Manuscript prepared for Nonlin. Processes Geophys.
with version 5.0 of the LaTeX class copernicus.cls.
Date: 22 December 2015

A sequential Bayesian approach for the estimation of the age–depth relationship of Dome Fuji ice core

S. Nakano[1,2], K. Suzuki[1], K. Kawamura[3,2], F. Parrenin[4], and T. Higuchi[1,2]

[1]The Institute of Statistical Mathematics, Tachikawa, 190–8562, Japan
[2]School of Multidisciplinary Science, SOKENDAI, Hayama, 240–0115, Japan.
[3]National Institute of Polar Research, Research Organization of Information and Systems, Tachikawa 190–8518, Japan
[4]Laboratoire de Glaciologie et Géophysique de l'Environnement, 38041, Grenoble, France

Correspondence to: S. Nakano (shiny@ism.ac.jp)

Abstract. A technique for estimating the age–depth relationship in an ice core and evaluating its uncertainty is presented. The age–depth relationship is determined by the accumulation of snow at the site of the ice core and the thinning process as a result of the deformation of ice layers. However, since neither the accumulation process nor the thinning process are fully understood, it is essential to incorporate observational information into a model that describes the accumulation and thinning processes. In the proposed technique, the age as a function of depth is estimated from age markers and δ^{18}O data. The estimation is achieved using the particle Markov chain Monte Carlo (PMCMC) method, in which the sequential Monte Carlo (SMC) method is combined with the Markov chain Monte Carlo method. In this hybrid method, the posterior distributions for the parameters in the models for the accumulation and thinning processes are computed using the Metropolis method, in which the likelihood is obtained with the SMC method. Meanwhile, the posterior distribution for the age as a function of depth is obtained by collecting the samples generated by the SMC method with Metropolis iterations. The use of this PMCMC method enables us to estimate the age–depth relationship without assuming either linearity or Gaussianity. The performance of the proposed technique is demonstrated by applying it to ice core data from Dome Fuji in Antarctica.

1 Introduction

Ice cores provide vital information on the climatic and environmental changes over the past hundreds of thousands of years. In order to make use of the chronological records from each slice of an ice core, it is crucial to accurately determine the age for each slice which should be obtained based on an accurate estimate of the relationship between age and depth in the ice cores. For de-

termining the age–depth relationship, dating methods based on glaciological modeling are widely used. However, since the glaciological processes controlling this relationship are not fully known, it is essential to reduce uncertainty by incorporating various types of observational information into the glaciological model. In particular, age markers sometimes provides significant constraints on the age–depth relationship in the dating of an ice core. It is thus important to effectively make use of the information of age markers to determine the age–depth relationship of the ice core. The Bayesian approach is a powerful way to combine a variety of observational information with a model, and it has been applied to the dating of ice cores in a number of studies. Parrenin et al. (2007) considered a glaciological process model which contains several uncertain parameters. They then estimated the parameters for that model using the Bayesian approach and the Markov chain Monte Carlo (MCMC) method, although they did not consider the errors in the glaciological process model in the estimation of the parameters. Klauenberg et al. (2011) took a Bayesian approach to estimate the accumulation and some parameters from δ^{18}O data and to evaluate the uncertainty of the estimate. However, their method was not designed to make use of the constraints of age markers to estimate the age–depth relationship. In order to effectively make use of age markers, it is essential to ensure the consistency of the estimated age within the whole ice core, and it is thus necessary to simultaneously consider a large number of variables to represent the age–depth relationship for the entire ice core. Hence, the Bayesian estimation of the age–depth relationship becomes a high-dimensional problem. Some existing methods handle this high-dimensionality by assuming Gaussianity. Dreyfus et al. (2007) used age markers and a penalized least square method, which assumes Gaussianity, to estimate the age as a function of depth. Lemieux-Dudon et al. (2009) also started by assuming that the uncertainties are Gaussian and that the model is approximately linear. However, if any of the relationships among the variables are nonlinear, Gaussianity does not hold in general. In this paper, we propose a dating method to estimate the age for the entire ice core without assuming either linearity or Gaussianity. The proposed method formulates the age–depth relationship based on a sequential Bayesian approach. The estimation is then achieved using the particle Markov chain Monte Carlo (PMCMC) method (Andrieu et al., 2010), which is applicable to nonlinear non-Gaussian problems formulated as a state space model. This method estimates the age by using the marginal distribution, in which the uncertainties of the parameters in the glaciological model are marginalized out. Hence, it evaluates the uncertainty of the estimated age after considering the effects of the uncertainties in the model parameters.

The remainder of the present paper is organized as follows. In Section 2, we describe the models of the accumulation and thinning processes that control the age–depth relationship. In Section 3, these models are formulated in a framework of the sequential Bayesian approach in order to estimate the age, accumulation rate, and the model parameters. The PMCMC algorithms are explained in Section 4. In Section 5, an application to the Dome Fuji ice core is demonstrated, and the performance of our method is evaluated. Finally, a summary and discussion are presented in Section 7.

2 Dating model

It is thought that the age–depth relationship is determined by two processes: the accumulation of snow at the site of the ice core and thinning due to long-term deformations within the ice sheet (e.g., Parrenin et al., 2001, 2007). We can thus consider the following differential equation describing the relationship between age and depth:

$$dz = A(z)\,\Theta(z)\,d\xi \tag{1}$$

where z [cm] denotes the depth from the surface of the ice sheet, ξ [year] is the age in year at the given z (past is positive), $A(z)$ [cm/year] is the annual rate of accumulation of snow, and $\Theta(z)$ (no unit) represents the thinning factor. It would be more meaningful to consider the accumulation rate as a function of age rather than depth. In this study, however, we first consider the accumulation rate as a function of depth for the convenience of computation. The accumulation with respect to age is then estimated after considering the uncertainty of age as described later. Equation (1) yields the age ξ in the following form:

$$\xi(z) = \int_0^z \frac{dz'}{A(z')\,\Theta(z')}. \tag{2}$$

This implies that the age ξ can be obtained by the integral from the surface at $z = 0$.

Assuming a steady state, the thinning factor $\Theta(z)$ in Eq. (2) can be written using the vertical velocity U:

$$\Theta(z) = U(z)/U(0). \tag{3}$$

Rescaling z and U as

$$\zeta = \frac{H - z}{H}, \qquad u(\zeta) = -\frac{U(z)}{H}, \tag{4}$$

Eq. (3) can be rewritten as:

$$\Theta(\zeta) = u(\zeta)/u(1). \tag{5}$$

In Eq. (4), H is the thickness of the ice sheet. We assume $H = 3031\mathrm{m}$ in this study. The variable ζ is a rescaled vertical coordinate which becomes 0 at the bottom and 1 at the surface, and u indicate the velocity in the ζ coordinate. We rewrite the rescaled vertical velocity $u(\zeta)$ in the following form (Parrenin et al., 2006):

$$u(\zeta) = u(0) + [u(1) - u(0)]\,\omega(\zeta), \tag{6}$$

where $\omega(\zeta)$ is a function satisfying $\omega(0) = 0$ and $\omega(1) = 1$. We assume the following form for $\omega(\zeta)$ (Lliboutry, 1979):

$$\omega(\zeta) = \zeta - \frac{1-s}{p+1}(1-\zeta)\left[1 - (1-\zeta)^{p+1}\right], \tag{7}$$

where s corresponds to the sliding ratio, which is the ratio of the basal horizontal velocity to the vertically averaged horizontal velocity. Denoting the accumulation at the surface by A_0 and the melting at the base of the ice sheet by m, A_0 and m correspond to the vertical velocity at $\zeta = 1$ and that at $\zeta = 0$, respectively, under a steady state. Eq. (6) can thus be rewritten as:

$$u(\zeta) = -\frac{1}{H}\left[m + (A_0 - m)\,\omega(\zeta)\right]. \tag{8}$$

Using Eq. (8), Eq. (5) becomes

$$\Theta(\zeta) = \frac{\omega(\zeta) + \mu}{1 + \mu}, \tag{9}$$

where μ is defined as $\mu = m/(A_0 - m)$. Using Eqs. (7) and (9), the thinning factor Θ can be determined if the parameters s, p, and μ are specified.

In order to obtain the age ξ using Eq. (2), it is also necessary to give the accumulation rate A. In this study, A is treated as an unknown variable to be estimated. However, since the accumulation rate A is related to the temperature in the Antarctica, it can be constrained by some proxy of the temperature. We used the δ^{18}O data taken at Dome Fuji (Watanabe et al., 2003), which is plotted in Figure 1, as a proxy for the temperature to estimate A. Since the vertical profile of the age ξ is associated with the profile of A, the information from the δ^{18}O data is also effective to improve the estimate of the age ξ.

At several depths, we can also use use more reliable age information. We used such age values as tie points when estimating the age–depth relationship. The age, depth, and uncertainty (2σ) for each tie point used in the paper are shown in Table 2. The first two points were given by Parrenin et al. (2007), and the other points were determined from the relationship between O_2/N_2 and the summer insolation (Kawamura et al., 2007). The age ξ was estimated considering both the δ^{18}O data and the tie points.

3 Bayesian model

Discretizing the vertical coordinate z with an interval Δz, the integral in Eq. (2) for any discretized z can be calculated using the following recurrence relation:

$$\xi_{z+\Delta z} = \xi_z + \frac{\Delta z}{A_z \Theta_z} + \nu_z \sqrt{\frac{\Delta z}{A_z \Theta_z}} \quad (z = 0, \Delta z, 2\Delta z, \ldots), \tag{10}$$

where ξ_z denotes the age at z, and Z denotes the depth at the bottom of the core ($Z = 2505\mathrm{m}$ in this study). At the surface ($z = 0$), ξ_0 is defined as zero. The accumulation rate and the thinning factor in the interval from z to $z + \Delta z$ are denoted by A_z and Θ_z, respectively. The term ν_z represents an unknown variation that are attributed to processes that are not taken into account in Eq. (2). We multiply ν_z by $\sqrt{\Delta z/(A_z\Theta_z)}$ in order that the variance of the unknown variation per year is not changed down to the bottom of the core. The thinning factor Θ_z can be obtained according to Eq.

(9) in a steady state. Eq. (9) does not consider all of the effects governing the thinning process Θ_z but some effects such as the effect of impurity (Freitag et al., 2013) are not specified in Eq. (9). The errors in Θ due to such unspecified effects would also be contained in ν_z in Eq. (10). The accumulation rate A_z is treated as an unknown variable, and its transition from z to $z+\Delta z$ is described by the following recurrence relation:

$$\log A_{z+\Delta z} = \log A_z + \eta_z \sqrt{\frac{\Delta z}{A_z \Theta_z}} \quad (z = 0, \Delta z, 2\Delta z, \ldots). \tag{11}$$

Note that the transition of A_z is described using its logarithm in Eq. (11) in order to guarantee $A_z > 0$. The term η_z represents the (unknown) variation in the accumulation rate. Eq. (11) allows non-steady variations in the accumulation rate A, which is not necessarily consistent with the steady-state assumption for the thinning factor Θ in Eq. (9). However, ν_z in Eq. (10) allows errors which might affect the age-depth relationship including the errors in the thinning function and the misestimation of the accumulation rate. Thus, the errors due to the inconsistency in the assumptions between A and Θ would be compensated by ν_z to some extent. We hereinafter assume $\Delta z = 1[\mathrm{m}]$. Eqs. (10) and (11) can thus be rewritten as follows:

$$\xi_{z+1} = \xi_z + \frac{1}{A_z \Theta_z} + \frac{\nu_z}{\sqrt{A_z \Theta_z}}, \tag{12}$$

$$\log A_{z+1} = \log A_z + \frac{\eta_z}{\sqrt{A_z \Theta_z}}, \quad (z = 0, \ldots, Z - \Delta z). \tag{13}$$

In order to apply a Bayesian approach, we introduce conditional probability density functions based on Eqs. (12), (13). We assume that ν_z and η_z obey the normal distributions $\mathcal{N}(0,\, \sigma_\nu^2)$ and $\mathcal{N}(0,\, \sigma_\eta^2)$, respectively, where we denote a normal distribution with mean μ and variance σ^2 by $\mathcal{N}(\mu, \sigma^2)$. Accordingly, the conditional distribution of ξ_{z+1} given ξ_z for each z becomes

$$p(\xi_{z+1}|\xi_z, \boldsymbol{\theta}) = \mathcal{N}\left(\xi_z + \frac{1}{A_z \Theta_z},\, \frac{\sigma_\nu^2}{A_z \Theta_z}\right), \tag{14}$$

and the conditional distribution of A_{z+1} given A_z for each z becomes a log-normal distribution as follows:

$$p(A_{z+1}|A_z, \boldsymbol{\theta}) = \log\mathcal{N}\left(A_z,\, \frac{\sigma_\eta^2}{A_z \Theta_z}\right), \tag{15}$$

where $\boldsymbol{\theta}$ indicates a collection of unspecified parameters such as p and s in Eq. (7). The full definition of $\boldsymbol{\theta}$ will be provided later. Because $p(\xi_{z+1}|\xi_z, \boldsymbol{\theta})$ and $p(A_{z+1}|A_z, \boldsymbol{\theta})$ are given, the joint distribution $p(\xi_{z+1}, A_{z+1}|\xi_z, A_z, \boldsymbol{\theta})$ can also be defined. We hereinafter combine ξ_z and A_z into one vector $\boldsymbol{x}_z$. Thus,

$$p(\boldsymbol{x}_{z+1}|\boldsymbol{x}_z, \boldsymbol{\theta}) = p(\xi_{z+1}, A_{z+1}|\xi_z, A_z, \boldsymbol{\theta}). \tag{16}$$

Estimates of ξ_z and A_z for each z are obtained on the basis of their posterior distributions given the tie points and the $\delta^{18}\mathrm{O}$ data. For the k-th tie point τ_k at depth z_k, we assume the following

relationship between τ_k and the modeled age ξ_{z_k}:

$$\tau_k = \xi_{z_k} + \varepsilon_k, \tag{17}$$

where ε_k is the discrepancy between the age at the tie point and the modeled age. We assume the tie points have no depth uncertainty. The depth uncertainty would not make essential effects on the estimate of the age for each slice of the ice core labeled with a depth value, even if its true depth is uncertain. The estimates of accumulation and thinning might be affected by the depth uncertainty. But the estimates of accumulation and thinning would not be sensitive to the depth uncertainty because accumulation and thinning are related with the increment of depth rather than the absolute depth from the surface. In addition, the uncertainty in age would compensate the possible effect of the depth uncertainty on the estimates of accumulation and thinning. Assuming that ε_k obeys the normal distribution $\mathcal{N}(0, \sigma_\varepsilon^2)$, the conditional distribution of τ_k given ξ_{z_k} becomes

$$p(\tau_k | \xi_{z_k}) = \mathcal{N}\left(\xi_{z_k},\, \sigma_\varepsilon^2\right). \tag{18}$$

The δ^{18}O data, which are associated with the accumulation rate, can be abundantly obtained from the ice core at Dome Fuji. Multiple data points for δ^{18}O are sometimes available within an interval of a single meter, and we used the mean δ^{18}O value for each such interval. It was assumed that A_z, the accumulation rate in the interval from z to $z + \Delta z$, is associated with δ^{18}O as follows:

$$\delta^{18}\mathrm{O}_z = a \log A_z + b + w_z, \tag{19}$$

which was also used by Klauenberg et al. (2011). Assuming that w_z obeys the normal distribution $\mathcal{N}(0, \sigma_w^2)$, the conditional distribution of $\delta^{18}\mathrm{O}_z$ given A_z becomes

$$p(\delta^{18}\mathrm{O}_z | A_z, \boldsymbol{\theta}) = \mathcal{N}\left(a \log A_z + b,\, \sigma_w^2\right). \tag{20}$$

Although we assume the regression coefficients a and b do not depend on age, it is not guaranteed that the accumulation rate and δ^{18}O have the same linear relationship over the entire period recorded in the ice core. Even if we could accept the linear assumption between the accumulation rate and δ^{18}O, a and b might change due to the variation of climatological conditions other than temperature. However, the uncertain variable η_z in Eq. (11) represents the variation of accumulation rate including not only the variation related with δ^{18}O but also the variation due to other unknown factors. Thus, errors in our assumption in the relationship between the accumulation rate and δ^{18}O would be absorbed by η_z to some extent. In addition, ν_z in Eq. (10) allows some errors in the relationship between the accumulation rate and δ^{18}O.

We define the vector of the available data for each z as $\boldsymbol{y}_z$. If both the tie point τ_{k_z} and the δ^{18}O data are available at z, then, $\boldsymbol{y}_z = (\tau_{k_z}, \delta^{18}\mathrm{O}_z)^T$. If the δ^{18}O data are available but a tie point is unavailable, we define $\boldsymbol{y}_z = \delta^{18}\mathrm{O}_z$. In the case that neither a tie point nor δ^{18}O data are available, we define $\boldsymbol{y}_z = \emptyset$. Using $\boldsymbol{y}_z$, the conditional distributions in Eqs. (18) and (20) can then be combined into the conditional distribution $p(\boldsymbol{y}_z | \boldsymbol{x}_z, \boldsymbol{\theta})$ for any z, where we define $p(\boldsymbol{y}_z = \emptyset | \boldsymbol{x}_z, \boldsymbol{\theta}) = 1$.

Our aim is to estimate $\boldsymbol{x}_{0:Z} = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_Z\}$ based on the sequence of the data $\boldsymbol{y}_{1:Z} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_Z\}$. If a set of the parameters $\boldsymbol{\theta}$ was given, we could obtain an estimate of $\boldsymbol{x}_{0:Z}$ from the posterior distribution $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z}, \boldsymbol{\theta})$. However, since the value of $\boldsymbol{\theta}$ is not specified, it is necessary to take into account the uncertainties of $\boldsymbol{\theta}$ in estimating $\boldsymbol{x}_{0:Z}$. We obtain an estimate from the marginal posterior distribution given $\boldsymbol{y}_{1:Z}$, where $\boldsymbol{\theta}$ is marginalized out:

$$p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z}) = \int p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z}, \boldsymbol{\theta})\, p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z})\, d\boldsymbol{\theta}. \tag{21}$$

Since $\boldsymbol{y}_z$ is conditionally independent of $\boldsymbol{x}_{z'}$ given $\boldsymbol{x}_z$ when $z' \neq z$,

$$p(\boldsymbol{y}_z|\boldsymbol{x}_{0:z}, \boldsymbol{\theta}) = p(\boldsymbol{y}_z|\boldsymbol{x}_z, \boldsymbol{\theta}). \tag{22}$$

Hence, $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z}, \boldsymbol{\theta})$ satisfies the following recurrence equation:

$$\begin{aligned} &p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z}, \boldsymbol{\theta}) \\ &\propto p(\boldsymbol{y}_z|\boldsymbol{x}_z, \boldsymbol{\theta})\, p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z-1}, \boldsymbol{\theta}) \\ &= p(\boldsymbol{y}_z|\boldsymbol{x}_z, \boldsymbol{\theta})\, p(\boldsymbol{x}_z|\boldsymbol{x}_{z-1}, \boldsymbol{\theta})\, p(\boldsymbol{x}_{0:z-1}|\boldsymbol{y}_{1:z-1}, \boldsymbol{\theta}). \end{aligned} \tag{23}$$

This equation expresses a sequential Bayesian model. By applying Eq. (23) recursively, we can obtain $p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z}, \boldsymbol{\theta})$ for any z. Thus, sampling from $p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z}, \boldsymbol{\theta})$ can be achieved using the sequential Monte Carlo (SMC) method (Doucet et al., 2001; Liu, 2001). If z in Eq. (23) is set at the depth at the bottom of the ice core (i.e., $z = Z$), we obtain $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z}, \boldsymbol{\theta})$, which provides the estimate of the age given all the data for the entire ice core.

A Bayesian approach also enables us to estimate the parameter $\boldsymbol{\theta}$. The posterior distribution of $\boldsymbol{\theta}$ given $\boldsymbol{y}_{1:Z}$ in Eq. (21) is calculated using the following equation:

$$p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z}) \propto p(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}) \tag{24}$$

The vector $\boldsymbol{\theta}$ contains all of the unspecified parameters used above. The full definition of $\boldsymbol{\theta}$ is as follows:

$$\boldsymbol{\theta} = (A_0\ a\ b\ \mu\ p\ s\ \sigma_\nu\ \sigma_\eta\ \sigma_w)^T. \tag{25}$$

An approximation of $p(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})$ can be calculated using the SMC method. Therefore, if the prior $p(\boldsymbol{\theta})$ is given, the posterior of $\boldsymbol{\theta}$ can readily be obtained. The prior distribution of each parameter was assumed to be a uniform distribution. Thus, the shape of the posterior corresponds to that of the likelihood function in this study.

Since the present accumulation A_0 is not specified in the above sequential model, A_0 is treated as one of unspecified parameters and is included in $\boldsymbol{\theta}$. The parameter vector $\boldsymbol{\theta}$ also contains three hyper-parameters σ_ν, σ_η, and σ_w, which represent the variabilities in the model. These hyper-parameters are estimated so as to well explain the variability observed in the data. For example, if σ_ν is taken to be too small, the estimated age would not be well fit to the data. On the contrary, if σ_ν is taken

to be too large, large variations of the age ξ are allowed. Thus, the result could be sensitive to the noises contained in the data. The posterior given the data provides an appropriate value of σ_ν, which is large enough to achieve a good fit but not too large. The posterior of σ_w indicates a typical magnitude of the dispersion of δ^{18}O data from the trend of the δ^{18}O variation. We did not include σ_ε in $\boldsymbol{\theta}$, but σ_ε for each tie point was set at a fixed value shown in Table 2, which was determined according to Kawamura et al. (2007).

4 Estimation algorithm

In order to approximately obtain the conditional distributions $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z},\boldsymbol{\theta})$ and $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z})$, we employ a non-Gaussian algorithm called the PMCMC method (Andrieu et al., 2010), which is a hybrid method combining the SMC method and the Markov chain Monte Carlo (MCMC) method. In this hybrid method, the posterior distributions for the uncertain parameters in the model are computed using the standard MCMC with the exception that the likelihood of the parameters is estimated using the SMC method. Meanwhile, the age–depth relationship is estimated by performing many repetitions of the SMC procedure under iterations of the MCMC. The SMC can be used only for obtaining $p(x_{0:Z}|y_{1:Z},\theta)$ under a given θ, but it can not be used for obtaining $p(\theta|y_{1:Z})$. In principle, the MCMC could be used for obtaining any probability distribution including $p(x_{0:Z}|y_{1:Z},\theta)$, $p(\theta|y_{1:Z})$, and $p(x_{0:Z}|y_{1:Z})$. However, it would require prohibitive computational cost for high dimensional problems. In practice, the MCMC is not applicable to obtain a high dimensional distribution like $p(x_{0:Z}|y_{1:Z},\theta)$ and $p(x_{0:Z}|y_{1:Z})$. Combining the SMC and the MCMC, we can obtain $p(x_{0:Z}|y_{1:Z},\theta)$, $p(\theta|y_{1:Z})$, and $p(x_{0:Z}|y_{1:Z})$ with acceptable computational cost. In the following, we first present the SMC method on which the PMCMC method is based. We then describe the PMCMC method and explain how approximations of $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z},\boldsymbol{\theta})$ and $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z})$ can be obtained.

4.1 Sequential Monte Carlo method

The SMC method, which is sometimes referred to as the particle filter/smoother in time-series analysis (Gordon et al., 1993; Kitagawa, 1996; Doucet et al., 2001), is used for sampling from the conditional distribution $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z},\boldsymbol{\theta})$. The SMC method approximates the density function $p(\boldsymbol{x}_{0:z-1}|\boldsymbol{y}_{1:z-1})$ by a set of particles $\{\boldsymbol{x}^{(i)}_{0:z-1|z-1}\}$ as

$$p(\boldsymbol{x}_{0:z-1}|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta}) \approx \frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{x}_{0:z-1}-\boldsymbol{x}^{(i)}_{0:z-1|z-1}), \tag{26}$$

where $\delta(\cdot)$ denotes the Dirac delta function, N denotes the number of samples, and $\boldsymbol{x}^{(i)}_{0:z-1|z-1}$ denotes the i-th sample from $p(\boldsymbol{x}_{0:z-1}|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta})$. Drawing a particle $\boldsymbol{x}^{(i)}_{z|z-1}$ according to

$$\boldsymbol{x}^{(i)}_{z|z-1} \sim p(\boldsymbol{x}_z|\boldsymbol{x}_{z-1}=\boldsymbol{x}^{(i)}_{z-1|z-1},\boldsymbol{\theta}), \tag{27}$$

the set of particles $\{\boldsymbol{x}^{(i)}_{0:z|z-1}\}$ provides an approximation of $p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta})$:

$$p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta}) \approx \frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{x}_{0:z}-\boldsymbol{x}^{(i)}_{0:z|z-1}). \tag{28}$$

An approximation of the distribution conditioned by the observation $\boldsymbol{y}_z$ at z can be obtained using the importance sampling scheme (e.g., Liu, 2001; Robert and Casella, 2004):

$$\begin{aligned} p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z},\boldsymbol{\theta}) &= \frac{p(\boldsymbol{y}_z|\boldsymbol{x}_z,\boldsymbol{\theta})\,p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta})}{p(\boldsymbol{y}_z|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta})} \\ &\approx \sum_{i=1}^{N}\beta_z^{(i)}\delta(\boldsymbol{x}_{0:z}-\boldsymbol{x}^{(i)}_{0:z|z-1}). \end{aligned} \tag{29}$$

The weight $\beta_z^{(i)}$ for each i is defined as

$$\beta_z^{(i)} = \frac{p(\boldsymbol{y}_z|\boldsymbol{x}^{(i)}_{z|z-1},\boldsymbol{\theta})}{\sum_{i=1}^{N}p(\boldsymbol{y}_z|\boldsymbol{x}^{(i)}_{z|z-1},\boldsymbol{\theta})}, \tag{30}$$

where $p(\boldsymbol{y}_z|\boldsymbol{x}^{(i)}_{z|z-1},\boldsymbol{\theta})$ is the likelihood of the particle $\boldsymbol{x}^{(i)}_{z|z-1}$ that indicates how well $\boldsymbol{x}^{(i)}_{z|z-1}$ explains the observation $\boldsymbol{y}_z$ at z.

Equation (29) indicates that $p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z},\boldsymbol{\theta})$ can be approximated by weighting the particles $\{\boldsymbol{x}^{(i)}_{0:z|z-1}\}$. However, the weights are usually highly unbalanced and many of the particles have only negligible weights. Because particles with negligible weights no longer contribute to the estimation, this destroys the efficiency of the approximation. In order to resolve the imbalance in the weights, a new set of N particles $\{\boldsymbol{x}^{(i)}_{0:z|z}\}$ is obtained by resampling the original particles $\{\boldsymbol{x}^{(i)}_{0:z|z-1}\}$ such that each $\boldsymbol{x}^{(i)}_{0:z|z-1}$ is drawn with a probability of $\beta_z^{(i)}$ (see Nakano et al., 2007; van Leeuwen, 2009). After resampling, the original particles in $\{\boldsymbol{x}^{(i)}_{0:z|z-1}\}$ that have low weights are removed, and those that have high weights are duplicated. The number of the duplicates of $\boldsymbol{x}^{(i)}_{0:z|z-1}$, $n_z^{(i)}$, becomes approximately equal to $N\beta_z^{(i)}$. The newly generated particles then provide an approximation of $p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z},\boldsymbol{\theta})$ as follows:

$$\begin{aligned} p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z},\boldsymbol{\theta}) &\approx \sum_{i=1}^{N}\beta_z^{(i)}\delta\left(\boldsymbol{x}_{0:z}-\boldsymbol{x}^{(i)}_{0:z|z-1}\right) \\ &\approx \sum_{i=1}^{N}\frac{n_z^{(i)}}{N}\delta\left(\boldsymbol{x}_{0:z}-\boldsymbol{x}^{(i)}_{0:z|z-1}\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\delta\left(\boldsymbol{x}_{0:z}-\boldsymbol{x}^{(i)}_{0:z|z}\right). \end{aligned} \tag{31}$$

Applying the procedure from Eq. (26) to Eq. (31) recursively up to $z=Z$, we obtain samples from the conditional distribution $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z},\boldsymbol{\theta})$. If only the marginal distribution $p(\boldsymbol{x}_z|\boldsymbol{y}_{1:z},\boldsymbol{\theta})$, where $\boldsymbol{x}_{0:z-1}$ is marginalized out, is of interest, it is not necessary to keep the whole sequence of $\boldsymbol{x}^{(i)}_{0:z|z}$ for each particle; instead, at each iteration, it is sufficient to keep only the element $\boldsymbol{x}^{(i)}_{z|z}$ and discard the remaining $\boldsymbol{x}^{(i)}_{1:z-1}$

4.2 Particle Markov chain Monte Carlo method

An approximation of the marginal likelihood $p(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})$ in Eq. (24) can be calculated using the SMC algorithm (Kitagawa, 1996). If we decompose $p(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})$ as

$$\begin{aligned} p(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta}) &= p(\boldsymbol{y}_{1:Z-1}|\boldsymbol{\theta})\, p(\boldsymbol{y}_Z|\boldsymbol{y}_{1:Z-1},\boldsymbol{\theta}) \\ &= p(\boldsymbol{y}_1|\boldsymbol{\theta}) \prod_{z=2}^{Z} p(\boldsymbol{y}_z|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta}), \end{aligned} \tag{32}$$

we can obtain $p(\boldsymbol{y}_z|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta})$ for each z, from the following equation:

$$\begin{aligned} &p(\boldsymbol{y}_z|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta}) \\ &= \int p(\boldsymbol{y}_z|\boldsymbol{x}_z,\boldsymbol{\theta})\, p(\boldsymbol{x}_z|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta})\, d\boldsymbol{x}_z \\ &= \int p(\boldsymbol{y}_z|\boldsymbol{x}_z,\boldsymbol{\theta})\, p(\boldsymbol{x}_{0:z}|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta})\, d\boldsymbol{x}_{0:z} \end{aligned} \tag{33}$$

Since samples from $p(\boldsymbol{x}_{0:z-1}|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta})$ can be obtained by the SMC, a Monte Carlo approximation of the integral in Eq. (33) can be obtained as follows:

$$\begin{aligned} &p(\boldsymbol{y}_z|\boldsymbol{y}_{1:z-1},\boldsymbol{\theta}) \\ &\approx \frac{1}{N}\sum_{i=1}^{N} \int p(\boldsymbol{y}_z|\boldsymbol{x}_z,\boldsymbol{\theta})\, \delta\left(\boldsymbol{x}_{0:z} - \boldsymbol{x}_{0:z|z-1}^{(i)}\right) d\boldsymbol{x}_{0:z} \\ &= \frac{1}{N}\sum_{i=1}^{N} p(\boldsymbol{y}_z|\boldsymbol{x}_{0:z|z-1}^{(i)},\boldsymbol{\theta}), \end{aligned} \tag{34}$$

where we used Eq. (22). We can then approximate the logarithm of $p(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})$:

$$\log \hat{p}(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta}) = \sum_{z=1}^{Z} \log\left[\frac{1}{N}\sum_{i=1}^{N} p(\boldsymbol{y}_z|\boldsymbol{x}_{0:z|z-1}^{(i)},\boldsymbol{\theta})\right], \tag{35}$$

and an approximation of the posterior $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z})$ in Eq. (24) can accordingly obtained. As a matter of fact, however, the approximation of the log-likelihood given in Eq. (35) is too sensitive to the parameter $\boldsymbol{\theta}$ because of the large amount of the δ^{18}O data. We then introduce the following relaxation in this study:

$$\log \hat{p}(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta}) = \sum_{z=1}^{Z} \log\left[\frac{1}{N}\sum_{i=1}^{N} p(\delta^{18}\mathrm{O}_z|\boldsymbol{x}_{0:z|z-1}^{(i)},\boldsymbol{\theta})^{\lambda} p(\tau_{k_z}|\boldsymbol{x}_{0:z|z-1}^{(i)},\boldsymbol{\theta})\right], \tag{36}$$

where we set $\lambda = 1/5$.

Using the Monte Carlo approximation of the marginal likelihood $\hat{p}(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})$, we can obtain an approximation of the marginal posterior distribution of $\boldsymbol{\theta}$ using the MCMC, which sequentially produces samples that obey the target distribution. This is the basic idea of the PMCMC method. There are some variants of the PMCMC method such as the particle Gibbs with ancestor sampling (Lindsten et al., 2014). In this study, we employ the Metropolis method to obtain an approximation of

$p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z})$ because of the ease of implementation. In the Metropolis method, at the k-th iteration, a proposal sample $\boldsymbol{\theta}^*$ is drawn from the proposal density $q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)})$, which is conditioned by the sample $\boldsymbol{\theta}^{(k-1)}$, which was obtained at the previous iteration:

$$\boldsymbol{\theta}^* \sim q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k-1)}). \tag{37}$$

In this paper, we use a zero-mean Gaussian distribution with a fixed variance for each element of $\boldsymbol{\theta}$ in order to satisfy

$$q(\boldsymbol{\theta}|\boldsymbol{\theta}') = q(\boldsymbol{\theta}'|\boldsymbol{\theta}) \tag{38}$$

for any $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$. The variance of $q(\boldsymbol{\theta}|\boldsymbol{\theta}')$ was tuned by preliminary runs. In obtaining the final results, the variances were set at $0.05, 0.1, 0.1, 0.001, 0.2, 0.01, 5.0, 0.0002, 0.005$ for the parameters $A_0, a, b, \mu, p, s, \sigma_\nu, \sigma_\eta, \sigma_w$, respectively. The proposal sample $\boldsymbol{\theta}^*$ is accepted with the following probability:

$$\min\left(1, \frac{\hat{p}(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta}^*)\,p(\boldsymbol{\theta}^*)}{\hat{p}(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta}^{(k-1)})\,p(\boldsymbol{\theta}^{(k-1)})}\right), \tag{39}$$

where $\hat{p}(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})$ is an approximation of the marginal likelihood obtained by the SMC. If $\boldsymbol{\theta}^*$ is accepted, we set $\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^*$; otherwise, we set $\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)}$ and thus $\hat{p}(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta}^{(k)}) = \hat{p}(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta}^{(k-1)})$. Using $\boldsymbol{\theta}^{(k)}$, the proposal sample at the next iteration can be obtained according to Eq. (37). Iterating the above procedure generates a large number of samples that obey the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z})$.

In the above algorithm, an approximated value of the marginal likelihood $p(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})$ is computed using the SMC method at each iteration of the Metropolis method. It should be noted that Eq. (33) can be modified as follows:

$$\begin{aligned} &p(\boldsymbol{y}_z|\boldsymbol{y}_{1:z-1}, \boldsymbol{\theta}) \\ &= \int p(\boldsymbol{y}_z|\boldsymbol{x}_z, \boldsymbol{\theta})\, p(\boldsymbol{x}_z|\boldsymbol{x}_{z-1}, \boldsymbol{\theta})\, p(\boldsymbol{x}_{z-1}|\boldsymbol{y}_{1:z-1}, \boldsymbol{\theta})\, d\boldsymbol{x}_{z-1}\, d\boldsymbol{x}_z. \end{aligned} \tag{40}$$

Thus, in calculating $p(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})$ in Eq. (32), it is not necessary to consider the joint distribution of the sequence $\boldsymbol{x}_{0:Z}$; it is sufficient to consider the marginal distribution $p(\boldsymbol{x}_z|\boldsymbol{y}_{1:z}, \boldsymbol{\theta})$ for each z. As mentioned above, sampling from $p(\boldsymbol{x}_z|\boldsymbol{y}_{1:z}, \boldsymbol{\theta})$ can be achieved when discarding $\boldsymbol{x}_{1:z-1}$. This greatly reduces the computational cost because it can skip some procedures for handling the whole sequence of 2510 time steps ($Z = 2510$ in this paper) for 5000 particles. In addition, the memory cost is also remarkably reduced. We then discard $\boldsymbol{x}_{1:z-1}$ in order to obtain an approximation of $p(\boldsymbol{y}_{1:Z}|\boldsymbol{\theta})$ at each iteration of the Metropolis method.

As mentioned in Section 3, if we retain the samples for the whole sequence $\boldsymbol{x}_{0:Z}$ from a run of the SMC with a given $\boldsymbol{\theta}$, we obtain samples from $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z}, \boldsymbol{\theta})$. The Metropolis procedure sequentially generates a large number of samples that obey the marginal posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z})$. By combining the SMC samples with various $\boldsymbol{\theta}$ values that obey $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z})$, we can obtain the

samples representing the marginal posterior distribution $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z})$ where $\boldsymbol{\theta}$ is marginalized out according to Eq. (21). If samples that obey $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z})$ are obtained in advance, the sampling procedures from $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z}, \boldsymbol{\theta})$ for various $\boldsymbol{\theta}$ can be performed in parallel, and an approximation of the marginal posterior distribution $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z})$ can be obtained efficiently.

5 Result

Following a burn-in period, we performed 250000 iterations of the Metropolis sampling, and we retained a sample every fifth iteration. We thus drew 50000 samples from the marginal posterior distribution of $\boldsymbol{\theta}$, $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:Z})$. For each run of the SMC, 5000 particles were used to obtain samples from $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z}, \boldsymbol{\theta}^{(k)})$.

Figure 2 shows the marginal histograms for the estimated posterior distribution for each parameter. The posterior mean and standard deviation of the accumulation at the surface $A_0 (= A(0))$ were 2.78 and 0.19 [cm/year], respectively. This result is in good agreement with the measurement by Kameda et al. (2008), who reported the surface mass balance at Dome Fuji to be $27.3 \pm 1.5 \mathrm{kg}/(\mathrm{m}^{-2} \cdot \mathrm{yr})$. The maxima of the posterior distributions for μ and s were estimated to be near zero. This result is similar to that obtained in a previous study that used the Metropolis-Hastings method (Parrenin et al., 2007), although this result estimated a smaller uncertainty for μ and a larger uncertainty for s. In the result by Parrenin et al. (2007), the posterior of p peaks around 3, and another peak was suggested around $p = 2$. On the other hand, the results obtained in this study suggest that the posterior of p peaks around 3, and it is not clear whether there is another mode. It should be noted that these two results were based on different modeling for the accumulation rate. In addition, the setting of the thinning factor in this study is different from that used by Parrenin et al.. Thus, it should not be expected that they would necessarily provide similar results.

In the posterior distribution, some of the parameters correlate with each other. Figure 3 shows the two-dimensional histogram indicating the marginal posterior distribution of a and b (a), the marginal posterior distribution of a and σ_η (b), and the marginal posterior distribution of b and σ_η (c). Close correlations among the three parameters a, b, and σ_η are observed in this posterior distribution. These three parameters are related with the accumulation rate and $\delta^{18}\mathrm{O}$ data. Thus, the accuracy of the estimation for these three parameters could be much improved if any of the three parameters were effectively constrained.

Figure 4 shows the estimated age as a function of depth. The red solid line indicates the median of the posterior distribution and the red dotted lines indicate the 10th and 90th percentiles of the posterior distribution. The result by Parrenin et al. (2007) is also indicated with a grey line for reference. The black crosses in this figure indicate the tie points used for the estimation. In order to verify the convergence of the SMC sampling, we repeated sampling from the marginal posterior distribution $p(\boldsymbol{x}_{0:Z}|\boldsymbol{y}_{1:Z})$ five times with different seeds and confirmed that there were no apparent

differences between the results of the five trials. (The figures shown in this paper show the result of one of the five trials.) Thus, the estimate shown in Figure 4 is sufficiently converged. The SMC method often suffers from the degeneracy problem, especially when the number of steps is large. In the PMCMC method, this degeneracy problem is overcome by collecting a large number of the SMC samples that are obtained by iterations of the Metropolis method. In Figure 4, it is difficult to discriminate the difference of the 10th and 90th percentiles from the median because the width of the posterior distribution is much smaller than the range of Figure 4. In order to make the posterior width visible, the difference of the 10th and 90th percentiles of the posterior distribution from the median of the posterior is shown with red dotted lines in Figure 5. A black cross shows the difference of each age marker from the median of the posterior. The uncertainty of the age is minimized at each tie point where the age is known with high accuracy. Note that the age at each tie point has some uncertainty. Accordingly, the uncertainty can not be reduced to zero even at a tie point. In Figure 5, the difference of the estimate by Parrenin et al. (2007) from the median of the posterior obtained by the proposed method is also shown with a grey line. The grey line tends to deviate from the zero line more highly than the black crosses. This means that the estimate with the proposed method fit to the tie points more closely than the estimate by Parrenin et al., although the difference between the two results is more than 3000 years at largest.

Figure 6 indicates the estimated thinning factor as a function of depth. Again, the red solid line indicates the median of the posterior distribution and the red dotted lines indicate the 10th and 90th percentiles of the posterior distribution. The estimate by Parrenin et al. (2007) is also plotted with a grey line. Since $\Theta = 1$ at the surface by definition, the width of the posterior distribution is almost zero near the surface, and the uncertainty becomes larger in the deeper core. The profile of the thinning factor seems to be different between the red solid line and the grey line. This difference would be caused by the assumption of a constant ice thickness. While Parrenin et al. (2007) allowed the variation of the ice thickness H, we assume that the ice thickness is constant in this study. It would be a useful future work to examine the effect of the assumption of a constant H. Figure 7 shows the estimated accumulation rate as a function of depth. As in Figure 4, the red solid line indicates the median of the posterior distribution, and the red dotted lines indicate the 10th and 90th percentiles of the posterior distribution. The difference between the 10th and 90th percentiles, which corresponds to 80% confidence interval, is also shown with blue dotted line. In this way, we can estimate the age and related variables, and we can also obtain information about the credibility of these estimates. The accumulation rate can also be considered as a function of age. As shown in Figure 7, we have the posterior distribution of the accumulation rate given depth $p(A|z)$. The accumulation rate with respect to age is estimated after considering the uncertainty of age:

$$p(A|\xi) = \int p(A|z)\, p(z|\xi)\, dz \tag{41}$$

where we assume $p(z)$ to be a uniform distribution in obtaining $p(z|\xi)$:

$$p(z|\xi) = \frac{p(\xi|z)p(z)}{\int p(\xi|z)p(z)dz}. \tag{42}$$

Figure 8 shows the estimate of the accumulation rate with respect to age.

6 Discussion

In order to evaluate the robustness of the estimation, we tried the estimation without using the last five age markers at below 2400m depth. We estimated the parameters and the age–depth relationship from the other 20 age markers and δ^{18}O data. Figure 9 shows the histograms of the marginal posterior distribution for the nine parameters. The result without using the five age markers is indicated with a red line and the result with all the age markers, which is the same as the result in Figure 2, is indicated with the blue line. The posterior distribution obtained without a part of the age markers is basically similar to the result shown in Figure 2. However, the posterior distribution of the three parameters a, b, and σ_η are slightly different. Since σ_η is estimated larger when the five age markers were not used, this might indicates that the variations of the accumulation rate was noisier near the bottom. However, more careful evaluations would be required to resolve the reason.

In Figure 10, red lines indicate the estimate of the age as a function of depth without using the last five age markers. The solid red line indicates the median of the posterior distribution, and the dotted red lines indicate the 10th and 90th percentiles of the posterior. In this figure, the estimate of the age obtained using all the age markers is also plotted with grey lines. However, the difference between the two results is not visible except that the two results are slightly different near the bottom of the ice core. This suggests that our model well represents the actual processes related with the age–depth relationship of the ice core.

The accumulation rate as a function of age was also estimated without using the five age markers near the bottom of the ice core. Figure 11 shows the estimate of the accumulation rate as a function of age. The red lines indicate the estimate without using the last five age markers, and the grey line indicates the estimate using all the age markers for reference. Each solid line indicates the median of the posterior, and the 10th and 90th percentiles are shown with dotted lines. The difference is remarkable below the depth where the age is 300,000 years. However, the difference was mostly within the uncertainty between the 10th and 90th percentile. Thus, this difference near the bottom would be acceptable.

In the SMC part, if the number of particles N is large, each run of the sampling requires high computational cost. Thus, it is preferable that N should be as small as possible. We evaluated the performance with various smaller values of N. Figure 12 shows the result when 1000 particles were used for the SMC run, but 1250000 iterations of the Metropolis sampling were performed in order to obtain 250000 samples. The SMC method with $N = 1000$ requires only about $1/5$ of the computational cost with $N = 5000$. Thus, 1250000 iterations with $N = 1000$ have approximately

the same computational cost as 250000 iterations with $N = 5000$. The averaged acceptance rate of the MCMC sampling was 7.6% with $N = 1000$ while it was 26.4% with $N = 5000$. The estimate with $N = 1000$, as shown in the cyan histogram, is somewhat similar to the estimate with $N = 5000$. However, even though these two estimates have similar computational costs, the histogram of the estimate with $N = 1000$ appear to deviate from that of the estimate with $N = 5000$. As far as we examined, the estimate did not reach the convergence in Figure 12. If we use a parallel computer, the computation of the particle filter with $N = 5000$ could be conducted as fast as that with $N = 1000$. Therefore, there seem to be no reasons to reduce the number of the particles in the SMC part.

7 Concluding remarks

We have developed a technique for the dating of an ice core by combining information obtained from age markers at various depths and with a model describing the accumulation of snow and glaciological dynamics. This technique provides estimates of unspecified parameters in the model from their posterior distributions as calculated with the PMCMC method. In the PMCMC method, the marginal posterior distributions of the parameters are obtained using the Metropolis method; this is similar to other existing techniques (Parrenin et al., 2007), but here the likelihood of the set of parameters is estimated with the SMC method. The age as a function of depth can also be estimated from the marginal posterior distributions where the parameters are marginalized out. The marginal posterior distribution of age at each depth is obtained by collecting the SMC samples produced by many iterations of the Metropolis method. We applied this PMCMC method to the data of the ice core at Dome Fuji. The estimates of the age–depth relationship and the parameters were successfully obtained.

The main advantage of the proposed technique is that it can be applied to general nonlinear non-Gaussian situations. Since the relationship between accumulation rate and a temperature proxy is typically nonlinear, it is not necessarily justified to assume linearity and Gaussianity in dating of an ice core using a temperature proxy. The PMCMC method allows us to use various kinds of data which are expected to have nonlinear relationship with model variables. Another advantage is that the PMCMC method estimates model parameters simultaneously with the age as a function of depth. The uncertainty of age is therefore evaluated after taking into account the uncertainties in the model parameters. However, the proposed technique requires high computational cost because the SMC sampling is performed at each iteration of the Metropolis method. At present, it takes about 43 hours to complete 250000 iterations of the Metropolis sampling with 5000 particles for the SMC on a workstation with two Intel Xeon processors (12 cores for each processor; 2.7Ghz). The efficiency could be improved by using a better proposal distribution (e.g., Doucet et al., 2001). This problem should be addressed in the future.

There might be a room for improvement in the model for the accumulation rate described by Eq. (11). Eq. (11) represents the transition of the accumulation rate by a random walk model with a Gaussian perturbation. However, we could consider another model such as an auto-regressive model

for the transition and another distribution for the perturbation. There are a large number of choices for the model for the accumulation rate and the goodness could be evaluated using some metric such as Bayes factors. However, it would take much time to make the selection among a variety of choices. Thus, we will not explore a better model in this study.

This study used the $\delta^{18}\mathrm{O}$ data and tie points deduced from $\mathrm{O}_2/\mathrm{N}_2$ data to estimate the age–depth relationships. However, data from other sources could also be used to improve the accuracy of the estimates. For example, deuterium-excess data have been used to estimate temperatures (Uemura et al., 2012), and this could be used for improving the accuracy of the accumulation rate. Some recent studies have provided simultaneous estimates of the age as a function of depth at multiple sites (e.g., Lemieux-Dudon et al., 2010; Veres et al., 2013). The extension of the SMC approach such that information at multiple sites could be combined would be a useful area of future work.

Acknowledgements. This work was conducted under the project "Exploration for Seeds of Integrated Research," supported by the Transdisciplinary Research Integration Center, Research Organization of Information and Systems.

References

Andrieu, C., Doucet, A., and Holenstein, R.: Particle Markov chain Monte Carlo methods, J. Roy. Statist. Soc. B, 72, 269–342, 2010.

Doucet, A., de Freitas, N., and Gordon, N., eds.: Sequential Monte Carlo methods in practice, Springer-Verlag, New York, 2001.

Dreyfus, G. B., Parrenin, F., Lemieux-Dudon, B., Durand, G., Masson-Delmotte, V., Jouzel, J., Barnola, J.-M., Panno, L., Spahni, R., Tisserand, A., Siegenthaler, U., and Leuenberger, M.: Anomalous flow below 2700m in the EPICA Dome C ice core detected using δ^{18}O of atmospheric oxygen measurements, Clim. Past, 3, 341–353, 2007.

Freitag, J., Kipfstuhl, S., and Laepple, T.: Core-scale radioscopic imaging: a new method reveals density–calcium link in Antarctic firn, J. Glaciology, 59, doi:10.3189/2013JoG13J028, 1009–1014, 2013.

Gordon, N. J., Salmond, D. J., and Smith, A. F. M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation, IEE Proceedings F, 140, 107, 1993.

Kameda, T., Motoyama, H., Fujita, S., and Takahashi, S.: Temporal and spatial variability of surface mass balance at Dome Fuji, East Antarctica, by the stake method from 1995 to 2006, J. Glaciology, 54, 107–116, 2008.

Kawamura, K., Parrenin, F., Lisiecki, L., Uemura, R., Vimeux, F., Severinghaus, J. P., Hutterli, M. A., Nakazawa, T., Aoki, S., Jouzel, J., Raymo, M. E., Matsumoto, K., Nakata, H., Motoyama, H., Fujita, S., Goto-Azuma, K., Fujii, Y., and Watanabe, O.: Northern Hemisphere forcing of climatic cycles in Antarctica over the past 360,000 years, Nature, 448, doi:10.1038/nature06 015, 912–916, 2007.

Kitagawa, G.: Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, J. Comp. Graph. Statist., 5, 1, 1996.

Klauenberg, K., Blackwell, P. G., Buck, C. E., Mulvaney, R., Röthlisberger, R., and Wolff, E. W.: Bayesian glaciological modelling to quantify uncertainties in ice core chronologies, Quaternary Sci. Rev., 30, 2961–2975, 2011.

Lemieux-Dudon, B., Parrenin, F., and Blayo, E.: A probabilistic method to construct an optimal ice core chronology for ice cores, in: Proceedings of the 2nd International Workshop on Physics of Ice Core Records (PICR-2), edited by Hondoh, T., pp. 233–245, Institute of Low Temperature Science, Hokkaido University, 2009.

Lemieux-Dudon, B., Blayo, E., Petit, J. R., Waelbroeck, C., Svenson, A., Ritz, C., Barnola, J. M., Narcisi, B. M., and Parrenin, F.: Consistent dating for Antarctic and Greenland ice cores, Quaternary Sci. Rev., 29, 8–20, 2010.

Lindsten, F., Jordan, M. I., and Schön, T. B.: Particle Gibbs with ancestor sampling, J. Mach. Learn. Res., 15, 2145–2184, 2014.

Liu, J. S.: Monte Carlo strategies in scientific computing, Springer-Verlag, New York, 2001.

Lliboutry, L.: Local friction laws for glaciers: a critical review and new openings, J. Glaciology, 23, 67–95, 1979.

Nakano, S., Ueno, G., and Higuchi, T.: Merging particle filter for sequential data assimilation, Nonlin. Process. Geophys., 14, 395–408, 2007.

Parrenin, F., Waelbroeck, J. J. C., Ritz, C., and Barnola, J.-M.: Dating the Vostok ice core by an inverse method,

J. Geophys. Res., 106, 31 837–31 851, 2001.

Parrenin, F., Hindmarsh, R. C. A., and Rémy, F.: Analytical solutions for the effect of topography, accumulation rate and lateral flow divergence on isochrone layer geometry, J. Glaciology, 52, 191–202, 2006.

Parrenin, F., Dreyfus, G., Durand, G., Fujita, S., Gagliardini, O., Gillet, F., Jouzel, J., Kawamura, K., Lhomme, N., Masson-Delmotte, V., Ritz, C., Schwander, J., Shoji, H., Uemura, R., Watanabe, O., , and Yoshida, N.: 1-D-ice flow modelling at EPICA Dome C and Dome Fuji, East Antarctica, Clim. Past, 3, 243–259, 2007.

Robert, C. P. and Casella, G.: Monte Carlo statistical methods, Second Edition, Springer Science+Business Media Inc., New York, 2004.

Uemura, R., Masson-Delmotte, V., Jouzel, J., Landais, A., Motoyama, H., and Stenni, B.: Ranges of moisture-source temperature estimated from Antarctica ice cores stable isotope records over glacial-interglacial cycles, Clim. Past, 8, 1109–1125, 2012.

van Leeuwen, P. J.: Particle filtering in geophysical systems, Mon. Wea. Rev., 137, 4089–4114, 2009.

Veres, D., Bazin, L., Landais, A., Kele, T. M., Lemieux-Dudon, B., Parrenin, F., Martinerie, P., Blayo, E., Blunier, T., Capron, E., Chappellaz, J., Rasmussen, S. O., Severi, M., Svensson, A., Vinther, B., and Wolff, E. W.: The Antarctic ice core chronology (AICC2012): an optimized multi-parameter and multi-site dating approach for the last 120 thousand years, Clim. Past, 9, 1733–1748, 2013.

Watanabe, O., Jouzel, J., Johnsen, S., Parrenin, F., Shoji, H., and Yoshida, N.: Homogeneous climate variability across East Antarctica over the past three glacial cycles, Nature, 422, 509–512, 2003.

z	Vertical coordinate
ξ	Age
A	Accumulation rate
Θ	Thinning factor
$\boldsymbol{x}_z$	State at depth z ($\boldsymbol{x}_z = (\xi_z\ A_z)^T$)
$\boldsymbol{y}_z$	Observation at depth z ($\boldsymbol{y}_z = (\tau_{k_z}\ \delta^{18}\mathrm{O}_z)^T$)
$\boldsymbol{\theta}$	Parameter vector ($\boldsymbol{\theta} = (A_0\ a\ b\ \mu\ p\ s\ \sigma_\nu\ \sigma_\eta\ \sigma_w)^T$)

Table 1. Definition of the variables used in this paper.

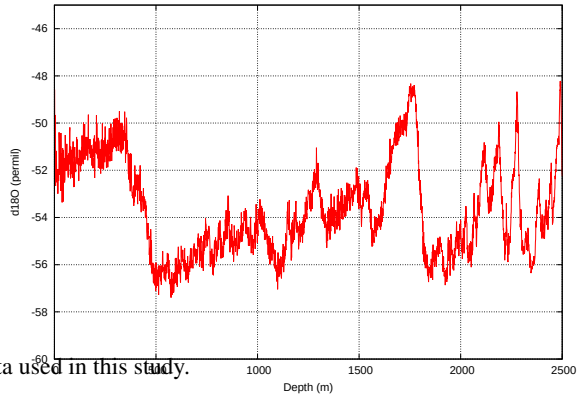

Fig. 1. δ^{18}O data used in this study.

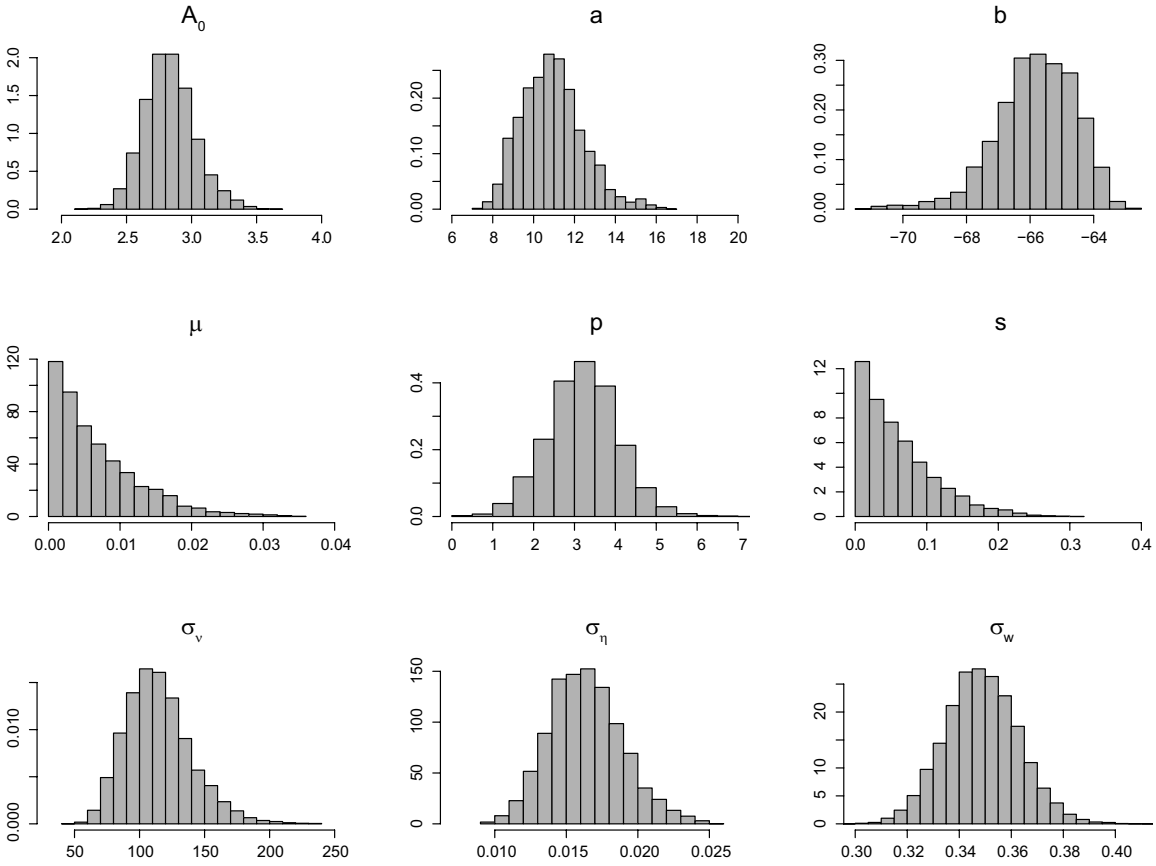

Fig. 2. Estimated marginal distributions of the posterior distributions for the nine parameters.

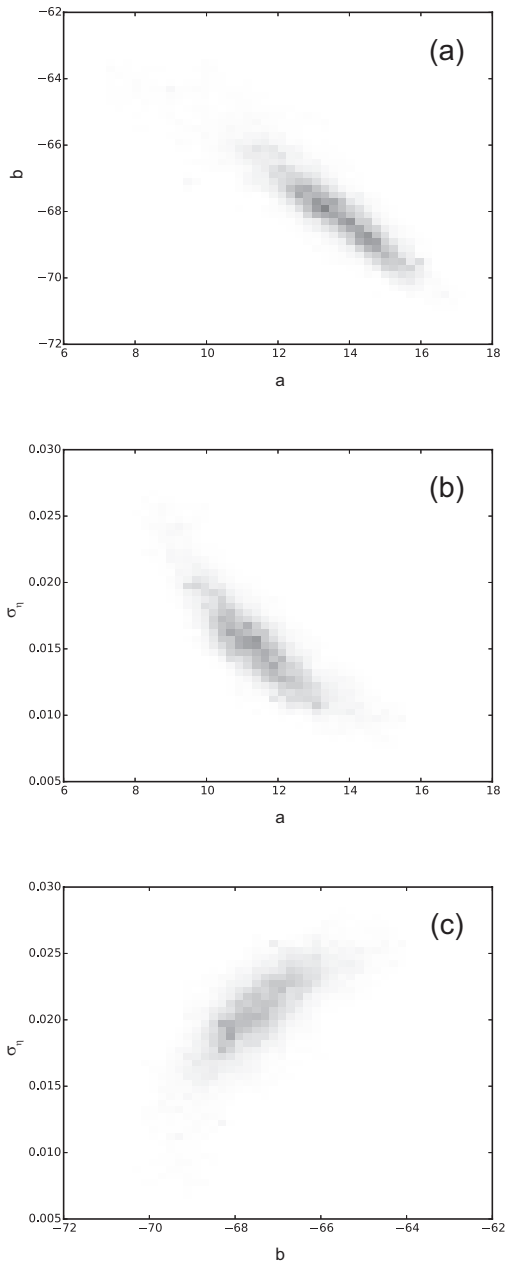

Fig. 3. Two dimensional histograms of the marginal posterior distribution of a and b (a), the marginal distribution of a and σ_η (b), and the marginal distribution of b and σ_η (c).

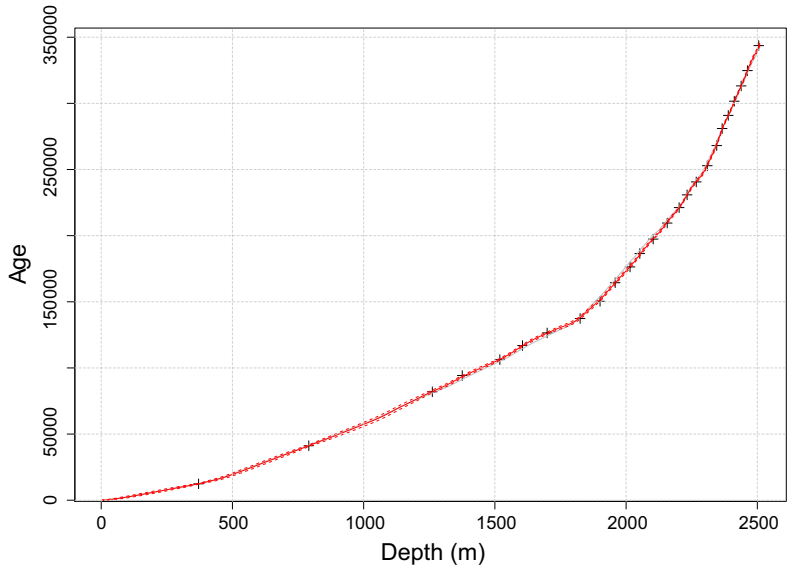

Fig. 4. Estimated age as a function of depth. The solid line indicates the median of the posterior distribution. The 10th and 90th percentiles of the posterior are indicated by red dotted lines. The black crosses indicates the age markers. The result obtained by Parrenin et al. (2007) is also plotted with a grey line.

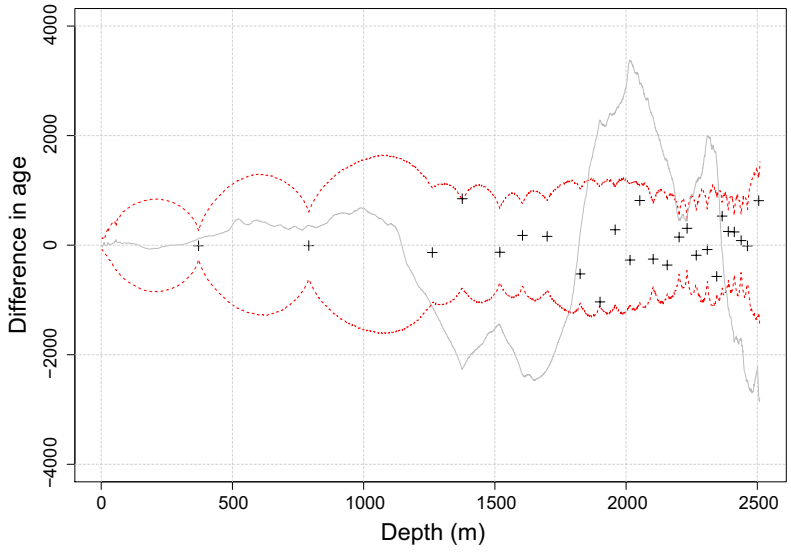

Fig. 5. Difference of the 10th and 90th percentiles of the posterior distribution from the median of the posterior (red dotted lines), difference of each age marker from the median of the posterior (black crosses), and difference of the estimate by Parrenin et al. (2007) from the median of the posterior obtained in this study (grey line).

Depth	Age	Uncertainty of the age ($2\sigma_\varepsilon$)
371.00	12390	400
791.00	41200	1000
1261.61	81973	2230
1375.67	94240	1410
1518.91	106263	1220
1605.27	116891	1490
1699.17	126469	1660
1824.80	137359	2040
1900.74	150368	2230
1958.31	164412	2550
2015.01	176353	2880
2052.23	186470	2770
2103.14	197394	1370
2156.67	209523	1980
2202.02	221211	890
2232.45	230836	780
2267.28	240633	1230
2309.35	252866	1160
2345.32	268105	1980
2366.01	280993	1600
2389.31	290909	1210
2412.25	301628	880
2438.37	313205	840
2462.36	324774	1110
2505.4	343673	2000

Table 2. The depth, the age, and the uncertainty of the age at each tie point.

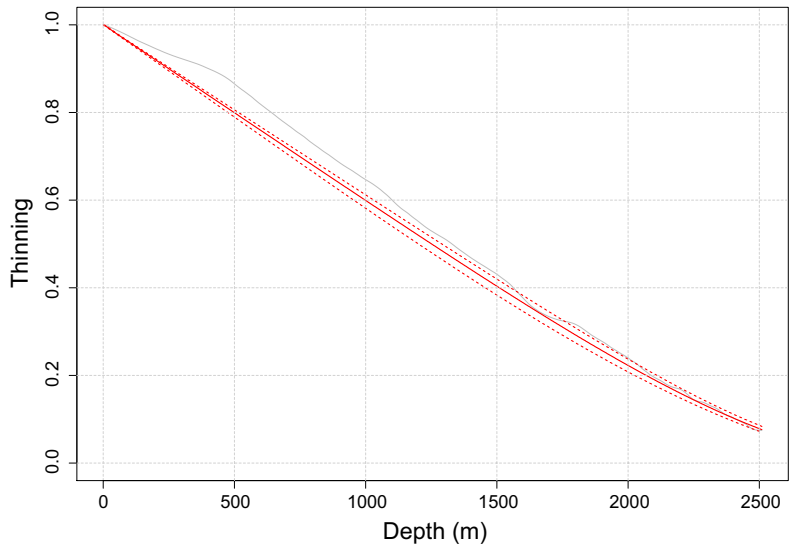

Fig. 6. Estimated thinning factor Θ as a function of depth. The median of the posterior is shown with a red solid line, the 10th and 90th percentiles are shown with red dotted lines, and the estimate by Parrenin et al. (2007) is shown with a grey line.

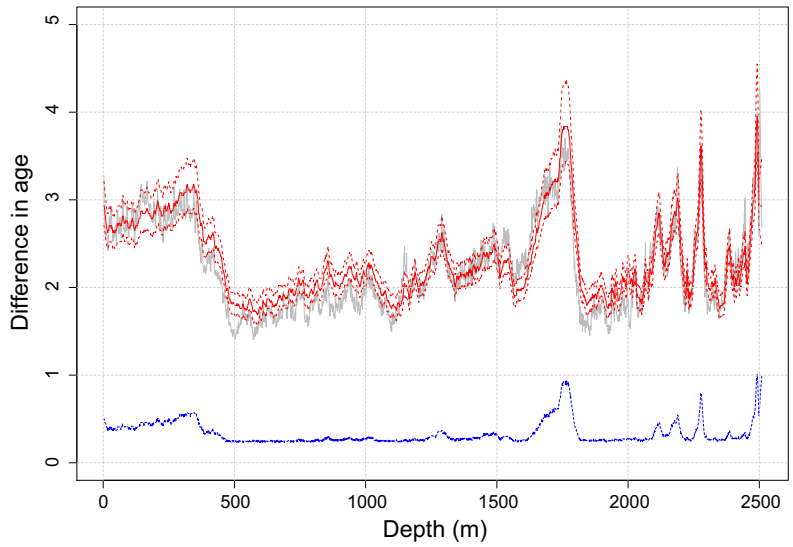

Fig. 7. Estimated accumulation rate as a function of depth. The median of the posterior is shown with a red solid line, the 10th and 90th percentiles are shown with red dotted lines, the difference between the 10th and 90th percentiles is shown with blue dotted line, and the estimate by Parrenin et al. (2007) is shown with a grey line.

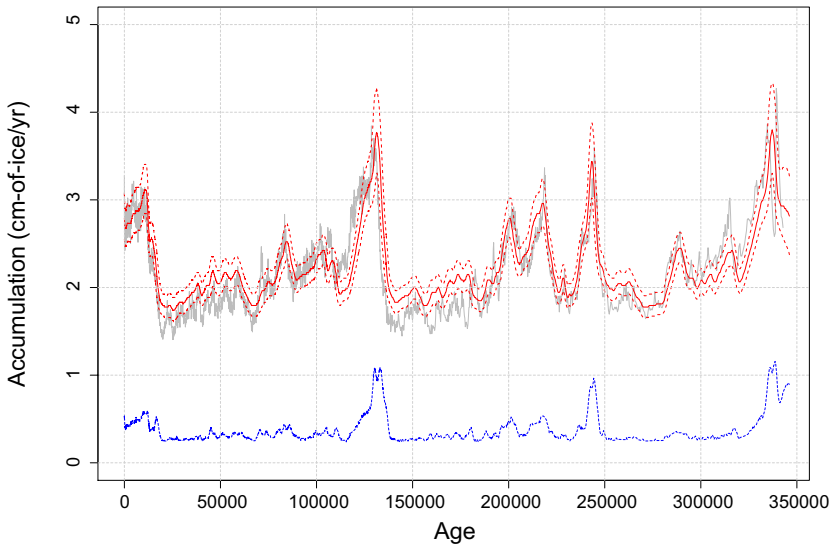

Fig. 8. Estimated accumulation rate as a function of age. The median of the posterior is shown with a red solid line, the 10th and 90th percentiles are shown with red dotted lines, the difference between the 10th and 90th percentiles is shown with blue dotted line, and the estimate by Parrenin et al. (2007) is shown with a grey line.

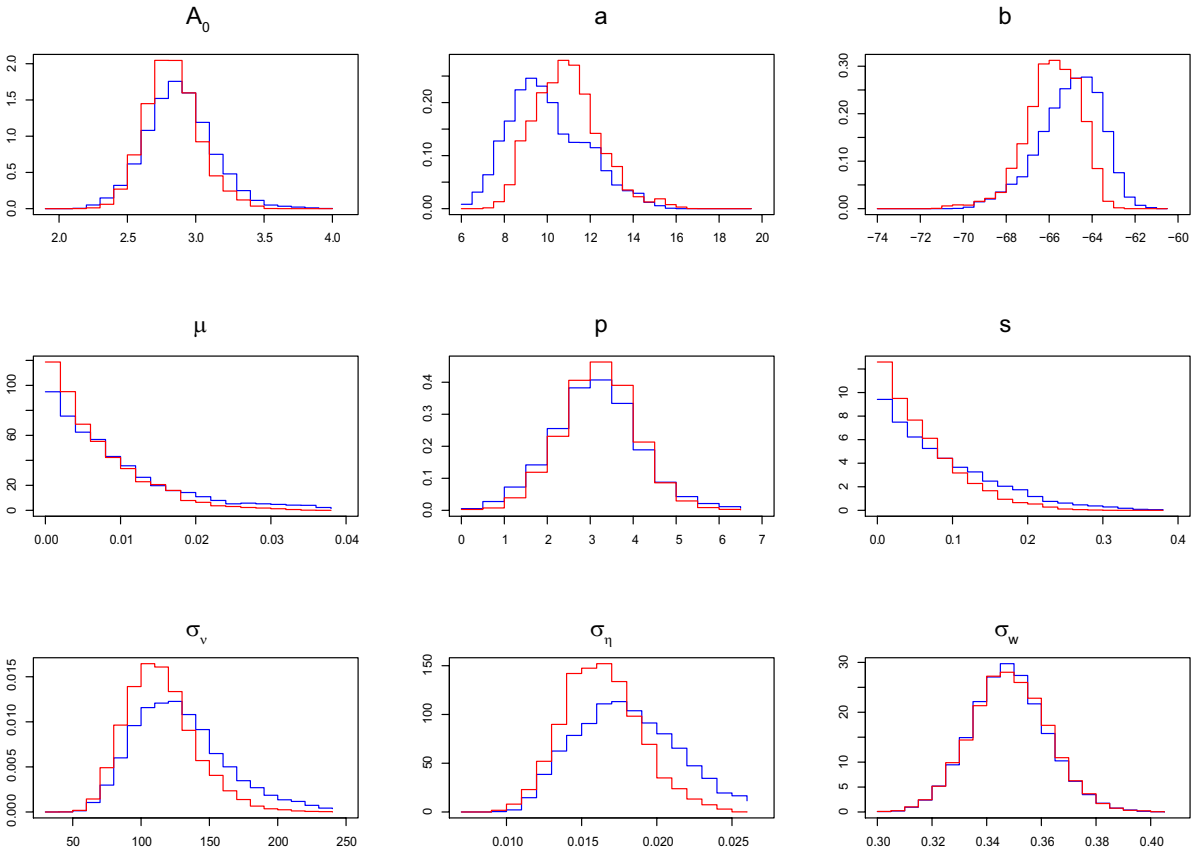

Fig. 9. Comparison of the estimated marginal distributions for the nine parameters between the result without using the last five age markers (red) and the result with all the age markers (blue).

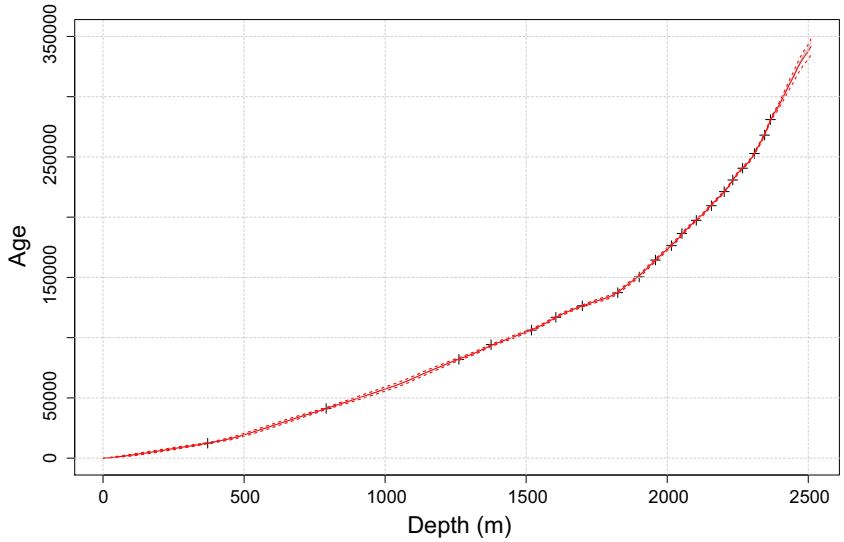

Fig. 10. Estimated age as a function of depth without using the last five agemarkers (red) and estimate using all the agemarkers (grey). Each solid line indicates the median of the posterior distribution. The 10th and 90th percentiles of the posterior are indicated by dotted lines.

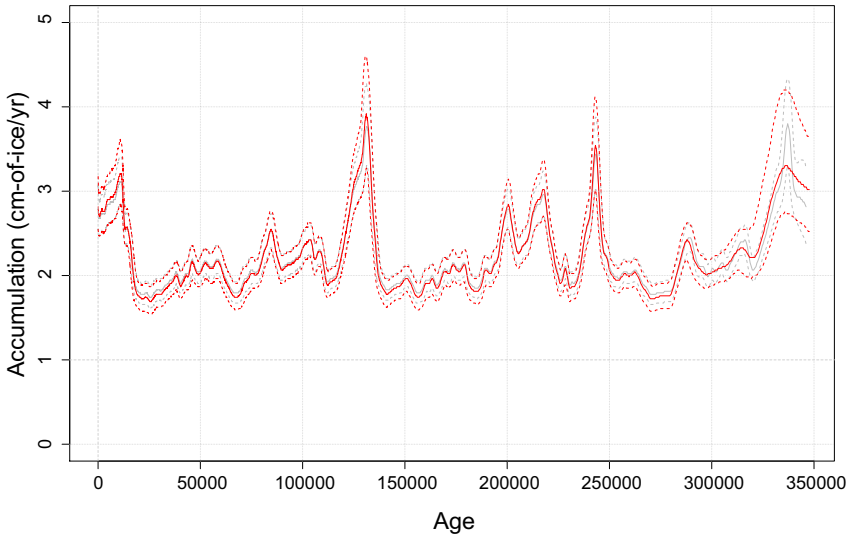

Fig. 11. Estimated accumulation rate as a function of age without using the last five agemarkers (red) and estimate using all the agemarkers (grey). Each solid line indicates the median of the posterior distribution. The 10th and 90th percentiles of the posterior are indicated by dotted lines.

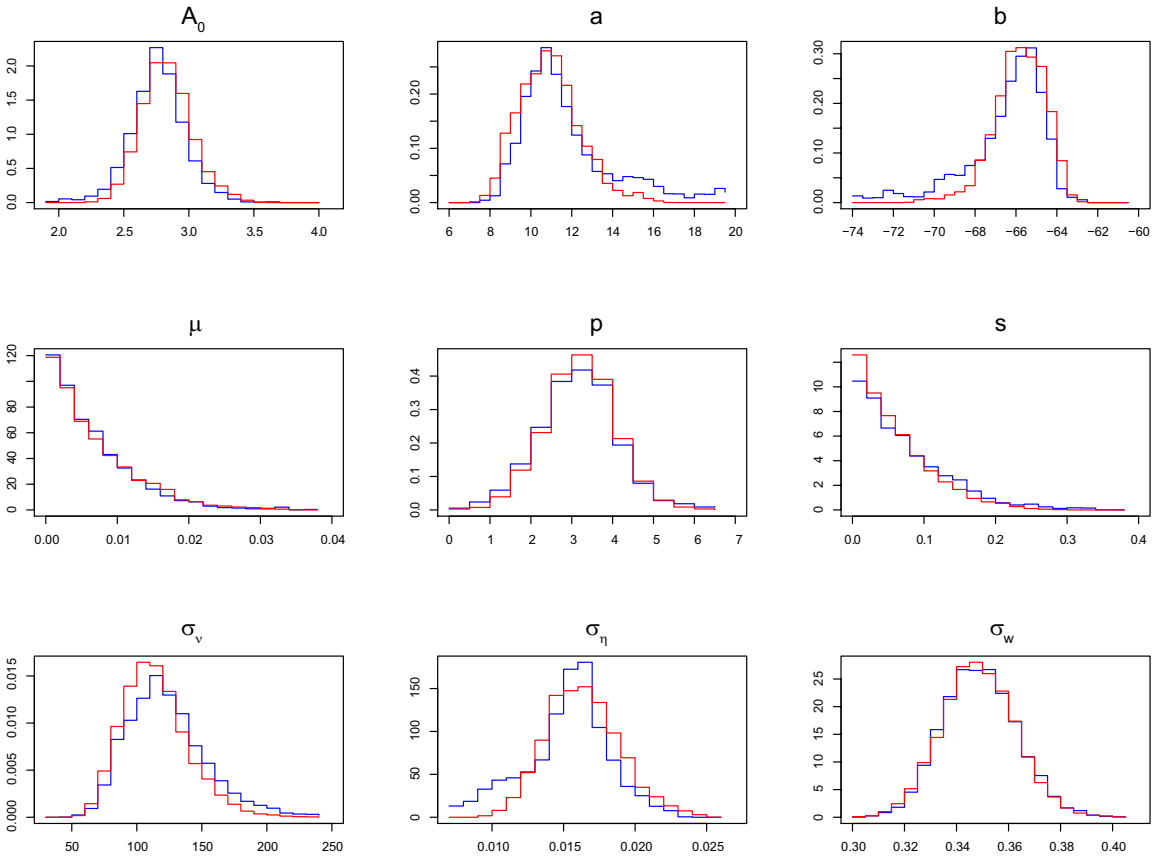

Fig. 12. Comparison of the estimated marginal distributions for the nine parameters between the result with 50000 iterations and 5000 particles (red line) and the result with 250000 iterations and 1000 particles (blue line).