Analysis and visualization of terrestrial water storage anomaly estimates from GRACE(-FO)

In this exercise, GRACE(-FO) Mascon data (JPL RL06M v04) will be used to generate time series of terrestrial water storage anomalies (TWSA) and investigate their temporal variability and long-term trends over major river basins. The target basins are the **Amazon, Mississippi, Ob, Danube, Po, and Yukon**. In addition, Greenland will be analyzed to investigate ice-mass changes.

1. Using the NetCDF file `jpl_global_mascons.nc`, which is downloaded from `https://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons`, generate a global map of TWSA for any selected month.

   NetCDF data

   NetCDF (*Network Common Data Form*) is a self-describing file format commonly used for multidimensional Earth-observation data. It stores the data together with information about dimensions, coordinates, units, and metadata.
   The file `jpl_global_mascons.nc` has the main dimensions

   $$\texttt{lon} = 720, \qquad \texttt{lat} = 360, \qquad \texttt{time} = 258.$$

   The main variable used in this exercise is

   $$\texttt{lwe_thickness(lon,lat,time)},$$

   with size $720 \times 360 \times 258$ and unit **cm**. It represents the GRACE/GRACE-FO terrestrial water storage anomaly as liquid water equivalent thickness.
   Use `lon`, `lat`, and `time` as the corresponding coordinates. Note that `time` is given in days since `2002-01-01`, and longitude ranges from 0° to 360°.

2. Calculate the linear trend over the time series for each Mascon/grid. Then visualize the resulting trends (in cm/year) on a global map.

   Instructions

   For every Mascon/grid cell, extract its complete TWSA time series and fit a linear model of the form

   $$\Delta S(t) = a + bt,$$

   where b represents the temporal trend.
   Make sure that the time variable is expressed consistently in years so that the resulting slope has units of storage change per year.
   GRACE(-FO) contains missing months. These should not be interpreted as no value. Exclude missing observations from the regression.
   After estimating the trend independently for all Mascon/grid cells, create a global trend map using a diverging color scale centered at zero.

   You can investigate the resulting spatial pattern:

   - Where are the strongest losses and gains observed?
   - Which regions show stable or only minor changes in water storage?

MJ Tourian
touriah@gis.uni-stuttgart.de

3. Generate the area-weighted TWSA time series for the following six river basins:

   Amazon, Mississippi, Ob, Danube, Po, and Yukon.

   Generate TWSA time series covering the entire available GRACE/GRACE-FO period for all six basins. Plot the time series and compare their temporal behavior.

Basin averaging

Read the provided shapefile `basin.shp`, which contains the boundaries of the river basins and Greenland. Select the target regions using the following IDs:

Region	CATCH_ID	Region	CATCH_ID
Amazon	193	Mississippi	29
Ob	183	Danube	24
Po	38	Yukon	78
Greenland	409		

Use the corresponding polygon boundaries to determine which GRACE(-FO) grid cells belong to each region. The basin-average TWSA should be calculated using an **area-weighted** mean:

$$\overline{\Delta S}_{\mathrm{b}}(t) = \frac{\sum_i A_i \Delta S_i(t)}{\sum_i A_i},$$

where A_i is the area of grid cell i and $\delta S_i(t)$ is its TWSA.
For a regular longitude–latitude grid, calculate the area of each grid cell on a spherical Earth as

$$A_i = R^2 \Delta\lambda \left(\sin\phi_{\mathrm{N}} - \sin\phi_{\mathrm{S}}\right),$$

where $R = 6371$ km is the mean Earth radius, $\Delta\lambda$ is the grid-cell width in longitude, and ϕ_{N} and ϕ_{S} are the northern and southern latitude boundaries of the grid cell. All angles in this equation must be expressed in radians. The variables `lat_bounds` and `lon_bounds` in the NetCDF file provide the grid-cell boundaries.

4. For each basin, determine the linear trend in TWSA using linear regression.

   You can investigate the following points:

   - Which basin shows the largest seasonal variations in terrestrial water storage?
   - Which basins show increasing or decreasing long-term water-storage trends?
   - Are there notable extreme wet or dry years in the individual basins?
   - How do the temporal characteristics of TWSA differ among the six basins, and what hydrological or climatic processes could explain these differences?

5. Calculate the TWSA for Greenland and determine the annual trend in terms of mass that Greenland uses per year.

Greenland mass change

Use the Greenland polygon from the provided shapefile to identify all GRACE Mascons/grid cells belonging to Greenland.
For Greenland, the quantity of interest is the **total mass change**, rather than only the mean equivalent water height.
Convert equivalent water height to water mass using

$$\Delta M(t) = \rho_w \sum_i A_i \, \Delta S_i(t),$$

where ρ_w is the density of water, A_i is the area of the corresponding Mascon/grid cell, and ΔS_i is expressed in metres.
Recall that

$$1 \text{ Gt} = 10^{12} \text{ kg}.$$

Generate the Greenland mass-anomaly time series in Gt and estimate its linear trend in Gt/year.
Pay attention to the sign: a negative trend corresponds to net mass loss.

Investigate:

- How much ice does Greenland lose per year (in gigatonnes)?
- Are the results consistent with values reported in the literature?

Terrestrial water balance closure using GRACE(-FO)

The terrestrial water balance can be written as

$$P - \text{ET} - R = \frac{dS}{dt}, \tag{1}$$

where P is precipitation, ET is evapotranspiration, R is runoff, and dS/dt represents the temporal change in terrestrial water storage. GRACE(-FO) equivalent water height (ΔS) is a state variable with the unit of [m], whereas the terms on the left-hand side of Equation 1 are fluxes with the unit of [m/month]. To employ GRACE(-FO) observations in the water-balance equation, the first temporal derivative of EWH needs to be calculated.
The numerical derivative can be approximated using the central difference:

$$\frac{dS}{dt}[t] = \frac{S[t + \Delta t] - S[t - \Delta t]}{2\Delta t}, \tag{2}$$

where, for monthly GRACE(-FO) solutions, Δt is equal to one month.
The numerical derivative based on the central difference introduces a temporal filtering effect. To ensure consistency between the GRACE-derived dS/dt and the hydrological fluxes, apply the following convolution filter to the water-balance fluxes:

$$h = \begin{bmatrix} \frac{1}{5} & \frac{3}{5} & \frac{1}{5} \end{bmatrix}. \tag{3}$$

For the hydrological fluxes, use two independent datasets for each component of the water balance:

- **Precipitation:** ERA5 and GPCC
- **Evapotranspiration:** ERA5 and GLEAM
- **Runoff:** ERA5 and LISFLOOD

1. Calculate monthly storage changes dS/dt from the GRACE/GRACE-FO TWSA time series for each of the six river basins.

   GRACE(-FO) storage change

   Start from the basin-averaged monthly GRACE(-FO) TWSA time series calculated in the previous section.
   Apply the central-difference approximation in Equation 2.
   Note that the central difference cannot be calculated for the first and last elements of a continuous time series. Special attention is also required around missing GRACE(-FO) months. Do not calculate a central difference across a data gap as if the observations were consecutive months.

2. Prepare the precipitation, evapotranspiration, and runoff time series for each basin.

   Hydrological fluxes

   The hydrological fluxes are provided as separate NetCDF files with a common structure. The data are already **basin-averaged**; therefore, no spatial aggregation is required.
   Each file contains the main dimensions

   $$\texttt{time} = N, \qquad \texttt{basin} = 6,$$

   where N depends on the temporal coverage of the dataset. The six basin columns correspond to

   Amazon, Mississippi, Ob, Danube, Po, Yukon.

   The main variables are `time`, `basin_id`, `basin_name`, and the corresponding hydrological flux. For example, the ERA5 precipitation file contains

   `precipitation_ERA5(time,basin).`

   The `time` variable is given in days since `1970-01-01`. The other flux files follow the same general structure, but may have different numbers of time steps and different temporal coverage.
   Before using the data, check the variable attributes and make sure that all fluxes have consistent units and sign conventions. Use only overlapping time periods when comparing or combining different datasets.

3. Apply the temporal filter in Equation 3 to the hydrological fluxes.

Temporal filtering

Apply the three-point convolution kernel

$$h = \left[\frac{1}{5}, \frac{3}{5}, \frac{1}{5}\right]$$

separately to the monthly P, ET, and R time series.
The purpose of this operation is to make the temporal characteristics of the hydrological fluxes compatible with the storage-change estimate derived from GRACE.
Treat the beginning and end of the time series consistently, and make sure that the filtered time series retains the correct time stamps.

4. Evaluate the water-balance closure for each river basin.

Water-balance combinations

Because two independent datasets are available for each water-balance component, evaluate different combinations of precipitation, evapotranspiration, and runoff.
For a given combination, calculate

$$\epsilon(t) = P(t) - \mathrm{ET}(t) - R(t) - \frac{dS}{dt}(t).$$

Only do this for months that are available in **all** datasets involved in the respective combination.
The datasets have different temporal coverage. Therefore, determine the common overlapping period separately for the required comparison rather than assuming that all datasets have the same start and end dates.

The following points can be investigated:

- In which basins does the water balance close best?
- Which combinations of P, ET, and R agree best with GRACE?
- Which component appears to contribute most strongly to differences between the dataset combinations?
- Are discrepancies mainly seasonal, associated with individual extreme events, or persistent over longer periods?
- What observational and methodological uncertainties could explain the remaining water-balance residual?