Indecis

Sectorial Climate Services

Integrated approach for the **d**evelopment across **E**urope of user-oriented **c**limate **i**ndicators for GFCS high-priority **s**ectors: Agriculture, disaster risk reduction, energy, health, water and tourism

Work Package 3
Deliverable 3.1.b

INDECIS Quality Control Software Suite for tall tower wind data (INDECIS-QCSS4TT)

J. Ramon[1], Ll. Lledó[1]

[1] Barcelona Supercomputing Center (BSC), Barcelona, Spain

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.

Table of contents

1. DESCRIPTION OF THE QCSS4TT

In order to ensure a minimum quality of tall tower wind data, a set of 18 sequential Quality Control (QC) tests have been created to be applied over tall tower wind speed and wind direction measurements. Two preliminary QC checks have been recommended to be performed before running the INDECIS Quality Control Software Suite for Tall Tower wind data (hereafter QCSS4TT). The QCSS4TT contains 16 QC functions have been coded using R language and can be accessed through this Git repository:

https://earth.bsc.es/gitlab/jramon/INDECIS-QCSS4TT/blob/master/R/qc_tests.R

Each of these QC routines flags each observation according to their level of confidence. Hence, every single measurement will have their associated flags so that any record is removed, modified or set to NA if it is deemed untrustworthy and the data user can impose their level of restriction. Some of these flags have been defined using threshold values based on World Meteorological Organization (WMO) guidances (WMO, 2007), QC software (Brower et al., 2012; IEC 2005, QARTOD 2017) manuals and scientific articles (Jiménez et al., 2010). Complete information of this software, as well as a guided example on how to use it, can be publicly accessed here:

https://earth.bsc.es/gitlab/jramon/INDECIS-QCSS4TT

2. THE BENCHMARK EXPERIMENT

A benchmark exercise has been designed to evaluate the performance of the QCSS4TT. The following sections present the data and methodology very briefly (Section 2.1), some results (Section 2.2) and the conclusions derived from the experiment (Section 2.3).

2.1 Data and methodology

Climate data series sampled from the KNMI's Regional Climate Model for Slovenia (SI) and South Sweden (SE) are used to create the INDECIS benchmark (also referred to as 'Baboon'). Different flavours for the Baboon have been created to evaluate different aspects of the homogenization and Quality Control (QC). We refer to D3.2 to find complete information on the benchmark creation.

In particular, the Tall Baboon flavour has been created by modifying the QC-free series (Clean World) by introducing a set of QC issues plus a layer of missing values (Corrupted World). The error seeding process has been carried out by an independent group of people from the creators of the QCSS4TT, thus being a blind benchmark exercise. However, by taking the differences between the Clean and Corrupted Worlds, we can detect the specific records that have been modified. We test the performance QCSS4TT by running the software over the daily wind speed series of the Corrupted World to check whether the QC tests can detect the QC problems. Then, we look at the amount of data that was corrupted and considered wrong by the QCSS4TT.

2.2 Results

After running the QCSS4TT, four types QC issues have been encountered in the time series (see Figures 1, 2, 3 and 4):

- **Records outside an allowable range of values**

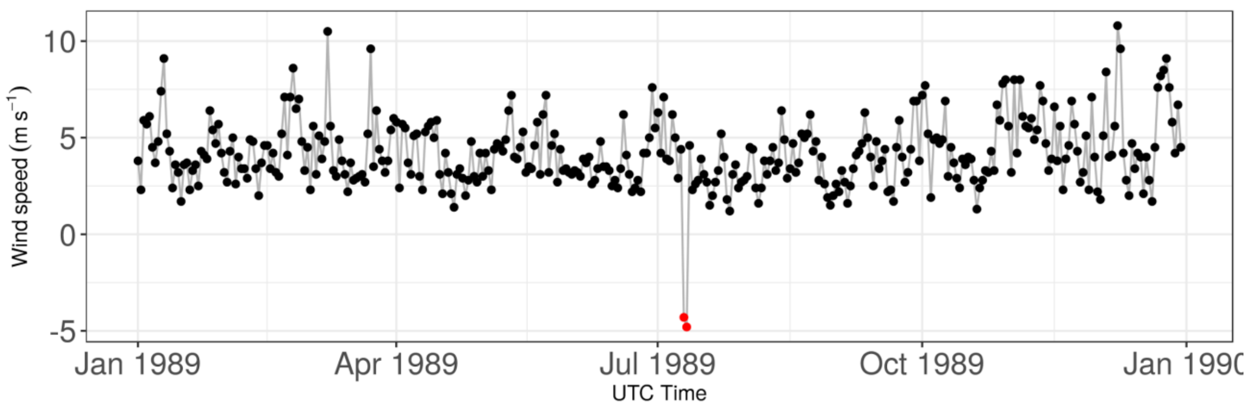

Figure 1.- Detection of negative values (red) in a series in South Sweden.

- **Duplicated sequences within the same series:**

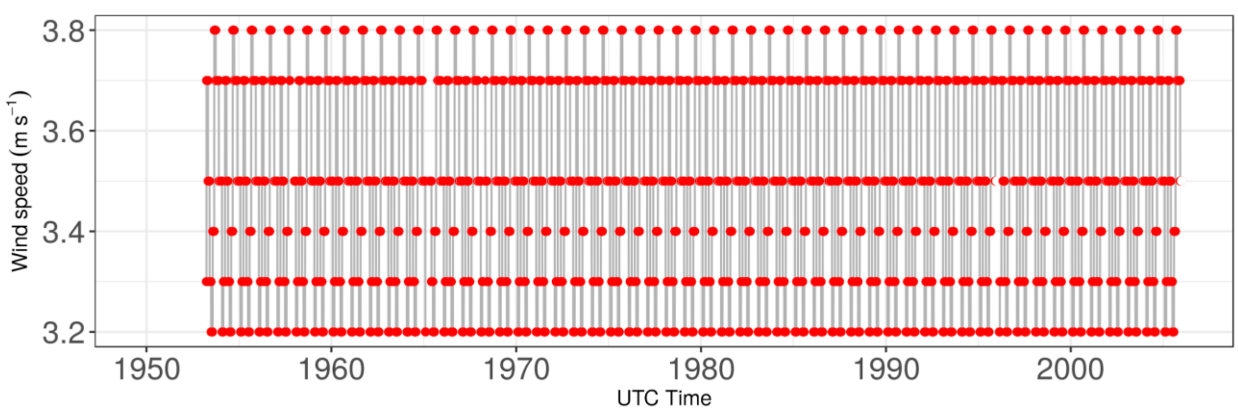

Figure 2.- Detection of a repeated sequence of the same values in a series in South Sweden.

- **Presence of spikes:**

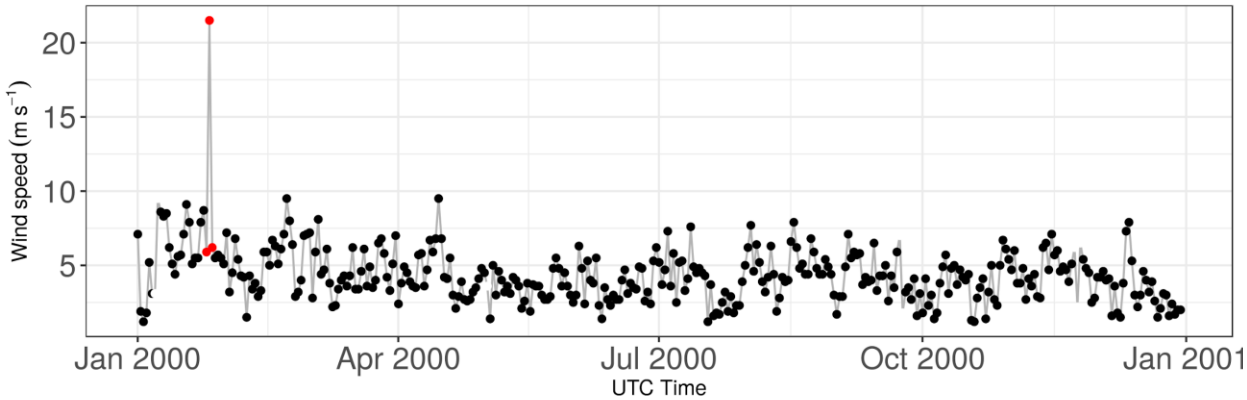

Figure 3.- Detection of a spike (red) in a series in South Sweden.

- **Flat lines:**

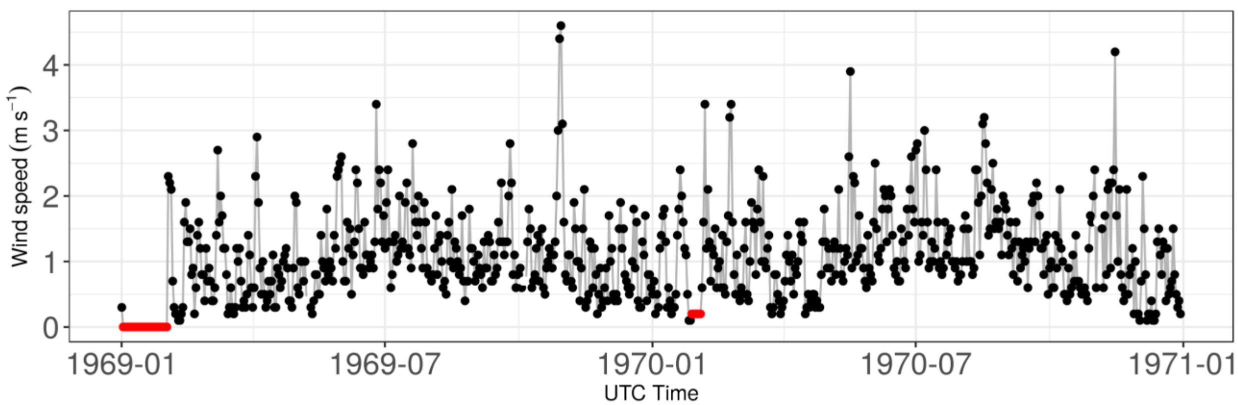

Figure 4.- Detection of flat lines (red) in a series in Slovenia.

3. CONCLUSIONS

The QCSS4TT has been applied over daily series sampled from an RCM, which have been previously corrupted by introducing a set of QC problems. Even though the QCSS4TT is specially created to clean sub-daily data, its application over the Corrupted World offered insights on the performance as well as helped to improve the original code.

The QCSS4TT is able to detect basic QC problems such as out-of-plausible range values, big spikes, and flat or duplicated sequences. However, some errors introduced into the Clean World have not been detected. Indeed, most of these skipped errors were seeded at random in the time series targeting sparse individual observations. In addition, the magnitude of those errors was not enough to produce detectable QC problems. Besides, the thresholds imposed in the QC tests have been set sufficiently permissible in order to not deem good data as wrong.

The benchmark exercise has also been used as a unit testing to detect errors in the QC code. The cost of finding a bug at the early stages of the code production is substantially lower than the cost of detecting, understanding, and correcting the bug later. In this particular experiment, no technical issues have been found so that the code can be useful provided that the input data is formatted correctly.

4. REFERENCES

Brower, M. C., B. H. Bailey, J. Doane, and M. J. Eberhard, 2012: Wind Resource Assessment: A Practical Guide to Developing a Wind Project.

IEC, 2005: International Standard - Wind Turbines. 2005.

Jiménez, P. A., J. F. González-Rouco, J. Navarro, J. P. Montávez, and E. García-Bustamante, 2010: Quality assurance of surface wind observations from automated weather stations. J. Atmos. Ocean. Technol., 27, 1101–1122, https://doi.org/10.1175/2010JTECHA1404.1.

QARTOD, 2017: Manual for Real-Time Quality Control of Wind Data.

WMO, 2007: Guide to the Global Observing System. Third edition. http://www.wmo.int/pages/prog/www/OSY/Manual/488_Guide_2007.pdf. Last accessed: 30th August 2019