Title: 'ERA5-based database of Atmospheric Rivers over Himalayas'

Author(s): Munir Ahmad Nayak et. al

Manuscript No.: essd-2020-397

Manuscript type: Data description paper

Special Issue: Extreme environment datasets for the three poles

Referee 3

The authors created a dataset of atmospheric rivers for the Himalayan region derived from ERA5 data. This dataset could be useful for the community for research into extreme precipitation and flooding. The manuscript is well written and presented.

Response:

We thank the Referee for appreciating our manuscript and his/her thoughtful comments. We agree with the Referee and believe that the dataset will advance AR studies over the unexplored Himalayas.

I think the manuscript could benefit from a bit more explanation on the choice of the AR detection algorithm and the chosen time step. Why did the authors decide to use 6-hourly data despite ERA5 being available on a higher temporal resolution? What made the authors choose this AR identification method over other available methods?

Response:

Thank you for the comments. We have used the modified version of Lavers et al., (2012) algorithm for AR identification in this study as the algorithm is region-specific and allows for a space-time varying threshold criteria. This allows identification of ARs of "weak" and "moderate" ARs in this region in all seasons; for example, in winter season in the western Himalaya, ARs rarely cross the 500 $kg\ m^{-1}s^{-1}$ (Figure 3) IVT, which will be disqualified by strict high-threshold algorithms. We wanted to include “weak” ARs in the database as these can have influence on this regional precipitation and important hydrological impacts as observed in some recent studies over other cold regions (Gorodetskaya et al., 2014; Nash et al., 2018; Wille et al., 2019). Also, we would like to highlight that two recent review studies on comparison of AR algorithms (Lora et al., 2020; Rutz et al., 2019) concluded that most of these algorithms identify ARs with fairly good agreement, especially the moderate to intense ARs, and that high-threshold algorithms ignore “weak ARs” from their records. The “weak” ARs are often the instances of ARs formation, dissipation, and merging (Lora et al., 2020), hence can provide important insights into the dynamical evolution of ARs formation and may have important societal impacts.

We have used 6-hourly ERA5 datasets because of four main reasons:

(1) This temporal resolution is commonly used in AR-detection algorithms when using global reanalysis products (Nash & Carvalho, 2020; Waliser & Guan, 2017),

(2) Our main goal is to identify ARs in the Himalayas and provide a ready-to-use and easily-manageable AR database for AR studies over this region for a sufficiently long period i.e., 37 years. We realize that for such a lengthy duration (including all the seasons) to reduce the data volume, 6-hourly analysis is sufficient to produce a distinct and manageable database that can be loaded in most of the software on a home desktop machine. In contrast, 1-hourly AR data will consume more RAM due to larger size, thereby reduce the system performance, while only adding marginal information than 6-hourly data.

(3) 6-hourly datasets provide sufficient temporal information to show the gradual evolution of AR over time (Nash et al., 2018; Ramos et al., 2015), rather than abrupt changes. It is worth mentioning that most climate model simulations for ARs are also archived at this temporal resolution.

(4) A previous study (Rutz et al., 2014) has found similar results in mean AR duration when 6-hourly ERA-Interim IVT dataset is used compared to 1-hourly observational based dataset used by an earlier study (Ralph et al., 2013) for the same study area in Bodega Bay, US West Coast. Another study (Dettinger, 2011) also observed similar results in AR duration when daily observations are used instead of 1-hourly in northern California.

When looking into the dataset I think there could be a bit more additional information on how the data is organised. I am not sure that someone downloading the dataset would be able to understand it in its current form. For example, it took me a while to figure out that a detected AR has a unique id but still has multiple rows as it consists of multiple timesteps. The description in the read me file is very short and could say more about the structure in the .csv files, e.g. that there is a line for every time step in an identified AR. The manuscript and meta data say that the covered period is 1982-2018 while the first detected AR in the files is from January 1979. For one AR timestep the IVT max says one value but when looking into the columns there is a higher IVT value. It seems a bit complicated organised that the longitudes and latitudes corresponding to the AR locations are in different files from the actual IVT values.

Response

We agree with the Reviewer, perhaps we were not detailed enough. We have updated the readme file, which now reads as:

We have also included a note regarding the 1979 to 1981 ARs, where we mention that cyclone dates have not been removed in these years due to unavailability of cyclone dates, so some cyclones may have been identified as ARs in this period.

-----Readme document text starts here-----

"Atmospheric Rivers (ARs) are long and narrow regions of intense moisture transport in the lower troposphere. The dataset comprises of Atmospheric Rivers that have happened over the Himalayan Basins from 1982 to 2018. It includes the dates and times, duration, intensity/magnitude, tracks, and categories of the ARs.

File Names and description:

1. ERA5_Persistant_Database2000km: This file includes the date, times, average Integrated Water Vapor Transport (IVT) magnitude ($kg.m^{-1}.s^{-1}$), starting IVT, maximum IVT, and duration of ARs. These terms are explained below in greater details.

Column "Date":

Gives the date and time (in Coordinated Universal Time UTC) of each AR timestep. The IVT data used to identify ARs is 6-hourly (00UTC, 06UTC, 12UTC and 18UTC).

Column "AR_ID":

Each identified persistent AR, lasting for at least 18 hours, is given a unique ID, which remains same for all timesteps of the AR. This column gives the ID of ARs. The ID of an AR is based on the year in which the AR occurred, the letters "AR", and the occurrence serial of the AR in the year. For example, the first AR in 1990 has ID 1980AR1. If the AR lasted for 10 timesteps, all 10 timesteps will have the same ID.

Column "Ind":

This column gives the python index of IVT data in 6-hour yearly data, giving the date and time of each AR timestep. This column can be ignored since the same information is more directly available in "Date" column.

Column "AvgIVT":

This column gives the average IVT magnitude ($kg.m^{-1}.s^{-1}$) along the AR major axis, i.e., the gridcells that have maximum IVT along the AR track. For example, the first value corresponds to the average of all values from column *"0"* to column *"88",* which give the IVT magnitude at each gridcell of the major axis of the first timestep.

Column "StartIVT":

This column gives the IVT magnitude ($kg.m^{-1}.s^{-1}$) at the initial gridcell on the first timestep when AR condition was identified.

Column "ARDuration":

This column gives duration of the AR in hours; for example, an AR lasting for three timesteps will have the duration of 18 hours, an AR lasting for four timesteps will have duration of 24 hours.

Column "MaxIVT":

This column gives the maximum of all IVT values ($kg.m^{-1}.s^{-1}$) at the starting gridcells on each timestep of an AR.

Column "ARCat":

This column gives category of the AR, based on IVT magnitude and duration of the ARs. Six categories have been defined, Cat0 denoting the weakest AR and Cat5 denoting the strongest AR. More details on this can be found in the accompanying paper.

Column "0" to the end.

These columns give the IVT magnitude ($kg.m^{-1}.s^{-1}$) at each gridcell of the major axis of each AR timestep.

Note that the cyclone dates were not available before 1982, so AR dates for 1979 to 1981 includes cyclonic IVT structures.

2. ERA5_Persistant_Database_lats_2000km: The file gives the latitudes of grid points of maximum IVT, i.e., the latitude of major axes of ARs throughout their duration.

Columns "Date", "AR_ID", "Ind", "AvgIVT", "StartIVT", "ARDuration", "MaxIVT", "ARCat" are the same as given above for "ERA5_Persistant_Database2000km.csv" file.

Column "0" to end.

These columns give the latitude (in degrees North) at each gridcell of the major axis of each AR timestep.

3. ERA5_Persistant_Database_lons_2000km: The file gives the longitudes of grid points of maximum IVT, i.e., the longitudes of major axes of ARs throughout their duration

Columns "Date", "AR_ID", "Ind", "AvgIVT", "StartIVT", "ARDuration", "MaxIVT", "ARCat" are the same as given above for "ERA5_Persistant_Database2000km.csv" file.

Column "0" to end.

These columns give the longitude (in degrees East) at each gridcell of the major axis of each AR timestep"

-----Readme document text end -----

Line 246-247: there is "southward" twice in this sentence, while I think one of them should be "eastward".

Response

Corrected, thank you for pointing this out.

References:

Dettinger, M. D. (2011). Climate change, Atmospheric Rivers, and floods in California - A Multimodel analysis of storm frequency and magnitude changes. *JAWRA Journal of the American Water Resources Association*, *47*(3), 514–523. https://doi.org/10.1111/j.1752-1688.2011.00546.x

Gorodetskaya, I. V., Tsukernik, M., Claes, K., Ralph, M. F., Neff, W. D., & Van Lipzig, N. P. M. (2014). The role of Atmospheric rivers in anomalous snow accumulation in East Antarctica. *Geophysical Research Letters*, *41*(17), 6199–6206. https://doi.org/10.1002/2014GL060881

Guan, B., & Waliser, D. E. (2017). Atmospheric rivers in 20 year weather and climate simulations: A multimodel, global evaluation. *Journal of Geophysical Research: Atmospheres*, *122*(11), 5556–5581. https://doi.org/10.1002/2016JD026174

Lavers, D. A., Vallarini, G., Allan, R. P., Wood, E. . F., & Wade, A. J. (2012). The detection of atmospheric reanalyses and their links to British winter floods and the large-scale climatic circulation. *Journal of Geophysical Research: Atmospheres*, *117*(D20). https://doi.org/10.1029/2012JD018027.

Lora, J. M., Shields, C. A., & Rutz, J. J. (2020). Consensus and disagreement in Atmospheric river detection: ARTMIP global catalogues. *Geophysical Research Letters*, *47*(20). https://doi.org/10.1029/2020GL089302

Nash, D., & Carvalho, L. M. V. (2020). Brief Communication: An electrifying atmospheric river – understanding the thunderstorm event in Santa Barbara County during March 2019.

Natural Hazards and Earth System Sciences, *20*(7), 1931–1940. https://doi.org/10.5194/nhess-20-1931-2020

Nash, D., Waliser, D., Guan, B., Ye, H., & Ralph, F. M. (2018). The role of Atmospheric rivers in extratropical and polar hydroclimate. *Journal of Geophysical Research: Atmospheres*, *123*(13), 6804–6821. https://doi.org/10.1029/2017JD028130

Ralph, F. M., Coleman, T., Neiman, P. J., Zamora, R. J., & Dettinger, M. D. (2013). Observed impacts of duration and seasonality of Atmospheric-river landfalls on soil moisture and runoff in Coastal Northern California. *Journal of Hydrometeorology*, *14*(2), 443–459. https://doi.org/10.1175/JHM-D-12-076.1

Ralph, F. M., Rutz, J. J., Cordeira, J. M., Dettinger, M. D., Anderson, M., Reynolds, D., et al. (2019). A scale to characterize the strength and impacts of Atmospheric rivers. *Bulletin of the American Meteorological Society*, *100*(2), 269–289. https://doi.org/10.1175/BAMS-D-18-0023.1

Ramos, A. M., Trigo, R. M., Liberato, M. L. R., & Tomé, R. (2015). Daily precipitation extreme events in the Iberian Peninsula and its sssociation with Atmospheric Rivers. *Journal of Hydrometeorology*, *16*(2), 579–597. https://doi.org/10.1175/JHM-D-14-0103.1

Rutz, J. J., Steenburgh, W. J., & Ralph, F. M. (2014). Climatological characteristics of Atmospheric rivers and their inland penetration over the western United States. *Monthly Weather Review*, *142*(2), 905–921. https://doi.org/10.1175/MWR-D-13-00168.1

Rutz, J. J., Steenburgh, W. J., & Ralph, F. M. (2015). The inland penetration of Atmospheric rivers over Western North America: A lagrangian analysis. *Monthly Weather Review*, *143*(5), 1924–1944. https://doi.org/10.1175/MWR-D-14-00288.1

Rutz, J. J., Shields, C. A., Lora, J. M., Payne, A. E., Guan, B., Ullrich, P., et al. (2019). The Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Quantifying uncertainties in Atmospheric river climatology. *Journal of Geophysical Research: Atmospheres*, *124*(24), 13777–13802. https://doi.org/10.1029/2019JD030936

Waliser, D., & Guan, B. (2017). Extreme winds and precipitation during landfall of Atmospheric rivers. *Nature Geoscience*, *10*(3), 179–183. https://doi.org/10.1038/ngeo2894

Wille, J. D., Favier, V., Dufour, A., Gorodetskaya, I. V., Turner, J., Agosta, C., & Codron, F. (2019). West Antarctic surface melt triggered by Atmospheric rivers. *Nature Geoscience*, *12*(11), 911–916. https://doi.org/10.1038/s41561-019-0460-1