history and thinning function are well constrained. A side product of glaciological modelling is the quantification of changes in surface accumulation rates, and the quantification of the initial geographical origin of ice. These additional informations are necessary to convert measurements of concentrations of chemical species in ice cores into deposition fluxes, and to correct ice core records from upstream origin effects (e.g., EPICA community members, 2006; Röthlisberger et al., 2008). Caveats are caused by unknown parameters of such glaciological models, such as amplitude of accumulation change between glacial and interglacial periods, the basal melting or the vertical velocity profile, which have a growing influence at depth.

A common and optimal chronology for several ice cores can be built through the combination of several of these methods in the frame of a probabilistic approach. The first attempts used absolute and orbitally-tuned age markers along one ice core to constrain the unknown parameters of an ice flow model (e.g., Parrenin et al., 2001, 2004; Petit et al., 1999). This method had however several limitations. First, the uncertainties associated with the ice flow model could not be taken into account, resulting in underestimated uncertainties. Second, the stratigraphic links between ice cores were not exploited, each ice core was dated separately resulting in inconsistent chronologies. A new probabilistic approach based on a Bayesian framework was subsequently introduced. The first tool, Datice, was developed by Lemieux-Dudon et al. (2010a, 2010b). It introduced modeling errors on three canonical glaciological quantities of the dating problem: the accumulation rate, the LID of air bubbles and the vertical thinning function (i.e. ratio between the in-situ annual layer vertical thickness on its initial surface thickness). This method starts from prior (also called "background") scenario for the three glaciological parameters corresponding to a prior chronology for each ice cores. These scenarios, deduced from a modeling of the sedimentation process, are associated with an uncertainty related to the degree of confidence in these prior scenarios. A minimization approach is then applied to find the best compromise between the prior chronological information for each ice core as well as absolute and relative age markers in the ice and in the air phases. This approach has been validated through the Datice tool and applied to build the Antarctic Ice Cores Chronology 2012 (AICC2012), producing coherent ice and air timescales for five different ice cores (Bazin et al., 2013; Veres et al., 2013): EPICA Dome C (EDC), Vostok (VK), Talos Dome (TALDICE), EPICA Dronning Maul Land (EDML) and NorthGRIP (NGRIP). Further developments of Datice were performed to incorporate additional dating constraints such as the depth intervals with known durations and correlation of errors (Bazin et al., 2014). Datice provides an excellent reference for this Bayesian approach. Still, because Datice has been developed over a long term period with a continuous effort in calculation optimization through methodological improvement, the final code is difficult to access for a non-expert and cannot easily be used as a community tool. We thus identified the need for an open and user-friendly program with a performance similar to Datice but that can be more easily used and implemented by different users within the ice core community.

In this paper, we present a new probabilistic model, IceChrono_v1, based exactly on the same approach as Datice but with improvements and simplifications in the mathematical, numerical and programming aspects. We first detail the IceChrono methodology highlighting the differences to Datice (Section 2). We then perform dating experiments described in section 3 using IceChrono1. We first replicate the AICC2012 experiment, and perform 4 additional experiments to test new functionalities of IceChrono1. The results of these experiments are discussed in section 4. We summarize our main findings in the conclusions, and describe perspectives for future developments of IceChrono in section 5.

11: "pros and cons" - > strengths and weaknesses

Ditto.

17: “but is generally accurate for event duration”. Would this not only be true when the accumulation history is well known also far back in time?

Ditto.

18: It seems like a circular argument here: surface accumulation is modelled in line 13 and suddenly it is a result that can be used for interpretation in line 18.

It is not circular. Modeling of snow accumulation is useful in itself (e.g. for determining chemical fluxes) but it is also used for dating.

24: low-accumulation sites.

The introduction has been entirely rewritten.

27: Add that these methods rely on the existence of climate-independent horizons or the assumption that the synchronized records indeed show the same changes synchronously.

The introduction has been entirely rewritten.

28: (4) The synchronization of ice-core records can be done.

Ditto.

6814:

8: “Optimal”. See above.

See answer above.

8: “therefore”: There is really no argument presented to support this statement.

The introduction has been entirely rewritten.

10: “calibrating 14C ages” is better than using the word “chronology”.

This sentence has been removed.

17: Please specify what these errors are.

This sentence has been rewritten.

23-24: Please replace 1 and 4 with one and four.

This sentence has been rewritten.

25: AICC ChronoloGY not IES

Corrected.

6816:

5: Remove “just”.

Removed.

5: “Un-thins” is modelling slang. Please revise.

We don't see which term could better represent what we mean here. We added the “virtually” word to clarify.

7: “Second member” . . . do you mean “term”? Or right-hand side?

Changed to “right-hand side”.

9-10: Very unclear sentence.

The thinning function is very close to 1 in the firn in all circumstances. So we can't make a big error on it.

6818:

26: J k ii is linked to ICE, right? If not, J k ii and J k ai seems to essentially be the same.

As it is explained in the text:
J k ii is the term linked to ice intervals with known durations.
J k ai is the term linked to air intervals with known durations.
So these are two different terms.

6820:

10: Annex -> Appendix

Changed.

6821:

24: The section starting here is unclear to me. In particular, is it possible to evaluate to which degree the residual vectors are indeed independent and whether their standard deviations are unity (which is what the word “unit” in line 27 means, right)?

We now give more details about this point:

We used the LM algorithm as implemented in the 'leastsq' function from the 'scipy.optimize' library, which also provides an automatic convergence criteria. It does not try to minimize directly the cost function, but rather a residual vector, the components of this residual vector being supposed independent from each others and with a unit standard deviation. Inside each term of the cost function:

$$\boldsymbol{R}^{\mathrm{T}}\mathbf{P}^{-1}\boldsymbol{R}\,, \qquad (1)$$

We allow defining a correlation matrix $\boldsymbol{P}$ *so that the residuals can actually be correlated. We thus used a Cholesky decomposition of* $\boldsymbol{P}$*:*

$$\mathbf{P}=\mathbf{P}^{1/2}\mathbf{P}^{1/2}, \quad (2)$$

and a change of variable:

$$\boldsymbol{R}'=\left(\mathbf{P}^{1/2}\right)^{-1}\boldsymbol{R}, \quad (3)$$

to transform, the residual vector into a vector composed of independent variables with unit standard deviation. The associated term of the cost function can now be written:

$$\left(\boldsymbol{R}'\right)^{T}\boldsymbol{R}', \quad (4)$$

that is, the residuals are now independent and with a unit standard deviation.

6822:

15: The use of annual-layer-counted intervals in Datice is described in a manuscript in revision for Climate of the Past by Bazin et al., so this is only partially true.

We now acknowledge that a new version of Datice to be published allows for observations as intervals with known durations.

22: What is meant by “development”? The rest of the sentence is very convoluted.

This paragraph has been clarified:

7) In IceChrono1, the Jacobian of the model is computed numerically by a finite difference approach while it is computed analytically in Datice. This Jacobian is needed by the minimizer to find the optimal solution X^{opt} and its uncertainty $\boldsymbol{C}^X$. When the optimal solution is found, it also allows to evaluate the uncertainty of the model $\boldsymbol{C}^G$ through equation (20). In Datice, analytical expressions of the Jacobian with respect to X have been derived and these expressions are used to numerically compute the Jacobian for a particular X. In IceChrono, each component of the X vector are alternatively perturbed and the forward model G is run to evaluate how the model G(X) is modified. In other words, the Jacobian is evaluated by a finite difference approach. While a numerical computation of the Jacobian leads to a slower computation time, it leads to a more flexible use of the model since if one modifies the formulation of the cost function, one does not need to derive again analytical expressions for the Jacobian, which is a complex task.

6823:

15: Can you really conclude that the matrices do not describe the physical reality well because they are hard to invert?

When we asked mathematicians about how to invert this kind of matrices, they told us: “If you try to invert such matrices, it is very likely that the formulation of your problem is wrong.”
But we changed a bit the argument, saying that the across-diagonal Gaussian shape leads to a too high correlation for neighboring points.
The definition of realistic correlation matrices is a complex problem that will be dealt in future studies.

6824:

2: It means that IceChrono is robust to a change of the resolution of the correction functions by a factor of 2. It may INDICATE robustness on a more general level.

Changed to “indicate”.

13: A more thorough analysis of WHY IceChrono and Datice differs at the Laschamp event would be useful.

We now use the a posteriori uncertainty estimates obtained directly from Datice and not the official AICC2012 uncertainties. The differences at the Laschamp event disappeared so we modified this paragraph.

19: The consistency of the results confirms that the codings of Datice and IceChrono are performing similarly, which can be taken as an indication that they are correct. It shows nothing about the validity of the assumptions or the method itself. The formulation (and especially the similar statement in the conclusion) should clearly reflect this.

We changed this paragraph into:
"The fact that IceChrono v1 and Datice codes, which have been developed independently using different programming languages and different numerical schemes, agree fairly well in this experiment indicates that both codes perform correctly. Note however that the main principles of these codes are the same, so this agreement is not a confirmation of these main principles."

6825:

1 and 7: repetition

Corrected.

Comments about the Berkner Island dating section:

As requested by other reviewers, the section on the Berkner ice core dating has been removed. It requires an in-depth study and it is not appropriate to include it in this methodological paper.

I'm not opposed to including the section on Berkner Island dating as an example of IceChrono, however if this is to be the first official timescale for the core, then more detail and figures are certainly required.

A discussion of possible reasons for the large accumulation correction around 80 kyr and (in particular) the physical realism of the reversal of the thinning function in fig. 7 below 830 m musty be included.

Moreover all assumptions and parameters that would be needed to replicate this dating by another user of IceChrono should be provided/tabulated or included as supplementary information. Examples of information that should be provided in a revised version as follows:

- p. 6826, line 22-23: 'values of alpha, beta, gamma etc have been chosen to obtain a good fit with independent age markers along the core'. These values should be listed. Please detail which independent age markers have been used and if they are different from the constraints applied later and listed in the tables (and if not, if they are independent from these constraints).
- Figure 3: On the scale of this figure it is difficult to see the differences, if any, between the background scenario and the corrected age. Two things would help: Use colours with more contrast, and add a subplot showing (1) the age difference versus age and (2) the age difference versus depth between the background scenario and the corrected age. It is hardly

relevant to state for each figure whether it has been produced by IceChrono. Mentioning it once the text is sufficient.

- Figure 4: As for Figure 3.
- Figure 5: As above use colours with greater contrast. Also please comment on the substantial differences between the background and corrected scenarios around 30-40 ka and 80-90 ka.
- Figure 6: As above for colours. Also please comment on the deviation from the background at 30-40ka is this due to a particular constraint?
- Figure 7, as above for colours.
- The isotope and gas records on the corrected timescale should also be shown in figures (see also below).

It is difficult to compare figure 7 (plotted versus depth) with figure 3-6 (plotted versus age). Please add a secondary age axis to the right axis of fig. 7.

Table 1 shows age ties based on ‘comparing the deuterium records’ of EDC and Berkner Island. The uncertainties attributed to these ties range between 150 and 300 years. Some more explanation and details should be provided here:

- Why is EDC used for the comparison, instead of EDML, which is closer to Berkner Island and better resolved?
- Also, how are these ties made? Is it by visual matching of Antarctic Isotope Maximum events, or perhaps by some statistical method? In any case, a figure illustrating this would be appropriate: i.e. a comparison of the deuterium records from EDC (or EDML etc) and Berkner Island and same for the CH_4 records.
- It seems optimistic to allocate centennial-scale time-scale uncertainties for deuterium ties. Noise and internal variability between Antarctic ice core sites is at least this large. The onset of deglacial warming in the deuterium record is a good example of this: at WAIS the onset of warming begins 2000 years earlier than at EDC. Yet the Table 1 is making the assumption that the onset of deglacial warming Berkner Island occur within 170 years of the onset of deglacial warming at EDC.
- Along the same lines: Please explain how the synchronization uncertainties in Table 2 are derived and why EDC was preferred over EDML or even a Greenland CH_4 record. Uncertainties in these methane ties as low as 50 years also seem highly optimistic. As above, figures illustrating these ties would be appropriate.
- It is not appropriate to give ages in Table 1 and 2 to two decimal points.

If this is to be the first official/recommended Berkner Island timescale then the ice and gas phase age-depth profiles must be provided as supplementary data and/or deposited in a well-established data repository.