|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Protocol**  Meeting Bachelor Thesis, FS 2024 | | | | | | | | | | | |
|  |  | | | | |  |  | | | | |
| **Protocol-No.:** | 08 | | | | | **Project name:** | Bachelor Thesis | | | | |
| **Meeting type:** | Discussion | | | | | **Location:** | GIUB, Bern | | | | |
| **Date / Time:** | 10.05.2024 / 10:00 | | | | |  |  | | | | |
| **Topic / Goals:** | Global Modelling | | | | | | | | | | |
| **Lead:** | Benjamin Stocker | | | | | **Logger:** | Patricia Gribi | | | | |
|  | | | | | | | | | | | |
| **Participants** | |  |  | | **E-mail** | | | **Present** | **Excused** | **Distribution** |
| Prof. Benjamin Stocker | |  | GECO-Group | | benjamin.stocker@unibe.ch | | | x |  |  |
| Patricia Gribi | |  | Unibe | | patricia.gribi@students.unibe.ch | | | x |  |  |
|  | | | | | | | | | | | |
| **Items discussed:** | | | | | | | | | | | |
| 1 global Modelling | | | | | | | | | | | |
| **Next meeting:** | | | | **Attachments:** | | | | | | | |
| * xx.05.22/15.00 | | | |  | | | | | | | |

| *(Legend for type: D = Decision, P = Pending, I = Information)* | Typ | Resp.: | Date: |
| --- | --- | --- | --- |
| Modelling |  |  |  |
| * **Reading information on NetCDF from linux shell:** ncdump -h file | I |  | 10.06 |
| * **Map2tidy:** The evspsbl data contained sometimes 2 records in one month. This was probably because the cmip6-ng netcdf files leap years are not considered.\* We adjusted the map2tidy function to calculate the years correctly.   + overwrite=True in map2tidy to overwrite the new files * **many-to-many relationship:** As mentioned above there are multiple rows in df\_evap containing 2 values per month or lacking a value per month. When joining with the other variables to one dataset there is a many-to-many relationship, resulting in double data in one day for example. What I did until now is:   + **A red numbers and black text      Description automatically generated with medium confidence**Aggregated double records by taking the mean   + The issue should be solved now with the implemented functionality of noleap | I |  | 10.06 |
| * **Next steps:**    + Test if the data shows plausible results based on a gridcell which is known and compared with other data. For example: De-Tha, comparison with the fluxnet data.   + **Annual maxima:** deficitgruppieren nach jahr, pro Jahr, pro gitterzelle, pro band das max zusammenfassen und diese dann zusammentragen (diese Funktion gibt es schon). Alles dann in ein NetCDF schreiben. Das Datum von maximum pro jahr sollte auch bekannt sein und ausgelesen werden. | I |  | 10.06 |
| Workflow |  |  |  |
| * **CWD outputs and map2tidy outputs here:** /data\_1/CMIP6 |  |  |  |
| * **Moved original data-download:** /data/scratch/CMIP6ng\_CESM2\_ssp585/cmip6-ng |  |  |  |
| * Rename the function cwd\_byilon to cwd\_byilon\_cmip6 | I |  | 10.06 |
| * **Github Issue:** There is an issue with my github. I cannot push to the cwd\_global repo. This problem needs to be solved. | P |  |  |
| Questions |  |  |  |

* How to adjust the paths to enable a reproducible workflow?
* Do I need to write in the methodology part what code adjustments I did?

\*Days in the NetCDF files without leap years, but we counted days with leap year, so the interpretation of the days was false (function helpers in map2tidy). If it was a leap year for example it would indicate the 31. December but the model counts as 1. January and we calculate it as 31. December. The model lies in advance.