Sensitivity of segmentation of GNSS IWV time series and trend estimates to data properties

Khanh Ninh Nguyen\textsuperscript{1,2}, Annarosa Quarello\textsuperscript{1,2}, Olivier Bock\textsuperscript{1,2}, and Emilie Lebarbier\textsuperscript{3}

\textsuperscript{1}Université de Paris, Institut de Physique du Globe de Paris, CNRS, IGN, F-75005 Paris, France (bock@ipgp.fr)
\textsuperscript{2}ENSG-Géomatique, IGN, F-77455 Marne-la-Vallée, France
\textsuperscript{3}Laboratoire Modal’X, UPL, Université Paris Nanterre, France

Homogenization is an important step to improve the quality of long-term observational data sets and estimate climatic trends. In this work, we use the GNSSseg/GNSSfast segmentation packages that were developed by Quarello et al., 2020, for the detection of abrupt changes in the mean of Integrated Water Vapour (IWV) data derived from GNSS measurements. The method works on the difference of the IWV time series (GNSS – reference) in order to cancel out the common climatic variations and enhance the discontinuities due to the inhomogeneities in the GNSS series. This segmentation method accounts for changes in the variance on fixed intervals (monthly) and a periodic bias (annual) due to representativeness differences between GNSS and the reference (in our case, a global atmospheric reanalysis).

The goal of this study is to analyze the sensitivity of the segmentation method to the data properties, particularly the GNSS data processing method. Two reprocessed GNSS solutions are considered: IGS repro1, covering the period 1995-2010, and CODE REPRO2015 + OPER, covering the period 1994-2018. Next, the impact of the length of time series and missing data are investigated. Finally, the use of two different reference series is considered (ERA-Interim and ERA5 reanalyses).

The segmentation results are screened for outliers (multiple detections occurring within a distance of 80 days) and validated with respect to known equipment changes (from GNSS metadata). The impact of the data properties is analyzed by comparing the number and position of detected change-points and the fraction of validated change-points. The influence of the variance of the IWV difference series and the magnitude of the periodic bias is examined. Finally, the results are compared in terms of estimated linear trends taking the detected change-points into account. From the multiple comparisons, we found that about 30 \% of change points are similar when the GNSS processing method changed, while 60 \% are similar when the CODE series is shortened to match the length of the repro1 series. These tests highlight that the segmentation results are processing-dependent and are affected by the length of the series. The impact of the data properties on the IWV trends and associated uncertainties are also quantified. Besides, it is important to note that the best segmentation result is found when the ERA5 reanalysis is used as a reference.