



Study on homogenization of synthetic GNSS-retrieved IWV time series and its impact on trend estimates with autoregressive noise

Anna Klos (1), Eric Pottiaux (2), Roeland Van Malderen (3), Olivier Bock (4), and Janusz Bogusz (1)

(1) Military University of Technology, Warsaw, Poland (anna.klos@wat.edu.pl), (2) Royal Observatory of Belgium, Brussels, Belgium, (3) Royal Meteorological Institute of Belgium, Brussels, Belgium, (4) IGN LAREG, University Paris Diderot, Sorbonne Paris, France

A synthetic benchmark dataset of Integrated Water Vapour (IWV) was created within the activity of “Data homogenisation” of sub-working group WG3 of COST ES1206 Action. The benchmark dataset was created basing on the analysis of IWV differences retrieved by Global Positioning System (GPS) International GNSS Service (IGS) stations using European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis data (ERA-Interim). Having analysed a set of 120 series of IWV differences (ERA-Interim-GPS) derived for IGS stations, we delivered parameters of a number of gaps and breaks for every certain station. Moreover, we estimated values of trends, significant seasonalities and character of residuals when deterministic model was removed. We tested five different noise models and found that a combination of white and autoregressive processes of first order describes the stochastic part with a good accuracy. Basing on this analysis, we performed Monte Carlo simulations of 25 years long data with two different types of noise: white as well as combination of white and autoregressive processes. We also added few strictly defined offsets, creating three variants of synthetic dataset: easy, less-complicated and fully-complicated. The ‘Easy’ dataset included seasonal signals (annual, semi-annual, 3 and 4 months if present for a particular station), offsets and white noise. The ‘Less-complicated’ dataset included above-mentioned, as well as the combination of white and first order autoregressive processes (AR(1)+WH). The ‘Fully-complicated’ dataset included, beyond above, a trend and gaps. In this research, we show the impact of manual homogenisation on the estimates of trend and its error. We also cross-compare the results for three above-mentioned datasets, as the synthesized noise type might have a significant influence on manual homogenisation. Therefore, it might mostly affect the values of trend and their uncertainties when inappropriately handled. In a future, the synthetic dataset we present is going to be used as a benchmark to test various statistical tools in terms of homogenisation task.