ORIGINAL ARTICLE

Investigating temporal and spatial patterns in the stochastic component of ZTD time series over Europe

Anna Klos[1] · Janusz Bogusz[1] · Rosa Pacione[2] · Vincent Humphrey[3] · Henryk Dobslaw[4]

Abstract
The EUREF Permanent GNSS Network (EPN) provides a unique atmospheric dataset over Europe in the form of Zenith Total Delay (ZTD) time series. These ZTD time series are estimated independently by different analysis centers, but a combined solution is also provided. Previous studies showed that changes in the processing strategy do not affect trends and seasonal amplitudes. However, its effect on the temporal and spatial variations of the stochastic component of ZTD time series has not yet been investigated. This study analyses the temporal and spatial correlations of the ZTD residuals obtained from four different datasets: one solution provided by ASI (Agenzia Spaziale Italiana Centro di Geodesia Spaziale, Italy), two solutions provided by GOP (Geodetic Observatory Pecny, Czech Republic), and one combined solution resulting from the EPN's second reprocessing campaign. We find that the ZTD residuals obtained from the three individual solutions can be modeled using a first-order autoregressive stochastic process, which is less significant and must be completed by an additional white noise process in the combined solution. Although the combination procedure changes the temporal correlation in the ZTD residuals, it neither affects its spatial correlation structure nor its time-variability, for which an annual modulation is observed for stations up to 1,000 km apart. The main spatial patterns in the ZTD residuals also remain identical. Finally, we compare two GOP solutions, one of which only differs in the modeling of non-tidal atmospheric loading at the observation level, and conclude that its modeling has a negligible effect on ZTD values.

Keywords GNSS · GPS · Zenith total delay · Empirical orthogonal functions · Autoregressive process · Non-tidal atmospheric loading

Introduction

The Global Positioning System (GPS) signal passing from the satellite to the receiver is delayed due to the transition through different layers of the atmosphere. This delay is a site-specific parameter that gives a measure of the integrated atmospheric condition due to the hydrostatic and wet contributions of the neutral atmosphere when the GPS signal is received. The signal delay along each satellite-receiver path is transferred into the zenith direction (Boehm et al. 2006) and forms the so-called Zenith Total Delay (ZTD). ZTD series are used in meteorological applications, climate studies, and numerical weather prediction (Mahfouf et al. 2015; Mile et al. 2019). Their quality influences these applications and there have been suggestions that it should even be considered along with ZTD values as an additional input variable (Lindskog et al. 2017).

The character and quality of ZTD time series are affected by several uncertainties related to the processes and models used to estimate them (Ning et al. 2016). GPS processing strategies vary among analysis centers, for example, in terms of the type of software, reference frame, the solution used, or background models, as there are no well-defined guidelines (Pacione et al. 2017). Although changes in processing strategy can cause differences in estimates of individual ZTD values or introduce unwanted periodic signals (Pacione and Vespe 2003; Watson et al. 2006), the deterministic parameters of ZTD time series that are used in climate analyses,

✉ Anna Klos
  anna.klos@wat.edu.pl

1 Military University of Technology, Warsaw, Poland

2 e-GEOS S.P.A, ASI/CGS, Matera, Italy

3 Institute for Atmospheric and Climate Science, ETH Zürich, Zurich, Switzerland

4 GFZ German Research Centre for Geosciences, Potsdam, Germany

i.e., trends and seasonal signals, remain almost unaffected (Baldysz et al. 2016; Dousa et al. 2017; Pacione et al. 2017). This is encouraging, as it suggests that trends and seasonal amplitudes can be interpreted as actual tropospheric variations, regardless of the technique or strategy used to estimate them (Bock et al. 2014). Although the influence of processing strategies on ZTD trends and seasonal signals has already been studied, its effect on the stochastic component of the ZTD time series and its spatiotemporal properties has not yet been addressed. The impact on the stochastic component may be significant, given that the magnitudes of the residuals are several times smaller than those of the seasonal signals (e.g., Baldysz et al. 2016; Klos et al. 2018).

We address this problem for the European daily ZTD time series. We use solutions from two EUREF GNSS Permanent Network (EPN) analysis centers and the official combined EPN Repro-2 solution (Pacione et al. 2017), which users use most often. (GNSS stands for the Global Navigation Satellite System.) The individual solutions are evaluated in terms of their temporal and spatial consistency. Temporal consistency is examined by determining the parameters of the ZTD residuals (trend and seasonal signals removed beforehand). Spatial consistency is evaluated using correlation coefficients and empirical orthogonal functions. To the best of our knowledge, this analysis is the first to demonstrate four things. First, the ZTD residuals from the combined solution are not consistent with the ZTD residuals from the individual analysis centers. Second, the correlation coefficients of the ZTD residuals estimated for station pairs are time-varying and dependent on the distance between stations. Third, the spatiotemporal patterns present in the ZTD residuals over Europe are consistent between solutions and are also physically meaningful. Finally, omitting the non-tidal atmospheric

loading (NTAL) model has a negligible effect on the ZTD residuals.

We introduce the ZTD time series employed, along with a description of its pre-analysis and homogenization. Then, we describe the methodology and briefly present the time series parameters. The main part of the research focuses on the temporal and spatiotemporal consistency of ZTD residuals. We summarize our research in the final section.

Datasets

We use four sets of ZTD time series. Three of them are produced by two EPN analysis centers: AS0 by Centro di Geodesia Spaziale G. Colombo (ASI), Matera, Italy; and GO1 and GO4 by Geodetic Observatory Pecny (GOP), Czech Republic, in the framework of the second EPN reprocessing campaign. Table 1 summarizes the processing options used in these three solutions. In addition to the three individual solutions, we also include in the analyses the combined solution resulting from the second reprocessing campaign, EPN Repro-2 (hereafter denoted as COMB). A detailed description and evaluation of the COMB along with individual tropospheric products from the analysis centers can be found in Pacione et al. (2017). The number of stations is 320 for AS0, 240 for GO1, GO4, and 217 for COMB. The AS0 solution, in addition to EPN stations, also includes IGS (International GNSS Service) core stations outside Europe. Since AS0 is a PPP (Precise Point Positioning) solution, the IGS core stations do not affect the ZTD estimates from the EPN stations. Therefore, we use the entire set of 320 stations for the AS0 solution. The ZTD data have time resolutions of 0.5 and 1 h and are stored in SINEX-TRO format. Both time resolutions are averaged to a daily sampling interval; the combined solution covers the period from the beginning of 1996 until the end of 2014. In all solutions, a period of 6.5 years is considered sufficient to obtain reliable estimates, but this choice is somewhat subjective. The median length of the time series is 14 years, with 19 stations having the longest time series of 19 years.

To assess the consistency of the solutions we use, we estimate the repeatability of the ZTD time series and note that it is similar to the series processed by ASI and GOP, while all are systematically shifted relative to the COMB solution (Fig. 1); the repeatability is determined for the residuals estimated using (1) and described further below. Furthermore, the repeatability of the series produced by GO1 and GO4, which differ only in the inclusion of the NTAL effect, is very similar, suggesting that its inclusion may have little effect on the individual ZTD values and does not affect the precision of the observations.

Table 1 Processing strategy adopted in the AS0 and GO1/GO4 solutions as part of the second EPN reprocessing campaign (reproduced from Pacione et al. 2017)

Analysis Center	AS0	GO1	GO4
Software	GIPSY 6.2	Bernese 5.2	
GNSS system	GPS	GPS	
Solution type	PPP	Network	
Stations	Full EPN + IGS core stations	Full EPN	
Orbits	JPL repro2	CODE repro2	
Antennas	IGS08	IGS08 and individual	
Gravity	EGM08	EGM08	
Mapping functions	VMF	VMF1	VMF
Ionosphere modelling	HOI included	HOI included	
Solid Earth tide model	IERS2010	IERS2010	
Elevation cut-off	3	3	
Tidal ocean loading (TOL)	Yes (FES2004)	Yes (FES2004)	
Tidal atmospheric loading (TAL)	No	No	
Non-tidal atmospheric loading (NTAL)	No	No	Yes (model from TU Vienna)

Statistical analysis methodology

For the ZTD value at the i-th epoch, the following mathematical description applies:

$$ZTD = a + b \cdot t + \sum_{o=1}^{p} c \cdot H(t - t_o)$$
$$+ \sum_{k=1}^{m} \left[A_k \cdot \sin(\omega_k \cdot t) + B_k \cdot \cos(\omega_k \cdot t) \right] + \varepsilon \tag{1}$$

where a is the initial value, b is the linear trend, c is the offset amplitude, $H(t)$ is a Heaviside step function, A and B are sine and cosine terms of periodic signals of angular frequency ω. ε is the residual (or stochastic component) that remains after removing the deterministic model. The periodic signals include the annual (with a period of 365.25 days, that is 1 cycle per year (cpy) frequency), semiannual (182.63 days, 2 cpy), triennial (121.75 days, 3 cpy) and quadrennial (91.31 days, 4 cpy) signals ($m=4$) (Baldysz et al. 2016). We identify outliers in the daily ZTD time series beforehand as the observations exceed three times the interquartile range. We estimate epochs of offsets (t_0) by performing a manual homogenization based on the epochs reported for the position time series. Using these epochs as a priori input, we fit the model in (1) using weighted least-squares estimation. We then iteratively exclude those epochs whose $1 - \sigma$ error is larger than their magnitude. On this basis, we perform three iterations and leave only those offsets for which the magnitudes are significant for the final analysis.

To study the time-correlated variations in the ZTD time series, consistent with Klos et al. (2018), we assume that

ε can be modeled with a white process combined with an autoregressive process (WN + AR(1)):

$$\varepsilon = \varepsilon_{WN} + \varepsilon_{AR(1)} \tag{2}$$

with $\varepsilon_{AR(1)}$ being defined using the following recurrence expression (Sowel 1992):

$$\varepsilon_{AR(1)_i} = \varphi \cdot \varepsilon_{AR(1)_{i-1}} + Z_i \tag{3}$$

where φ is an autoregressive coefficient of the first order and Z_i is the Gaussian noise. We do not test alternative noise models since it has been done in previous works. Note that the appropriateness of the WN + AR(1) model is highlighted in Fig. 2, which depicts the power spectral density of the ZTD residual time series of individual EPN stations against that predicted by the model. The power spectral density of the ZTD residuals approximately shows a constant power at frequencies below 10 cpy, a decrease in power between 10 and 100 cpy, and a flattening above 100 cpy.

The WN + AR(1) model is characterized by three parameters: the variance factor, the noise fraction, and the coefficient φ of the AR(1) process. The noise fractions of white and first-order autoregressive processes indicate their relative contribution to the model. The coefficient of the AR(1) process points to how much the current value is dependent on the previous one. We use Hector software (Bos et al. 2013) to estimate both deterministic and stochastic parameters with the Maximum Likelihood Estimation approach (Langbein 2017).

To study the spatially correlated variations, we analyze both the Pearson correlation coefficient estimated from pairs of residual ZTD time series and the spatiotemporal patterns retrieved by the probabilistic Principal Component Analysis of the ZTD residuals (pPCA; applied as in Gruszczynski et al. 2018). This pPCA approach is based on the PCA but re-formulated to handle the missing values in the time series without interpolation (Tipping and Bishop 1999). In particular, we estimate individual Principal Components (PCs), the total variance of series explained by individual PCs, and normalized responses for individual PCs. Spatial dependencies are retrieved by using normalized responses of individual stations to a particular PC. The total variance of series explained by individual PC is employed to decide on the significance of the contribution of this PC in the time series residuals.

Temporal correlation of ZTD residuals

ZTD residuals of the stations included in the temperate zone (denoted as C in Fig. 3) are characterized by the highest noise standard deviation, exceeding 3.5 mm (not shown here). Lower standard deviations, around 3 mm, are observed for the continental zone covering central Europe
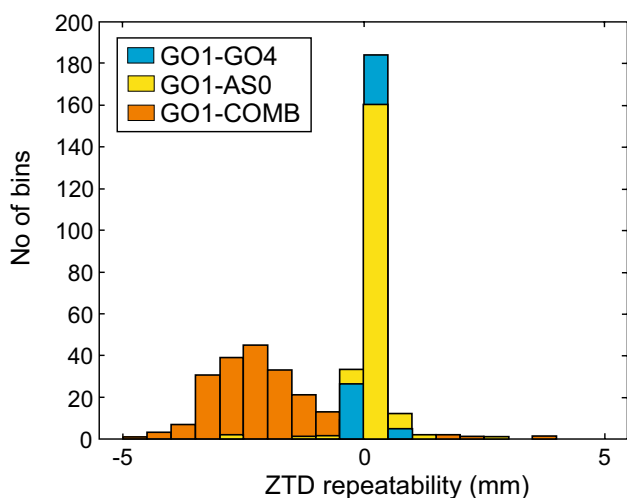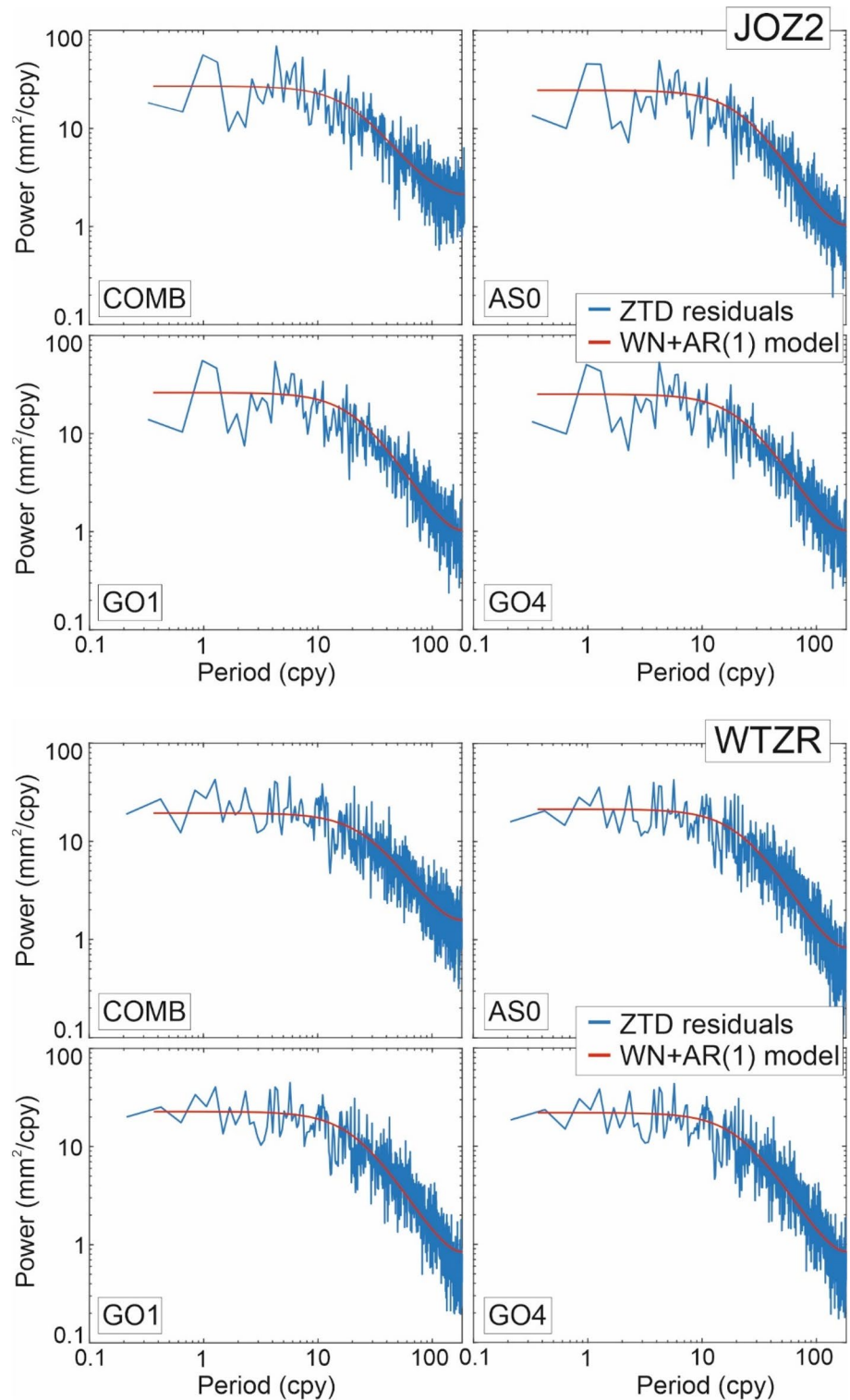

Fig. 1 Repeatability of ZTD residuals plotted for differences between GO1 and GO4 (in blue), GO1 and AS0 (in yellow) and GO1 and COMB (in orange)

Fig. 2 Power spectral densities of ZTD residuals (in blue) and WN + AR(1) model (in red) plotted for two randomly selected GPS stations: JOZ2 (Poland) and WTZR (Germany). The plots are presented for the four solutions we use


(denoted as D). In contrast, the Scandinavian and Alpine regions (denoted as E) show the lowest scatter of ZTD residuals. Figure 3 presents a contribution of first-order autoregressive process in the WN + AR(1) model, which is only slightly smaller than 100% for the entire set of stations,

regardless of the climate zone. This means there is almost no white noise in the ZTD residuals for AS0, GO1, and GO4 solutions. Unlike the fraction of AR(1), the AR(1) coefficient values depend on the climate zones and are latitude-dependent. The smallest AR(1) coefficient values are found for latitudes within $35 - 43°$, covering the Mediterranean

Sea, Black Sea, and North Atlantic coastal areas. Central European stations located in the temperate zone (denoted as *C*) are characterized by larger values of AR(1) coefficients. The largest estimates are observed in continental and polar climate zones (denoted as D and E, respectively).

Generally, ZTD residuals should exhibit some temporal autocorrelation but not on a time scale longer than atmospheric circulation and moisture transport. For instance, ground-based microwave measurements of precipitable water vapor made at the Madeira Archipelago, in the

Fig. 3 AR(1) fraction (contribution) estimates for the WN + AR(1) noises combination and AR(1) coefficient, plotted on the top of the Köppen-Geiger climate classification. The main Köppen-Geiger classes are as follows: A (tropical), B (dry), C (temperate), D (continental), and E (polar). Class A does not occur in Europe. Each class is split into sub-classes. A fraction of 100% means that ZTD residuals contain pure AR(1) noise
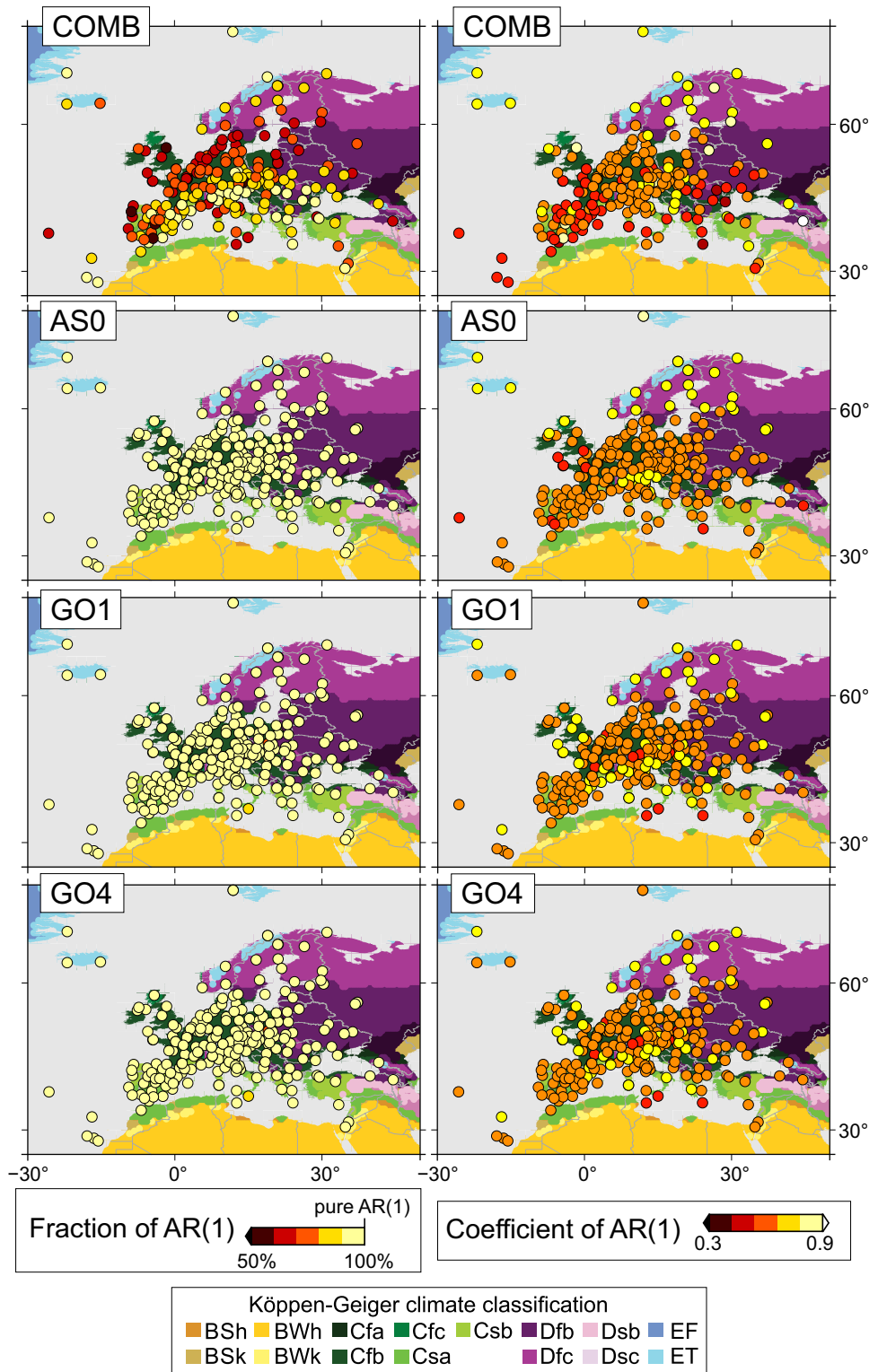
Atlantic Ocean, suggest a decorrelation time of up to 24 h (Snider 2000). Overland, large-scale averages of reanalysis-based daily atmospheric water vapor transport anomalies have a decorrelation time of up to 5 days (Gutowski et al. 1997), approximately corresponding to a 1-day lag correlation of at most 0.66 (von Storch 1999). In addition to the variability of atmospheric water vapor, other physical processes might influence the autocorrelation of ZTD residuals with daily sampling. As anomalies in land water storage have a much longer persistence than anomalies in tropospheric water vapor, they could potentially increase the serial autocorrelation of the ZTD residuals to a level that cannot be explained by atmospheric water vapor variability alone. Indeed, our analysis shows daily ZTD serial autocorrelation coefficients φ higher than 0.7 for most stations (Fig. 3), which is higher than what is suggested in the literature (Gutowski et al. 1997; Snider 2000).

The above dependencies are not valid for the combined solution COMB. Neither dependency on latitude nor dependency on climate zones is observed for the combined product. The fraction of AR(1) noise process is lower than that of individual solutions, meaning that some part of AR(1) noise is lost in favor of white noise, which is recognized as a processing noise. The AR(1) noise coefficient estimated for the combined solution COMB is also lower than the estimates provided for individual EPN analysis centers, meaning that the combination reduces the signal for individual stations.

Spatial correlation of ZTD residuals

Spatial consistency between ZTD residuals can be demonstrated using Pearson's correlation coefficient ρ. The correlation coefficients estimated for all station pairs for the COMB solution are higher than 0.8 for stations less than 200 km apart (Fig. 4). The highest spatial correlation is observed in the alpine region in northern Italy. This correlation decreases at further distances and drops to zero for distances greater than 2,000 km. Since the AS0 solution also includes stations outside Europe that were not included in the combined solution, we can see a further flattening of the spatial dependence for stations more than 6,000 km apart. Different spatial correlations between station pairs is noted for GO1 and GO4 solutions. The coefficients show very large discrepancies even for similar separation distances. A few stations from GO1 and GO4 solutions show similar characteristics to the COMB and AS0 solutions, i.e., the correlation decreases with distance, but most of the ZTD residuals do not reflect this relationship. There are ZTD residuals that are highly correlated with each other, but there are also some for identical distance that are negatively correlated.

Averaging the correlation coefficients for 63 equal distance classes of 100 km (red dots in Fig. 4) clearly shows how correlation coefficients differ between COMB or AS0 and GO1 or GO4 solutions. Even for the shortest distances between stations, the average correlation coefficient differs by 0.3. We use the polynomial function of degree 4 to

Fig. 4 Spatial correlation identified by the correlation coefficients ρ between the residuals ε for any pair of stations as a function of the distance d between them, given in kilometers. The red dots represent correlation coefficients averaged over equal distance classes of 100 km, and the black line is the adjusted spatial correlation function, i.e., a fourth-degree polynomial. The blue line represents the exponential fit into correlation coefficients; no averaging was performed in this case
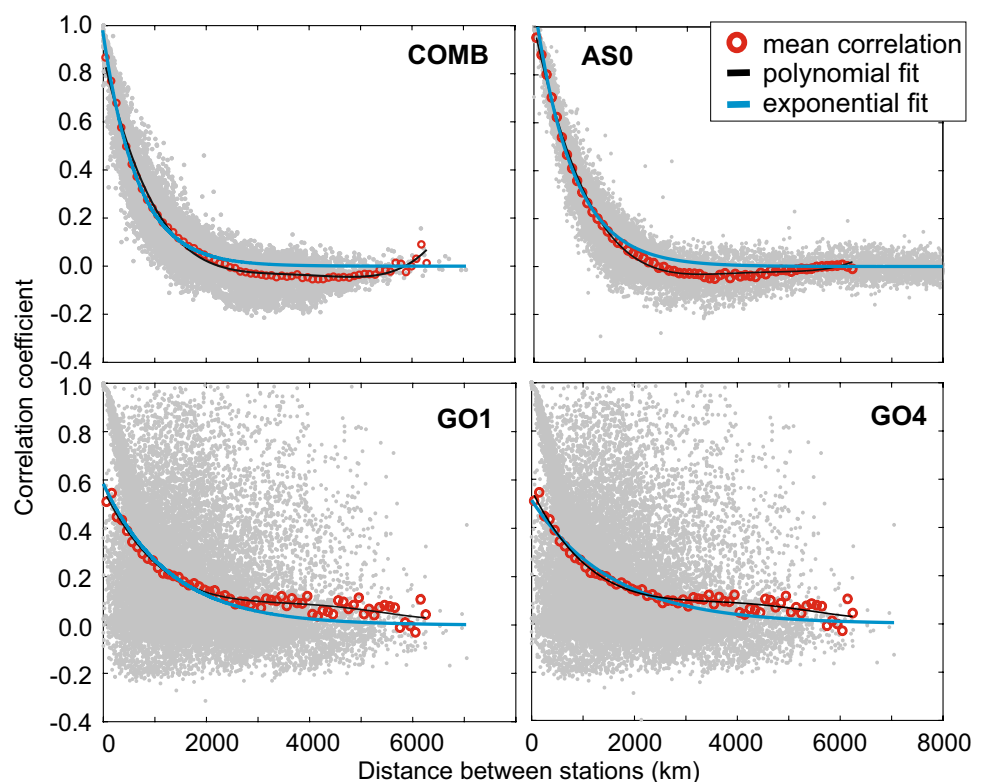
characterize the dependence of average correlation coefficients $\bar{\rho}$ on the average distance between stations $\bar{d}$:

$$\bar{\rho} = p_0 + p_1\bar{d} + p_2\bar{d}^2 + p_3\bar{d}^3 + p_4\bar{d}^4 \qquad (4)$$

We find that this function is efficient enough to describe this relationship. Since we know that polynomials will fit into any type of data, we compare it with an exponential model fitted into unaveraged correlation coefficients ρ:

$$\rho = a_e e^{b_e d} \qquad (5)$$

where a_e and b_e are coefficients of exponential model, d stands for the distance between any pair of stations. Parameters of both models are provided in Table 2 for all solutions we use. We point out the advantages of using an exponential function over a polynomial function; the parametric function makes it possible to reproduce the value of the coefficient for any distance between stations and does not cause it to adjust artificially to places where there is less data. In contrast, Fig. 4 clearly shows that the exponential function, being an asymptotic function, cannot reproduce the negative correlation coefficients for distances between 2,000 and 4,000 km for AS0 and COMB solutions.

The spatial correlations between the ZTD residuals for all station pairs contained in the AS0 solution are almost identical to those for the COMB solution, meaning that although some AR(1) noise was removed from the time series during the combination procedure, correlation coefficients estimated between station pairs seem to be unaffected.

We now examine the variability of correlation coefficients over time for 63 equal distance classes of 100 km (Fig. 5). We observe that the spatial correlation coefficients change with time. In particular, they show clear annual oscillations with a maximum in January from the shortest distances up to 1,000 km. For larger distances, from 1,000 to 1,500 km, we only note a small annual component for the AS0 solution. The correlation coefficients for distances above 1,500 km

have no significant features and oscillate around zero. A similar spatial correlation seasonality in GNSS vertical position time estimates was presented in Gobron et al. (2021). They showed that spatial correlation is reduced when non-tidal loading corrections are applied, suggesting that the main contributor to the annual signal present in the correlation coefficient is the intrinsic property of atmospheric variability. The remainder of the annual signal can be explained by imperfections in physical models, especially those derived from orbit modeling. We also observe a clear correlation reduction at short distances for GO1 and GO4 solutions from 1996 to 2003. This is due to the fact that before 2003, ZTD estimates are only available for stations that are highly correlated with each other. After 2003, there are more stations

Table 2 Parameters of the polynomial and exponential functions fitted to the correlation coefficients estimated between pairs of stations depending on the distance between them

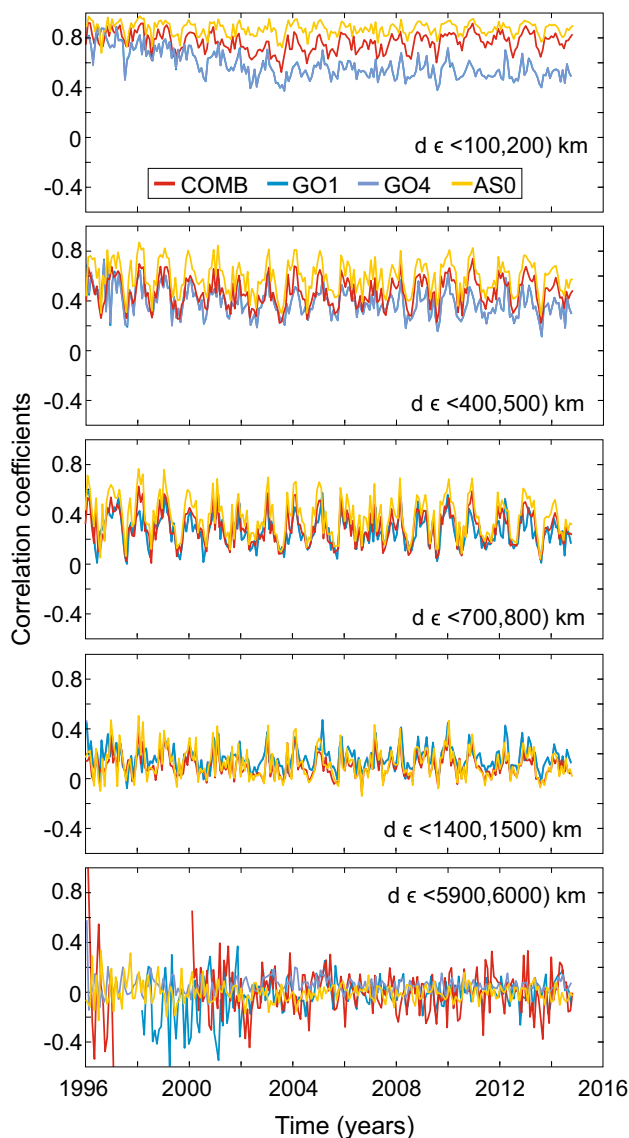|	COMB	AS0	GO1	GO4
Polynomial fit				
p_0	0.92	1.01	0.55	0.55
p_1	$-0.08\text{E}-02$	$-0.10\text{E}-02$	$-0.04\text{E}-02$	$-0.04\text{E}-02$
p_2	3.21E-07	3.58E-07	1.49E-07	1.52E-07
p_3	$-5.47\text{E}-11$	$-5.57\text{E}-11$	$-2.32\text{E}-11$	$-2.36\text{E}-11$
p_4	3.20E-15	3.22E-15	1.29E-15	1.31E-15
Exponential fit				
a_e	0.981	1.105	0.515	0.514
b_e	-0.0015	-0.0014	-0.0006	-0.0006

Fig. 5 Time-variable correlation coefficients estimated for station pairs whose distance was within d, averaged over a given month

for which the correlation is negative, hence decreasing the average.

We examine the residuals of the ZTD time series using the pPCA approach to assess the spatiotemporal patterns present in each dataset. Figure 6 presents spatial responses associated with the first 3 PCs with explained variance higher than 10%. The responses to the 1st PC are positive for the majority of EPN stations included in all solutions we use, also including combined solution COMB, and they are the largest for the temperate and part of the continental zones; both in central Europe (Fig. 3). Although the combination procedure reduces the signal for individual stations, spatial patterns retrieved from ZTD residuals remain unaltered up to the 4th PC. Correlation coefficients estimated between

1st PCs for all solutions we use are higher than 0.97. Signal included in 1st PC explains around 43% of the total variance of ZTD residuals variability. Among possible explanations, this is most likely related to the barycenter of the EPN network despite the various software packages employed, or data processing types applied. However, since AS0 is a PPP solution, including also IGS core stations situated outside Europe, it should not be biased by the barycenter effect. Therefore, alternatively, this 1st PC pattern might also indicate a continental-level response to variability in weather conditions. GO4 solution is computed including NTAL during the GPS processing, so it should not respond in the same way as other solutions do unless there are still some mismodeled effects in the NTAL model which leak to the ZTD
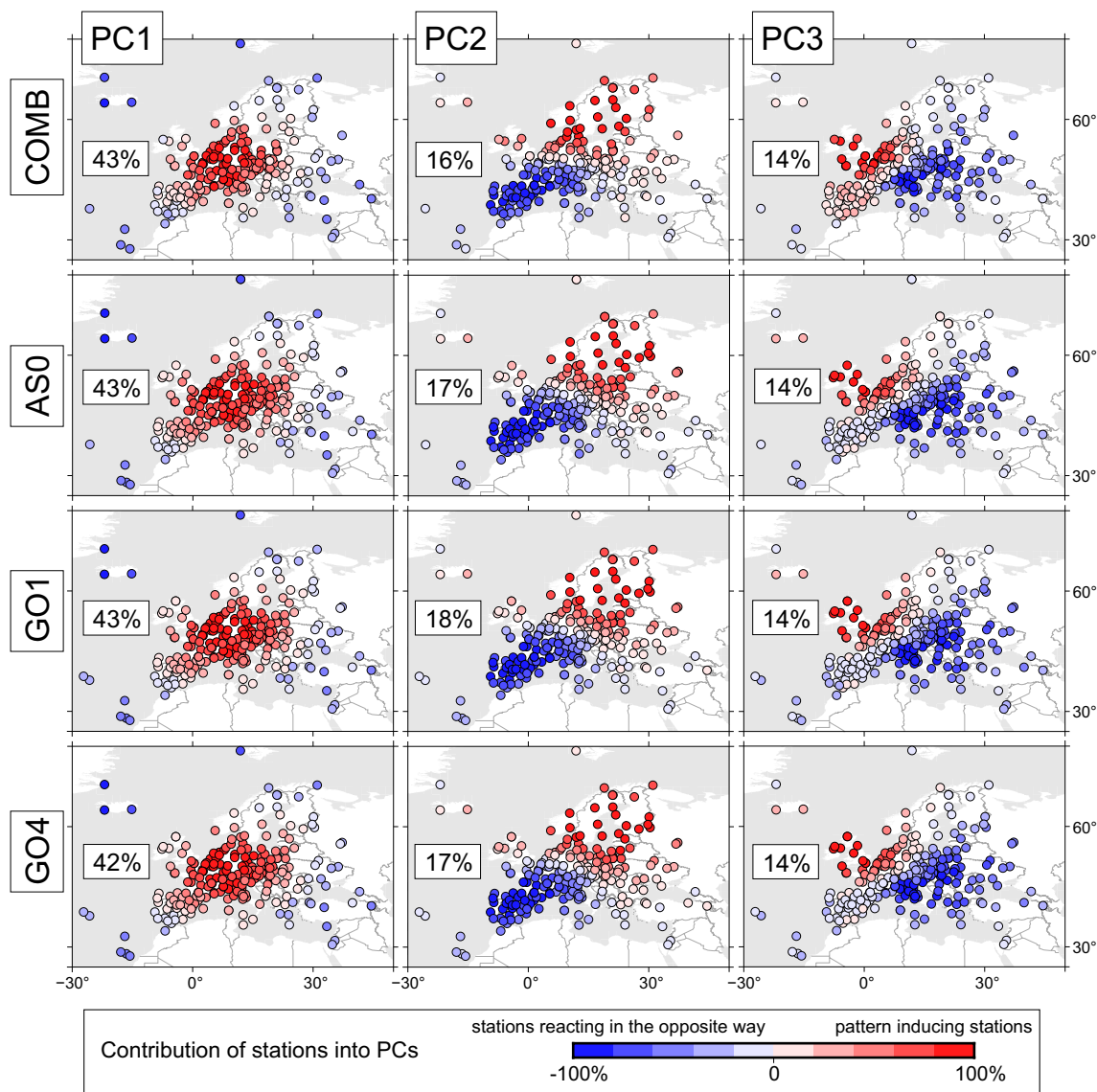

Fig. 6 Spatial responses to first three PCs computed for ZTD residual time series obtained from the COMB, AS0, GO1 and GO4 solutions. For each solution, the percentage of total variance explained by PC is provided (100% means total variance)

residuals. If so, their impact must be large and constitutes almost half of the observations' variance, i.e., 43%.

Higher PCs reveal a similar character of spatial responses for the combined solution COMB as well as individual contributions from AS0, GO1, and GO4. Temporal responses for all solutions are highly correlated with each other, with more than 0.90 for PCs from 2nd to 4th. For higher PCs, although their spatial responses are generated by a similar set of stations for all solutions, we notice a decrease in correlation between COMB and individual solutions for the temporal responses. For the 5th PC, the correlation is equal to -0.93, whereas for the 6th PC it ranges between 0.80 and 0.86, for AS0 and GO1/GO4 solutions, respectively. This confirms that some part of the information is lost during the combination procedure. The combination procedure affects common signals present in ZTD residuals of explained variance lower than 10%. However, if a combined ZTD product (that is, a weighted mean of the individual contributing solutions) is validated against other independent techniques such as radiosondes or VLBI (Very-Long Baseline Interferometry), its standard deviation is lower than that of the contributing solutions (Pacione et al. 2011).

The 2nd PC, explaining about 17% of the variance, displays a strong gradient from the southeast to the northwest (Fig. 6). The responses for the 2nd PC are imposed by the North and Baltic Seas stations, for which the signals are positive. Only the Mediterranean and western European stations respond negatively to the 2nd PC.

The 3rd PC, explaining 14% of the variance, displays a northeast to southwest gradient, for all solutions (Fig. 6). This mode is imposed by the North Sea stations. Both 2nd and 3rd spatial patterns correspond well to the main modes of winter and summer atmospheric circulation in Europe. Similar results can be found in the literature concerning empirical orthogonal functions analyses of precipitable water vapor (Zveryaev et al. 2008; Wypych et al. 2018) and sea level pressure (Moron and Plaut 2003; Folland et al. 2009). These modes are related to atmospheric circulation, which influences European weather through the interaction of atmospheric blocking system variations and North Atlantic storm tracks.

Spatial responses derived for PCs higher than 3rd explain less than 10% of the total variance. They are generated by the same sets of stations, meaning that the continental-level responses are related to real atmospheric processes, reaching signals of small variance. No significant difference can be found between PCs derived for GO1 and GO4 solutions, meaning that ZTD residuals retrieved from a dense European network are not systematically affected by NTAL corrections. The pattern of ZTD for all PCs largely concurs for the different individual solutions and also for the combined one. This confirms that the major features of ZTD were

robustly estimated in the past and might be done so in the future even without NTAL corrections.

## Analysis of non-tidal atmospheric loading impact

In this section, we focus on the differences between GO1 and GO4 residuals (GO1-GO4, Fig. 7), which differ only in including the NTAL effect. We exclude MDVO station (Russia) here, whose ZTD estimates after 2002 cannot be reliably explained (an abrupt increase of root-mean-square value with large interannual changes).

In particular, we investigate the correlation between GO1 and GO4 differences and the NTAL vertical surface deformation, as computed from geophysical fluid models by the Earth-System-Modelling group of Deutsches Geo-ForschungsZentrum (ESMGFZ; Dill and Dobslaw 2013). We estimate GO1-GO4 residual differences and compute the spatiotemporal patterns using pPCA. It is worth noting that the GOP analysis center used the NTAL provided by TU Vienna when estimating GO4 solution. The TU Vienna product is based on the same ECMWF surface pressure data as the ESMGFZ product. Thus, the differences are only due to the elastic Earth model and perhaps to other (secondary) data processing choices. Therefore, we assume that the type of NTAL used when processing the GO4 solution should not affect the comparison with the ESMGFZ model we present in this paper.

The 1st PC explains 38% of the total variance of GO1-GO4 differences. It is induced by inland stations, with coastal stations responding negatively. It is similar to the 1st PC from Fig. 6. The correlation estimated between 1st PC of GO1 and GO4 differences and 1st and 3rd PC of NTAL loading model is significantly negative, reaching, respectively, -0.43 and -0.54 (Fig. 8). Similar to ZTD residuals, the 1st PC of the GO1 and GO4 differences is positively correlated with the 6th PC of NTAL, reaching 0.43. 2nd PC explains 16% of the total variance of GO1-GO4 differences, with clear southeast to northwest stripes. This PC may be NTAL-generated, as its spatial responses look similar to the combination of 1st and 2nd PCs of NTAL. However, their correlation in the temporal domain is negative, between -0.10 and -0.20 (Fig. 8). 4th PC explains 11% of GO1-GO4 differences' variance and it is generated by North Sea stations, with no significant correlation to the NTAL model noticed.

We notice that individual PCs of the NTAL model might not necessarily be correlated to GO1-GO4 differences at the same level (1st PC to 1st PC, etc.), as they are sorted according to their decreasing variance. Effects that induce high variance in loading might not unavoidably induce the same amount of variance in GO1-GO4 differences. Therefore, we compute correlation coefficients and compare the PCs on different levels (Fig. 8). We find that almost all values

Fig. 7 Spatial responses to PCs computed for differences between GO1 and GO4 (GO1-GO4) and residuals of NTAL. The percentage of total variance explained by each PC is provided (100% means total variance)
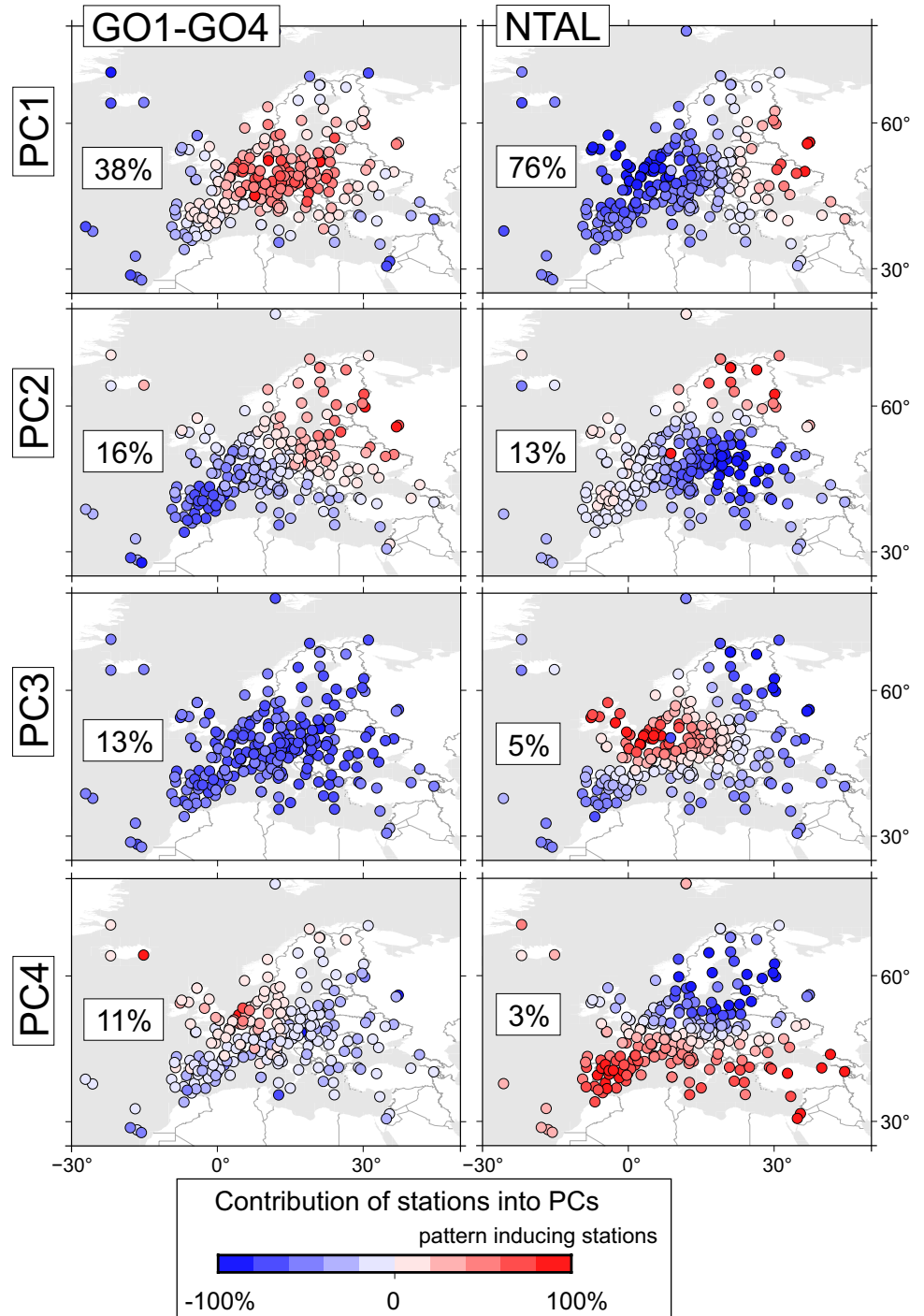
are insignificant, confirming the robustness of ZTD residuals with respect to the application of NTAL at the present level of accuracy. This is also confirmed by summing the explained variance for the 1st and 2nd PCs of GO1-GO4 differences, which correlate positively with NTAL residuals, i.e., 38% and 16%. We find that 54% of GO1-GO4 differences may be impacted by un- or mismodeled NTAL. However, since this signal appears only in the GO1-GO4 differences, we suggest that it will affect only the error spectra of ZTD, with no influence on linear trends or amplitudes of seasonal signals. The above has implications on the significance of ZTD trends (their errors), and, therefore, we recommend that the NTAL is applied at the observation level during GPS observation processing.
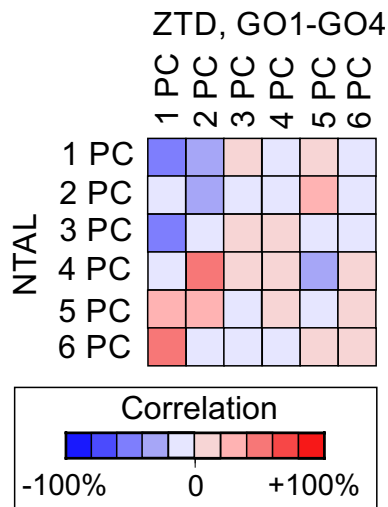
ZTD, GO1-GO4

Fig. 8 Correlation estimated for individual PCs between NTAL residuals and GO1-GO4 differences

Summary

We examine the temporal and spatial properties of the stochastic component of the ZTD time series over Europe. We use the three homogeneously processed solutions provided in the second EPN Reprocessing campaign, namely AS0 and GO1/GO4, together with the combined EPN Repro-2 solution. We remove the deterministic model and examine ZTD residuals. We find that its temporal correlation is similar for all AS0, GO1 and GO4 solutions, which suggests that ZTD reflects real atmospheric conditions in the entire frequency band. They differ, however, from the combined product of EPN. For the combined solution, some part of the autoregressive process is lost in favor of white noise during the combination. This could influence analyses of the significance of ZTD trends or future recommendations on the minimum period required for reliable estimates of ZTD trends. This should, however, not be used as an argument on which solution one should use or not. We should keep in mind that ZTD values are estimated during the processing, and their accuracy is strongly dependent on the input used. Combination changes the stochastic process that best describes individual ZTD time series, but, as we also show, the physical spatial properties remain intact. This means that ZTD is estimated robustly, even for signals of small variance.

We show that the ZTD residuals for different locations are strongly spatially correlated; correlation coefficients are higher than 0.8 for stations situated less than 500 km from each other. The highest spatial correlation is observed for the alpine region in northern Italy. This is true for the COMB and AS0 solutions. The complete opposite situation occurs for GO1 and GO4 solutions, for which the ZTD residuals are strongly correlated or

completely uncorrelated for stations that are the same distance apart; no evident spatial correlation function is observed. The reason for this weaker correlation with respect to the COMB and AS0 solutions requires further research. Moreover, we also show that the spatial correlation between station pairs is time-varying. We observe clear annual oscillations in the time series of correlation coefficients for stations up to 1,000 km apart. Then, these oscillations diminish, and the correlation coefficients oscillate around zero. A similar situation was noticed for the position time series by Gobron et al. (2021).

We use empirical orthogonal functions and show that the modes of ZTD residuals of all four solutions are very similar. We show that the main mode of ZTD residuals in the European area is largely impacted by the continental-level response to atmospheric loading or affected by a large-scale network-related signal. We also have evidence that the other modes are consistent over the European area, which is strongly coupled with atmosphere processes, covering even signals of the smallest variance.

To search for possible contributions of NTAL to the ZTD residuals, we compute GO1-GO4 differences and cross-compare them with spatiotemporal responses of NTAL residuals. Correlation coefficients are barely significant for corresponding levels of PC, but this is related to the method itself, which orders PCs according to the amount of total variance they explain. A slight positive correlation is noted at different levels of PC. We conclude that NTAL effects will be disclosed in ZTD; however, due to the small variance they explain (they are present only in GO1-GO4 differences), they will only contribute to the error spectra of GPS-derived ZTDs, affecting the error of ZTD trend. This could be important given the wide range of applications in which ZTDs are employed nowadays.

References

Baldysz Z, Nykiel G, Araszkiewicz A, Figurski M, Szafranek K (2016) Comparison of GPS tropospheric delays derived from two consecutive EPN reprocessing campaigns from the point of view of climate monitoring. Atmos Meas Tech 9(9):4861–4877. https://doi.org/10.5194/amt-9-4861-2016

Bock O, Willis P, Lacarra M, Bosser P (2014) An inter-comparison of zenith tropospheric delays derived from DORIS and GPS data. Adv Space Res 46(12):1648–1660. https://doi.org/10.1016/j.asr.2010.05.018

Boehm J, Werl B, Schuh H (2006) Troposphere mapping functions for GPS and very long baseline interferometry from European centre for medium-range weather forecasts operational analysis data. J Geophys Res 111:B02406. https://doi.org/10.1029/2005JB003629

Bos MS, Fernandes RMS, Williams SDP, Bastos L (2013) Fast error analysis of continuous GNSS observations with missing data. J Geod 87(4):351–360. https://doi.org/10.1007/s00190-012-0605-0

Dill R, Dobslaw H (2013) Numerical simulations of global-scale high-resolution hydrological crustal deformations. J Geophys Res: Solid Earth 118(9):5008–5017. https://doi.org/10.1002/jgrb.50353

Dousa J, Vaclavovic P, Elias M (2017) Tropospheric products of the second GOP European GNSS reprocessing (1996–2014). Atmos Meas Tech 10:3589–3607. https://doi.org/10.5194/amt-10-3589-2017

Folland CK, Knight J, Linderholm HW, Fereday D, Ineson S, Hurrell JW (2009) The summer North Atlantic oscillation: past, present, and future. J Clim 22(5):1082–1103. https://doi.org/10.1175/2008JCLI2459.1

Gobron K, Rebischung P, Van Camp M, Demoulin A, de Viron O (2021) Influence of aperiodic non-tidal atmospheric and oceanic loading deformations on the stochastic properties of global GNSS vertical land motion time series. J Geophys Res Solid Earth 126:e2021JB022370. https://doi.org/10.1029/2021JB022370

Gruszczynski M, Klos A, Bogusz J (2018) A filtering of incomplete GNSS position time series with probabilistic principal component analysis. Pure Appl Geophys 175(5):1841–1867. https://doi.org/10.1007/s00024-018-1856-3

Gutowski WJ Jr, Chen Y, Ötles Z (1997) Atmospheric water vapor transport in NCEP–NCAR reanalyses: comparison with river discharge in the central United States. B Am Meteorol Soc 78(9):1957–1970. https://doi.org/10.1175/1520-0477

Klos A, Hunegnaw A, Teferle FN, Abraha KE, Ahmed F, Bogusz J (2018) Statistical significance of trends in Zenith wet delay from reprocessed GPS solutions. GPS Solut 22:51. https://doi.org/10.1007/s10291-018-0717-y

Langbein J (2017) Improved efficiency of maximum likelihood analysis of time series with temporally correlated errors. J Geod 91:985–994. https://doi.org/10.1007/s00190-017-1002-5

Lindskog M, Ridal M, Thorsteinsson S, Ning T (2017) Data assimilation of GNSS zenith total delays from a Nordic processing centre. Atmos Chem Phys 17:13983–13998. https://doi.org/10.5194/acp-17-13983-2017

Mahfouf J-F, Ahmed F, Moll P, Teferle FN (2015) Assimilation of zenith total delays in the AROME France convective scale model: a recent assessment. Tellus Dyn Meteorol Oceanogr 67:1. https://doi.org/10.3402/tellusa.v67.26106

Mile M, Benáček P, Rózsa S (2019) The use of GNSS zenith total delays in operational AROME/Hungary 3D-Var over a central European domain. Atmos Meas Tech 12:1569–1579. https://doi.org/10.5194/amt-12-1569-2019

Moron V, Plaut G (2003) The impact of El Niño-Southern oscillation upon weather regimes over Europe and the North Atlantic during boreal winter. Int J Climatol 23(4):363–379. https://doi.org/10.1002/joc.890

Ning T, Wang J, Elgered G, Dick G, Wickert J, Bradke M, Sommer M, Querel R, Smale D (2016) The uncertainty of the atmospheric integrated water vapour estimated from GNSS observations. Atmos Meas Tech 9:79–92. https://doi.org/10.5194/amt-9-79-2016

Pacione R, Araszkiewicz A, Brockmann E, Dousa J (2017) EPN-Repro2: a reference GNSS tropospheric data set over Europe. Atmos Meas Tech 10:1689–1705. https://doi.org/10.5194/amt-10-1689-2017

Pacione R, Vespe F (2003) GPS Zenith Total Delay estimation in the mediterranean area for climatological and meteorological applications. J Atmos Ocean Tech 20(7):1034–1042. https://doi.org/10.1175/1520-0426

Pacione R, Pace B, de Haan S, Vedel H, Lanotte R, Vespe F (2011) Combination methods of tropospheric time series. Adv Space Res 47:323–335. https://doi.org/10.1016/j.asr.2010.07.021

Snider JB (2000) Long-term observations of cloud liquid, water vapor, and cloud-base temperature in the North Atlantic Ocean. J Atmos Ocean Tech 17(7):928–939. https://doi.org/10.1175/1520-0426

Sowel F (1992) Maximum likelihood estimation of stationary univariate fractionally integrated time series models. J Econom 53:165–188. https://doi.org/10.1016/0304-4076(92)90084-5

von Storch H (1999) Misuses of statistical analysis in climate research. In: von Storch H, Navarra A (eds) Analysis of Climate Variability. Springer, Berlin Heidelberg, pp 11–26. https://doi.org/10.1007/978-3-662-03744-7_2

Tipping ME, Bishop CM (1999) Probabilistic principal component analysis. J Roy Stat Soc 61B:611–622. https://doi.org/10.1111/1467-9868.00196

Watson C, Tregoning P, Coleman R (2006) Impact of solid earth tide models on GPS coordinate and tropospheric time series. Geophys Res Lett 33(8):L08306. https://doi.org/10.1029/2005GL025538

Wypych A, Bochenek B, Rozycki M (2018) Atmospheric moisture content over Europe and the Northern Atlantic. Atmos 9(1):18. https://doi.org/10.3390/atmos9010018

Zveryaev II, Wibig J, Allan RP (2008) Contrasting interannual variability of atmospheric moisture over Europe during cold and warm seasons. Tellus A 60(1):32–41. https://doi.org/10.1111/j.1600-0870.2007.00283.x

Anna Klos is an associate professor at the Military University of Technology, Poland. Her main research area is related to time series analysis in terms of earth's deformations, loading effects, tropospheric delay and their applications. She is focused on the geodetic deliverables, which may be helpful for climate change studies.

Vincent Humphrey is a Senior Scientist at ETH Zürich in the Institute for Atmospheric and Climate Science. From 2019 to 2020, he was a SNSF Postdoctoral Fellow at the California Institute of Technology. His research focuses on the interactions between the atmosphere, land hydrology, and terrestrial ecosystems in the context of recent climate change.

Janusz Bogusz is a full professor of geodesy at the Military University of Technology, Warsaw, Poland. His research interests are in physical geodesy, geodynamics, and geodetic time series analysis. He is currently the President of the Commission 3 "Earth Rotation and Geodynamics" of the International Association of Geodesy (IAG) and a Fellow of the IAG.

Henryk Dobslaw is leading the Working Group "Hydrosphere Dynamics" within the Department "Geodesy" of the German Research Centre for Geosciences (GFZ) in Potsdam. He is developing numerical models of large-scale mass redistributions in the atmosphere, oceans, and the terrestrial hydrosphere to aid the processing and interpretation of geodetic observations taken both at the ground and from space.

Rosa Pacione is currently working at e-GEOS S.p.A. as project manager of the space geodetic activities performed in the framework of the ASI contract to operate the ASI Centro di Geodesia Spaziale (CGS) in Matera (Italy). She is a member of the EUREF Governing Board and EPN Troposphere Coordinator with the task of generating the combined EPN tropospheric solutions. Her main research area focuses on improving GNSS processing strategy for atmospheric water vapor estimation and regional reference frame.