[Rainfall data](https://eur03.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdata.earthobservation.vam.wfp.org%2Fpublic-share%2Fclimdata-sahel%2Fsahel_rfh_20240618.csv&data=05%7C02%7CA.Hema%40cgiar.org%7Cfd311602b7ae4418636a08dc903717c4%7C6afa0e00fa1440b78a2e22a7f8c357d5%7C0%7C0%7C638543813536021628%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=kM5bVKA%2FNd46e7ZreKPOPx84g8dCQQZHUs6E01DEYgw%3D&reserved=0):

This is based on CHIRPS dekad and includes

* *rfh:* the rainfall (mm)
* *rfq*: the anomaly in percent
* *rfh\_avg:* the long term average calculated over the period 1989 - 2018 (also the reference for the anomaly)
* *rn{h,q,s}*: these indicators are the aggregation (*h*), aggregation anomalies (q) and SPIs for the duration (*n*) in months (which is one of 1,2,4,6,8, y) with *y* being the 12-month aggregation period. The aggregations also come with the respective long term averages (*\_avg*) used for the anomalies.

[NDVI data](https://eur03.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdata.earthobservation.vam.wfp.org%2Fpublic-share%2Fclimdata-sahel%2Fsahel_vim_20240618.csv&data=05%7C02%7CA.Hema%40cgiar.org%7Cfd311602b7ae4418636a08dc903717c4%7C6afa0e00fa1440b78a2e22a7f8c357d5%7C0%7C0%7C638543813536032047%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=vpV37wf66gjHUPf7N8gUfXAzMxlqTpcPZkm6hCJxaro%3D&reserved=0):

This is based on our cloud-filtered and gapfilled MODIS 5k NDVI in dekads, and includes

* *vim*: the NDVI
* *vim\_avg:* the NDVI long term average (calculated over the period 2002 - 2018)
* *viq*: the NDVI anomaly (in percent)

[Land Surface Temperature (LST)](https://eur03.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdata.earthobservation.vam.wfp.org%2Fpublic-share%2Fclimdata-sahel%2Fsahel_lst_20240618.csv&data=05%7C02%7CA.Hema%40cgiar.org%7Cfd311602b7ae4418636a08dc903717c4%7C6afa0e00fa1440b78a2e22a7f8c357d5%7C0%7C0%7C638543813536041047%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=7nd75Mwx8iWDBLsKFPcWb6reX5f1%2F8PVTAU1t5twzZA%3D&reserved=0):

This is based on our cloud-filtered and gapfilled MODIS 5k LST in dekads, and includes

* *tda*: Daytime LST
* *tna:* Nightime LST
* *taa*: Diurnal difference (aka amplitude) - so difference between Day & Night

[Reference Evapotranspiration:](https://eur03.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdata.earthobservation.vam.wfp.org%2Fpublic-share%2Fclimdata-sahel%2Fsahel_et0_20240618.csv&data=05%7C02%7CA.Hema%40cgiar.org%7Cfd311602b7ae4418636a08dc903717c4%7C6afa0e00fa1440b78a2e22a7f8c357d5%7C0%7C0%7C638543813536051617%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=08WlwYguAEyS4U8o6quWslNjs9kNq4%2BwvGRFMCH9x9s%3D&reserved=0)

This is based on the FAO WAPOR V3 Reference Evapotranspiration dekad, and includes

* *et0*: Evapotranspiration (mm)
* et0\_avg: long term average of ET (calculated over the period 1989 - 2018)
* *etq*: Anomaly of Evapotranspiration (in percent)

Please note the following:

All these datasets contain the columns

* *time*: timestamp for each observation. Since all of the above are dekads, these map to the 1st , 11th and 21st for each respective month.
* *location*: these are the PCODES used for aggregation based on the shapefile provided by you.
* *valids*: the number of valid pixels per location - some of the admin-2s are fairly small for 5km data, so if you see a low number of valid pixels, you should take this into consideration when using these datapoints.

I have three quick and (I guess) very easy questions for you:

1. We are unable to exactly retrieve the long-term average and anomalies for the different variables using the data at adm2 level. Are we correct that both are first derived at the pixel-level, before they were aggregated to the adm2 level?

      For this application, we calculate the LTA and anomalies from the spatially averaged values - so per admin area we average the observation across the timeseries and then calculate the downstream indicators.

You should be therefore able to reproduce the computations if you account for:

* LTAs for rainfall indicators are calculated over the period from 1989-01-01 to 2018-12-31
* LTAs for NDVI indicators are calculated over the period from 2002-07-01 to 2018-07-01

1. The anomalies are expressed as percentages. Am I correct that this is just the ratio of the observed value divided by its corresponding long-term average (at pixel level); or is there another procedure being followed?

      The anomalies are a ratio for the rainfall & NDVI products, but as per above between the spatially averaged observations and the respective LTA values. In addition to that, we add a small constant to the computation to avoid excessive anomalies with small observations:

* for rainfall / et0 we add 5 (mm)
* for NDVI we ad 0.05

The anomalies for LST are just the difference in degrees, without any constants.

1. Is there any way we can reference your work. For example, you mention the use of cloud-filter & gap-filler techniques on the MODIS database. Can this be cited somehow?

We do have a fairly old paper for the Whittaker smoother we developed for the cloud filtering. You can find the citation info here: [https://ieeexplore.ieee.org/document/8076705](https://eur03.safelinks.protection.outlook.com/?url=https%3A%2F%2Fieeexplore.ieee.org%2Fdocument%2F8076705&data=05%7C02%7CA.Hema%40cgiar.org%7Ceedb94cf0bed42182bf408dc91756b2f%7C6afa0e00fa1440b78a2e22a7f8c357d5%7C0%7C0%7C638545180724312902%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=o7029vahZsfS8sXwsATimXpnGIOg4mmrwmR%2BS68Klpo%3D&reserved=0)

Otherwise, I don't think there's much to reference - of course you can always refer to WFP, and more specifically the GIS & Remote Sensing unit. If you want to get very specific, this is all a product of our data processing infrastructure - the Humanitarian Data Cube - which allows us to process EO data at a global scale and derive these downstream products rather quickly. .