Using supervised machine learning and global datasets (ERA5-Land, CMIP6) to predict climate indices from monthly data

Supervisor: Marianne Bügelmayer (AIT), Herbert Formayer (BOKU)
Beratungsteam:
Marianne Bügelmayer-Blaschek
Herbert Formayer
Harald Rieder (BOKU)

Kristofer Hasel, Marianne Bügelmayer-Blaschek, Herbert Formayer
kristofer.hasel@ait.ac.at, kristofer.hasel@boku.ac.at
AIT: Giefinggasse 6, 1210 Vienna, BOKU: Gregor-Mendel-Straße 33, 1180 Vienna

Background

- Climate change indices describe the **extremes in climate** while climate itself is mainly described by its mean.
- To estimate possible **future climate states,** climate change indices **are computed from computationally intensive climate models**, which demand considerable computational resources resulting in a high energy demand.
- Climate models are complex systems needing a great amount of background knowledge while **climate change indices** are able to **display the impact of a changing climate** in an **easy to understand** way.

Energy consumption **VSC-3** under full load (596.01 TFlop/s)[1]	450 kW
Average energy consumption of a **household** in Austria in 2016[2]	0.41 kW
Energy consumption of a usual **notebook**[3]	60 W

[1] Source: Top500 Liste - https://www.top500.org/
[2] Source: Strom- und Gastagebuch – Statistik Austria, www.statistik.at
[3] Source: My notebook

Aim

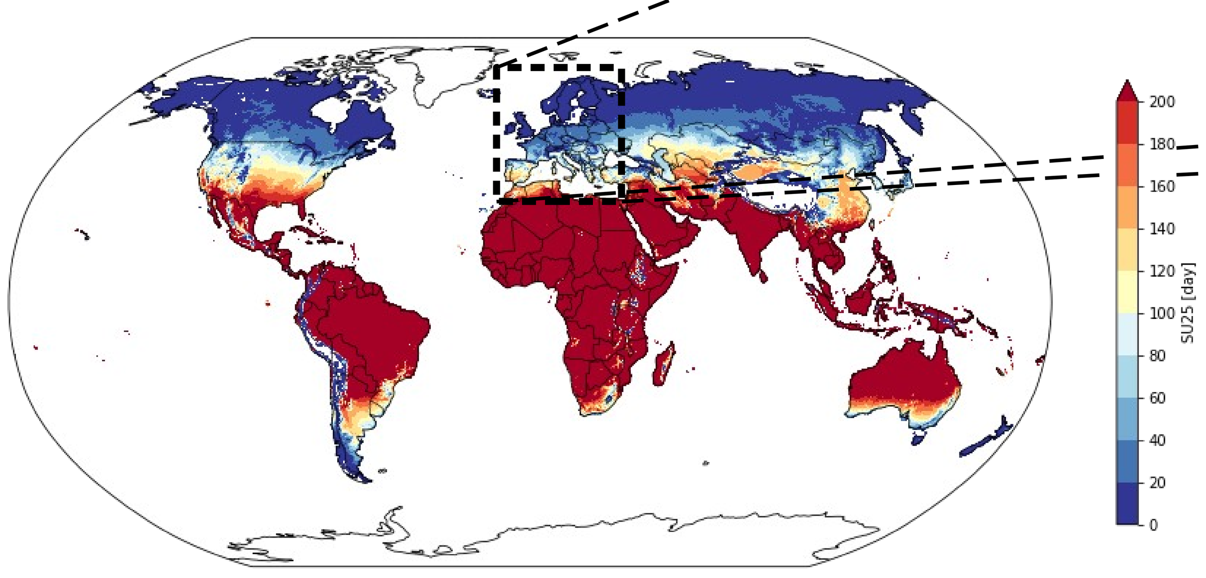
- Derive a **mathematical expression** connecting the peak over threshold climate indices with the monthly mean temperature or monthly precipitation sum.
- The mathematical expression should fullfill the following needs:
  - **Simplicity:** One should be able to use it with a calculator.
  - **Reproduction of climatologys:** For the comparison of different climate models, the index climatology should be reproduced well.
- The method therefore has the following goals:
  - **Prediction of climate change indices** based on monthly means and sums
  - **Easy and efficient comparison** of climate model ensembles

Data

- Era5-Land global reanalysis dataset with spatial resolution of 9 km and hourly data from 1981 – 2020
- SSP5-8.5 scenario from the CMIP6 global climate model MPI-M MPI-ESM1.2-50 HR with a spatial resolution of 100 km and daily data from 1981 – 2100

Methodology

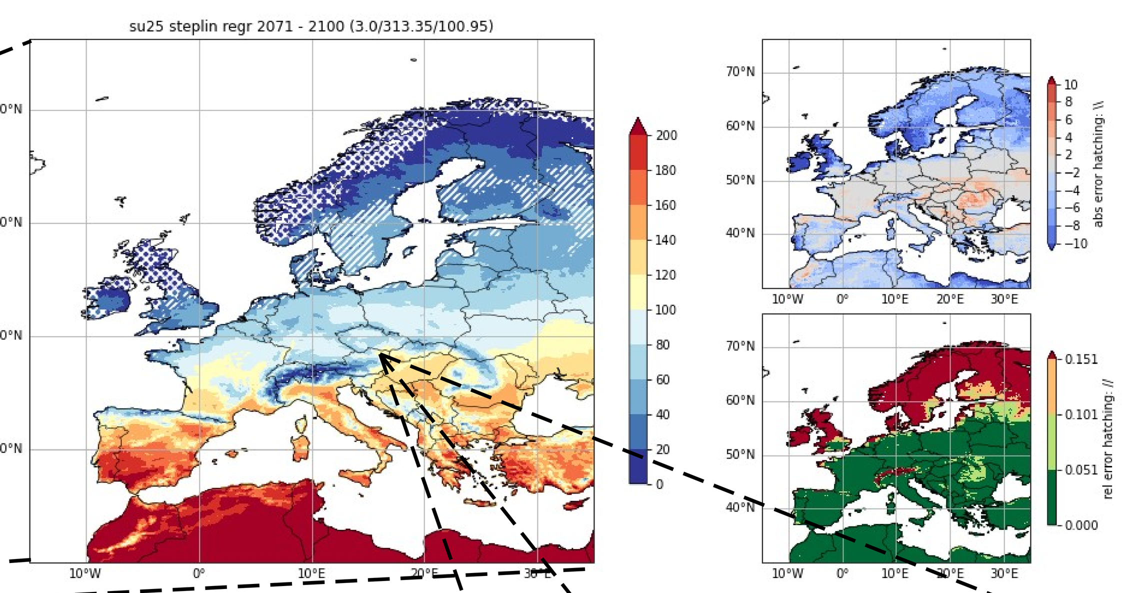
- **Calculate monthly peak over threshold indices** for the global Era5-Land dataset.
- **Add anomalys calculated from CMIP6** global climate model SSP5-8.5 pathway to cover the impacts of a possible future climate.
- **Produce regression model** for every gridpoint on the dataset for 3 different mathematical expressions:
  - Linear function: $f(x) = b_1 * x + b_2$
  - Stepwise linear function $f(x) = b_1 * x + b_2 * (x - z_i)_+ + b_3$
  - Exponential function $f(x) = b_1 * \exp(b_2 * x)$
- **Find the best fitting regression function** by R^2, RMSE, relative and absolute error in climatologys.



Climatology of summer days SU25 derived from Era5-Land for the period 1981-2020.

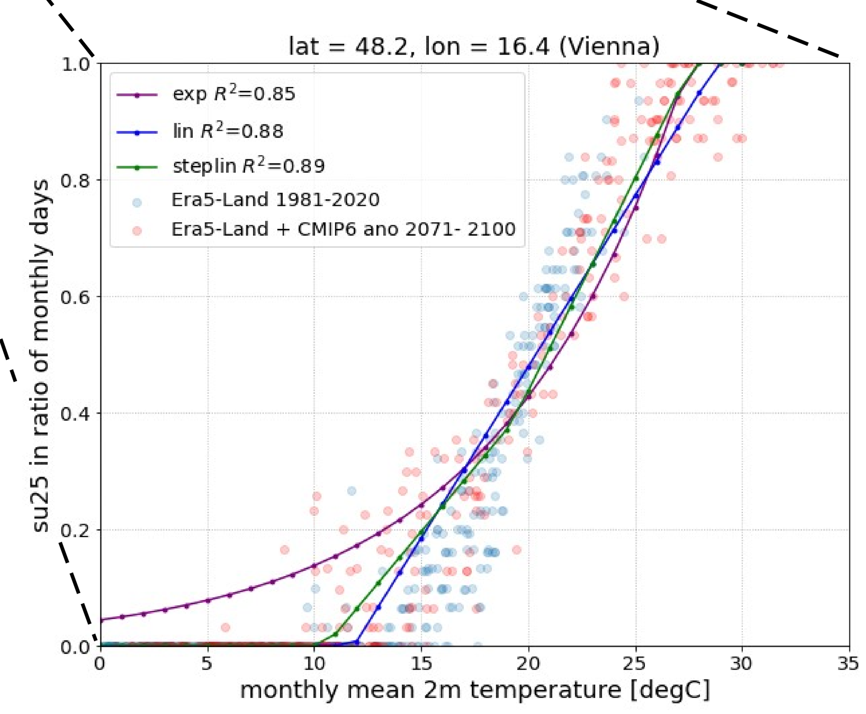
Preliminary results

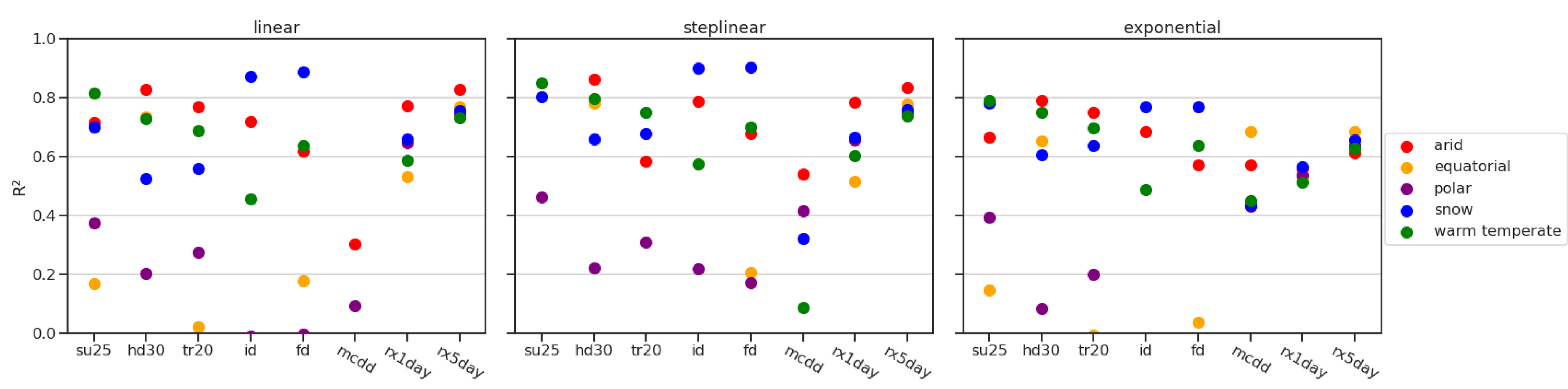
- From the three regression functions the **linear and stepwise linear function works best for temperature indices** like summer days, tropical nights, heat days, ice and frost days.
- The precipitation indices **rx1day, rx5day are reproduced well with linear regression function** while the **maximum consecutive dry days are best displayed by the exponential function**.
- The indices **climatology can be reproduced almost globally**. In areas where a index is allways present, like summer days in the tropics, the regression can not resolve the data's variance.
- Areas where the index occurs rarely, the regression has a big relative error, but usually a small absolute error.



Summer days SU25 climatology for the future ssp5-8.5 scenario from 2071 – 2100 produced by the regression model with stepwise linear function (left). Absolute error top right and relative error lower right. Areas with hatching are not performing well enough. Areas with a R^2 lower than 0.5 are hatched with 'x', Areas with an relative error greater than 0.15 are hatched with '\\' and areas with a climatology bellow 20 days and an absolute error greater than 3 are hatched with '//'.

The figure to the right presents the different regression functions for the climate index summer days SU25 in Vienna. The x axis displayes the monthly mean temperature, the y axis the index summer days SU25. The blue and orange points show the Era5-Land data and the Era5-Land data with added anomalies from the CMIP6 ssp5-8.5 model scenario. The purple, blue and green lines represent the different regression functions.





Evaluation of regression model metrics (R^2) based on the 5 basic Köppen-Geiger climate zones for the calculated climate indices.

Literature

- Formayer, H., I. Nadeem, and I. Anders, 2015: Climate Change Scenario: From Climate Model Ensemble to Local Indicators. Springer Climate, Springer, 55–74, https://doi.org/10.1007/978-3-319-12457-5_5
- Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. Journal of Geophysical Research Atmospheres, 111 (5), 1–22, https://doi.org/10.1029/2005JD006290.107
- Cornes, R. C., G. van der Schrier, E. J. van den Besselaar, and P. D. Jones, 2018: An ensemble version of the e-obs temperature and precipitation data sets. Journal of Geophysical Research: Atmospheres, 123 (17), 9391–9409.110
- Haylock, M., and N. Nicholls, 2000: Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910-1998. International Journal of Climatology, 20 (13), 1533–1541,https://doi.org/10.1002/1097-0088(20001115)20:13<1533::AID-JOC586>3.0.CO;2-J.
- James, G., D. Witten, T. Hastie, and R. Tibshirani, 2013: An introduction to statistical learning, Vol. 112. Springer.
- Jungclaus, J., and Coauthors, 2019: MPI-M MPI-ESM1.2-HR model output prepared for CMIP6 CMIP historical. Version 20190825. Earth System Grid Federation, URL https://doi.org/10.22033/ESGF/CMIP6.6594, https://doi.org/10.22033/ESGF/CMIP6.6594.
- Muñoz Sabater, J., 2021: Era5-land hourly data from 1981 to present. copernicus climate change service (c3s) climate data store (cds). https://doi.org/10.24381/cds.e2161bac, accessed on 06.05.2021.
- Schupfner, M., and Coauthors, 2019: DKRZ MPI-ESM1.2-HR model output prepared for CMIP6 ScenarioMIP ssp585. Version 20190721. Earth System Grid Federation, URL https://doi.org/10.22033/ESGF/CMIP6.4403, https://doi.org/10.22033/ESGF/CMIP6.4403.
- Sillmann, J., V. Kharin, X. Zhang, F. Zwiers, and D. Bronaugh, 2013: Climate extremes indices in the cmip5 multimodel ensemble: Part 1. model evaluation in the present climate. Journal of Geophysical Research: Atmospheres, 118 (4), 1716–1733.
- Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel, 2006: World Map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift, 15 (3), 259–263, https://doi.org/10.1127/0941-2948/2006/0130, URL http://dx.doi.org/10.1127/0941-2948/2006/0130.