Indecis Sectorial Climate Services

Integrated approach for the development across Europe of user oriented climate indicators for GFCS high-priority sectors: Agriculture, disaster risk reduction, energy, health, water and tourism

Work Package 3

Deliverable 3.2c

Recommended Homogenisation Techniques based on Benchmarking Results

Javier Sigró (C3-URV), Erik Engström (SMHI) and Mary Curley (MetÉireann)





This report arises from the Project INDECIS which is part of ERA4CS, an ERA-NET initiated by JPI Climate, and funded by FORMAS (SE), DLR (DE), BMWFW (AT), IFD (DK), MINECO (ES), ANR (FR), with co-funding by the European Union's Horizon 2020 research and innovation programme

1. PARALLEL SERIES DATA SET

Also the generation of a complete and extensive Benchmark, with the objective of testing different homogeneity methods on a daily scale with real stations, the southern region of Sweden (under latitude 62^o), coinciding with the Benchmark, was chosen to identify nearby stations that had common measurement periods.

From these pairs of "parallel" stations, composite series have been built combining both stations. In this way we obtain a composite series from a station 1, which on a given date we replace by station 2, within the common measurement period.

We obtain a series with one identified break that we have generated and we also have the series corresponding to station 1 unmodified. In this way we can test how the composite series is adjusted and what is the difference between series 1 ("real", unmodified) and the correction made in the composite series.

1.2 Selection of the series.

We have selected data series from ECA&D daily non blended data set for the South Sweden region for 8 variables: maximum temperature (Tmax), minimum temperature (Tmin), precipitation (rr), sea level pressure (pp), humidity (hu), sunshine duration (ss), wind speed (fg) and snow deep (sd).

Table 1. Criteria used in the selection of series for diferent variables to be candidates to the parallel series data set.

criteria/variable	Tmax	Tmin	FG	HU	SP	RR	SS	SD
overlap period	20	20	20	20	20	20	20	20
% missing data	15	15	15	15	10	15	15	
maximum distance between stations	0.5º	0.5º	1º	1º	1º	0.1º	1º	0.02º



1

others	50 m altitude difference	50 m altitude difference	orientation of terrain			100 m altitude dif.		
ECA&D stations	317	317	56	52	42	601	9	543
Composite series	94	94	16	16	16	31	3	34

Table 1 show the different criteria used to selection stations candidates to be integrated into the parallel data set for every variable. For all the variables, 20 years of common period is needed, to have enough years to make the composition between both stations, several years before and after the break. The number of missing data has also been limited, with maximums of 15% of the total (exception made of the variable snow deep, where many missing data correspond to periods in which the event does not occur).

The maximum distance between two series has been adapted according to the variable and the density of existing stations, with the most restrictive cases being precipitation and snow deep, because they are the variables that are most variable with distance and with local conditions.

Other aspects taken into account are the orientation in the case of wind and the difference in altitude between both stations for the temperature and precipitation variables.

Applying these criteria we have obtained the following pairs of stations: 94 pairs for maximum temperature and 94 for minimum temperature; 31 for precipitation;, 34 for Snow deep; 16 para wind speed, humidity and surface pressure; and 3 for sunshine. Spatial distribution can be observed in Figure 1.2.

2.2 Content of the database

The database is available on the INDECIS project (<u>http://www.indecis.eu/benchmarking.php</u>) website.

The data base "parallel series Sweden.rar" contains the parallel series for South Sweden. The files are:

- 8 txt files "coordenadas XX selected.txt"" with the codes and metadata of composite stations for every climate variable
- 8 txt files "report parallel XX.txt" with the information about the composition. Code 1 and code2 are the code stations of parallel stations (code 1 is the near to present station and code 2 the early station), the number of non-missing data in every section of composition, and the date of the breakpoint (the point where station 1 changes to station 2)
- Files with "p" at the beginning of the name are the composed series
- Files with "r" are the series only with the information of code 1 stations.
- Readme file





TEMPERATURE

61-60-

59-58-57-56



HUMIDITY

15 16 18

14

13

SUNSHINE



18 19

Indecis

Work Package 3 / Deliverable 3.2c

19

17 18

16

3



SURFACE PRESSURE

Figure 1.2. Spatial distribution of stations integrate in Parallel Data Base



Work Package 3 / Deliverable 3.2c