

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

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

- 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:

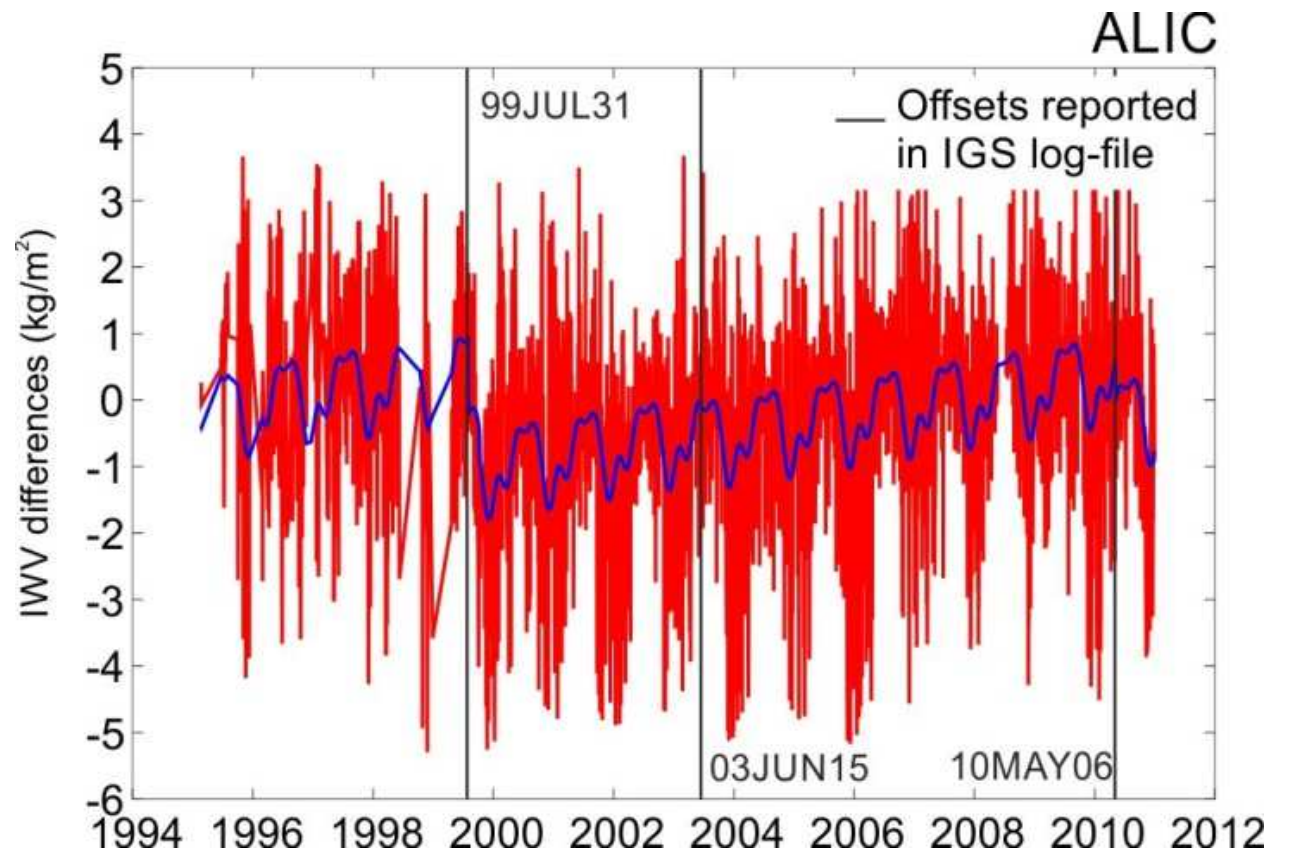
How does it look in practice?

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

**Can anyone see more offsets?**

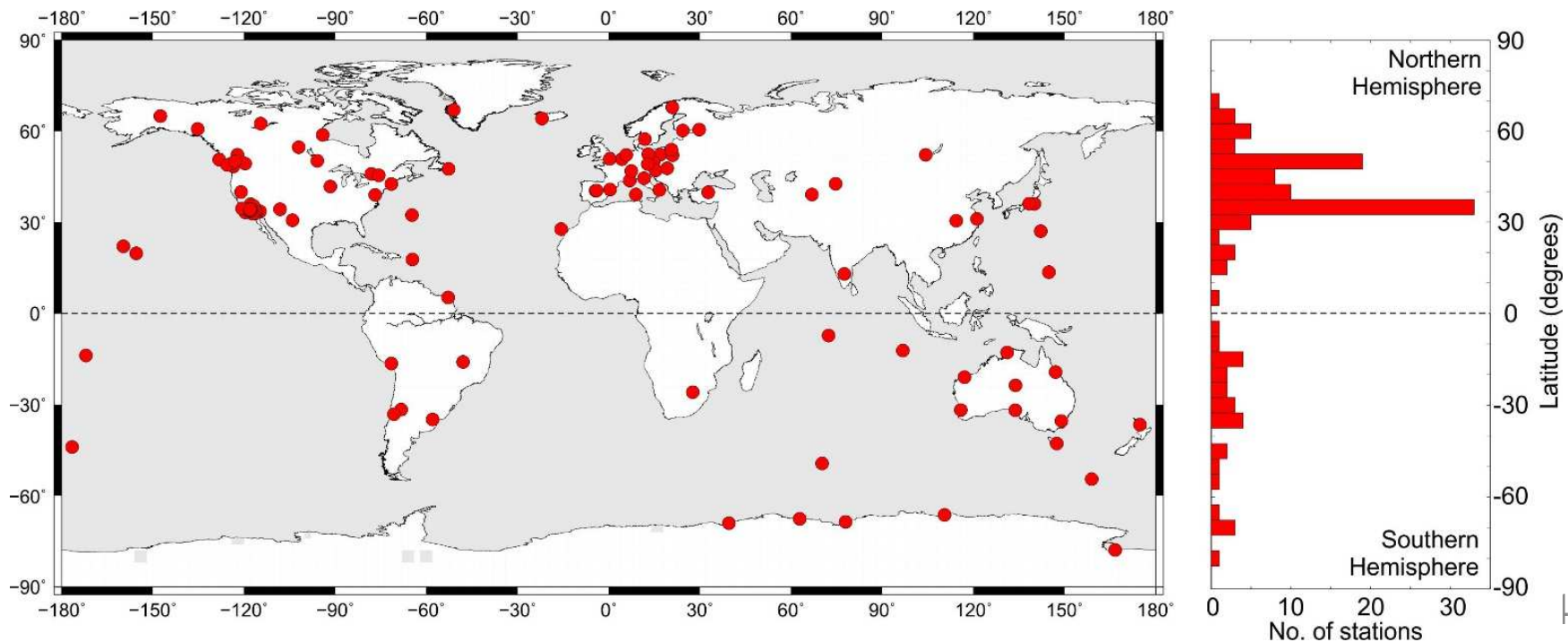
**What we aim at?**

Only real breakpoints not regime-like shifts should be corrected!



## Data:

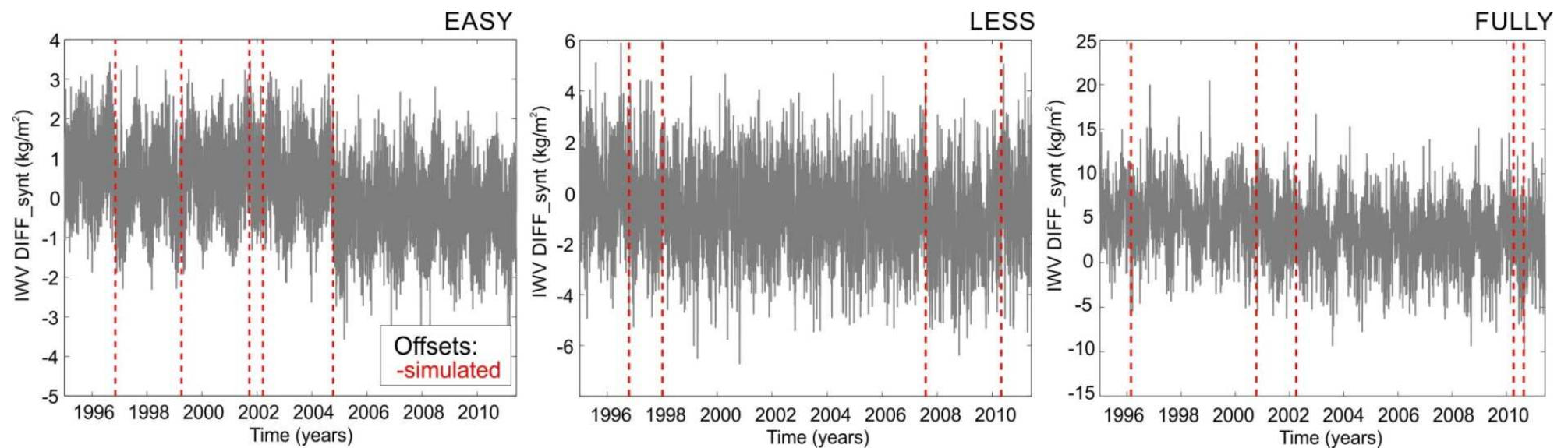
1. IGS „repro1” troposphere products screened and converted to Integrated Water Vapor (IWV) by O. Bock.
2. 120 stations, daily observations, a period of 1995-2010.
3. The IWV differences: ERAI-GPS were employed.



## Generation of the benchmark:

3 variants of synthetic time series were generated:

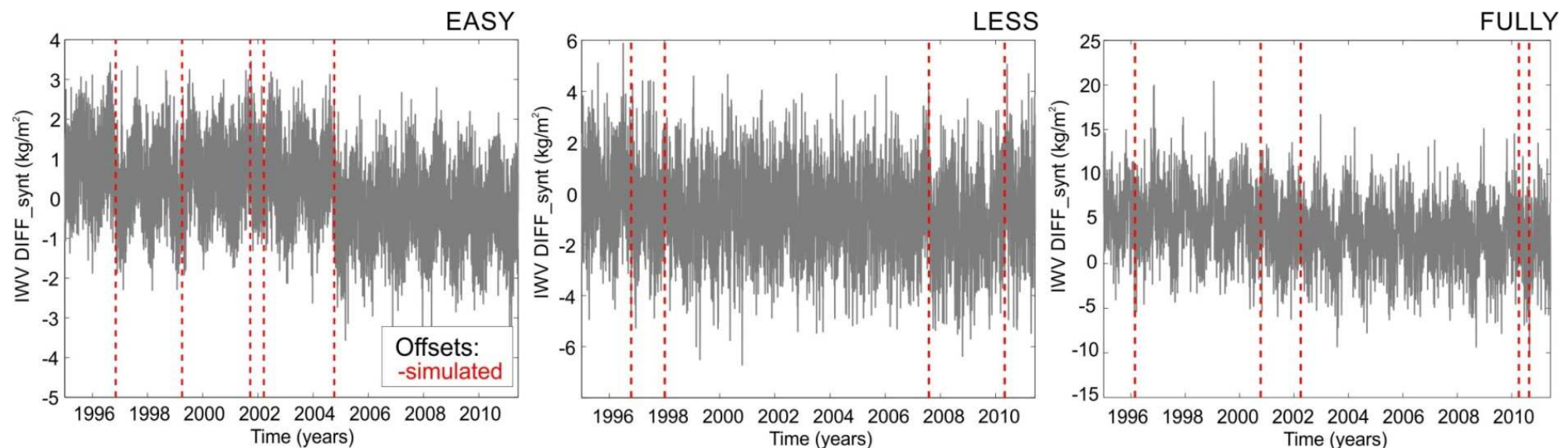
1. **EASY** dataset: **seasonal signals + offsets + white noise (WN)**,
2. **LESS COMPLICATED** dataset: same as 1. + autoregressive process of the first order (noise model = **AR(1)+WN**),
3. **FULLY COMPLICATED** dataset: same as 2. + **trend + gaps** (up to 20% of missing data).



## 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?**



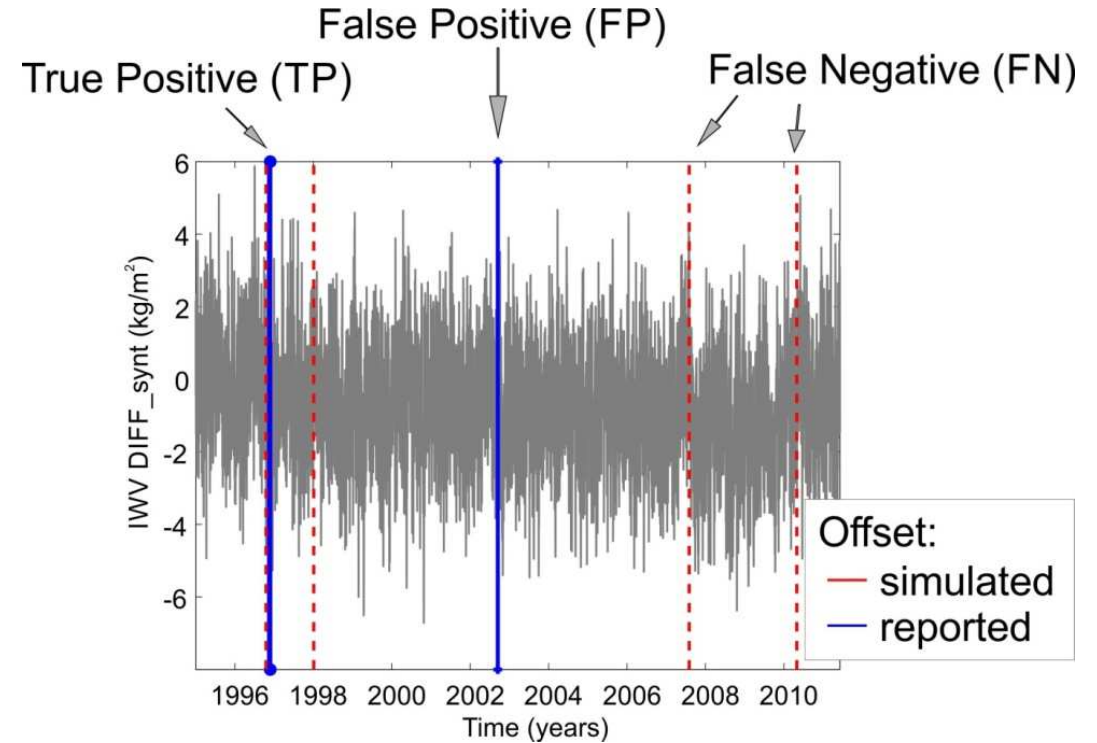
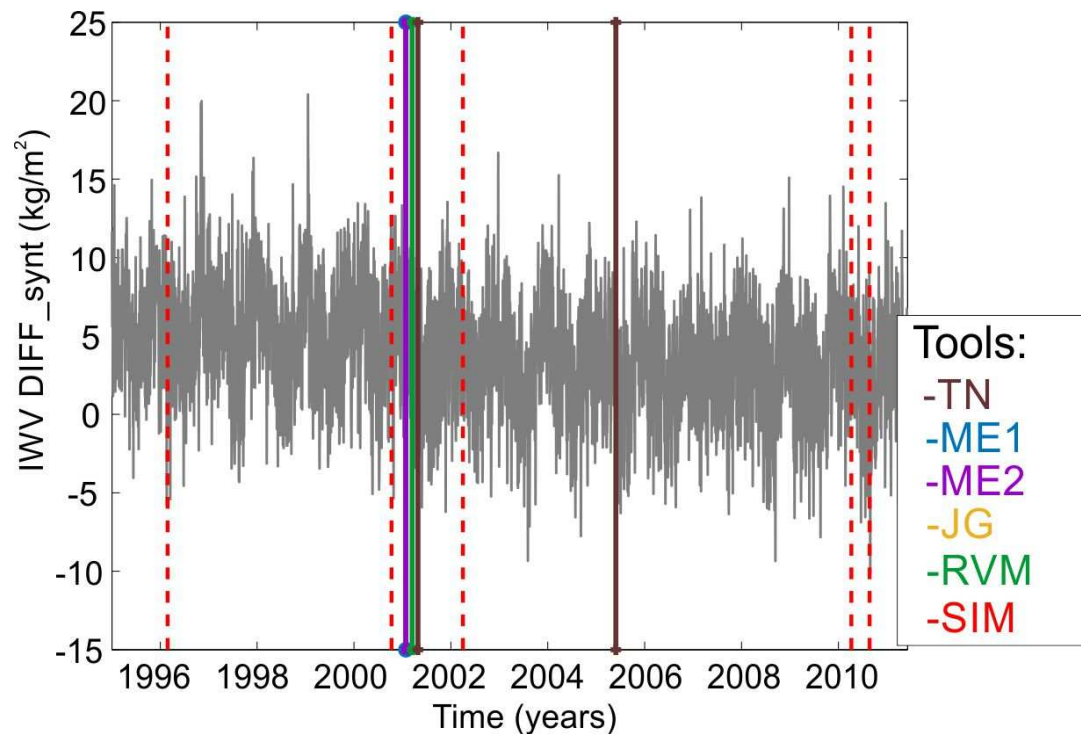
## Algorithms:

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

## How to classify breaks?

Defining a proper time window - 2 months





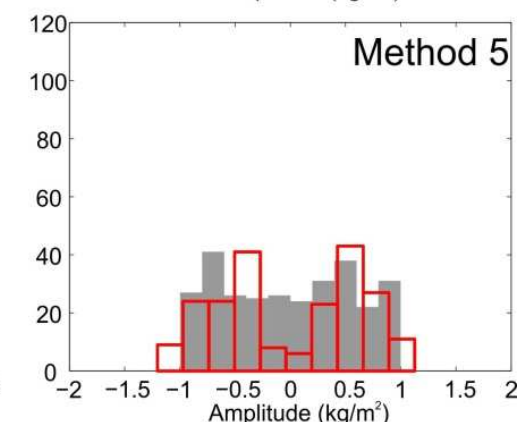
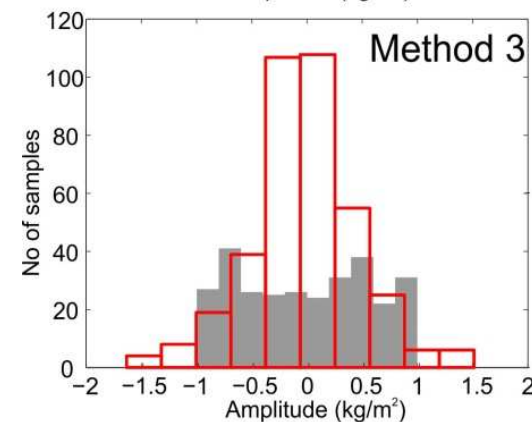
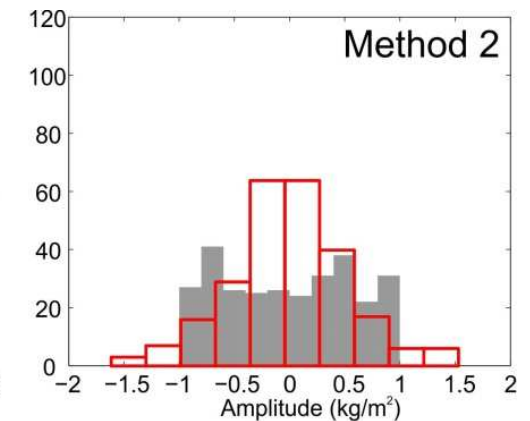
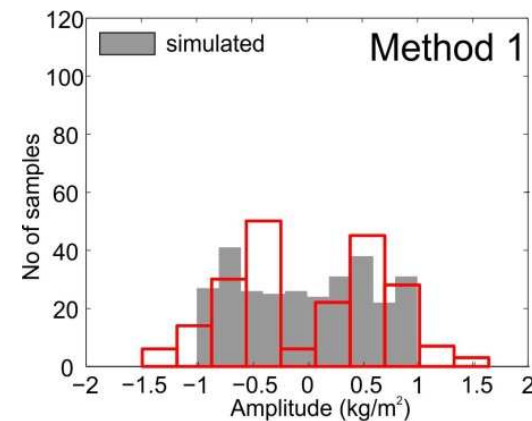
Offsets:

	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.



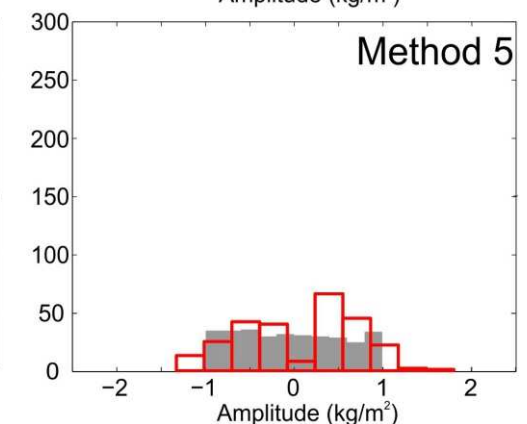
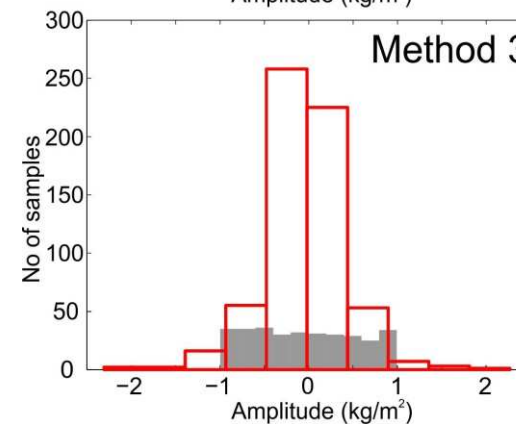
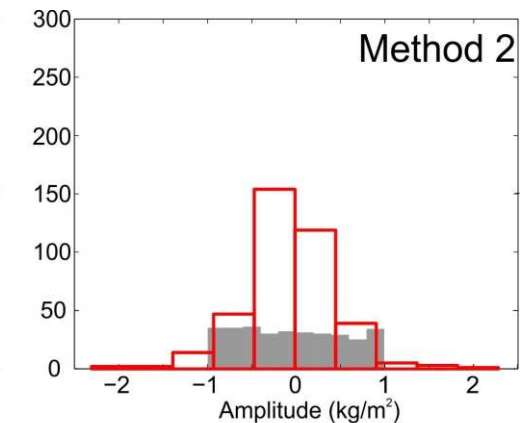
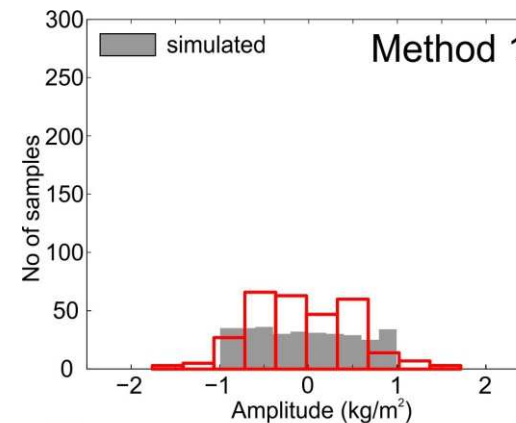
Offsets:

	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:

FULLY-COMPLICATED,  
DAILY (SIM: 317):

- method 1: 295,
- method 2: 386,
- method 3: 622,
- method 5: 264,
- method 8: 433.

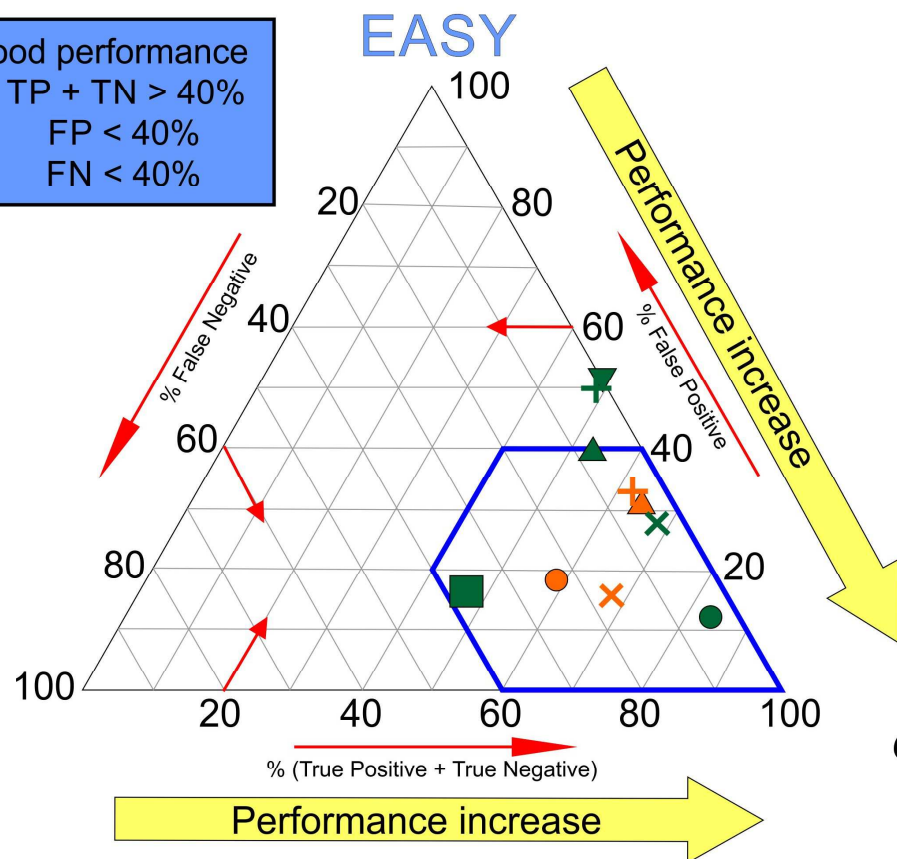


Introduction Data Synthetic benchmark Tools Results Summary

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  
 $TP + TN > 40\%$   
 $FP < 40\%$   
 $FN < 40\%$



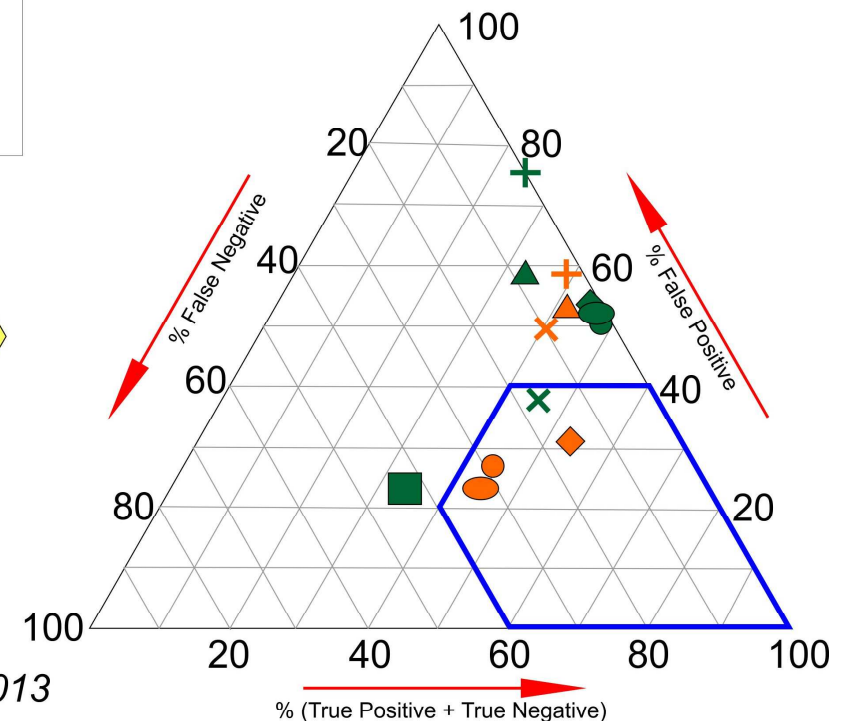
Observations:

■ daily  
 ■ montly

+ AR(1) noise  
 + gaps & trend

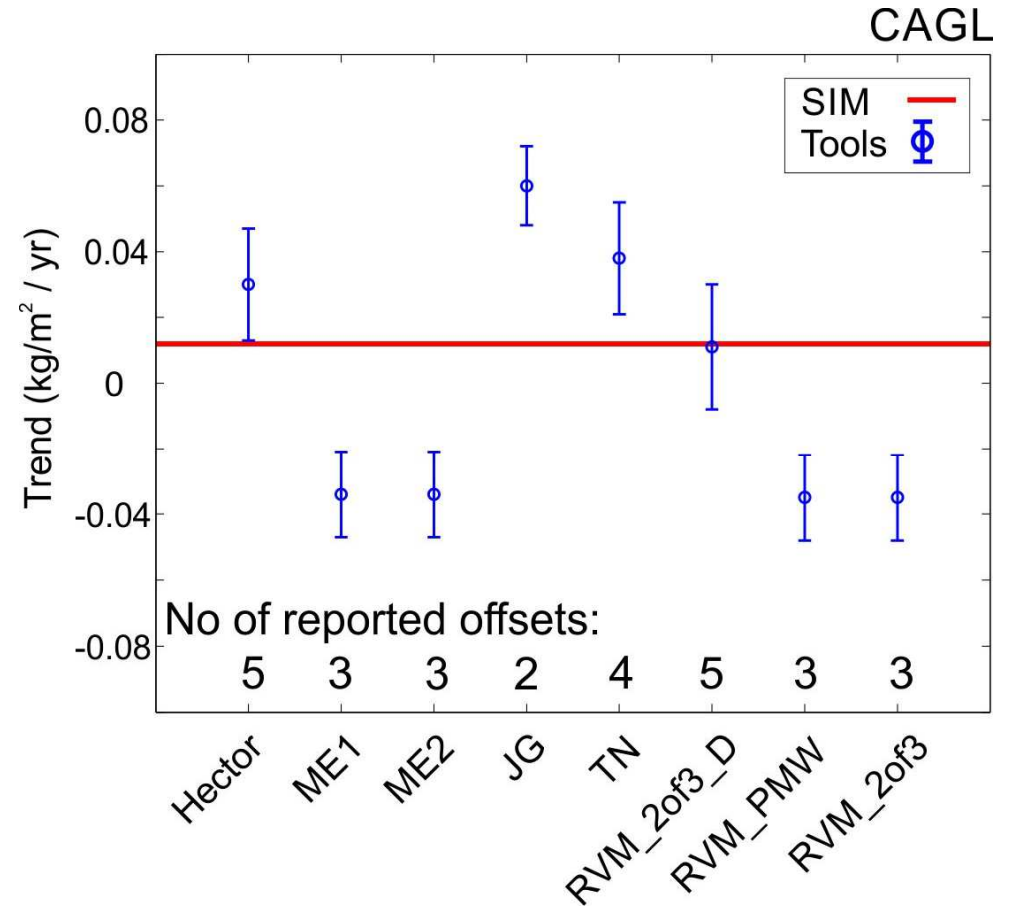
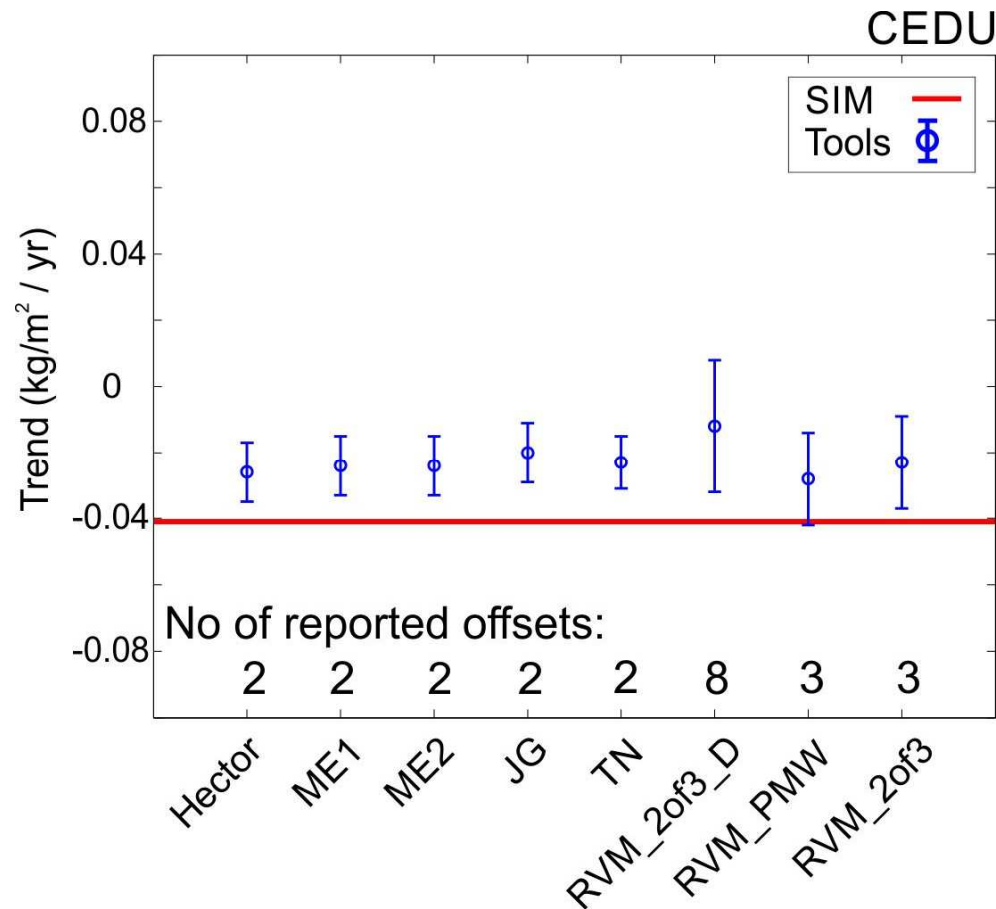
Adapted from  
 Gazeaux et al., 2013

**FULLY COMPLICATED**



## Changes in Trends:

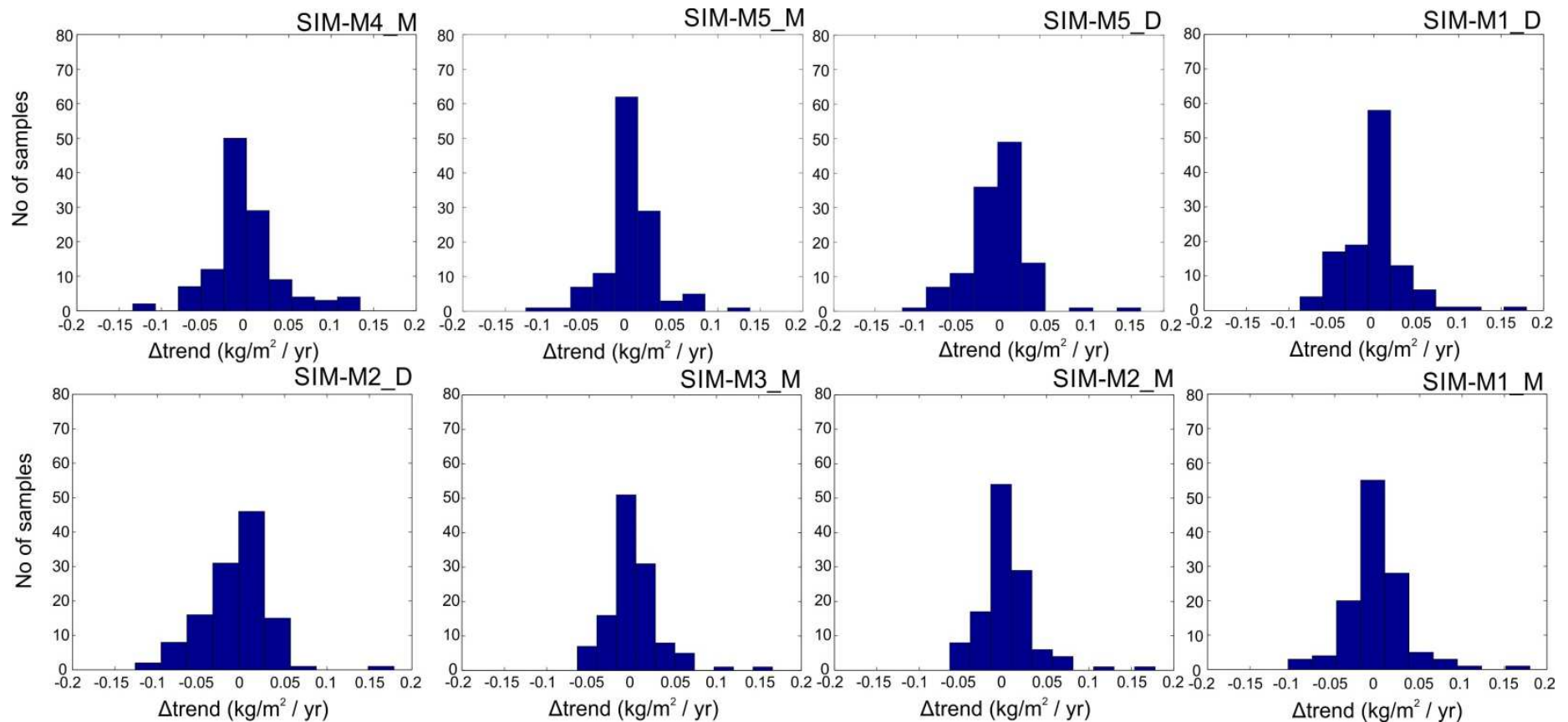
**FULLY-COMPLICATED**



Introduction Data Synthetic benchmark Tools Results 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



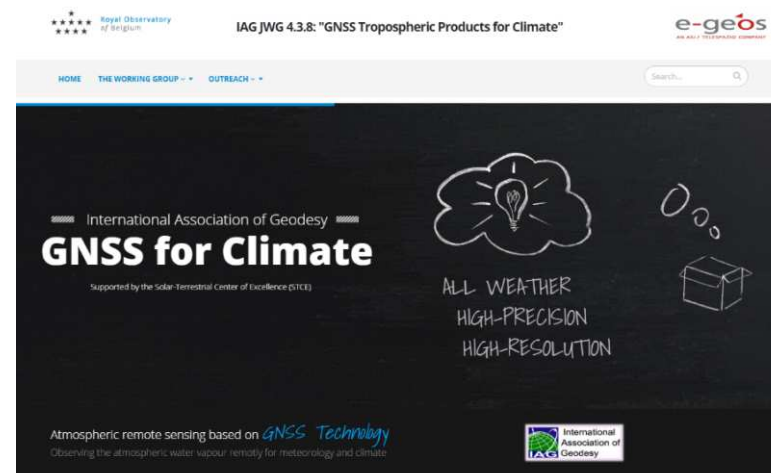
## Now:

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.

## 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>



### Author Words.

This website is dedicated to the research and developments carried out in the framework of the International Association of Geodesy (IAG) Joint Working Group (JWG) 4.3.8 "GNSS Tropospheric Products for Climate". This JWG belongs to the Sub-Commission 4.3: "Atmosphere Remote Sensing" under IAG Commission 4 "Positioning and applications" (2015 - 2019) and aims at exploiting Global Navigation Satellite Systems (GNSS) signals to support applications in Climate Sciences.

Activities developed in this JWG have strong links with the European COST Action ES1206 "Advanced Global Navigation Satellite Systems Tropospheric Products for Monitoring Severe Weather Events and Climate" (GNSS4WECC), the International GNSS Service (IGS) and its Troposphere Working Group, as well as with the EUREF and its associated EPN Troposphere Product.

### Latest News Posts

06 [Outreach]: New Oral Presentation.

Sep '18 Description: The IAG JWG 4.3.8 "GNSS Tropospheric Products for Climate" activities will be presented at the IAG Commission 4 Symposium in Wrocław, Poland.

## Acknowledgments

Anna Klos was supported by COST Action ES1206 GNSS4SWEC ([www.gnss4swec.knmi.nl](http://www.gnss4swec.knmi.nl)) during her stay at the Royal Observatory of Belgium (ROB) and the Royal Meteorological Institute of Belgium (RMI).

This research was supported by the Polish National Science Centre, grant No. UMO-2016/21/B/ST10/02353.

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

# Thank you!