Towards the Homogenization of GNSS Tropospheric Delay Time Series

STATUS AND RECENT DEVELOPMENTS

E. Pottiaux[1], *R. Van Malderen*[2], *A. Klos*[3], *F. Alshawaf*[4], *O. Bock*[5], *J. Bogusz*[3], *B. Chimani*[6], *P. Domonkos*, *M. Elias*[7], *M. Gruszczynska*[3], *J. Guijarro*[8], *T. Ning*[9], *V. Tornatore*[10], *S. Z. Kazanci*[11], *M. Hoseini*[12]

1 Royal Observatory of Belgium
2 RMI
3 WAT
4 GFZ Helmholtz-Zentrum Potsdam
5 IGN Institut National de l'Information Géographique et Forestière
6 ZAMG Zentralanstalt für Meteorologie und Geodynamik
7 VÚGTK
8 AEMet Agencia Estatal de Meteorología
9 Karadeniz Teknik Üniversitesi 1955
10 Lantmäteriet
11 Politecnico Milano 1863
12 NTNU Norwegian University of Science and Technology

Royal Observatory of Belgium – Solar-Terrestrial Centre of Excellence

Motivation and Introduction

ALL STARTED WITH…

Context and Motivation

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"

- It turned out that several groups were showing results from time series analyses, sometimes based on the same datasets.

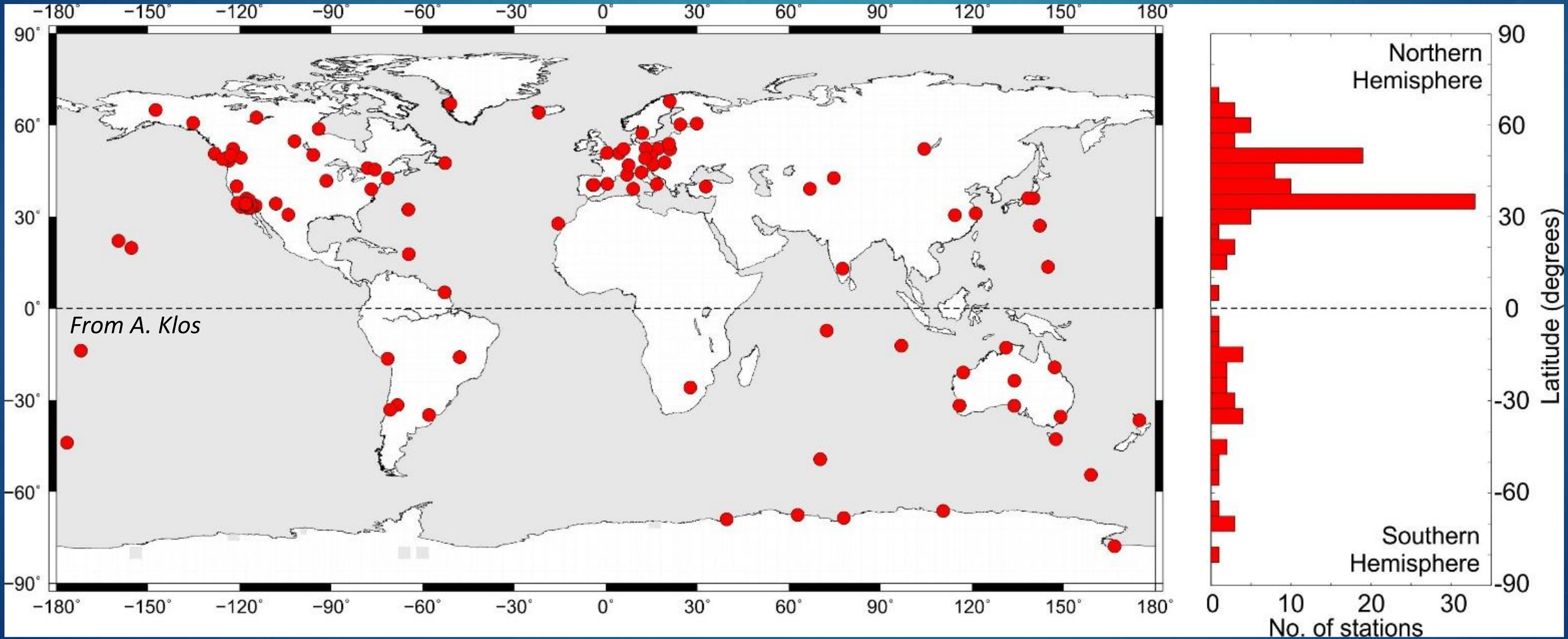
- They were dealing/struggling with the homogenization of their datasets.

Common Homogenization Activity

Common dataset: IGS "repro1" troposphere products

ZTD → IWV	Daily Observations	120 Stations	Period: 1995-2010

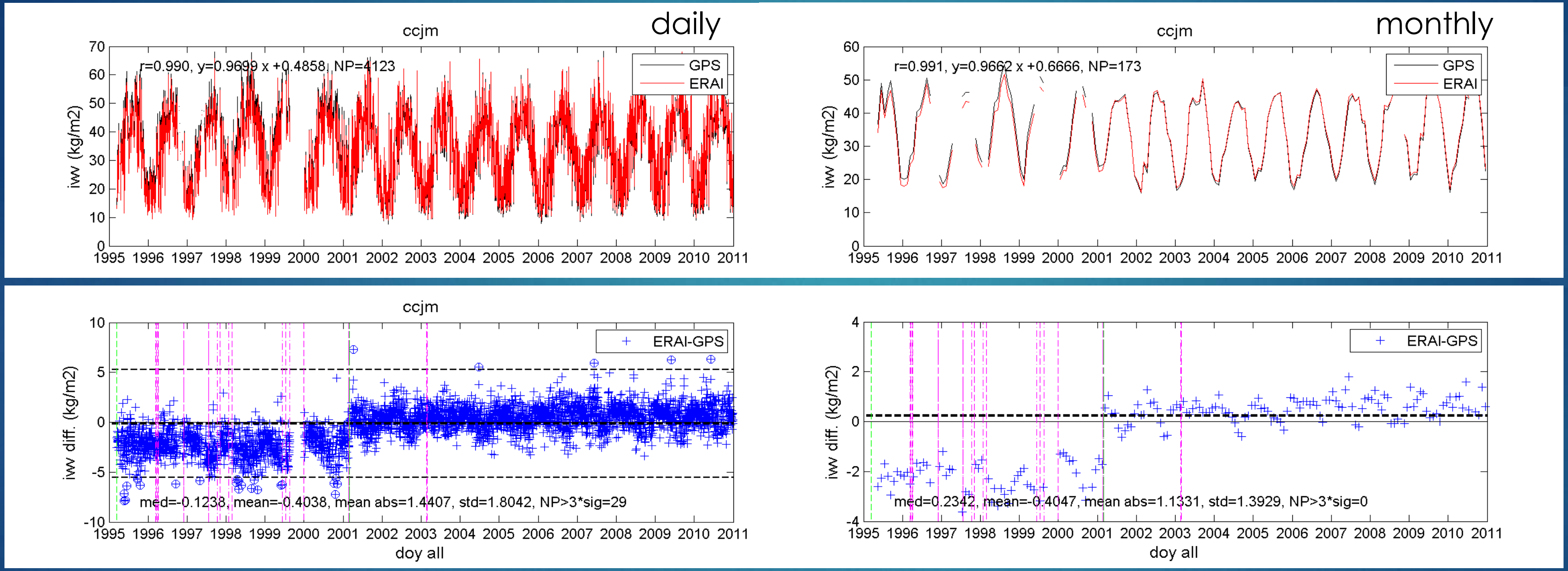
Screened and converted to Integrated Water Vapor (IWV) by O. Bock.



Dataset and Reference Series

Targeted dataset: IGS "repro1" troposphere products

CCJM: Ogasawara, Japan



From O. Bock

→ We will look for break points/change points in the ERA-interim-GPS IWV differences series.

Dataset and Reference Series

How to assess performance of the homogenization tools?

Generating Synthetic datasets

Blind Homogenization Benchmarking Activity

Real IWV Diff. (ERAI-GPS)		Synthetic IWV Diff.
Manual Homogenization IGS log files	→	Characterization of the number and amplitude of offsets (randomly inserted)
Power Spectra Density Analysis	→	Significant Frequencies (annual, semi-annual…)
Noise Analysis	→	Noise Model: AR(1) + W.N.
Non-climatic Trend Analysis	→	Characterization of non-climatic trends (reference series)

Synthetic Datasets Variants

Blind Homogenization Benchmarking Activity

We wanted to assess the performances of the homogenization tools w.r.t. dataset characteristics

EASY

- Seasonal signals
- Offsets
- White noise (WN)

LESS COMPLICATED

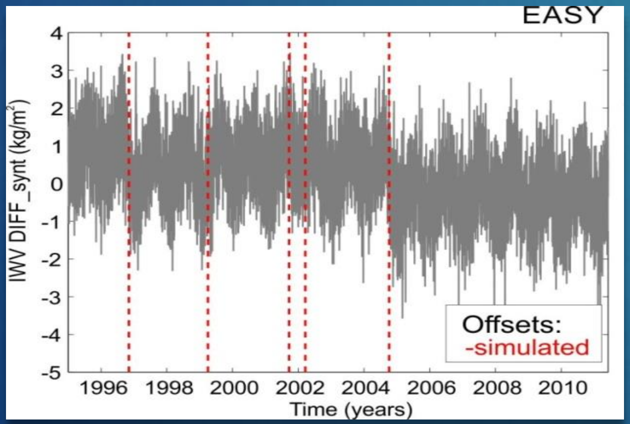
Similar to EASY but +

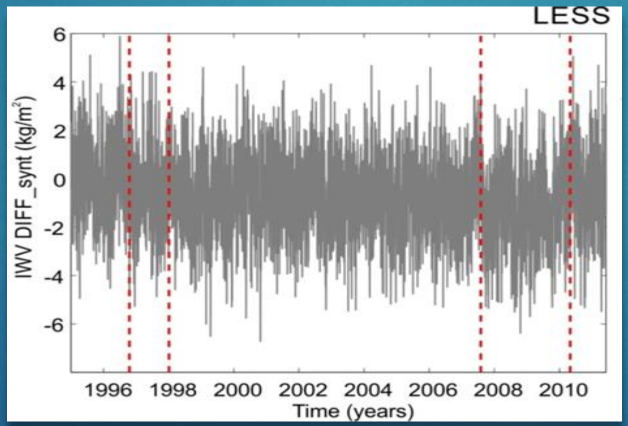
- Autoregressive process of the first order (noise model = AR(1)+WN)

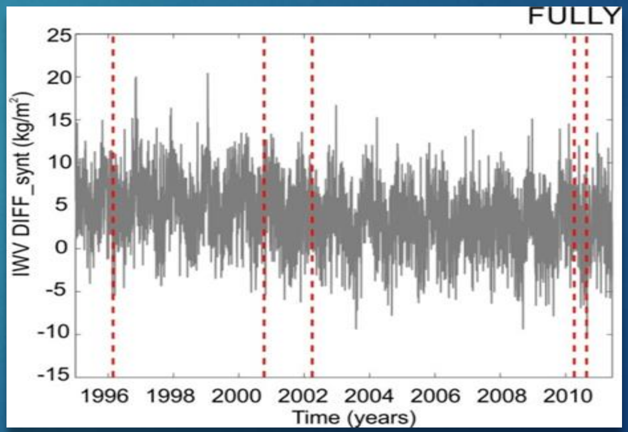
FULLY COMPLICATED

Similar to LESS but +

- Gaps (up to 20% of missing data)
- Non-climatic Trend (Ref. Series)







Participating Homogenization Tools

WG Contribution Summary

Climatol **J. Guijarro**	HOMOP **B. Chimani**	PMTred **T. Ning**
Non-parametric tests **R. Van Malderen**	2-sample t-test **M. Elias**	ACMANT **P. Domonkos**
IGN **O. Bock**	GFZ **F. Alshawaf**	

- 8 homogenization operators
- 13 break detection methods (daily+monthly)
- Applied on EASY/LESS/FULLY complicated synthetic datasets

4 main types of break detection methods:

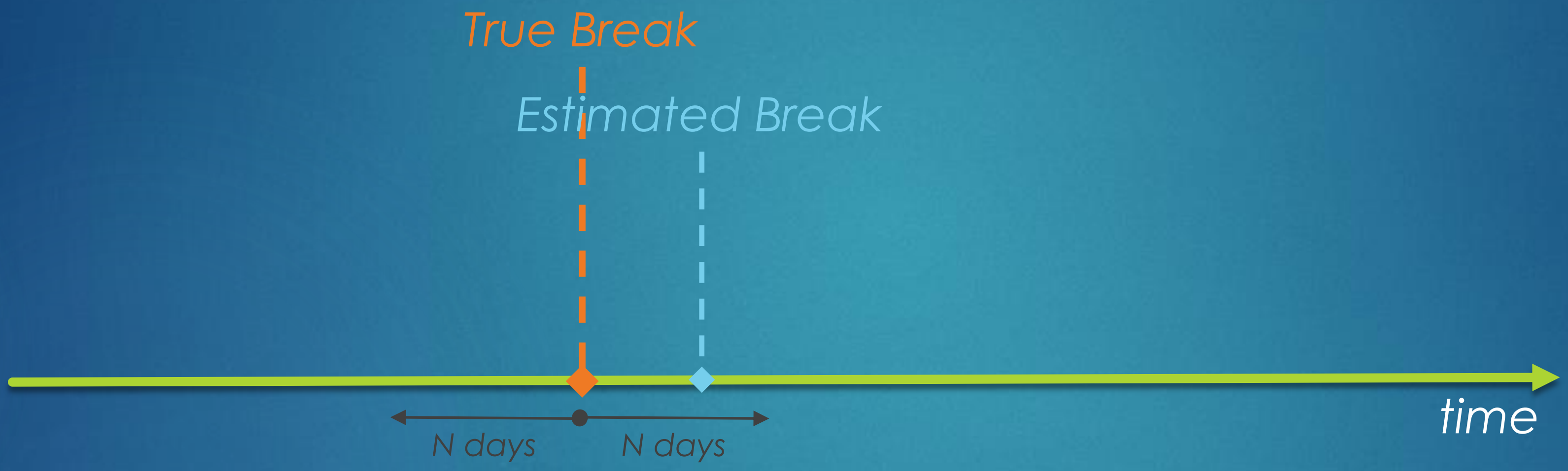
- *t-test with cutting algorithm*
- *Maximum Likelihood (ML) multiple break methods*
- Singular Spectrum Analysis (SSA)
- Non-parametric methods

Homogenization Tool Performance

ESTIMATED OFFSET CLASSIFICATION AND TIMING - SCORES AND SKILLS

Did I find a true break point?

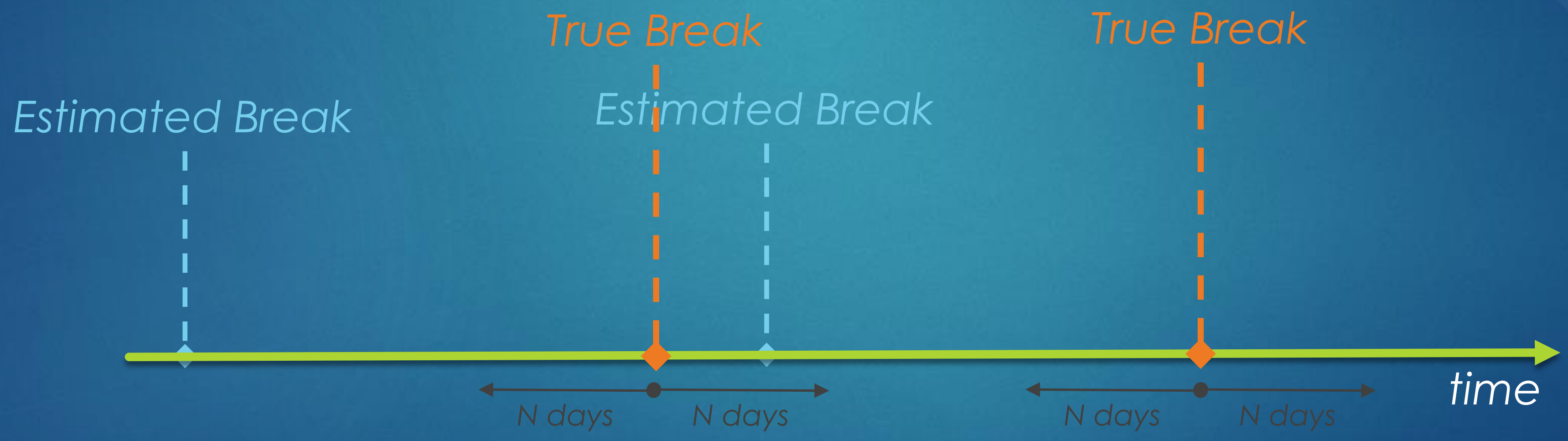
Assessing breakpoint identification requires defining a time window



We work with daily but also monthly time series → define a time window of **2 months**

Estimated Offsets Classification

False Alarms	Matches	Misses	
False Positive	True Positive	False Negative	True Negative



From Classification to Score and Skill Analysis

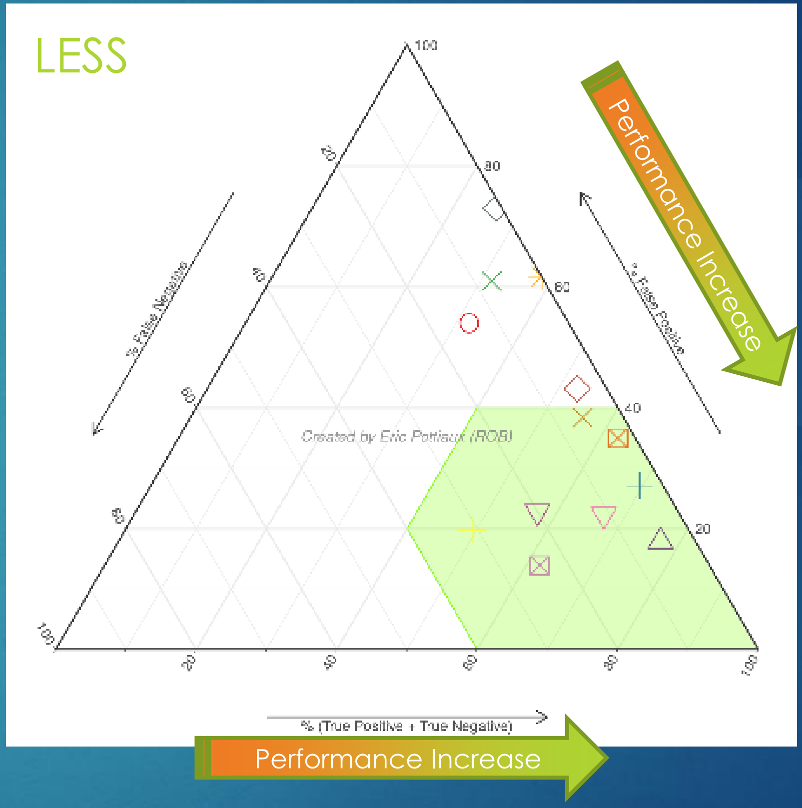
Green zone == Good performance == if ((TP + TN > 40%) && (FP < 40%) && (FN < 40%))

Ternary graph adapted from Gazeaux et al. 2013

Score and Skill Analysis

Performance Summary – Ternary Graphs

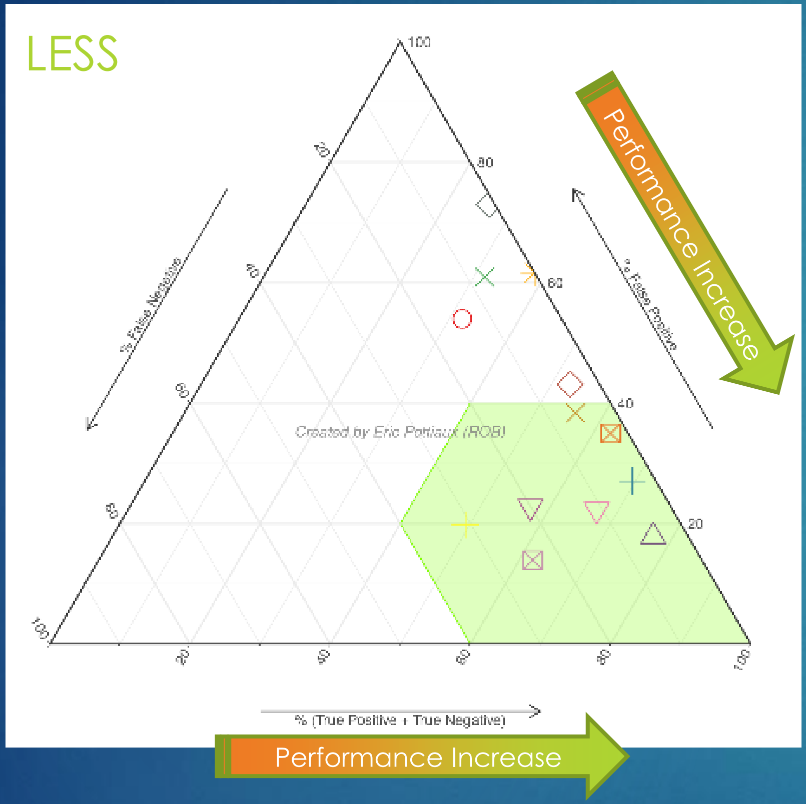
+ A.R. Noise



→ Good performance for the majority of the tools for the easy and less complicated dataset

Score and Skill Analysis

Performance Summary – Ternary Graphs



+ gaps and trend

→ Performance decreases drastically for almost all the tools when adding gaps and a trend in the benchmark time series

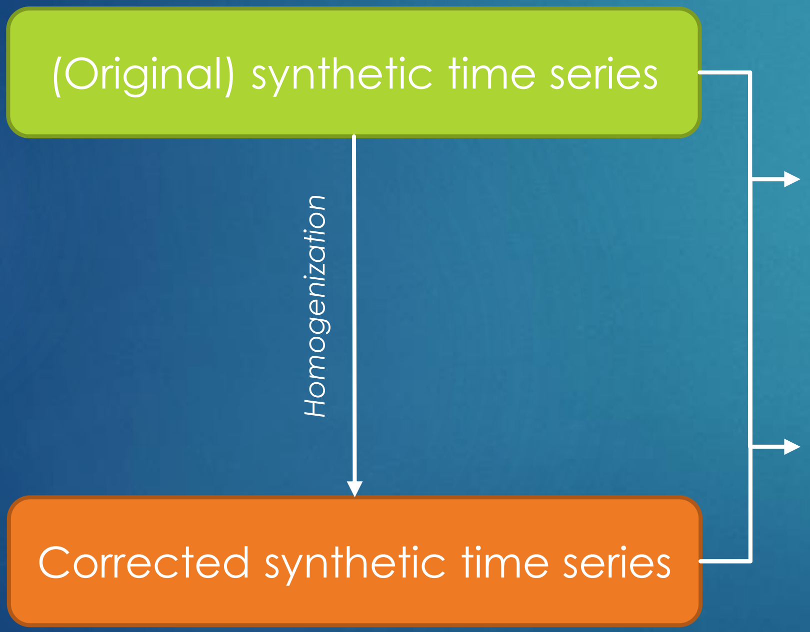
Homogenization Tool Performance

IMPACT ON THE TARGETED APPLICATION(S)

Impact on the targeted application(s)

CRMSE and Trend Analysis - Principle

- For each synthetic dataset, each homogenization tool contribution, and each time series we have

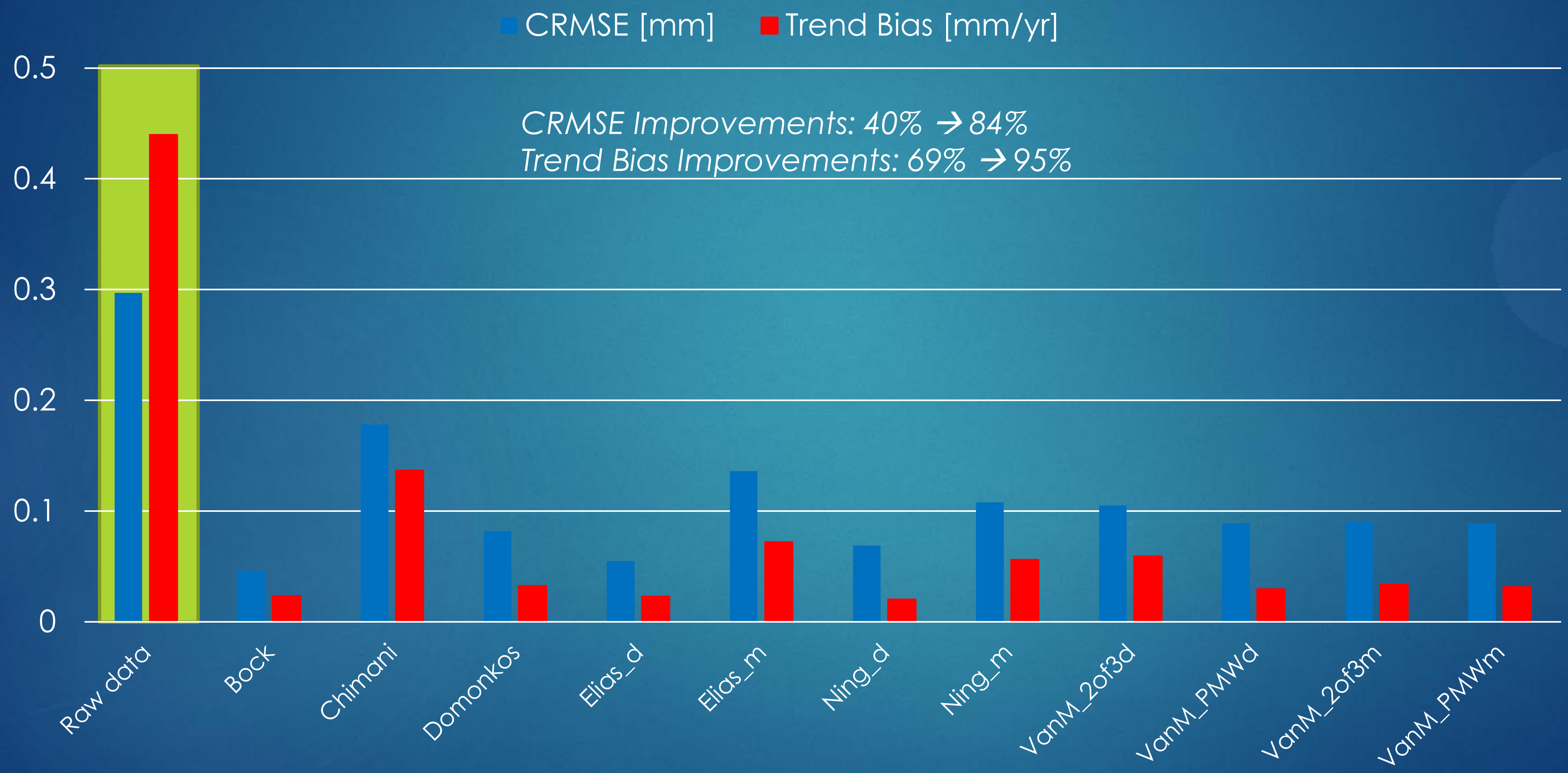


$$CRMSE \equiv \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\left(X_{i,orig} - \bar{X}_{orig}\right) - \left(X_{i,corr} - \bar{X}_{corr}\right)\right]^2}$$

$$\text{Abs. Trend Bias} \equiv abs\left(Trend_{orig} - Trend_{corr}\right)$$

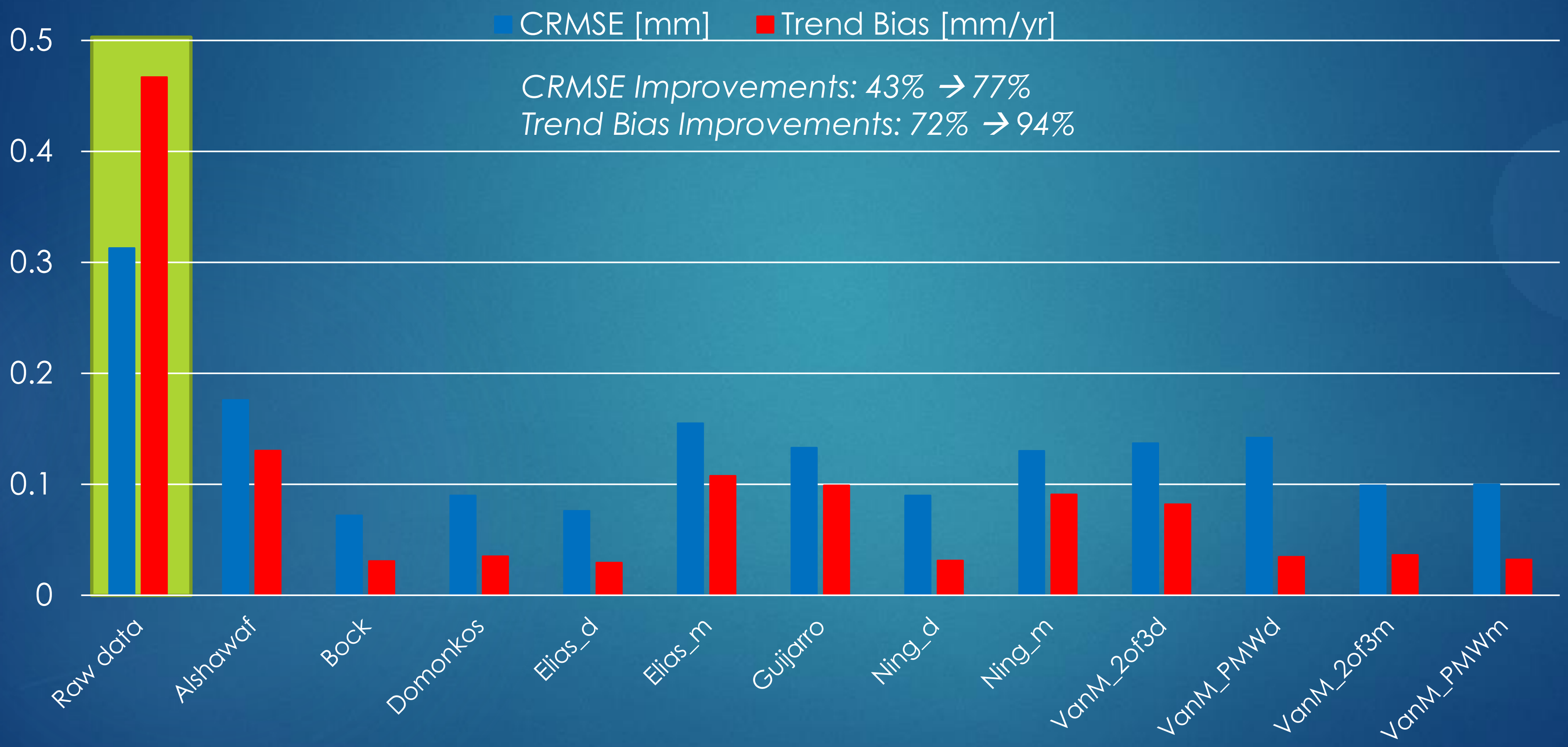
CRMSE and Trend Analysis

Synthetic Dataset "EASY" *(Arithmetic mean over all stations)*



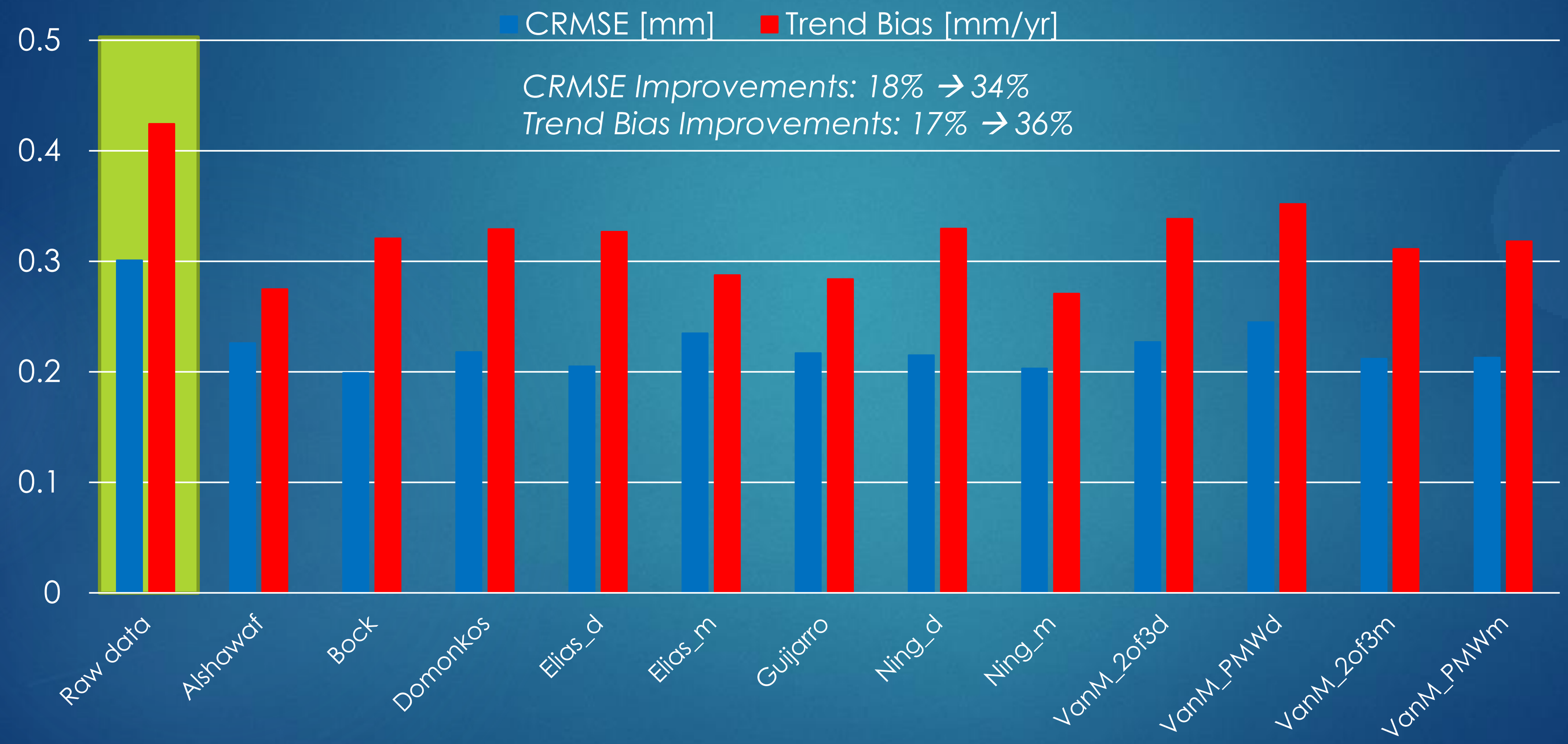
CRMSE and Trend Analysis

Synthetic Dataset "LESS COMPLICATED" *(Arithmetic mean over all stations)*



CRMSE and Trend Analysis

Synthetic Dataset "FULLY COMPLICATED" *(Arithmetic mean over all stations)*



First Conclusions and Future Steps

- EASY and LESS Complicated:
  - Most considered homogenization perform well in terms of scores and skill (timing of the offset), and show a large improvement in terms of CRMSE and trend bias (application side).
- FULLY complicated (+gaps and trends):
  - There is a drastic decrease in improvement, for all methods, with a large increase of false alarms (scores and skill , timing of the offset), and also a very reduced improvement in terms of CRMSE and trend bias. Reason is unclear (gap or trend) and must be further investigated.
  - The variation of performances within a single method increase when looking at individual time series.
- Next major steps?
  - Prepare next benchmark & blind homogenization test campaign ?
  - Determine a proper strategy for correcting the (real) IGS repro 1 dataset and apply it (and possibly to other datasets e.g. the EPN repro 2).

Thank you…