

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

Shin'ya Nakano^{1,2}, Kazue Suzuki¹, Kenji Kawamura^{3,2}, Frédéric Parrenin⁴, and Tomoyuki Higuchi^{1,2}

¹The Institute of Statistical Mathematics, Tachikawa, 190–8562, Japan

²School of Multidisciplinary Science, SOKENDAI, Hayama, 240–0115, Japan.

³National Institute of Polar Research, Research Organization of Information and Systems, Tachikawa 190–8518, Japan

⁴Laboratoire 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 rate nor the thinning process is fully known, 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 by making use of age markers and $\delta^{18}\text{O}$ data. The age markers provides reliable age information at several depths. The data of $\delta^{18}\text{O}$ is used as a proxy of the temperature for estimating the accumulation rate. The estimation is achieved using the particle Markov chain Monte Carlo (PMCMC) method, which is a combination of the sequential Monte Carlo (SMC) method and the Markov chain Monte Carlo method. In this hybrid method, the posterior distributions for the parameters in the models for the accumulation and thinning process are computed using the Metropolis method, in which the likelihood is obtained with the SMC method, and 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 relationship between age and depth in the ice cores. Many of dating methods for determining the age–depth relationship rely on glaciological modeling. However, since the glaciological processes controlling the age–depth relationship are not fully known, it is essential to reduce uncertainty by incorporating various types of observational information into the glaciological model. In particular, it is important to effectively make use of the information of age markers, which provides significant constraints on the age–depth relationship. 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 age–depth relationship and to improve some parameters in their dating model by using $\delta^{18}\text{O}$ data. The uncertainty of the estimate was also evaluated in a Bayesian manner. However, their method was not designed to make use of the constraints of age markers when estimating 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-squares 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 state space model to apply a sequential Bayesian approach. The estimation is then achieved using the particle Markov chain Monte Carlo (PMCMC) method (Andrieu et al., 2010), which is a sequential Bayesian 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 provide a model of the age–depth relationship. In Section 3, the age–depth relationship is formulated in a framework of a state space model in order to apply PMCMC for the estimation of 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. The

discussion on the proposed method is presented in Section 6. Finally, a summary is presented in Section 7. For reference, symbols used in this paper are listed in Table 1.

60 2 Dating model

The age–depth relationship is determined by two processes. One is accumulation of snow at the site of the ice core and the other is thinning process due to long-term deformations within the ice sheet (e.g., Parrenin et al., 2001, 2007). In this section, a model for describing the age–depth relationship is introduced. Basic ideas about how to estimate the contributions from the snow accumulation and
65 thinning are also provided.

Denoting the annual rate of snow accumulation by $A(z)$ [m/year] and the thinning factor by $\Theta(z)$ (dimensionless), the relationship between age ξ [year] and depth from the surface z [m] is described by the following differential equation:

$$dz = A(z)\Theta(z) d\xi. \quad (1)$$

70 In this equation, the accumulation rate $A(z)$ is written as a function of depth. This means that $A(z)$ indicates the accumulation rate at the time when the ice at depth z was deposited. It would be more natural to consider the accumulation rate as a function of age ξ rather than depth z . In this study, however, we first consider the accumulation rate as a function of depth for the convenience of computation. The accumulation rate with respect to age is then estimated after considering the
75 uncertainty of age as described later. Equation (1) yields the age ξ in the following form:

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

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

In order to model the thinning factor $\Theta(z)$ in Eq. (2), we adopt the pseudo-steady hypothesis (Parrenin et al., 2006; Parrenin and Hindmarsh, 2007), which assumes steady geometry of the icesheet
80 and steady vertical profile for velocity. Assuming a pseudo-steady state, the thinning factor $\Theta(z)$ in Eq. (2) can be written using the vertical velocity U :

$$\Theta(z) = U(z)/U(0). \quad (3)$$

Rescaling z and U as

$$85 \zeta = \frac{H-z}{H}, \quad u(\zeta) = -\frac{U(z)}{H}, \quad (4)$$

Eq. (3) can be rewritten as follows:

$$\Theta(\zeta) = u(\zeta)/u(1). \quad (5)$$

In Eq. (4), H is the thickness of the ice sheet, which is constant in the pseudo-steady state. The variable ζ is a rescaled vertical coordinate which becomes 0 at the bottom and 1 at the surface, and u indicates the velocity in the ζ coordinate. We rewrite the rescaled vertical velocity $u(\zeta)$ in the following form (Parrenin et al., 2006):

$$u(\zeta) = u(0) + [u(1) - u(0)]\omega(\zeta). \quad (6)$$

where $\omega(\zeta)$ is a function satisfying $\omega(0) = 0$ and $\omega(1) = 1$. In the pseudo-steady state, the function $\omega(\zeta)$ is unchanged in time. In this study, $\omega(\zeta)$ is described by the Lliboutry equation (Lliboutry, 1979):

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

where s corresponds to the sliding ratio, which is the ratio of the basal horizontal velocity to the vertically averaged horizontal velocity. In the pseudo-steady state, the vertical velocity profile $u(\zeta)$ is assumed to be steady. Thus, the parameters p and s does not vary over time. Equation (7) derives from the Shallow Ice Approximation (SIA), and is based on a linearization of the temperature profile. Martin and Gudmundsson (2012) have shown that the Lliboutry equation is not appropriate for a steady dome. However, we can expect that the domes in central Antarctica are non-steady because the Raymond bumps have never been observed. We thus assume that the Lliboutry equation can still be used. Denoting the melting rate at the base of the ice sheet by m , A and m correspond to the vertical velocity at $\zeta = 1$ and that at $\zeta = 0$, respectively, under a pseudo-steady state. Equation (6) can thus be rewritten as follows:

$$u(\zeta) = -\frac{1}{H}[m + (A - m)\omega(\zeta)]. \quad (8)$$

Setting $\mu = m/(A - m)$, Eqs. (5) and (8) yields

$$\Theta(\zeta) = \frac{\omega(\zeta) + \mu}{1 + \mu}. \quad (9)$$

We assume μ is constant, which means the ratio m/A is constant. This assumption would be approximately justified because m is typically much smaller than A . 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 profile of the accumulation rate A . In this study, A is treated as unknown to be estimated. Since the accumulation rate is related to the Antarctic temperature, we can use proxies of the temperature for constraints when estimating the profile of A . As a proxy for the temperature, we used the $\delta^{18}\text{O}$ data taken at Dome Fuji (Watanabe et al., 2003), which is plotted in Figure 1. Since the vertical profile of the age ξ is associated with the profile of A , the information from the $\delta^{18}\text{O}$ data is also effective for improving the estimate of the age ξ .

At several depths, we can also use reliable age values given by age markers. 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 point was determined from the ACR-Holocene transition and the second one was determined from the Beryllium 10 peak. These two points were given by Parrenin et al. (2007). The subsequent 23 points were determined from the relationship between O_2/N_2 and the summer insolation (Kawamura et al., 2007). Both the $\delta^{18}O$ data and the tie points were considered when estimating the age ξ .

3 State space model

In this section, the age–depth relationship is formulated in a framework of a state space model on the basis of the model described in the previous section. The state space model represents the evolution of variables by a recurrence equation. The state space model provides a platform for the sequential Bayesian estimation using PMCMC, which will be explained in the next section.

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

$$\xi_{z+\Delta z} \approx \xi_z + \frac{\Delta z}{A_z \Theta_z} \quad (z = 0, \Delta z, 2\Delta z, \dots, Z - \Delta z), \quad (10)$$

where ξ_z denotes the age at depth z , and we denote the accumulation rate and the thinning factor in the interval from z to $z + \Delta z$ by A_z and Θ_z , respectively. At the surface ($z = 0$), ξ_0 is defined as zero. The depth at the bottom of the core is denoted by Z .

Equation (10) would contain an error due to the discretization of Eq. (2). In addition, since we can not accurately know A_z and Θ_z for each z , an estimate of the age–depth relationship might also be affected by errors in A_z and Θ_z . We represent these unspecified errors by ν_z . Equation (10) is thus modified as follows:

$$\xi_{z+\Delta z} = \xi_z + \frac{\Delta z}{A_z \Theta_z} + \nu_z \sqrt{\frac{\Delta z}{A_z \Theta_z}} \quad (z = 0, \Delta z, 2\Delta z, \dots, Z - \Delta z). \quad (11)$$

We assume that ν_z obeys the normal distributions $\mathcal{N}(0, \sigma_\nu^2)$, where we denote a normal distribution with mean μ and variance σ^2 by $\mathcal{N}(\mu, \sigma^2)$. We multiply ν_z by $\sqrt{\Delta z / (A_z \Theta_z)}$ in order that the variance of the unknown variation per year remain the same to the bottom of the core. In a pseudo-steady state, the thinning factor Θ_z can be obtained according to Eq. (9). However, Eq. (9) does not consider all of the effects governing the thinning process Θ_z ; for example, it omits the effect of impurities (Freitag et al., 2013). The errors in Θ due to such unspecified effects would also be considered by ν_z in Eq. (11).

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:

$$\log A_{z+\Delta z} = \log A_z + \eta_z \sqrt{\frac{\Delta z}{A_z \Theta_z}} \quad (z = 0, \Delta z, 2\Delta z, \dots, Z - \Delta z). \quad (12)$$

Note that in Eq. (12), the transition of A_z is described by using its logarithm in order to guarantee $A_z > 0$. The term η_z represents the (unknown) variation in the accumulation rate. We assume that η_z

155 obeys $\mathcal{N}(0, \sigma_\eta^2)$. We hereinafter assume $\Delta z = 1$ [m]. Equations (11) and (12) can thus be rewritten as follows:

$$\xi_{z+1} = \xi_z + \frac{1}{A_z \Theta_z} + \frac{\nu_z}{\sqrt{A_z \Theta_z}}, \quad (z = 0, \dots, Z-1) \quad (13)$$

$$\log A_{z+1} = \log A_z + \frac{\eta_z}{\sqrt{A_z \Theta_z}}, \quad (z = 0, \dots, Z-1). \quad (14)$$

160 Based on Eqs. (13) and (14), we introduce conditional probability density functions. Since we assumed ν_z and η_z obey $\mathcal{N}(0, \sigma_\nu^2)$ and $\mathcal{N}(0, \sigma_\eta^2)$, respectively, the conditional distribution of ξ_{z+1} given ξ_z and that of A_{z+1} given A_z for each z become

$$p(\xi_{z+1} | \xi_z, \boldsymbol{\theta}) = \mathcal{N}\left(\xi_z + \frac{1}{A_z \Theta_z}, \frac{\sigma_\nu^2}{A_z \Theta_z}\right), \quad (15)$$

$$165 \quad p(A_{z+1} | A_z, \boldsymbol{\theta}) = \log \mathcal{N}\left(A_z, \frac{\sigma_\eta^2}{A_z \Theta_z}\right), \quad (16)$$

respectively, 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.

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 $\delta^{18}\text{O}$ data. For the k -th tie point τ_k at depth z_k , we assume the following
170 relationship between τ_k and the modeled age ξ_{z_k} :

$$\tau_k = \xi_{z_k} + \varepsilon_k, \quad (17)$$

where ε_k is the discrepancy between the age at the tie point and the modeled age. We assume that ε_k obeys the normal distribution $\mathcal{N}(0, \sigma_\varepsilon^2)$. While we consider the uncertainty in the age of tie points, we assume there to be no uncertainty in the depths of tie points. This is because the depth
175 uncertainty would not make essential effects on the estimate of the age for each slice of the ice core. The effects of the depth uncertainty on the estimates of accumulation and thinning are also expected to be minor. because the accumulation and thinning are related with the increment of depth rather than the absolute depth from the surface. In this study, the uncertainty in the age increment is taken into account by ν_z in Eq. (13). This would compensate the possible effect of the depth uncertainty
180 on the estimates of accumulation rate and thinning factor.

The $\delta^{18}\text{O}$ data, which are associated with the accumulation rate, can be abundantly obtained from the ice core at Dome Fuji. Multiple data points for $\delta^{18}\text{O}$ are sometimes available within an interval of a single meter, and we used the mean $\delta^{18}\text{O}$ 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 $\delta^{18}\text{O}$ as follows:

$$185 \quad \delta^{18}\text{O}_z = a \log A_z + b + w_z; \quad (18)$$

this relation was also used by Klauenberg et al. (2011). We assume that w_z obeys the normal distribution $\mathcal{N}(0, \sigma_w^2)$. Although we assume the regression coefficients a and b do not depend on age,

it is not guaranteed that the accumulation rate and $\delta^{18}\text{O}$ 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 $\delta^{18}\text{O}$, a and b might change due to the variation of climatological conditions other than temperature. However, the uncertain variable η_z in Eq. (12) represents the variation of accumulation rate including not only the variation related with $\delta^{18}\text{O}$ but also the variation due to other unknown factors. Thus, to some extent, errors in our assumption about the relationship between the accumulation rate and $\delta^{18}\text{O}$ would be absorbed by η_z .

Like Eqs. (15) and (16), we introduce conditional probability density functions based on Eqs. (17) and (18). Since we assumed ε_z and w_z obey $\mathcal{N}(0, \sigma_\varepsilon^2)$ and $\mathcal{N}(0, \sigma_w^2)$, respectively, the conditional distribution of τ_k given ξ_{z_k} and that of $\delta^{18}\text{O}_z$ given A_z become

$$p(\tau_k|\xi_{z_k}, \boldsymbol{\theta}) = \mathcal{N}(\xi_{z_k}, \sigma_\varepsilon^2), \quad (19)$$

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

We hereinafter combine ξ_z and A_z into one vector \mathbf{x}_z as $\mathbf{x}_{z+1} = (\xi_z A_z)^T$. 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. Thus,

$$p(\mathbf{x}_{z+1}|\mathbf{x}_z, \boldsymbol{\theta}) = p(\xi_{z+1}, A_{z+1}|\xi_z, A_z, \boldsymbol{\theta}). \quad (21)$$

We also define the vector of the available data for each z as \mathbf{y}_z . If both the tie point τ_{k_z} and the $\delta^{18}\text{O}$ data are available at z , then, $\mathbf{y}_z = (\tau_{k_z}, \delta^{18}\text{O}_z)^T$. If the $\delta^{18}\text{O}$ data are available but a tie point is unavailable, we define $\mathbf{y}_z = \delta^{18}\text{O}_z$. If neither a tie point nor $\delta^{18}\text{O}$ data are available, we define $\mathbf{y}_z = \emptyset$. Using \mathbf{y}_z , the conditional distributions in Eqs. (19) and (20) can then be combined into the conditional distribution $p(\mathbf{y}_z|\mathbf{x}_z, \boldsymbol{\theta})$ for any z , where we define $p(\mathbf{y}_z = \emptyset|\mathbf{x}_z, \boldsymbol{\theta}) = 1$.

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

$$p(\mathbf{x}_{0:Z}|\mathbf{y}_{1:Z}) = \int p(\mathbf{x}_{0:Z}|\mathbf{y}_{1:Z}, \boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathbf{y}_{1:Z}) d\boldsymbol{\theta}. \quad (22)$$

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

$$p(\mathbf{y}_z|\mathbf{x}_{0:z}, \boldsymbol{\theta}) = p(\mathbf{y}_z|\mathbf{x}_z, \boldsymbol{\theta}). \quad (23)$$

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

$$\begin{aligned} p(\mathbf{x}_{0:z}|\mathbf{y}_{1:z}, \boldsymbol{\theta}) \\ &\propto p(\mathbf{y}_z|\mathbf{x}_z, \boldsymbol{\theta}) p(\mathbf{x}_{0:z}|\mathbf{y}_{1:z-1}, \boldsymbol{\theta}) \\ &= p(\mathbf{y}_z|\mathbf{x}_z, \boldsymbol{\theta}) p(\mathbf{x}_z|\mathbf{x}_{z-1}, \boldsymbol{\theta}) p(\mathbf{x}_{0:z-1}|\mathbf{y}_{1:z-1}, \boldsymbol{\theta}). \end{aligned} \quad (24)$$

220 By applying Eq. (24) recursively, we can obtain $p(\mathbf{x}_{0:z}|\mathbf{y}_{1:z}, \boldsymbol{\theta})$ for any z . Thus, sampling from $p(\mathbf{x}_{0:z}|\mathbf{y}_{1:z}, \boldsymbol{\theta})$ can be achieved using the sequential Monte Carlo (SMC) method (Doucet et al., 2001; Liu, 2001). If z in Eq. (24) is set at the depth at the bottom of the ice core (i.e., $z = Z$), we obtain $p(\mathbf{x}_{0:Z}|\mathbf{y}_{1:Z}, \boldsymbol{\theta})$, which provides the estimate of the age given all the data for the entire ice core.

225 We can also estimate the parameter $\boldsymbol{\theta}$. The posterior distribution of $\boldsymbol{\theta}$ given $\mathbf{y}_{1:Z}$ in Eq. (22) is calculated using the following equation:

$$p(\boldsymbol{\theta}|\mathbf{y}_{1:Z}) \propto p(\mathbf{y}_{1:Z}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \quad (25)$$

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

230
$$\boldsymbol{\theta} = (A_0 \ a \ b \ \mu \ p \ s \ \sigma_\nu \ \sigma_\eta \ \sigma_w)^T . \quad (26)$$

An approximation of $p(\mathbf{y}_{1:Z}|\boldsymbol{\theta})$ 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. In this study, we use flat prior distributions. Since it is unreasonable to allow the parameters except a and b to be negative, the prior distributions for these non-negative parameters was assumed to be a uniform distribution on the non-negative
 235 real line. The prior distributions for the other parameters a and b were assumed to be a uniform distribution on the real line. The shape of the posterior thus 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 the unspecified parameters and is included in $\boldsymbol{\theta}$. The parameter vector $\boldsymbol{\theta}$ also contains
 240 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 fit the data well. On the other hand, 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 noise contained in the data. The posterior given the data provides an appropriate value of
 245 σ_ν so that it is large enough to achieve a good fit, but not too large. The posterior of σ_w indicates the typical magnitude of dispersion of $\delta^{18}\text{O}$ data from the predicted $\delta^{18}\text{O}$ based on the estimated accumulation rate. We did not include σ_ε in $\boldsymbol{\theta}$, but we set a fixed value for σ_ε for each tie point, as shown in Table 2; the values were determined according to Kawamura et al. (2007).

4 Estimation algorithm

250 In order to approximate the conditional distributions $p(\mathbf{x}_{0:Z}|\mathbf{y}_{1:Z}, \boldsymbol{\theta})$ and $p(\boldsymbol{\theta}|\mathbf{y}_{1:Z})$, we employ the PMCMC method (Andrieu et al., 2010), a non-Gaussian hybrid method that combines the SMC and the MCMC methods. In this method, the posterior distributions for the uncertain parameters in the model are computed using standard MCMC, with the exception that the likelihood of the parameters

is approximated using the SMC method. Meanwhile, the age–depth relationship is estimated by performing many repetitions of the SMC procedure under iterations of MCMC. The SMC method can be used to obtain $p(x_{0:z}|y_{1:z}, \theta)$ under a given θ , but it cannot be used to obtain $p(\theta|y_{1:z})$. In principle, MCMC could be used to obtain 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, this would require prohibitive computational cost for high dimensional problems. Thus, use of MCMC is not practical for obtaining high-dimensional distributions like $p(x_{0:z}|y_{1:z}, \theta)$ and $p(x_{0:z}|y_{1:z})$. By combining SMC and 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 an acceptable computational cost. Below, we first present the SMC method, on which the PMCMC method is based. We then describe the PMCMC method and explain how approximations of $p(x_{0:z}|y_{1:z}, \theta)$ and $p(\theta|y_{1:z})$ 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 $p(x_{0:z}|y_{1:z}, \theta)$. The SMC method approximates a probability distribution by a set of N particles, which are the samples drawn from the distribution. Let $\mathbf{x}_{0:z-1}^{(i)}$ be the i -th sample from $p(x_{0:z-1}|y_{1:z-1}, \theta)$, we have the following approximation:

$$p(x_{0:z-1}|y_{1:z-1}, \theta) \approx \frac{1}{N} \sum_{i=1}^N \delta(x_{0:z-1} - \mathbf{x}_{0:z-1}^{(i)}), \quad (27)$$

where $\delta(\cdot)$ denotes the Dirac delta function. If we draw a particle $\mathbf{x}_{z|z-1}^{(i)}$ for each i according to

$$\mathbf{x}_{z|z-1}^{(i)} \sim p(\mathbf{x}_z | \mathbf{x}_{z-1} = \mathbf{x}_{z-1}^{(i)} | y_{1:z-1}, \theta), \quad (28)$$

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

$$p(x_{0:z}|y_{1:z-1}, \theta) \approx \frac{1}{N} \sum_{i=1}^N \delta(x_{0:z} - \mathbf{x}_{0:z}^{(i)}). \quad (29)$$

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

$$\begin{aligned} p(x_{0:z}|y_{1:z}, \theta) &= \frac{p(\mathbf{y}_z | \mathbf{x}_z, \theta) p(x_{0:z}|y_{1:z-1}, \theta)}{p(\mathbf{y}_z | y_{1:z-1}, \theta)} \\ &\approx \sum_{i=1}^N \beta_z^{(i)} \delta(x_{0:z} - \mathbf{x}_{0:z}^{(i)}). \end{aligned} \quad (30)$$

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

$$\beta_z^{(i)} = \frac{p(\mathbf{y}_z | \mathbf{x}_{z|z-1}^{(i)}, \theta)}{\sum_{i=1}^N p(\mathbf{y}_z | \mathbf{x}_{z|z-1}^{(i)}, \theta)}, \quad (31)$$

280 where $p(\mathbf{y}_z | \mathbf{x}_{z|z-1}^{(i)}, \boldsymbol{\theta})$ is called the likelihood of the particle $\mathbf{x}_{z|z-1}^{(i)}$.

Equation (30) indicates that $p(\mathbf{x}_{0:z} | \mathbf{y}_{1:z}, \boldsymbol{\theta})$ can be approximated by weighting the particles $\{\mathbf{x}_{0:z|z-1}^{(i)}\}$. 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
 285 set of N particles $\{\mathbf{x}_{0:z|z}^{(i)}\}$ is obtained by resampling the original particles $\{\mathbf{x}_{0:z|z-1}^{(i)}\}$ such that each $\mathbf{x}_{0:z|z-1}^{(i)}$ is drawn with a probability of $\beta_z^{(i)}$ (see Nakano et al., 2007; van Leeuwen, 2009). After resampling, the original particles in $\{\mathbf{x}_{0:z|z-1}^{(i)}\}$ that have low weights are removed, and those that have high weights are duplicated. The number of the duplicates of $\mathbf{x}_{0:z|z-1}^{(i)}$, $n_z^{(i)}$, becomes approximately equal to $N\beta_z^{(i)}$. The newly generated particles then provide an approximation of $p(\mathbf{x}_{0:z} | \mathbf{y}_{1:z}, \boldsymbol{\theta})$ as
 290 follows:

$$\begin{aligned}
 p(\mathbf{x}_{0:z} | \mathbf{y}_{1:z}, \boldsymbol{\theta}) &\approx \sum_{i=1}^N \beta_z^{(i)} \delta(\mathbf{x}_{0:z} - \mathbf{x}_{0:z|z-1}^{(i)}) \\
 &\approx \sum_{i=1}^N \frac{n_z^{(i)}}{N} \delta(\mathbf{x}_{0:z} - \mathbf{x}_{0:z|z-1}^{(i)}) \\
 &= \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_{0:z} - \mathbf{x}_{0:z|z}^{(i)}).
 \end{aligned} \tag{32}$$

Applying the procedure from Eq. (27) to Eq. (32) recursively up to $z = Z$, we obtain samples from the conditional distribution $p(\mathbf{x}_{0:Z} | \mathbf{y}_{1:Z}, \boldsymbol{\theta})$. If only the marginal distribution $p(\mathbf{x}_z | \mathbf{y}_{1:z}, \boldsymbol{\theta})$, where $\mathbf{x}_{0:z-1}$ is marginalized out, is of interest, it is not necessary to keep the whole sequence of $\mathbf{x}_{0:z|z}^{(i)}$ for
 295 each particle; instead, at each iteration, it is sufficient to keep only the element $\mathbf{x}_{z|z}^{(i)}$ and discard the remaining $\mathbf{x}_{1:z-1}^{(i)}$

4.2 Particle Markov chain Monte Carlo method

An approximation of the marginal likelihood $p(\mathbf{y}_{1:Z} | \boldsymbol{\theta})$ in Eq. (25) can be calculated using SMC (Kitagawa, 1996). If we decompose $p(\mathbf{y}_{1:Z} | \boldsymbol{\theta})$ as

$$\begin{aligned}
 p(\mathbf{y}_{1:Z} | \boldsymbol{\theta}) &= p(\mathbf{y}_{1:Z-1} | \boldsymbol{\theta}) p(\mathbf{y}_Z | \mathbf{y}_{1:Z-1}, \boldsymbol{\theta}) \\
 &= p(\mathbf{y}_1 | \boldsymbol{\theta}) \prod_{z=2}^Z p(\mathbf{y}_z | \mathbf{y}_{1:z-1}, \boldsymbol{\theta}),
 \end{aligned} \tag{33}$$

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

$$\begin{aligned}
 p(\mathbf{y}_z | \mathbf{y}_{1:z-1}, \boldsymbol{\theta}) \\
 = \int p(\mathbf{y}_z | \mathbf{x}_z, \boldsymbol{\theta}) p(\mathbf{x}_{0:z} | \mathbf{y}_{1:z-1}, \boldsymbol{\theta}) d\mathbf{x}_{0:z}
 \end{aligned} \tag{34}$$

Since samples from $p(\mathbf{x}_{0:z-1}|\mathbf{y}_{1:z-1}, \boldsymbol{\theta})$ can be obtained by SMC, a Monte Carlo approximation of the integral in Eq. (34) can be obtained as follows:

$$\begin{aligned}
& p(\mathbf{y}_z|\mathbf{y}_{1:z-1}, \boldsymbol{\theta}) \\
305 \quad & \approx \frac{1}{N} \sum_{i=1}^N \int p(\mathbf{y}_z|\mathbf{x}_z, \boldsymbol{\theta}) \delta(\mathbf{x}_{0:z} - \mathbf{x}_{0:z|z-1}^{(i)}) d\mathbf{x}_{0:z} \\
& = \frac{1}{N} \sum_{i=1}^N p(\mathbf{y}_z|\mathbf{x}_{0:z|z-1}^{(i)}, \boldsymbol{\theta}),
\end{aligned} \tag{35}$$

where we used Eq. (23). We can then approximate the logarithm of $p(\mathbf{y}_{1:Z}|\boldsymbol{\theta})$:

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

and an approximation of the posterior $p(\boldsymbol{\theta}|\mathbf{y}_{1:Z})$ in Eq. (25) can accordingly be obtained. As a matter of fact, however, the approximation given in Eq. (36) for the log-likelihood is too sensitive to the parameter $\boldsymbol{\theta}$ because of the large amount of $\delta^{18}\text{O}$ data. We thus introduce the following relaxation:

$$\log \hat{p}(\mathbf{y}_{1:Z}|\boldsymbol{\theta}) = \sum_{z=1}^Z \log \left[\frac{1}{N} \sum_{i=1}^N p(\delta^{18}\text{O}_z|\mathbf{x}_{0:z|z-1}^{(i)}, \boldsymbol{\theta})^\lambda p(\tau_{k_z}|\mathbf{x}_{0:z|z-1}^{(i)}, \boldsymbol{\theta}) \right], \tag{37}$$

where we set $\lambda = 1/5$ so that the information of tie points becomes effective enough.

Using the Monte Carlo approximation of the marginal likelihood $\hat{p}(\mathbf{y}_{1:Z}|\boldsymbol{\theta})$, we can obtain an approximation of the marginal posterior distribution of $\boldsymbol{\theta}$ using 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 method with ancestor sampling (Lindsten et al., 2014). In this study, because of the ease of implementation, we employ the Metropolis method to obtain an approximation of $p(\boldsymbol{\theta}|\mathbf{y}_{1:Z})$. In the Metropolis method, at the k -th iteration, a proposal sample $\boldsymbol{\theta}^*$ is drawn from the proposal density $q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k-1)})$, which is conditioned by the sample $\boldsymbol{\theta}^{(k-1)}$ obtained at the previous iteration:

$$\boldsymbol{\theta}^* \sim q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k-1)}). \tag{38}$$

In this paper, the proposal distribution q was taken to be a zero-mean Gaussian distribution with a fixed variance for each element of $\boldsymbol{\theta}$. This means we assume a symmetrical proposal distribution satisfying

$$q(\boldsymbol{\theta}|\boldsymbol{\theta}') = q(\boldsymbol{\theta}'|\boldsymbol{\theta}) \tag{39}$$

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, in order that the Markov chain rapidly moves around in the parameter space. The proposal sample $\boldsymbol{\theta}^*$ is accepted with the following probability:

$$\min \left(1, \frac{\hat{p}(\mathbf{y}_{1:Z}|\boldsymbol{\theta}^*) p(\boldsymbol{\theta}^*)}{\hat{p}(\mathbf{y}_{1:Z}|\boldsymbol{\theta}^{(k-1)}) p(\boldsymbol{\theta}^{(k-1)})} \right), \tag{40}$$

where $\hat{p}(\mathbf{y}_{1:Z}|\boldsymbol{\theta})$ is an approximation of the marginal likelihood obtained by SMC as written in Eq. (37). If $\boldsymbol{\theta}^*$ is accepted, we set $\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^*$; otherwise, we set $\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)}$ and thus $\hat{p}(\mathbf{y}_{1:Z}|\boldsymbol{\theta}^{(k)}) = \hat{p}(\mathbf{y}_{1:Z}|\boldsymbol{\theta}^{(k-1)})$. Using $\boldsymbol{\theta}^{(k)}$, the proposal sample at the next iteration can be obtained according to Eq. (38). Iterating the above procedure generates a large number of samples that obey the posterior
335 distribution $p(\boldsymbol{\theta}|\mathbf{y}_{1:Z})$. A short summary of PMCMC is also found in a pseudo-code in the original paper by Andrieu et al. (2010).

In the above algorithm, an approximated value of the marginal likelihood $p(\mathbf{y}_{1:Z}|\boldsymbol{\theta})$ is computed using the SMC method at each iteration of the Metropolis method. It should be noted that Eq. (34) can be modified as follows:

$$\begin{aligned}
& p(\mathbf{y}_z|\mathbf{y}_{1:z-1}, \boldsymbol{\theta}) \\
340 \quad & = \int p(\mathbf{y}_z|\mathbf{x}_z, \boldsymbol{\theta}) p(\mathbf{x}_z|\mathbf{x}_{z-1}, \boldsymbol{\theta}) p(\mathbf{x}_{z-1}|\mathbf{y}_{1:z-1}, \boldsymbol{\theta}) d\mathbf{x}_{z-1} d\mathbf{x}_z.
\end{aligned} \tag{41}$$

Thus, in calculating $p(\mathbf{y}_{1:Z}|\boldsymbol{\theta})$ in Eq. (33), it is not necessary to consider the joint distribution of the sequence $\mathbf{x}_{0:Z}$; it is sufficient to consider the marginal distribution $p(\mathbf{x}_z|\mathbf{y}_{1:z}, \boldsymbol{\theta})$ for each z . As mentioned above, sampling from $p(\mathbf{x}_z|\mathbf{y}_{1:z}, \boldsymbol{\theta})$ can be achieved when discarding $\mathbf{x}_{1:z-1}$. This greatly reduces the computational cost because it can skip some procedures for handling the whole
345 sequence of 2510 time steps ($Z = 2510$ in this paper) for 5000 particles. In addition, the memory cost is also remarkably reduced, although the memory cost could be reduced by using another efficient algorithm by Jacob et al. (2015). We then discard $\mathbf{x}_{1:z-1}$ in order to obtain an approximation of $p(\mathbf{y}_{1:Z}|\boldsymbol{\theta})$ at each iteration of the Metropolis method.

As mentioned in Section 3, if we retain the samples for the whole sequence $\mathbf{x}_{0:Z}$ from a run of
350 SMC with a given $\boldsymbol{\theta}$, we obtain samples from $p(\mathbf{x}_{0:Z}|\mathbf{y}_{1:Z}, \boldsymbol{\theta})$. The Metropolis procedure sequentially generates a large number of samples that obey the marginal posterior distribution $p(\boldsymbol{\theta}|\mathbf{y}_{1:Z})$. By combining the SMC samples with various $\boldsymbol{\theta}$ values that obey $p(\boldsymbol{\theta}|\mathbf{y}_{1:Z})$, we can obtain the samples representing the marginal posterior distribution $p(\mathbf{x}_{0:Z}|\mathbf{y}_{1:Z})$ where $\boldsymbol{\theta}$ is marginalized out according to Eq. (22). If samples that obey $p(\boldsymbol{\theta}|\mathbf{y}_{1:Z})$ are obtained in advance, the sampling proce-
355 dures from $p(\mathbf{x}_{0:Z}|\mathbf{y}_{1:Z}, \boldsymbol{\theta})$ for various $\boldsymbol{\theta}$ can be performed in parallel, and an approximation of the marginal posterior distribution $p(\mathbf{x}_{0:Z}|\mathbf{y}_{1:Z})$ can be obtained efficiently.

5 Results

We applied the PMCMC method to the Dome Fuji ice core. In this study, the thickness of the ice sheet H is assumed to be 3031m. The bottom of the core Z is 2505m. Following a burn-in period,
360 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 $\boldsymbol{\theta}$, $p(\boldsymbol{\theta}|\mathbf{y}_{1:Z})$. For each run of SMC, 5000 particles were used to obtain samples from $p(\mathbf{x}_{0:Z}|\mathbf{y}_{1:Z}, \boldsymbol{\theta}^{(k)})$.

Figure 2 shows the marginal histograms for the estimated posterior distribution for each parameter. The posterior mean and standard deviation of the present accumulation rate $A_0(= A(0))$

365 were 0.0278 and 0.0019 [m/year], respectively. This result is in good agreement with the mea-
 surements by Kameda et al. (2008), who reported the surface mass balance at Dome Fuji to be
 $27.3 \pm 1.5 \text{ kg}/(\text{m}^{-2} \cdot \text{year})$, which corresponds to about 0.0273m/year. The maxima of the poste-
 rior 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). The result in
 370 Figure 2 suggests that μ is most likely between 0 and 2% of the accumulation rate. Considering that
 the accumulation rate A were mostly less than 0.03m/year, we can guess that the basal melting rate
 is mostly less than 0.0006m/year(= 0.6mm/year). This roughly agrees with the result by Parrenin
 et al., who showed that the basal melting rate is likely to be less than 0.4mm/year. Such a small
 value of m is consistent with our assumption of the pseudo-steady state, in which the ratio m/A is
 375 constant as described in Section 2. 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 models of the accumulation
 rate. In addition, the setting of the thinning factor in this study is different from that used by Parrenin
 et al. as discussed later. Thus, it should not be expected that they would necessarily provide similar
 380 results.

In the posterior distribution, some of the parameters are correlated with each other. Figure 3
 shows two-dimensional histograms of 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
 385 correlations between the three parameters a , b , and σ_η are observed in this posterior distribution.
 These three parameters are related to the accumulation rate and $\delta^{18}\text{O}$ data. Thus, the accuracy of the
 estimation for these three parameters could be much improved if any of the three parameters could
 be effectively constrained.

Figure 4 shows the estimated age as a function of depth. The red solid line indicates the median
 390 of the posterior distribution and the red dotted lines indicate the 10th and 90th percentiles of the
 posterior distribution. For reference, the result by Parrenin et al. (2007) is indicated by a grey
 line. 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 $p(\mathbf{x}_{0:Z}|\mathbf{y}_{1:Z})$ five times with different seeds and confirmed that there were no apparent
 395 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 considered reliable. The SMC
 method often suffers from the degeneracy problem, especially when the number of steps is large. In
 PMCMC, this problem is overcome by collecting a large number of SMC samples from the iterations
 of the Metropolis method. In Figure 4, it is difficult to discriminate the 10th and 90th percentiles
 400 from the median because the width of the posterior distribution is much smaller than the range of
 Figure 4. In order to make the width of the posterior visible, the 10th and 90th percentiles of the

posterior distribution are indicated by red dotted lines around the median of the posterior distribution in Figure 5. Black crosses show the difference between each tie point and the median of the posterior. The uncertainty of age is minimized at each tie point, where the age is known with high accuracy, although it is not possible to completely remove any uncertainty. In Figure 5, a grey line indicates the difference between the estimate by Parrenin et al. (2007) and the median of the posterior obtained by the proposed method. Note that this line tends to deviate further from the median than do the black crosses. This means that the estimate with the proposed method fits the tie points more closely than does the estimate by Parrenin et al., although the difference between the two results is about 3000 years at the greatest.

Figure 6 shows 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 indicated by a grey line. Since, by definition, $\Theta = 1$ at the surface, the width of the posterior distribution is almost zero near the surface, and the uncertainty becomes larger in the deeper core. The profiles of the thinning factor indicated by the red solid line and the grey line differ, and this is probably caused by the assumption of a constant ice thickness. While Parrenin et al. (2007) allowed the ice thickness H to vary, we assumed that it was constant; it would be instructive to examine the effect of this assumption in a future work. 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 the 80% confidence interval, is indicated by a blue dotted line. In this way, we can estimate the age and related variables, and we can also obtain information about the accuracy 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:

$$p(A|\xi) = \int p(A|z)p(z|\xi) dz \quad (42)$$

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

$$p(z|\xi) = \frac{p(\xi|z)p(z)}{\int p(\xi|z)p(z) dz}. \quad (43)$$

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

6 Discussion

In order to evaluate the robustness, we obtained the estimate without using the last five tie points at below 2400m depth. We estimated the parameters and the age–depth relationship from the other 20 tie points and the $\delta^{18}\text{O}$ data. Figure 9 shows the histograms of the marginal posterior distributions

of the nine parameters. The results without using the five tie points are indicated by red lines and the results with all the tie points, which are the same as the results in Figure 2, are indicated as blue lines. The posterior distributions obtained without some of the tie points are very similar to the result shown in Figure 2. However, the posterior distributions of the three parameters a , b , and σ_η are slightly different. Since σ_η was estimated to be larger when the five tie points were not used, this might indicate that the variation in the accumulation rate was noisier near the bottom. The marginal posterior distribution for a and b could accordingly be modified due to the correlation with σ_η shown in Figure 3. However, more careful evaluation would be required to determine the reason.

Figure 10 shows the estimates of the age as a function of depth are compared between the result without using the last five tie points and that with all the tie points. In order to make the differences visible, this figure shows the differences from the median of the posterior without the last five tie points like in Figure 5. The red lines indicate the estimate without the last five tie points, and the grey lines indicate the estimate with all the tie points. The dotted lines indicate the 10th and 90th percentiles of the posterior distributions. The tie points used for the estimation are shown with black crosses. The deviation of the grey lines tended to be large near the bottom of the ice core. However, the grey lines were within the range of uncertainty shown with the red dotted lines. This suggests that our model does a good job of representing the uncertainties due to the lack of the information.

The accumulation rate as a function of age was also estimated without using the five tie points at 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 tie points, and for reference, the grey line indicates the estimate using all the tie points. The solid lines indicate the median of the posterior, and the 10th and 90th percentiles are indicated by 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 is acceptable.

The proposed technique requires a 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 used in SMC (e.g., Doucet et al., 2001). This problem should be addressed in the future.

There may be a room for improvement in the model for the accumulation rate described by Eq. (12). Equation (12) 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 of fit could be evaluated using some metric such as Bayes factors. However, it would require a great deal of time to evaluate a wide variety of choices, and so such a search is beyond the scope of this study.

This study used the $\delta^{18}\text{O}$ data and tie points deduced from O_2/N_2 data to estimate the age–depth relationships. However, PMCMC allows us to use various kinds of data. Thus, data from
475 other various 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 SMC approach could be extended to include information at
480 multiple sites; this would be a useful area for future work.

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 with a model describing the accumulation of snow and glaciological dynamics. This technique provides estimates of unspecified parameters in the model from the
485 posterior distributions 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
490 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
495 typically nonlinear, it is not necessarily justified to assume linearity and Gaussianity when using a temperature proxy to date an ice core. The PMCMC method allows us to use various kinds of data which are expected to have a nonlinear relationship with the model variables. Another advantage is that the PMCMC method estimates the 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
500 the model parameters.

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

- 505 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.,
510 Panno, L., Spahni, R., Tisserand, A., Siegenthaler, U., and Leuenberger, M.: Anomalous flow below 2700m in the EPICA Dome C ice core detected using $\delta^{18}\text{O}$ of atmospheric oxygen measurements, *Clim. Past*, 3, 341–353, 2007.
- Freitag, J., Kipfstuhl, S., and Laepple, T.: Core-scale radiosopic imaging: a new method reveals density–calcium link in Antarctic firn, *J. Glaciology*, 59, doi:10.3189/2013JoG13J028, 1009–1014, 2013.
- 515 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.
- Jacob, P. E., Murray, L. W., and Rubenthaler, S.: Path storage in the particle filter, *Stat. Comput.*, 25, 487–496, 2015.
- Kameda, T., Motoyama, H., Fujita, S., and Takahashi, S.: Temporal and spatial variability of surface mass
520 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
525 over the past 360,000 years, *Nature*, 448, doi:10.1038/nature06015, 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–
530 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.
- 535 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.
- 540 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.
- Martin, C. and Gudmundsson, G. H.: Effects of nonlinear rheology, temperature and anisotropy on the relation-

- ship between age and depth at ice divides, *The Cryosphere*, 6, 1221–1229, 2012.
- 545 Nakano, S., Ueno, G., and Higuchi, T.: Merging particle filter for sequential data assimilation, *Nonlin. Process. Geophys.*, 14, 395–408, 2007.
- Parrenin, F. and Hindmarsh, R. C. A.: Influence of a non-uniform velocity field on isochrone geometry along a steady flowline of an ice sheet, *J. Glaciology*, 53, 612–622, 2007.
- Parrenin, F., Waelbroeck, J. J. C., Ritz, C., and Barnola, J.-M.: Dating the Vostok ice core by an inverse method, 550 *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.: 555 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, 560 *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 565 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
Z	Vertical coordinate at the bottom
H	Thickness of the ice sheet
ξ	Age
A	Accumulation rate
Θ	Thinning factor
μ, p, s	Parameters for Θ
U	Vertical velocity
ζ	Rescaled vertical coordinate
u	Rescaled velocity
ω	Flux shape function
m	Basal melting rate
$\nu_z / \sqrt{A_z \Theta_z}$	System noise for age
σ_ν^2	Variance of ν_z
$\eta_z / \sqrt{A_z \Theta_z}$	System noise for accumulation rate
σ_η^2	Variance of η_z
τ_k	Tie point (age marker)
ε_k	Observation noise for tie point
σ_ε	Variance of τ_k
$\delta^{18}\text{O}$	$\delta^{18}\text{O}$
a, b	Parameters of the observation model for $\delta^{18}\text{O}$
w_z	Observation noise for $\delta^{18}\text{O}$
σ_w	Variance of w
\mathbf{x}_z	State at depth z ($\mathbf{x}_z = (\xi_z A_z)^T$)
\mathbf{y}_z	Observation at depth z ($\mathbf{y}_z = (\tau_{k_z} \delta^{18}\text{O}_z)^T$)
$\boldsymbol{\theta}$	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.

Depth	Age	Uncertainty of 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. Depth, age, and uncertainty of age at each tie point.

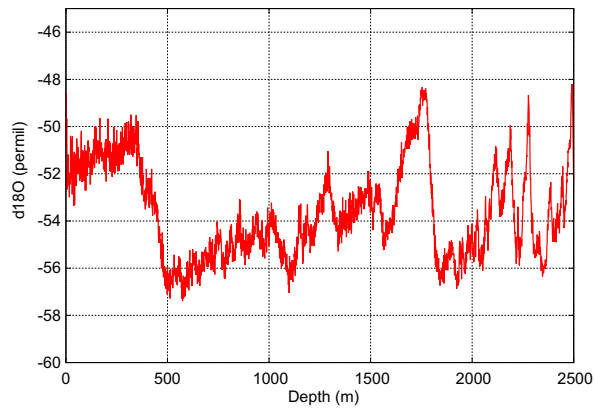


Fig. 1. $\delta^{18}\text{O}$ data used in this study.

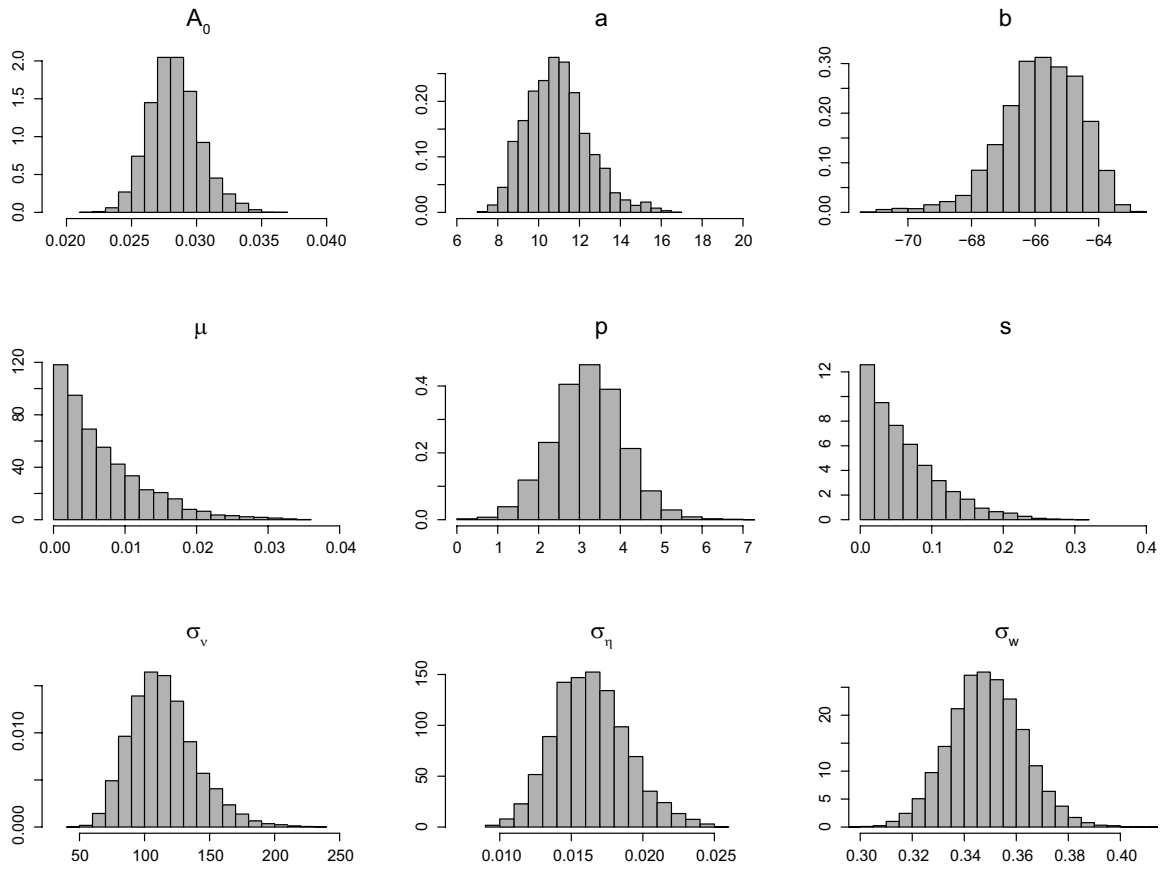


Fig. 2. Estimated marginal distributions of the posterior distributions for the nine parameters.

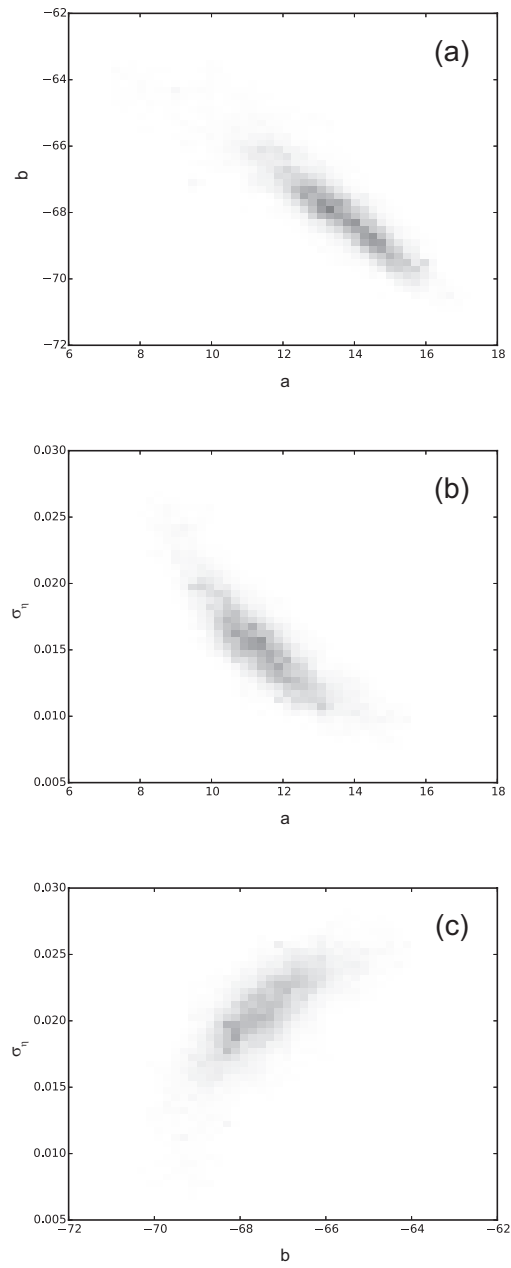


Fig. 3. Two-dimensional histograms of the marginal posterior distribution of a and b (a), the marginal distribution of a and σ_η (b), and the marginal distribution of b and σ_η (c).

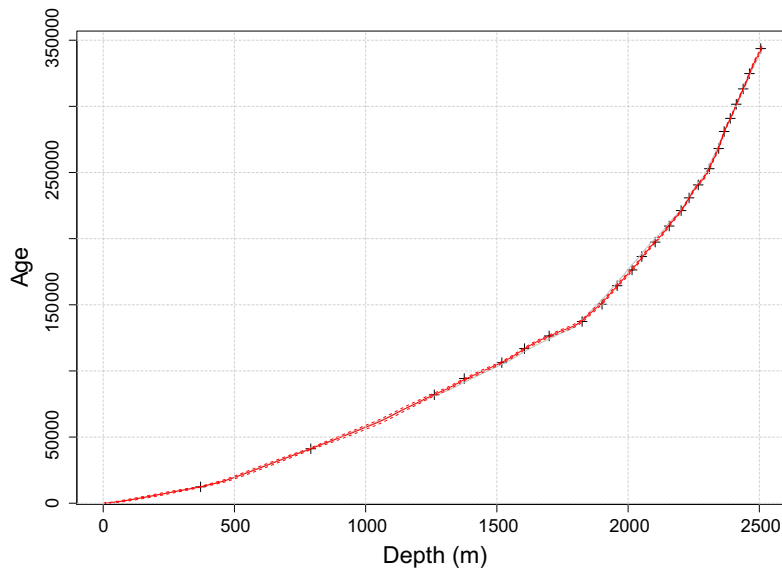


Fig. 4. Estimated age as a function of depth. The solid line indicates the median of the posterior distribution. The 10th and 90th percentiles of the posterior are indicated by red dotted lines. The black crosses indicates the tie points. The result obtained by Parrenin et al. (2007) is also indicated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

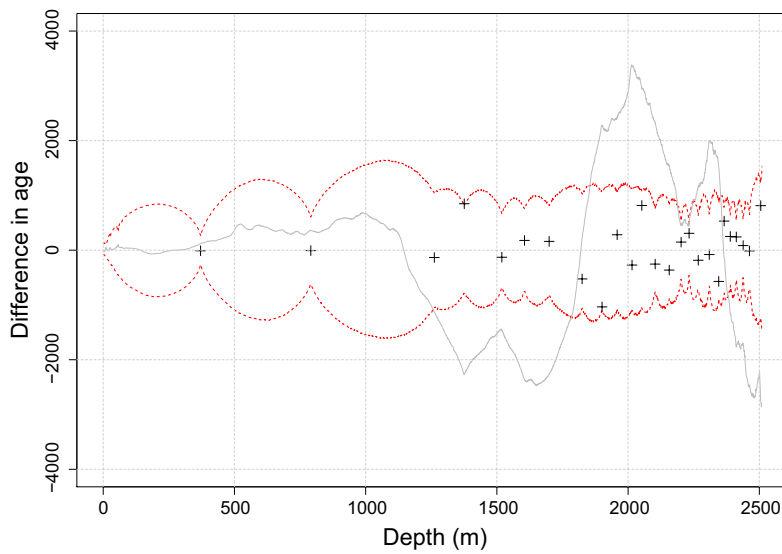


Fig. 5. Difference of the 10th and 90th percentiles of the posterior distribution from the median of the posterior (red dotted lines), difference of each tie point from the median of the posterior (black crosses), and difference of the estimate by Parrenin et al. (2007) from the median of the posterior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

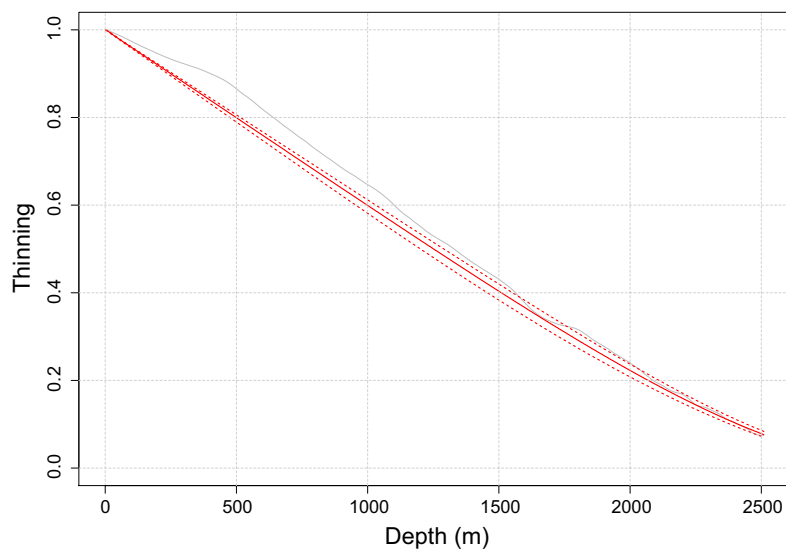


Fig. 6. Estimated thinning factor Θ as a function of depth. The median of the posterior is indicated by a red solid line, the 10th and 90th percentiles are indicated by red dotted lines, and the estimate by Parrenin et al. (2007) is indicated by a grey line.

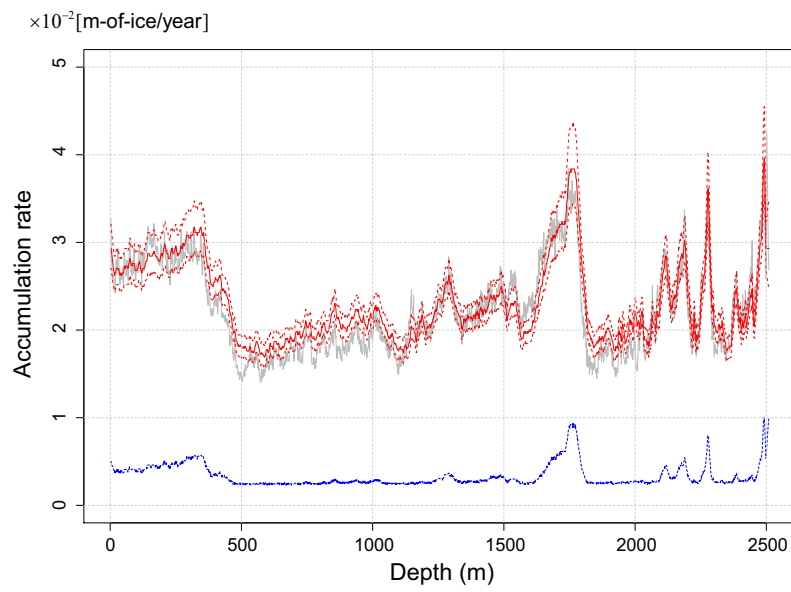


Fig. 7. Estimated accumulation rate as a function of depth. The median of the posterior is indicated by a red solid line, the 10th and 90th percentiles are indicated by red dotted lines, the difference between the 10th and 90th percentiles is indicated by a blue dotted line, and the estimate by Parrenin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

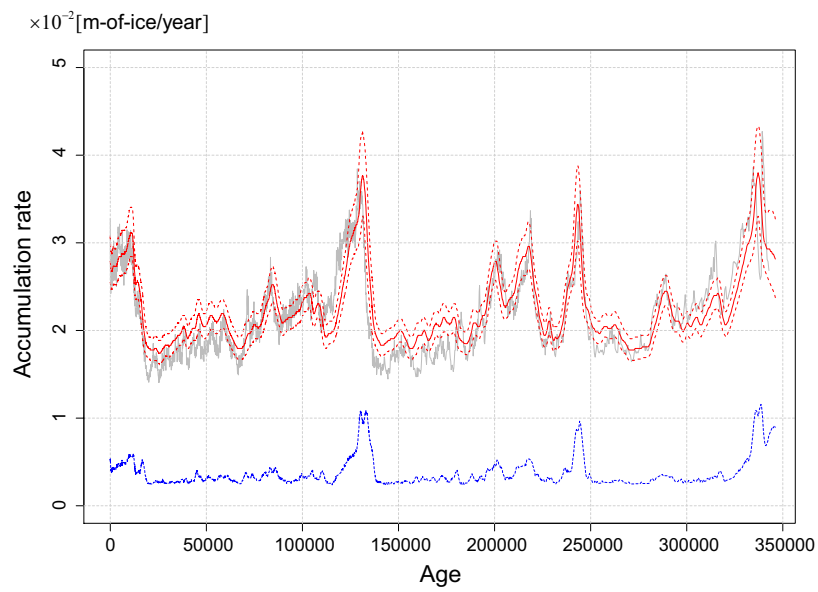


Fig. 8. Estimated accumulation rate as a function of age. The median of the posterior is indicated by a red solid line, the 10th and 90th percentiles are indicated by red dotted lines, the difference between the 10th and 90th percentiles is indicated by a blue dotted line, and the estimate by Parrenin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

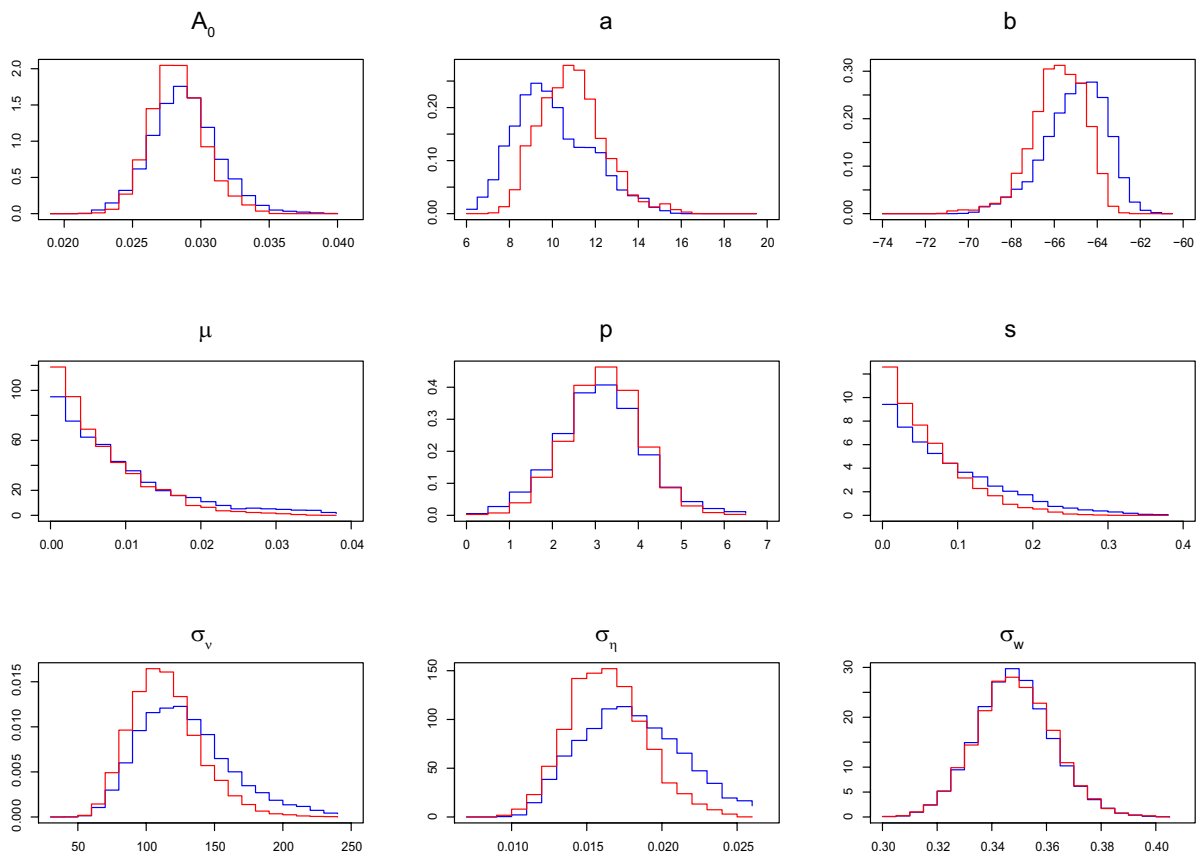


Fig. 9. Estimated marginal distributions of each of the nine parameters; without using the last five tie points (red) and with all the tie points (blue).

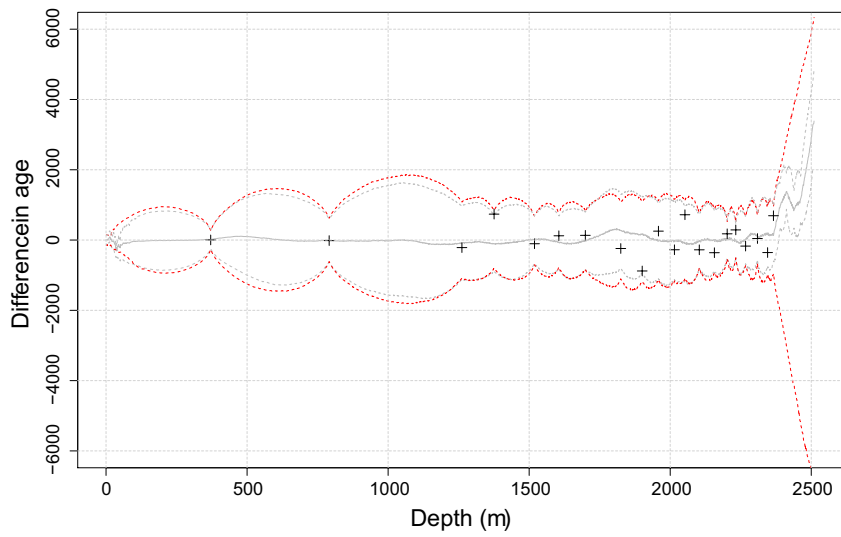


Fig. 10. Difference of the 10th and 90th percentiles of the posterior distribution from the median of the posterior (red dotted lines) for the result without using the last five age markers. The difference of each tie point from the median of the posterior (black crosses), and the difference between the estimate with all the tie points and the estimate without using the last five tie points (grey line) are also shown.

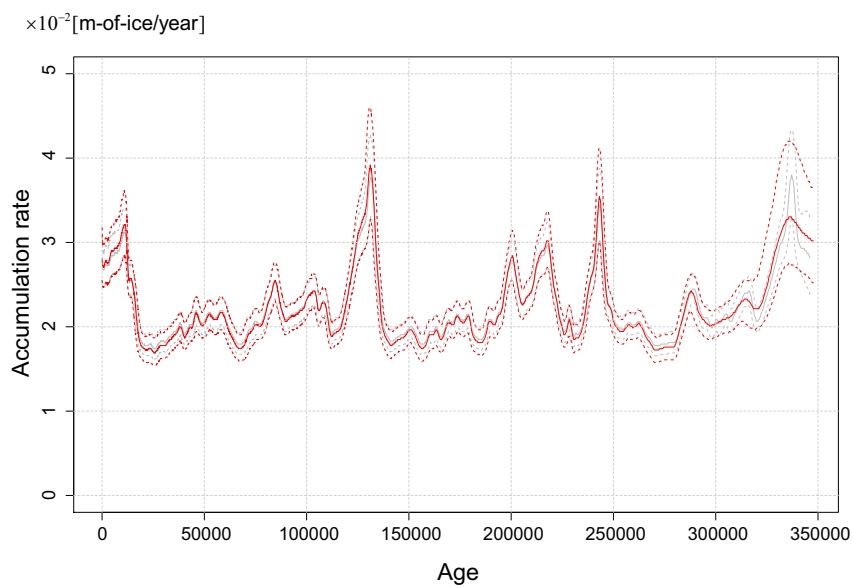


Fig. 11. Estimated accumulation rate as a function of age: without using the last five tie points (red) and with all the tie points (grey). The solid lines indicate the median of the posterior distribution. The 10th and 90th percentiles of the posterior are indicated by dotted lines.