

Q1. The principal advantage claimed for the proposed method is that it can determine D_w without observing the complete isotope-exchange process. However, the manuscript does not compare the characteristic-time method with conventional full-curve fitting using the same experimental datasets. The authors should apply both methods to the available experiments and compare the resulting diffusion coefficients, uncertainties, and required observation times.

5 Agreement with previous literature alone is not sufficient to validate the new retrieval method. Alternatively, the authors should explain well why such experiments on full-curve fitting has not been done, or why the current methodology is enough to support your statements.

We thank the reviewer for these detailed comments. In response, we have applied both the characteristic-time method and the conventional full-curve fitting method to the available experimental datasets and compared the retrieved diffusion coefficients, their uncertainties, and the required observation times. The additional analyses are presented below.

1.1 Comparison of retrievals and uncertainties between characteristic-time and full-curve fitting methods

We applied the characteristic-time (CT) method to nine experiments conducted at different relative humidities to determine the water diffusion coefficient, D_w , over 30%–45% RH. To further assess its performance relative to the conventional full-curve fitting (FCF) method, we also fitted the same nine experimental datasets using the FCF method and compared the resulting D_w values and retrieval uncertainties.

Both methods use the same D₂O fraction measurements. In the CT method, the experimental data are first smoothed, and the resulting curve is used to identify the characteristic time, from which D_w is calculated. We select the smoothing parameter objectively using blocked cross-validation. In the FCF method, the complete experimental curve, from the initial time to the end of the record, is fitted with the spherical diffusion model to obtain the optimal D_w . Thus, the two methods share the same physical model but differ in how they use the temporal information in the experimental data. The CT method primarily uses the time at which the experimental curve reaches the D₂O fraction corresponding to the characteristic time, whereas the FCF method uses the full temporal evolution of the D₂O fraction to retrieve D_w . Figures 1.1 and 1.2 show results from a representative experiment at 35% RH. Figure 1.1 presents the CT retrieval, with the smoothed curve shown in blue. Figure 1.2 presents the FCF retrieval, with the fitted curve shown in red.

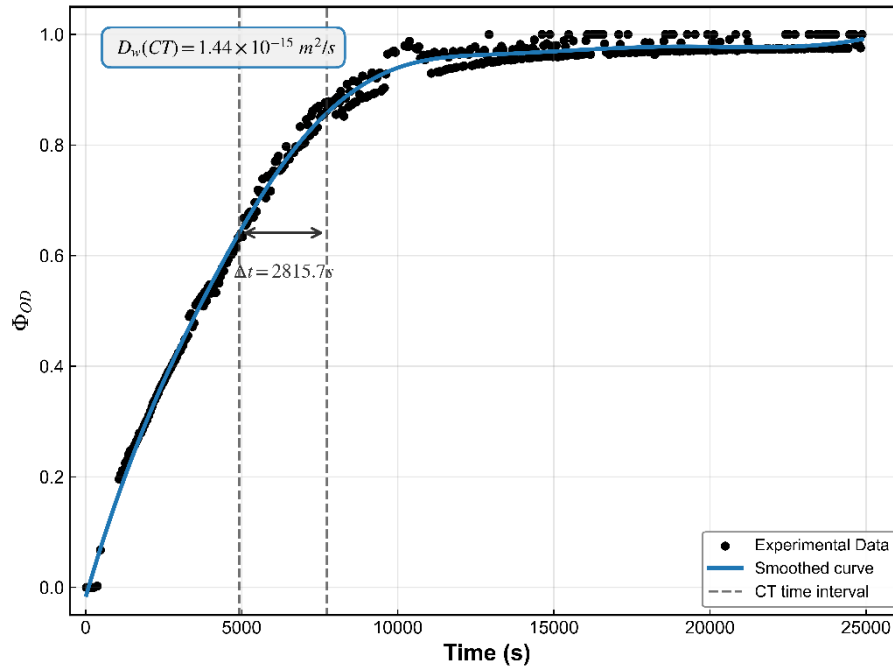


Figure 1.1: Smoothed curve, characteristic-time interval, and D_w retrieved using the CT method.

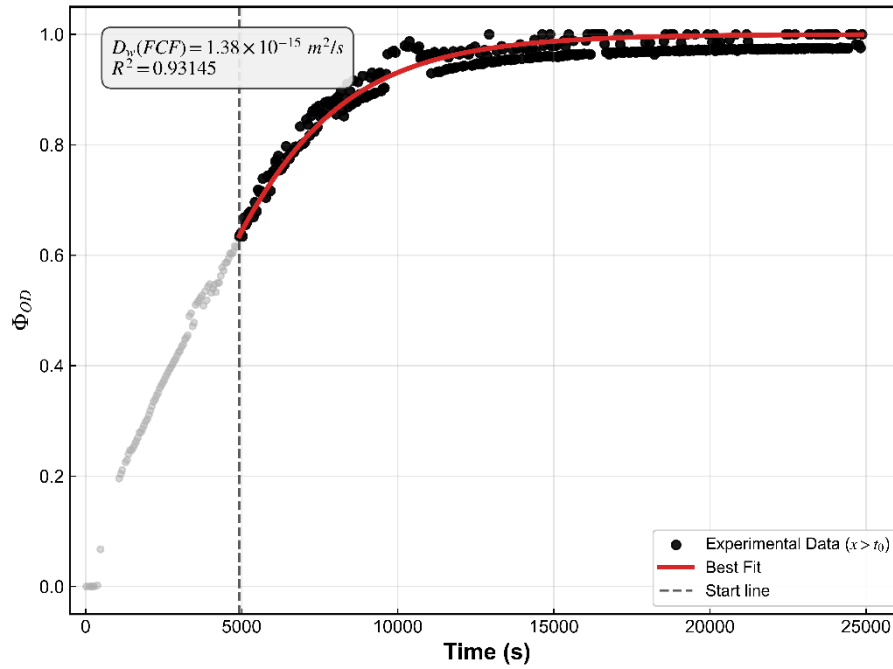


Figure 1.2: Fitted curve and D_w retrieved using the FCF method.

30 To compare the sensitivity of the two retrieval methods to fluctuations in the experimental data and assess the resulting uncertainty in D_w , we evaluated their robustness using residual bootstrap. First, we constructed a trend curve using the same smoothing procedure as in the CT method, and used the residuals relative to this curve to represent random fluctuations in the experimental data. The residuals were then resampled and added to the original trend curve to generate a new perturbed dataset for each bootstrap replicate. We applied both CT and FCF to the same perturbed dataset in each replicate. This
 35 ensures that the uncertainties are compared under identical data perturbations and avoids introducing additional variability from different random samples. For each experiment, 1000 bootstrap replicates were generated to obtain the bootstrap distributions of D_w for the CT and FCF methods. The 2.5th and 97.5th percentiles of each distribution were taken as the lower and upper bounds of the 95% confidence interval. Because D_w spans a wide range of orders of magnitude, we defined the equivalent half-width of the 95% confidence interval in $\log_{10}^{D_w}$ space to facilitate comparisons of uncertainty across
 40 experiments and between the two methods:

$$H_{95,\log} = \frac{1}{2}[\log_{10} D_{97.5} - \log_{10} D_{2.5}], \quad (1.1)$$

This half-width was then converted to a relative percentage:

$$H_{95,\%} = (10^{H_{95,\log}} - 1) \times 100\%. \quad (1.2)$$

Figure 1.3 compares the D_w retrieved using the CT and FCF methods for the nine experiments. The results generally follow
 45 the 1:1 relationship, with all data points lying within the $\pm 20\%$ deviation band or close to its boundaries. Although D_w spans a wide numerical range across the experimental conditions, the differences between the two methods show no clear systematic dependence on D_w . This indicates that the CT method yields absolute values of D_w similar to those obtained by FCF and captures the relative variation in D_w with RH across the experimental conditions, demonstrating broad consistency between the two methods.

50 Figure 1.4 further compares the half-widths of the 95% confidence intervals on the logarithmic scale for the two methods across the nine experiments, expressed as relative percentages, $H_{95,\%}$, to provide a more intuitive measure of retrieval uncertainty. Overall, the CT and FCF methods exhibit comparable uncertainties. The CT method yields a smaller $H_{95,\%}$ than FCF in six of the nine experiments. These results indicate that using only the characteristic-time information, rather than the full time-evolution curve as in FCF, does not reduce retrieval precision. Combined with the D_w comparison in Figure 3,
 55 these results indicate that the CT method can simplify data processing while yielding diffusion coefficients broadly consistent with those obtained by conventional FCF and comparable uncertainty levels.

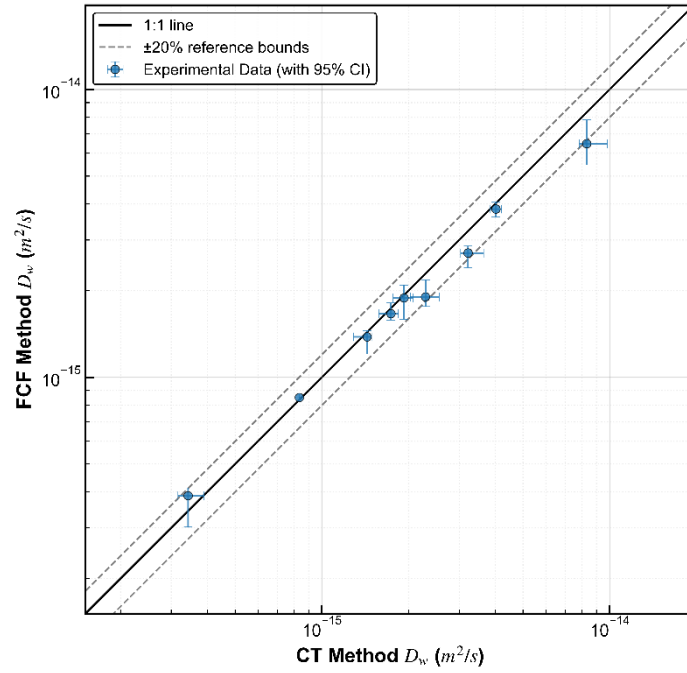


Figure 1.3: Comparison of D_w retrieved using the CT and FCF methods.

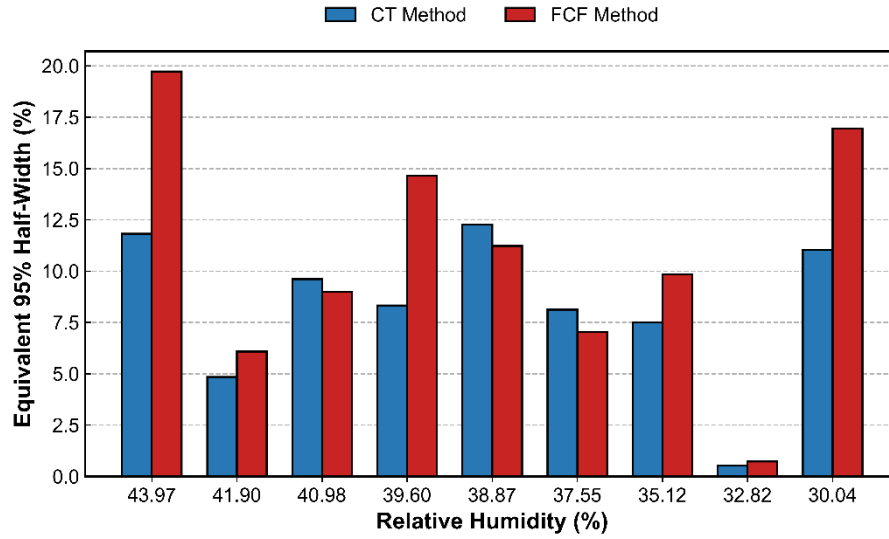


Figure 1.4: Comparison of retrieval uncertainties for the CT and FCF methods.

1.2 Comparison of the observation times required by CT and FCF

We further compared the observation durations required by the CT and FCF methods. For CT, the required duration follows directly from the definition of the characteristic time. The experimental record must extend from the initial time to the time at which the corresponding D_2O fraction is reached. The method therefore has a clearly defined time criterion. We denote the

65 observation duration required by CT as Δt_{CT} . For FCF, however, no clear, unified criterion currently defines the required observation duration. Because FCF retrieves the diffusion coefficient by fitting the full diffusion process, the required observation time must be assessed from the convergence of the fitting results. We therefore fitted progressively longer portions of each experimental record and compared the resulting diffusion coefficients to determine the minimum observation time needed for stable, reliable FCF retrievals.

70 Specifically, starting from the initial fitting time, t_0 , we progressively advanced the end time of the fitted data, increasing the observation duration by 50 s at each step. For each truncated dataset, we repeated the FCF retrieval and the uncertainty analysis described above to obtain $D_w(t_i)$ and its uncertainty, where t_i denotes the endpoint of the i th truncated record. Taking the retrieval based on the complete record as the reference, we constrained the minimum observation time required by FCF using the stability of both $D_w(t_i)$ and its uncertainty.

75 For the diffusion coefficient, the 95% confidence interval of D_w obtained by fitting the complete experimental record was used as the reference range. We defined t_{D_w} as the earliest time at which $D_w(t_i)$ enters this range and remains within it for all subsequent observation durations.

For uncertainty, each fit to a truncated record was evaluated using the bootstrap resampling procedure described above:

$$H_{95,log}(t_i) = \frac{1}{2} [\log_{10} D(t_i)_{97.5} - \log_{10} D(t_i)_{2.5}]. \quad (1.3)$$

80 Here, $D(t_i)_{2.5}$ and $D(t_i)_{97.5}$ are the 2.5th and 97.5th percentiles of the FCF bootstrap distribution for the current truncated record, respectively. The quantity $H_{95,log}(t_i)$ represents the half-width of the 95% confidence interval on the logarithmic scale and quantifies uncertainty. The uncertainty of the fit to the complete record is denoted by $H_{95,log}(t_{full})$, which is identical to $H_{95,log}$ as defined above, with the record duration added to the notation. To establish an allowable upper limit for uncertainty, we generated 2000 resamples from the 1000 bootstrap diffusion coefficients obtained using the complete record.

85 Each resample contained 1000 values, from which the half-width of the 95% confidence interval on the logarithmic scale was calculated using the definition above. This procedure yielded a distribution of 2000 half-width estimates. Their standard deviation, denoted by $SE(H_{95,log}(T_{full}))$, was used as the standard error of the uncertainty estimate. The allowable upper limit was then defined as the sum of the uncertainty for the complete record, $H_{95,log}(T_{full})$, and its standard error, $SE(H_{95,log}(T_{full}))$:

$$90 \quad H_{max} = H_{95,log}(T_{full}) + SE(H_{95,log}(T_{full})). \quad (1.4)$$

Here, H_{max} denotes the allowable upper limit for uncertainty. For the fits to records truncated at different times, we defined t_H as the earliest time at which $H_{95,log}(T_i)$ does not exceed H_{max} and this condition remains satisfied for all subsequent observation durations.

Finally, the later of t_{D_w} and t_H was taken as the earliest observation endpoint satisfying both criteria:

$$95 \quad t_{FCF} = \max(t_{D_w}, t_H). \quad (1.5)$$

Thus, we considered an observation time sufficient for reliable FCF retrieval only when both the retrieved diffusion coefficient and its uncertainty had stabilised. The minimum observation duration required for the FCF method was denoted as Δt_{FCF} and calculated as the difference between t_{FCF} and the initial time.

Figure 1.5 compares the observation durations required by CT and FCF at different RH values. In our observations, CT consistently required substantially shorter records than FCF. To quantify this advantage, we defined the fractional reduction in observation duration achieved by CT as:

$$\eta = \frac{\Delta t_{FCF} - \Delta t_{CT}}{\Delta t_{FCF}} \times 100\%. \quad (1.6)$$

Across the eight experiments where comparison was possible, CT reduced the required observation duration by 54.0%–73.8% relative to FCF, with a mean reduction of 65.7% and a median of 66.7%. Under the criteria adopted in this study, CT therefore saves approximately two-thirds of the observation time on average, with a reduction achieved in every experiment included in the comparison. The absolute time savings are particularly substantial for experiments involving slow isotope exchange. At 35.1% RH, FCF required 8200 s, whereas CT required only 2815.7 s, a reduction of 65.7%. At the lower RH of 32.8%, the required duration decreased from 235.0 min for FCF to 69.6 min for CT, a 70.4% reduction. Of the nine experiments conducted, the droplet at 30% RH escaped from the optical trap before isotope exchange was complete, terminating the observation prematurely. The subsequent record was therefore insufficient to establish the minimum sufficient observation duration for FCF using the criteria above, and this experiment was excluded from the time comparison. Nevertheless, the record acquired before escape from the optical trap covered the characteristic time required by CT and could still be used to retrieve the water diffusion coefficient. This example further illustrates that CT reduces the dependence of diffusion measurements on late-stage exchange data. As long as the record covers the characteristic time, CT can still extract diffusion information from an experiment that terminates prematurely.

Together with the consistent diffusion coefficients and comparable retrieval uncertainties reported above, this comparison shows that the CT method can substantially reduce the required record length while maintaining results broadly comparable to those obtained by FCF.

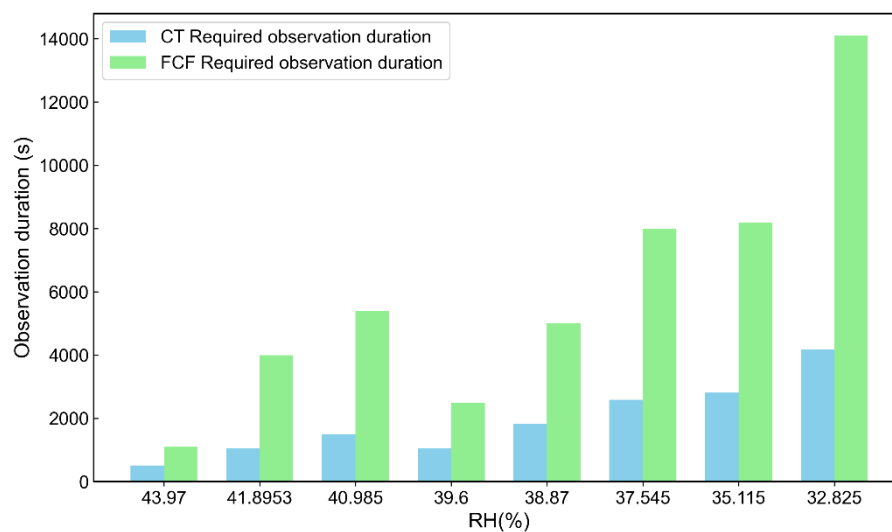


Figure 1.5: Comparison of the observation durations required by the CT and FCF methods.

It should be noted that, in this revision, we modified the procedure for selecting the smoothing parameter in the CT method and improved the determination of the time at which RH stabilises in response to comments from the other reviewer. Consequently, the final D_w values differ slightly from those in the original submission.

Q2. Section 5.1 primarily describes similarities and differences between the present results and previous studies, but does not sufficiently explain what these comparisons imply for the reliability of the proposed method. The discussion should focus more directly on the authors' results, methodological assumptions, and limitations. Claims regarding non-spherical systems, different experimental platforms, low temperatures, and RH conditions below the measured range should be moderated because these applications were not demonstrated in the present study.

We thank the reviewer for this suggestion. We have revised and expanded the Results and Discussion in Section 5 to focus more directly on our experimental results, methodological assumptions, and limitations, and to clarify how the comparisons inform the assessment of the proposed method's reliability. Comparison with previously reported water diffusion coefficients provides an external reference for evaluating the magnitude of our measurements and their variation with RH.

To further assess the CT retrievals, we added a comparison of CT and FCF results from the same experimental datasets, an analysis of uncertainties arising from fluctuations in the experimental data, and a quantitative assessment of the reduction in observation duration achieved by CT, as detailed in our response to Q1. These analyses provide a basis for evaluating the consistency and robustness of CT retrievals and their advantage in terms of observation time. We have also expanded the discussion of the basis for approximating the initial concentration profile with a quadratic function and adopting radial averaging, together with the associated conditions of applicability and limitations. The additions are presented below.

For the initial condition, idealising a zero initial concentration field requires establishing a stable external D₂O environment before isotope exchange begins within the droplet. In practice, however, D₂O replacement and the experimental system's response take time, and D₂O begins to enter the droplet before the external environment stabilises. We therefore take the time at which RH stabilises as the starting point for the retrieval and introduce a non-zero, non-uniform initial concentration profile so that the assumed constant surface concentration is consistent with the droplet's internal state at that time.

We represent the initial concentration profile with a quadratic function. For a spherically symmetric system, the initial concentration profile must have zero radial gradient at the droplet centre while retaining radial dependence to represent a non-uniform concentration distribution. A quadratic function is the lowest-order polynomial that satisfies both requirements. The two coefficients of the quadratic function can then be determined from the surface concentration and the measured radially averaged fraction. The normalised expression of the initial concentration profile is:

$$\frac{C(r,0)}{C_s} = \frac{3(1-\Phi_0)}{2R^2} \cdot r^2 + \frac{(3\Phi_0-1)}{2}. \quad (2.1)$$

Here, Φ_0 is the initial radially averaged D₂O fraction, R is the droplet radius, and C_s is the surface concentration. The initial concentration must remain non-negative at every point within the droplet, which requires $\frac{(3\Phi_0-1)}{2} \geq 0$, or $\Phi_0 \geq \frac{1}{3}$.

The initial fractions in all experiments satisfy this condition, so the assumed initial concentration profile remains non-negative throughout each droplet. However, the experiment did not directly measure the radial concentration profile. The quadratic function remains an approximation to the concentration distribution in the experimental droplet. This approximation incorporates the instrumental response time into the initial condition while satisfying the constraints imposed by the stable surface boundary, symmetry at the centre, and the measured initial fraction.

The choice between radial and volume averaging should reflect the measurement's spatial sampling characteristics. When the internal concentration is non-uniform, radial and volume averaging assign different weights to concentrations at different positions. They can therefore yield different average fractions and temporal evolution even for the same concentration field. As discussed above, the transverse extent of the laser beam in an optical tweezers system is typically smaller than the droplet diameter. Therefore, radial averaging better represents the composition sampled by the Raman signal. The optical tweezers provide stable three-dimensional confinement of the trapped droplet, which limits the displacement of the droplet from the beam axis under stable trapping conditions. To further examine how the retrieval would be affected if a small laser-beam offset from the droplet centre were present, we considered an offset corresponding to 20% of the droplet radius. Even for this comparatively large offset, the retrieved D_w changed only slightly, providing further support for the radial-averaging approximation under the present experimental conditions. The corresponding analysis and discussion are presented below and will be included in the Supplement of the revised manuscript.

2.1 Radial-averaging model for an off-centre laser beam

Although the droplet is confined as described above under ideal optical trapping conditions, weak gas-flow perturbations and other non-ideal experimental factors may still cause small, transient misalignment between the beam axis and the droplet centre. To evaluate the sensitivity of the radial-averaging treatment used in this study to such potential misalignment, we considered a deliberately large lateral displacement,

$$\beta = \frac{b}{R} = 0.2, \quad (2.2)$$

where β denotes the dimensionless displacement, b is the lateral distance between the beam axis and the droplet centre, and R is the droplet radius. This configuration was introduced solely as a relatively conservative sensitivity test to determine the extent to which this geometrical effect could influence the retrieved D_w , even when the beam axis is displaced from the droplet centre by as much as 20% of the droplet radius.

When the laser-beam axis is laterally displaced from the droplet centre by a distance b , the beam traverses the spherical droplet along a chord that does not pass through the droplet centre. The half-length of this chord is

$$L = \sqrt{R^2 - b^2} = R\sqrt{1 - \beta^2}. \quad (2.3)$$

Defining the z -axis along the direction of the laser beam, the radial distance from any point along the beam path to the droplet centre is

$$r = \sqrt{b^2 + z^2}. \quad (2.4)$$

Accordingly, under an off-centre sampling geometry, the averaged D_2O fraction represented by the Raman signal can be expressed as

$$\Phi_\beta(t) = \frac{1}{C_s} \frac{1}{2L} \int_{-L}^L C(\sqrt{b^2 + z^2}, t) dz. \quad (2.5)$$

where C_s is the concentration at the droplet surface. We retain the same spherical diffusion equation, constant surface concentration boundary condition, and quadratic initial radial concentration profile as in the main model, but reformulate the

initial condition to account for the off-centre sampling geometry described above. Based on the initial path-averaged fraction, $\Phi_{\beta,0}$, along the off-centre beam path, the corresponding initial condition is

$$C(r, 0) = C_s + \frac{3(1-\Phi_{\beta,0})C_s}{2(1-\beta^2)R^2} \cdot \left(\frac{r^2}{R^2} - 1\right). \quad (2.6)$$

The modified model therefore retains the same underlying physical assumptions as the model presented in the main text; the only change is the laser beam's sampling path through the spherical concentration field. When $\beta = 0$, the off-centre chord reduces to a diameter passing through the droplet centre, and the formulation correspondingly reduces to the original no-offset model used in the main text. Following the same procedure as in the main text, the spherically symmetric diffusion equation can be solved using the initial concentration profile together with the boundary condition, yielding the D_2O concentration distribution

$$\frac{C(r,t)}{C_s} = 1 + \frac{R}{r} \sum_{n=1}^{\infty} \frac{18(1-\Phi_{\beta,0})(-1)^n}{n^3\pi^3(1-\beta^2)} \sin\left(\frac{n\pi r}{R}\right) \exp\left(-\frac{n^2\pi^2 D_w t}{R^2}\right). \quad (2.7)$$

Integrating this concentration field along the off-centre beam path defined above, and introducing the geometric function

$$I_n(\beta) = \frac{1}{\sqrt{1-\beta^2}} \int_0^{\sqrt{1-\beta^2}} \frac{\sin[n\pi\sqrt{\beta^2+s^2}]}{\sqrt{\beta^2+s^2}} ds, \quad (2.8)$$

gives the path-averaged D_2O fraction under the displaced-beam condition:

$$\Phi_{\beta}(t) = 1 + \sum_{n=1}^{\infty} \frac{18(1-\Phi_{\beta,0})(-1)^n}{n^3\pi^3(1-\beta^2)} I_n(\beta) \exp\left(-\frac{n^2\pi^2 D_w t}{R^2}\right). \quad (2.9)$$

As in the original model, the characteristic time is defined as

$$\tau = \frac{R^2}{\pi^2 D_w}. \quad (2.10)$$

At $t = \tau$, the off-centre model retains a linear relationship between the D_2O fraction at the characteristic time and the initial fraction. The averaged D_2O fraction at the characteristic time is therefore

$$\Phi_{\beta}(\tau) = 1 + \sum_{n=1}^{\infty} \frac{18(1-\Phi_{\beta,0})(-1)^n}{n^3\pi^3(1-\beta^2)} I_n(\beta) \exp(-n^2). \quad (2.11)$$

For $\beta = 0.2$, numerical evaluation gives

$$\Phi_{0.2}(\tau) = 0.6142 + 0.3858\Phi_{0.2,0}. \quad (2.12)$$

2.2 Impact on experimentally retrieved D_w

We subsequently applied the characteristic-time relationship rederived for $\beta = 0.2$ to the full experimental dataset and recalculated D_w for each experiment. The resulting values were then compared with those obtained using the original model ($\beta = 0$).

As shown in Fig. 2.1, across the full range of RH conditions investigated, the relative differences in D_w between the 20% offset model and the no-offset model were approximately 1.0–5.0%, with all experimental cases remaining below 5%. Moreover, we observed no systematic increase or decrease in deviation as RH changed. These results indicate that, even when a relatively large lateral offset is imposed as a sensitivity test, the resulting change in the retrieved D_w remains limited.

220 In summary, the single-beam optical trap itself constrains a stably trapped droplet to remain close to the beam centre. On this basis, we further imposed a lateral offset as large as 20% of the droplet radius to test the robustness of the radial-averaging assumption under a deliberately non-ideal sampling geometry. Even under this relatively large prescribed displacement, the differences between the recalculated D_w values and those obtained using the no-offset model remained below 5% for all experiments. We therefore conclude that the radial-averaging approximation adopted in this study is insensitive to the small
225 beam-droplet misalignments that may occur in practice.

We added the corresponding sensitivity analysis and discussion to the Supplement of the revised manuscript.

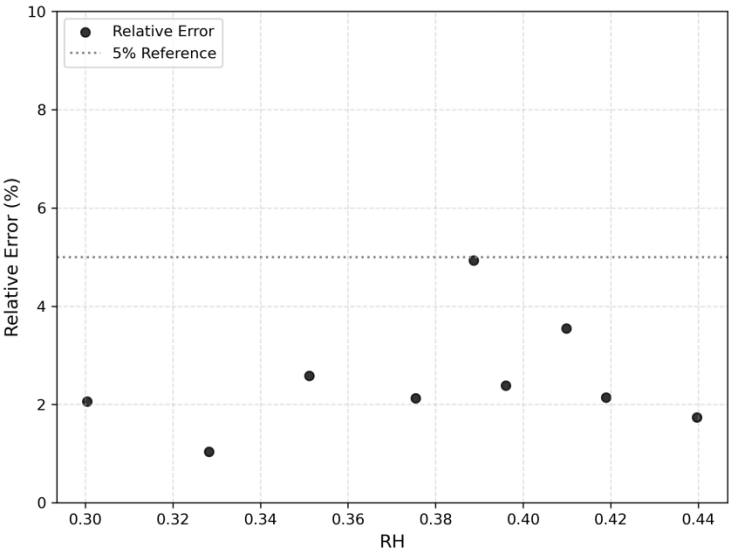


Figure 2.1: Relative difference in retrieved D_w between the lateral-offset and no-offset models.

Finally, we revised statements concerning non-spherical systems, different experimental platforms, and low temperatures
230 and RH conditions below the range measured in this study. The revised discussion places greater emphasis on the conditions that must be met when applying the method on other experimental platforms, and its applicability to these systems remains to be validated.

Q3. Figure 4 includes a curve labelled “Fit,” but the fitting method and fitted parameters are not explained. It is also unclear whether the reported D_w was derived from this fit or from the characteristic-time method. The authors should define the fit and use it to compare the two retrieval approaches where possible. The reduction in observation time should also be quantified rather than described only qualitatively.

We thank the reviewer for this important comment. We agree that this part was not clearly described in the original manuscript and that the label “Fit” in Figure 4 was inaccurate and could cause confusion about how D_w was retrieved.

In the original Figure 4, the red curve labelled “Fit” was not obtained by fitting the experimental data with the physical diffusion model. It was a smoothed curve derived directly from the experimental data points and used to identify the characteristic time required by the CT method. Thus, both the D_w value shown in the original figure and the D_w values reported in the manuscript were obtained by applying the CT method to the smoothed experimental curves, rather than by full-curve fitting. The CT retrieval of D_w is therefore independent of the full-curve-fitting procedure.

In the revised manuscript, we have corrected this potentially misleading description and no longer label the smoothed curve as “Fit”. We have also described the smoothing procedure in detail, including how the smoothing parameter is selected, to clearly define the full procedure from the experimental data points to the characteristic time and, ultimately, D_w . Details are provided in our response to Q1 and in the revised figure and caption.

Following the reviewer’s suggestion, we have also applied the conventional FCF method to the same experimental datasets and directly compared the D_w values retrieved by CT and FCF. This comparison further distinguishes the two retrieval procedures: CT determines D_w from the characteristic time identified on the smoothed experimental curve, whereas FCF determines the optimal D_w by fitting the full temporal evolution of the experimental data with the diffusion model. We have also quantitatively assessed the required observation times, as detailed in our response to Q1.

We also thank the reviewer for the minor comments, all of which have been addressed in the revised manuscript.