

## INDECIS WP3. QC & Homogenization Benchmark

INDECIS Benchmark will be based on KNMI's Climate Regional Model. The series will be sampled for the regions of South Sweden (SE) and Slovenia (SI)

### Important Dates:

- Deadline for Contributions: January, 31<sup>st</sup>, 2019
- Homogenization of the ECA&D dataset: June, 30<sup>th</sup>, 2019

### Benchmark Definition:

- The Clean Worlds will be created sampling from the model grid points
- The number of stations composing the network will be 100 for SE and 30 for SI
- The Corrupted Worlds will include an average of 1 breakpoint every 20 years.
- The size and shape of the breakpoints will be defined by generating differences/ratios between neighboring gridpoints and applying them to the required segment of the data
- To simulate networkwide trends, a network-wide trend is introduced (e.g. 0.4°C/100 years in temperature, -5%/100 years in precip and wind, -2%/100 years in Relative Humidity)
- INDECIS will apply the missing values mask produced for the MULTITEST project
- INDECIS will introduce a set of QC problems into the Corrupted World
- Data formats mimic ECA&D

### Benchmark Flavors

- In order to evaluate different aspects of homogenization and its main procedures (detection/correction), previous requirements (QC) and usual problems (missing values) the INDECIS' Benchmark (aka Baboon) will be presented in different flavors: (explain better)
  - o **Baboon:** Clean
  - o **Tall Baboon:** Clean + QC layer + Missing Values (MV) layer
  - o **Angry Baboon:** Clean + MV + Breakpoints (BP)
  - o **Ballistic Baboon:** Clean + MV + QC + BP
  - o **Weird Baboon:** Clean + BP
- Each flavor contains a dataset for SE and SI and for the ECA&D variables TX, TN, PP, RR, SD, FG, SS, RH, CC
- Contributors are not requested to work with each flavor of Baboon, and the priority is to submit results for Ballistic Baboon, i.e. the most complex benchmark. The rest of the flavors have also scientific value, and contributions will also be considered and summarized in a different set of statistics.
- Similarly, contributions should cover all variables, although analysis covering a subset of variables or a single variable will also be accepted and summarized.

- Finally, it is necessary that, for each returned variable, that all the stations in the benchmark have been assessed. Keep in mind that the final goal is to homogenize a large dataset (ECA&D) so methods which are not suitable to be applied to large datasets (or possibilities to evolve into it) have little use for INDECIS' purposes. If, for some reason, (e.g. lack of well correlated references or too many missing values) one or several stations cannot be adjusted, return as "adjusted" the original series. If the breakpoints have been detected, but no adjustment is possible, keep them in the list of detected breaks.

### Benchmarking Protocol

- INDECIS' Benchmark will function understanding that its users are bona-fide scientists and it will not be a blind benchmark, i.e., upon release, we will make public the Clean Worlds and the Corrupted Worlds and the break points and qc issues lists.
- We will release also evaluation software for self-evaluation.
- Contributions to the benchmarking exercise are expected to be done using open source code. We cannot accept contributions created with black-box-like software.
- The contributions will be uploaded into specific folders of the INDECIS' WP3 owncloud.
- Upon benchmark release, INDECIS will distribute instructions on formats for communicating contributions (e.g. formats, files, bp files, qc files, etc.).
- Once participants are ready to submit their contributions they should e-mail to [joanramon.coll@urv.cat](mailto:joanramon.coll@urv.cat) and they will receive a link to INDECIS' OwnCloud server and instructions to upload.

### FEEDING BABOON's results into the homogenization of the ECA&D Dataset

- The homogenization of ECA&D for INDECIS will be done with the method that provides the best statistical results and is technically applicable to such a large dataset and meeting INDECIS' deadlines.
- CLIMATOL has already been applied to homogenize the ECA&D dataset with apparently good results (although with no full benchmark based validation) and is also technically applicable to large datasets in time for INDECIS' purposes, so it is regarded as the default method for INDECIS homogenization unless other approaches outperform it.