EGU General Assembly 2022

IPGP

Université Paris Cité

IGN
INSTITUT NATIONAL
DE L'INFORMATION
GÉOGRAPHIQUE
ET FORESTIÈRE

A new method for the attribution of breakpoints in segmentation of IWV difference time series

This presentation participates in OSPP

EGU

Outstanding Student & PhD candidate Presentation contest

PhD candidate NGUYEN Ninh (IPGP -IGN)
Dr. BOCK Olivier (IPGP - IGN)
Prof. LEBARBIER Emilie (Université Paris-Nanterre)

27 May 2022

Sharing is encouraged

Introduction

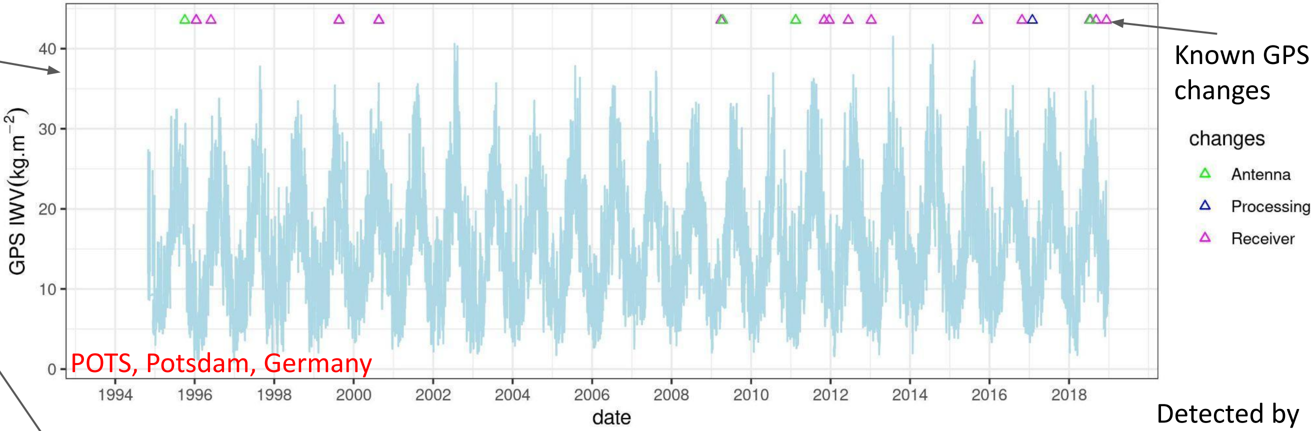
1. GPS series contains known equipment changes, but it's hard to see any induced IWV changes

2. Differenced series (GPS-ERA) is segmented using the GNSSseg method*

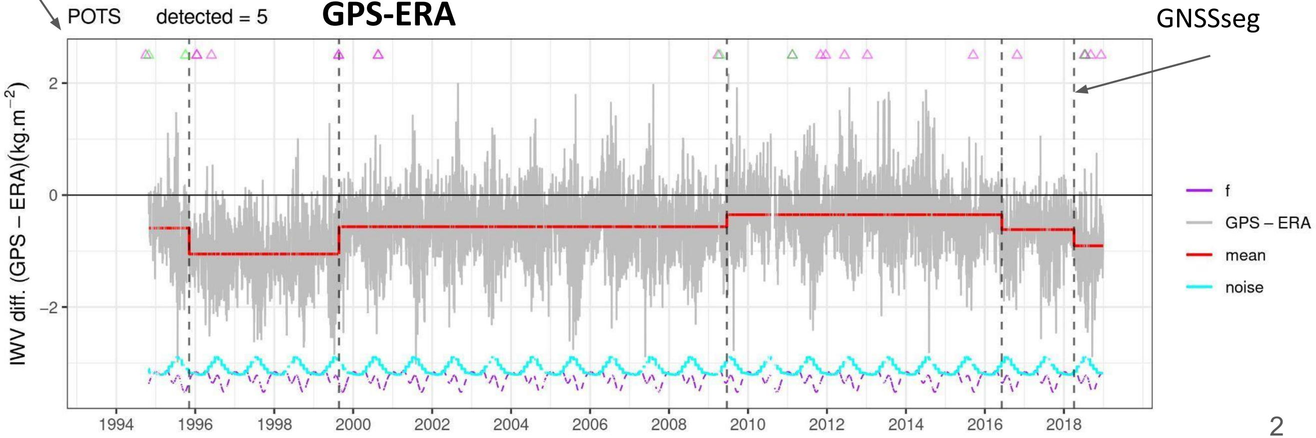
Problem: are the change-points due to GPS or to ERA ?

3. Some detected change points are "close" to known equipment changes and others are not…

=> Attribution = procedure used to chose between GPS and ERA



GPS-ERA



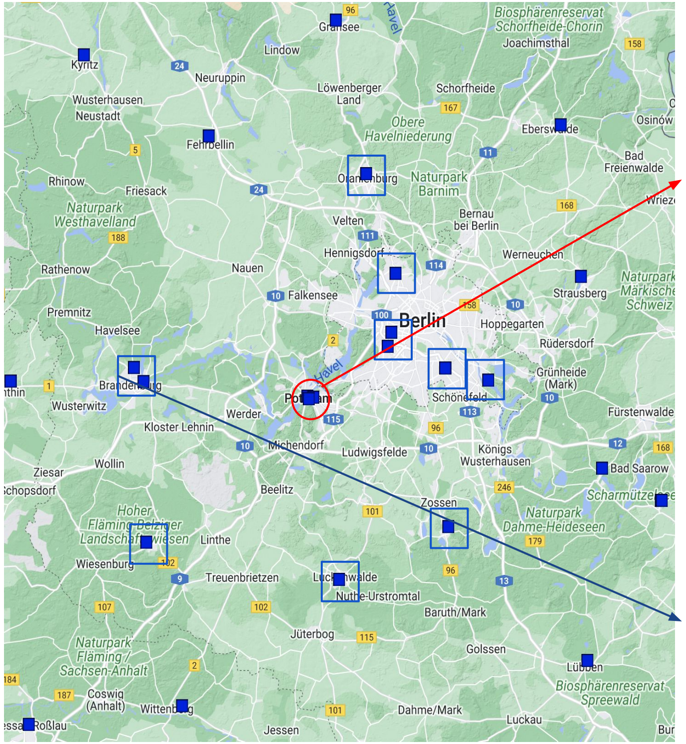
* GNSSseg R package available on the CRAN

Attribution method

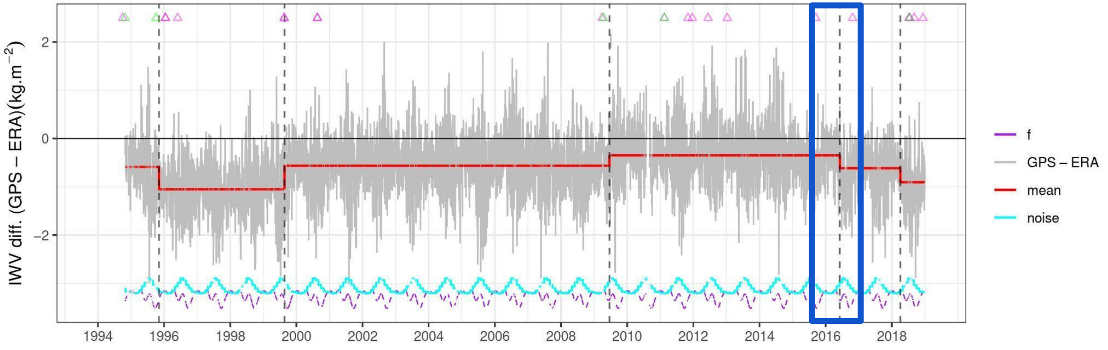
General idea:

1. Use data from **nearby stations** supposedly not affected by the same change-points
2. Each nearby station provides 2 additional series: GPS’ and ERA’
3. Form **6 series of differences** : GPS-ERA, GPS-GPS’, GPS-ERA’, ERA-ERA’, GPS’-ERA’, GPS’-ERA.
4. Apply a **statistical test** for a change in the mean to the 6 series of differences in a +/- 1 year window around each detected change-point
5. Determine in which of the 4 series (GPS, ERA, GPS’, or ERA’) the change point occured using a **predictive rule**.

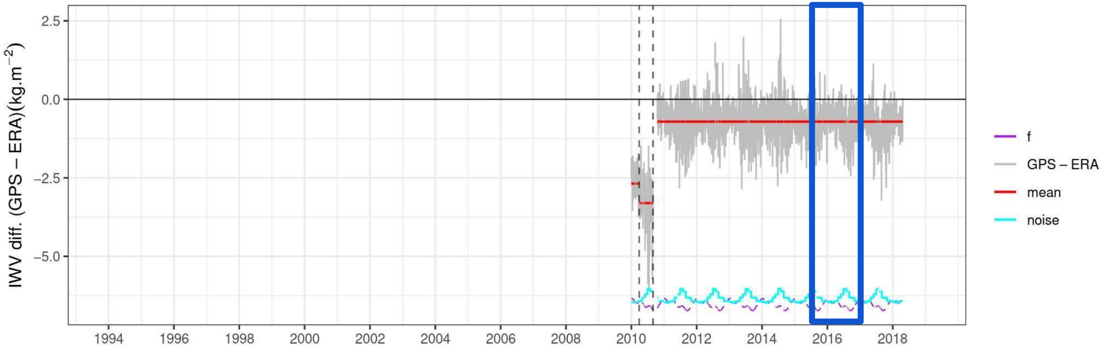
Select nearby station



Main station: POTS **GPS - ERA**



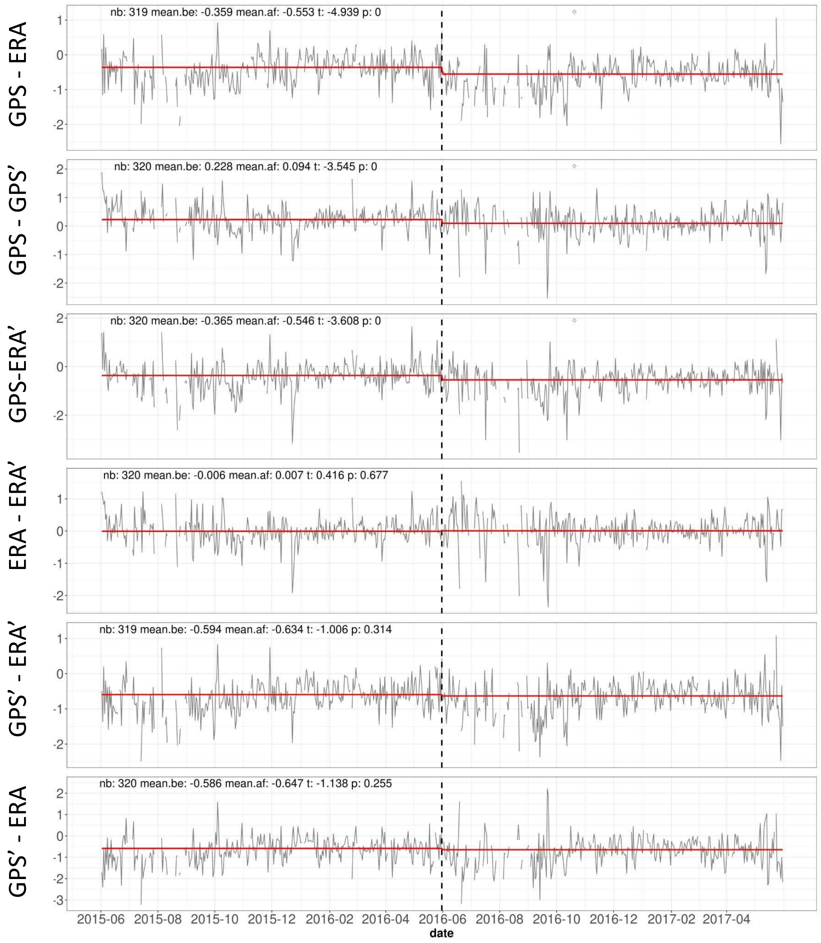
Nearby station: D001 **GPS' – ERA'**



Horizontal distance: 33km

Vertical distance: 57m

Form 6 series of differences and apply test



Test results

-1

-1

-1

0

0

0

Look-up table

Truth				Test results					
A	B	C	D	1	2	3	4	5	6
GPS	ERA	GPS'	ERA'	GPS-ERA	GPS-GPS'	GPS-ERA'	ERA-ERA'	GPS'-ERA'	GPS'-ERA
-1	0	0	0	-1	-1	-1	0	0	0
-1	0	0	1	-1	-1	-1	-1	-1	0
-1	0	0	-1	-1	-1	0	1	1	0
-1	0	1	0	-1	-1	-1	0	1	1
-1	0	1	1	-1	-1	-1	-1	0	1
0	1	0	0	-1	0	0	1	0	-1
0	1	1	1	-1	-1	-1	0	0	0

solution

- The look-up table provides the correct solution given that the combination of 6 test results exists in the table.
- If the combination is not in the table a more sophisticated search is necessary, e.g. based on prior probabilities for each combination or a statistical predictive rule.

Application to global GPS network

- Data sets:
- GPS = 81 IGS stations, CODE REPRO2015 solution
- GPS' = 704 nearby stations*, NGL repro3 solution
- ERA, ERA' = ECMWF reanalysis ERA5

- 156 breakpoints in 56 main stations can be tested with 1 to 10 nearby stations

- 109 breakpoints have at least 1 nearby station with combination of 6 tests results in the table

- Finally, 52% breakpoints are attributed to GPS and 48% are attributed to ERA.

Perspectives

1. Work in progress:
   Build a predictive rule using supervised classification

2. Final objective:
   After attribution, correct offsets in GPS series

More information:
https://meetingorganizer.copernicus.org/EGU22/EGU22-6390.html
knguyen@ipgp.fr

Thank you for your attention!

* Horizontal distance < 200km, vertical distance < 500m

Supplemental material

1. GPS data sets used in this study

	Main stations: CODE REPRO2015	**Nearby stations**: NGL
Software	Bernese GNSS software v5.3	GipsyX Version 1.0
Strategy	Double-difference solution of a global network	PPP
Orbits, clocks, ERPs	CODE repro2 (1994.0-2015.0) + CODE final (2015.0-2019.0)	daily Repro3.0 (JPL)
Reference frame	IGb08	IGS14
Antenna calibration	igs08_1852.atx until 28 January 2017, igs14.atx from 29 January 2017	igs14_www.atx
Window length	72h	30 h
Elevation cutoff angle	3	7
Observations	GPS (1994.0-2002.0), GPS+GLONASS (2002.0-2019.0)	GPS
Observation sampling	3 min	5 min
Observation weighting	1/cos(z)**2	1/sin(z)
Tropospheric model	6-hourly ECMWF analysis (provided by TUV). VMF1 mapping functions (hydrostatic and wet). Piece-wise linear model for ZWD with constraints: 5 m absolute and 5 m relative. Sampling : 2h (ZWD), 24h (gradients).	VMF1 gridded ECMWF tropo parameters from TU Vienna Mapping Function: VMF1 (hydrostatic and wet). Estimation: Zenith delay and gradients as random walk every 5 minutes
Tropo files	ZTD and gradient estimates provided in SINEX files (resampled to 01, 03,... 23 UTC)	ZTD and gradient estimates provided in SINEX files (0000, 0005,...2345 UTC)
Ambiguity resolution	Fixed	Fixed

2. Details of the attribution method

For each detected change-point in the main station:

Step 1: Select nearby stations and form 6 series of differences

- Search for nearby stations with limit distance of 200 km and height difference < 500m.
- Apply vertical correction(Bock et al., AMT, 2022) to the IWV series from the nearby stations => GPS' and ERA'.
- Check homogeneity of all series in the +/- 1 yr window (if other change points were detected, reduce the window)
- Form the 6 series of differences
- Screen the series of difference to remove outliers
- Pair the data on the left and right of the change-point
- If the final number of points is lower than 100 on each side, discard this nearby station.

Step 2: Characterize series and apply weighted t-test

- Heteroskedasticity: estimate the monthly variance of the difference series over the +/- 1 yr period with a robust estimator
- Autocorrelation: identify a stochastic model, ARMA of max order (1,1), for the difference series.
- Compute the t-statistic for the 6 differences series, taking into account the heteroskedasticity and autocorrelation, and test their significance.

Step 3: Apply a table search or a prediction rulte for the results of the 6 tests

Step 4: Summarize the results for the given breakpoint if several nearby stations have been used.

Step 2: Significance test on mean difference

- Use the weighted-t-test with the monthly variance. The weighted mean and the t-statistics can be computed as follow:

$$t = \frac{\hat{\mu_1} - \hat{\mu_2}}{\sqrt{\frac{1}{\sum_{t \in I_1} \frac{1}{\sigma_t^2}} + \frac{1}{\sum_{t \in I_2} \frac{1}{\sigma_t^2}}}} \qquad \hat{\mu_1} = \frac{\sum \frac{Y_t}{\sigma_t^2}}{\sum_{t \in I_1} \frac{1}{\sigma_t^2}}, \; \hat{\mu_2} = \frac{\sum \frac{Y_t}{\sigma_t^2}}{\sum_{t \in I_2} \frac{1}{\sigma_t^2}}$$

- The variance σ_t^2 is computed using a robust estimator
- The denominator in the t-statistic and is rescaled using the Effective Sample Size concept (Zwiers and von Storch, 1995):
- The ratio between the actual and the effective sample size is computed from the autocorrelation function:

$$\frac{n}{n_e} = 1 + 2\sum_{1}^{n-1}(1 - \frac{h}{n})\rho(h)$$

- The autocorrelation function rho(h) is the theoretical ACF for the identified ARMA(1,1) model with model parameters estimated by maximum likelihood method (Box and Jenkins, 2017).

Step 3: Table search

1. In total, there are 54 combinations of interest from GPS, ERA, GPS’, ERA’ (see left part of table)
→ 46 different combinations of 6 differences including values -2, -1, 0, 1, 2 (see logical table)

2. From test results, the only possible values are -1, 0, 1
→ 38 different combination of 6 test results can be detected

3. Remove the duplicate combinations in the *test results* table based on prior probabilities (see right part of table)

4. Use of the table: search for the combination in the *test result* table and deduce the corresponding ABCD values. If the combination is not in the table, search for the "nearest" one (e.g. with smallest number of differences). If several solutions exist, use again the prior probabilities to disentangle them.

	Truth				Logical table						Test results							Prior joint probability
	A	B	C	D	1	2	3	4	5	6	1	2	3	4	5	6		P(A,B,C,D)
	GPS	ERA	GPS'	ERA'	GPS-ERA	GPS-GPS'	GPS-ERA'	ERA-ERA'	GPS'-ERA'	GPS'-ERA	GPS-ERA	GPS-GPS'	GPS-ERA'	ERA-ERA'	GPS'-ERA'	GPS'-ERA		
1	1	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	0.1944
2	1	0	0	1	1	1	0	-1	-1	0	1	1	0	-1	-1	0	2	0.0108
3	1	0	0	-1	1	1	1	1	1	0	1	1	1	1	1	0	3	0.0108
4	1	0	1	0	1	0	1	0	1	1	1	0	1	0	1	1		0.0108
5	1	0	1	1	1	0	0	-1	0	1	1	0	0	-1	0	1		0.0006
6	1	0	1	-1	1	0	1	1	1	1	1	0	1	1	1	1	4	0.0006
7	1	0	-1	0	1	1	1	0	-1	-1	1	1	1	0	-1	-1	5	0.0108
8	1	0	-1	1	1	1	0	-1	-1	-1	1	1	0	-1	-1	-1	6	0.0006
9	1	0	-1	-1	1	1	1	1	0	-1	1	1	1	1	0	-1	7	0.0006
10	0	-1	0	0	1	0	0	-1	0	1	1	0	0	-1	0	1	8	0.0108
11	0	-1	0	1	1	0	-1	-1	-1	1	1	0	-1	-1	-1	1	9	0.0108
12	0	-1	0	-1	1	0	1	0	1	1	1	0	1	0	1	1	10	0.1944
13	0	-1	1	0	1	-1	0	-1	1	1	1	-1	0	-1	1	1	11	0.0006
14	0	-1	1	1	1	-1	-1	-1	0	1	1	-1	-1	-1	0	1	12	0.0006
15	0	-1	1	-1	1	-1	1	0	1	1	1	-1	1	0	1	1	13	0.0108
16	0	-1	-1	0	1	1	0	-1	-1	0	1	1	0	-1	-1	0		0.0006
17	0	-1	-1	1	1	1	-1	-1	-1	0	1	1	-1	-1	-1	0	14	0.0006
18	0	-1	-1	-1	1	1	1	0	0	0	1	1	1	0	0	0		0.0108
19	-1	0	0	0	-1	-1	-1	0	0	0	-1	-1	-1	0	0	0	15	0.1944
20	-1	0	0	1	-1	-1	-2	-1	-1	0	-1	-1	-1	-1	-1	0	16	0.0108
21	-1	0	0	-1	-1	-1	0	1	1	0	-1	-1	0	1	1	0	17	0.0108
22	-1	0	1	0	-1	-2	-1	0	1	1	-1	-1	-1	0	1	1	18	0.0108
23	-1	0	1	1	-1	-2	-2	-1	0	1	-1	-1	-1	-1	0	1	19	0.0006
24	-1	0	1	-1	-1	-2	0	1	2	1	-1	-1	0	1	1	1	20	0.0006
25	-1	0	-1	0	-1	0	-1	0	-1	-1	-1	0	-1	0	-1	-1		0.0108
26	-1	0	-1	1	-1	0	-2	-1	-2	-1	-1	0	-1	-1	-1	-1	21	0.0006
27	-1	0	-1	-1	-1	0	0	1	0	-1	-1	0	0	1	0	-1		0.0006
28	0	1	0	0	-1	0	0	1	0	-1	-1	0	0	1	0	-1	22	0.0108
29	0	1	0	1	-1	0	-1	0	-1	-1	-1	0	-1	0	-1	-1	23	0.1944
30	0	1	0	-1	-1	0	1	2	1	-1	-1	0	1	1	1	-1	24	0.0108
31	0	1	1	0	-1	-1	0	1	1	0	-1	-1	0	1	1	0		0.0006
32	0	1	1	1	-1	-1	-1	0	0	0	-1	-1	-1	0	0	0		0.0108
33	0	1	1	-1	-1	-1	1	2	2	0	-1	-1	1	1	1	0	25	0.0006
34	0	1	-1	0	-1	1	0	1	-1	-2	-1	1	0	1	-1	-1	26	0.0006
35	0	1	-1	1	-1	1	-1	0	-2	-2	-1	1	-1	0	-1	-1	27	0.0108
36	0	1	-1	-1	-1	1	1	2	0	-2	-1	1	1	1	0	-1	28	0.0006
37	1	-1	0	0	2	1	1	-1	0	1	1	1	1	-1	0	1	29	0.0009
38	1	-1	0	1	2	1	0	-2	-1	1	1	1	0	-1	-1	1	30	0.0009
39	1	-1	0	-1	2	1	2	0	1	1	1	1	1	0	1	1	31	0.0162
40	1	-1	1	0	2	0	1	-1	1	2	1	0	1	-1	1	1	32	0.00005
41	1	-1	1	1	2	0	0	-2	0	2	1	0	0	-1	0	1		0.00005
42	1	-1	1	-1	2	0	2	0	2	2	1	0	1	0	1	1		0.0009
43	1	-1	-1	0	2	2	1	-1	-1	0	1	1	1	-1	-1	0	33	0.00005
44	1	-1	-1	1	2	2	0	-2	-2	0	1	1	0	-1	-1	0		0.00005
45	1	-1	-1	-1	2	2	2	0	0	0	1	1	1	0	0	0		0.0009
46	-1	1	0	0	-2	-1	-1	1	0	-1	-1	-1	-1	1	0	-1	34	0.0009
47	-1	1	0	1	-2	-1	-2	0	-1	-1	-1	-1	-1	0	-1	-1	35	0.0162
48	-1	1	0	-1	-2	-1	0	2	1	-1	-1	-1	0	1	1	-1	36	0.0009
49	-1	1	1	0	-2	-2	-1	1	1	0	-1	-1	-1	1	1	0	37	0.00005
50	-1	1	1	1	-2	-2	-2	0	0	0	-1	-1	-1	0	0	0		0.0009
51	-1	1	1	-1	-2	-2	0	2	2	0	-1	-1	0	1	1	0		0.00005
52	-1	1	-1	0	-2	0	-1	1	-1	-2	-1	0	-1	1	-1	-1	38	0.00005
53	-1	1	-1	1	-2	0	-2	0	-2	-2	-1	0	-1	0	-1	-1		0.0009
54	-1	1	-1	-1	-2	0	0	2	0	-2	-1	0	0	1	0	-1		0.00005