Performance of various homogenization tools on a synthetic benchmark dataset of GPS and ERA-interim IWV differences

Presentation · August 2017

11 authors, including:

Eric Pottiaux
Royal Observatory of Belgium
65 PUBLICATIONS **219** CITATIONS
SEE PROFILE

Olivier Bock
Institut national de l'information géographiq…
169 PUBLICATIONS **1,821** CITATIONS
SEE PROFILE

Janusz Bogusz
Military University of Technology
169 PUBLICATIONS **329** CITATIONS
SEE PROFILE

Michal Elias
Research Institute of Geodesy, Topography a…
8 PUBLICATIONS **6** CITATIONS
SEE PROFILE

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

Project — IAG Joint Working Group 4.3.8: "GNSS Tropospheric Products for Climate" View project

Project — MULTITEST (Multiple verification of automatic softwares homogenizing monthly temperature and precipitation series) View project

Performance of various homogenization tools on a synthetic benchmark dataset of GPS and ERA-interim IWV differences

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

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) IGN LAREG, University Paris Diderot, Sorbonne Paris, France,
5) Central Institute for Meteorology and Geodynamics, Austria,
6) Research Institute of Geodesy, Topography and Cartography, Czech Republic,
7) AEMET (Spanish Meteorological Agency), Spain,
8) Karadeniz Technical University, Turkey,
9) Lantmäteriet, Sweden.

GNSS 4 SWEC

Motivation & 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.

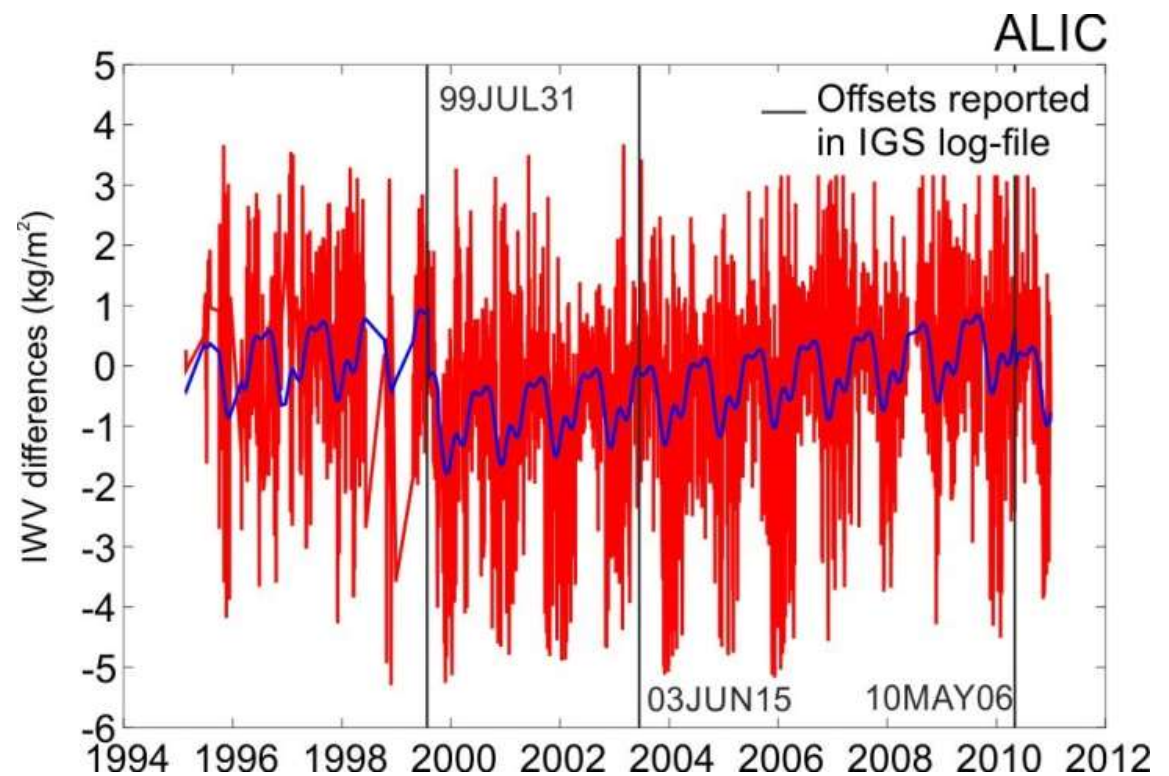**Motivation** Data Synthetic benchmark Tools Results Summary

Motivation & 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

Motivation   Data   Synthetic benchmark   Tools   Results   Summary

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.


Dataset available at: https://doi.org/10.14768/06337394-73a9-407c-9997-0e380dac5590.

4
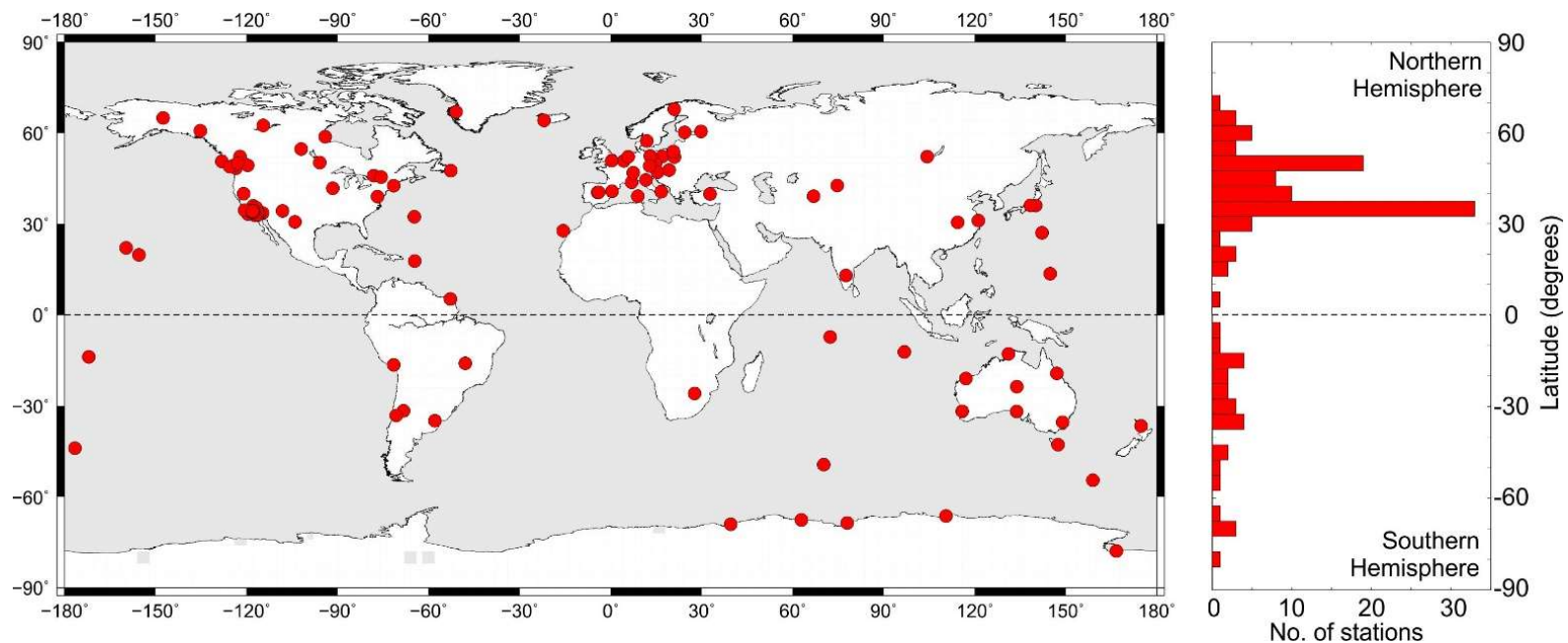
Data:

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.

5

Motivation Data Synthetic benchmark Tools Results Summary
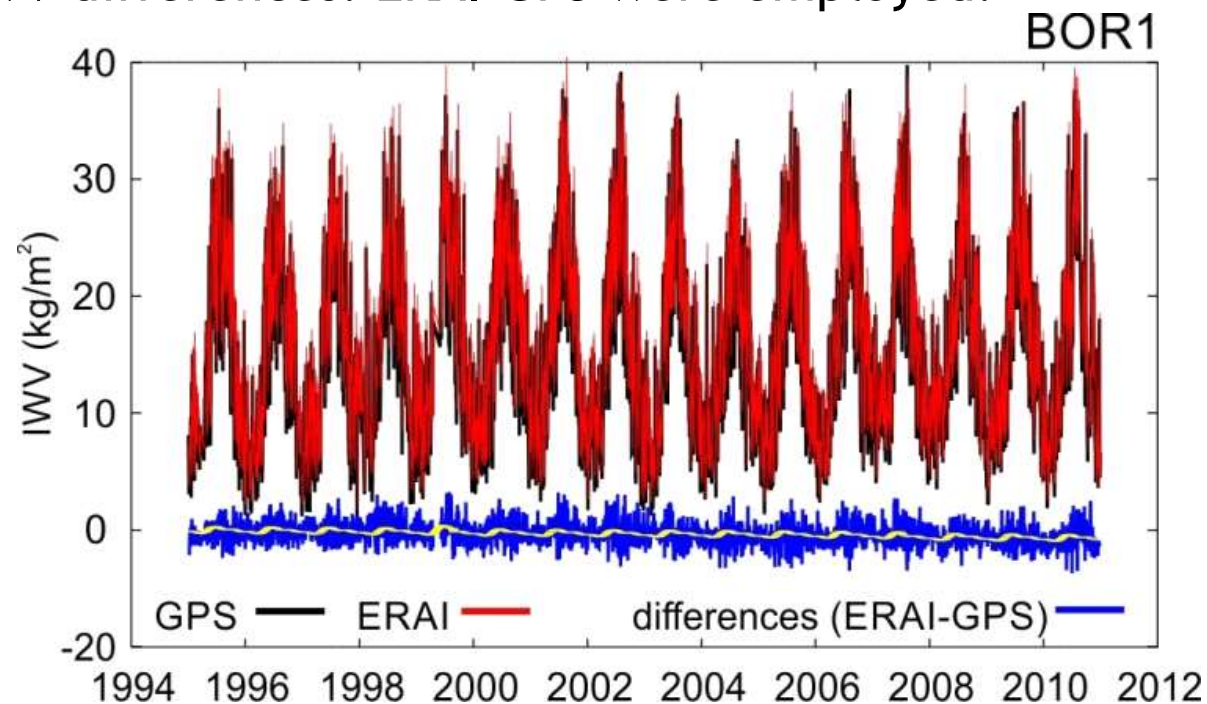
Real data analysis:

1. Manual homogenization + IGS log files: **221 epochs**.
2. Analysis of significant frequencies: **Power Spectral Densities.**
3. **Maximum Likelihood Estimation** (MLE) employed.
4. Noise analysis: **AR(1)+WN** chosen as the preferred noise model.
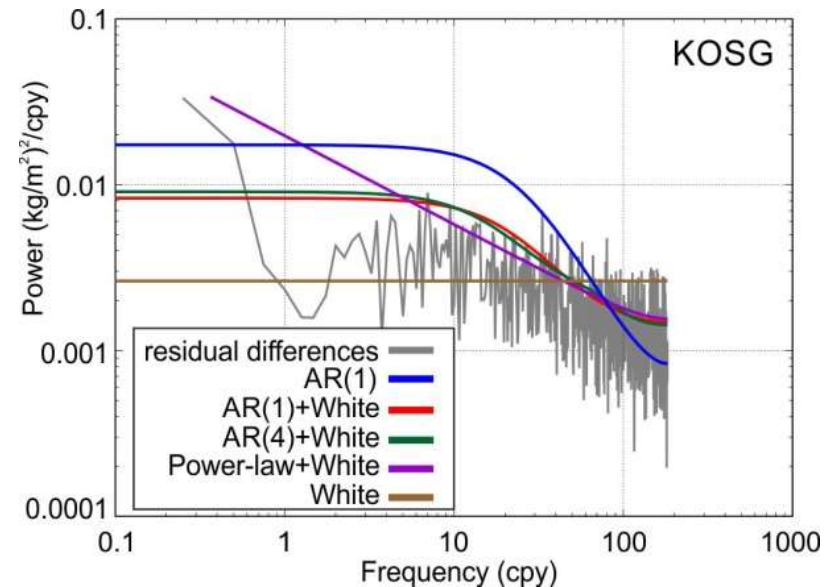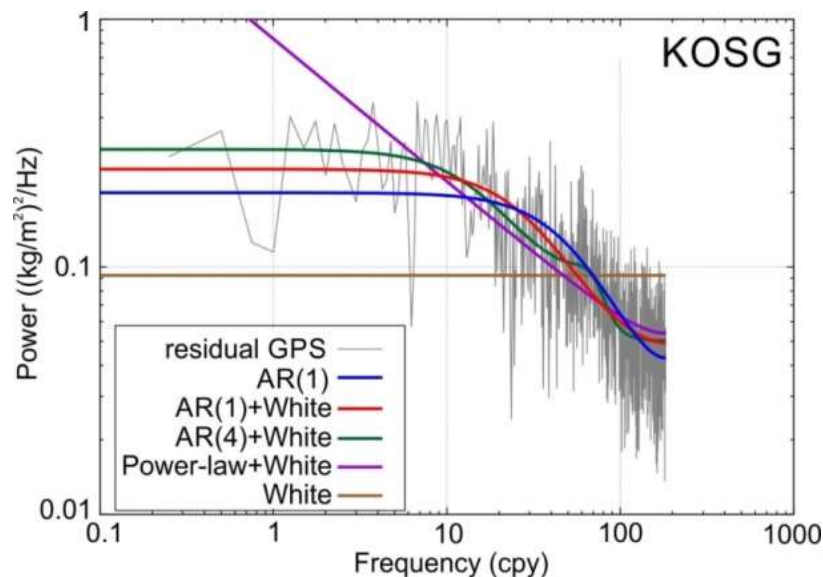
Generation of the benchmark:

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).
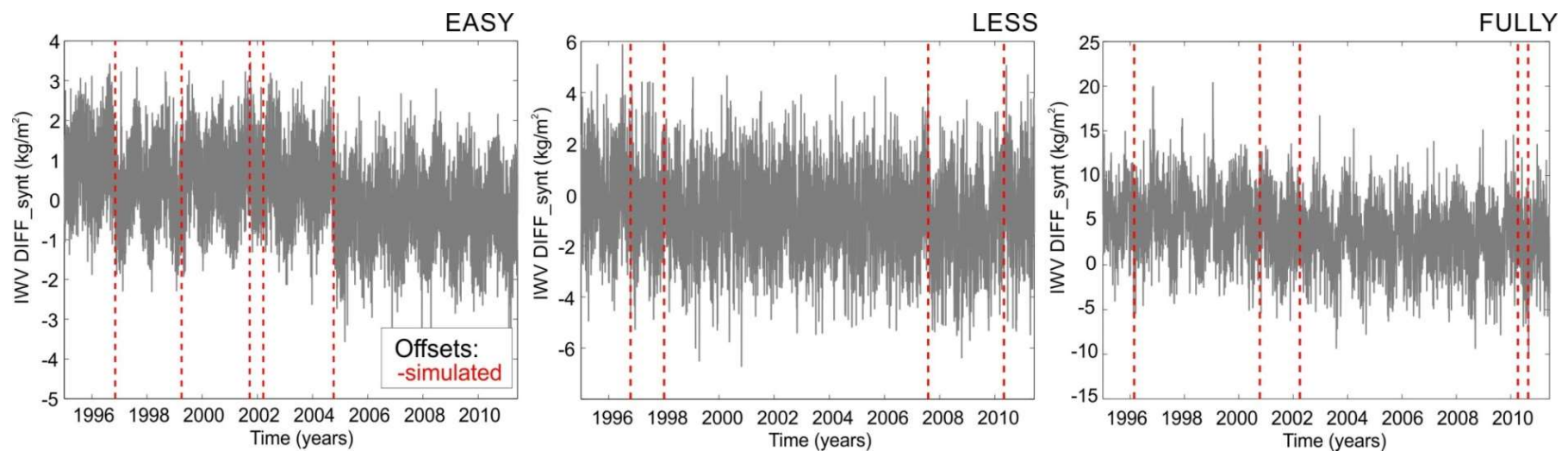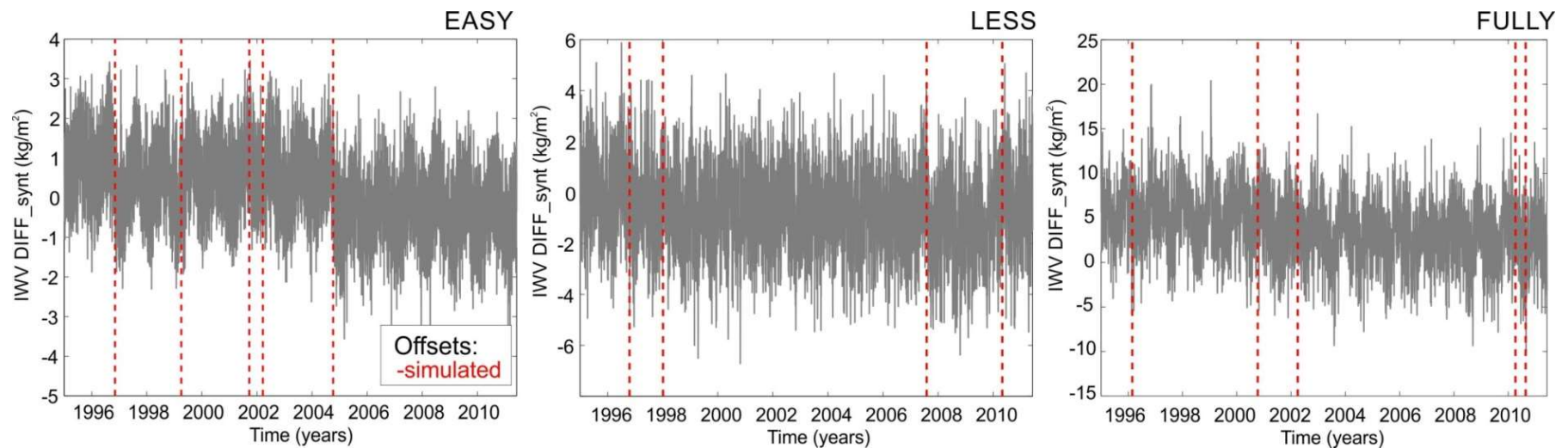
Generation of the benchmark:

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**.

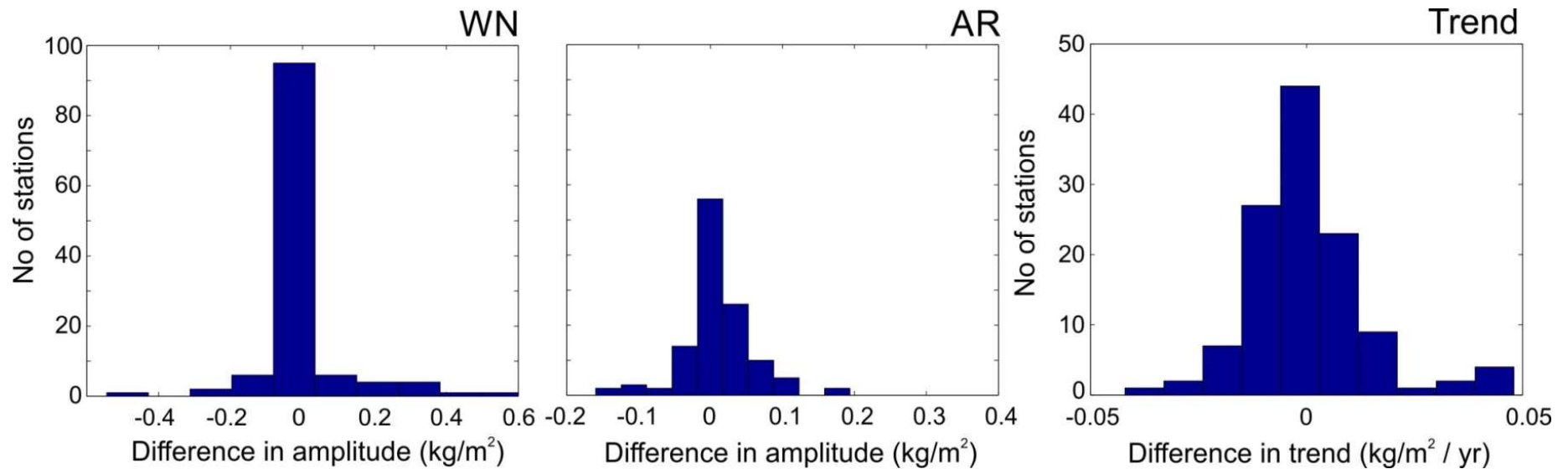**Motivation   Data   Synthetic benchmark   Tools   Results   Summary**

Verfication of the benchmark:

FULLY-COMPLICATED

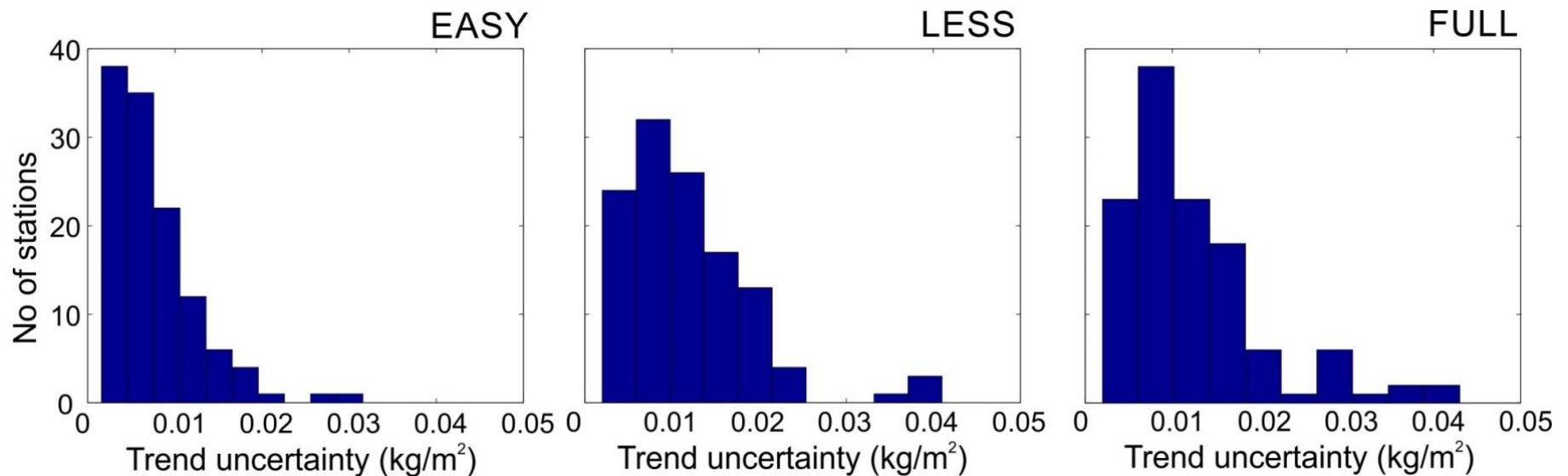Coefficients of AR agree within 0.05.

Amplitudes & trends agree with – see histograms.

Expected trend uncertainty:

Preliminary estimation of which trend uncertainty can be expected from real differences, based on synthetic differences.

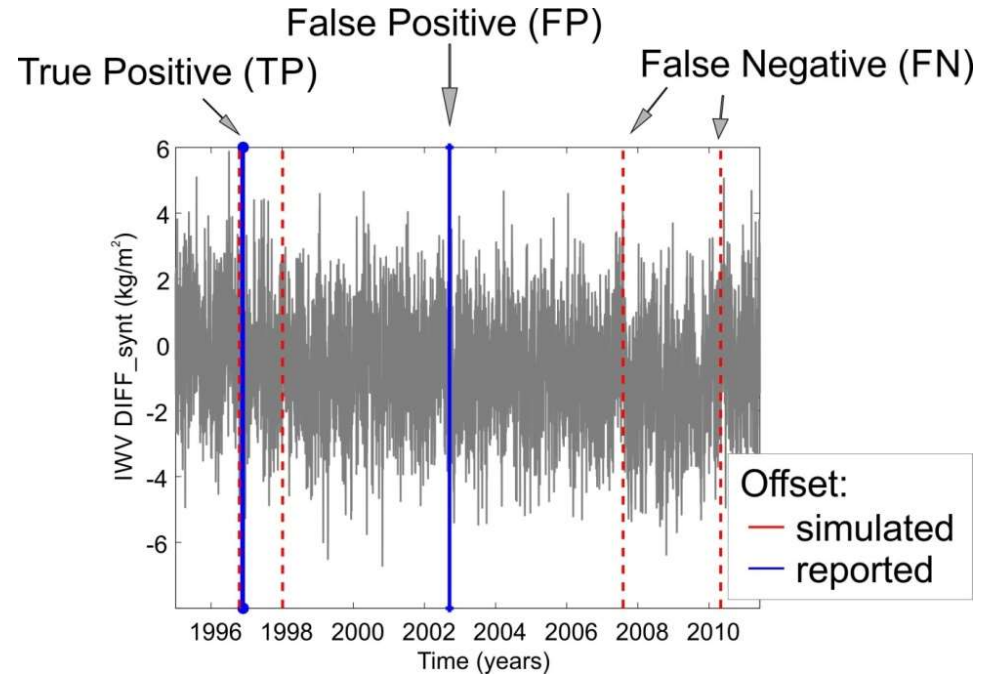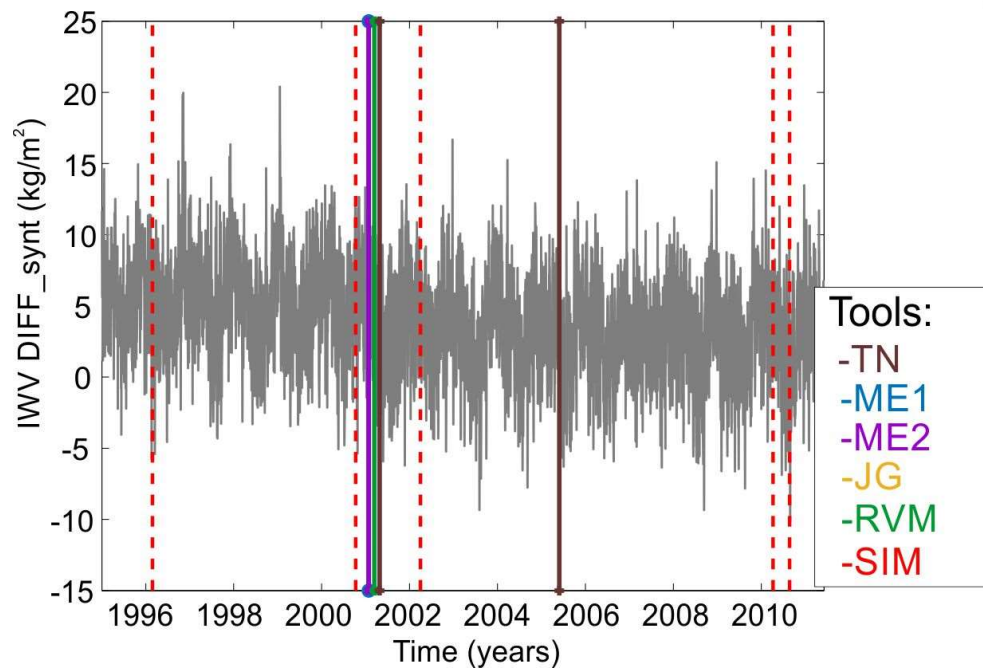**Motivation Data Synthetic benchmark Tools Results Summary**

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
Symbol	⬭⬭	△	✚	✕	◆	▽	-
Operator	M. Elias	R. Van Malderen	R. Van Malderen	J. Guijarro	T. Ning	S. Zengin	B.Chimani
Method / SW	2-sample t-test	2 of 3 (non-parametric tests)	PMW	CLIMATOL	PMTred	Pettitt test	HOMOP
Daily/Monthly	D+M	D+M	D+M	D+M	D+M	D	X
Easy/Less/Full	E+L+F	E+L+F	E+L+F	L+F	E+L+F	E+L+F	E+F

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

<u>How to classify breaks?</u>

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

Motivation   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
Symbol	⬭ ⬮	▲	✛	✕	◆	▼	-
Operator	M. Elias	R. Van Malderen	R. Van Malderen	J. Guijarro	T. Ning	S. Zengin	B.Chimani
Method / SW	2-sample t-test	2 of 3 (non-parametric tests)	PMW	CLIMATOL	PMTred	Pettitt test	HOMOP
Daily/Monthly	D+M	D+M	D+M	D+M	D+M	D	X
Easy/Less/Full	E+L+F	E+L+F	E+L+F	L+F	E+L+F	E+L+F	E+F

Amplitudes of reported offsets:

EASY, DAILY (SIM: 291):

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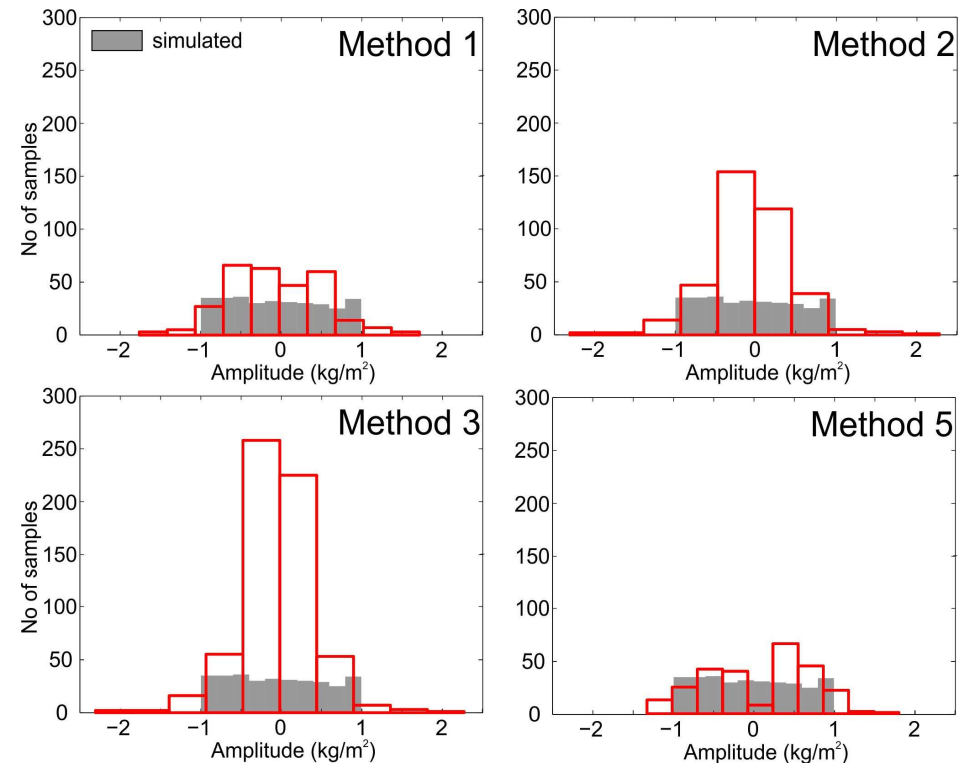- method 2:   252,
- method 3:   377,
- method 5:   216.

Motivation   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
Symbol	⬭ ⬭	▲	＋	✕	◆	▼	-
Operator	M. Elias	R. Van Malderen	R. Van Malderen	J. Guijarro	T. Ning	S. Zengin	B.Chimani
Method / SW	2-sample t-test	2 of 3 (non-parametric tests)	PMW	CLIMATOL	PMTred	Pettitt test	HOMOP
Daily/Monthly	D+M	D+M	D+M	D+M	D+M	D	X
Easy/Less/Full	E+L+F	E+L+F	E+L+F	L+F	E+L+F	E+L+F	E+F

Amplitudes of reported offsets:

FULLY-COMPLICATED, DAILY (SIM: 317):

- method 1:   295,
- method 2:   386,
- method 3:   622,
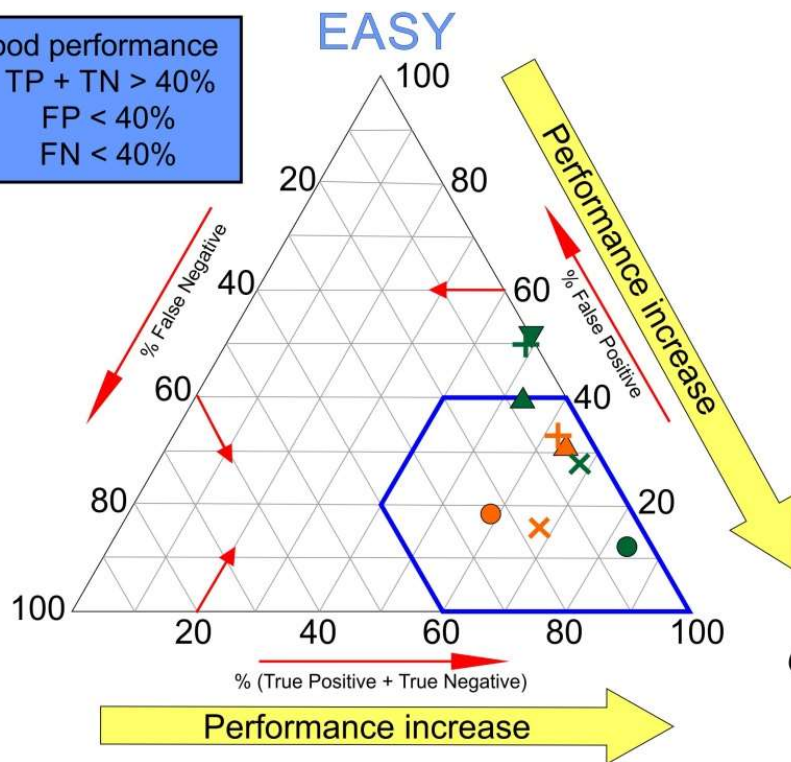- method 5:   264.

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

Tools perfor-mance:

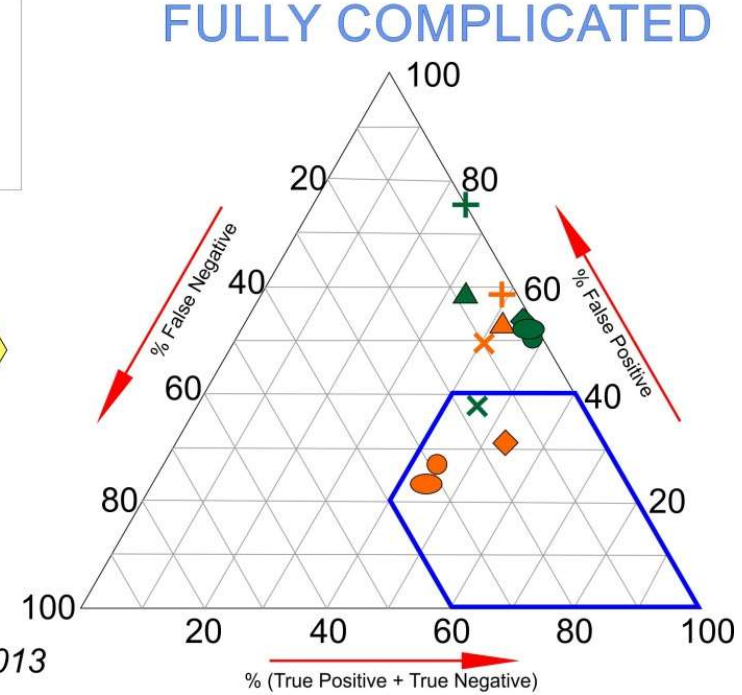|	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7
Symbol	● ⬭	▲	+	✕	◆	▼	-
Operator	M. Elias	R. Van Malderen	R. Van Malderen	J. Guijarro	T. Ning	S. Zengin	B.Chimani
Method / SW	2-sample t-test	2 of 3 (non-parametric tests)	PMW	CLIMATOL	PMTred	Pettitt test	HOMOP
Daily/Monthly	D+M	D+M	D+M	D+M	D+M	D	X
Easy/Less/Full	E+L+F	E+L+F	E+L+F	L+F	E+L+F	E+L+F	E+F

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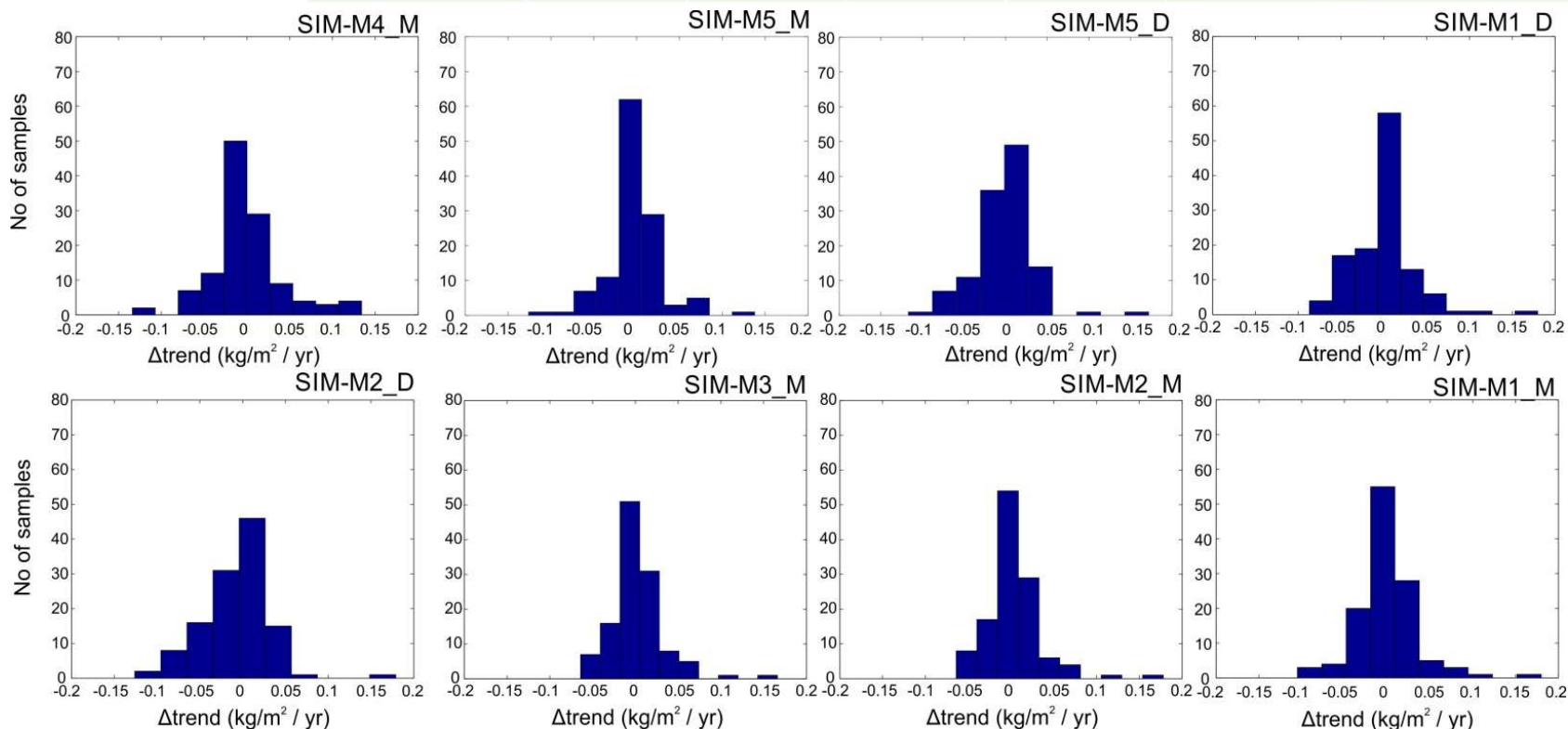

Observations:
daily
montly

+ AR(1) noise +gaps & trend

Adapted from Gazeaux et al., 2013

Motivation   Data   Synthetic benchmark   Tools   Results   Summary

Trends:

	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7
Symbol	● ⬭	▲	✛	✕	◆	▼	-
Operator	M. Elias	R. Van Malderen	R. Van Malderen	J. Guijarro	T. Ning	S. Zengin	B.Chimani
Method / SW	2-sample t-test	2 of 3 (non-parametric tests)	PMW	CLIMATOL	PMTred	Pettitt test	HOMOP
Daily/Monthly	D+M	D+M	D+M	D+M	D+M	D	X
Easy/Less/Full	E+L+F	E+L+F	E+L+F	L+F	E+L+F	E+L+F	E+F

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

<u>What's next?</u>

1. A detailed assessment of tools – their sensitivity.

2. Now – epochs were given to the participants to fine-tune their methods.

3. Still looking for other contributions. Interested? Contact Roeland: roeland@meteo.be, Eric: eric.pottiaux@observatoire.be or me: anna.klos@wat.edu.pl

4. A next generation of synthetic benchmark is ongoing.

5. New results of blind homogenization by the September/October.

6. Next homogenization workshop following two previous ones in Brussels (in 2016) and Warsaw (in 2017). If interested in partcipating – please, contact us ☺

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

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

Motivation   Data   Synthetic benchmark   Tools   Results   Summary

Acknowledgments

The map was drawn in the Generic Mapping Tool (Wessel et al., 2013).

Thank you!