climAPCREGION-01a: Regional Climate Variability (Exercises) (climAPC) (500018)

statistical methods code (python or matlab): (20% of final mark)

Hand in a code (function) that performs the assigned statistical analysis to an unknown time series data set. The unknown test dataset will be either one time sieries, two timeseries, or a 2D temporal changing field, depending on the assigned statistical method. The code should be written in Python or Matlab and only use standard libraries that do not require extra installations / licenses.

Publish the code on a public repository and hand in the link. (e.g. github, …)

All codes have to be public for all to use for the following analysis.

Codes are marked only if running and giving correct results (with a flawless test dataset). Marks will be given for good commentary and explanations within the code and error-robustness (with datasets with errors, gaps).

climAPCREGION-01a: Regional Climate Variability (Exercises) (climAPC) (500018)

time series analysis methods (Part I)

Climate indices and other time series data have characteristics that can be separated. Some climate indices are only one of those characteristics. These characteristic variations include seasonality, trend, outlier and rare events (possibly cyclically) as well as disruptions and step changes and random variations.

The time series can can be broken down into their variations by various methods and several climate indices are represented by only one or two of these variations

1) Seasonal variations (2) Trend variations (3) Cyclical variations, and (4) Random variations.

There are several statistically methods to analyze a time series / climate index. These include but are not limited to: EOF – Empirical Orthogonal Function and PCA/SVD - Principal Component Analysis / Single Value Decomposition which are closely related, Wavelet analysis, (V/S)ARIMA - Autoregressive integrated moving average, Fourier Transformation and model fitting by non-linear least squares fitting, optimal interpolation and Krigging.

climAPCREGION-01a: Regional Climate Variability (Exercises) (climAPC) (500018)

time series analysis methods (Part I)

The website:
https://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml
contains several climate indices, data, and explanations on how rotated EOF are used to derive these and lists references describing them.
These climate modes will also be used in the later part of this exercise in the following weeks.

Statistical methods for the analysis can be found in many text books and as code in many repositories. Feel free to copy and modify them according to your needs and code copyright regulations (since you will have to publish you code, please take copyright issues serious).
You can use the climate modes from the above website or any other environmental dataset as input, please reference accordingly and publish dataset with code as example. The code should be published on a global repository like “github”, feel free to use another one that is freely accessible and suitable. Since code is available and know – marking will be on detailed commentary in the code and error robustness in programming (code can deal with missing values, false values, non linear time vector, gaps, …) Your code may be interactive or analyse certain aspects. Some methods (the simpler ones) are suited for error analysis, this should be included and will be marked.

time series analysis methods (Part I)

Methods are randomly assigned. Since any method has several students, feel free to discuss the approaches. Do not hand in the same files. - choose a data set of your liking.

EOF / PCA – Empirical Orthogonal Function and.or Principal Component Analysis
take rotated EOF into account, make sure code includes explanation why the 'rotation' is helpful for physical parameters/connections. Use any dataset you see suitable, reference correctly, give links to dataset.

Wavelet
make sure the code can run with gaps (you have to interpolate, choose wise methods), check versus known noises (red noise/ white noise)

(**V/S)ARIMA** - Autoregressive integrated moving average
make sure the code can run with gaps (you have to interpolate, choose wise methods), find a good model and choose setting you expect to suit.

Fourier Transformation
simple method, perform thorough check versus known noises (red noise/ white noise), compute 95%/98% confidence intervals. MonteCarlo is a possible approach.

least squares fitting & optimal interpolation (optional Krigging)
find a good model to analyse time series of choice, find a way to estimate the error.

time series analysis methods (Part I)

This part of the exercise is about creating useful code (Matlab/Python) for the second part of the exercise, not about interpreting the results / climate signals you choose as input. This will be done later. It will be helpful if you code is well structured so that you and others from this group can use/modify it after submission. Following students have been randomly assigned to following methods:

EOF / PCA
C. Posern, P. Damke, M. Schulz, A. Vasudevan

Wavelet
F. Kirch, B. Friedrich, Q. Cunningham, C. McKellar

(**V/S)ARIMA**
T. Knoop, A. Andrae, N. Hocke, H. Pavanan

Fourier Transformation
P. Rethmeier, C. Ziska, C. Babu, T. Sieker

least squares fitting & optimal interpolation (optional Krigging)
L. Aroucha, M. Pohl, V. Volkova, R. Herbst