

Climatol 2: Interactive and automatic R functions for homogenisation of climatological series

José A. Guijarro (jaguijarro@inm.es)

Meteorological State Agency, Balearic Islands Office (Spain)

Sixth Seminar for Homogenization and Quality Control in Climatological Databases (Budapest, 26-30/May-2008)

Homogenisation...

- ▶ Is indispensable when dealing with climatic series
- ▶ Is a very time consuming process
- ▶ And must be repeated as series become longer
- ▶ ⇒ Some level of automation is required

Homogenisation...

- ▶ Is indispensable when dealing with climatic series
- ▶ Is a very time consuming process
- ▶ And must be repeated as series become longer
- ▶ ⇒ Some level of automation is required

Homogenisation...

- ▶ Is indispensable when dealing with climatic series
- ▶ Is a very time consuming process
- ▶ And must be repeated as series become longer
- ▶ ⇒ Some level of automation is required

Homogenisation...

- ▶ Is indispensable when dealing with climatic series
- ▶ Is a very time consuming process
- ▶ And must be repeated as series become longer
- ▶ \Rightarrow Some level of automation is required

Climatol 1.x has a function that allows:

- ▶ Compute synthetic reference series by interpolation of standardized data from neighboring stations
- ▶ Analyse the differences observed-computed in search of outliers, shifts in the mean and trends
- ▶ Automatic missing data filling and outlier correction

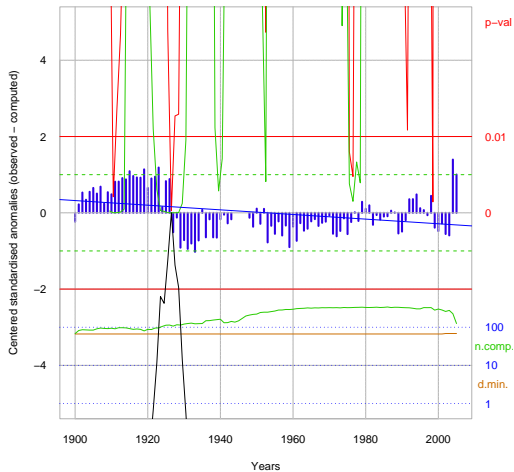
Climatol 1.x has a function that allows:

- ▶ Compute synthetic reference series by interpolation of standardized data from neighboring stations
- ▶ Analyse the differences observed-computed in search of outliers, shifts in the mean and trends
- ▶ Automatic missing data filling and outlier correction

Climatol 1.x has a function that allows:

- ▶ Compute synthetic reference series by interpolation of standardized data from neighboring stations
- ▶ Analyse the differences observed-computed in search of outliers, shifts in the mean and trends
- ▶ Automatic missing data filling and outlier correction

44: 48-MUENCHEN (s=0.78) AvTemp



Diagnostic graphic

Climatol 2.0 is providing a new function that iteratively:

- ▶ Computes reference series, fills missing data and correct outliers
- ▶ Applies several tests of shifts in the average, giving points when significance is achieved
- ▶ When maximum shift score is over a prescribed threshold, the series is split and the child joins the dataset as an independent series
- ▶ The process is repeated with different threshold levels, allowing a progressive correction from the greater inhomogeneities to the lesser ones

Climatol 2.0 is providing a new function that iteratively:

- ▶ Computes reference series, fills missing data and correct outliers
- ▶ Applies several tests of shifts in the average, giving points when significance is achieved
- ▶ When maximum shift score is over a prescribed threshold, the series is split and the child joins the dataset as an independent series
- ▶ The process is repeated with different threshold levels, allowing a progressive correction from the greater inhomogeneities to the lesser ones

Climatol 2.0 is providing a new function that iteratively:

- ▶ Computes reference series, fills missing data and correct outliers
- ▶ Applies several tests of shifts in the average, giving points when significance is achieved
- ▶ When maximum shift score is over a prescribed threshold, the series is split and the child joins the dataset as an independent series
- ▶ The process is repeated with different threshold levels, allowing a progressive correction from the greater inhomogeneities to the lesser ones

Climatol 2.0 is providing a new function that iteratively:

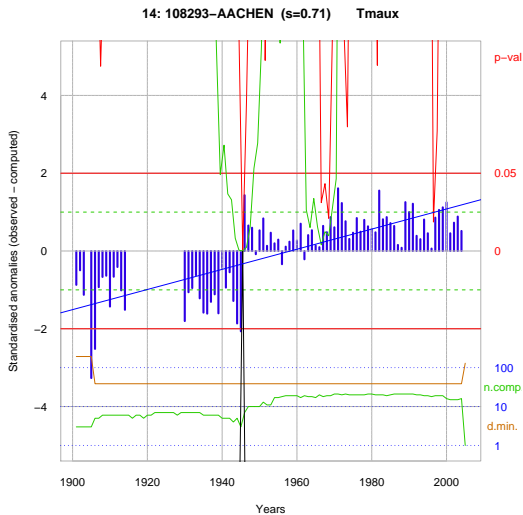
- ▶ Computes reference series, fills missing data and correct outliers
- ▶ Applies several tests of shifts in the average, giving points when significance is achieved
- ▶ When maximum shift score is over a prescribed threshold, the series is split and the child joins the dataset as an independent series
- ▶ The process is repeated with different threshold levels, allowing a progressive correction from the greater inhomogeneities to the lesser ones

- ▶ The automatic procedure seems to work well with annual data.
- ▶ But: Cannot be applied to seasonal or monthly data the same way, because the change points must be time consistent!
- ▶ Strategy tested so far: Accumulate monthly (or seasonal) shift scores over a whole year.
- ▶ Problem: Signal to noise ratio decreases dramatically when stepping from annual to seasonal and monthly data. (Do not dare to try with daily data!).

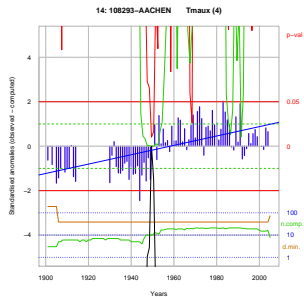
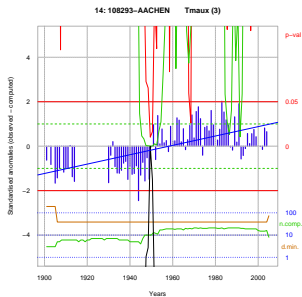
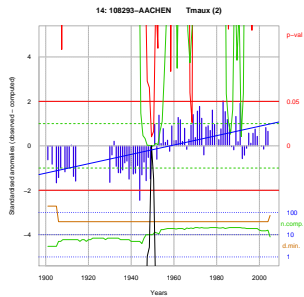
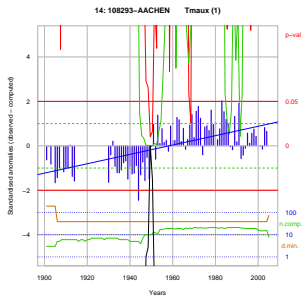
- ▶ The automatic procedure seems to work well with annual data.
- ▶ But: Cannot be applied to seasonal or monthly data the same way, because the change points must be time consistent!
- ▶ Strategy tested so far: Accumulate monthly (or seasonal) shift scores over a whole year.
- ▶ Problem: Signal to noise ratio decreases dramatically when stepping from annual to seasonal and monthly data. (Do not dare to try with daily data!).

- ▶ The automatic procedure seems to work well with annual data.
- ▶ But: Cannot be applied to seasonal or monthly data the same way, because the change points must be time consistent!
- ▶ Strategy tested so far: Accumulate monthly (or seasonal) shift scores over a whole year.
- ▶ Problem: Signal to noise ratio decreases dramatically when stepping from annual to seasonal and monthly data. (Do not dare to try with daily data!).

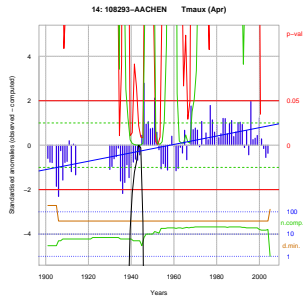
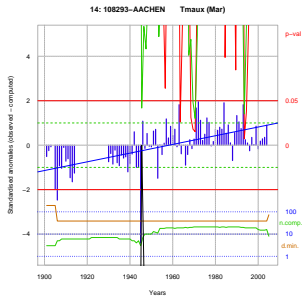
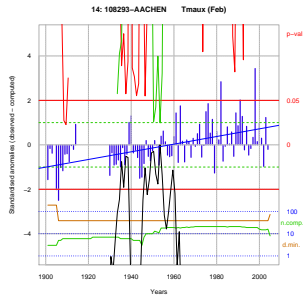
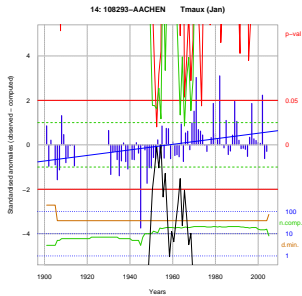
- ▶ The automatic procedure seems to work well with annual data.
- ▶ But: Cannot be applied to seasonal or monthly data the same way, because the change points must be time consistent!
- ▶ Strategy tested so far: Accumulate monthly (or seasonal) shift scores over a whole year.
- ▶ Problem: Signal to noise ratio decreases dramatically when stepping from annual to seasonal and monthly data. (Do not dare to try with daily data!).



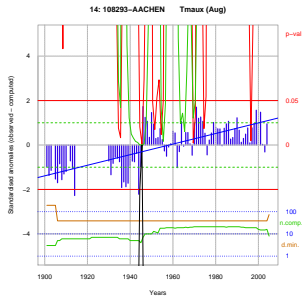
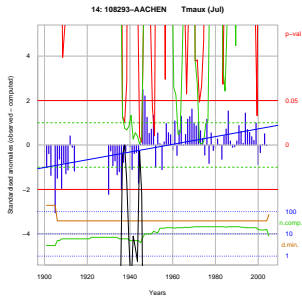
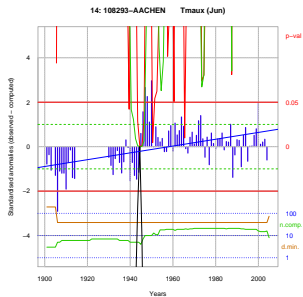
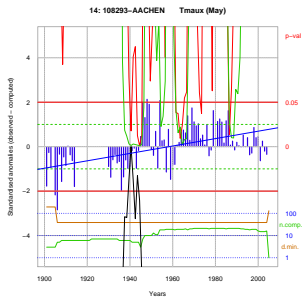
Annual mean temperature anomalies at Aachen



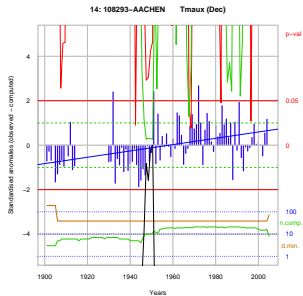
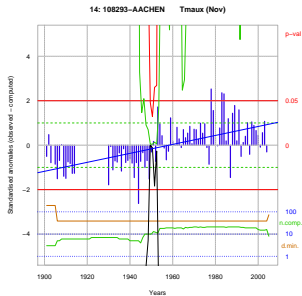
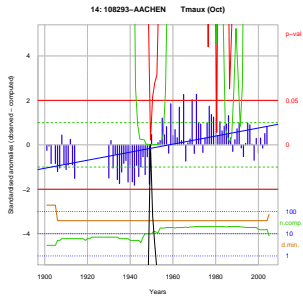
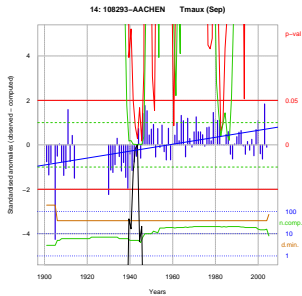
Quarterly mean temperature anomalies at Aachen



January to April mean temperature anomalies at Aachen

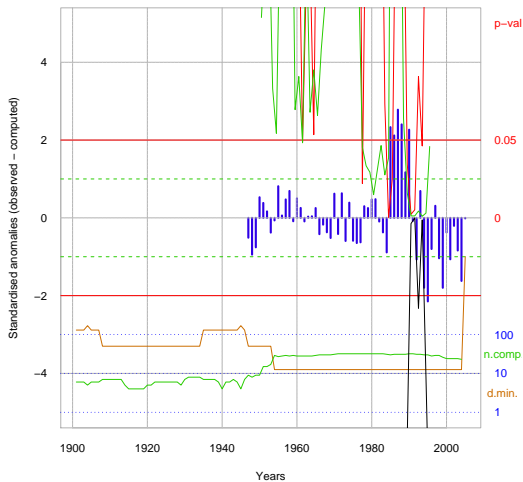


May to August mean temperature anomalies at Aachen

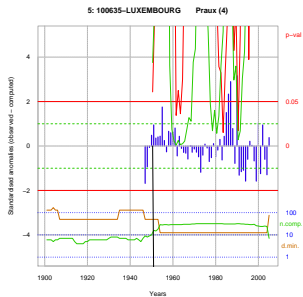
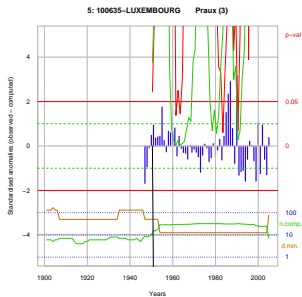
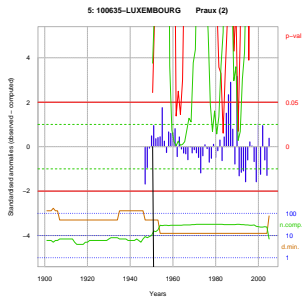
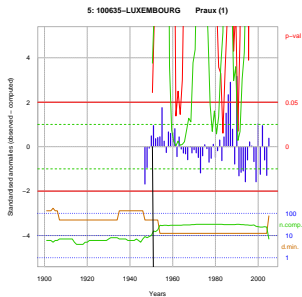


Sep. to December mean temperature anomalies at Aachen

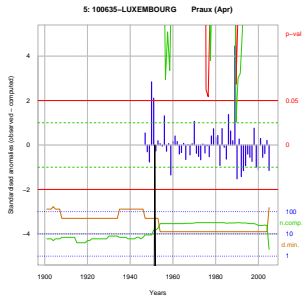
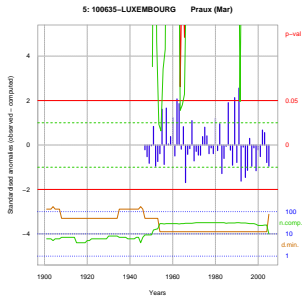
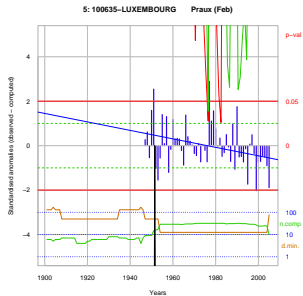
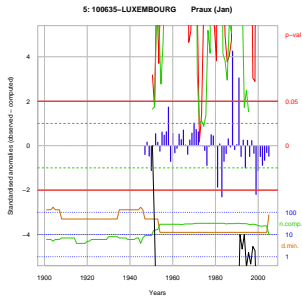
5: 100635-LUXEMBOURG Praux



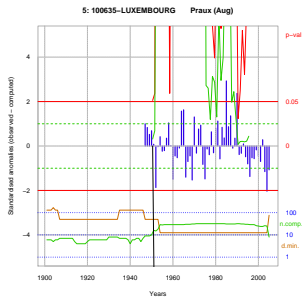
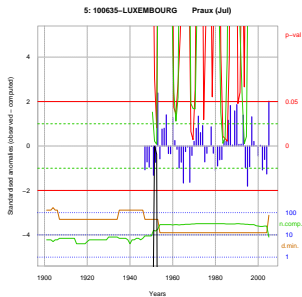
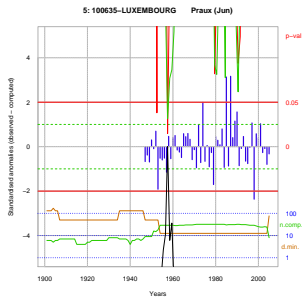
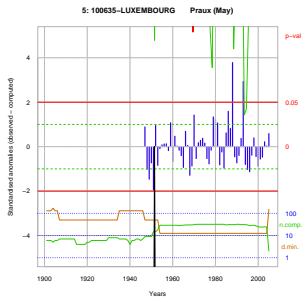
Annual precipitation anomalies at Luxembourg



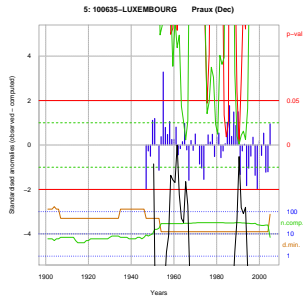
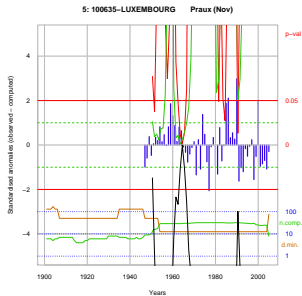
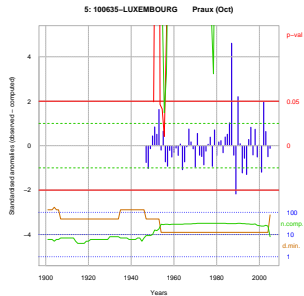
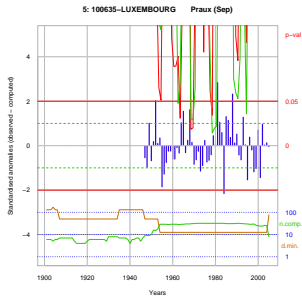
Quarterly precipitation anomalies at Luxembourg



January to April precipitation anomalies at Luxembourg

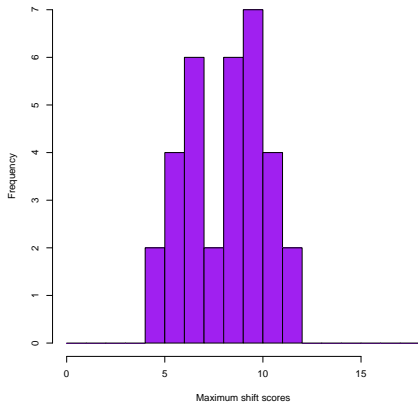


May to August precipitation anomalies at Luxembourg

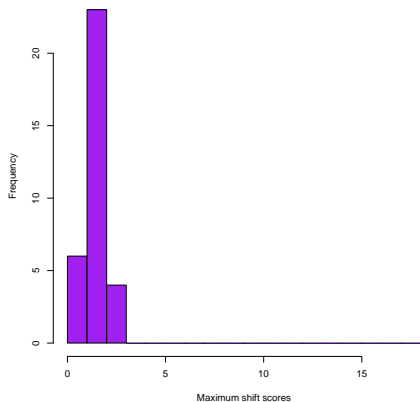


Sep. to December precipitation anomalies at Luxembourg

Station's maximum shift scores (Stage 1)



Station's maximum shift scores (Stage 1)



Maximum shift scores of annual (left) and monthly (right) precipitation

- ▶ Next task: Change the strategy to detecting shift points in annual data, and then looking for the monthly (or seasonal) shift point only in that year.
- ▶ Climatol 2.0 release is delayed, but anybody willing to test it will be provided with the beta version. (Just ask to jaguijarro@inm.es)

- ▶ Next task: Change the strategy to detecting shift points in annual data, and then looking for the monthly (or seasonal) shift point only in that year.
- ▶ Climatol 2.0 release is delayed, but anybody willing to test it will be provided with the beta version. (Just ask to jaguijarro@inm.es)