SPECIAL FEATURE: ISOSCAPES

Using atmospheric trajectories to model the isotopic composition of rainfall in central Kenya

KEIR SODERBERG,† STEPHEN P. GOOD, MOLLY O’CONNOR, LIXIN WANG, KATHLEEN RYAN, AND KELLY K. CAYLOR

1Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey 08544 USA
2S. S. Papadopulos & Associates, Inc., Bethesda, Maryland 20814 USA
3Department of Earth Sciences, Indiana University-Purdue University, Indianapolis (IUPUI), Indianapolis, Indiana 46202 USA
4Water Research Centre, School of Civil and Environmental Engineering, University of New South Wales, Sydney, New South Wales 2052 Australia


Abstract. The isotopic composition of rainfall (δ²H and δ¹⁸O) is an important tracer in studies of the ecohydrology, plant physiology, climate and biogeochemistry of past and present ecosystems. The overall continental and global patterns in precipitation isotopic composition are fairly well described by condensation temperature and Rayleigh fractionation during rainout. However, these processes do not fully explain the isotopic variability in the tropics, where intra-storm and meso-scale dynamics may dominate. Here we explore the use of atmospheric back-trajectory modeling and associated meteorological variables to explain the large variability observed in the isotopic composition of individual rain events at the study site in central Kenya. Individual rain event samples collected at the study site (n = 41) range from −51‰ to 31‰ for δ²H and the corresponding monthly values (rain volume-weighted) range from −15‰ to 15‰. Using the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model, we map back-trajectories for all individual rain hours occurring at a research station in central Kenya from March 2010 through February 2012 (n = 544). A multiple linear regression analysis demonstrates that a large amount of variation in the isotopic composition of rainfall can be explained by two variables readily obtained from the HYSPLIT model: (1) solar radiation along the trajectory for 48 hours prior to the event, and (2) distance covered over land. We compare the measurements and regression model results to the isotopic composition expected from simple Rayleigh distillation along each trajectory. The empirical relationship described here has applications across temporal scales. For example, it could be used to help predict short-term changes in the isotopic composition of plant-available water in the absence of event-scale sampling. One can also reconstruct monthly, seasonal and annual weighted mean precipitation isotope signatures for a single location based only on hourly rainfall data and HYSPLIT model results. At the study site in East Africa, the annual weighted mean δ²H from measured and modeled values are −7.6‰ and −7.4‰, respectively, compared to −18‰ predicted for the study site by the Online Isotopes in Precipitation Calculator.

Key words: back-trajectory; East Africa; HYSPLIT; Kenya; precipitation; Special Feature: Isoscapes; stable isotope.

Received 1 June 2012; revised 15 October 2012; accepted 23 October 2012; final version received 8 February 2013; published 11 March 2013. Corresponding Editor: G. Bowen. Copyright: © 2013 Soderberg et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. http://creativecommons.org/licenses/by/3.0/
† E-mail: soderbrg@princeton.edu
INTRODUCTION

The stable isotope composition of water ($\delta^{2}H$ and $\delta^{18}O$) has become an important tool for investigating ecosystem processes (Dawson et al. 2002). Understanding, predicting, and modeling the isotopic composition of rainfall is useful in studies of current and past ecosystems (Hendricks et al. 2000). From the earliest measurements of isotopes in precipitation, the spatial and temporal patterns have been linked to a few dominant processes (Craig 1961, Dansgaard 1964). Principally, temperature during phases changes and the Rayleigh-type isotope fractionation during rainout, can explain many of the large-scale patterns (Gat 2005). Post-condensation fractionation processes have been increasingly accepted as important drivers of isotopic composition (Friedman et al. 1962, Dansgaard 1964, Stewart 1975, Gat 2005, Risi et al. 2010a, Noone 2012) and have been included in isotopically-enabled General Circulation Models (GCMs) (Joussaume et al. 1984, Jouzel et al. 1987, Noone and Simmonds 2002, Lee et al. 2007, Risi et al. 2010c). These processes include equilibration of raindrops with the water vapor at the cloud base and while they are falling as well as evaporation of raindrops and re-entrainment of this evaporated moisture into clouds.

The meteorological history of an air parcel describing the conditions from the source vapor to cloud formation and through the course of rainout has also been recognized as an important predictor of isotopic composition (Gat and Carmi 1970, Lawrence et al. 1982, Rozanski 2005). Elevation, temperature, rainout volumes and air mass trajectory are known to influence the composition of precipitation. However, the effect of each specific factor may vary with location and season (Rozanski 2005). While some analysis of the influence of these confounding factors has been conducted in the tropics (Gonfiantini et al. 2001, Poage and Chamberlain 2001, Lawrence and Gedzelman 2003), little research has addressed the influence of atmospheric trajectories on East African rainfall signatures. Here we present an analysis relating isotopic composition to the spatial and meteorological history of rain events occurring at a research station in central Kenya over a 2-year period from 2010 to 2012.

Previous studies suggest that the isotopic composition of tropical rainfall is influenced more by local dynamics than by large-scale Rayleigh-type distillation effects that control global isotope patterns. Such meso-scale processes include the contribution of moisture from surface fluxes (Hendricks et al. 2000), convective dynamics (Lawrence and Gedzelman 2003), and post-condensation processes (Lee et al. 2007, Risi et al. 2010b). The importance of local dynamics in low latitudes indicates that models driven by high spatial and temporal resolution data are needed to capture the processes driving the isotopic composition of rainfall in these regions. Comparisons between sampled (e.g., the IAEA’s Global Network for Isotopes in Precipitation, “GNIP”) and GCM precipitation isotopes are acceptable for applications looking at seasonal and annual aggregate values (Werner et al. 2011). Many relevant ecohydrological studies, however, require knowledge of variations in isotopic composition at shorter time scales.

Event-scale precipitation isotopic composition data are useful for studying ecosystem processes from sub-daily to weekly scales such as evapotranspiration partitioning (Wang et al. 2010), radiation control on the isotopic composition of transpiration (Wang et al. 2012), hydraulic redistribution (Dawson 1993) and soil water dynamics (Gazis and Feng 2004, Brooks et al. 2010). Such higher frequency precipitation isotope investigations have been undertaken in the past to help identify the processes underlying patterns in aggregate data (Noone and Simmonds 2002). Hoffmann et al. (1998) found that the variability of modeled daily values was higher than the observed variability, whereas Noone and Simmonds (2002) found better agreement. Interestingly, GCM modeling has found that the best correlations between temperature and the isotopic composition of precipitation are for temperatures on precipitation days, although the interest in paleoclimate reconstructions is in the annual mean temperature (Noone and Simmonds 2002). Thus, additional event-scale comparisons with aggregated data will help to define underlying uncertainties in relationships between the isotopic composition of precipitation and climatic variables. This information will be especially useful in low-latitude environments where temperature-isotope relationships are weak.
Previous air mass trajectory models of isotopic composition have related variability in the rainfall isotopic composition to changes in the origin of trajectories as well as air mass mixing. These studies have been undertaken in North America (Lawrence et al. 1982, Friedman et al. 2002, Burnett et al. 2004, Sjostrom and Welker 2009, Ersek et al. 2010, Sinclair et al. 2011), Europe (Gat and Carmi 1970, Dirican et al. 2003, Baldini et al. 2010), Asia (Fudeyasu et al. 2011, Liu et al. 2011), Australia (Barras and Simmonds 2008, 2009) and one looked at precipitation origins in Lake Tanganika, Africa (Lewis et al. 2010). Related modeling efforts include atmospheric dynamics and land-atmosphere exchange (Rozanski et al. 1982, Yoshimura et al. 2003, Henderson-Sellers et al. 2006, Levin et al. 2009, Yoshimura et al. 2010, Zhao et al. 2012). East African precipitation remains understudied despite its importance for this region that depends largely on rain-fed agriculture (Funk et al. 2008). East Africa has steep spatial gradients in the isotopic composition of rainfall, suggesting that this could be a tool to help understand precipitation dynamics in the region (Bowen and Revenaugh 2003, Levin et al. 2009). There is a high degree of uncertainty in GCM precipitation predictions for East Africa, although a drying trend is becoming apparent (Williams and Funk 2011).

Here we use meteorological data along back-trajectories of air parcels associated with rain events, as well as spatial information about the trajectories, to develop a regression model of rainfall isotopic composition at the event-scale. This effort is not meant to develop a universal model of rainfall isotopic composition, but rather to test the utility of easily obtainable atmospheric trajectory spatial and meteorological information to predict the isotopic composition of rainfall at the study site. Any such model will necessarily be site-specific, but highly relevant to studies of isotope hydrology at that site given that the model yields both event-scale and aggregate volume-weighted isotopic composition. The high temporal resolution information provided by such a local model will then help to interpret, for example, continuous water vapor isotope measurements on an eddy covariance flux tower. If this procedure proves useful, it can be repeated at other sites with sufficient measurements of daily or event-scale rainfall isotopic composition.

**METHODS**

The influence of atmospheric trajectory on East African rainfall isotopic composition was studied using sampled rainwater and the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Draxler and Hess 2004) developed by the National Oceanic and Atmospheric Administration (NOAA). Rainwater samples were collected at the Princeton University eddy flux tower located at the Mpala Research Center (MRC), in Laikipia County, Kenya (0°29’8.15” N, 36°52’12.32” E, elevation 1619 m). The stable isotopic composition of water ($\delta$) is defined in Eq. 1:

$$\delta = \left( \frac{R}{R_{\text{std}}} - 1 \right)$$

where $R$ is the ratio of a rare (denoted i, e.g., $^{18}$O) to common isotope ($^2$H/$^1$H or $^{18}$O/$^{16}$O) in sample water and $R_{\text{std}}$ is the same ratio of the international standard, VSMOW (Gonfiantini 1978).

**Sample collection and analysis**

We collected 41 rain samples from a period of March 2010 through February 2012 (Supplement 1). The sample collection was designed to prevent evaporation. When investigators were present during a rain event, samples were collected immediately from an open container ($n = 18$). Otherwise, a ball-in-funnel collector was used: when water is in the funnel, the ball floats, and water can enter the otherwise sealed container; when water is not in the funnel, the ball covers the funnel opening and prevents vapor exchange (Deshpande et al. 2010). The samples represent 35% of the rain volume for the two-year period. During this time, 18% of the rain was not sampled, and the remaining 47% failed quality control criteria for analysis (3%) and holding times (44%).

The isotopic composition of the rain samples ($\delta^{18}$O and $\delta^2$H) was measured at MRC using an off-axis integrated cavity output spectroscopy (ICOS) system coupled to a liquid water vaporizer (DLT-100 & WVISS, Los Gatos Research Inc.). The ICOS was calibrated daily using multiple liquid water isotope standards that span...
the range of the rain samples. Typical standard deviations during the time of analysis were 0.3‰ and 2.0‰ for $\delta^{18}O$ and $\delta^2H$, respectively. Hourly rainfall amounts were measured with tipping bucket rain gauges at the flux tower and the adjacent UHURU experiment (Pringle 2012).

**HYSPLIT model**

The HYSPLIT model calculates air mass position through time using pressure, temperature, wind speed, vertical motion, and solar radiation inputs from the NOAA FNL meteorological database. There is a web-based interface for HYSPLIT (Draxler and Rolph 2011), but the model can also be downloaded and automated (contact the authors for details). Air parcel trajectories were modeled 10 days backwards in time starting from every hour of rainfall that occurred (544 rain hrs at the site over 2 years). To compute a trajectory, the HYSPLIT model requires a starting time, location, altitude as well as the NOAA meteorological data files. The starting location was the sampling site and the altitude used was a Lifting Condensation Level (LCL) based on micrometeorological conditions and the tower (Eq. 2):

$$\text{LCL} = \frac{T - T_{\text{dew}}}{\Gamma_d - \Gamma_{\text{dew}}}$$

where $T$ ($^\circ$C) is the ambient temperature, $T_{\text{dew}}$ ($^\circ$C) is the dewpoint temperature, $\Gamma_d$ (${\text{cm}}\,{\text{m}}^{-1}$) is the dry adiabatic lapse rate, and $\Gamma_{\text{dew}}$ (${\text{cm}}\,{\text{m}}^{-1}$) is the wet adiabatic lapse rate.

The HYSPLIT model uses a three-dimensional Lagrangian air mass velocity algorithm to determine the position of the air mass and reports these values at an hourly time-resolution over the trajectory. The model also outputs meteorological variables along the trajectory: potential temperature, ambient temperature, precipitation, mixing depth, relative humidity, solar radiation, terrain height and altitude of the air parcel. For the purpose of analysis, we computed a mean trajectory for each rain sample ($n=41$) by weighing the trajectories ($n=185$ total) by the proportion of the rain sample that had fallen during their respective starting hours. Utilizing all rain hours that contributed to a rain event represents an advance in this type of modeling (Sjostrom and Welker 2009, Ersek et al. 2010).

**Developing a local model of rainfall isotopic composition**

The utility of spatial and meteorological variables along the back-trajectories for providing local isotopic information was tested by selecting independent variables based on the strength of their correlations to the measured isotopic composition of rain samples ($\deltaV$). The “deuterium excess” value ($d$) was calculated (Eq. 3) and included as a measured variable given its utility for comparing both $\delta^{18}O$ and $\delta^2H$ to the Global Meteoric Water Line “GMWL” (Dansgaard 1964)

$$d = \delta^2H - 8 \times \delta^{18}O.$$  

Geometric mean regression (also known as Type II or reduced major axis regression) was then used to develop linear relationships between a primary variable and $\delta_V$, and subsequently between a secondary variable and the residuals of the primary relationship. Geometric mean regression is an appropriate regression model when both variables are random (Pataki et al. 2003). These calculations were made using a MATLAB script (Trujillo-Ortiz 2010).

The regression model was parameterized in three scenarios for each isotope to investigate the effects of sampling on model fitness: (A) all data points for the two-year period were used ($n=41$), (B) the median of 50 cross-validation models each built from a random selection of 18 data points, and (C) data points from the first year only were used to train the model ($n=19$). For (B) and (C), the data points not involved in constructing each model were used as test data. Accuracy was assessed using the RMS of model residuals (known value minus predicted value) and bias was assessed using the median residual value.

To help interpret the regression model, we compared the measured and modeled variables to a process-based calculation for the trajectory hours associated with each rain sample ($n=185$). In this first-order approach, we used the simplest case of Rayleigh fractionation along each trajectory (Dansgaard 1964, Noone 2012). During a Rayleigh process, condensate is removed from the system, leaving the remaining vapor isotopically depleted. We calculated the decrease in the the water vapor isotopic composition ($\deltaV$) for each trajectory hour that experienced a decrease in water vapor mixing ratio (Eq. 4):
\[
\delta_{V_{n+1}} - \delta_{V_n} = (\alpha_{LV} - 1) \ln \left( \frac{q_{n+1}}{q_n} \right)
\]

with \( q \) as the mixing ratio (mmol mol\(^{-1}\)), and \( \alpha_{LV} \) as the equilibrium isotopic fractionation factor (Horita and Wesolowski 1994). The subscripts “\( n \)” and “\( n + 1 \)” refer to the values at the previous and current trajectory point, respectively.

For this calculation, the starting point was taken as the last trajectory point over the ocean for all trajectories with an ocean component (\( n = 178 \)). The trajectories with no ocean component (\( n = 7 \)) followed the coastline closely for most of their 240-hr trajectories, and the first point close to the coast (i.e., the nearest in time to the rain event) was taken as the starting point for this calculation. The starting \( \delta_V \) was assumed to be in thermal equilibrium (Eq. 5) with liquid water of a mean ocean composition (0\% for \( \delta^{18}\)O and \( \delta^2\)H).

\[
\delta_V = \left( \frac{\delta_L + 1}{\alpha_{LV}} \right) - 1.
\]

The changes in isotopic composition along the trajectory from Eq. 4 were then summed, and the resulting final \( \delta_V \) value was used to calculate a \( \delta_L \) value in thermal equilibrium. This liquid value is taken as the rainfall isotopic composition that would have resulted from Rayleigh fractionation only (\( \delta_{L_i} \)). This calculation ignores post-condensation processes, advection and mixing with land surface fluxes (Noone 2012), but it can be a first step towards determining the relative importance of various fractionating and mixing processes.

**RESULTS**

*Measured rainfall isotopic composition*

The isotopic composition of rainfall at the MRC flux tower over a two-year period shows a large amount of variability, with individual samples ranging from −51\% to 30\% for \( \delta^2\)H. Monthly means weighted by the rain volume associated with each sample have a range of −15\% to 10\% for \( \delta^2\)H (Fig. 1). These monthly means allow for comparison with data from the Online Isotopes in Precipitation Calculator (OIPC) (Bowen and Revenaugh 2003, Bowen et al. 2005, Bowen 2012), which is based on data from the ongoing GNIP sampling program (Dansgaard 1964, Rozanski et al. 1993). The OIPC monthly values span −33\% to 6\% for \( \delta^2\)H, representing a slightly larger range than the sample-based monthly mean values. The OIPC value for annual weighted mean (−18\% for \( \delta^2\)H) is much more depleted than the corresponding sample-based value (−7.6\%). The local meteoric water line (LMWL) derived from MRC samples (\( \delta^2\)H = 8.3 × \( \delta^{18}\)O + 11.0) lies between that for all African GNIP stations (\( \delta^2\)H = 7.4 × \( \delta^{18}\)O + 10.1) (Levin et al. 2009) and that calculated from the OIPC monthly values for the study site (\( \delta^2\)H = 8.1 × \( \delta^{18}\)O + 14.0). Consistent LWML slopes of 7.4 and 7.5 have been defined for tropical montane systems in Africa and South America (Gonfiantini et al. 2001). It should be noted that the LMWL defined above with coefficients of 8.3 and 11.0 was derived using a geometric mean regression. A least squares regression gives coefficients of 7.7 and 10.0, which are much closer to those of Levin et al. (2009) and Gonfiantini et al. (2001).

**HYSPLIT model**

The 10-day back trajectories for every rain hour at the MRC flux tower over the two-year study period (\( n = 544 \)) are shown in Fig. 2. There is a distinctly different trajectory for the dry season (27 December to 24 March), with nearly all trajectories coming from the northeast. Trajectories from the other seasons come from the southeast, with a more easterly component associated with the short rains (1 October to 26 December), and a more southerly component in the continental rains (22 May to 31 September). Season cutoff dates are the mean dates from Camberlin et al. (Camberlin et al. 2009), with a modification of the short rains start date from 23 October to 1 October based on local anecdotal evidence.

**RIST model**

The strongest relationship found between the isotopic composition of rainfall (\( \delta_L \)) and both local and trajectory-based variables was with the solar flux along the trajectory during the 48 hrs preceding the rain event (\( I_{48} \) [W m\(^{-2}\); Fig. 3]). Combining \( I_{48} \) with the total distance traveled over land for the air parcel results in the Rainfall Isotopes Solar Terrain model (RIST). The other solar flux time periods evaluated were 12 hrs, 24 hrs and 240 hrs. Many other local and trajectory-based variables were considered, and a correlation matrix is presented in Supplement 2. The solar flux relationship was modeled using log-
transformed data and the coefficients were determined through geometric mean regression. The RIST model is a function of $I_{48}$ and the horizontal distance covered over land ($x_{\text{land}}$ in meters). The relationship between the residuals of the $I_{48}$ vs. $d_L$ regression and $x_{\text{land}}$ to be inverse (i.e., longer distances were associated with lower residuals), and a better correlation was found between residuals and the inverse of $x_{\text{land}}$ (Eq. 6):

$$
\delta_L = c_1 \ln(I_{48\text{hrs}}) + c_2 \left( \frac{1}{x_{\text{land}}} \right) + c_3 \quad (6)
$$

with $\delta_L$ referring to either $\delta^2$H or $\delta^{18}$O. The empirical coefficients ($c_1$, $c_2$, and $c_3$) are listed in Table 1.

The RIST regression model explains the majority of the variations in the measured values ($r^2 = 0.63$ and 0.62 for $\delta^2$H and $\delta^{18}$O, respectively), with each isotope having a slope around 1.1 (modeled vs. measured values) indicating that the model slightly overestimates the observed variability (Fig. 3). However, the corresponding intercepts are 0.1%o for both $\delta^2$H and $\delta^{18}$O, indicating little overall bias when considered in relation to the analytical uncertainties of 2 and 0.3‰, respectively.

The different model parameterization scenarios (A, B and C) result in slightly different predictions (Table 1). The $\delta^2$H cross-validation results (B) show a similar RMS value to (A) in the average training set (14.3‰), and a slightly higher RMS in the average test set (15.6‰). The $\delta^{18}$O cross-validation results mirror the $\delta^2$H results, with RMS values of 2.0‰ and 2.2‰ for the training and test sets, respectively, compared with 2.0‰ in scenario A. There is very little bias in the $\delta^{18}$O cross-validation results. Overall, the cross-validation test data sets exhibit only slightly reduced accuracy relative to their respective

---

Fig. 1. Monthly rainfall isotope values based on two years of event-scale sampling (“Sample”, n = 41), and values obtained from the Online Isotopes in Precipitation Calculator (“OIPC”) for the study site in central Kenya. The global meteoric water line (“GMWL” $\delta^2$H = $8 \times \delta^{18}$O + 10) is shown for reference along with local meteoric water lines (LMWLs) from least-squares regression of the monthly values: Sample ($\delta^2$H = $8.3 \times \delta^{18}$O + 11.0), OIPC ($\delta^2$H = $8.1 \times \delta^{18}$O + 14.0). The “Sample” monthly values are volume-weighted means for samples collected in a given month.
training data sets. The temporal extrapolation shown in (C), however, shows a large difference in RMS and bias between training and test data sets, which could be expected from forcing the training and test data to come from two separate populations of data (i.e., year 1 and year 2).

RIST model (A) was used to construct monthly weighted mean isotopic composition of rainfall throughout the study period using 537 of the 544 total rain hours. The 7 excluded rain hours had very low solar flux values (<120 W/m²) which would have required significant extrapolation from the linear regression model (spanning 133 to 323 W/m²). The excluded rain hours represent 2% of the total rainfall for the study period and occurred in May and July 2010. When plotted over time with the sample-based and OIPC monthly mean values (Fig. 4), it is clear that a significant difference in the RIST and OIPC models occurs in the short rains season (October–December). During the short rains, OIPC predicts much more depleted values than observed in the samples. The other striking feature is the highly variable sample and RIST model results for the continental rains, particularly in July and August. The continental rains exhibited the largest amount of variability of the four seasons in terms of solar flux, air temperature, and pressure. There is relatively better agreement among the three sources in the dry season and long rains.

Connecting the modeled isotope values for each rain hour back to their respective 10-day trajectories reveals an additional pattern (Fig. 5). Although isotopically depleted or enriched rain can come from all directions, the most enriched quintile of rain hours (δ²H > 17.4‰) tends to come from a more easterly direction (mean of 122° from north, p < 0.025 for a two-sided t-test on the lowest and highest isotope quintile trajectories), whereas the most depleted quintile (δ²H < -26.7‰) tends to come from a more southerly direction (152°).

The Rayleigh fractionation calculation resulted in...
in, as expected, highly depleted \( \delta_{LT} \) values relative to the measured values \( (\delta_{LM}) \). Plotting the difference between the two \( (\Delta_{m-r}) \) reveals a strong relationship with the sum of changes in

\[
\delta^2H_{\text{RIST}} = 1.06 \delta^2H_m + 0.132
\]

\[
\frac{\delta^2H_m}{\text{RMS}} = 96.22\ln(I_{48}) - 536.8
\]

\[
\frac{\delta^{18}O}{\text{Bias}} = \text{median residual (known value – predicted value), in } \%.
\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\(|\text{Bias}| = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]

\[\text{Bias} = \frac{\text{median residual (known value – predicted value), in } \%}{\text{Bias}}\]
mixing ratio for trajectory hours that experienced an increased humidity ($R_{qp}$) (Fig. 6). This correlation mainly reflects the underlying balance between the increases and decreases in $q$ ($\Sigma q_p$ and $\Sigma q_n$). The intercept of the regressions ($\Delta_{m-r}$ vs. $R_{qp}$) are positive for both $\delta^2$H and $\delta^{18}$O (13.4% and 0.41%, respectively). The residuals of this relationship are strongly correlated with $I_{48}$, as with the original measured values described above. Following a similar procedure as the RIST model, a regression model using the Rayleigh fractionation results along with $I_{48}$ produces the RISR model (Fig. 6). The RISR model does not perform as well as the RIST model, with $r^2$ values of 0.43 and 0.47 for $\delta^{18}$O and $\delta^2$H, respectively (RMS values of 2.3% and 17.8%; bias of 0.2% and 0.1%, respectively).

**DISCUSSION**

The annual weighted mean isotopic composition of rainfall from the samples and the RIST model are much more enriched than expected from the OIPC model. The sample-based isotopic composition of rainfall for MRC lies between the long-term GNIP stations located to the south and north of the study site (Muguga, Kericho, Asela, Addis Ababa; Fig. 7). Based on the GNIP station data (IAEA 2002), Central Kenya, and the Laikipia region to the northwest of Mt Kenya in particular, occurs in a transition from depleted rainfall isotopic composition to the southwest ($-16\%$ to $-21\%$ for $\delta^2$H), and enriched values to the northeast ($0.0\%$ to $4.1\%$ for $\delta^2$H). The fact that the annual value for the study site ($-7.6\%$ for $\delta^2$H) falls between these long-term stations is consistent with its location. A recent study
Fig. 5. Atmospheric back-trajectories (240 hr) for the most enriched (Panel A; >17.4‰ for δ²H) and most depleted (Panel B; <−26.7‰ for δ²H) quintiles of the rainfall isotopic composition modeled for each rain hour observed on site over the two-year study period. The RIST model (Eq. 6, parameterization scenario A) was used to generate the isotope values associated with each trajectory.
reported on several rain, surface water and groundwater samples collected across Kenya and Ethiopia (Levin et al. 2009). The un-weighted mean value for $\delta^{2}H$ from Kenyan rain samples in the Levin et al. (2009) study is $-1.8\%$ ($n = 18$). The lack of long-term GNIP data across large portions of this region can probably explain the large differences with the OIPC value ($-18\%$ for $\delta^{2}H$).

The RIST model highlights the importance of two variables on the isotopic composition of rainfall (solar flux and distance travelled over land), each ultimately linked to some contributing physical processes. The Rayleigh fractionation calculations give some insight into these processes. The calculated rainfall isotopic composition based solely on Rayleigh fractionation ($\delta_{LR}$) is highly depleted relative to measured values ($\delta_{LM}$), and the difference between the two ($\Delta_{m-r}$) should be related to the other physical processes involved. The trajectory hours that experienced an increase in humidity, and the magnitude of this increase, are directly related to $\Delta_{m-r}$ (Fig. 6). This relationship is expected given that the highly depleted $\delta_{LR}$ values (i.e., the largest $\Delta_{m-r}$) are associated with longer distances travelled over land, allowing the air mass to experience complementary increases in humidity. Interestingly,
however, the controlling variable in the RIST model, \(I_{48}\), is larger for values falling above the regression line for \(\Delta_{\text{m-r}}\) vs. \(\Sigma_{\text{qp}}\), with \(r^2\) values of 0.56 and 0.58 (\(\delta^{18}\)O and \(\delta^2\)H, respectively) for the correlation between residuals and \(I_{48}\) (Fig. 6). One explanation for this pattern is that trajectories with higher \(I_{48}\) receive a flux of isotopically enriched water vapor (advection or surface flux), resulting in higher \(\delta_{\text{lm}}\) values in onsite rainfall.

Post-condensation processes such as evaporation of falling raindrops could be another explanation worth pursuing. The intercept of the \(\Delta_{\text{m-r}}\) vs. \(\Sigma_{\text{qp}}\) regression (Fig. 6) is positive for both \(\delta^2\)H (13.4\(\%\)) and \(\delta^{18}\)O (0.35\(\%\)). This implies that even with no trajectory hours with humidity increases (i.e., no substantial advection or surface flux input), the Rayleigh fractionation predicts a depleted rainfall isotopic composition relative to the measured values. With no external inputs, this enrichment must be related to a post-condensation process. Interestingly, for a raindrop starting out with an isotopic composition on the GMWL (i.e., \(d = 10\)), experiencing enrichment in \(\delta^2\)H of 13.4\(\%\) and \(\delta^{18}\)O of 0.35\(\%\) would result in a final deuterium excess value of 21\(\%\) (Eq. 3). The magnitude of this fractionation will be variable, but the relative enrichment of \(\delta^2\)H and \(\delta^{18}\)O is more conservative, and has been a useful marker of post-condensation processes in the tropics since global precipitation patterns were first investigated (Dansgaard 1964).

A recent study of isotope dynamics associated with the Madden-Julian Oscillation found a maximum \(d\) in rainfall correlated with a maximum in stratiform cloud rainfall. However, this observation represents an alternative explanation to raindrop evaporation because the high \(d\) rainfall values were attributed to the subsidence of high \(d\) water vapor from the upper troposphere (Kurita et al. 2011). Another post-condensation process is exchange between water droplets and surrounding vapor, which could occur with isotopically enriched vapor recently derived from evaporation. This latter process has been shown to be a dominant mechanism at low latitudes (Field et al. 2010), and would result in
lower $d$ values.

The equilibration of raindrops with locally-derived vapor during light storms is one possible explanation for the enriched precipitation values seen at the study site. Levin et al. (2009) note that this enrichment could be enhanced if transpiration is a major source of vapor for equilibration. Transpired vapor tends to be isotopically similar to the liquid water drawn into its roots (Harwood et al. 1998, Wang et al. 2012), whereas vapor is highly depleted relative to its liquid water source due to equilibrium and kinetic fractionation (Craig and Gordon 1965). With a given isotope profile in the soil, vapor derived from transpiration will be more enriched than vapor derived from soil evaporation. Analysis of wind and potential temperature fields indicated the Congo basin as a moisture source for isotopically enriched Ethiopian rainfall (Levin et al. 2009). This westerly source of vapor has the potential to bring moisture to parts of Kenya during the continental rains of June-September. The HYSPLIT back-trajectories presented here are highly variable during the continental rains, leading to large intra-seasonal differences (e.g., June through September 2010, Fig. 4). Perhaps some intermittent role of westerly, transpiration-dominated, vapor could explain some of these differences. Such a process would be consistent with a recent global evaluation of continental precipitation recycling which placed a steep gradient across Kenya, from around 40% recycled precipitation in the west to around 10% along the coast (van der Ent et al. 2010).

The relationships utilized by the RIST model also may change over time as precipitation dynamics in the region appear to be changing. Although annual rainfall has shown no trend, rain events are becoming less frequent and more intense over time in the Laikipia region (Franz et al. 2010). There has been a regional pattern of decreased rainfall totals during the long rains, Kenya’s most important growing season, since 1980 (Funk et al. 2008, Williams and Funk 2010), although the short rains appear to be less affected and continue to be related to the El Nino Southern Oscillation (Camberlin et al. 2009). During the study period investigated here, the Short Rains was the most significant season overall, contributing 34% of the total rainfall for the two-year period.

The RIST model and the current parameterization have some limitations and could be further improved in the future. First, it is site-specific, making its utility outside the flux tower footprint unknown. We have two other collection sites ongoing across MRC, each about 30km from the study site, and plan to assess the broader applicability of the model using data from these sites. Second, we sampled only 35% of the total rainfall (see Methods). Ideally, a cumulative long-term collector would be deployed at the event-scale collection site and sampled monthly in addition to the event-scale sampling. In testing the effect of sampling on the model fitness, we show that although the model is not sensitive to removing samples (Table 1, scenario B), it is sensitive to using only one year of data to predict another (scenario C). This could be expected given the variability of the isotopic composition of rainfall at this site, but highlights the need to consider inter-annual, inter-seasonal as well as intra-seasonal variability when designing a sampling campaign. A new technology is being developed which will allow for the automated field measurement of rainfall isotopic composition in real-time (Stocker et al. 2012), which will vastly improve the coverage of rain events. Despite these limitations, the RIST model represents a first step towards integrating the spatial and meteorological history of an air parcel to associated rainfall isotopic composition. With similar sampling and modeling efforts at different sites, a regional empirical isoscape could be developed to support investigations of, e.g., evapotranspiration and moisture source modeling.

One future direction of this research is to better characterize the relationship between the isotopic composition of atmospheric water vapor and rainfall during different types of precipitation. Interactions between falling raindrops and the surrounding vapor have long been recognized (Stewart 1975, Field et al. 2010), and acknowledged since the first isotope-enabled GCM (Joussaume et al. 1984). For example, one GCM applies 100% equilibration between liquid droplets and vapor in stratiform precipitation, and 50% equilibration in convective precipitation (Field et al. 2010). A new formulation of this process incorporates both temperature and rainfall rate, with modeled raindrop size (Lee et al.
2007). Understanding such a process will help modeling studies (Risi et al. 2010c), which have already demonstrated the utility of isotopes in quantifying confounding water cycle components such as precipitation/transpiration ratios or the influence of evaporation during rainfall (Peng et al. 2005, Worden et al. 2007, Risi et al. 2010c, Yoshimura et al. 2010). Ground-based studies of high temporal resolution vapor and liquid isotopic composition are beginning to enhance our understanding of local and synoptic climatology (Deshpande et al. 2010, Risi et al. 2010a, Risi et al. 2010b, Wen et al. 2010, Noone et al. 2011, Tremoy et al. 2012). The development of remotely sensed water vapor isotopic composition measurements, such as the Tropical Emission Spectrometer (TES) and Scanning Imaging Absorption Spectrometer for AtmosphericChar- tography (SCIAMACHY), will help connect the ground-based studies to the global water cycle (Worden et al. 2007, Frankenberg et al. 2009, Risi et al. 2010c). Information about the isotopic composition of atmospheric water vapor, in combination with other gridded meteorological datasets presents the opportunity to model with greater detail the dynamic progression of water vapor in air masses over their lifetime (Noone et al. 2011). A recent description of the theoretical framework for undertaking these investigations by Noone (2012) provides a useful roadmap for future work.

CONCLUSION

We have demonstrated that highly variable isotopic composition of rainfall can be explained by incorporating the spatial and meteorological history of the associated air masses. We find that the solar flux along the atmospheric trajectory for 48 hrs preceding rainfall is correlated with the isotopic composition of rainfall. The solar flux and the distance traveled over land are combined together in a simple model for predicting the isotopic composition of rainfall at the study location in central Kenya. Such a model can be readily created from hourly rainfall data, two years of event-scale isotope sampling, and back-trajectories provided by the HYSPLIT model. This type of model may only apply to certain geographical areas, but it is worth attempting if highly variable rainfall isotopic compositions are observed. Particular candidate sites would be in the tropics or other areas where inter and intra-seasonal variations cannot be easily explained by temperature.

Incorporating the spatial history of an air mass is a first step towards determining the relative contributions of various factors to the isotopic composition of precipitation (e.g., storm dynamics, orographic effects, differential water vapor sources). The regression model described here is simple, and will likely perform poorly relative to more complex land-atmosphere exchange models that incorporate mechanistic explanations of isotopic variation. The model is also limited by being site-specific and based on only two-years of data. However, this exercise has highlighted the importance of certain parameters that could help guide more complex models to better represent precipitation isotopes for areas that are currently underperforming in the models. Better representation of isotopic mixing and fractionation processes will assist in resolving outstanding issues in global Earth science research.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the NOAA Air Resources Laboratory (ARL) for the provision of the HYSPLIT transport and dispersion model and used in this publication. This project was funded by NSF through a CAREER award to K. K. Caylor (EAR847368). Lixin Wang also acknowledges the financial support from the vice-chancellor’s postdoctoral research fellowship at the University of New South Wales. We greatly appreciate the field assistance of Ekomwa Akuwam, John Gitonga, Boniface Mukoma, Moses Musyoka, and the general support staff at Mpala Research Center. We thank the two reviewers for their very constructive comments.

LITERATURE CITED


Draxler, R. R. and G. D. Hess. 2004. Description of the HYSPLIT_4 modeling system. NOAA Technical Memorandum ERL ARL-224. NOAA Air Resources Laboratory, Silver Spring, Maryland, USA.


Harwood, K. G., J. S. Gillon, H. Griffiths, and M. S. J. Broadmeadow. 1998. Diurnal variation of D\textsuperscript{13}CO\textsubscript{2}, D.C.\textsuperscript{18}O and evaporative site enrichment of dH\textsubscript{2}O\textsuperscript{16}O in Piper aduncum under field conditions in Trinidad. Plant Cell and Environment 21:269–283.


Risi, C., S. Bony, F. Vimeux, C. Frankenberg, D. Noone,


**Supplement 1**

Rainfall sample data including the measured stable isotope values, rainfall amounts, collection dates and times, and meteorological data at the collection site (Ecological Archives C004-003-S1).

**Supplement 2**

A matrix of correlation coefficients for variables considered in the construction of the RIST and RISR regression models (Ecological Archives C004-003-S2).