https://doi.org/10.5194/egusphere-2026-3558
Preprint. Discussion started: 30 June 2026

Brief communication: Estimating diffusion length from low resolution data

Fyntan Shaw[1], Torben Kunz[2], and Thomas Laepple[1, 3]

[1]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany
[2]Deutscher Wetterdienst, Offenbach, Germany
[3]University of Bremen, MARUM - Centre for Marine Environmental Sciences and Faculty of Geosciences, Bremen, Germany

Correspondence: Fyntan Shaw (fyntan.shaw@awi.de)

Abstract.

Data-derived estimates of water isotope diffusion lengths in ice cores enable corrections for diffusion-induced signal attenuation and are also used to infer past firn temperatures, but rely on fitting the correct spectral model to the observed power spectrum. Established approaches do not account for additional smoothing and aliasing introduced by discrete sampling, which can bias diffusion length estimates. We show that this bias increases with coarser sampling and exceeds 10% when the sampling interval is greater than approximately twice the diffusion length. By explicitly incorporating sampling effects into the spectral model, we derive an unbiased estimator that improves diffusion length estimates from coarse, on-line measurements and from ice core sections with small diffusion lengths.

1 Introduction

Since the discovery of the relationship between stable water isotopes and temperature (Dansgaard, 1964), numerous ice cores have been drilled to obtain temperature reconstructions dating back as far as 800 kya (Petit et al., 1999; EPICA community members, 2004; NGRIP members, 2004). Ice cores are uniquely valuable due to their high temporal resolution over millennial timescales compared to other proxy archives. However, water isotope signals are affected by diffusion, which smooths high-frequency variability (Johnsen, 1977; Johnsen et al., 2000). Diffusion length, defined as the average vertical displacement of water molecules from their initial deposition layer, can be estimated from the data by fitting a model to the estimated power spectrum (Johnsen et al., 2000; Gkinis et al., 2014; Kahle et al., 2018; Holme et al., 2018). Knowledge of the diffusion length not only enables methods to restore isotopic variability loss, but also offers an alternative, independent temperature proxy (Simonsen et al., 2011; Gkinis et al., 2014; Holme et al., 2018).

Once an ice core arrives at the laboratory, two common measurement approaches are used to extract paleoclimate records: continuous flow analysis (Gkinis et al., 2011) and traditional discrete analysis. The latter involves dividing the ice core into smaller samples, which are then melted and measured individually. This sampling procedure introduces additional artifacts into the data that affect the power spectrum and, consequently, the reliability of established diffusion length estimation methods.

https://doi.org/10.5194/egusphere-2026-3558
Preprint. Discussion started: 30 June 2026

In this paper, we derive theoretical expressions describing the effects of discrete sampling on the data, and incorporate them into the established diffusion length model fit. Using surrogate data with varying sampling intervals, we apply a Monte Carlo approach to demonstrate that the resulting estimator is unbiased. We compare these results with estimates obtained using conventional methods and an alternative approach previously proposed to account for sampling effects. Finally, we highlight the practical relevance of the unbiased estimator for regions with small diffusion lengths and for obtaining early, reliable estimates from coarse, on-line data.

2 Theory

The effect of diffusion on some initial water isotope record, $X_0(z)$, can be described as a convolution of the record with a Gaussian filter, $G(z)$,

$$X_d(z) = X_0(z) * G(z), \tag{1}$$

where $X_d(z)$ is the diffused isotope signal, z is the depth relative to the initial layer of deposition and $*$ denotes the convolution operator. The filter $G(z)$ is given by,

$$G(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-z^2}{2\sigma^2}}. \tag{2}$$

The parameter σ is the diffusion length and describes the width of the Gaussian filter and thus the magnitude of diffusion. According to the convolution theorem, the Fourier transform of two convolved signals is equal to the product of their Fourier transforms (Cochran et al., 1967). Thus, Eq. 1 can be expressed in the frequency domain as,

$$\hat{X}_d(f) = \hat{X}_0(f) \cdot \hat{G}(f) \tag{3}$$

$$= \hat{X}_0(f) \cdot e^{-0.5(2\pi f\sigma)^2}, \tag{4}$$

where the hat operator represents the Fourier transform. The squared magnitude of this (i.e., the power spectral density, PSD), here denoted by P, forms the basis for conventional diffusion length estimation from the power spectrum, with an additional term P_n to account for measurement noise,

$$P_{X_d}(f) = P_{X_0}(f) \cdot e^{-(2\pi f\sigma)^2} + P_n(f). \tag{5}$$

An additional smoothing effect arises from the discrete sampling process, as each sample represents an average over a finite interval rather than a point measurement. For a sampling interval a, this averaging can be described as a convolution of the original signal $X_0(z)$ with a rectangular window,

https://doi.org/10.5194/egusphere-2026-3558
Preprint. Discussion started: 30 June 2026

$$X_r(z) = X_0(z) * \Pi\left(\frac{z}{a}\right), \tag{6}$$

where $X_r(z)$ is the averaged record and $\Pi(\frac{z}{a})$ is a rectangular function of width a given by,

$$\Pi\left(\frac{z}{a}\right) = \begin{cases} \frac{1}{a}, & \text{if } |z| \leq \frac{a}{2} \\ 0, & \text{if } |z| > \frac{a}{2}. \end{cases} \tag{7}$$

Since the Fourier transform of a rectangular function is a sinc function, Eq. 6 in the frequency domain becomes,

$$\hat{X}_r(f) = \hat{X}_0(f) \cdot \hat{\Pi}(f) \tag{8}$$

$$= \hat{X}_0(f) \cdot \mathrm{sinc}(a\pi f). \tag{9}$$

In addition to these smoothing effects, the discretisation of the continuous record $X_0(z)$ must also be considered. This can be described as a multiplication of the continuous signal with a Dirac comb,

$$X_s(z) = X_0(z) \cdot \text{Ш}(z), \tag{10}$$

where $X_s(f)$ is the discretised record and $\text{Ш}(z)$ denotes the Dirac comb function defined as,

$$\text{Ш}(z) = \sum_{k=-\infty}^{+\infty} \delta(z - ka), \tag{11}$$

Since the Fourier transform of a Dirac comb function is another Dirac comb function, the Fourier transform of the discretised record is,

$$\hat{X}_s(f) = \hat{X}_0(f) * \text{Ш}(f), \tag{12}$$

which simplifies, using the shifting property of the Dirac delta function, to

$$\hat{X}_s(f) = \sum_{k=-\infty}^{+\infty} \hat{X}_0\left(f - \frac{k}{a}\right). \tag{13}$$

This can be conceptualised as the sum of infinite repeated copies of $\hat{X}_0(f)$ centred around frequencies equal to multiples of k. The frequencies beyond the Nyquist frequency are thus reflected into $\hat{X}_s(f)$, an effect known as spectral aliasing. Fig. 1 provides a simple example to help illustrate how this aliasing arises.

https://doi.org/10.5194/egusphere-2026-3558
Preprint. Discussion started: 30 June 2026

Combining all of these effects gives an expression for the Fourier transform of a discretely measured diffused water isotope record $X_m(z)$ with sample size and spacing a,

$$\hat{X}_m(f) = \hat{X}_0(f) \cdot \hat{G}(f) \cdot \hat{\Pi}(f) * \text{Ш}(f) \tag{14}$$

$$= \hat{X}_0(f) \cdot e^{-0.5(2\pi f\sigma)^2} \cdot \text{sinc}(a\pi f) * \sum_{k=-\infty}^{+\infty} \delta\left(f - \frac{k}{a}\right) \tag{15}$$

$$= \sum_{k=-\infty}^{+\infty} \hat{X}_0\left(f - \frac{k}{a}\right) \cdot e^{-0.5(2\pi\left(f-\frac{k}{a}\right)\sigma)^2} \cdot \text{sinc}\left(a\pi\left(f - \frac{k}{a}\right)\right), \tag{16}$$

and the expected power spectrum of $X_m(z)$ is,

$$P_{X_m}(f) = \left\langle \left| \hat{X}_m(f) \right|^2 \right\rangle \tag{17}$$

$$= \left\langle \left| \sum_{k=-\infty}^{+\infty} \hat{X}_0\left(f - \frac{k}{a}\right) \cdot e^{-0.5(2\pi\left(f-\frac{k}{a}\right)\sigma)^2} \cdot \text{sinc}\left(a\pi\left(f - \frac{k}{a}\right)\right) \right|^2 \right\rangle \tag{18}$$

$$= \sum_{k=-\infty}^{+\infty} \left\langle \left| \hat{X}_0\left(f - \frac{k}{a}\right) \cdot e^{-0.5(2\pi\left(f-\frac{k}{a}\right)\sigma)^2} \cdot \text{sinc}\left(a\pi\left(f - \frac{k}{a}\right)\right) \right|^2 \right\rangle \tag{19}$$

$$= \sum_{k=-\infty}^{+\infty} P_{X_0}\left(f - \frac{k}{a}\right) \cdot e^{-(2\pi\left(f-\frac{k}{a}\right)\sigma)^2} \cdot \text{sinc}^2\left(a\pi\left(f - \frac{k}{a}\right)\right), \tag{20}$$

where again P represents the PSD and $\langle . \rangle$ denotes the expected value operator. The step from Eq. 18 to Eq. 19 is possible, because for a stochastic process the phases of the Fourier transforms at different frequencies are independent. This is analogous to the fact that the variance of the sum of independent stochastic processes equals the sum of their variances.

The shaded region in Fig. 1 indicates the resolved frequency range. Starting from a white-noise spectrum (black), diffusion attenuates high frequencies (red). Finite sample averaging further smooths the spectrum (purple), while discrete sampling introduces aliasing (blue). The aliasing arises from the sum of shifted spectral copies (grey, solid), generated by convolution with the Dirac comb. This is equivalent to the power spectral density being reflected at the Nyquist frequencies (grey, dashed lines).

For the model fit, the infinite summation in Eq. 20 can be approximated by retaining only the terms $k = 0, 1$, as the power within the resolved frequency range (shaded region in Fig. 1) is negligible for $k \geq 2$. In addition, the power spectrum of the measurement noise, P_n, must be included, which for discretely sampled records can be approximated as white noise. This yields the final model expression,

$$P_{X_m}(f) = P_{X_0}(f) \cdot e^{-(2\pi f\sigma)^2} \cdot \text{sinc}^2(a\pi f) + P_{X_0}\left(f - \frac{1}{a}\right) \cdot e^{-(2\pi\left(f-\frac{1}{a}\right)\sigma)^2} \cdot \text{sinc}^2\left(a\pi\left(f - \frac{1}{a}\right)\right) + P_n, \tag{21}$$

https://doi.org/10.5194/egusphere-2026-3558
Preprint. Discussion started: 30 June 2026

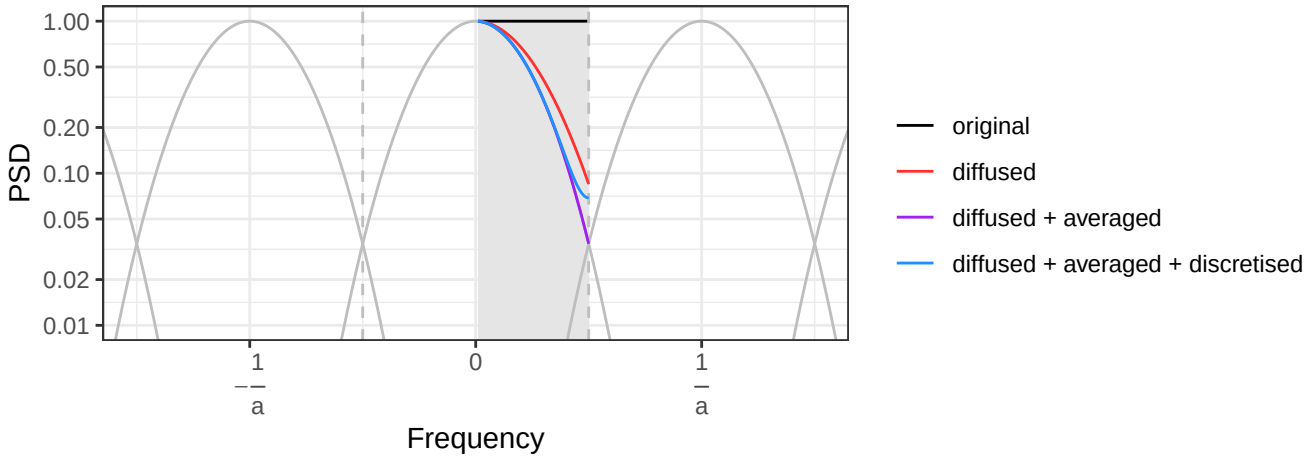

Figure 1. Illustration of the effects of diffusion and discrete sampling on the power spectrum of isotopic records. The shaded region indicates the resolved frequency range for a sampling interval of a. Starting from a white-noise spectrum (black), diffusion attenuates high frequencies (red). Finite sample averaging further reduces high frequency variability (purple), while discrete sampling introduces aliasing (blue). The aliasing arises from the sum of shifted spectral copies (grey, solid), generated by convolution with the Dirac comb. This is equivalent to the power spectral density being reflected at the Nyquist frequencies (grey, dashed lines).

For very small sample intervals, the second term becomes negligible, and $\mathrm{sinc}^2(a\pi f)$ approaches unity. In this limit, the expression reduces to the conventional diffusion length model (Eq. 5), in which sampling effects are absent.

3 Sample resolution vs diffusion length

The significance of the additional sampling effects outlined in Sect. 2 depends not only on the sample resolution, but also on the diffusion length. For large diffusion lengths relative to the sample interval, the effect of finite sample averaging becomes negligible, as high-frequency variability has already been attenuated by diffusion. Similarly, since spectral aliasing primarily affects the highest frequencies, its impact is also negligible in strongly diffused records. To explore this relationship, we first take a qualitative approach by comparing the idealised power spectra of a diffused isotope record, before and after sampling, using Eq. 5 and Eq. 21 respectively (Fig. 2). For simplicity, we assume a constant initial PSD of $P_{X_0} = 1$, equivalent to a white-noise time series with standard deviation 1. We compute the PSDs obtained using 4 different sample intervals relative to the diffusion length, $a = 0.5\sigma$, $a = \sigma$, $a = 2\sigma$, $a = 4\sigma$.

https://doi.org/10.5194/egusphere-2026-3558
Preprint. Discussion started: 30 June 2026

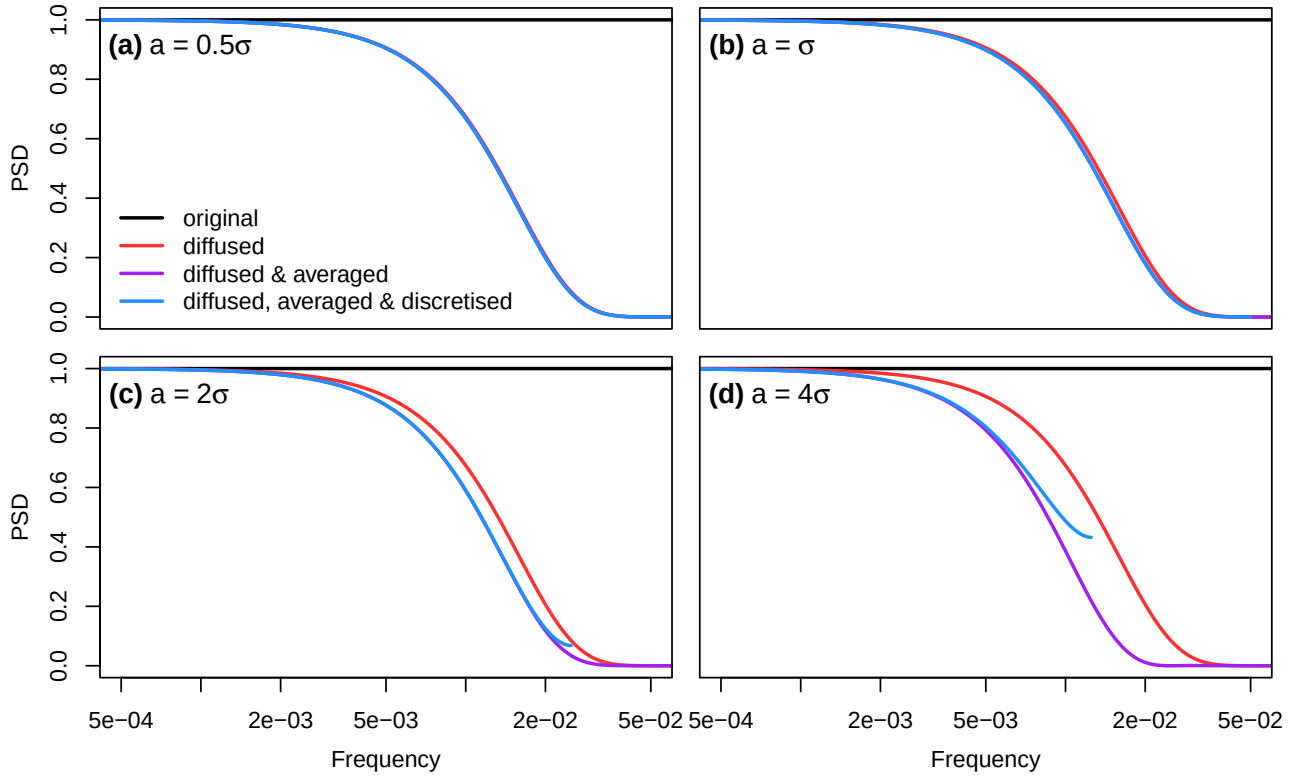

Figure 2. PSDs of a white-noise timeseries at 4 different stages, for sampling sizes of **(a)** 0.5σ, **(b)** σ, **(c)** 2σ and **(d)** 4σ. The original record (black) is diffused (red), then averaged (purple) and then discretised (blue).

Fig. 2 shows that for sampling sizes of 0.5σ (or smaller), the additional effects of sampling are negligible. A slight reduction relatvie to the diffused PSD appears $a = \sigma$ (Fig. 2b), but aliasing effect remains insignificant. For coarser sampling ($a = 2\sigma$ Fig. 2c; $a = 4\sigma$ Fig. 2d), smoothing and aliasing become clearly visible. In these cases, the Nyquist frequency falls below frequencies that still contain substantial variability.

4 Diffusion length estimation

To assess the impact of sampling on diffusion length estimates, we adopt a Monte Carlo approach, simulating multiple records across a range of sample intervals. Synthetic isotope records are generated by creating white noise, scaling its Fourier transform to match a target power spectrum, and applying the inverse transform to obtain depth-domain signals. We first simulate firn isotope profiles using Eq. 5 with $P_{X_0} = 0.32‰^2\text{m}$ and a typical diffusion length of $\sigma = 8$ cm. This value of P_{X_0} is taken from the short-timescale regime of the Dome C composite spectrum (Shaw et al., 2025). One hundred realisations are generated at high resolution (0.2 cm), after which sampling is simulated by binning and averaging over intervals a, ranging from 0.25σ to

https://doi.org/10.5194/egusphere-2026-3558
Preprint. Discussion started: 30 June 2026

2.5σ (10 values per realisation). Measurement noise (0.1‰ SD) is then added. The procedure is repeated for a deep-ice case using $P_{X_0}(f) = 0.15 f^{-2.2}$‰^{2}m and σ = 30 cm, allowing a coarser initial resolution of 1 cm.

We then estimate diffusion lengths from the power spectra of each record using both the conventional model fit (Eq. 5) and the new model fit (Eq. 21). We apply the Bayesian fitting approach of Shaw et al. (2024), which allows flexible forms of $P_{X_0}(f)$ and accounts for the gamma-distributed errors of Fourier-based PSD estimates (Bloomfield, 2004). For each sample interval, we compute the mean and standard deviation of the estimated diffusion lengths (Fig. 3).

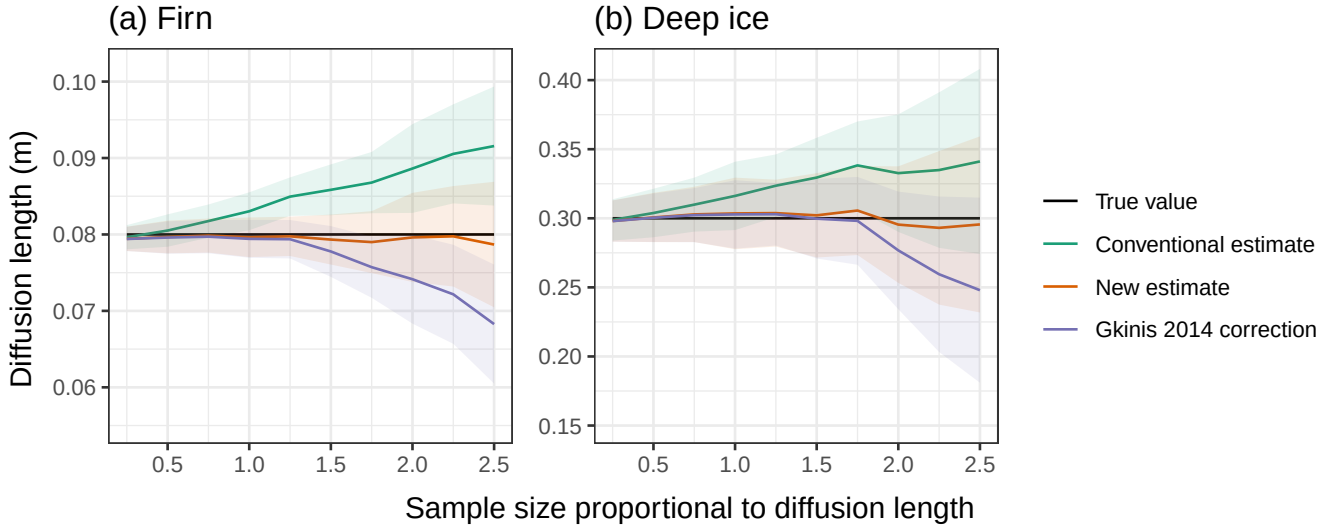

Figure 3. Comparison of diffusion length estimates obtained using the original method (green), the new method (orange), and the correction proposed by Gkinis et al. (2014) (purple), across different sample intervals. Values represent the mean of 100 realisations for **(a)** a white-noise firn record and **(b)** a power-law deep-ice record. Shaded regions indicate ± 1 standard deviation. The true diffusion length is shown in black for reference.

For the conventional model fit, diffusion length estimates remain unbiased with a low uncertainty up to a sampling resolution of $a \approx 0.5\sigma$, but become increasingly positively biased and show an increased uncertainty at coarser resolutions. By contrast, the new method shows no significant bias and accurately recovers the diffusion length across all tested sample resolutions, albeit with similar uncertainty to the conventional approach. These results are consistent for both firn and deep-ice scenarios. For comparison, we also include the correction method of Gkinis et al. (2014), which approximates the transfer function of a rectangular filter by that of a Gaussian filter. This approach effectively reduces bias up to $a \approx 1.5\sigma$, but over-corrects at larger sample intervals.

https://doi.org/10.5194/egusphere-2026-3558
Preprint. Discussion started: 30 June 2026

5 Discussion

In this study, we examine the effect of discrete sampling on diffused water isotope records using analytical expressions and Monte Carlo simulations. For sufficiently coarse sampling, additional smoothing similar to diffusion and spectral aliasing at high frequencies are introduced. Although small relative to the diffusion signal, these effects become more significant as the sample interval increases. In particular, the theoretical PSD of the diffused record before and after sampling begins to diverge for $a \gtrsim \sigma$ (Fig. 2). This is consistent with the results from the conventional diffusion length estimator, which become positively biased beyond a similar threshold (Fig. 3), as the additional smoothing from the rectangular sampling filter is misinterpreted as increased diffusion. The correction proposed by Gkinis et al. (2014) reduces this bias for $a \lesssim \sigma$, but over-corrects for coarser sampling. By explicitly incorporating sampling effects into the model fit, we obtain unbiased diffusion length estimates across all tested sample resolutions.

The magnitude of sampling effects depends on the sampling resolution and can therefore be mitigated by finer sampling. However, achieving sufficiently high sampling resolution is often difficult or impractical. In deep ice cores, the diffusion length typically increases rapidly in the firn to values of $\sigma \approx 8$ cm before firn diffusion ceases at pore close-off, at depths of approximately 50 – 110 m (Gkinis et al., 2021; Burr et al., 2018). The subsequent densification and layer thinning gradually reduces the diffusion length with depth until ice diffusion becomes dominant. As a result, large sections of deep ice, often exceeding 1000 m, can exhibit diffusion lengths below 5 cm, with minima of 2 – 3 cm (Gkinis et al., 2014; Grisart et al., 2022). To avoid bias in the conventional model fit, this would require sampling intervals below approximately 2.5 cm over these intervals. In contrast, the model proposed here does not require such high resolution, substantially reducing processing time and effort.

A similar argument applies to other settings where very small diffusion lengths are expected. Ice cores from high-accumulation sites often exhibit smaller maximum firn diffusion lengths than the typical 8 cm, as diffusion acts only over a limited time before pore close-off (Fernandoy et al., 2018). Small diffusion lengths can also occur at warmer sites, where refrozen ice layers inhibit vapour exchange within the firn.

In practice, ice cores are often first analysed at low sampling resolution, both in the field and during initial laboratory measurements, to provide a rapid overview of large-scale variability. Such coarse data typically preclude reliable diffusion length estimates using conventional methods, as sample intervals often exceed the diffusion length. The approach presented here enables unbiased diffusion length estimates even from such preliminary data. This provides early insight into isotopic signal preservation and can help guide the choice of sampling resolution for subsequent high-resolution analyses. In addition, it allows diffusion length estimation for legacy datasets or core sections only sampled at coarse resolutions, such as the GISP2 ice core (Stuiver and Grootes, 2000).

https://doi.org/10.5194/egusphere-2026-3558
Preprint. Discussion started: 30 June 2026

The results presented here are not limited to water isotopes. Any proxy record based on measurements averaged over discrete samples will be subject to similar smoothing and aliasing effects. Therefore, power spectral analyses of other ice-core proxies, as well as of other similarly sampled proxy archives, should account for the additional artifacts introduced by the sampling process.

6 Conclusions

In conclusion, we examined the effects of discrete sampling and sample resolution on diffused water isotope records. We show that discrete sampling introduces additional smoothing and spectral aliasing, which can bias diffusion length estimates when sample intervals exceed the diffusion length. By incorporating these sampling effects into the model fit, this bias is removed, yielding accurate diffusion length estimates across all tested sample resolutions. This unbiased estimator is particularly relevant for ice cores with small diffusion lengths and enables diffusion length analysis of low-resolution datasets that were previously unsuitable. The sampling effects identified here should be considered in future diffusion length estimation methods and are likely relevant for a broader range of paleoclimate proxy records. To facilitate application, we provide the code implementation of the proposed estimator to facilitate its adoption in future studies.

Code availability. The code used in this manuscript is available upon request.

Author contributions. TL and FS designed the study. TL contributed to the methods and results. TK contributed to the theory. FS performed the analysis and wrote the manuscript. All authors contributed to the discussion of the results and to the preparation of the final manuscript.

Competing interests. There are no competing interests

https://doi.org/10.5194/egusphere-2026-3558
Preprint. Discussion started: 30 June 2026

References

Bloomfield, P.: Fourier analysis of time series: an introduction, John Wiley & Sons, 2004.

Burr, A., Ballot, C., Lhuissier, P., Martinerie, P., Martin, C. L., and Philip, A.: Pore morphology of polar firn around closure revealed by X-ray tomography, The Cryosphere, 12, 2481–2500, 2018.

Cochran, W. T., Cooley, J. W., Favin, D. L., Helms, H. D., Kaenel, R. A., Lang, W. W., Maling, G. C., Nelson, D. E., Rader, C. M., and Welch, P. D.: What is the fast Fourier transform?, Proceedings of the IEEE, 55, 1664–1674, 1967.

Dansgaard, W.: Stable isotopes in precipitation, tellus, 16, 436–468, 1964.

EPICA community members: Eight glacial cycles from an Antarctic ice core, Nature, 429, 623–628, 2004.

Fernandoy, F., Tetzner, D., Meyer, H., Gacitúa, G., Hoffmann, K., Falk, U., Lambert, F., and MacDonell, S.: New insights into the use of stable water isotopes at the northern Antarctic Peninsula as a tool for regional climate studies, The Cryosphere, 12, 1069–1090, 2018.

Gkinis, V., Popp, T. J., Blunier, T., Bigler, M., Schüpbach, S., and Johnsen, S.: Water isotopic ratios from a continuously melted ice core sample, Atmospheric Measurement Techniques Discussions, 4, 4073–4104, 2011.

Gkinis, V., Simonsen, S. B., Buchardt, S. L., White, J., and Vinther, B. M.: Water isotope diffusion rates from the NorthGRIP ice core for the last 16,000 years–Glaciological and paleoclimatic implications, Earth and Planetary Science Letters, 405, 132–141, 2014.

Gkinis, V., Holme, C., Kahle, E. C., Stevens, M. C., Steig, E. J., and Vinther, B. M.: Numerical experiments on firn isotope diffusion with the Community Firn Model, Journal of Glaciology, 67, 450–472, 2021.

Grisart, A., Casado, M., Gkinis, V., Vinther, B., Naveau, P., Vrac, M., Laepple, T., Minster, B., Prié, F., Stenni, B., et al.: Sub-millennial climate variability from high-resolution water isotopes in the EPICA Dome C ice core, Climate of the Past, 18, 2289–2301, 2022.

Holme, C., Gkinis, V., and Vinther, B. M.: Molecular diffusion of stable water isotopes in polar firn as a proxy for past temperatures, Geochimica et Cosmochimica Acta, 225, 128–145, 2018.

Johnsen, S. J.: Stable isotope profiles compared with temperature profiles in firn with historical temperature records., 1977.

Johnsen, S. J., Clausen, H. B., Cuffey, K. M., Hoffmann, G., Schwander, J., and Creyts, T.: Diffusion of stable isotopes in polar firn and ice: the isotope effect in firn diffusion, in: Physics of ice core records, pp. 121–140, Hokkaido University Press, 2000.

Kahle, E. C., Holme, C., Jones, T. R., Gkinis, V., and Steig, E. J.: A generalized approach to estimating diffusion length of stable water isotopes from ice-core data, Journal of Geophysical Research: Earth Surface, 123, 2377–2391, 2018.

NGRIP members: High-resolution record of Northern Hemisphere climate extending into the last interglacial period, Nature, 431, 147–151, 2004.

Petit, J.-R., Jouzel, J., Raynaud, D., Barkov, N. I., Barnola, J.-M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., et al.: Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica, Nature, 399, 429–436, 1999.

Shaw, F., Dolman, A. M., Kunz, T., Gkinis, V., and Laepple, T.: Novel approach to estimate the water isotope diffusion length in deep ice cores with an application to Marine Isotope Stage 19 in the Dome C ice core, The Cryosphere, 18, 3685–3698, 2024.

Shaw, F., Münch, T., Gkinis, V., and Laepple, T.: The impact of measurement precision on the resolvable resolution of ice core water isotope reconstructions, The Cryosphere, 19, 4913–4928, 2025.

Simonsen, S. B., Johnsen, S., Popp, T., Vinther, B., Gkinis, V., and Steen-Larsen, H.: Past surface temperatures at the NorthGRIP drill site from the difference in firn diffusion of water isotopes, Climate of the Past, 7, 1327–1335, 2011.

Stuiver, M. and Grootes, P. M.: GISP2 oxygen isotope ratios, Quaternary Research, 53, 277–284, 2000.