Homogenization of tropospheric data: evaluating the algorithms under the presence of autoregressive process

Presentation · June 2018

14 authors, including:

Anna Klos
Military University of Technology
81 PUBLICATIONS **175** CITATIONS

SEE PROFILE

Eric Pottiaux
Royal Observatory of Belgium
68 PUBLICATIONS **280** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

geodynamical studies in the North of Algeria View project

Noise characteristics in Zenith Total Delay time series View project

Introduction   Data   Synthetic benchmark   Tools   Results   Summary

Homogenization of tropospheric data: evaluating the algorithms under the presence of autoregressive process

Anna Klos[1], Roeland Van Malderen[2], Eric Pottiaux[3],
Fadwa Alshawaf[4], Olivier Bock[5], Janusz Bogusz[1],
Barbara Chimani[6], Peter Domonkos[7], Michal Elias[8], Marta Gruszczynska[1],
José Guijarro[9], Selma Zengin Kazancı[10], Tong Ning[11], Vincenza Tornatore[12]

1) Military University of Technology, Warsaw, Poland,
2) Royal Meteorological Institute of Belgium (RMI), Brussels, Belgium,
3) Royal Observatory of Belgium (ROB), Brussels, Belgium,
4) GFZ German Research Centre for Geosciences, Potsdam, Germany
5) IGN LAREG, University Paris Diderot, Sorbonne Paris, France,
6) Central Institute for Meteorology and Geodynamics, Austria,
7) Climate Researcher, Tortosa, Spain,
8) Research Institute of Geodesy, Topography and Cartography, Czech Republic,
9) AEMET (Spanish Meteorological Agency), Spain,
10) Karadeniz Technical University, Turkey,
11) Lantmäteriet, Sweden.
12) DICA Politecnico di Milano, Italy

Introduction:

1. COST action **GNSS4SWEC** „Advanced Global Navigation Satellite Systems tropospheric products for monitoring severe weather events and climate", WG3: Use of GNSS tropospheric products for climate monitoring.

2. A **proper homogenization** of tropospheric dataset is indispensable, as the parameters of deterministic part, e.g. **trend** will be influenced by undetected breaks.

3. Different groups / different methods / different estimates – **the truth is not known**.

4. A synthetic benchmark dataset: a way to quantify results given by various algorithms.

Introduction:

How does it look in practice?

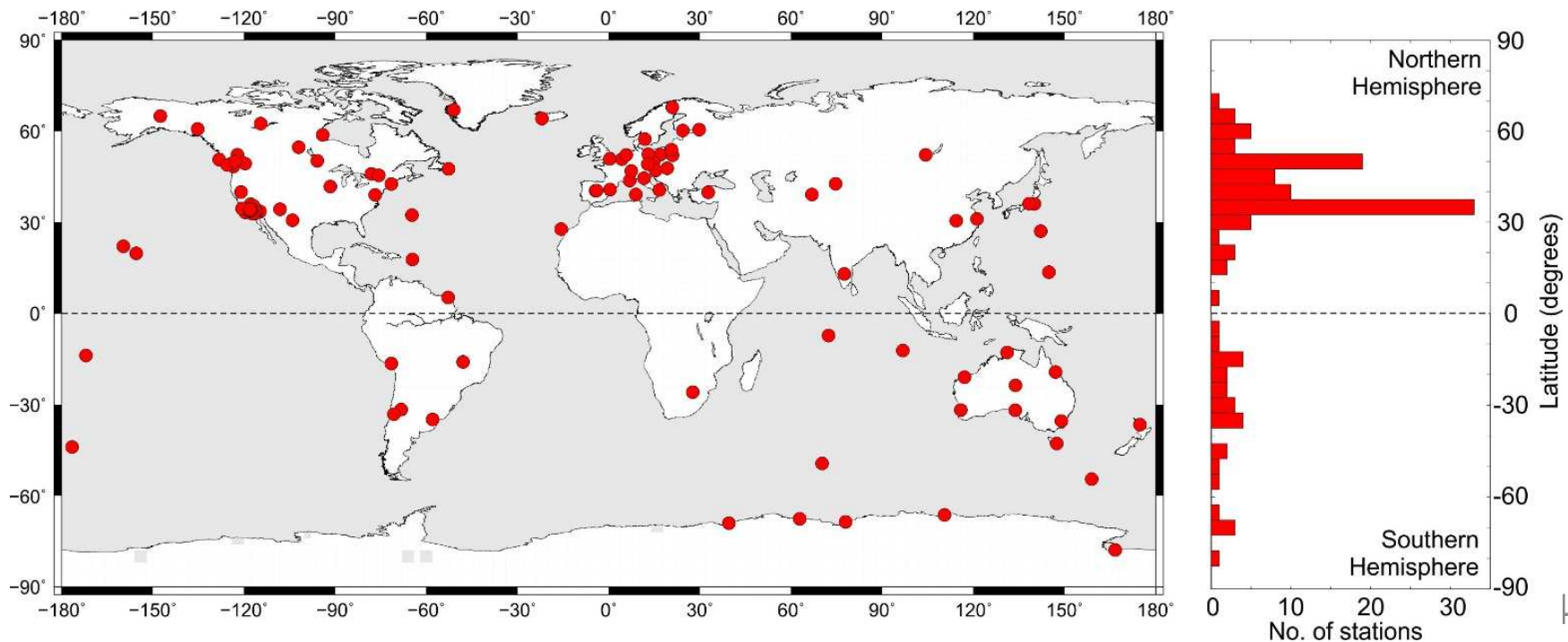A change in trend possible and **very likely!**

Can anyone see more offsets?

What we aim at?
Only real breakpoints not regime-like shifts should be corrected!


3
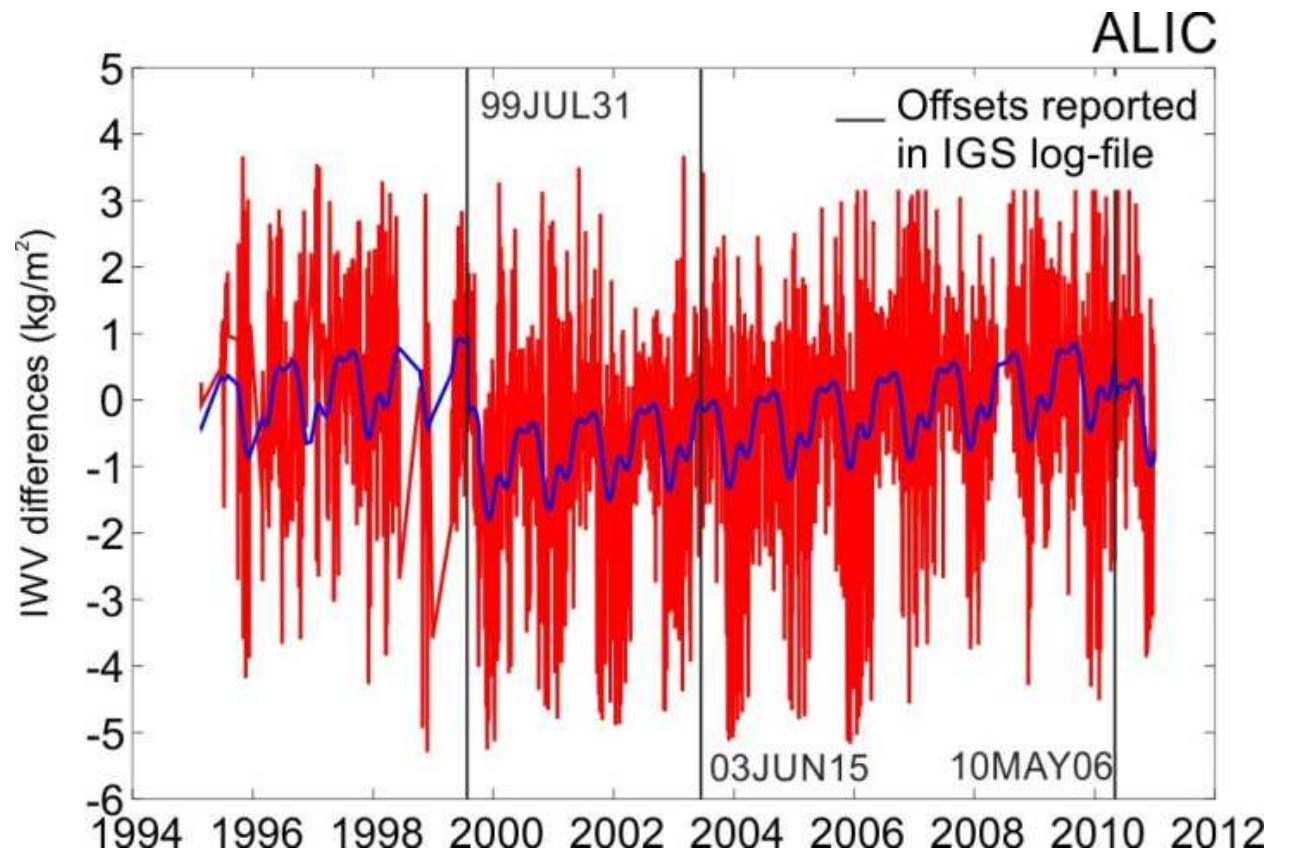
Data:

1. IGS „repro1" troposphere products screened and converted to Integrated Water Vapor (IWV) by O. Bock.

2. 120 stations, daily observations, a period of 1995-2010.

3. The IWV differences: ERAI-GPS were employed.

Generation of the benchmark:

3 variants of synthetic time series were generated:

1. **EASY** dataset: seasonal signals + offsets + white noise (WN),

2. **LESS COMPLICATED** dataset: same as 1. + autoregressive process of the first order (noise model = AR(1)+WN),

3. **FULLY COMPLICATED** dataset: same as 2. + trend + gaps (up to 20% of missing data).
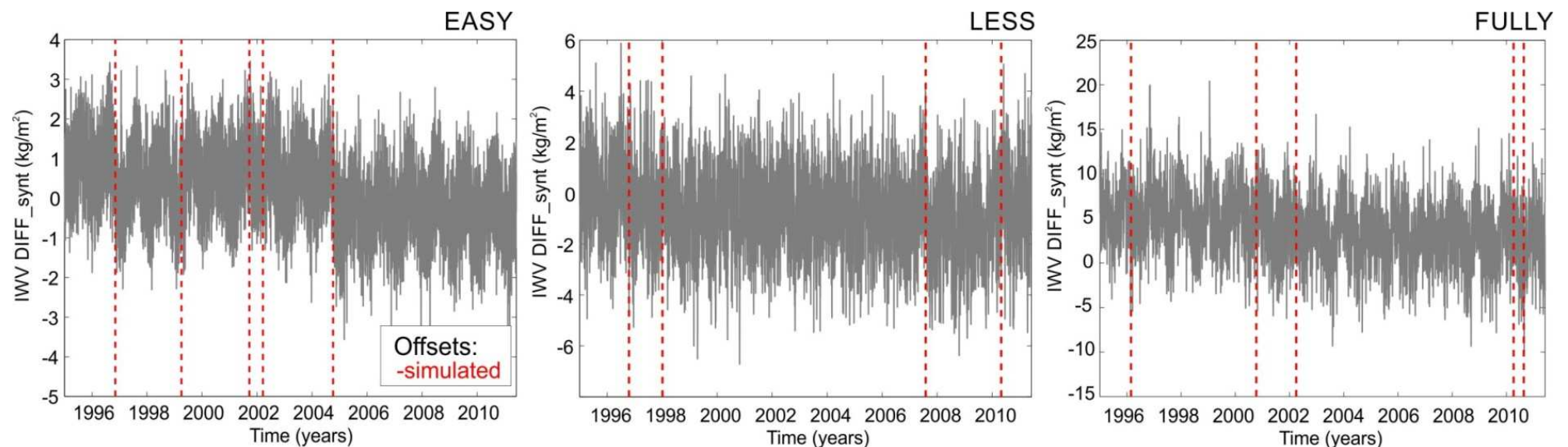
Generation of the benchmark:

1. **120 series in each synthetic dataset** simulated.

2. Deterministic model of data taken directly from real differences: trend, seasonal signals, noise.

3. Offsets simulated **randomly**.

4. Number of offsets and exact epochs are **blinded**.

Can anyone see more offsets?

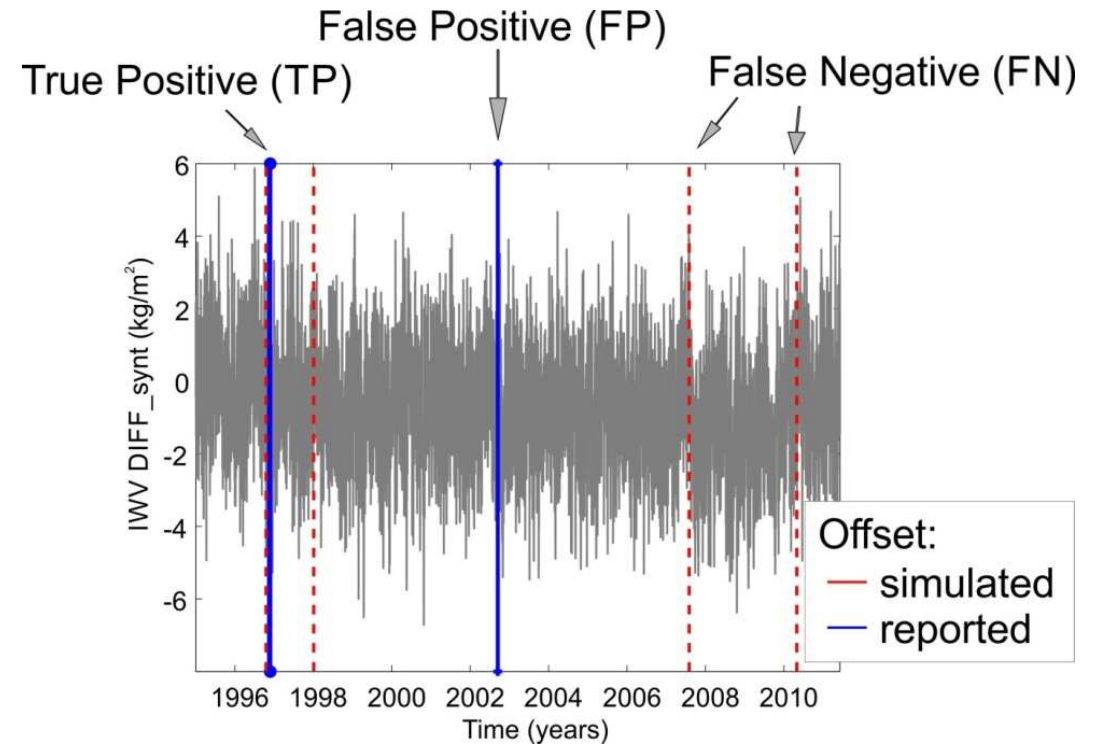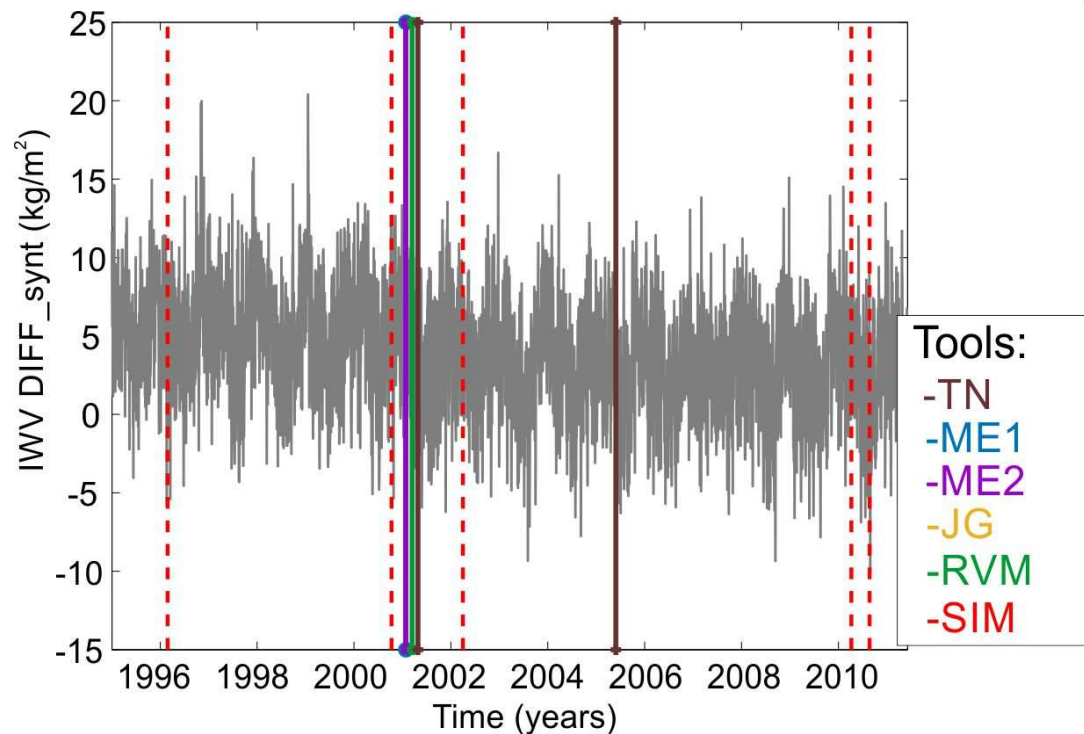**Introduction   Data   Synthetic benchmark   Tools   Results   Summary**
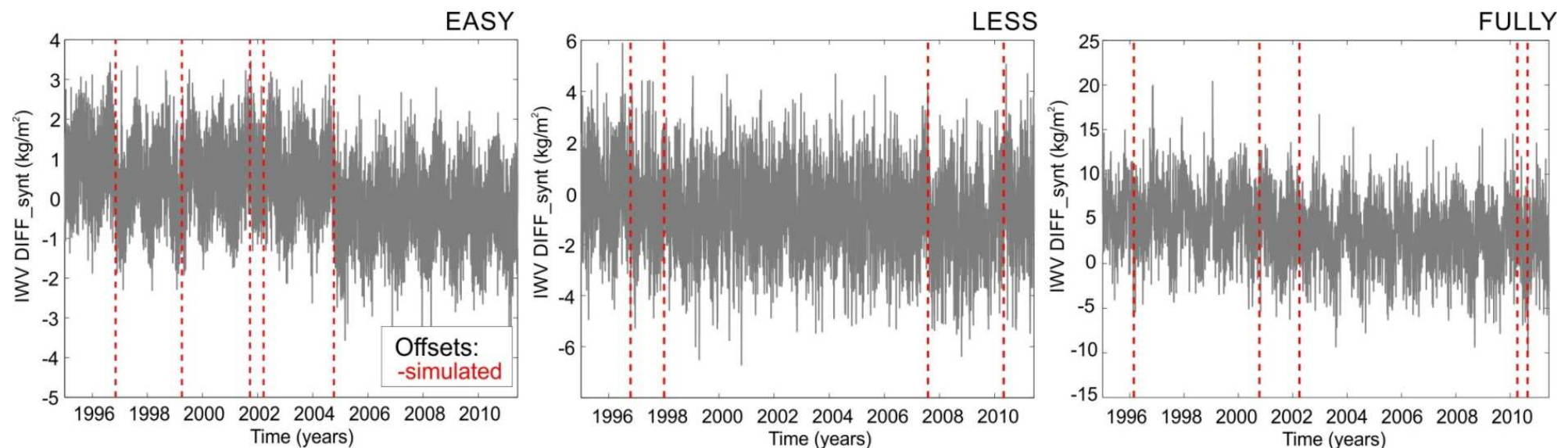
Algorithms:

1. Sensitivity analysis: the identification of the **epochs** of the inserted breakpoints.

2. Estimates of the **trends** of the 3 sets of synthetic IWV differences.

	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8
Symbol	🟡	🔺	➕	✖	🔶	🔻	-	🟧
Operator	M. Elias	R. Van Malderen	R. Van Malderen	J. Guijarro	T. Ning	S. Zengin Kazanci	B. Chimani	M. Gruszczynska
Method / SW	2-sample t-test	2of3	PMW	CLIMATOL	PMTred	Pettitt test	HOMOP	STARS
Daily / Monthly	D+M	D+M	D+M	D+M	D+M	D	X	D
Easy / Less / Fully	E+L+F	E+L+F	E+L+F	L+F	E+L+F	E+L+F	E+F	E+L+F

Introduction   Data   Synthetic benchmark   Tools   Results   Summary

How to classify breaks?

Defining a proper time
window – **2 months**

Introduction   Data   Synthetic benchmark   Tools   <u>Results</u>   Summary
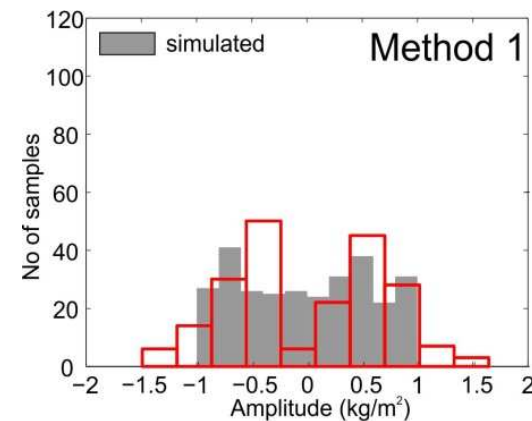
Offsets:

	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8
Symbol	●	▲	✚	✖	◆	▼	-	◼
Operator	M. Elias	R. Van Malderen	R. Van Malderen	J. Guijarro	T. Ning	S. Zengin Kazanci	B. Chimani	M. Gruszczynska
Method / SW	2-sample t-test	2of3	PMW	CLIMATOL	PMTred	Pettitt test	HOMOP	STARS
Daily / Monthly	D+M	D+M	D+M	D+M	D+M	D	X	D
Easy / Less / Fully	E+L+F	E+L+F	E+L+F	L+F	E+L+F	E+L+F	E+F	E+L+F

Amplitudes of reported offsets:

EASY, DAILY (SIM: 291):

- method 1:   211,
- method 2:   252,
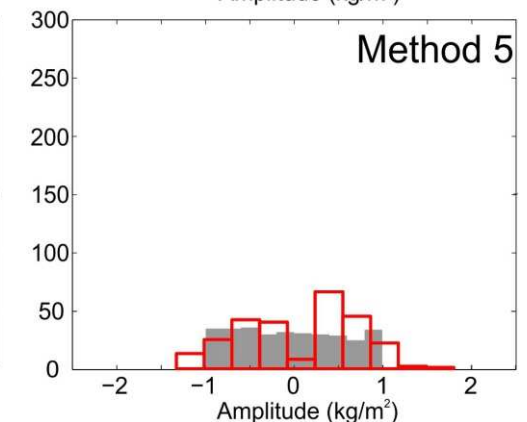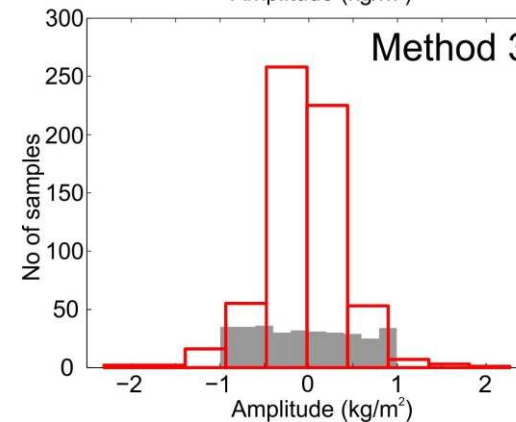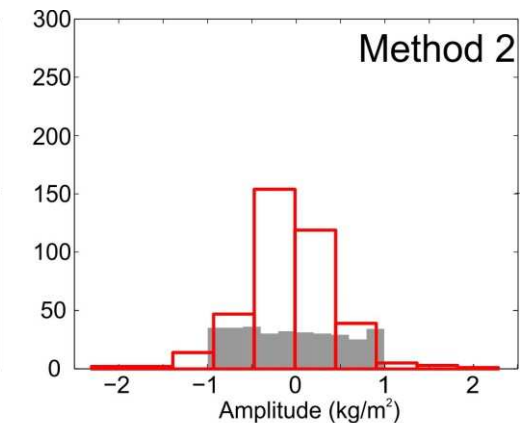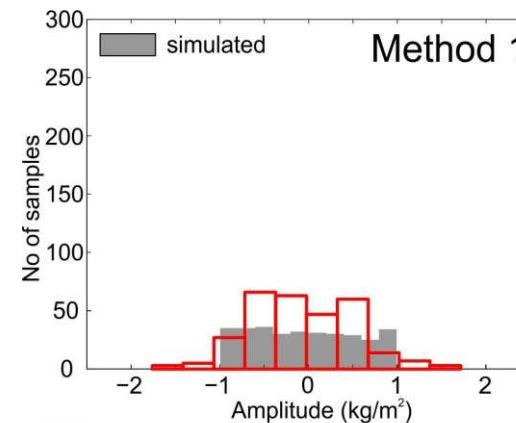- method 3:   377,
- method 5:   216,
- method 8:   347.

Introduction   Data   Synthetic benchmark   Tools   <u>Results</u>   Summary

<u>Offsets:</u>

	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8
Symbol	●	▲	✚	✖	◆	▼	-	■
Operator	M. Elias	R. Van Malderen	R. Van Malderen	J. Guijarro	T. Ning	S. Zengin Kazanci	B. Chimani	M. Gruszczynska
Method / SW	2-sample t-test	2of3	PMW	CLIMATOL	PMTred	Pettitt test	HOMOP	STARS
Daily / Monthly	D+M	D+M	D+M	D+M	D+M	D	X	D
Easy / Less / Fully	E+L+F	E+L+F	E+L+F	L+F	E+L+F	E+L+F	E+F	E+L+F

Amplitudes of reported offsets:

FULLY-COMPLICATED, DAILY (SIM: 317):
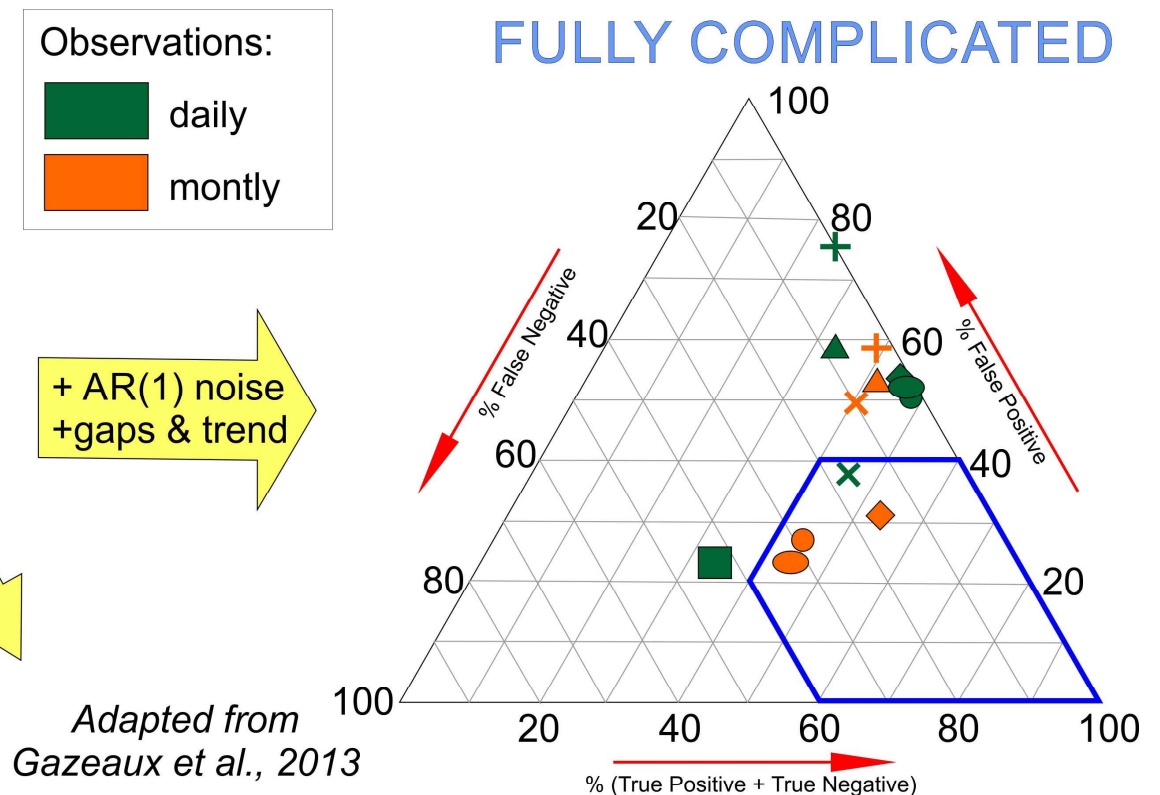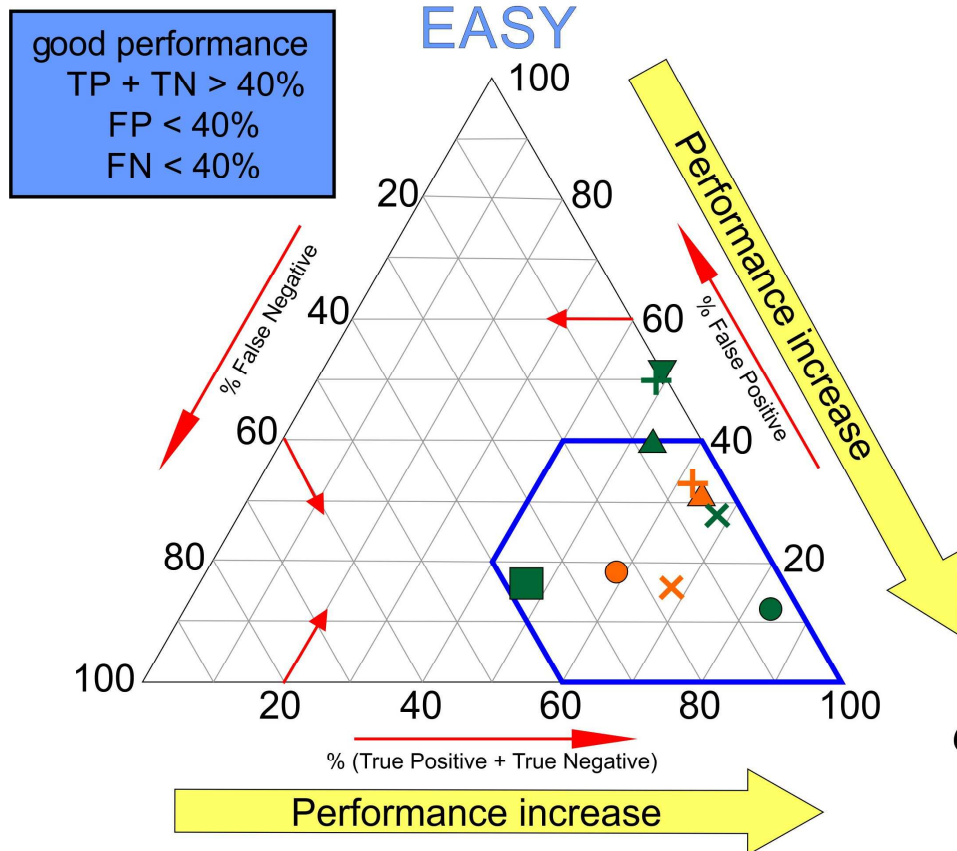
- method 1:   295,
- method 2:   386,
- method 3:   622,
- method 5:   264,
- method 8:   433.

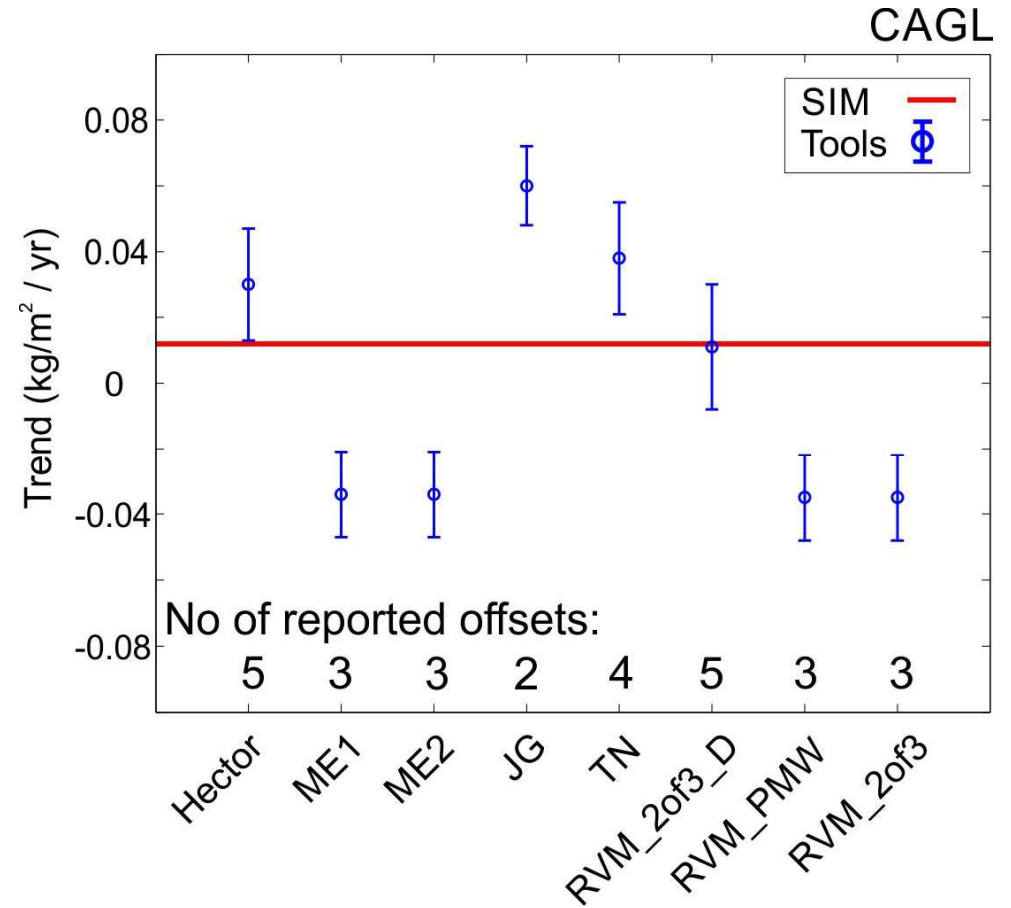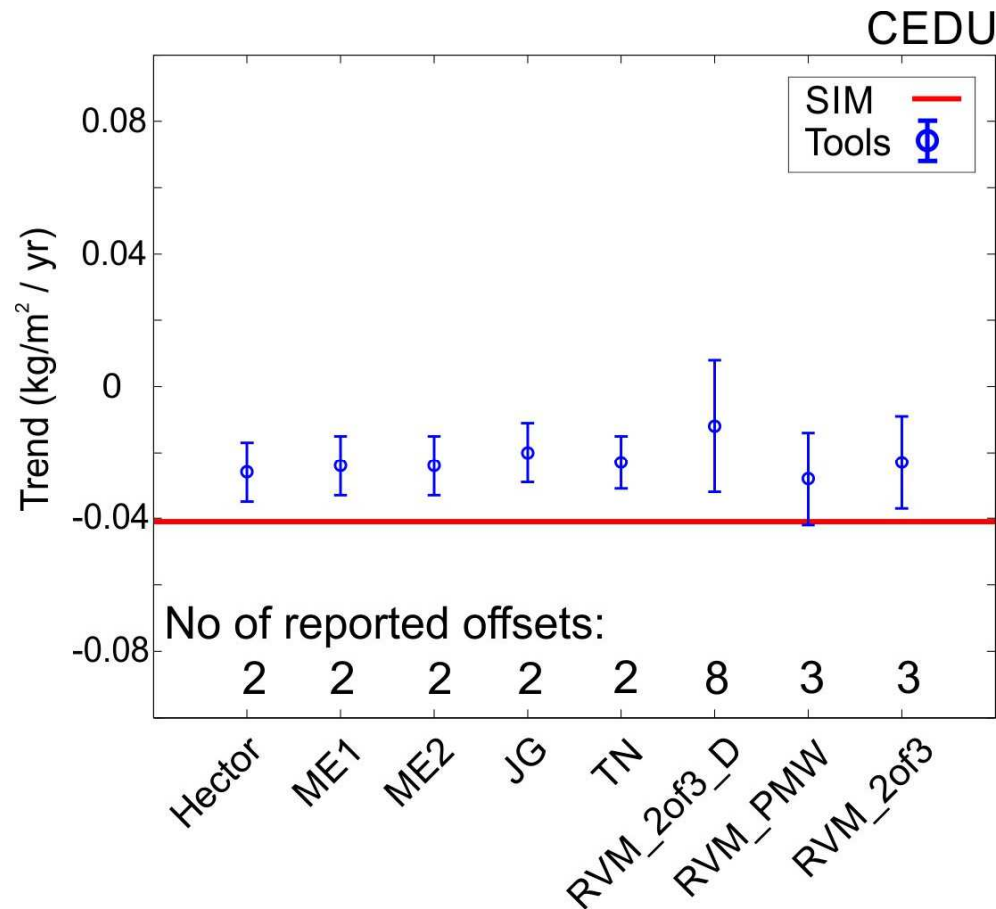**Introduction   Data   Synthetic benchmark   Tools   <u>Results</u>   Summary**

Tools performance:

	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8
Symbol	●	▲	✚	✖	◆	▼	-	■
Operator	M. Elias	R. Van Malderen	R. Van Malderen	J. Guijarro	T. Ning	S. Zengin Kazanci	B. Chimani	M. Gruszczynska
Method / SW	2-sample t-test	2of3	PMW	CLIMATOL	PMTred	Pettitt test	HOMOP	STARS
Daily / Monthly	D+M	D+M	D+M	D+M	D+M	D	X	D
Easy / Less / Fully	E+L+F	E+L+F	E+L+F	L+F	E+L+F	E+L+F	E+F	E+L+F

good performance
TP + TN > 40%
FP < 40%
FN < 40%

EASY

Observations:
daily
montly

FULLY COMPLICATED

+ AR(1) noise
+gaps & trend

Performance increase

% False Negative

% False Positive

% (True Positive + True Negative)

Adapted from Gazeaux et al., 2013

Introduction   Data   Synthetic benchmark   Tools   <u>Results</u>   Summary

<u>Changes in Trends:</u>
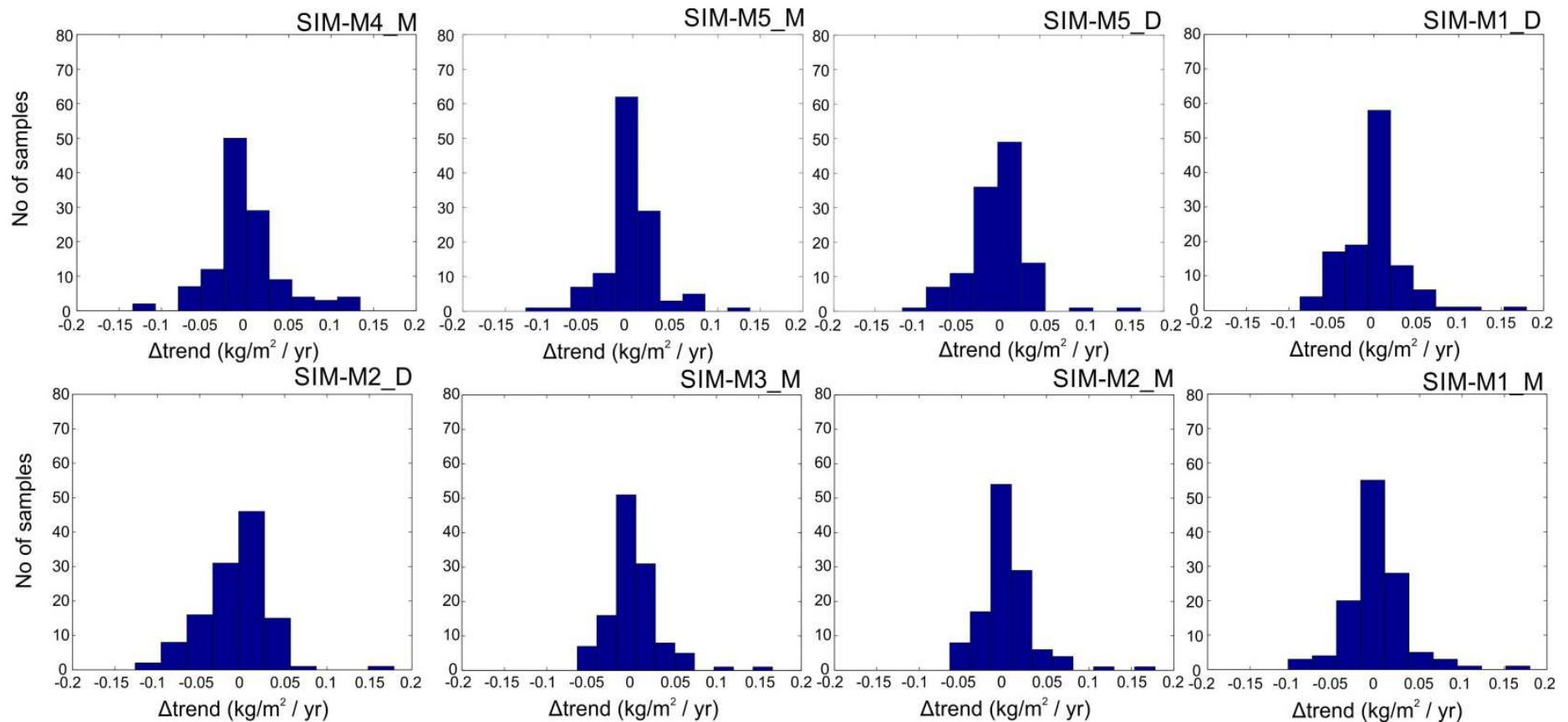
FULLY-COMPLICATED

Introduction   Data   Synthetic benchmark   Tools   <u>Results</u>   Summary

Trends:

	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8
Symbol	●	▲	✚	✖	◆	▼	-	■
Operator	M. Elias	R. Van Malderen	R. Van Malderen	J. Guijarro	T. Ning	S. Zengin Kazanci	B. Chimani	M. Gruszczynska
Method / SW	2-sample t-test	2of3	PMW	CLIMATOL	PMTred	Pettitt test	HOMOP	STARS
Daily / Monthly	D+M	D+M	D+M	D+M	D+M	D	X	D
Easy / Less / Fully	E+L+F	E+L+F	E+L+F	L+F	E+L+F	E+L+F	E+F	E+L+F

Introduction   Data   Synthetic benchmark   Tools   Results   <u>Summary</u>

<u>Now:</u>

1. A detailed assessment of tools sensitivity.

2. The epochs of breaks were given to the participants to fine-tune their methods.

3. A next generation of synthetic benchmark is ongoing.

Introduction   Data   Synthetic benchmark   Tools   Results   Summary

And then...

1. The best performing tools are going to be employed to homogenize the IGS repro1.

2. A need to define the reliable strategy for homogenization.

http://iag-gnssclimate.oma.be/index.php

Introduction   Data   Synthetic benchmark   Tools   Results   Summary

Acknowledgments

Thank you!