AUTOREGRESSIVE PROCESS IN HOMOGENIZATION OF GNSS TROPOSPHERIC DATA

Anna Klos[1], Olivier Bock[2], Eric Pottiaux[3], Roeland Van Malderen[3], Janusz Bogusz[1]

1) Military University of Technology, Faculty of Civil Engineering and Geodesy, Warsaw, Poland
2) IGN LAREG, University Paris Diderot, Sorbonne Paris, France
3) Royal Observatory of Belgium, Brussels, Belgium

AGU FALL MEETING
San Francisco | 12–16 December 2016

MAIN OBJECTIVES

1. To analyse the Integrated Water Vapour (IWV) time series derived from IGS „repro1" re-processing to end with estimates of trend, seasonal part and a character of noise.
2. To use these parameters to create a synthetic benchmark dataset to be homogenized with different statistical tools.
3. To decide on the most appropriate tool to be used for homogenization of tropospheric data.

DATASET

Data: 1-day sampled differences of Integrated Water Vapour (IWV) from GPS and ERA-Interim for period from January 1st, 1995 to December 31st, 2010 (Figure 1).
Author: Olivier Bock.
Brief description of processing and post-processing:
- GPS Zenith Total Delay (ZTD) source: for 1995-2007: IGS repro1 (reprocessed by JPL with GIPSY OASIS software in May 2010), for 2008-2010: IGS tropo_new (reprocessed by JPL with GIPSY OASIS software in July 2010);
- GPS ZTD screening based on range check and outliers check of ZTD and formal errors at 5 min resolution;
- GPS ZTD resampling to 1 hour interval (mean of all values in the interval t-0.5h to t+0.5h, with at least 4 values);
- GPS ZTD to IWV conversion: ZHD & Tm computed from ERA-Interim pressure level data (with bi-linear horizontal interpolation from 4 grid points);
- IWV differences computed when time-match is within +/-1 hour;
- IWV differences screened with normality test;
- IWV differences resampled to daily mean values when all 6-hourly values are available (mean computed with weights [0.5 1 1 1 0.5] for values at 00, 06, 12, 18, 24UTC).

METHODOLOGY

1. We manually homogenized IWV differences from IGS „repro1" solution.
2. We employed Maximum Likelihood Estimation (MLE) to derive parameters of trend, seasonal part and noise character.
3. We used spectral analysis to clearly demonstrate the differences between different noise models.
4. We used station-by-station parameters to create three synthetic benchmark datasets: EASY, LESS COMPLICATED and FULLY COMPLICATED to be tested with different tools.

CHARACTERISTICS OF IWV DIFFERENCES

Homogenisation of IWV differences:
- 1028 epochs reported in IGS log files (ftp://igs.ign.fr/pub/igs/igscb/station/log/ and ftp://igs.ign.fr/pub/igs/igscb/station/oldlog/) arising from changes in receiver, antenna or radome. We averaged them using 30-day window and ended with 970 epochs. These were validated manually: 174 offsets remained;
- 54 offsets reported manually in differences of IWV (ERAI-GPS) (Figure 2a);
- few stations included in a special class: cases where visual detection of offsets is not reliable (GOPE: Czech Republic, KOUR: French Guiana, LONG: USA and MCM4: Antarctica) (Figure 2b).

A **model fitted** to IWV differences:

$$IWV(t_i)=a+b\cdot(t_i-t_0)+c\cdot\sin(2\pi\cdot(t_i-t_0))+d\cdot\cos(2\pi\cdot(t_i-t_0))+$$
$$+e\cdot\sin(4\pi\cdot(t_i-t_0))+f\cdot\cos(4\pi\cdot(t_i-t_0))+g\cdot\sin(6\pi\cdot(t_i-t_0))+h\cdot\cos(6\pi\cdot(t_i-t_0))+$$
$$+i\cdot\sin(8\pi\cdot(t_i-t_0))+j\cdot\cos(8\pi\cdot(t_i-t_0))+\sum_{j=1}^{n}d_jH+\varepsilon_{IWV_i}$$

We examined on amplitudes of seasonal changes c, d: annual + e, f: semi-annual + g, h: 3 months + I, j: 4 months), values of trend: b and character of stochastic part ε_{IWV}: AR(1)+WN with Maximum Likelihood Estimation (Figure 5) in Hector Software (Bos et al., 2013). The amplitudes of offsets are estimated along with above mentioned parameters in a least-squares procedure.

Autoregressive of first order plus white noise (AR(1)+WN) chosen as a most appropriate for daily IWV data, basing on AIC criterion plus time of computation:

$$\varepsilon_{IWV_i}=\phi_1\varepsilon_{IWV_{i-1}}+a_i$$

GENERATING OF THE BENCHMARK DATASET

We simulated three types of data. The first, referred to as EASY, includes pure white noise, seasonal terms, trend and offsets. The second, referred to as LESS COMPLICATED, includes all above mentioned characteristics plus an autoregressive process of the first order. Finally, the FULLY COMPLICATED dataset includes all the above mentioned characteristics from "less complicated" plus gaps. All these 3 datasets provide a nice amount of series (360 series) to be tested with different homogenization tools. The homogenization group can examine on the effectiveness of different approaches when white noise prevails in stochastic part and when autoregressive behavior is added to a series. Also, they can study the impact of the gaps on their approach.

THIS IS A BLIND TEST: EPOCHS OF OFFSETS ARE NOT KNOWN...

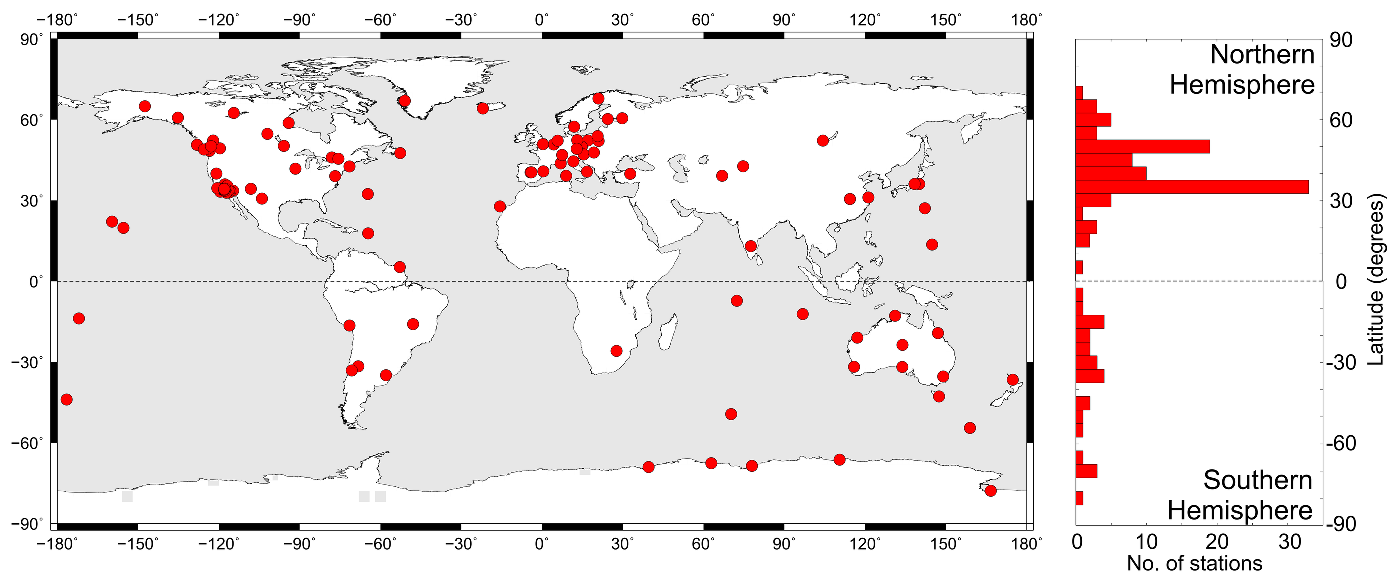


Figure 1. Stations used in a WG3 sub-working group "Data Homogenisation" activity.

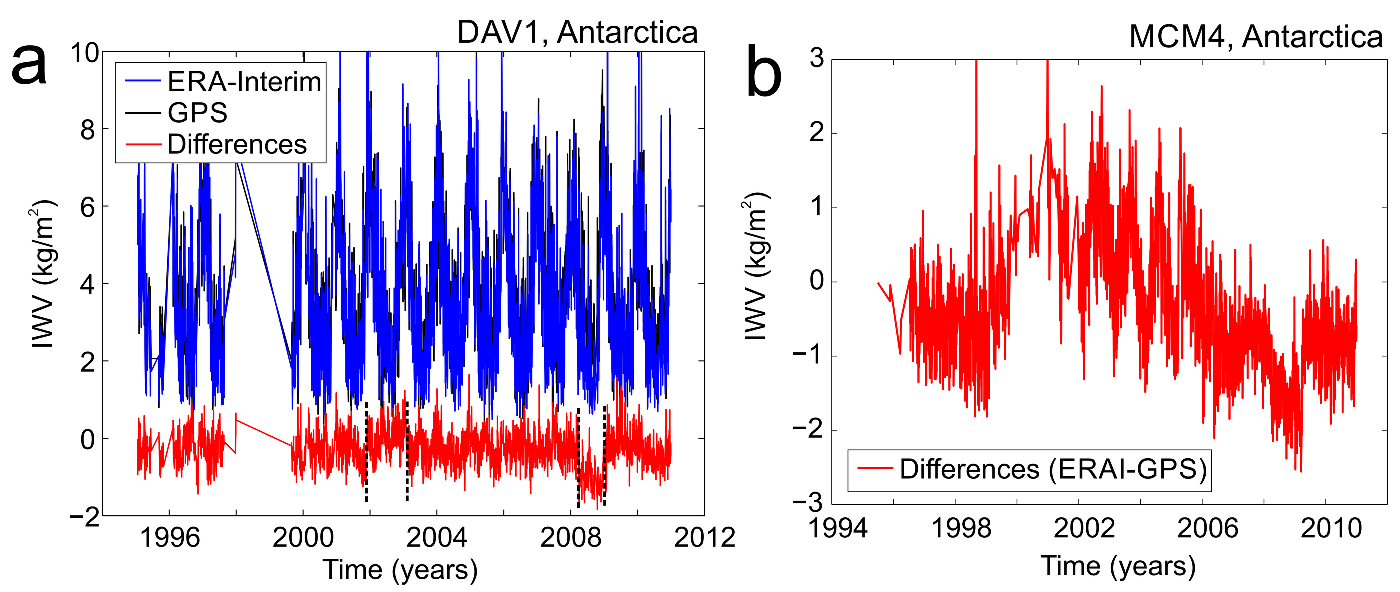


Figure 2. A manual homogenization of IWV differences. 54 epochs in total were discovered manually in differences of IWV (ERAI-GPS). An example of 4 offsets reported in DAV1 station (a). Time series from MCM4 station included to a special class (b).

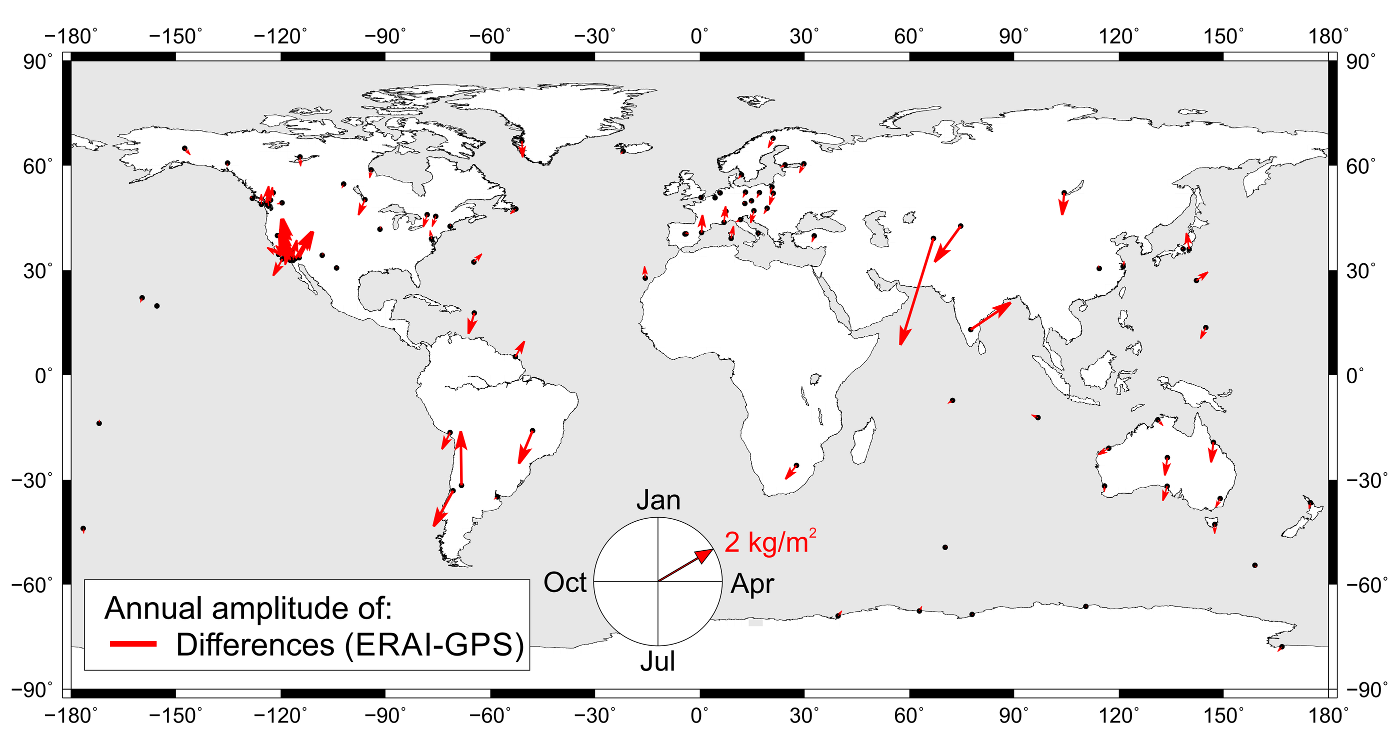


Figure 3. Annual amplitudes derived from IWV differences between ERA-Interim model and GPS.

ACKNOWLEDGMENTS

Anna Klos and Janusz Bogusz are supported by the Faculty of Civil Engineering and Geodesy statutory research funds.
Part of this work has been supported by COST Action ES1206 GNSS4SWEC

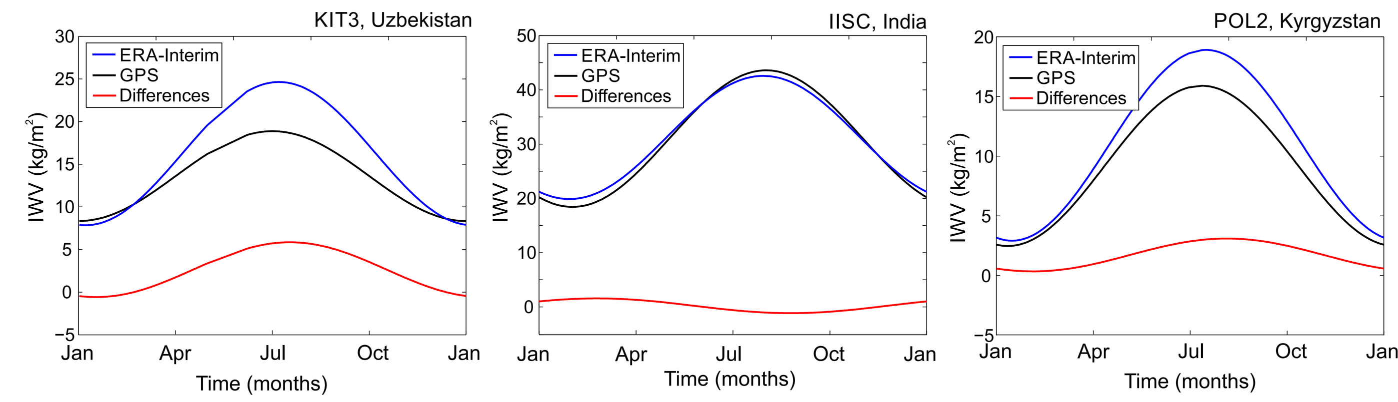


Figure 4. The largest differences between seasonal oscillations derived from IWV series retrieved by ERA-Interim model and GPS.

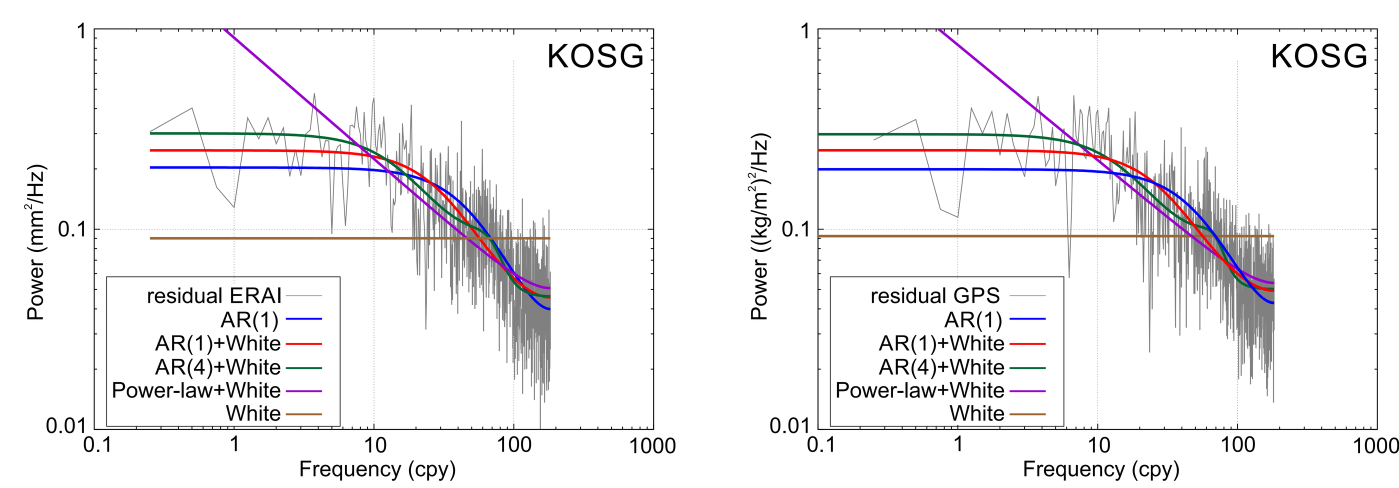
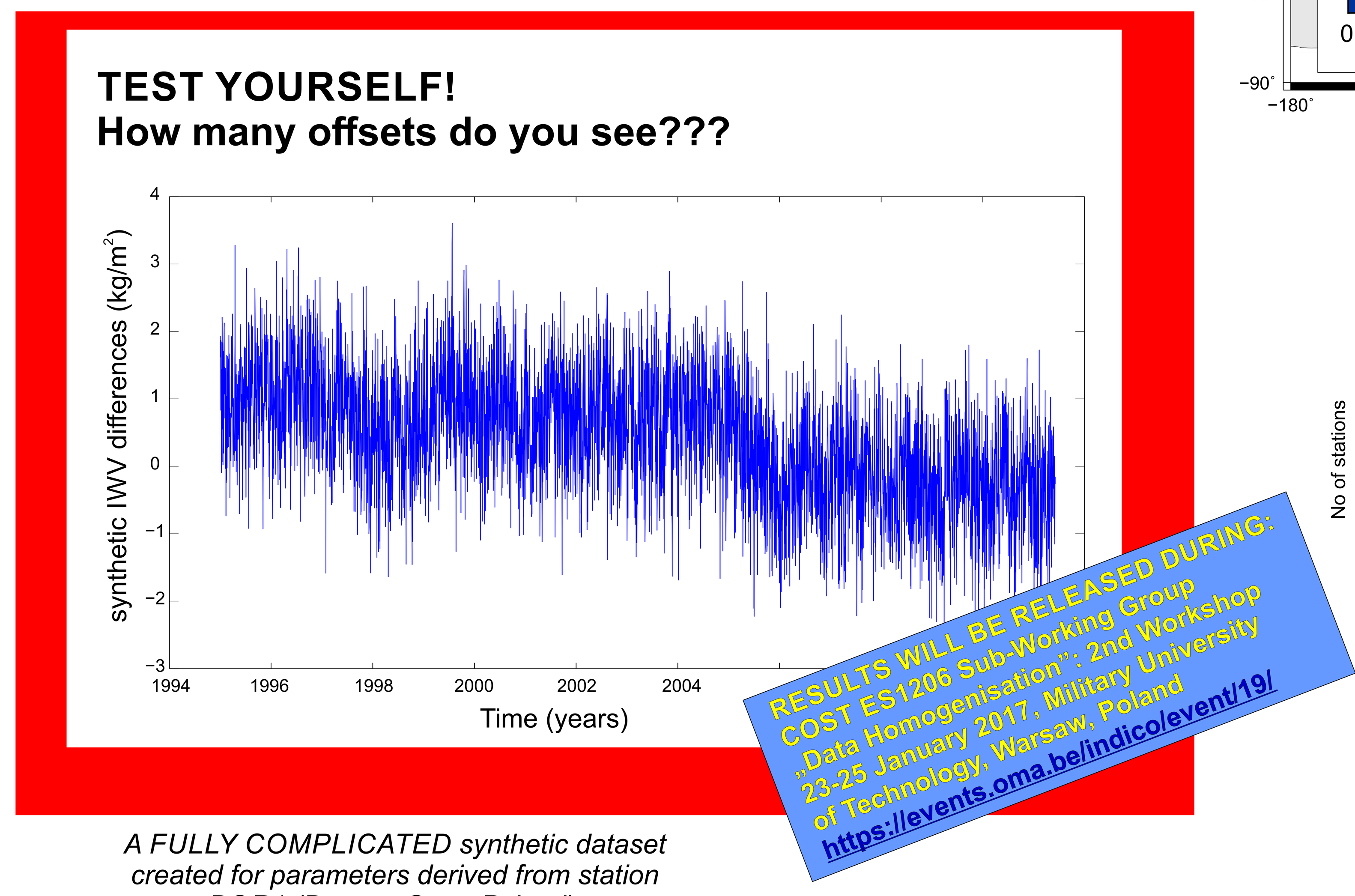


Figure 5. Tests on an optimal character for stochastic part of IWV differences. Autoregressive of first order plus white noise (AR(1)+WN) was chosen for daily differences of IWV data as a compromise between Bayesian Information Criterion (BIC) (or a goodness of fit) and a time of computation. This one is also the best fit for both ERAI and GPS-retrieved IWV (paper by Klos et al., in review).

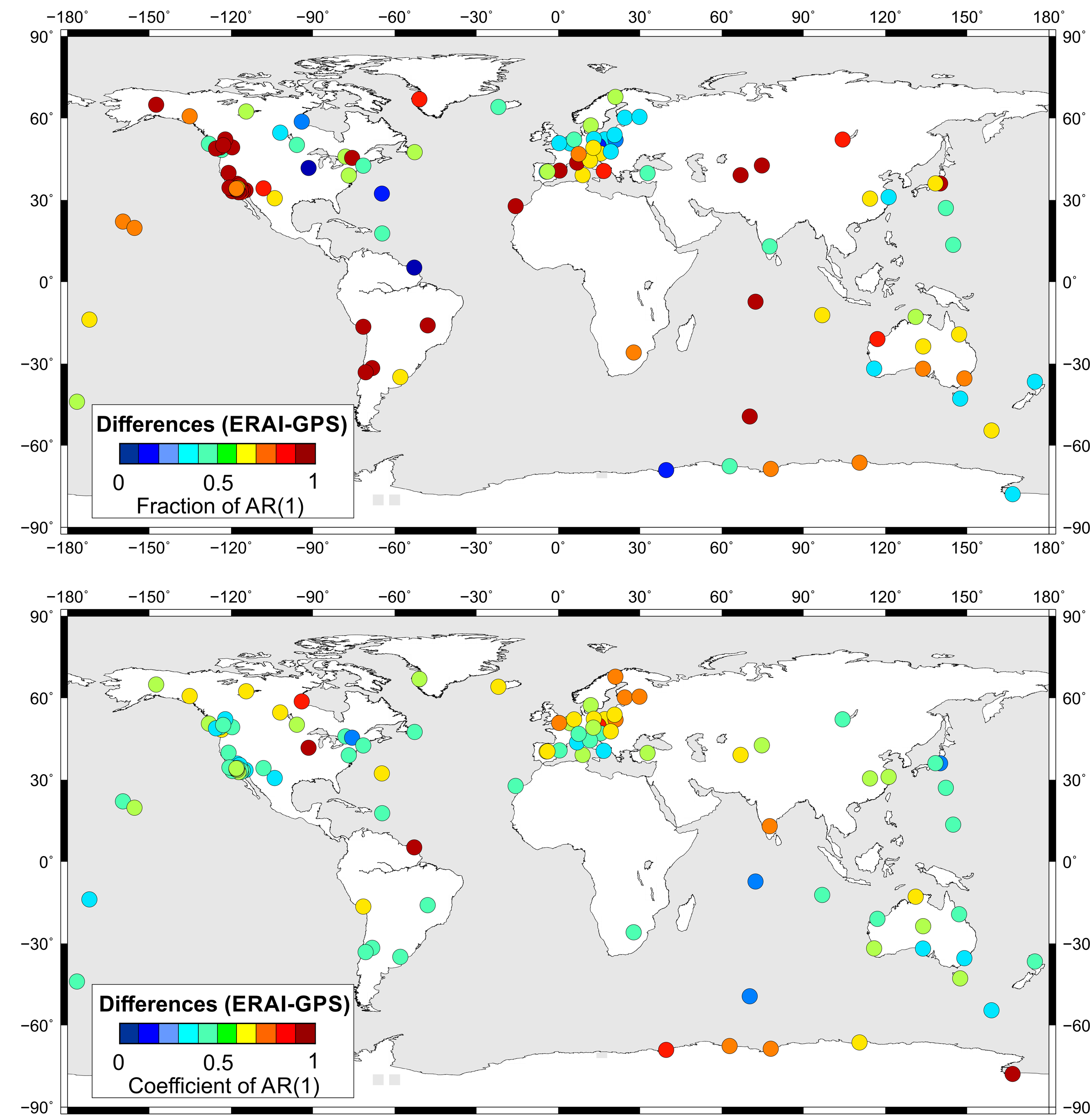


A FULLY COMPLICATED synthetic dataset created for parameters derived from station BOR1 (Borowa Gora, Poland).

REFERENCES

Bos, M.S., Fernandes, R.M.S., Williams, S.D.P., and Bastos, L.: 2013, Fast Error Analysis of Continuous GNSS Observations with Missing Data. J. Geod., 87, No. 4, 351-360. DOI:10.1007/s00190-012-0605-0.

Klos A., Hunegnaw A., Teferle F.N., Abraha K.E., Ahmed F., Bogusz J.: "Noise characteristics in Zenith Total Delay from homogeneously reprocessed GPS time series". Submitted to the Atmospheric Measurement Techniques.



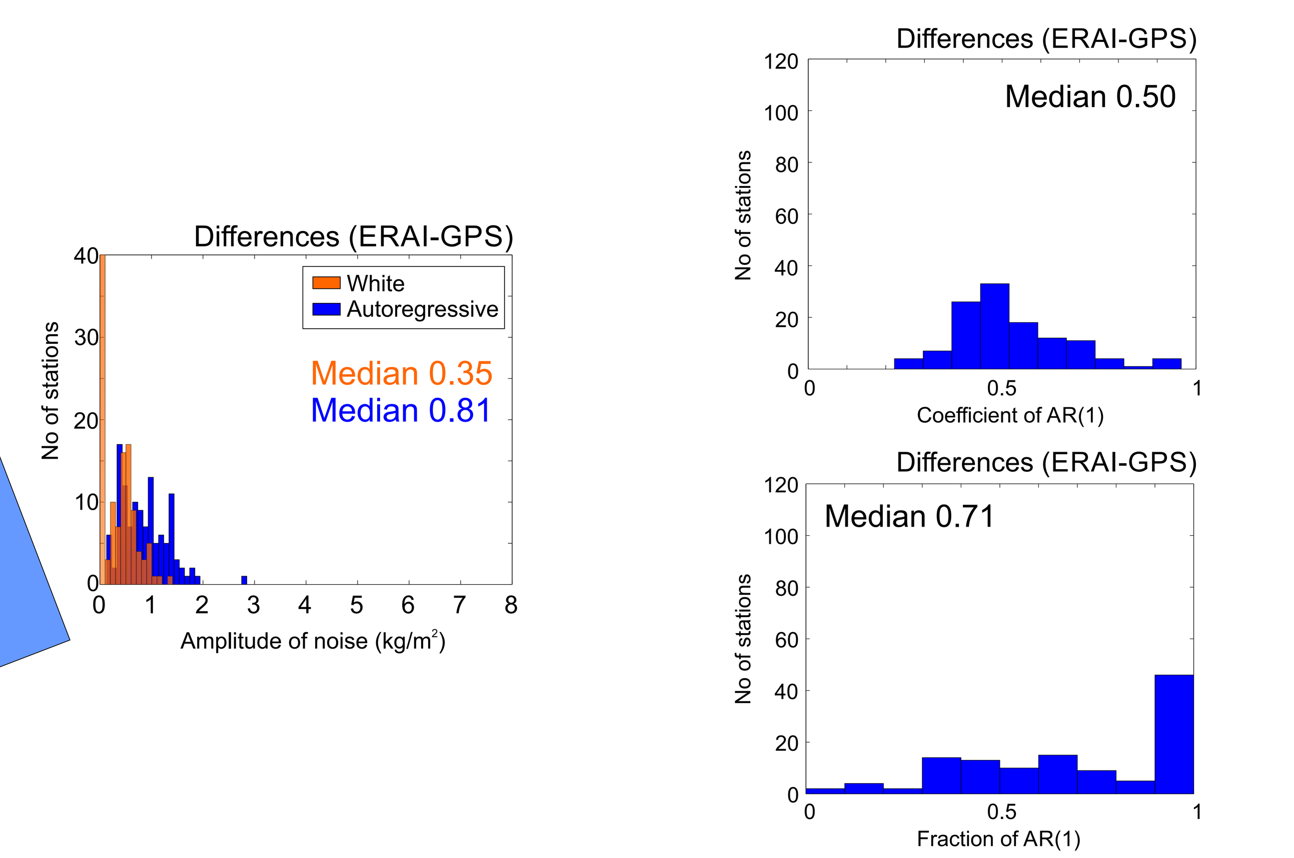


Figure 6. Parameters of autoregressive noise derived for differences of IWV. Left: histograms of fraction, coefficients and amplitude of white and autoregressive processes, right: coefficient and fraction of autoregressive noise.

NOW...

Test different statistical approaches to give a clear answer if we are able (or not) to report offsets with statistical approaches, being aware of the autoregressive noise in a stochastic part of the IWV data.