This document contains a list and description of all data tables, files, and code included in this project. Unless otherwise mentioned, all files are contained on the external hard drive in "/Elements/Hendricks model/".

Appendix A. Large Data Tables and Files

A number of large data tables and files are alluded to in the text or are used in the code for the model. These tables and files are outlined below and are included in the electronic version of this submission.

A.1 Large Data Tables

The data tables containing modern isotopic values for soils and waters can be found in the following attached files:

1) "modern soils.csv"
2) "modern water d18O dD – no USGS.csv"
3) "modern water d18O dD.xlsx"
4) "I&Ttable.xlsx"
5) "Cotton 2012.csv"
6) "Cotton 2013.csv"

1) "modern soils.csv" contains isotopic values for six soil profiles, including some not currently published.
2-3) "modern water d18O dD – no USGS.csv" contains isotopic values for modern waters, mostly collected during the winter months. Many of the referenced publications contain further non-winter values that were not included in the compilation. "modern water d18O dD.xlsx" is similar, except that it also includes many data from USGS, many of which are from groundwater and/or were not collected during winter months.
4) "I&Ttable.xlsx" contains the data from Ingraham and Taylor's (1991) Traverse II, used for plot generation.
5-6) "Cotton 2012.csv" and "Cotton 2013.csv" contain data for soil respiration – mean annual precipitation relationships. These are neither addressed nor relevant to this project, but they are included because they are utilized in the attached code.

A.2 Large Data Files

A number of large data files were created and used throughout this project, attached here to limit time from re-creating or re-downloading.

1) "CaliforniaNevadaStormTrack_NARR_July2.RData"
2) "NARRgriddata.RData"
3) "NARR_RData" – contained in "/Elements/NARR_RData"
4) "HYSPLIT_output_NARR" – contained in "/Elements/HYSPLIT_output_NARR"

1) "CaliforniaNevadaStormTrack_NARR_July2.RData" contains the NARR monthly-mean values for the climatological variables input to the vapor transport and soil moisture models.
2) "NARRgriddata.RData" contains all of the latitude, longitude, and land mask cover values for the entire NARR dataset.
3) "NARR_RData" is a folder containing a file for each year of 3-hourly NARR data in files named "NARRyyyy.RData." Each of these files contains the Damköhler number, precipitable water, relative humidity, air temperature, and timestamp (hours since 1800-1-1 00:00:0.0) associated with each NARR observation for that year. Note that each of these files is quite large (~2.3 GB) and *they should not all be loaded into the R environment at the same time*!
4) "HYSPLIT_output_NARR" is a folder containing all of the direct HYSPLIT output for each location, the .nc histogram files produced by "Trajectory Plotter-NARR_AJR" (*e.g.* "histhay.DJF.nc"), the .RData files generated by "Trajectory Plotter-NARR_AJR" (*e.g.* "hay.DJF.precip.RData"), the lists of HYSPLIT output integrated with appropriate NARR data (*e.g.* "list.Hay.DJF.withNARR.RData"), and the model results along each individual trajectory (*e.g.* "Hay.DJF.modelresults.RData").

Appendix B. Description of Code

Many R files were used and included in this project. Those used to run the model and to create the plots shown in this document are described below.

B.1 "bulk code.R"

This file contains the bulk of the code used to generate plots throughout this project. Ten hypothetical NARR trajectories are considered and averaged, and they are used to feed the vapor transport model and soil moisture models. This file also contains the code for soil carbon isotopic values, though those are not addressed in this project. "bulk code.R" contains many sub-sections:

1) “parameters and inputs” – This section defines constants and establishes vectors for evaporation/transpiration partitioning ratios and depths in a soil profile. This section also defines the initial values for δ^{18}O to initialize model runs.
2) “necessary equations” – This section defines equations for fractionations and carbon concentration calculations used later in the code.
3) “basic atmosphere model for multiple storm tracks” – This section runs the Hendricks model for both advection-only and eddies-only cases along the ten storm tracks input from the NARR data.
4) “average tracks together by distance from coast” – This section runs the same model (the code is copy-pasted from section 3), except along a mean storm track where the NARR values are averaged by their distance from the coast.
5) “model plots” – This section produces plots of model output, without any data.
6) “data time!” – This section loads, organizes, and plots data (both as maps and in conjunction with model output).
7) “Ingraham and Taylor, 1991 plots” – This section creates plots using just the Ingraham and Taylor (1991) Traverse II data.
8) “resp-MAP relationships” – This section defines three respiration-mean annual precipitation relationships relevant to carbon isotope calculations. This section is not relevant to this project.
9) “soil carbonate function” – This section defines a function that calculates δ^{13}C of soil carbonate from CO_2 concentrations and mean annual precipitation rates. This is not relevant to this project.
10) “soil carbon profile calculations” – This section calculates profiles of soil carbon isotope values along the mean storm track. This section is not relevant to this project.
11) “plot model soil carbonate profiles” – This section plots the soil carbon profiles. This section is not relevant to this project.
12) “soil water component” – This section calculates the δ^{18}O of soil water profiles along the mean storm track.
13) “plot model soil water profiles” – This section plots the soil water profiles.
14) “modern soil profiles” – This section plots modern soil δ^{18}O data in conjunction with the model-output profiles.
15) “remaking map” – This section replots the map, including locations for compiled water and

soil data and the NARR data locations used.

16) “porosity sensitivity analysis” – This section performs a sensitivity analysis on the initial value of porosity used.

B.2 “non-dimensional model.R”

This file contains the code to predict $\delta^{18}O$, d-excess, and ^{17}O-excess along hypothetical storm tracks with non-dimensional distance. A plot is created for these values at varying E/ET partitioning ratios. All climatological inputs are mostly arbitrary, but the equations are otherwise the same as those in the “bulk code.R” file.

B.3 HYSPLIT/Model Integration

All code relevant to the integration of the model with individual HYSPLIT trajectories is described in Appendix C.

Appendix C. HYSPLIT and Model Integration

A large portion of this project was devoted to integrating the model with individual storm path trajectories output by HYSPLIT, but the results thus far have been unsatisfactory. HYSPLIT trajectories are output as a set of coordinates associated with the number of hours since trajectory initialization (note that these are negative hours due to the back-trajectory calculation performed by HYSPLIT). NARR also contains data as a grid of coordinates and data values associated with a timestamp (hours since 1800-1-1 00:00:0.0). Integrating these for use in the model requires placing NARR data, of appropriate time and location, in the format of a trajectory. The following outlines the steps required to integrate the two dataset and to run the model on individual trajectories. Note that the majority of the code mentioned below requires a computer with access to large volumes of memory (RAM).

C.1 Prepare HYSPLIT trajectories for use in R: “Trajectory Plotter-NARR_AJR.R”

The “Trajectory Plotter-NARR_AJR.R” file is used to manipulate the raw HYSPLIT output. It is used to separate out the precipitating trajectories, generate their histograms, plot them, and output RData files for all of the precipitation-producing trajectories from a given initialization point. These RData files, stored in “HYSPLIT output” (see Appendix A.2), will be used later for integration.

C.2 Prepare the NARR dataset for use in R: "dataset processing.R"

Our first step in integrating the model with HYSPLIT trajectories was to make the HYSPLIT and NARR datasets compatible with each other, firstly by making all of the NARR data available for use in R. This was done in "dataset processing.R," in which all of the NARR NetCDF files are opened and saved as .RData files by year (stored in "NARR_RData"; see Appendix A.2). It is important to mention that each of these is a very large data file, and care must be taken so as not to exceed a computer's memory capacity. This code functions with on a computer with 16 GB RAM (lower RAM values were not tested).

C.3 Convert HYSPLIT timestamps to NARR-compatible timestamps: "timechanger.R"

The next step is to make the timestamps for HYSPLIT (year/month/day/hour) compatible with those for the NARR dataset (hours since 1800-1-1 00:00:0.0). This is done through the function defined in "timechanger.R." This function takes a dataframe of meteorological variables (the direct HYSPLIT output), and adds an additional column for the modified timestamp to match the hours since 1800-1-1 00:00:0.0 format used in the NARR data.

C.4 Convert HYSPLIT coordinates to NARR-compatible grid: "latlonNARR.R"

Though the HYSPLIT output is based on the raw NARR data, the output coordinates are interpolated values from the NARR grid values. Thus, to integrate the two datasets, we could either interpolate the NARR dataset or set the HYSPLIT output to be compatible with the NARR grid. We choose the latter both to avoid increased error due to interpolation, and because interpolating the NARR dataset to the resolution of HYSPLIT output is computationally quite expensive. Thus, we use the function defined in "latlonNARR.R" to use a nearest-neighbor search to select the NARR coordinate closest in space to each HYSPLIT trajectory coordinate. This function takes in a dataframe of a time-corrected HYSPLIT output trajectory and replaces the latitude and longitude columns with the appropriate NARR coordinates for each point in the trajectory.

C.5 Integrate NARR data into trajectory points: "NARRgrabber.R"

The final step in integrating the NARR and HYSPLIT output datasets is to associate each point in each trajectory with the appropriate NARR climatological data at that location and time.

This is handled by the functions defined in “NARRgrabber.R,” which take in a dataframe of corrected coordinates and timestamps and append columns for the NARR climatological variables needed for the vapor transport model (Damköhler number, precipitable water, relative humidity, air temperature, and land mask cover). This function needs access to all of the NARR dataset; each year is about 2.3 GB of data. The number of files that are in use at any given point is minimized, but large volumes of memory are still required. This function will not crash a computer with 16 GB of RAM, but lower RAM values were not tested.

C.6 Establish equations necessary to run the model: “model.sourcefunctions.R”

Several functions to calculate isotopic values are used by the vapor transport model. Those for fractionation and for calculating the integrated isotopic composition of evapotranspiration are defined in “model.sourcefunctions.R.” This file also contains functions for soil carbon calculations, which are not currently relevant to this aspect of the project.

C.7 Run the vapor transport model along a single trajectory: “Hendricksmodel.onetraj.R”

Once the coordinates and timestamps have been corrected and the appropriate NARR data has been extracted for each trajectory, we can run the model along a single trajectory using the function defined by “Hendricksmodel.onetraj.R.” This function takes in a dataframe of chronologically ordered meteorological values (each row is a time point, each column is an input), runs the model along this trajectory, and appends columns for the model-output isotopic values. The current setup of the function only outputs for $\delta^{18}O$, but including δD and $\delta^{17}O$ only requires uncommenting one line of code and is clearly demarcated. It is also worth noting that the “Hendricksmodel.onetraj.R” file also contains the values for constants needed for the model. This function is meant to be used as part of the script “atmosphere.fromHYSPLIT.R,” described below with a more complete description of the necessary input format.

C.8 Full application of model to HYSPLIT output: “atmosphere.fromHYSPLIT.R”

The final aspect of integrating the vapor transport model to individual HYSPLIT trajectories was to run the model along multiple individual trajectories. This is handled in “atmosphere.fromHYSPLIT.R.”

First, all appropriate function files are loaded. Next, a list of dataframes as generated: the list represents all of the trajectories for a given location/season, and each list element contains a dataframe corresponding to an individual trajectory. For each dataframe in the list, the coordinates and timestamps are corrected. This will take time. Next, for each list element (each dataframe trajectory), the "NARRgrabber" function is called to append appropriate NARR data to each row of each dataframe. This step is extremely computationally expensive and will take quite some time. The list is then saved (*e.g.* "list.Hay.DJF.withNARR.RData" as in Appendix A.4) such that re-computing is not necessary. Next, all NA points (those without NARR data) are removed from the dataframes, and each dataframe is re-ordered to chronological order (because original trajectory outputs are in reverse chronological order).

To finalize the dataframes before input into the model, we must append values of "t," defined as the ratio of transpiration to total evapotranspiration (T/ET), to each point along each trajectory/dataframe in the list. The current setup of the code appends a single value to all points along each trajectory, though it is possible to assign unique values to each point (not handled here) to allow for changing T/ET values along a trajectory. Now, we have a list of dataframes to which we can apply the "Hendricksmodel.onetraj" function. Each dataframe is formatted as follows: Each row represents a point along a single trajectory, where rows are in chronological order (earliest timestamp at first row). There are 11 columns: Year, Back.Hour, Lat, Lon, Time (hours since 1800-1-1 00:00:0.0), Land (land mask cover; 0 or 1), Nd (Damköhler number), Prw (precipitable water), rhum (relative humidity as percent), Temp_2m (2m air temperature, Kelvin), t (T/ET ratio). The output is a list of dataframes, where each dataframe represents a single trajectory with meteorological variables and the isotopic values output by the vapor transport model for each point along the trajectory.

While we were successfully able to produce code for the computationally expensive problem of integrating individual trajectories with NARR data and our vapor transport model, the results are unsatisfactory and suggest room for improvement due to the following three limitations. First, the NARR dataset is imperfect, particularly at local scales. The presence of negative precipitation rates and otherwise questionable values at some locations introduces many potential problems when these data are used at the scale of individual complexities. Second, evapotranspiration must be better constrained. At this point, we are assuming ratios that are likely not fully representative of the actual values experienced along a storm path. Third, and perhaps most importantly, this model assumes constant initial isotopic compositions for each trajectory. Parts of the model will have to be adjusted to consider spatially and temporally variable initial compositions.