Ganguly, A., Padhya, V., Oza, H., **Strauch, G.**, Rajendrakumar, D. (2023):
New insights into diffusive kinetic fractionation during liquid condensation under supersaturated environment: an alternative approach for isotope tagging of ground-level water vapour
Isot. Environ. Health Stud. **59** (1), 1 - 26

http://dx.doi.org/10.1080/10256016.2022.2153126

New insights into diffusive kinetic fractionation during liquid condensation under supersaturated environment: an alternative approach for isotope tagging of ground level water vapour

Akash Ganguly[a,b], Virendra Padhya[a], Harsh Oza[a], Gerhard Strauch[c] and Rajendrakumar D. Deshpande[a]

[a]Geosciences Division, Physical Research Laboratory, Navrangpura, Ahmedabad, India; [b]Indian Institute of Technology Gandhinagar, Gandhinagar, India; [c]Department of Hydrogeology, Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany

Contact Akash Ganguly Email: akashg@prl.res.in

Supplemental data for this article can be accessed at DOI…

Abstract

Stable water isotopes in ground level vapour are key to estimating water exchange between geospheres. Their sampling, however, is limited to laser-absorption spectrometers and satellite observations, having inherent shortcomings. This study investigates diffusive kinetic fractionation during liquid condensation under supersaturated environment, providing a cost-effective, reliable way of sampling ground level vapour isotopes (^{18}O, ^{2}H). Experiments were undertaken at three different climatic zones in India with 'liquid' samples collected from condensation of ambient air at 0 °C. Simultaneously, pristine 'vapour' was sampled via cryogenic-trapping using liquid nitrogen–alcohol slush at –78 °C. The 'liquid' condensed under supersaturation was progressively more depleted in ^{18}O, and less enriched in ^{2}H than expected under equilibrium fractionation, with increasing degree of supersaturation expressed as saturation index (S_i). This study revealed: (1) S_i, molecular density, Rh, T together control the extent of isotopic kinetic fractionation. (2) The presence of diffusive concentration gradient inhibits the flow of heavier isotopes during liquid condensation. (3) The non-linear stochastic nature of the process cannot be adequately explained using physics-based model alone. The artificial neural network model is hence deployed to sample $\delta^{18}O$ ($\delta^{2}H$) within –0.24 ± 1.79 ‰ (0.53 ± 11.23 ‰) of true value. (4) The approach can be extended to ground-validate isotope-enabled general circulation models and satellite observations.

Keywords: Diffusive gradient; hydrogen-2; isotope hydrology; kinetic isotope fractionation; machine learning; oxygen-18; supersaturation; vapour sampling

1. Introduction

Atmospheric water vapour is the most important component of the hydrological cycle and the global climate system, as water molecules transfer and redistribute solar energy through latent heat of fusion, vaporization and condensation during phase change and transport through hydrological cycle [1,2]. Oxygen and hydrogen isotopes are the most important tracers widely used to understand various hydrological processes concerning the atmospheric component of hydrology. However, a major knowledge gap and a prominent limitation is the non-availability of adequate ground-based isotope data of water vapour. The sampling of ground level water vapour, for subsequent measurement of isotopic composition using isotope ratio mass spectrometry (IRMS) or laser absorption-based spectroscopy, requires cryogenic (using liquid nitrogen or dry ice) trapping of water vapour from ambient air-stream and longer sampling duration for pumping at a slow rate (~ 450 cm^3/min) in a laboratory set up to collect one sample [3]. On the other hand, direct isotope measurement of ambient atmospheric water vapour as well as its vertical profile can also be done continuously using either conventional cavity ringdown spectroscopy or off-axis integrated cavity output spectroscopy (OA-ICOS) [4–6]. However, both IRMS and laser spectroscopy measurements cannot be done simultaneously at a large number of places owing to their high cost, which is necessary to detect local variations and provide regional picture. Due to these limitations of ground-based measurements of isotopic composition of water vapour, there is tremendous dependence on satellite-based observations and model-based estimates. However, there are substantial inaccuracies in these estimates which are essentially required for computing the amount of continental recycling of precipitated water and upward transport of ground level water vapour to mix with tropospheric moisture [7–10].

The inaccuracies in the estimated isotopic composition of evaporated and transpired water vapour from different geographic regions in the world are carried forward into isotope-based estimation of evapotranspiration component of hydrological cycle regionally. Due to limitations in correctly estimating the isotopic composition of ground level water vapour in vegetated area, the dynamics of water vapour exchange between vegetation and atmosphere, and the hydrological processes concerning formation of dew, fog, cave condensation water etc. are poorly understood.

Inaccuracy in the estimated isotopic composition of ground level water vapour, which is also carried forward into the isotope based computational models, cannot be ignored anymore because the water vapour governs the water availability on earth and the global temperature, both of which affect the life on earth. The increasing surface air temperature increases the potential for evaporation, and consequent water vapour loading in atmosphere, the heat and water distribution over earth is expected to be affected, about which there is considerable uncertainty [11–13].

The deuterium isotopic composition (δ^2H) of atmospheric water vapour estimated by the Tropospheric Emission Spectrometer on board the Aura spacecraft of NASA has provided insight into the mechanism governing the recycling of water vapour [8,14]. It has been estimated that around 20 % of rainfall typically gets evaporated and fed back to the source whereas the values can go up to 50 % in case of convective clouds [7]. The estimated isotopic composition of atmospheric water vapour, based on which above inferences are drawn, are average values of the 550–800 hPa mean layer. The isotope values in these layers are susceptible to the maximum variations and are extremely sensitive and the estimated δ^2H values have a precision of around 10 ‰ in the tropics and 24 ‰ in case of the poles [8]. Moreover, the isotopic abundance ratio H^2HO/H_2O suffers due to spectral interference of H_2O on H^2HO due to pressure broadening. To add to this the space borne interferometers and spectrometers suffer from serious limitations in resolving thermal emissions in the short-wave infrared range thus failing to provide much insight into the composition of lower tropospheric vapour where much of the dynamic processes within the hydrological cycle take place. In addition to above limitations, the variation in relative humidity and its corresponding effect on the isotopic composition of ground level water vapour is not yet well understood [15,16].

The water isotopes incorporated into the atmospheric general circulation models (AGCM) have provided valuable insights into surface evaporation, condensation, super saturation and vertical distribution of isotopes [17–19]. These AGCM numerically represent and interpret the physical processes that may affect isotopic variation in vapour but encounter certain limitations that stem from biases in temperature or precipitation simulation [20]. The AGCMs work based on numerical simulations using certain approximations which fail for example when a sharp boundary is encountered such as e.g. the Tibetan Plateau [21]. Most AGCM simulate water vapour transport pathways by simulating the convective updrafts and downdrafts which involves solving the cloud microphysics using numerical integration. Other than being computationally expensive, these methods cannot accurately represent instantaneous localised atmospheric phenomena such as cloud burst, land fall of cyclonic storms etc., which has tremendous societal relevance.

Another possible option of estimating average monthly isotopic composition of ground level water vapour is from the monthly composite isotopic composition of rain, assuming that the ambient atmospheric water vapour is in isotopic equilibrium with rain [22,23]. Estimating isotopic composition of ground level vapour from that of average monthly rainfall has a serious drawback because it has been shown in a study in a semi-arid region that ground level water vapour tends to attain isotopic equilibrium with falling raindrops but is not necessarily in equilibrium. On cessation of rainfall, ground level vapour returns to its baseline isotopic composition [24]. It has also been observed in certain cases that the signature of tropospheric vapour differs significantly from that of precipitation in tropical areas indicating the significant role of evapotranspiration [7]. The biggest limitation of estimating vapour composition assuming it to be in isotopic equilibrium with rain is that this method is not applicable in the non-rainy season and hence creates a data gap.

The technological and practical constraints behind inaccuracies in satellite-based and model estimates, and limited ground measurements of vapour described above clearly highlight the urgent need of an alternative approach for accurate estimates of ground level water vapour.

One such initiative of isotopically characterizing the ground level water vapour was made under the aegis of a multi-institutional collaborative National Programme on Isotope Fingerprinting of Waters of India (IWIN National Programme) aimed at understanding hydrological processes concerning India which houses over 1391 million people (17.6 % of the world population). With a view to understand vapour source variation and rain–vapour interaction, ambient water vapour was sampled simultaneously using complete cryogenic trapping and liquid condensation on ice-cooled surface, for subsequent isotope analyses [24,25]. The liquid condensation of ambient vapour on ice cooled surface (see Section 2) is a simple, cost effective and easily replicable method for sampling ambient vapour for isotope analyses. The condensation method was, however, found to involve a kinetic fractionation due to molecular diffusion through a supersaturated layer just over the ice-cooled surface [25], similar to that involved in vapour deposition over ice [26]. The kinetic fractionation due to diffusion through supersaturated microenvironment generated during liquid condensation results in the isotopic composition of liquid condensate which is different from the true isotopic composition of vapour. Since the effective degree of supersaturation expressed as saturation index (S_i) during liquid condensation cannot be ascertained experimentally, and owing to the stochastic nature of the diffusive process, a simple linear regression-based approach [27] could not correct the vapour data. Thus, due to non-availability of any better correction method, even the un-corrected vapour data has been used in some of the studies [28,29].

In this study we have used an innovative approach of an artificial neural network (ANN) based machine learning (ML) model to estimate the isotopic composition of ground level vapour from liquid condensate collected on ice-cooled surface. We have also explained why the concept of saturation index expounded earlier by Jouzel and Merlivat [26] to explain the kinetic fractionation for vapour to solid phase change, and by Deshpande et al. [25] for vapour to liquid phase change, is not enough to estimate the unfractionated true values of vapour. We have also noticed that there can be different values of molecular densities for a given value of saturation index (S_i). Hence, temperature, relative humidity as well as the extent of supersaturation – all three are important in correcting the fractionating effects. The ANN and ML approach used in this study provides more accurate values of ground level vapour than possible through simple linear regression.

2. Materials and methods

The samples of ambient atmospheric water vapour were collected simultaneously by two different methods, namely, (1) complete cryogenic trapping; and (2) liquid condensation on ice-cooled surface.

2.1. Complete cryogenic trapping

In this method (Figure 1a), water vapour from a stream of ambient atmospheric air is cryogenically trapped in a glass condenser maintained at –78 °C by immersing in liquid nitrogen + alcohol slush. A stream of ambient air is diverted into a glass condenser immersed in a liquid nitrogen/alcohol slush maintained at –78 °C. The slush temperature of –78 °C is chosen so that the CO_2 from ambient air does not condense along with the water vapour. The air-flow rate is also maintained at around 450 cm^3/min to ensure that there is no isotopic fractionation due to preferential loss of lighter isotopic molecular species of water vapour. Since the entire vapour in the air stream is completely condensed, this method provides true isotopic composition of ambient water vapour without any isotope fractionation. Depending on the Rh and temperature of ambient air, it takes from 2 to 5 hours to collect about 2 ml of the unfractionated liquefied vapour by this method. Isotopic composition of this liquid sample represents the true isotopic composition of ambient vapour, and is now on referred to as ‘vapour’ in this paper. Detailed sampling procedure and sampling devices for complete cryogenic trapping is discussed by Deshpande et al. [25].

2.2. Liquid condensation on ice-cooled surface

In this method (Figure 1b), water vapour from the ambient atmospheric air is sampled by condensing it at 0 °C on a conical ice-cooled surface. An aluminium cone (diameter 15 cm, height 18 cm) is filled with ice cubes and covered with a PVC lid so that its external conical surface cools down to 0 °C and ambient water vapour condenses on it. This condensation is faster compared to cryogenic trapping. Depending on the Rh and temperature of ambient air it takes from a few tens of minutes to just less than an hour to collect as much as 15 ml of liquid condensate. The liquid condensation in this method takes place at ≈ 0 °C throughout the experiment, and involves kinetic fractionation related to diffusion through supersaturated layer and preferential condensation of isotopically lighter water molecule due to their higher diffusive velocities [25,26]. The liquid condensate obtained in this method is always isotopically fractionated and does not represent the true isotopic composition of ambient water vapour. This liquid condensate is referred to as ‘liquid’ in this study. Detailed sampling procedure and sampling devices for liquid condensation method is discussed by

Deshpande et al. [25].

The sampling for both ‘liquid’ as well as ‘vapour’ samples was performed simultaneously and for the same time duration. Moreover, the vapour sampling by complete cryogenic trapping and the liquid sampling by condensation on ice-cooled surface are done in the open environment, at ambient temperature, pressure and Rh conditions. Therefore, it is representative of the ambient environment without any laboratory artefact.

2.3. Sampling locations

The sampling was done at three stations (Ahmedabad: 23.02°N, 72.57°E; Hyderabad: 17.38°N, 78.49°E; and Roorkee: 29.85°N, 77.89°E;) located far away from each other in different parts of India and having different weather regimes and hydrometeorological aspects. The sampling at Ahmedabad, Hyderabad and Roorkee was carried out respectively by the Physical Research Laboratory (PRL), the National Geophysical Research Institute (NGRI) and the National Institute of Hydrology (NIH) under the aegis of IWIN National Programme [24,25,30] using the two procedures as described above.

The samples from Ahmedabad were collected mostly during May to September of 2005–2008 followed by from May 2013 to October 2014. Few samples were collected in the years 2009, 2010 and 2012 as well. At Hyderabad, the samples were collected mostly from July 2008 to October 2009. At Roorkee, the collection was done mostly during May to October 2009 and then again in January 2010.

Ahmedabad is a city in warm semi-arid western India with annual average rainfall of 800 mm received on ~ 35 rainy days during rainy season of southwest (SW) summer monsoon in the months of June to September [31]. Hyderabad is a city in south-eastern India and has predominantly a tropical wet and dry climate with average annual rainfall of 766 mm. It receives its 76 % of annual rain during southwest summer monsoon and the remaining 24 %

during northeast (NE) winter monsoon. Roorkee is in north India and has a warm and temperate climate. It receives annual rainfall of 1170 mm of which 59–84 % is received in SW monsoon and the remainder in NE winter monsoon.

A total of 538 pairs of vapour and liquid samples collected by above two methods have been used in this study.

2.4. Isotope analyses

The oxygen and hydrogen isotope analyses were carried out by standard gas equilibration method using an isotope ratio mass spectrometer in continuous flow mode of gas bench [32]. Based on analyses of multiple aliquots of secondary laboratory standards the precision of measurement was better than 0.10 ‰ for $\delta^{18}O$ and 1.00 ‰ for $\delta^{2}H$.

The oxygen and hydrogen isotopic composition is expressed in terms of per mil (‰) deviation of abundance ratio of heavier to lighter isotopes with reference to international standard reference material. Isotopic composition is defined in terms of δ (‰) notations as: [$\delta^{18}O$ or $\delta^{2}H = (R_{sample}/R_{std} - 1)$]. R_{sample} refers to the abundance ratio ($^{18}O/^{16}O$ or $^{2}H/^{1}H$) for the sample, and R_{std} refers to similar ratio for international standard reference material Vienna Standard Mean Ocean Water (VSMOW) [33,34].

2.5. Artificial neural network model

ANNs are a class of computational models which aim to mimic the biological working of neurons in brain by using a complex and dense network of artificial neurons, which can be perceived as a transfer function with a set of input being mapped to a particular output depending on the problem statement [35,36]. Inside the network, we have interconnected networks of hidden layers and neurons, which are sequentially inter-connected and produce output based on the input from the previous layer and the activation function, with which it assigns weights to each neural synapse, not unlike the coefficients in a fitting problem. The usage of an interconnected web coupled with use of specific transfer function helps in achieving higher degree of non-linearity and hence can capture more complex and stochastic processes.

In our case, we have a regression problem where capturing the underlying physics is of utmost significance. The variation of the vapour isotope values was previously modelled based on the equation provided by Jouzel and Merlivat [26] as follows:

$$1+\delta_l=\frac{D'\times(e_v\times(1+\delta_v)-e_i\times\frac{(1+\delta_l)}{\alpha_{eq}})}{D\times(e_v-e_i)}, \tag{1}$$

where δ_l and δ_v stand for the liquid and vapour δ values respectively, e_v and e_i denote the partial vapour pressure and saturated vapour pressure over water. The D and D' stand for the diffusivity coefficients for lighter and heavier isotopologues for oxygen.

To overcome the limitations in the S_i based models (Equation (1)), arising from the non-linearity between S_i values and molecular densities, which govern the fractionation, we have designed a simple ML model. A schematic diagram for the ANN model for the case of ^{18}O is shown in Figure 2. We incorporate this non-linearity in the physical system by inputting T, Rh, saturation and isotopic composition of liquid into the input layer of the neural network, in the form of Rh, T, $\frac{1}{\mathrm{Si}}$, $\frac{\delta_l}{\mathrm{Si}}$ and δ_l instead of simply as Rh, T, δ_l and S_i. The reason for using these specific terms in the ML model is explained in the Section 3.

3. Theory: Signatures of kinetic fractionation and theoretical consideration

In case of complete cryogenic trapping of water vapour from a stream of ambient air, all the H_2O molecules are converted from vapour to liquid phase and hence isotopic composition of liquefied water vapour is the same as that of ambient water vapour. In case of the isotope equilibrium between liquid and vapour phases, the condensed liquid is isotopically enriched in heavier isotopes (^{18}O and ^{2}H) compared to the vapour from which it is condensed. Isotope enrichment in liquid under equilibrium with vapour phase can be estimated using the equilibrium fractionation factor. In contrast, it is observed in our experiments that the liquid condensate is depleted in ^{18}O and less enriched in ^{2}H compared to vapour phase, than that expected under equilibrium condition. Consequently, the resultant liquid condensate has high *d*-excess, due to kinetic isotope fractionation involved in it.

Under the equilibrium condition ambient water vapour pressure equals the saturation vapour pressure and the rate of forward and backward reaction is similar. However, in the liquid condensation method of this study, in which condensation occurs on ice-cooled aluminium surface at 0 °C, the actual vapour pressure at ice-cooled surface is far more than the saturated vapour pressure at 0 °C. This is because the unsaturated ambient air at ambient temperature (~ 30 °C) becomes supersaturated with water vapour when it cools down to 0 °C. Thus, on the ice-cooled metallic surface there is a micron-thick boundary layer of air supersaturated with water vapour, and hence condensation of liquid from ambient air takes place under supersaturated environment. We wish to clarify that the micron-thick supersaturated boundary layer that we are referring to in this study is primarily governed by ground level T and Rh. This is different from the atmospheric boundary layer in which the isotopic composition is governed by soil moisture, transpiration and advected vapour isotopic signature. The atmospheric boundary layer dynamics may perturb the vapour isotopic signature at higher altitude in short time duration, but we are sampling ground level water vapour which does not change drastically at a short time (unless it starts raining). Furthermore, even if there is a significant impact in ground level vapour isotopic makeup in short time span, the dynamic processes will equally influence both the liquid and vapour samples. Since we are interested in studying the difference between isotopic compositions of vapour and liquid, and not their absolute isotopic composition, the sanctity of our experimental observations is still preserved.

The theory of kinetic effect associated with liquid condensation under supersaturated environment is discussed in detail by Deshpande et al. [25], similar to the theory for the solid condensation discussed by Jouzel and Merlivat [26]. It is noteworthy that both these studies explain why condensate formed under supersaturated environment has different isotopic composition compared to that expected under equilibrium condition. These two studies have also shown that the isotopic difference between vapour and the condensate formed under supersaturated condition strongly depends on the degree of supersaturation given by the saturation index (S_i):

$$S_i = \frac{partial\ pressure\ of\ water\ vapour\ at\ ambient\ temperature\ and\ relative\ humidity}{saturation\ vapour\ pressure\ at\ condensation\ temperature\ (0°C)} \qquad (2)$$

In spite of the works of Jouzel and Merlivat [26] and Deshpande et al. [25] it has not been possible to calculate back the true isotopic composition of vapour from the measured isotopic composition of condensate. Simple linear regression between isotopic composition of condensed liquid and cryogenically trapped vapour [27] can only provide seasonally variable empirical relationships between isotopic composition of liquid and vapour, with large deviations of regression based on predicted isotopic composition of vapour from its true value.

The scientific factors behind failure in estimating the true isotopic composition of vapour from liquid condensed from ambient vapour under supersaturation environment is that only the maximum possible value of the saturation index (S_i) can be computed from the measured values of ambient temperature, Rh and the condensation temperature (~ 0 °C). However, the actual effective degree of supersaturation prevalent at the condensing surface cannot be estimated due to uncertainties about prevalent condensation temperature. The actual temperature of condensation at condensing surface can be slightly more than the ice temperature (~ 0 °C) due to latent heat of condensation added to the ice-cooled condensation surface. Moreover, the ice inside the cone also melts slowly during experiment period which may slightly increase its temperature. Thus, temperature of condensation is expected to be slightly more than 0 °C, though the measured values of temperature inside the condensing cone were found to be between 0 to 0.5 °C in our experiments. Further, at molecular scale, removal of water molecules from vapour to liquid phase due to condensation would reduce the effective degree of supersaturation. Also, the effective diffusive velocities at prevalent degree of supersaturation cannot be estimated accurately. These are the scientific reasons why it has not been possible to calculate back theoretically the true isotopic composition of ambient vapour from its liquid condensate, in spite of the fact that this phenomenon is known to isotope geochemists [25,26].

Despite these uncertainties in precisely estimating degree of supersaturation, we have observed that the δ^{18}O and δ^{2}H values of liquid condensate collected on different days have a strong inverse relationship with saturation index (S_i). A plot of isotopic difference ($\Delta = \delta_{\text{liquid}} - \delta_{\text{vapour}}$) between liquid and vapour phases versus S_i shows that the isotopic difference between liquid and vapour increases with increasing values of S_i (Figure 3a,b). The liquid condensed under supersaturation was progressively more depleted in ^{18}O, and less enriched in ^{2}H than expected under equilibrium fractionation. A progressive trend of increasing *d*-excess in liquid when compared to vapour with increasing values of saturation index was also observed (Figure 3c).

In order to explain the underlying physics involved in the kinetic fractionation, the saturation index S_i (Equation (2)) can also be expressed as a ratio of the partial pressure of the vapour (e_v) in ambient air to the saturation vapour pressure (e_i) over water at condensing surface. For the sake of understanding this process, the immediate surrounding of the conical condensing surface can be schematically divided into three discrete zones A, B and C (Figure 4). At time $t = t_o$, the ice cubes were introduced into the cone hence reducing condensation temperature of condensing surface to ~ 0 °C, which in turn reduces saturation vapour pressure over water from e_i to e_{i3}. The magnitude of reduction in saturation vapour pressure (e_i) after introducing ice cubes depends on the difference between ambient and the condensation temperature (~ 0 °C). When saturation vapour pressure (e_i) drops below actual vapour pressure (e_v), the value of S_i increases to values greater than 1, which means that supersaturated condition ($e_{v3}>e_{i3}$, i.e., $S_i = e_{v3}/e_{i3} > 1$) is generated on the condensing surface (layer S in Figure 4). It is noteworthy that supersaturated condition generated at the condensation surface does not mean increase in the absolute humidity. It only means that there are more water molecules present in the air than it can hold at condensing temperature (~ 0 °C). Since the air cannot hold any more water vapour, H_2O molecules condense from vapour to liquid and tend to reduce the degree of supersaturation in zone A. Consequently, the actual water vapour content (absolute humidity) reduces just on the condensing surface because H_2O molecules condense from vapour to liquid and are removed from the environment. To compensate for this removal of water molecules from zone A, there is a mass flow from zone B to A and zone C to B, such that actual vapour pressure reduces from C to B to A ($e_{v3} > e_{v2} > e_{v1}$). Thus, we have a concentration gradient from zone C to zone A which drives vapour to come and condense on the ice cooled surface. In this process, isotopic water molecules ($H_2{}^{16}O$, H^2HO and $H_2{}^{18}O$) must diffuse from zone C to B to A to S. Isotopically lighter molecules have higher diffusivities and consequently, lighter molecules reach faster at condensing surface and get removed in liquid faster than heavier molecules. This diffusivity-based discrimination of molecules in favour of lighter mass becomes more prominent with increasing degree of supersaturation. This conceptual framework (Figure 4) explains the observation (Figure 3) of progressively increasing isotopic difference between liquid and vapour with increasing Si.

Jouzel and Merlivat [26] provided a model to explain the kinetic isotope effect in terms of the diffusive velocities of the various isotopologues through air. According to their formulation, the kinetic fractionation is related to ratio of diffusion coefficients of lighter to heavier isotopologues, the equilibrium fractionation factor, and the saturation index as:

$$\alpha_{kin}=\frac{S_i}{[\alpha_{equil}\times\frac{D}{D'}\times(S_i-1)]+1}, \tag{3}$$

where α_{equil} is the equilibrium fractionation factor, α_{kin} is the kinetic fractionation factor, D/D' represents the ratio of diffusivities for the lighter (^{16}O) to the heavier (^{18}O) isotope of oxygen or hydrogen (^{2}H, ^{1}H). For a given condensation temperature, α_{equil} and D/D' are both constants; therefore, the value of kinetic fractionation α_{kin} strongly depends on the value of S_i. Using this model, Deshpande et al. [25] have tried to explain the observed isotopic difference between vapour and liquid condensate using extrapolated D/D' values at 0° C.

The values of α_{equil} can be computed from the regression equation given by Horita and Wesolowski [37]. It is to be noted from Equation (1) above that the value of α_{kin} for any supersaturated condition (i.e. $S_i>1$) is always less than unity; therefore, total fractionation ($\alpha = R_l/R_v = \alpha_{kin} \cdot \alpha_{equil}$) involved in the liquid condensation works out to be less than equilibrium fractionation as explained in Deshpande et al. [25]. Consequently, liquid condensate is less enriched in heavy isotopes compared to ambient vapour than that expected under equilibrium condition. In highly supersaturated condition, beyond a certain critical value of S_{ic}, when $\alpha_{kin} < 1/\alpha_{equil}$, the liquid condensate will become isotopically depleted in heavier isotopes compared to ambient vapour. Since the values of α_{equil} and D/D' at 0 °C are close to 1 [25], the value of α_{kin} very strongly depends on S_i values. However, we observe the isotopic depletion of liquid is far more noticeable in ^{18}O than for ^{2}H, with more depleted values of liquid observed in ^{18}O even for lower values of S_i, while for ^{2}H this depletion is observed for higher values of S_i (~ >3). This behaviour can be attributed to the peculiar variation of diffusivity ratios D/D' for the different isotopologues ($H_2{}^{16}O$/ $H_2{}^{18}O$, $H_2{}^{16}O$/ H^2HO) at 0 °C. The computation of α_{kin} requires these diffusivity ratios for both molecular as well as turbulent diffusion. Several studies have experimentally and theoretically arrived at diffusivity ratios for molecular diffusion [38,39]. Luz et al. [39] and Merlivat [38] found that the diffusivity ratios for $^{18}O/^{16}O$ remain steady in the temperature range of 10–40 °C, whereas for $^{2}H/^{1}H$ there is a marked decrease in the diffusivity ratio with decreasing temperature hence resulting in a wider range of α_{kin} with variations in temperature. However, to the best of the authors' knowledge, few studies have attempted to investigate such kinetic fractionation processes occurring under supersaturated conditions for condensation of liquid from ambient vapour [25].

In our study, we have already highlighted the role played by a concentration gradient as well as a gradient in temperature starting from 0 °C at the surface of metallic cone ranging to ambient temperature the further away we go in causing preferential condensation of lighter isotopes in the collected liquid sample. Due to the unique difference in diffusivity characteristics between ^{2}H and ^{18}O, we can expect a wider variation in range of diffusivity ratios, α_{kin} and in turn the S_{ic} for ^{2}H when compared to ^{18}O, even by just assuming pure molecular diffusion due to variations in ambient temperature on different days as well as seasonally. Moreover, we also need to consider the possible impact of turbulent diffusion around the metallic cone. However, values of diffusivity ratios for turbulent diffusion are not readily available in literature. Deshpande et al. [25] have theoretically computed these ratios for the full spectrum (molecular to turbulent diffusion), and those ratios along with values obtained from sparse available literature have been considered for the purpose of this study [40]. The estimated range of S_{ic} for ^{18}O and ^{2}H are 1.24–1.42 and 3.27–31.50, respectively, thus implying a wide range of S_{ic} for ^{2}H [25]. Hence, a small change in trend from molecular diffusion towards turbulent transport is expected to have a far more pronounced effect on ^{2}H

as compared to ^{18}O. Moreover, ^{2}H has a higher value for the lower threshold of S_{ic}. This can explain why in the case of ^{2}H, we observe relatively more enriched isotopic signature in liquid for some samples even for higher values of S_i (Figure 3b).

In spite of past work [25,26], there has been an experimental limitation that the prevalent degree of supersaturation at the point of condensation in the micron thick boundary layer (layer S in Figure 4) formed during liquid condensation cannot be experimentally captured or accurately estimated. Moreover, the stochastic nature of diffusivity ratios makes it difficult for conventional physics-based model to explain the underlying processes or offer accurate predictions hence necessitating a simple ANN model.

To overcome the limitations in the S_i based models (Equation (1)), arising from the non-linearity between S_i values and molecular densities, which govern the fractionation, we have designed a simple ML model. We incorporate this non-linearity in the physical system by inputting the T, Rh, saturation and isotopic composition of liquid into the input layer of the neural network, in the form of Rh, T, $\frac{1}{\mathrm{Si}}$, $\frac{\delta_l}{\mathrm{Si}}$ and δ_l instead of simply as Rh, T, δ_l and S_i. The reason for using these specific terms in the ML model is explained in the following. We must understand that the aim of the ML model is to mimic the physical process happening in the natural system. Hence, inputting the parameters as Rh, T, δ_l and S_i directly may not represent the natural system well. To demonstrate this, we shall expand Equation (1) as:

Say, D/D' = A,

then Equation (1) reduces to [since $\frac{e_v}{e_i} = \mathrm{Si}$]

$$\delta_l + 1 = \frac{\alpha_{equil} \times S_i \times (\delta_v + 1)}{\alpha_{equil} \times \mathrm{A} \times (S_i - 1) + 1}$$

Solving for Y we get:

$$\delta_v = A \times \delta_l + \frac{\delta_l}{S_i} \times \left(\frac{1}{\alpha_{equil}} - A\right) + \frac{1}{S_i} \times \left(\frac{1}{\alpha_{equil}} - A\right) + A - 1$$

In the above equation, A (= D/D') and α_{equil} are both constants for a given temperature. Solving for δ_v in terms of δ_l and S_i, we get:

$$\delta_v = A \times \delta_l + B \times \left(\frac{\delta_l}{S_i}\right) + C \times \left(\frac{1}{S_i}\right) + K, \tag{4}$$

where A, B, C are terms made up of D/D' and α_{equil} and can be replaced with constants, and where K is a dimensionless constant. Hence, it is intuitive to use terms of the form $\frac{1}{\mathrm{Si}}$, $\frac{\delta_l}{\mathrm{Si}}$ and δ_l as they will better mimic the natural processes happening in the system. The physical interpretation for this is explained below:

There can be possible two extreme scenarios of fractionation during liquid condensation from vapour: (1) pure equilibrium fractionation, and (2) purely diffusive fractionation involving turbulent exchange.

(1) In case of equilibrium fractionation, the actual vapour pressure equals saturation vapour pressure i.e., $S_i = 1$, hence kinetic fractionation, $\alpha_{kin} = 1$ (Equation (3)). Therefore, the total fractionation is just the equilibrium fractionation ($\alpha_{total} = \alpha_{kin} \times \alpha_{equil} = \alpha_{equil}$). In this case, Equation (1) can be simplified to a linear form connecting both δ_l and δ_v as:

$$1+\delta_l = \alpha_{equil}\,(1+\delta_v) \tag{5}$$

This explains the need for the δ_l term as one of the input parameters.

(2) In case of purely diffusive fractionation in highly supersaturated condition, $S_i >> 1$ and $e_v >> e_i$. In this situation, Equation (1) can be reduced to:

$$1 + \delta_l \approx \frac{D' \times (1+\delta_v)}{D} \tag{6}$$

Here, the dependence of the isotopic makeup of resultant liquid is completely independent of S_i. Hence, this explains the term of the form of $\frac{1}{\mathrm{Si}}$ in Equation (4), since when $S_i >> 1$, this term will tend to vanish.

In our experiment, samples lie between the above two extreme cases which explains the term $\frac{\delta_l}{\mathrm{Si}}$ should be included in the formulation to compute true isotopic composition of vapour (δ_v) from that of liquid condensate (δ_l).As discussed earlier, S_i is a function of both T and Rh, although accurate measurement of S_i is subject to certain limitations. Moreover, even small pressure variations can lead to significant variation in environmental conditions, especially in humid areas. However, these pressure variations are mostly hydrostatic in nature and associated with temperature variation in the air column. Thus, temperature is the most important parameter, representing the corresponding minute pressure variations in the air column. Therefore, both T and Rh are inputted along with $\frac{1}{\mathrm{Si}}$, $\frac{\delta_l}{\mathrm{Si}}$ and δ_l in their respective form into the input layer of the neural network to help train the model to predict δ_v. Inputting data into the model in this form yielded much better model performance and convergence.

4. Results and discussions

4.1. Data cleaning

To ensure that ML algorithm is not adversely affected by any artefact from sampling, storage, or isotope analyses in mass spectrometer, it is essential to clean and filter out any spurious data from the 538 pairs of measured isotopic composition of vapour and liquid. In order to achieve this data quality, a unique criterion of data acceptance was defined for both vapour and liquid. If any of the vapour–liquid pair did not meet the required threshold of data acceptance, then the concerned pair was discarded to ensure the consistency of data. One of the parameters used is the *d*-excess (*d*-excess = δ^2H – 8 · δ^{18}O). The vapour samples collected from the Indian subcontinent are reported to have *d*-excess within an approximate range of 0-25 ‰ for vapour samples [25,41–44]. But some studies have reported even slightly higher *d*-excess values of vapour in rare cases. Hence, considering relaxation of 5 ‰ on either side of reported range (0–25 ‰) of *d*-excess, the range of –5 to 30 ‰ is considered as acceptable range for vapour samples. A total of 26 samples outside this range were detected and hence discarded.

Due to lack of sufficient published experimental data on isotope fractionation during liquid condensation occurring under supersaturated conditions, it is not possible to decide an appropriate range of acceptable *d*-excess values for filtering out spurious liquid samples based on experimental data. Therefore, we have used the theoretically estimated critical saturation index ($S_{i,c}$) as a limiting value beyond which the liquid condensate is expected to be depleted compared to the vapour [25]. The range of $S_{i,c}$ for stable oxygen and hydrogen isotopes has been theoretically estimated using diffusivity ratios considering both turbulent and molecular diffusion extremes and intermediate scenarios. The theoretically computed values of $S_{i,c}$ for oxygen isotope (^{18}O) have a narrow range from 1.24–1.42, whereas those for hydrogen range from 3.27–31.50 [25]. This implies that oxygen isotope fractionation under supersaturation condition is more consistent under different diffusive regimes (turbulent and molecular). Therefore, we have used the $S_{i,c}$ values for oxygen to reliably filter out spurious data points for liquid condensate. Those liquid samples were filtered out which were depleted in $\delta^{18}O$ compared to vapour even though the saturation index (S_i) was less than the lower threshold of its critical value ($S_{i,c}$ = 1.24) for ^{18}O. Similarly, those liquid samples were also filtered out, which were enriched in $\delta^{18}O$ even though saturation index (S_i) was more than the higher threshold of its critical value for ^{18}O ($S_{i,c}$ = 1.42). A total of 31 liquid samples were filtered out based on above consideration. After filtering out possibly spurious data of vapour and liquid as per above considerations, a total of 481 out of 538 data pairs were used to train and validate our ML model.

4.2. Model architecture, performance and new insights

Two separate ML algorithms, one each for ^{18}O and ^{2}H, have been used since the molecular weights of respective species and hence diffusivity ratios for heavier to lighter isotopes vary significantly for each one. Since the input features have varying ranges and standard deviations, variables having larger magnitude may affect the architecture of the model while updating weights, hence a standard scaler was used to standardize input data to have mean 0 and standard deviation of 1.

The whole dataset was split randomly into 70–30 % with 70 % data for training the model and 30 % for model testing and validation. Furthermore, while training the model, 5-fold cross validation was used in order to ensure model generalisation and stability and prevent overfitting [45,46]. GridSearchCV was employed for hyperparameter tuning by running maximum permutations and combinations of different model architecture with different number of hidden layers and varying activation functions [47–49]. The model with most robust and stable performance was chosen at the end. Root mean squared error as well as mean absolute error were used as metrics for model validation whereas mean squared error was chosen as the loss function. Adaptive momentum optimization (Adam) optimizer was chosen for this study [50]. The detailed model architecture and results are explained in Supplementary section, Table 1.

The input layer dimensions (5 nodes) represent the number of feature vectors used for training while the final layer in case of a regression problem always has just 1 layer with a linear activation function. Rectified linear unit (RELU) was chosen for the hidden layers since it gave best convergence [51,52]. We trained the model with small batch size with the number of iterations optimized using GridSearchCV. Since the non-linearity demonstrated for the oxygen isotope (^{18}O) with respect to variation in diffusivity and other experimental limitations or constraints was less, a shallower neural network was chosen and additional regularization was not deemed necessary (Supplementary section, Table 1).

For the case of hydrogen isotopes, due to the intrinsic non-linearity in the problem statement owing to large range in $S_{i,c}$, a denser more complex network gave the best performance. Additionally, we used Dropout regularization for the hidden layers [53,54]. We chose Dropout = 0.1 for purpose of regularization for the hidden layers.

The ML model was found to be robust when trained and validated on data from three stations with varying climatic regime. The tendency of mean error to lie near 0 indicates the absence of a strong negative or positive bias in either case. Hence, the mean error coupled with its standard deviation can effectively be used to get a general estimate of uncertainty in predicting vapour isotopic composition given the required input variables. Hence given a liquid sample, Rh and T, we can arrive at vapour isotope values within –0.24 ± 1.69 ‰ for $\delta^{18}O$ and 0.53 ± 10.19 ‰ for $\delta^{2}H$ which describe the mean of error distribution with 1 standard deviation (See Table 1 in Supplementary section). It must be noted that this uncertainty mentioned is only that of the ANN model and does not include the analytical uncertainty arising out of IRMS measurements. Hence accounting for the analytical uncertainty of 0.10 ‰ for $\delta^{18}O$ and 1.00 ‰ for $\delta^{2}H$, we can arrive at vapour isotope values within -0.24 ± 1.79 ‰ for $\delta^{18}O$ and 0.53 ± 11.23 ‰ for $\delta^{2}H$ (See Table 1 in Supplementary section). This level of accuracy of ground level water vapour provides significant improvements over satellite-based measurements.

Furthermore, the relatively comparable values of both RMSE (root mean squared error) and mean of absolute error for both models hint at a robust fit without overfitting. The RMSE for $\delta^{18}O$ is almost an order of magnitude lower than that for $\delta^{2}H$, but we must realise

RMSE is not robust to outliers and is also dependent on the range of the target variable. For $\delta^{18}O$, we have a range of 19.16 ‰ for vapour, for which RMSE around 1.7 ‰ is quite reliable. For $\delta^{2}H$, we have a range in vapour of 147.73 ‰, for which an RMSE score of ≈ 10 ‰ is still quite reasonable.

The predicted isotopic composition of vapour based on ANN models and the measured isotopic composition of vapour has been compared (Figure 8) for 145 data pairs (30 % of data set) used for testing and validation of models, which shows closeness of modelled and true values confirming the robustness of model performance and validity of this approach. The results confirm that the ML algorithms have fairly captured the stochastic non-linear processes. More than this, it is important to recognize the fact that the pair of ANN models used in this study have provided isotope values of ground level vapour which is an improvement over the available satellite-based measurements. This approach provides a solution for a long pending search of a suitable method to tag isotopically the ground level water vapour.

It is important to note that the saturation index S_i is a measure of degree of vapour saturation for a given T and Rh combination, but it does not represent the absolute abundance of H_2O molecules in the boundary layer. The molecular density within concentration gradient decides the degree of molecular resistance exerted on diffusing molecules. Although dependence of kinetic fractionation on saturation index is proven in this study (Figure 3), as also reported in earlier studies, its dependence on molecular density has not been studied so far. Therefore, it is necessary to understand the relationship between the kinetic fractionation, molecular density, and the saturation index S_i.

Figure 5 shows the progressive increase in the number density of H_2O molecules with increasing value of saturation index S_i ranging from 1 to 13 for different T–Rh combinations, with temperature ranges from 15 to 40 °C at an interval of 5 °C. For each of these temperature values, the number density of H_2O molecules increases with increasing values of Si, and the values of S_i increases with increasing temperature and increasing Rh (Figure 5). It is very important to notice now, since it was not explicitly shown in earlier studies, that for the same value of S_i (for example, Si = 4.5 in Figure 5b), the lower temperature and higher Rh conditions can generate higher molecular density and higher concentration gradient, offering higher resistance to diffusing molecules, which can result in a higher degree of diffusive kinetic fractionation under supersaturated environment.

With a view to demonstrate the higher degree of kinetic fractionation under low temperature and high Rh conditions for a given range of S_i, the dataset with S_i values ranging from 3.5 to 5.5 (316 liquid–vapour pairs, the highest density of samples in S_i wise distribution) was chosen. Furthermore, a still narrow bin of S_i from 3.5 to 4.5 was also constructed to study kinetic fractionation for samples having similar degree of supersaturation and hence identify other contributing factors. The isotopic difference ($\Delta = \delta_{\text{liquid}} - \delta_{\text{vapour}}$) between liquid and vapour for each pair of samples is plotted against its temperature and Rh for S_i range of 3.5 to 4.5 in Figure 6. The similar plot for the broader S_i range of 3.5 to 5.5 is provided in the Supplementary material. It is obvious from Figure 6 that, even within a narrow range of S_i, the isotopic difference ($\Delta = \delta_{\text{liquid}} - \delta_{\text{vapour}}$) is more pronounced under low temperature and high Rh conditions. This confirms higher degree of kinetic fractionation under low temperature and high Rh combinations for a given narrow range of S_i. It is also noteworthy that the degree of supersaturation may be similar but the number density of molecules could be quite different (Figure 5).

To further confirm that the observed increased isotopic difference ($\Delta = \delta_{\text{liquid}} - \delta_{\text{vapour}}$) under low temperature–high Rh conditions (Figure 6) is due to pronounced kinetic fractionation, and not due to change in vapour composition, the isotopic difference is plotted against the isotopic composition of liquid for S_i ranges of 3.5 to 4.5 (Figure 7), and of 3.5 to

5.5 (Supplementary material). Both cases show a positive correlation between the isotopic difference and the isotopic composition of liquid. A similar plot for isotopic difference versus vapour isotopic composition (Supplementary material) shows no correlation thus indicating that the initial isotopic composition of vapour plays no role in determining extent of kinetic fractionation. This proves that even for similar degree of supersaturation, we observe varying degree of kinetic fractionation, which is governed by the gradient of absolute molecular abundance in the diffusive layers.

Several past studies have attempted to study kinetic diffusion-based fractionation occurring during phase change from vapour to snow, liquid to ice and liquid to vapour [26,40,55–62]. However, to the best of the authors' knowledge, few studies have tried to observe and understand such processes occurring in liquid condensation from vapour under supersaturated conditions. Jouzel and Merlivat [26] noticed conspicuous effect of kinetic fractionation during deposition of snow from vapour under supersaturated conditions. Similar kinetic fractionation during freezing of water to ice owing to differences in diffusivity ratios has also been reported [61,62]. Several studies have attempted to explain the diffusion based kinetic fractionation during evaporation from liquid to vapour while emphasising the role played by a concentration gradient as well as turbulent diffusion [57,63–65]. Recent studies involving state-of-the-art quantum mechanical calculations have also made significant breakthrough in bridging the gap between theoretical understanding and experimental observations for the case of molecular diffusion [66]. While the role of molecular diffusion has been studied in detail, the role of turbulent diffusion transport is not yet well understood and is usually assumed to not impart any isotopic fractionation in the turbulent sublayer [67]. However, the experimental results from this study cannot be arrived at by assuming only equilibrium and molecular diffusion-based isotopic fractionation, hence implying a strong role played by turbulent diffusion under supersaturated conditions, in imparting isotopic fractionation. The turbulent diffusive processes are very difficult to characterize and are extremely stochastic in nature, which is probably why the physical based models used in past studies have failed to accurately predict the extent of such kinetic fractionation. Various studies have attempted to address the issue of isotope fractionation in clouds under supersaturated conditions either by using physics-based models or trying a linear fit to address the non-linear deviations in isotopes. This study does not only highlight the limitations of earlier approach [68–71] but demonstrates the robustness of the ML model approach in capturing a non-linear stochastic process, with accurate predictions on data taken from three sites having vastly differing climate and geographical setup.

Thus, it must be noted that the processes driving isotopic composition of the liquid are dependent only on the degree of supersaturation and the consequent molecular density generated in the micron thin boundary layer over the ice-cooled surface. They are independent of regional characteristics and weather conditions. In other words, independent of the geographical location, whenever a particular degree of supersaturation is generated on the condensing surface, there will be corresponding kinetic fractionation. It is important to note that the physical basis for the kinetic fractionation, even in the Jouzel and Merlivat [26] model, is also the same, i.e. the degree of supersaturation (saturation index Si). They have explained the observed kinetic fractionation during vapour deposition in terms of the degree of supersaturation in the environment over the ice. In our study, it is the degree of supersaturation over the ice-cooled surface. In addition, we have also introduced the concept of number density of H_2O molecules, which was not previously considered by Jouzel and Merlivat [26]. Figures 5 to 7 provide new insights about specific aspects such as molecular density, concentration gradient, temperature–Rh combination and liquid–vapour isotopic differences associated with the kinetic fractionation under supersaturated environment. The important new insight emerging from our experiments is that the kinetic fractionation under supersaturated environment is primarily governed by the molecular density and the concentration gradient that develops on the condensing surface. The magnitude of molecular density and the concentration gradient depends on the combination of T and Rh (Figure 5) which turns out to behave in a non-linear and stochastic way which cannot be captured accurately by models dependent on S_i alone. It is precisely for this reason that so far it has not been possible to estimate accurately the true isotopic composition of vapour from the liquid condensate.

To overcome the above problem a simple machine learning algorithm based on ANN with backpropagation was deployed to map water vapour values to liquid condensate values systematically and incorporate the non-linearities, which a simple physics-driven model or multi-regression statistical models are unable to achieve.

5. Summary and conclusions

Isotope tagging of ground level water vapour is very important to understand hydrological processes in different geographical regions having varied eco-hydrological, agro-climatic and water resource situations. In particular, the current level of understanding about the recycling of precipitated water, and transport and addition of ground level water vapour into rain-forming cloud systems, is restricted because isotopic composition of ground level water vapour is not accurately represented in the isotope based AGCM.

There are technological and infrastructural limitations in obtaining reliable isotopic signatures of ground level water vapour from ground based or remotely sensed observations. In this study a simple, cost efficient and reliable methodology is discussed, in which ambient water vapour is collected by liquid condensation on ice-cooled metallic surface and cryogenic trapping, and using an ANN based ML model for computing the true isotopic composition of ground level water vapour from measured isotopic composition of liquid condensate.

This study has significant implications in improving the scientific understanding of several natural processes like cloud bursts, dew formation, water in ecological pool, recycling of ground level moisture through evaporation and transpiration, which requires accurate tagging of outgoing ground level flux as well as precipitation resultant from condensation under supersaturated environment. The isotope fractionation under supersaturated environment, for example during cloud burst or thunder storm, or dew formation where prior studies have used either physics-based models or tried to account for the observed excess depletion in isotopes using simple linear fits could not account for the stochastic nature of the variation in the natural system.

Isotopically tagging the ground level water vapour near the soil level (dominated by the soil moisture) and within vegetation canopy (dominated by the transpired vapour) has been a challenge because ambient moisture is a mixture of soil moisture, transpired vapour and the advected vapour. To be able to isotopically tag these vapour sources it is necessary to sample the vapour from appropriate heights above the ground so that vapour dominated by the three different sources is isotopically characterised. It may be noted that at present such studies sample soil water, twig xylem water and rainwater to study soil–plant–atmosphere continuum [72,73]. This study offers an approach in which the experimental setup can be placed at different heights to sample the vapour for isotope analyses, which is very important in advancing the current understanding about the soil–plant–atmosphere continuum.

The robustness of the ML model in predicting vapour isotope values in three regions with different climatic conditions demonstrated in this study offers an extremely simple and cost-effective method of sampling ground level water vapour isotopes, which was difficult earlier due to high cost and laboratory infrastructural requirement especially in remote areas, and particularly for developing countries. The present study offers a solution to these limitations since it only involves recording Rh, T and collecting water samples using a simple metal cone and ice. Further, the sampling method coupled with the ANN-based method developed in this study could have a wide variety of applicability, ranging from ground-validation of AGCM and understanding of the underlying hydrological cycle, to modelling of supersaturated fractionation of clouds under supersaturated conditions.

This approach could also help validate satellite-based observations of water vapour isotopes and also help overcome the limited network of δ^{18}O measurements in vapour since satellite-derived measurements are unable to sample oxygen isotopes. However, owing to the limitations in satellite-based techniques in measuring lower tropospheric stable water isotopes in vapour, this method can only be employed in regions of high elevation where the satellite measurements are viable (between 500–800 hPa). Moreover, the uncertainties reported can be further improved upon by addition of more data points from future continued experiments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The work reported here is carried out under the aegis of a National Programme on Isotope Fingerprinting of Waters of India (IWIN) which was initially funded jointly by the Department of Science and Technology (DST), Government of India, vide Grant No. IR/S4/ESF-05/2004, and the Physical Research Laboratory (PRL), a Unit under the Department of Space, Government of India. The IWIN National Programme is currently sustained exclusively by the PRL.

Data

The research data base is attached as a Supplement and can be used to reproduce the results. The code used to perform this study can also be shared upon request.

Authors’ contributions

RDD conceived the initial problem statement and the significance of the same. AG was responsible for the data analyses, modelling and scientific evaluation. VP was responsible for processing the samples. GS, RDD and AG drafted the manuscript. All authors contributed in finalizing the manuscript.

Disclaimer

An earlier version of this manuscript is available in the form of pre-print at SSRN pre-print server [72].

References

1] Lieberman RS, Ortland DA, Yarosh ES. Climatology and interannual variability of diurnal water vapor heating. J Geophys Res Atmos. 2003;108(D3):4123.

[2] Hoffmann G, Cuntz M, Jouzel J, et al. how much climatic information do water isotopes contain? In: Aggarwal PK, Gat JR, Froehlich KFO, editors. Isotopes in the water cycle: Past, present and future of a developing science. Dordrecht: Springer Netherlands; 2005. p. 303–320.

[3] Gupta SK, Deshpande RD. Synoptic hydrology of India from the data of isotopes in precipitation. Curr Sci. 2003;85(11):1591–1595.

[4] Aemisegger F, Pfahl S, Sodemann H, et al. Deuterium excess as a proxy for continental moisture recycling and plant transpiration. Atmos Chem Phys. 2014;14(8):4029–4054.

[5] Bonne J-L, Behrens M, Meyer H, et al. Resolving the controls of water vapour isotopes in the Atlantic sector. Nat Commun. 2019;10(1):1632.

[6] Steen-Larsen HC, Sveinbjörnsdottir AE, Peters A, et al. Climatic controls on water vapor deuterium excess in the marine boundary layer of the North Atlantic based on 500 days of in situ, continuous measurements. Atmos Chem Phys. 2014;14(15):7741–7756.

[7] Worden J, Noone D, Bowman K, et al. Importance of rain evaporation and continental convection in the tropical water cycle. Nature. 2007;445(7127):528–532.

[8] Worden J, Bowman K, Noone D, et al. Tropospheric Emission Spectrometer observations of the tropospheric HDO/H_2O ratio: Estimation approach and characterization. J Geophys Res Atmos. 2006;111:D16309.

[9] Brown D, Worden J, Noone D. Comparison of atmospheric hydrology over convective continental regions using water vapor isotope measurements from space. J Geophys Res Atmos. 2008;113:D15124.

[10] Yoshimura K, Frankenberg C, Lee J, et al. Comparison of an isotopic atmospheric general circulation model with new quasi-global satellite measurements of water vapor isotopologues. J Geophys Res Atmos. 2011;116:D19118.

[11] Lawrence JR, Gedzelman SD, Dexheimer D, et al. Stable isotopic composition of water vapor in the tropics. J Geophys Res Atmos. 2004;109:D06115.

[12] Wild M, Ohmura A, Gilgen H, et al. On the consistency of trends in radiation and temperature records and implications for the global hydrological cycle. Geophys Res Lett. 2004;31:L11201.

[13] Oza H, Ganguly A, Padhya V, et al. Hydrometeorological processes and evaporation from falling rain in Indian sub-continent: Insights from stable isotopes and meteorological parameters. J Hydrol. 2020;591:125601.

[14] Beer R, Glavich TA, Rider DM. Tropospheric emission spectrometer for the Earth Observing System's Aura satellite. Appl Opt. 2001;40(15):2356.

[15] Trenberth KE, Dai A, Rasmussen RM, et al. The changing character of precipitation. Bull Am Meteorol Soc. 2003;84(9):1205–1218.

[16] Bosilovich MG, Schubert SD, Walker GK. Global changes of the water cycle intensity. J Climate. 2005;18(10):1591–1608.

[17] Werner M, Langebroek PM, Carlsen T, et al. Stable water isotopes in the ECHAM5 general circulation model: Toward high-resolution isotope modeling on a global scale. J Geophys Res Atmos. 2011 2011;116:D15109.

[18] Uemura R, Yonezawa N, Yoshimura K, et al. Factors controlling isotopic composition of precipitation on Okinawa Island, Japan: Implications for paleoclimate reconstruction in the East Asian Monsoon region. J Hydrol. 2012;475:314–322.

[19] Xi X. A review of water isotopes in atmospheric general circulation models: Recent advances and future prospects. Int J Atmos Sci. 2014;2014:250920.

[20] Noone D, Sturm C. Comprehensive dynamical models of global and regional water isotope distributions. In: West JB, Bowen GJ, Dawson TE, et al., editors. Isoscapes: Understanding movement, pattern, and process on Earth through isotope mapping. Dordrecht: Springer Netherlands; 2010. p. 195–219.

[21] Yao T, Masson-Delmotte V, Gao J, et al. A review of climatic controls on δ^{18}O in precipitation over the Tibetan Plateau: Observations and simulations. Rev Geophys. 2013;51(4):525–548.

[22] Araguás-Araguás L, Froehlich K, Rozanski K. Deuterium and oxygen-18 isotope composition of precipitation and atmospheric moisture. Hydrol Process. 2000;14(8):1341–1355.

[23] Gammons CH, Poulson SR, Pellicori DA, et al. The hydrogen and oxygen isotopic composition of precipitation, evaporated mine water, and river water in Montana, USA. J Hydrol. 2006;328(1):319–330.

[24] Deshpande RD, Maurya AS, Kumar B, et al. Rain–vapor interaction and vapor source identification using stable isotopes from semiarid western India. J Geophys Res Atmos. 2010;115:D23311.

[25] Deshpande R, Maurya A, Kumar B, et al. Kinetic fractionation of water isotopes during liquid condensation under super-saturated condition. Geochim Cosmochim Acta. 2013;100:60–72.

[26] Jouzel J, Merlivat L. Deuterium and oxygen 18 in precipitation: Modeling of the isotopic effects during snow formation. J Geophys Res Atmos. 1984;89(D7):11749–11757.

[27] Purushothaman P, Rao MS, Kumar B, et al. Comparison of two methods for ground level vapour sampling and influence of meteorological parameters on its stable isotopic composition at Roorkee, India. Hydrol Process. 2014;28(3):882–894.

[28] Krishan G, Rao MS, Kumar B. Study of climatological conditions using isotopic signature of air moisture at Roorkee, Uttarakhand, India. In: Proceedings India Water Week 2012 – Water, Energy and Food Security: Call for Solutions, 10–14 April 2012, New Delhi; 2012.

[29] Saranya P, Krishan G, Rao MS, et al. Controls on water vapor isotopes over Roorkee, India: Impact of convective activities and depression systems. J Hydrol. 2018;557:679–687.

[30] Deshpande RD, Gupta SK. Oxygen and hydrogen isotopes in hydrological cycle: new data from IWIN national programme. Proc Indian Nat Sci Acad. 2012;78(3):321–331.

[31] Oza H, Padhya V, Ganguly A, et al. Hydrometeorological processes in semi-arid western India: insights from long term isotope record of daily precipitation. Clim Dyn. 2020;54(5):2745–2757.

[32] Maurya AS, Shah M, Deshpande RD, et al., editors. Protocol for δ^{18}O and δD analyses of water sample using Delta V plus IRMS in CF Mode with Gas Bench II for IWIN National Programme at PRL. Ahmedabad: Indian Society for Mass Spectrometry Hyderabad; 2009.

[33] Gat JR. The isotopes of hydrogen and oxygen in precipitation. In: Fritz P, Fontes JC, editors. Handbook of environmental isotope geochemistry, Vol. 1A. Amsterdam: Elsevier; 1980. p. 21–47.

[34] Clark ID, Fritz P. Environmental isotopes in hydrogeology. Boca Raton: CRC Press; 1997.

[35] Xin Y. Evolving artificial neural networks. Proc IEEE. 1999;87(9):1423–1447.

[36] Sathya P, Malathi L. Classification and segmentation in satellite imagery using back propagation algorithm of ANN and K-means algorithm. Int J Mach Learn. 2011;1(4):422–426.

[37] Horita J, Wesolowski DJ. Liquid–vapor fractionation of oxygen and hydrogen isotopes of water from the freezing to the critical temperature. Geochim Cosmochim Acta. 1994;58(16):3425–3437.

[38] Merlivat L. Molecular diffusivities of $H_2{}^{16}O$, $HD^{16}O$, and $H_2{}^{18}O$ in gases. J Chem Phys. 1978;69(6):2864–2871.

[39] Luz B, Barkan E, Yam R, et al. Fractionation of oxygen and hydrogen isotopes in evaporating water. Geochim Cosmochim Acta. 2009;73(22):6697–6703.

[40] Cappa CD, Hendricks MB, DePaolo DJ, et al. Isotopic fractionation of water during evaporation. J Geophys Res Atmos. 2003;108(D16):4525.

[41] Rahul P, Ghosh P, Bhattacharya SK, et al. Controlling factors of rainwater and water vapor isotopes at Bangalore, India: Constraints from observations in 2013 Indian monsoon. J Geophys Res Atmos. 2016;121(23):13,936–13,952.

[42] Unnikrishnan Warrier C, Praveen Babu M, Sudheesh M, et al. Studies on stable isotopic composition of daily rainfall from Kozhikode, Kerala, India. Isot Environ Health Stud. 2016;52(3):219–230.

[43] Jeelani G, Deshpande R. Isotope fingerprinting of precipitation associated with western disturbances and Indian summer monsoons across the Himalayas. J Earth Syst Sci. 2017;126(8):108.

[44] Oza H, Padhya V, Ganguly A, et al. Investigating hydrometeorology of the Western Himalayas: Insights from stable isotopes of water and meteorological parameters. Atmos Res. 2022;268:105997.

[45] Rodriguez JD, Perez A, Lozano JA. Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE Trans Pattern Anal Mach Intell. 2010;32(3):569–575.

[46] Yadav S, Shukla S, editors. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. IEEE 6th International Conference on Advanced Computing (IACC); 2016.

[47] Karlik B, Olgac AV. Performance analysis of various activation functions in generalized MLP architectures of neural networks. Int J Artif Intell Expert Syst. 2010;1(4):111–122.

[48] Abiodun OI, Jantan A, Omolara AE, et al. State-of-the-art in artificial neural network applications: A survey. Heliyon. 2018;4(11):e00938.

[49] Ranjan GSK, Kumar Verma A, Radhika S, editors. K-nearest neighbors and grid search CV based real time fault monitoring system for industries. IEEE 5th International Conference for Convergence in Technology (I2CT); 2019.

[50] Bock S, Weiß M, editors. A proof of local convergence for the Adam optimizer. 2019 International Joint Conference on Neural Networks (IJCNN); 2019.

[51] Dahl GE, Sainath TN, Hinton GE, editors. Improving deep neural networks for LVCSR using rectified linear units and dropout. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing; 2013.

[52] Zeiler MD, Ranzato M, Monga R, et al., editors. On rectified linear units for speech processing. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing; 2013.

[53] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res. 2014;15(1):1929–1958.

[54] Gal Y, Ghahramani Z. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In: ICML'16: Proceedings of the 33rd International Conference on International Conference on Machine Learning, Volume 48. 2016; p. 1050–1059.

[55] Merlivat L, Jouzel J. Global climatic interpretation of the deuterium-oxygen 18 relationship for precipitation. J Geophys Res Oceans. 1979;84(C8):5029–5033.

[56] Gibson J, Edwards T. Regional water balance trends and evaporation-transpiration partitioning from a stable isotope survey of lakes in northern Canada. Glob Biogeochem Cycl. 2002;16(2):10-1–10-14.

[57] Stewart MK. Stable isotope fractionation due to evaporation and isotopic exchange of falling waterdrops: Applications to atmospheric processes and evaporation of lakes. J Geophys Res. 1975;80(9):1133–1146.

[58] Gonfiantini R. Environmental isotopes in lake studies. In: Fritz P, Fontes JC, editors. Handbook of environmental isotope geochemistry, Vol. 2. Amsterdam: Elsevier; 1986. p. 113–168.

[59] Hendricks M, DePaolo D, Cohen R. Space and time variation of $\delta^{18}O$ and δD in precipitation: Can paleotemperature be estimated from ice cores? Glob Biogeochem Cycl. 2000;14(3):851–861.

[60] Rozanski K, Sonntag C. Vertical distribution of deuterium in atmospheric water vapour. Tellus. 1982;34(2):135–141.

[61] Souchez R, Tison JL, Jouzel J. Freezing rate determination by the isotopic composition of the ice. Geophys Res Lett. 1987;14(6):599–602.

[62] Souchez R, Jouzel J, Lorrain R, et al. A kinetic isotope effect during ice formation by water freezing. Geophys Res Lett. 2000;27(13):1923–1926.

[63] Gat JR, Bowser C, editors. The heavy isotope enrichment of water in coupled evaporative systems. In: Tayor HP, O'Neill JR, Kaplan IR, editors. Stable isotope geochemistry: A tribute to Samuel Epstein. St Louis, Missouri: Geochemical Society; 1991. p. 159–168. (Geochemical Society Special Publication; No.3).

[64] Cappa CD, Smith JD, Drisdell WS, et al. Interpreting the H/D isotope fractionation of liquid water during evaporation without condensation. J Phys Chem C. 2007;111(19):7011–7020.

[65] Cappa CD, Drisdell WS, Smith JD, et al. Isotope fractionation of water during evaporation without condensation. J Phys Chem B. 2005;109(51):24391–24400.

[66] Hellmann R, Harvey AH. First-principles diffusivity ratios for kinetic isotope fractionation of water in air. Geophys Res Lett. 2020;47(18):e2020GL089999.

[67] Pierchala A, Rozanski K, Dulinski M, et al. Quantification the diffusion-induced fractionation of $^{1}H_2{}^{17}O$ isotopologue in air accompanying the process of water evaporation. Geochim Cosmochim Acta. 2022;322:244–259.

[68] Xinping Z, Tandong Y, Jingmiao L, et al. Simulations of stable isotopic fractionation in mixed cloud in middle latitudes–Taking the precipitation at Ürümqi as an example. Adv Atmos Sci. 2003;20(2):261–268.

[69] Lowenthal D, Hallar AG, McCubbin I, et al. Isotopic fractionation in wintertime orographic clouds. J Atmos Oceanic Technol. 2016;33(12):2663–2678.

[70] Lamb KD, Clouser BW, Bolot M, et al. Laboratory measurements of HDO/H_2O isotopic fractionation during ice deposition in simulated cirrus clouds. Proc Natl Acad Sci USA. 2017;114(22):5612–5617.

[71] Dütsch M, Blossey PN, Steig EJ, et al. Nonequilibrium fractionation during ice cloud formation in iCAM5: Evaluating the Common parameterization of supersaturation as a linear function of temperature. J Adv Model Earth Syst. 2019;11(11):3777–3793.

[72] Ganguly A, Padhya V, Oza H, et al. New insights into diffusive kinetic fractionation during liquid condensation under supersaturated environment: A novel approach for isotope tagging of ground level water vapor. 39 p. Available at SSRN: https://dx.doi.org/10.2139/ssrn.3978642

Figure captions

Figure 1. (a) Experimental setup for sampling vapour from ambient atmospheric air by complete cryogenic trapping. (i) Dewar flask containing slush (liquid nitrogen + ethyl alcohol) maintained at –78 °C; (ii) thermometer for monitoring slush temperature; (iii) glass condenser bulb; (iv) digital thermo–hygrometer to monitor temperature and Rh of ambient air; (v) aquarium pump to push the ambient air through the glass condenser; (vi) absolute alcohol for preparing the slush. (b) Experimental setup for collecting liquid condensate from the ambient air. (1) aluminum cone which contains ice at 0 °C to cool the outer surface; (2) thick PVC lid to cover the cone when filled with ice; (3) aluminum stand with adjustable screws to hold the cone such that the condensed liquid drops directly into small plastic vial placed in the groove on the base-plate; (4) coarse wire-mesh protective cover enclosing the sample vial and the cone; (5) the assembled condensation system.

Figure 2. Schematic representation of the ANN model architecture to reconstruct δ^{18}O in vapour with the help of Rh, T, δ^{18}O in liquid and S_i. The use of feature engineering for input parameters is done to better mimic the physical processes. We use 2 hidden layers with 10 nodes each in this network.

Figure 3. (a) Difference between liquid and vapour ($\Delta = \delta_{\text{liquid}} - \delta_{\text{vapour}}$) for δ^{18}O plotted against the saturation index S_i showing the progressive depleting trend of liquid isotopic signature with increase in S_i. (b) Similar plot for δ^2H. (c) Difference between *d*-excess of liquid and vapour highlighting the progressive increasing trend of *d*-excess in liquid with increase in S_i.

Figure 4. Schematic diagram explaining the development of the concentration gradient from initial time $t = t_o$ to final time $t = t_f$. The layer S denotes the microns thin supersaturated

boundary layer near the outer wall of conical surface. Here we have $e_{v3} > e_{v2} > e_{v1}$ and $e_{i3} > e_{i2} > e_{i1}$, where e_v denotes the partial pressure of water vapour at ambient temperature and relative humidity and e_i denotes the corresponding saturation vapour pressure for that layer or zone. The quantities e_{v3}, e_{v2}, e_{v1} and e_{i3}, e_{i2}, e_{i1} denote the evolution of partial vapour pressure as well as saturation vapour with time for the corresponding layer.

Figure 5. (a) Number density of H_2O molecules/m^3 of air for different S_i values ranging from 1 to 13, for different T–Rh combinations. The temperatures range from 15 to 40 °C at an 5 °C interval. The Rh range from 20 to 100 % at an interval of 20 %. (b) For narrow range of S_i values from 3.5 to 5.5. For a given value of S_i, for example 4.5, the highest number density is found for low T-high Rh combinations.

Figure 6. (a) Isotopic difference ($\Delta = \delta_{\mathrm{liquid}} - \delta_{\mathrm{vapour}}$) in δ^{18}O between each pair of liquid and vapour samples is plotted against its T and Rh, within a narrow range of S_i (3.5 to 4.5). The greater extent of isotope depletion due to more kinetic fractionation is observed under higher Rh condition, even within this narrow range of supersaturation. (b) Similar plot for δ^2H. The darker shade represents more depleted liquid compared to vapour, representing more fractionation.

Figure 7. (a) Isotopic difference ($\Delta = \delta_{\mathrm{liquid}} - \delta_{\mathrm{vapour}}$) in δ^{18}O between each pair of liquid and vapour samples is plotted against δ^{18}O of liquid, for a narrow range of S_i (3.5 to 4.5) showing positive correlation between the two which confirms that isotope depletion varies even under similar extent of supersaturation. (b) Similar plot for δ^2H.

Figure 8. ANN model predicted δ^2H and δ^{18}O vapour isotopic values plotted against the observed values of δ^2H and δ^{18}O showcasing the closeness of predicted values to the true values.