Environment Canada

Environment Canada

Overview and updates of the RHtestsV5 data homogenization software package

Xiaolan L. Wang and Y. Feng

(Contact: Xiaolan.Wang@ec.gc.ca or Yang.Feng@ec.gc.ca)

Climate Research Division, Science & Technology Branch, Environment Canada

❖What tests can be done with the RHtests package?

The different versions of RHtests (currently RHtestsV5) is a software package for homogenization of climate data that can be made to approximate a normal/Gaussian distribution, such as annual surface air temperature (SAT), de-seasonalized SAT, atmospheric pressure, etc..

This software allows users to perform four statistical tests:

(1) the Penalized Maximum t (PMT) test for detection of unknown changepoints (Wang et al. 2007) using a reference series (FindU.wRef);

(2) the Student t test for determining the statistical significance of known changepoints using a reference series (FindUD.wRef or StepSize.wRef);

(3) the Penalized Maximum F (PMF) test for detection of unknown changepoints (Wang 2008b) without using a reference series (FindU);

(4) the regular F test for determining the statistical significance of known changepoints without using a reference series (FindUD or StepSize).

It also allows users to convert the daily data series in the RClimDex standard format to the monthly mean series in the RHtests format.

❖The unique features of the RHtests packages include:

(1) This and the RHtests_dlyPrcp package (presented earlier this afternoon) are the only existing data homogenization software that allows users to test both known and unknown changepoints.

(2) The RHtestsV5 allows users to make Quantile-Matching (QM) adjustments (Wang et al. 2010, Vincent et al. 2012) to daily or subdaily (up to hourly) data series for the changepoints already identified in the corresponding annual or monthly data series.

(3) The lag-1 autocorrelation in the data series being tested is accounted for, which greatly minimizes the false alarm rate (Wang 2008a);

(4) The annual cycle and lag-1 autocorrelation (and linear trend of the base series when no reference is used) are modelled in tandem while accounting for all identified shifts (Wang 2008a);

(5) Both the mean-adjusted and QM-adjusted data series, along with plots of the series and the resulting regression fit are provided in the output.

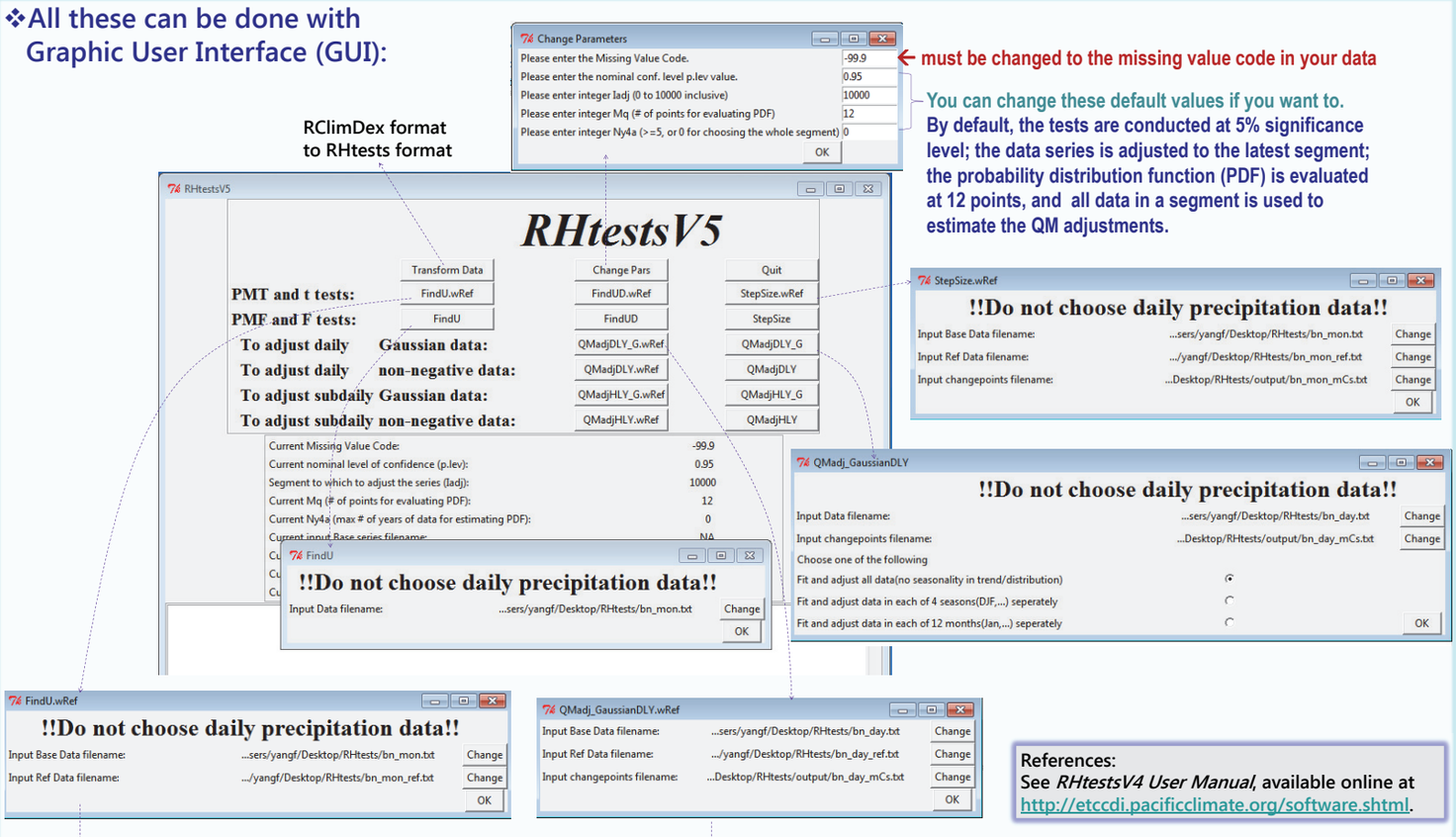
(6) users can also

(i) choose the segment to which the base series is to be adjusted;

(ii) choose to use the whole or part of the segments before and after a shift to estimate the QM-adjustments;

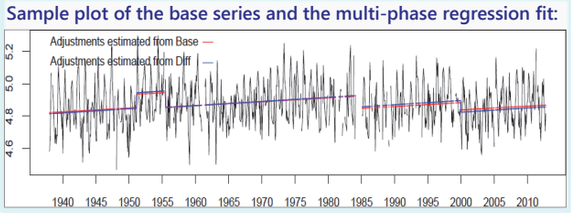
(iii) choose the level of significance at which to conduct the tests

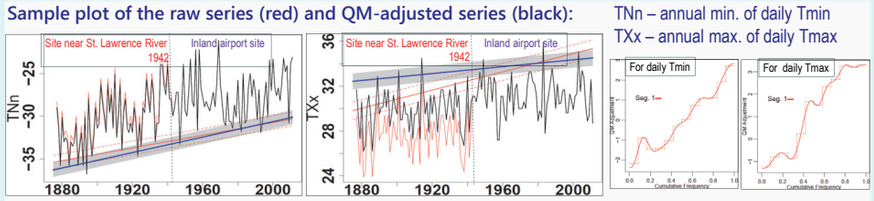
❖All these can be done with Graphic User Interface (GUI):



References:
See *RHtestsV4 User Manual*, available online at http://etccdi.pacificclimate.org/software.shtml.

Sample plot of the base series and the multi-phase regression fit:



Sample plot of the raw series (red) and QM-adjusted series (black):

TNn – annual min. of daily Tmin
TXx – annual max. of daily Tmax