Title: Dataset associated with "Analysis of Kenya's Atmospheric Moisture Sources and Sinks"

Abstract: Achievement of the United Nations Sustainable Development Goals (SDGs) are contingent on understanding the potential interactions among human and natural systems. In Kenya, the goal of conserving and expanding forest cover to achieve SDG 15 'Life on land' may be related to other SDGs because it plays a role in regulating some aspects of Kenyan precipitation. We present a 40-year analysis of the sources of precipitation in Kenya, and the fate of the evaporation that arises from within Kenya. Using MERRA2 climate reanalysis and the Water Accounting Model 2-layers, we examine the annual and seasonal changes in moisture sources and sinks. We find that most of Kenya's precipitation originates as oceanic evaporation, but that 10% of its precipitation originates as evaporation within Kenya. This internal recycling is concentrated in the mountainous and forested Kenyan highlands, with some locations recycling more than 15% of evaporation, to Kenyan precipitation. We also find that 75% of Kenyan evaporation falls as precipitation elsewhere over land, including 10% in Kenya, 25% in the Democratic Republic of the Congo, and around 5% falling in Tanzania and Uganda. Further, we find a positive relationship between increasing rates of moisture recycling and fractional forest cover within Kenya. By beginning to understand both the seasonal and biophysical interactions taking place, we may begin to understand the types of leverage points that exist for integrated atmospheric water cycle management. These findings have broader implications for disentangling environmental management and conservation and have relevance for large-scale discussions about sustainable development.

Contact: Patrick Keys, patrick.keys@colostate.edu

Recommended data citation – Keys, P.W., Warrier R., van der Ent, R.J., Galvin, K. A. and Boone, R.B. (2021). Dataset associated with "Analysis of Kenya's Atmospheric Moisture Sources and Sinks." Colorado State University. Libraries. http://dx.doi.org/10.25675/10217/235433

Associated article citation – Keys, P. W., Warrier, R., van der Ent, R. J., Galvin, K. A., & Boone, R. B. (2022). Analysis of Kenya's Atmospheric Moisture Sources and Sinks, *Earth Interactions, 26*(1), 139-150. https://doi.org/10.1175/EI-D-21-0016.1

Format of data files – .mat files

File Information — The moisture recycling data for the 40-year analysis of the sources of Kenyan precipitation and the sinks of Kenyan evaporation. There are two files in the repository:

1. Back_tracking_data_1980-2019.mat
2. Forward_tracking_data_1980-2019.mat

Each file is a four-dimensional matrix in terms of: [YEAR, MONTH, X-COORDINATE, Y-COORDINATE] The files are global in domain, and the years proceed in the order 1980, 1981….2018, 2019. The month are from Jan, Feb… Nov, Dec. The data span 80 degrees north to 80 degrees south, in increments of 0.5 degrees. The data span from 0 degrees west to 360 degrees west, in increments of 0.625 degrees. The units of all data files are in cubic meters of water.

Method(s) –
****Please note that the following text is taken from the Methods section of the linked publication. See "Associated article citation" above:*

"We employ the Modern Era Retrospective Reanalysis version 2 (MERRA2) climate reanalysis dataset (Gelaro et al. 2017), specifically variables pertaining to the atmospheric water cycle (Supplementary Table 1). The MERRA2 water cycle representation has been examined globally and in regional detail, and reveals both improvements and continued challenges, compared to both past and contemporary reanalysis data products (Bosilovich et al. 2017).

The Water Accounting Model 2 layers (WAM-2layers) is a global, Eulerian moisture tracking model that reconstructs the atmospheric water cycle - from water's evaporative origin on the planet, through the atmosphere, and to its fate elsewhere as precipitation (van der Ent et al. 2014). The WAM-2layers is a flexible model that can be used with gridded climate data of varying resolutions (Findell et al. 2019; Guo et al. 2020). Past work compared global and regional results of using the WAM-2layers with the ERA-Interim climate reanalysis (Dee et al. 2011) relative to the MERRA version 1.0 climate reanalysis (Rienecker et al. 2011), and found strong global fidelity between the two datasets, as well as strong regional similarities, with some isolated differences (Keys et al. 2014).

There are eleven MERRA2 variables that are used as input for the updated WAM-2layers and are summarized in the Supplemental Table 1. The three-dimensional data were downloaded on a model-level grid, rather than pressure-level, to avoid issues related to moisture being lost in high altitude regions during the WAM-2layers runs (van der Ent et al. 2014; Keys et al. 2014). All data are publicly accessible and were downloaded from the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC). All data were downloaded at the 0.5° latitude x 0.625° longitude MERRA2 spatial resolution, for the period 1980-2019. The two-dimensional variables were downloaded at the 1-hour temporal resolution, and the three-dimensional variables were downloaded at the 6-hour resolution. We detail key metadata for each of the data types used in the analysis in Supplemental Table 1.

The WAM-2layers was modified to accommodate the different input data. Specifically, changes were made to how the model reads in the vertical flux data (i.e., column water, evaporation, and precipitation), to the timestep of some calculations (reflecting the MERRA2 time discretization), and to the atmospheric boundary separating the upper and lower levels of the atmosphere. The WAM2-layers has two key steps. First, we calculated 'fluxes and states' of the water balance globally. This step tracked water for every gridcell as it originated on the land surface as evaporation, its trajectory through the lower and upper layers of the atmosphere, and where the water precipitated (either in the same gridcell or elsewhere). To preserve that atmospheric water balance, the input data are discretized from either 6- or 1-hourly data, down to a 7.5-minute timestep. Longer timesteps would risk missing water passing quickly through small gridcells, since the horizontal distance between degrees of longitude becomes smaller in moving poleward from the equator. Shorter timesteps allow the WAM2-layers to account for water moving through smaller gridcells (van der Ent et al., 2010).

In the second step, we used the fluxes and states information to either (a) track the evaporation of a specific location forward to its fate as precipitation (i.e., stepping forward in time), or (b) track the precipitation of a specific location backward to its origin as evaporation (i.e., stepping backward in time). Thus, the WAM-2layers was first run for 40 years at the global scale, to calculate the fluxes and states for all locations. Second, the WAM-2layers was used to track a country-specific run for Kenya, tracking moisture both forward and backward in time.

Previous work introduced the idea of terrestrial moisture recycling ratios. Evaporation recycling ratios are defined as the fraction of precipitation arising from a specific location's evaporation, c. In general, terrestrial evaporation recycling ratios are defined as,

$$\epsilon_c = \frac{E_c}{E_c + E_o}$$

where, Ec is the evaporation arising from terrestrial (i.e., land) sources, Eo is the evaporation arising from oceanic sources, and c is the ratio of terrestrial to oceanic sources. Precipitation recycling ratios are defined as the fraction of evaporation that falls as precipitation on a specific location, c. In general, terrestrial precipitation recycling ratios are defined as,

$$\rho_c = \frac{P_c}{P_c + P_o}$$

where, Pc is the evaporation arising from terrestrial (i.e., land) sources, Po is the evaporation arising from oceanic sources, and c is the ratio of terrestrial to oceanic sources. We note specifically in this paper when the definition of moisture recycling ratios departs from these definitions."

Original MERRA2 data links

The original MERRA 2 data that were used are summarized in the Supplement of the linked article, and the files can be found at the NASA website here:

DATA TYPE	MERRA2 DATA NAME	SOURCE DATA FILE	TEMPORAL RESOLUTION	DOI
Surface pressure	Surface pressure	inst1_2d_asm_Nx	1 hour	10.5067/3Z173KIE2TPD
Evaporation	EV Land	tavg1_2d_lnd_Nx	1 hour	10.5067/RKPHT8KC1Y1T
Precipitation	Bias corrected precipitation	tavg1_2d_flx_Nx	1 hour	10.5067/7MCPBJ41Y0K6
Zonal wind	U	inst6_3d_ana_Nv	6 hour	10.5067/IUUF4WB9FT4W
Meridional wind	V	inst6_3d_ana_Nv	6 hour	10.5067/IUUF4WB9FT4W
Specific humidity	Q	inst6_3d_ana_Nv	6 hour	10.5067/IUUF4WB9FT4W
Total column water	Column liquid water	inst1_2d_int_Nx	1 hour	10.5067/G0U6NGQ3BLE0
	Column ice water	inst1_2d_int_Nx	1 hour	10.5067/G0U6NGQ3BLE0
	Column water vapor	inst1_2d_int_Nx	1 hour	10.5067/G0U6NGQ3BLE0
Corrected eastward moisture flux	Ecorr flux	tavg1_2d_int_Nx	1 hour	10.5067/Q5GVUVUIVGO7
Corrected northward moisture flux	Ncorr flux	tavg1_2d_int_Nx	1 hour	10.5067/Q5GVUVUIVGO7