The ACMANT2 Software Package

Peter Domonkos
Centre for Climate Change
University Rovira i Virgili, Campus Terres de l’Ebre, Tortosa, Spain

ACMANT2

- Six programs for the automatic homogenization of air surface temperature and precipitation time series
-
- a) Tmean or Tmax from mid- or high latitudes – **changes in the algorithm relative to ACMANT1**
- b) Other temperature series – **new**
- c) Precipitation total – **new**

- For all of a) – c) monthly homogenization and downscaling from monthly to daily when daily data is available – **new**

Homogenization methods in recent climatic studies

- In climatic studies of the first half of 2013:

8% used HOME-recommended methods,

3% used multiple break methods (HOMER, MASH, PRODIGE, ACMANT)

(Domonkos and Efthymiadis, 2013, EMS2013)

Most recent tendencies:

- **HOMER** is more frequently applied
- ACMANT seems to be unknown

Correct reference to ACMANT

- Domonkos, P. 2011: Adapted Caussinus-Mestre Algorithm for Networks of Temperature series (ACMANT). International Journal of Geoscience, 2, 293-309, doi: 10.4236/ijg.2011.23032.

Why about ACMANT?

- ACMANT is a multiple break method. It was one of the most successful methods in the blind tests of HOME.
- The newer versions have even higher efficiency in large test datasets
- It is a fully automatic, easy-to-use method
- Freely accessible from web (http://www.c3.urv.cat/data.html)

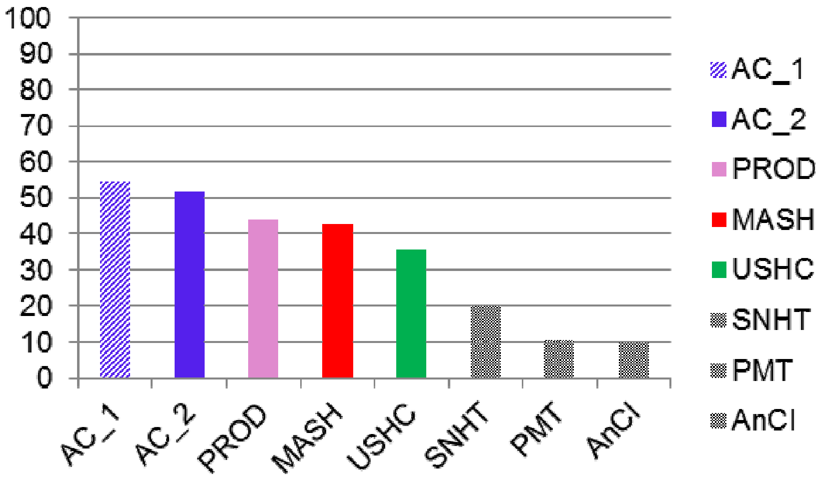
==========================

- Sponsor: FP7 project of Uncertainties in Ensembles of Regional ReAnalyses (UERRA)

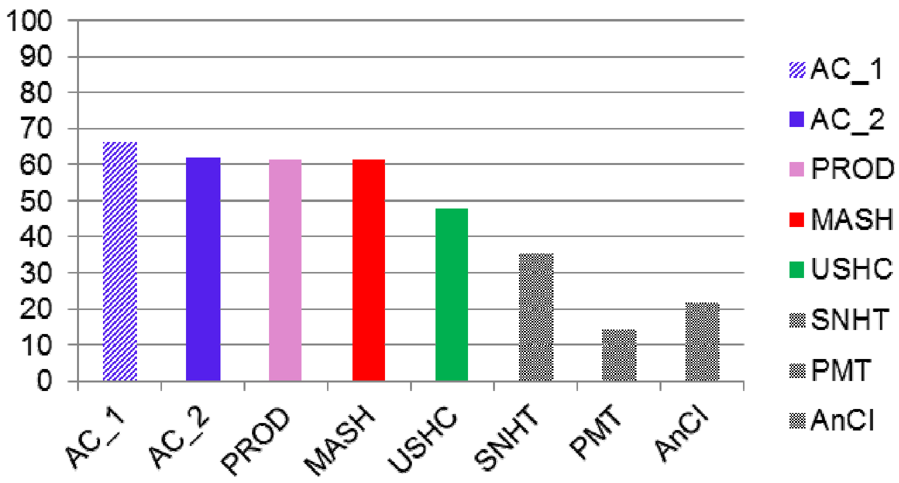
Efficiencies (%) with the HOME benchmark, in reducing RMSE and trend biases

(AC_1 = ACMANTv1, AC_2 = ACMANTv2, PROD = PRODIGE)

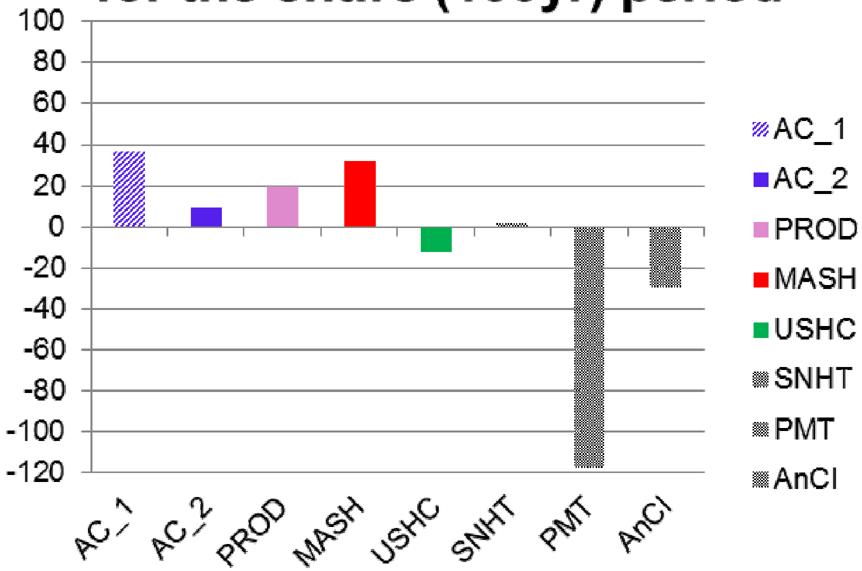
Monthly RMSE



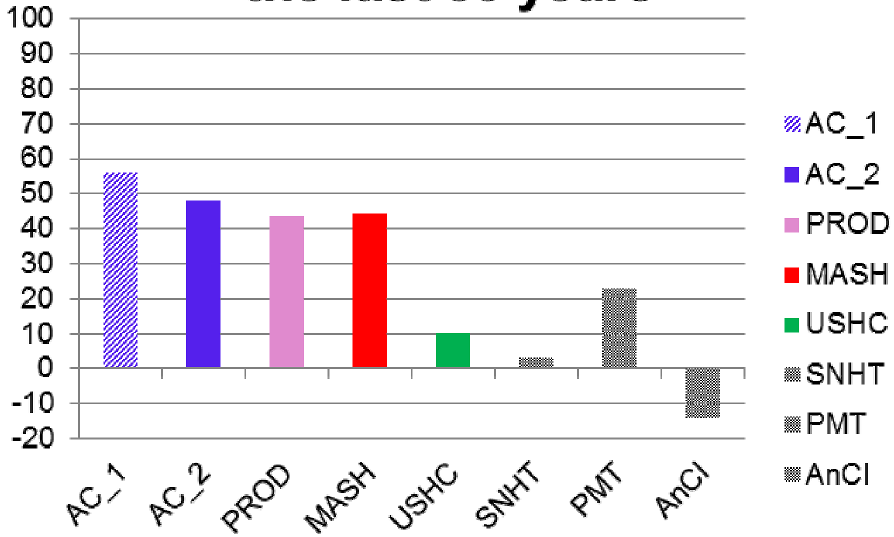
Annual RMSE



Network-mean trend bias for the entire (100yr) period

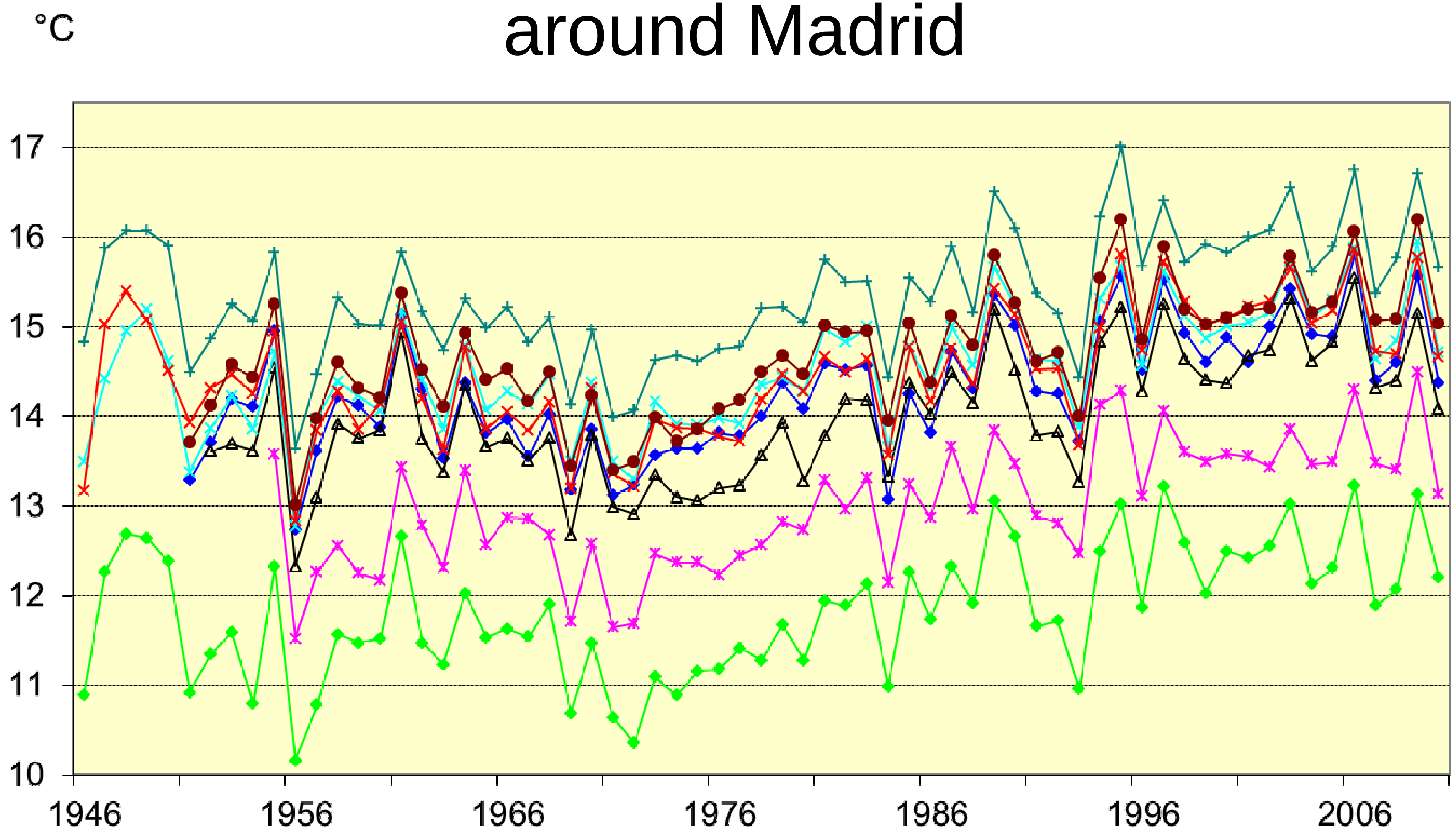


Network-mean trend bias in the last 50 years



New tests are needed

Homogenized observed temperatures in and around Madrid



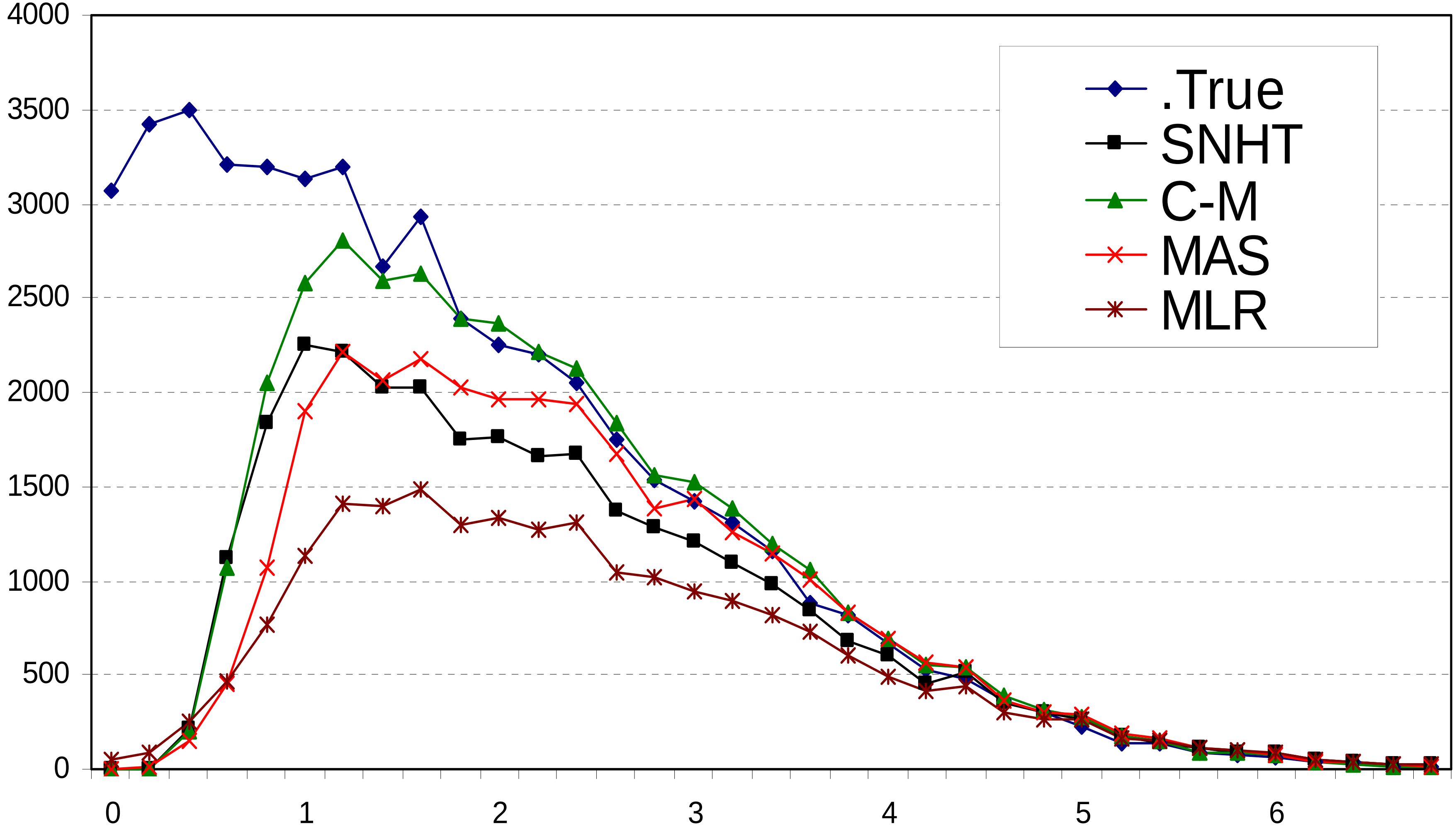
New tests are needed

Benchmarking problems

- Homogenizing global datasets is a combined task of regionalization and homogenization
- In observational temperature datasets short-term biases are frequent
- The occurrence of large fraction of missing data (of continuous sections or fragmented) is frequent

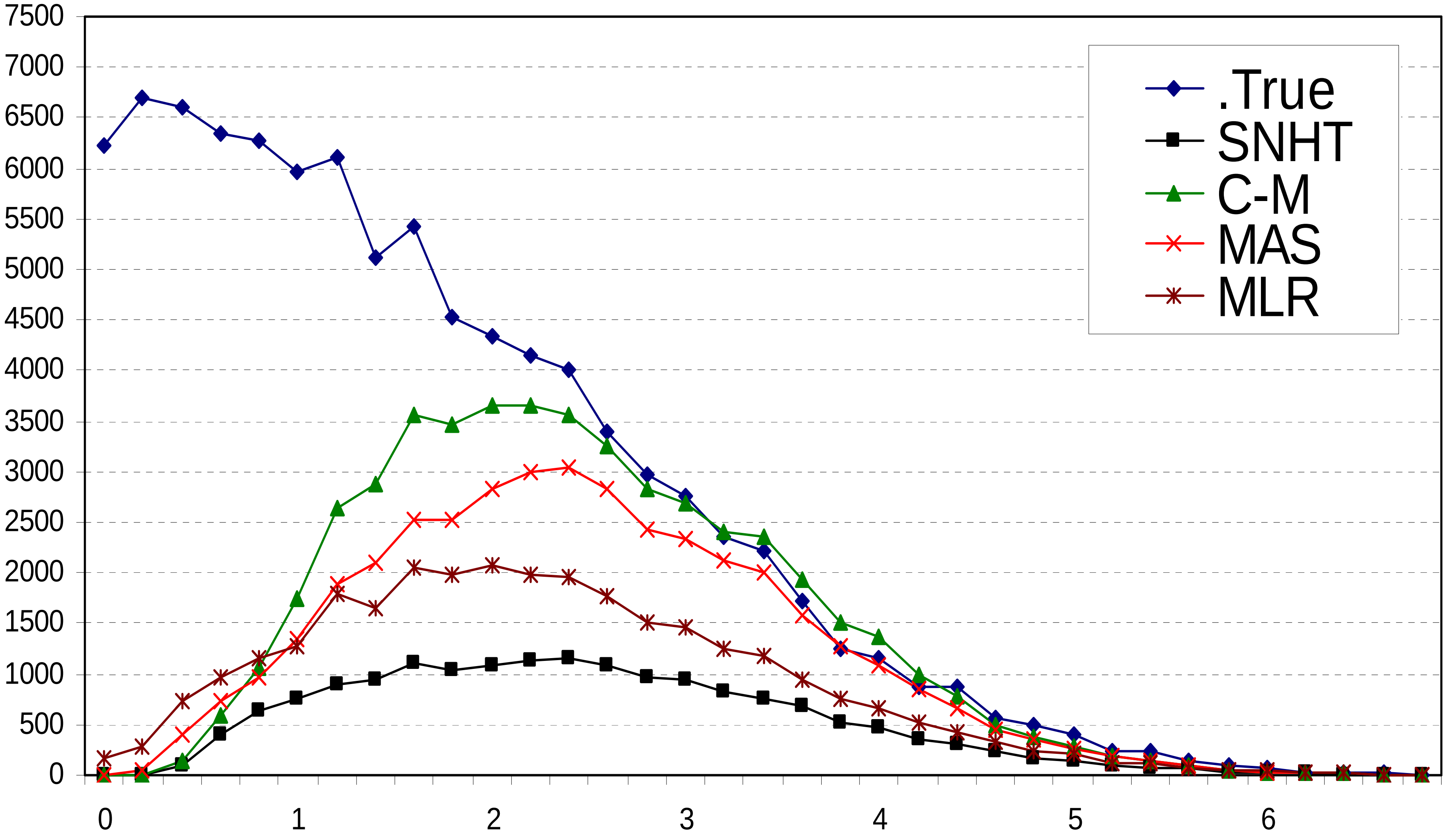
True and detected frequencies of breaks

Randomly positioned breaks



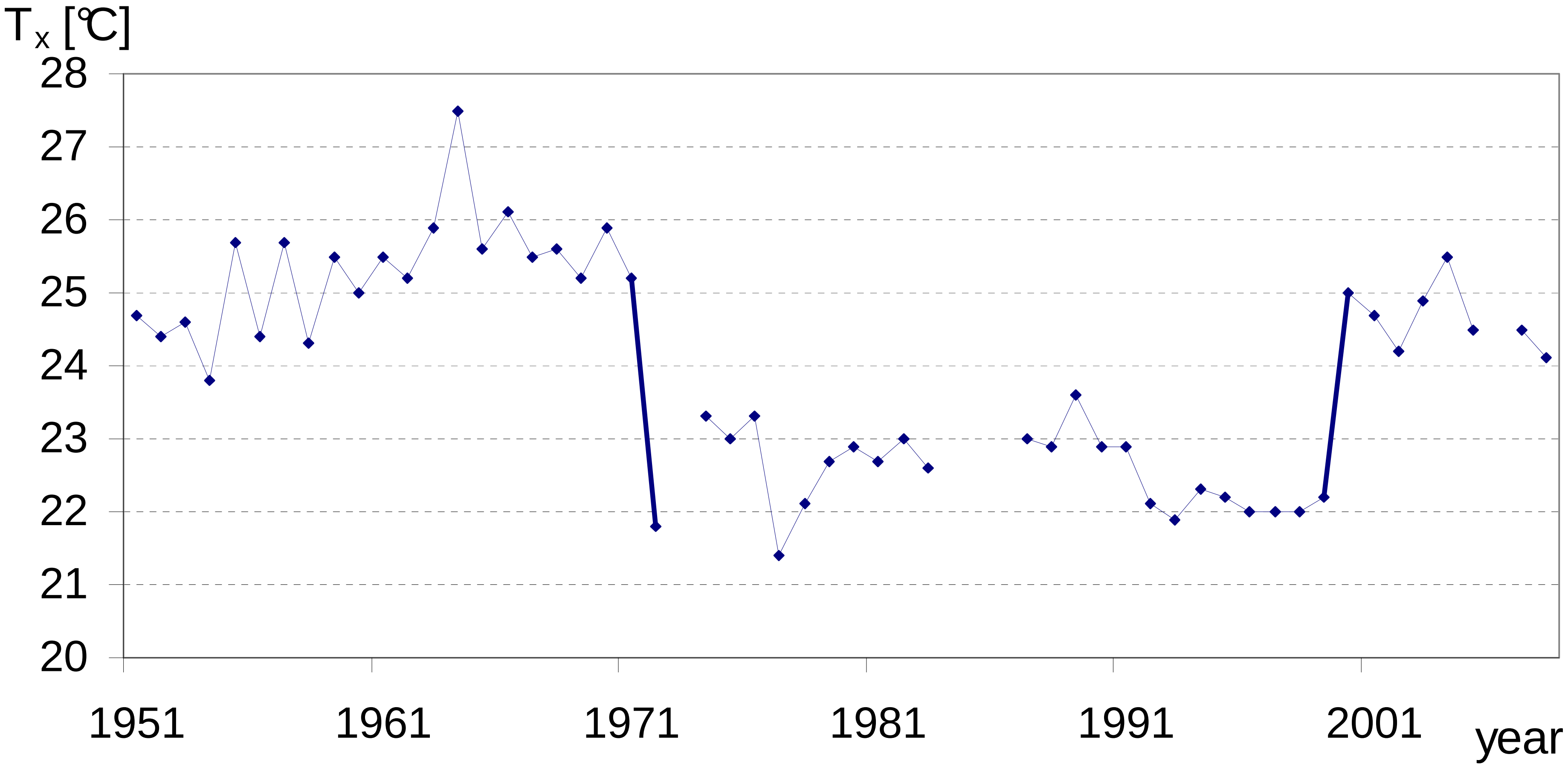
True and detected frequencies of breaks

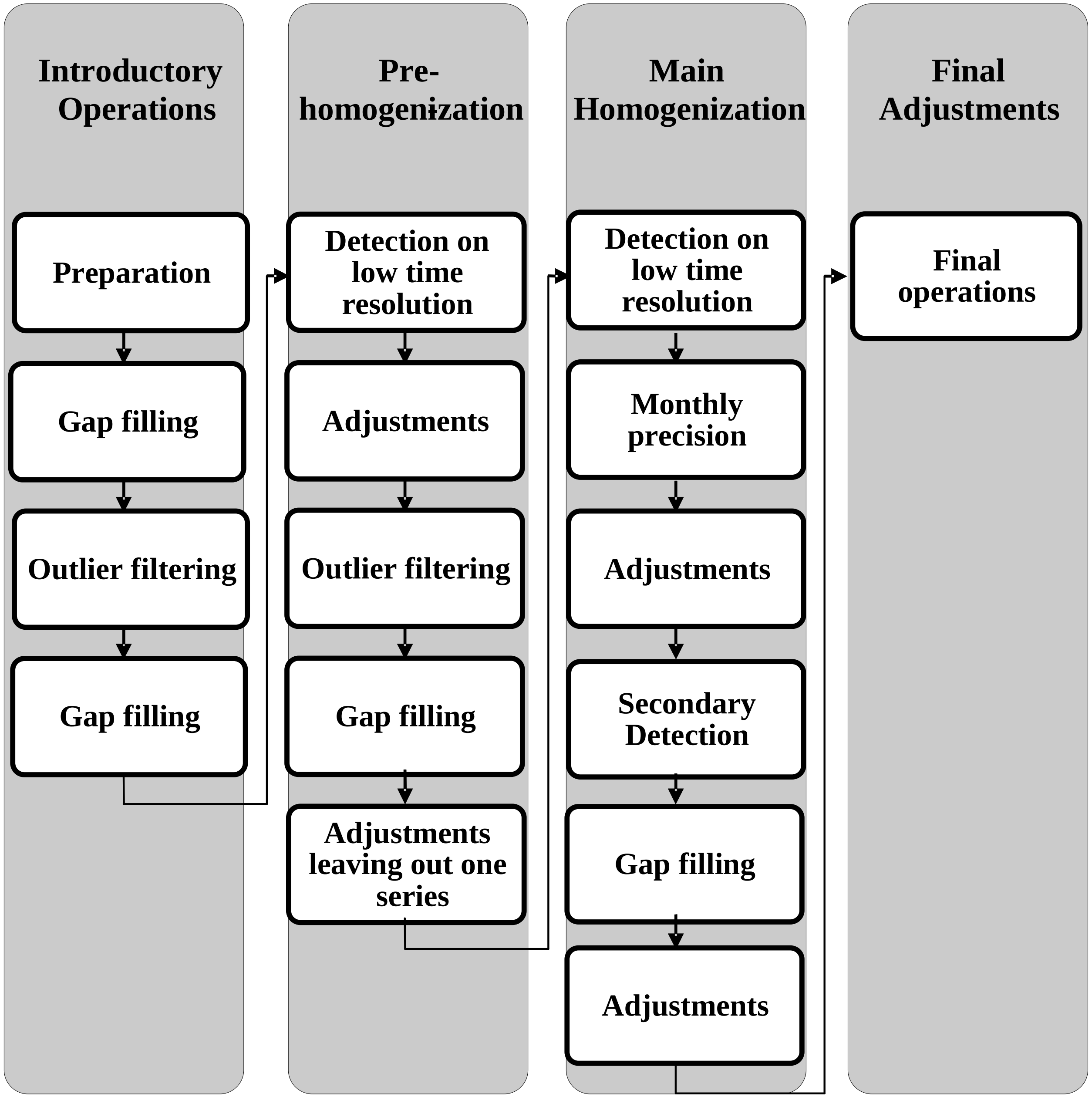
Short-term, platform shaped biases



Inhomogeneous time series with missing data

Monthly mean Tmax in summer, Tarifa, Spain





Introductory Operations
Pre-homogenization
Main Homogenization
Final Adjustments
Preparation
Gap filling
Outlier filtering
Gap filling
Detection on low time resolution
Adjustments
Outlier filtering
Gap filling
Adjustments leaving out one series
Detection on low time resolution
Monthly precision
Adjustments
Secondary Detection
Gap filling
Adjustments
Final operations

Bivariate detection

- In Tmean and Tmax series of mid- and high latitudes:

Station-effects often have seasonal cycle, since they are often related to radiation changes. Therefore in ACMANT (and also in HOMER) **breaks are jointly searched for two annual variables: a) annual mean temperature, b) amplitude of summer - winter difference**.

Bivariate detection

Fitting optimal step function with simultaneous steps to two variables (TM and TD), the number of steps is K'.

$$\min_{[j_1, j_2, \ldots j_{K'}]} :$$

$$\left\{ \sum_{k=0}^{K'} \sum_{i=j_k+1}^{j_{k+1}} (tm_i - \overline{\mathbf{TM_k}})^2 + c_0^2 (td_i - \overline{\mathbf{TD_k}})^2 \right\}$$

Bivariate detection

- **Caussinus –Lyazrhi criterion** for assessing K’:

$$\ln\left\{1-\frac{\sum_{k=0}^{K'}(j_{k+1}-j_k)\cdot\left[(\overline{\mathbf{TM_k}}-\overline{\mathbf{TM}})^2+c_0^2(\overline{\mathbf{TD_k}}-\overline{\mathbf{TD}})^2\right]}{\sum_{i=1}^{L}(tm_i-\overline{\mathbf{TM}})^2+c_0^2(td_i-\overline{\mathbf{TD}})^2}\right\}+G$$

$$G=\frac{2K'}{L-1}\ln(L)$$

j – timing of break, L – length of series

Bivariate detection

Effective when the breaks of TM and TD are simultaneous

- Their detection is more confident and more accurate with bivariate detection than with separate detection processes
- Higher signal-to-noise ratio for 2 annual variables only than for monthly or seasonal variables

Bivariate detection

Effective, even when only a part of the breaks are simultaneous

- If a strongly significant break occurs in any of TM and TD, the break will be detected
- If the break is only in one of the two variables, it may result in a small error (only) in the other variable
- If there is break only in one of the two variables and the break-size is close to the significance-threshold, its detection might fail due to the additional noise term of the other variable

Comparison of HOMER and ACMANT

Common for both methods:

- Fitting optimal step function (Hawkins, 1972)
- Caussinus – Lyazrhi criterion (Caussinus and Lyazrhi, 1997, Caussinus and Mestre, 2004)
- ANOVA adjustments (Caussinus and Metsre, 2004)
- Bivariate detection (Domonkos, 2011)
- Monthly precision of breaks (Domonkos, 2011)

Comparison of HOMER and ACMANT

HOMER	ACMANT
• Pairwise comparisons	• Composite refer. series
• Joint detection for all breaks of the network	• Pre-homogenization
• Gap filling with ANOVA	• Gap filling with interpol.
• ANOVA: separately for seasons	• ANOVA: for annual variables only
• Does not have routines for break detection on monthly scale	• Has routines for break detection on monthly scale

Comparison of HOMER and ACMANT

HOMER

- For small or medium-size networks, particularly with metadata use

ACMANT

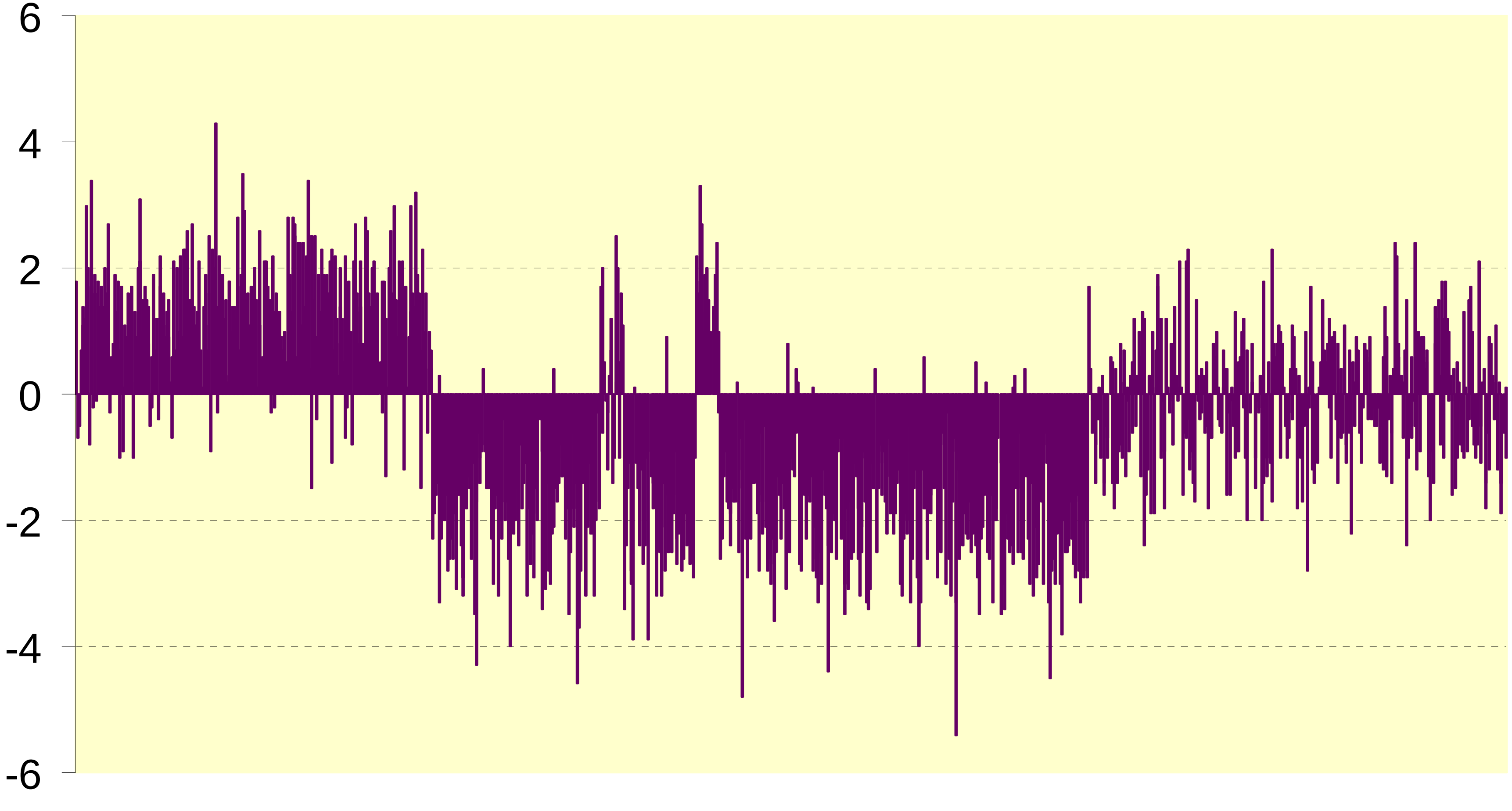
- For homogenizing without metadata
- Particularly for very large networks

  (> 50 stations)

Properties of ACMANT2

- ACMANT2 is a relative homogenization method, at least 4 spatially correlated time series are needed for its use.
- ACMANT2 uses composite reference series for spatial comparisons. In its model the spatial gradients of climate are temporally constant, but the reference composites are not expected to be homogeneous.
- ACMANT2 intends to separate the work on annual and monthly scales in order to minimize the problem of interference between problems on different time scales.

A series with long-term and short-term biases



Time series comparison

- Composite reference series are built from the other series of the network, with which the candidate series has at least 0.4 correlation. The composites are weighted with the squared spatial correlation of the increment series of the monthly values.
- If the time series cover different time periods, different reference series are built for homogenizing various sections of the same candidate series. ACMANT always selects the best reference series, i.e. which is built from the highest number of available composites.

More about break detection

- Breaks are searched in difference series between the candidate and reference series. In homogenizing precipitation, first the values are transformed to make them similar to additive variables.

- Two-phase homogenization is performed. The first round is named pre-homogenization. When reference series are built for the second homogenization, the composites are already adjusted, but the later candidate series is always excluded from the calculation of adjustment terms of pre-homogenization.

More about break detection

- For Tmin and for any temperature variable in the tropics and under monsoon climate, univariate detection is performed and seasonal cycle of biases is not considered.
- For precipitation, bivariate detection is applied under cool or cold climates, the two variables are the precipitation total of snowy season and that of the rainy season.
- For precipitation, breaks are searched only on annual scale and no automatic outlier filtering is performed.

Calculation of adjustment-terms

- It has been proven that the minimization of the common variance of the homogenized values (ANOVA) provides the optimal estimation with step function models if a) the climate signal is the same for the network and b) the timings of breaks are correctly detected. (Caussinus and Mestre, 2004.)
- Experiments with lists of the detected breaks (produced in HOME benchmark homogenization with various methods) proved that the inclusion of ANOVA generally raises the efficiency (Domonkos et al., 7th Homogenization Seminar).

Problems with using ACMANT or HOMER

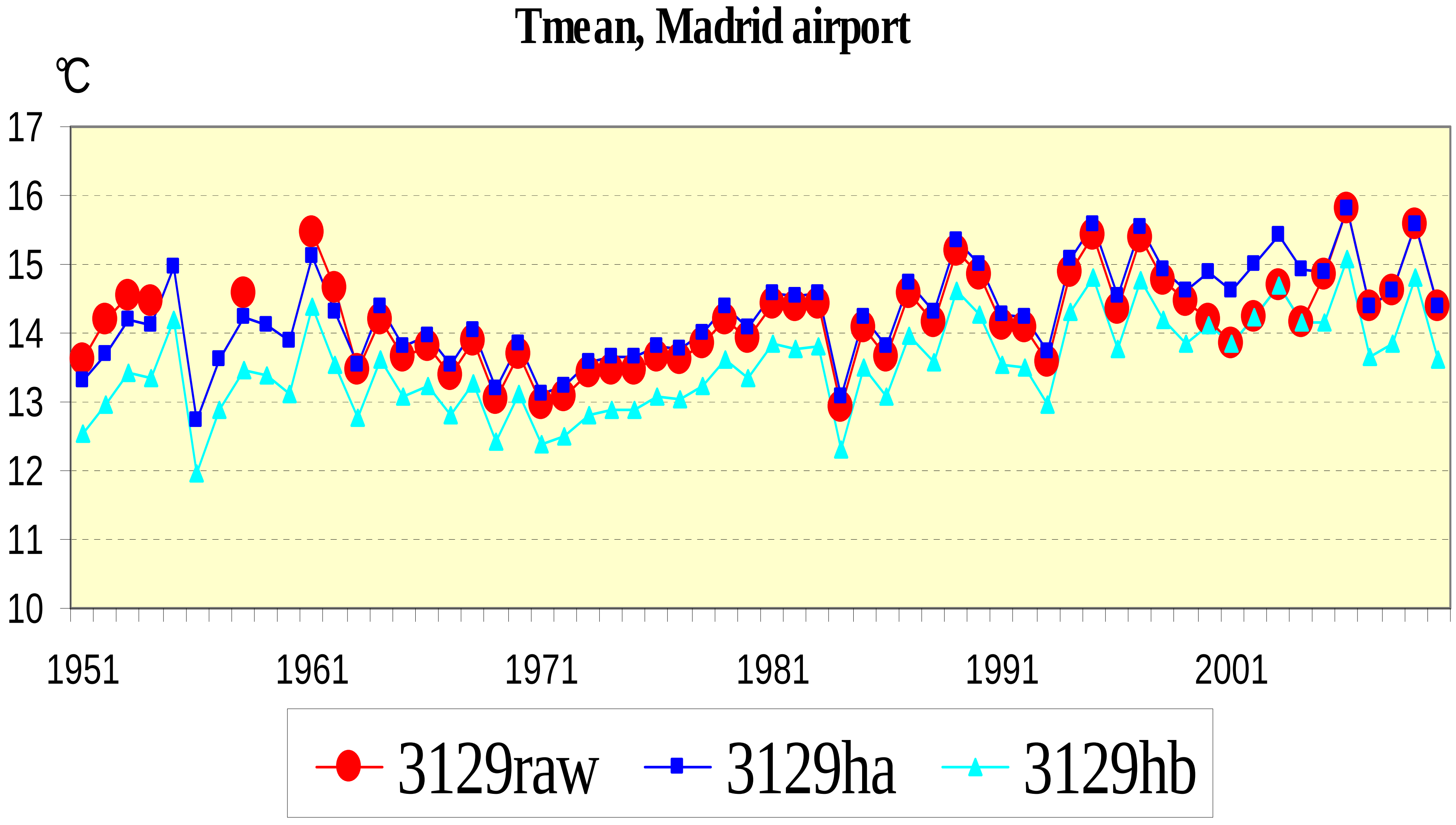
- Raw data is expected to be free of obvious errors, physical outliers.
- Zero values instead of missing data code in precipitation series
- Temporally concordant breaks. – In this case the spatial comparison of the data does not help to find the inhomogeneities
- Known biases quantified by parallel measurements and known concordant shifts of the means are expected to be corrected before running ACMANT.

Problems with using ACMANT or HOMER

- Breaks are adjusts relative to the last homogeneous section of time series. If this section does not represent well the true climate of the site, the spatial variance of the homogenized data will be biased. – For the same reason, the inclusion of quantiles method in downscaling to daily data adjustments, is not planned within automatic homogenization.

The impact of reference period on the homogenisation result

Dark blue line is fitted to the last homogeneous section, while light blue line to the last but one section of the series. The two blue lines are parallel.



Error sources of ACMANT2 results

- Bad allocation of inhomogeneities due to coincidental problems in more than 1 series *
- Inappropriate model of the seasonality of biases
- Zero values instead of missing data code
- Low number of simultaneously observed data
- Spatial heterogeneity of climate *

Thank you for your attention!