Balidakis, K., Nilsson, T., Zus, F., Glaser, S., Heinkelmann, R., Deng, Z., Schuh, H. (2018): Estimating Integrated Water Vapor Trends From VLBI, GPS, and Numerical Weather Models: Sensitivity to Tropospheric Parameterization. - *Journal of Geophysical Research*, *123*, 12, pp. 6356—6372.

DOI: http://doi.org/10.1029/2017JD028049

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1029/2017JD028049

Key Points:
- VLBI data analysis benefits from employing homogenized in situ meteorological data
- Alternating the meteorological data source in VLBI analysis impacts on IWV rates, whereas the mapping functions does not
- Good agreement between IWV rates from VLBI, GPS, and ERA-Interim employing a variant of the MIDAS estimator

Correspondence to:
K. Balidakis,
kyriakos.balidakis@gfz-potsdam.de

Citation:
Balidakis, K., Nilsson, T., Zus, F., Glaser, S., Heinkelmann, R., Deng, Z., & Schuh, H. (2018). Estimating integrated water vapor trends from VLBI, GPS, and numerical weather models: Sensitivity to tropospheric parameterization. *Journal of Geophysical Research: Atmospheres*, *123*, 6356–6372. https://doi.org/10.1029/2017JD028049

Received 11 NOV 2017
Accepted 17 MAY 2018
Accepted article online 24 MAY 2018
Published online 20 JUN 2018

Estimating Integrated Water Vapor Trends From VLBI, GPS, and Numerical Weather Models: Sensitivity to Tropospheric Parameterization

Kyriakos Balidakis[1], Tobias Nilsson[1], Florian Zus[1], Susanne Glaser[1], Robert Heinkelmann[1], Zhiguo Deng[1], and Harald Schuh[1,2]

[1]GFZ, German Research Centre for Geosciences, Space Geodetic Techniques, Potsdam, Germany, [2]Institute for Geodesy and Geoinformation Science, Technische Universität Berlin, Berlin, Germany

Abstract In this study, we estimate integrated water vapor (IWV) trends from very long baseline interferometry (VLBI) and global navigation satellite systems (GNSS) data analysis, as well as from numerical weather models (NWMs). We study the impact of modeling and parameterization of the tropospheric delay from VLBI on IWV trends. We address the impact of the meteorological data source utilized to model the hydrostatic delay and the thermal deformation of antennas, as well as the mapping functions employed to project zenith delays to arbitrary directions. To do so, we derive a new mapping function, called Potsdam mapping functions based on NWM data and a new empirical model, GFZ-PT. GFZ-PT differs from previous realizations as it describes diurnal and subdiurnal in addition to long-wavelength variations, it provides harmonic functions of ray tracing-derived gradients, and it features robustly estimated rates. We find that alternating the mapping functions in VLBI data analysis yields no statistically significant differences in the IWV rates, whereas alternating the meteorological data source distorts the trends significantly. Moreover, we explore methods to extract IWV given a NWM. The rigorously estimated IWV rates from the different VLBI setups, GNSS, and ERA-Interim are intercompared, and a good agreement is found. We find a quite good agreement comparing ERA-Interim to VLBI and GNSS, separately, at the level of 75%.

1. Introduction

Quantifying water vapor is a crucial task because (i) it is the most efficient and abundant greenhouse gas as it enforces global warming (via albedo) and reduces the absorption of solar energy (via clouds), (ii) the spatiotemporal variations in its content have a dominant impact on weather systems and climate change, and (iii) obtaining accurate and precise results from the analysis of microwave-based space geodetic observations highly depends on the successful modeling of the propagation delay effects it induces. Integrated water vapor (IWV) increases with increasing temperature at a rate of 5–7%/°C (e.g., Alshawaf et al., 2017; Nilsson & Elgered, 2008). Due to the average relative humidity being conserved, IWV rates are global warming indicators, and hence, there is a crucial importance in studying them. The IWV can be precisely deduced from space geodetic techniques observing at microwave frequencies, such as very long baseline interferometry (VLBI) and global navigation satellite systems (GNSS). Throughout the years, many studies have been dedicated to the estimation of IWV trends from VLBI and GNSS (e.g., Alshawaf et al., 2017; Heinkelmann et al., 2007; Nilsson & Elgered, 2008; Schuh et al., 2006).

The nature of the analysis of network-based space geodetic techniques such as VLBI and GNSS is such that deficiencies in station-specific models not only affect the local parameters of the station they refer to but also contaminate the entire network. This issue is emphasized through a small number of stations and at stations participating in the datum definition via minimum constraints. Moreover, to compensate for disregarded effects, additional parameters are estimated causing increased sensitivity to observations' noise.

Thus, to ensure the high accuracy of the main geodetic products as well as the by-products that are essential for better comprehending the physical system where the observations are conducted (such as IWV), it is necessary to successfully model systematic and random errors. In modern microwave-based geodesy, a large contribution to the error budget stems from the propagation delay induced by the volatile water vapor content.

The current study is motivated by the relatively few investigations clearly addressing the implications of the modeling and parameterization of the tropospheric delay on IWV trends. The impact of a priori zenith hydrostatic delay as well as the hydrostatic and nonhydrostatic mapping function is of main interest. To the authors' knowledge, comparisons between rigorously estimated IWV trends from GNSS, VLBI, and numerical weather models (NWMs) featuring up to 23, 34, and 38 years, respectively, have not been carried out. Furthermore, very few studies have been carried out assessing IWV trends derived from NWM IWV series that were extracted employing different approaches.

In the following, we succinctly describe how we homogenize the meteorological data recorded in the vicinity of VLBI stations (section 2), we develop the Potsdam mapping functions (hereafter PMF) and gradients (section 3), and we create a new empirical model for meteorological parameters of geodetic interest, GFZ-PT (section 4). We study how to extract IWV series from NWMs (section 5). The VLBI and GNSS data analysis carried out here are outlined in section 6. We assess the impact of alternating the meteorological data source and the mapping functions within VLBI data analysis (section 7). Rigorously estimated IWV rates from VLBI, GNSS, and ERA-Interim (Dee et al., 2011) are intercompared in section 8. Section 9 summarizes the results and provides recommendations. Further details on PMF and GFZ-PT are provided in Appendices A and B, respectively.

2. Homogenization of In Situ Meteorological Observations

From the meteorological data recorded in the vicinity of VLBI stations the barometric pressure is used for the calculation of the zenith hydrostatic delay (d_h^z) and the air temperature for the thermal deformation of VLBI antennas. These series often experience spurious offsets and drifts that deteriorate the ensuing geodetic analysis and eventually induce artifacts to the zenith nonhydrostatic delays (d_{nh}^z), that render the physical interpretation of trends in IWV uncertain.

Temperature variations cause deformation of the telescope structure, thus inducing group delay fluctuations depending on the antenna characteristics and its location. Those typically do not exceed 15 ps, even for the largest telescopes. Since this effect is 3–4 orders of magnitude smaller than the atmospheric delay, and the temperature lapse rate is much smaller than the pressure lapse rate, emphasis is placed on the impurities of the pressure series. Reasonable meteorological station relocations will yield small changes in the temperature records and almost no impact on the VLBI analysis.

Conventionally, the a priori atmospheric model consists of only the slant hydrostatic delay, which is usually the product of d_h^z with the hydrostatic mapping function mf_h. Since mf_h differs significantly from the nonhydrostatic mapping function mf_{nh} as $\varepsilon \to 0°$, estimating the remaining zenith atmospheric delay (nonhydrostatic part, theoretically) cannot fully compensate impurities in the a priori d_h^z. This is the point where employing a total mapping function is of some advantage, as the geodetic results are not affected by imprecise d_h^z (Böhm et al., 2006); notwithstanding, a shortcoming of this approach is that it fails to describe variations in the atmospheric state at time scales of hours or shorter, should the temporal resolution of the related coefficients be lower than 3 hr (Nilsson et al., 2013).

To detect systematic spurious effects in the recursive average, we performed the penalized maximal t test (e.g., Wang et al., 2007) on the anomalies that were formed as the difference between the raw observations and series rigorously extracted from the model levels of ERA-Interim (hereafter ERAinML). Figure 1 illustrates a case (HOBART26) where a drift and jumps were detected in the pressure time series. The reason we resort to employing the model levels and not just the surface fields lies in the fact that extracting series from the model levels of a NWM, the orography-related errors are minimized. For instance, during the data assimilation of ERA-Interim, ERA5, and the current European Centre for Medium-Range Weather Forecasts's (ECMWF) operational model, different orographies are being utilized. This results in different georeferencing in the height dimension of the data retrieved (temperature, specific humidity, pressure, etc.), even when the spatial resolution is identical. For the investigations carried out herein the spatial resolution of both the orography employed during the data assimilation for the generation of the NWM fields, and the orography at which the data are retrieved, is crucial. Despite the fact that series extracted from the surface fields thereof are relatively biased as a function of the height differences between them, when extracting from the model levels there is no bias. Proof that employing the surface fields is not appropriate for geodetic analysis has been given by Heinkelmann, Balidakis, et al. (2016). This indicates that for the current purpose it is not necessary to employ reanalysis products that have a latency of several months, but the homogenization process can be conducted operationally, as is for the tropospheric combination carried out in the framework of the International VLBI

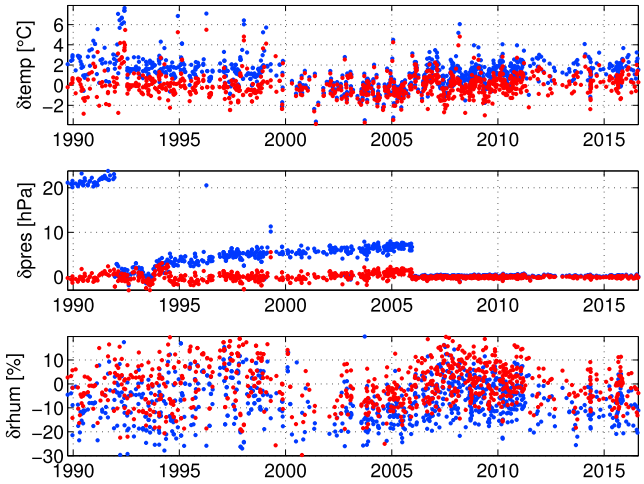

Figure 1. The differences between the raw in situ meteorological data recorded at HOBART26 and ERAinML (blue) and the differences between the homogenized in situ meteorological data and ERAinML (red).

Service for Geodesy and Astrometry, hereinafter IVS (ftp://ivs.bkg.bund.de/pub/vlbi/ivsproducts/trop/). To confirm that the inhomogeneities stem from the in situ records rather than from ERAinML, we employed series from the model levels of Modern-Era Retrospective Analysis for Research and Applications, version 2 MERRA-2, (Gelaro et al., 2017).

The detection of inhomogeneities is also possible without utilizing reference series, provided that they are observed regularly (e.g., Killick et al., 2012). Nevertheless, in addition to the fact that employing reference time series to perform the homogenization facilitates the detection of change points, in the context of the current study such an approach cannot be applied to all stations owing to the sparsity of the time series.

3. Development of the Potsdam Mapping Functions and Gradients

The propagation delay that microwave signals experience traversing the electrically neutral atmosphere is a function of the time-dependent atmospheric density field, and receiver and source location. At a given epoch, the total delay at a station is approximated by a function of azimuth angle α and elevation angle ε:

$$\begin{aligned} d_{\text{atm}} = {} & mf_h(\varepsilon) d_h^z + mf_{\text{nh}}(\varepsilon) d_{\text{nh}}^z \\ & + mf_g(\varepsilon) \left(G_{\text{NS}} \cos(\alpha) + G_{\text{EW}} \sin(\alpha) \right. \\ & \left. + G_{\text{NN}} \cos^2(\alpha) + G_{\text{NE}} \sin(\alpha) \cos(\alpha) + G_{\text{EE}} \sin^2(\alpha) \right), \end{aligned} \tag{1}$$

where the subscript h stands for hydrostatic and nh for nonhydrostatic; d_i^z denote the zenith delays; G_{NS} and G_{EW} are the total linear horizontal delay gradients; and G_{NN}, G_{NE}, and G_{EE} are the second-order gradients. The hydrostatic delay is mainly induced by dry gases, whereas the nonhydrostatic arises mainly from water vapor and hydrometeors. We employ higher-order terms because retaining only first-order (O^1) gradients implicitly assumes that the refractivity only varies linearly in the horizontal directions, which is not necessarily a sound assumption. This point is illustrated in Figure 2, where a weather front over Tsukuba (Japan) was captured by ERA-Interim. One should notice that while the differences between ray-traced delays and assembled delays (equation (1)) employing gradients of first and second orders at $\varepsilon = 5°$ could reach a couple of millimeters, omitting the second-order terms can induce modeling errors as large as 10 cm. The improvement stemming from employing second-order gradients is usually larger for the slant nonhydrostatic delay.

Employing second-order gradients estimated from state-of-the-art NWMs, such as ERA-Interim and MERRA-2 in the analysis of space geodetic observations, potentially eliminates the need for parameter space expansion to better describe weather variations. This is not the standard case neither in meteorology nor in geodesy but a rather new concept. In space geodetic data analysis the estimated parameters as well as the intervals thereof are to a certain extent determined by the lack of understanding of the dynamics of the individual effects (e.g., weather and Earth orientation). In most cases, applying the fully fledged PMF in VLBI data analysis without the estimation of zenith delays and gradients —that is fixing them to a priori—yields a WRMS in the residuals as low as when these parameters are estimated without applying the PMF but solving for these parameters,

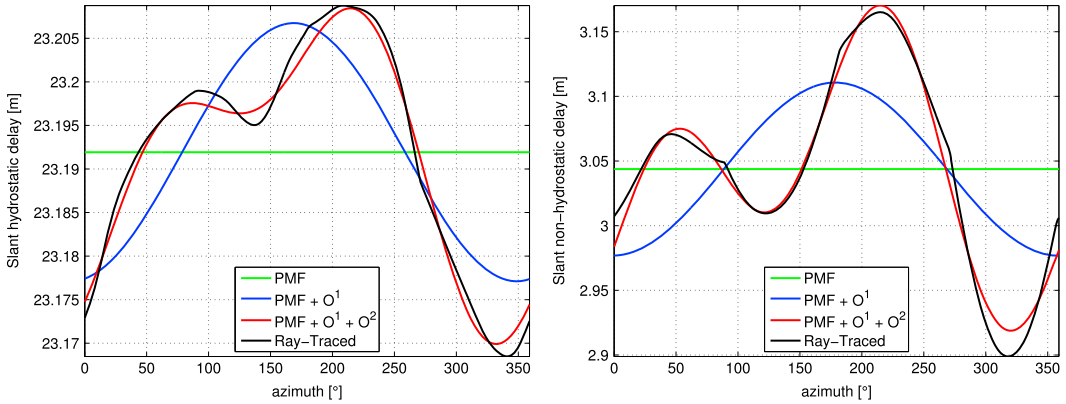

Figure 2. Slant hydrostatic (left) and nonhydrostatic (right) delays at 5° by different modeling approaches: PMF (green), PMF and linear horizontal gradients (blue), PMF and high-order gradients (red), and ray-traced delays (black) calculated employing data from the model levels of ERA-Interim. The station shown is Tsukuba (Japan) on 15 August 2003 06:00. PMF = Potsdam mapping functions.

at the same time reducing the number of estimated parameters at least by 80%. The presentation of our results in this regard exceeds the scope of this work. Further details on the calculation of PMF are provided in Appendix A.

To compare the symmetric delay components alone, we generate PMF on the orography utilized by the Vienna Mapping Functions 1, hereinafter VMF1 (Böhm et al., 2006). For the comparison to be fair, the same d_h^z and d_{nh}^z were utilized, so that the discrepancies between the mapping functions will be revealed. Comparing the symmetric part alone yields considerable differences. Figure 3 displays the differences expressed in terms of the expected relative height errors (Böhm, 2004) ensuing a VLBI data analysis where the mapping functions were either VMF1 or PMF. For that, we assume that we have a VLBI station installed at every grid node that observes at elevation angles down to 5°. The discrepancies (usually below 3 mm for the hydrostatic and below 5 mm for the nonhydrostatic) mainly stem from the different radii of curvature (1–2 mm at the equator and polar regions), the underlying NWMs and the spatial resolution thereof, the ray tracing algorithm, the different refractivity (and other physical) constants, and the parameterization of *a*, *b*, and *c* rather than for *a* alone.

The poor observation geometry especially during the early VLBI sessions as well as the intensive sessions should be ameliorated by imposing suitable constraints on a priori gradients (e.g., Nilsson et al., 2017). Since these constraints should be fairly tight, the a priori gradients should be of high quality. To assess the accuracy of PMF, we calculated the assembled atmospheric delays (equation (1)) for VLBI observations spanning the period 1999–2014 employing the current development, VMF1 together with GRAD (Landskron et al., 2017), and ray-traced delays by the Goddard Space Flight Center (GSFC) VLBI group (Eriksson et al., 2014; https://vlbi.gsfc.nasa.gov/services/tropodelays/tropodelays/) based in the Goddard Earth Observing System Model (GEOS) 5.9.1 tavg3_3d_asm_Nv FP-IT model

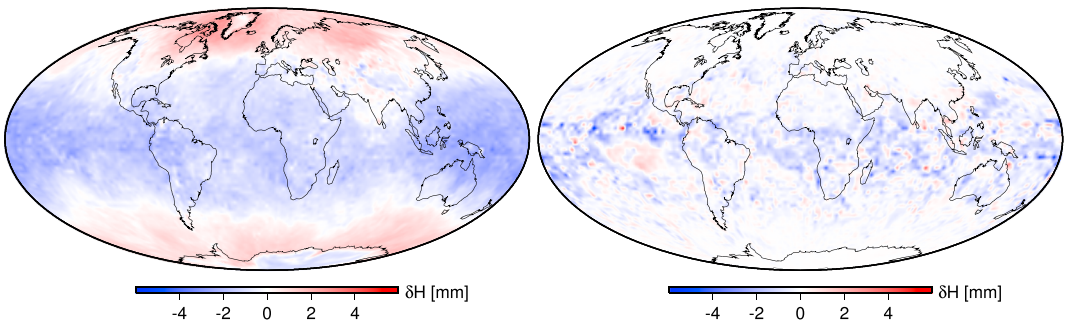

Figure 3. Expected equivalent relative height errors ensuing VLBI data analysis where the mapping functions are either VMF1 or PMF (1 January 2017 00:00). Shown are the differences due to the hydrostatic (left) and the nonhydrostatic component (right). PMF = Potsdam mapping functions; VLBI = very long baseline interferometry.

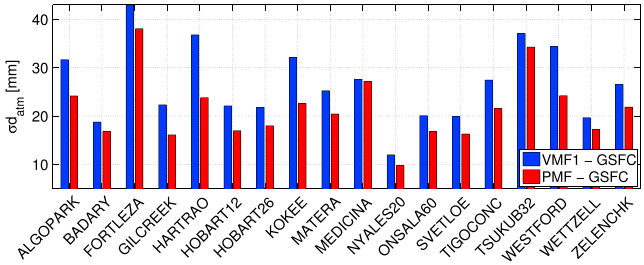

Figure 4. Comparison of the total atmospheric delays assembled (symmetric part and asymmetric part to the first order) at the actual (ε, α, t) of all VLBI observations within the time span 1999–2014 between PMF, VMF1+GRAD, and ray-traced delays by the GSFC VLBI group. Shown is the standard deviation of the differences between assembled delays and the ray-traced delays at the stations that participated in more than 50,000 scans. GSFC = Goddard Space Flight Center; PMF = Potsdam mapping functions; VLBI = very long baseline interferometry.

(https://gmao.gsfc.nasa.gov/products/GEOS-5_FP-IT_details.php). The latter serve as the reference as in (Eriksson et al., 2014) a performance better than VMF1 was proven. To perform the interpolation from the regular 6-hourly PMF and VMF1 series and irregular GRAD series in the time domain, we employed Lagrange polynomials. Figure 4 illustrates the scatter of the series of the differences between the assembled slant delays from PMF and VMF1 and the ray-traced delays from GSFC. In every case PMF is closer to the ray-traced delays, reducing the scatter by 19% on average.

4. Development of GFZ-PT

To describe meteorological parameters (in the broad sense) that are interesting for geodetic applications as harmonic functions instead of discrete series, we have developed GFZ-PT. Compared with other empirical models (e.g., GPT2w, Böhm et al., 2015), GFZ-PT differs in that it additionally features high-frequency harmonic terms, robustly estimated rates, and harmonic functions of gradient components estimated rigorously from ray-traced delays. In particular, it has harmonic coefficients at annual (S_a), semiannual (S_{sa}), terannual (S_{ta}), diurnal (S_1), semidiurnal (S_2), and terdiurnal (S_3) frequencies. Figure 5 illustrates the amplitudes for the S_1 pressure and temperature waves simulated by GFZ-PT. For instance, the barometric tides' contribution to d_h^z induces peak-to-peak systematic variations as large as 13 mm, prompting that the additional parameters featured in GFZ-PT are essential. Figure 6 shows the average linear gradient components (top row) and the annual amplitudes thereof (bottom row). For G_{NS} the largest amplitudes are found in the tropics, whereas for G_{EW} they are found in coastal areas. Previous attempts to model the gradient vector components provided only offsets either in a rather low spherical harmonic expansion, that is 9 for Böhm et al. (2013), or only at selected VLBI sites (MacMillan, 1995; MacMillan & Ma, 1997). In certain locations (e.g., the tropics) the magnitude of the S_a variations can be as large as 40% of the time average of the signal, a fact that renders the inclusion of seasonal signals more precise. The spatial resolution of GFZ-PT allows G_{EW} orographic features to be resolved, a fact that is not possible via a spherical harmonics representation. Also the fact that the gradient components, the physical meaning of which involves the spatial partial derivatives of the zenith delays, in previous models are modeled as offsets without a trend implies that there are either no trends in the zenith delays or the trends thereof are globally constant, which is untrue. To determine the long-wavelength variations (offset, rate, and seasonals), 38 years of ERA-Interim

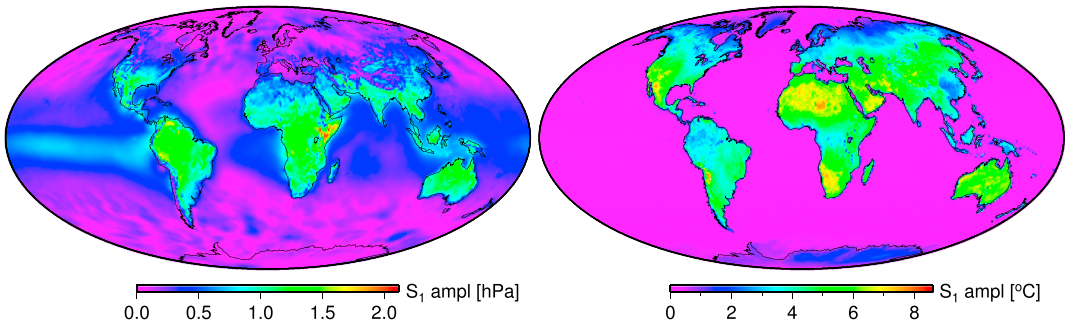

Figure 5. The amplitude of the diurnal (S_1) pressure (left) and temperature (right) fluctuations, estimated from ERA5.

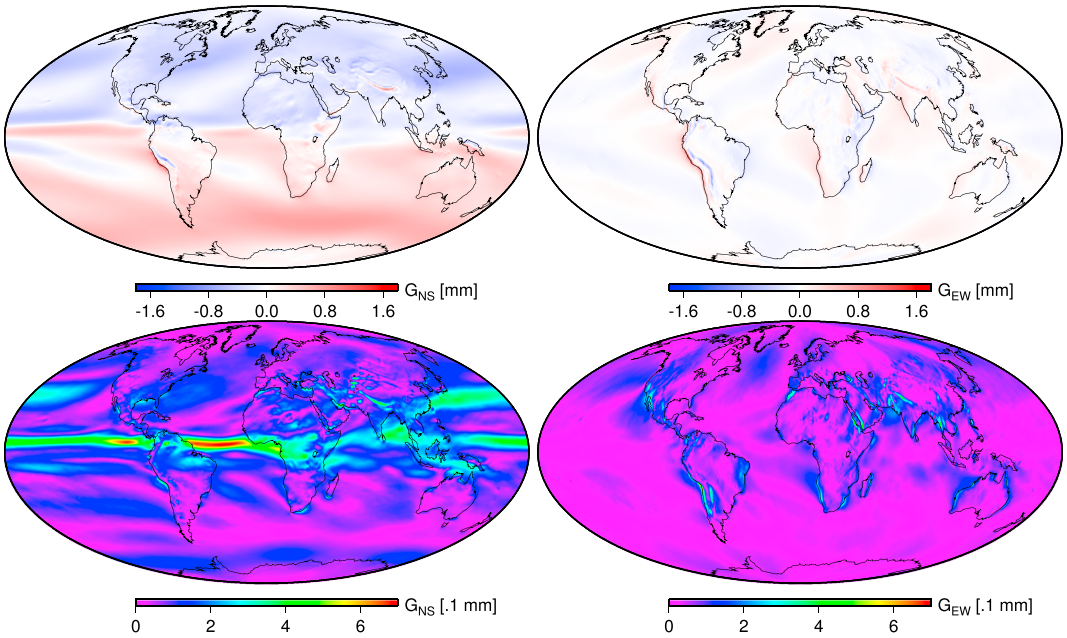

Figure 6. Average G_{NS} (left column) and G_{EW} (right column) gradient components (top row), and the S_a amplitudes thereof (bottom row), from ray tracing in ERA-Interim.

(0.5° × 0.5° on 60 model levels) were employed, and to estimate the high-frequency terms, 7 years of hourly ERA5 (0.25° × 0.25° on 137 model levels) data were employed. GFZ-PT includes zenith delays, mapping function coefficients (*a*, *b*, and *c*), gradient components of the first and second orders, pressure (necessary for the calculation of d_h^z), temperature (necessary to account for thermal deformation of VLBI antennas), and relative humidity and water vapor-weighted mean temperature (equation (4)) for climate studies. Further details on the calculation of GFZ-PT are provided in Appendix B.

Since a portion of the seasonal variations (mainly S_a) the VLBI height (and baseline length) time series experience is of thermal origin, we elaborate further upon the temperature component of GFZ-PT. Thermal deformation of VLBI antennas if not corrected for induces an annual signal to the scale of the implied reference frame as large as 1 mm (e.g., Wresnik et al., 2007). The transition to VLBI Global Observing System, with smaller antennas, will attenuate this systematic effect. To achieve 1-mm precision, this effect should be accounted for. To determine the number of terms of the temperature harmonic expansion of GFZ-PT that are necessary for the reduction of VLBI observations, the time delay of the thermal deformation of the antennas's elements was evaluated under different scenarios. Alternating the reference temperature (GFZ-PT over Global Pressure and Temperature, GPT) introduces virtual displacements in the radial coordinate component that are biased with less than 0.2 mm (e.g., BADARY and HOBART26) and have an RMS well below 0.2 mm for most stations. Employing GFZ-PT in lieu of in situ temperature does not introduce bias at any of the stations. Neglecting the S_{sa} and S_{ta} waves rarely introduces relative errors larger than 0.1 mm; hence, they may be neglected, if a target precision of 0.1 mm is to be met. Omitting the S_1 term introduces a relative error that fluctuates above the 0.15-mm level at most stations; thus, it should be retained. The vertical deformation induced by overtones of the main short-period temperature wave (S_2 and S_3) is below 0.05 mm, so they may be dismissed.

For the investigations conducted in the framework of this study, the complete harmonic expansion of GFZ-PT was used.

5. Calculation of IWV From NWMs

Despite the remarkable accuracy of IWV retrieved from VLBI and GNSS, relying on these methods alone does not allow to infer climate signals at regions lacking such equipment having been installed decades beforehand and meticulously maintained ever since. To address this, we avail ourselves to NWMs.

We explored five approaches to extract IWV given a NWM, ERA-Interim herein:

1. ERAinSFC,
2. ERAinML,

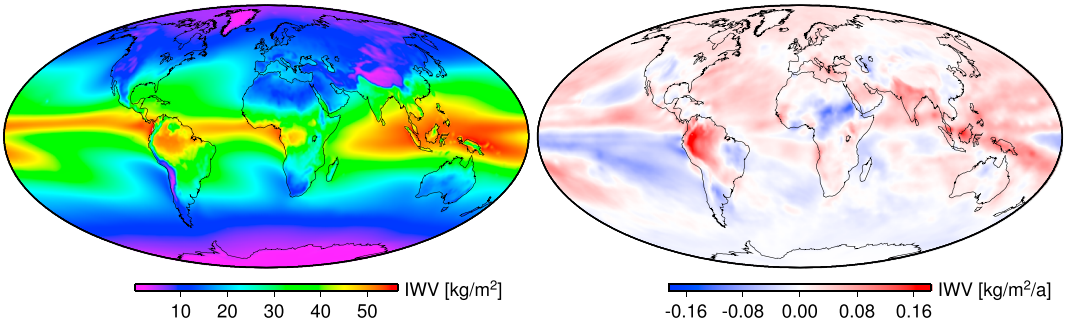

Figure 7. Average IWV (left) and the linear rates (right) of series spanning the era of geodetic VLBI (1979–2017), calculated by numerical integration along the model levels of ERA-Interim (ERAinML). IWV = integrated water vapor; VLBI = very long baseline interferometry.

3. RT(NI),
4. RT(GFZ-PT), and
5. RT(Bevis92).

ERAinSFC and ERAinML consist of performing numerical integration from the surface of the model and from the antenna reference point of the station of interest, respectively:

$$\mathrm{IWV} = \frac{1}{g}\int Q\mathrm{d}P, \tag{2}$$

where g denotes the WMO-defined gravity acceleration, Q the specific humidity, and P the total pressure. Additionally, we have retrieved IWV from ray-traced d^z_{nh}, alternating the water vapor-weighted mean temperature source, T_m. The transfer function reads

$$\mathrm{IWV} = \frac{10^6}{\left(k'_2 + \frac{k_3}{T_m}\right)R_w}\mathrm{d}^z_{\mathrm{nh}}, \tag{3}$$

where k'_2 and k_3 are refractivity constants and R_w is the specific gas constant of water vapor. We have obtained T_m by performing numerical integration in the model levels (RT(NI), equation (4)), by a harmonic function (RT(GFZ-PT)), and following Bevis et al. (1992) by an empirical function of surface temperature (RT(Bevis92)).

$$T_m = \frac{\int \frac{P_w}{T} Z^{-1}\mathrm{d}h}{\int \frac{P_w}{T^2} Z^{-1}\mathrm{d}h}, \tag{4}$$

where P_w is the partial pressure of water vapor, T is the temperature, Z^{-1} is the inverse compressibility factor, and h is the geodetic height. Theoretically, being the most straightforward, ERAinML is the most rigorous approach; hence, it serves as the reference. Figure 7 depicts the linear rates of IWV estimated from the ERAinML series. Here we compare rates of IWV series obtained at the positions of all VLBI stations, assuming uninterrupted observations during 1979–2017.

Solving for the linear rate as well as seasonal signals (e.g., S_a) within a weighted least squares adjustment, we found significant (1σ level throughout this paper) relative rate errors for 7.5% of the stations with respect to ERAinSFC, 4.5% with respect to RT(NI), 4.0% with respect to RT(GFZ-PT), and 3.5% with respect to RT(Bevis92). As expected, due to the finite spatial representativeness of the underlying NWM, the largest differences (30 $\mathrm{g\cdot m^{-2}\cdot a^{-1}}$) were found between ERAinML and ERAinSFC at regions with steep orographic gradients. For instance, the height of KOKEE and KP-VLBA in the surface of ERA-Interim is more than 1-km smaller compared to the actual one. For MK-VLBA the height difference exceeds 3 km. Only a 20% of the VLBI stations have a height difference larger than 250 m. Nevertheless, these discrepancies will be abridged with the ever improving spatial resolution of state-of-the-art NWMs. Proof for that is provided by ECMWFs operational model the spatial resolution of which gradually changed from 79 km (the data assimilation system employed by ERA-Interim) to the current of 9 km. Among else, this improvement is reflected on the reduction of the aforementioned height discrepancies. The rate errors estimated with the least squares approach are deflated

Table 1
Percentage of Stations That Exhibit Differences Between the Aforementioned Approaches to Calculate IWV ($g{\cdot}m^{-2}{\cdot}a^{-1}$) and ERAinML, Within a Bin

$g{\cdot}m^{-2}{\cdot}a^{-1}$	ERAinSFC		RT(NI)		RT(GFZ-PT)		RT(Bevis92)	
< 2.5	81.1	66.2	71.1	40.8	71.6	44.3	77.6	42.2
[2.5, 5)	12.4	19.4	22.9	26.9	21.4	27.9	16.9	32.3
[5, 10)	6.5	9.5	5.0	27.4	6.0	23.4	4.5	20.9
[10, 20]	0.0	2.0	1.0	4.5	1.0	4.0	1.0	4.0
> 20	0.0	3.0	0.0	0.5	0.0	0.5	0.0	0.5

Note. Two rate estimation approaches were utilized: least squares accounting for systematic harmonic variations (left) and MIDAS (right).

due to the fact that we did not account for the pervasive autocorrelation in the IWV time series; that is, the poleward increasing autocorelation in the IWV series reduces the number of independent observations, thus yielding overoptimistic rate error estimates. To obtain a more realistic estimate of the IWV rate errors, we employed a variant of the MIDAS (median interannual difference adjusted for skewness) estimator (equation (B2); e.g., Blewitt et al., 2016). Considering more realistic errors, only at 3.0% of the stations we find significant trend differences with respect to ERAinSFC, whereas only at SC-VLBA (St. Croix, USA) do the d^z_{nh}-based data sets display a statistically significant difference. Table 1 illustrates the binned differences. We note that the agreement between the different approaches seems to deteriorate when MIDAS is employed. However, due to the fact that rate errors are approximately 4 times larger, the differences cannot be regarded as important.

As long as the height assigned to the projection of a site in the surface of a NWM is not too different from the actual height, we can conclude that the IWV rate differences obtained between the approaches tested herein can be neglected.

6. VLBI and GPS Data Analysis

The least squares module of the VieVS@GFZ VLBI software (Nilsson et al., 2015) was utilized to analyze interferometric group delay data from all global VLBI multibaseline sessions (5,870) observed in the framework of IVS (Nothnagel et al., 2015), from 1979 to 2017. Employing the modeling approaches outlined in sections 2–4, two series of solutions were produced alternating the meteorological data source and the mapping function, respectively. The VLBI analysis performed here abides by the International Earth Rotation and Reference Systems Service Conventions (2010; Petit & Luzum, 2010), with the exception of utilizing nontidal geophysical loading models generated by GFZ Potsdam (Dill & Dobslaw, 2013) at the observation equation level.

Alternating the pressure and temperature source utilized for the VLBI analysis, four solutions were generated:

1. GFZ-PT (cf. section 4);
2. ERAinML, series extracted from the model levels of ERA-Interim;
3. RawInSitu, in situ (series extracted from the meteorological sensors); and
4. HomogInSitu, homogenized in situ employing ERAinML (cf. section 2).

The parametrization of PMF is more precise than the parametrization of VMF1 (cf. Appendix A). Therefore, we utilize PMF in all four solutions. Alternating both the hydrostatic and nonhydrostatic mapping functions, three solutions were generated:

1. GFZ-PT (cf. section 4),
2. VMF1, and
3. PMF (cf. section 3).

We expect that the HomogInSitu meteorological data set is conceptually the most appropriate; therefore, we utilize HomogInSitu in all three solutions. In total six VLBI solutions were produced.

For the GPS-based products (estimated zenith delays and the uncertainties thereof), GFZ's solution for the second Tide Gauge Benchmark Monitoring was used, where VMF1 and GPT2 were employed. It was generated employing the least squares module of GFZ's GNSS analysis software EPOS.P8 (Earth Parameter and Orbit

determination System). For analysis options as well as the processing strategy, the interested reader is referred to Deng et al. (2016).

We explicitly mention that to avoid tectonic plate motion manifesting into IWV trends, coordinate corrections were estimated for all VLBI solutions performed herein.

7. Geodetic Analysis Results

Since there is no absolute measure to distinguish the best between similar IWV series, we resort to assessing the geodetic results to infer the superiority of a solution over another.

7.1. VLBI Baseline Length Repeatability

A unique quantitative measure of the quality of a multiyear solution is provided by studying the changes in the baseline length repeatability. The WRMS scatter of weighted baseline length estimates was calculated. Given a baseline length time series $\{b\}_{i=1}^{N}$, it is

$$\mathrm{WRMS} = \sqrt{\sum_{i=1}^{N} \frac{v_{b_i}^2}{\sigma_{v_{b_i}}^2} \Big/ \sum_{i=1}^{N} \frac{1}{\sigma_{v_{b_i}}^2}}, \tag{5}$$

where v_{b_i} are the baseline length residuals from a straight-line fit and $\sigma_{v_{b_i}}$ are the formal errors thereof.

PMF marginally improves the baseline length repeatability with respect to VMF1; 55% of the baselines show an improvement of 3% on average. With respect to GFZ-PT, PMF reduces the baseline length scatter in 71% of the baselines and 9% on average.

To assess the overall impact of the atmospheric delay modeling on the baselines, we submit the WRMS of the baseline length series to a fit over the lengths thereof using $y(x) = \sqrt{A^2 + (Bx)^2}$ as the fitting ansatz (e.g., MacMillan & Ma, 1994). In most cases the longer the baseline, the lower the elevation angles at which the stations at its ends observed, and therefore the larger the d_{atm}. Such a configuration yields larger atmosphere-propagation-related uncertainties. Of course, other effects (e.g., source structure) contribute to the increase of the WRMS scatter with increasing baseline length but herein any changes in the quadratic decay B stem from tropospheric modeling changes. The choice of y is common in VLBI data analysis. Considering all baselines with a sufficient number of observations, employing PMF in lieu of VMF1 slightly improves A by 1% and B by 0.5%. Employing PMF instead of the harmonic function GFZ-PT reduces A by 20% and B by 2%.

Alternating the meteorological data source has a larger impact. Employing homogenized series improves A by 9% and B by 3% compared to utilizing the raw records, whereas using ERAinML instead of harmonic functions improves A by 26% and B by 2%. Thus, in the absence of in situ meteorological records, if obtaining the most accurate results is intended, discrete series obtained from the model levels of accurate NWMs should be preferred to empirical models. If the pressure is recorded at the geodetic sites, the series should be checked for impurities before being employed for the reduction of geodetic observations.

7.2. Helmert Transformation

Let $\mathbf{X}$ and $\mathbf{X}'$ be the station coordinate estimates of a global network from two VLBI solutions. The L_2-norm estimate of the 3-D similarity transformation (also known as Helmert transformation in geodetic literature) between them reads (e.g., Kotsakis et al., 2014)

$$\begin{aligned} \hat{\theta} &= \mathbf{C}_{\hat{\theta}} \mathbf{J}^{\top} \left(\mathbf{C}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}'} \right)^{-1} \left(\mathbf{X} - \mathbf{X}' \right) \\ \mathbf{C}_{\hat{\theta}} &= \left(\mathbf{J}^{\top} \left(\mathbf{C}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}'} \right)^{-1} \mathbf{J} \right)^{-1}, \end{aligned} \tag{6}$$

where $\hat{\boldsymbol{\theta}} = \left[T_X \;\; T_Y \;\; T_Z \;\; R_X \;\; R_Y \;\; R_Z \;\; D \right]^{\top}$ contains the standard Euclidean similarity parameters: a translation vector in meter ($\mathbf{T} = \left[T_X \;\; T_Y \;\; T_Z \right]^{\top}$), an angular rotation matrix in radian ($[\left[R_X \;\; R_Y \;\; R_Z \right]^{\top} \times]$), and a scale factor in parts per billion (1ppb $\approx$ 6.4mm on Earth's surface) D. The design matrix is denoted by $\mathbf{J}$, the covariance matrices of the station coordinates estimated from the two different solutions are denoted by $\mathbf{C}_{\mathbf{X}}$ and $\mathbf{C}_{\mathbf{X}'}$, and the covariance matrix of the estimated similarity parameters is denoted by $\mathbf{C}_{\hat{\theta}}$. For the investigations carried out herein, we have considered fully populated covariance matrices.

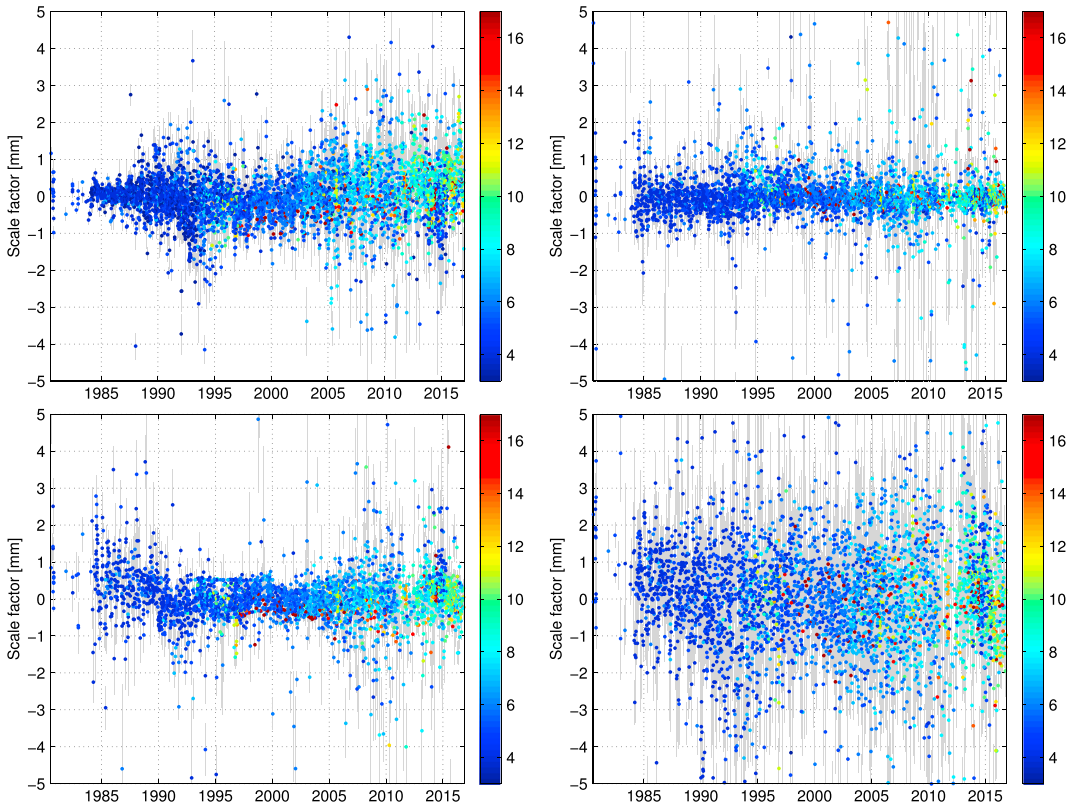

Figure 8. The scale factor of the session-wise seven-parameter Helmert transformation between the VLBI solutions employing VMF1 and PMF (top left). Additionally, shown is the scale between the VLBI solution where homogenized pressure and temperature were employed and raw in situ data (top right), ERAinML (bottom left) and GFZ-PT (bottom right). The dot color indicates the number of stations that participated in each session, and the length of the gray vertical lines indicates the uncertainty of the scale estimate. PMF = Potsdam mapping functions; VLBI = very long baseline interferometry.

To assess the impact of the mapping function on the implied reference frame, session-wise Helmert transformations were performed between the different solutions. The time series of the factor indicating the scale difference between the frame realized by the VMF1 and by the PMF session-wise solutions is illustrated in Figure 8 (top left) along with the scale differences between the VLBI solutions alternating meteorological data. To robustly assess the results from a statistical viewpoint, we employed the least median of squares estimator (e.g., Rousseeuw & Leroy, 2005) to calculate the time average of the scale series as well as the WRMS thereof. The scale from the Helmert transformation between the empirical model GFZ-PT and HomogInSitu yields the highest scatter from all pairs of solutions generated herein (1.8 mm). As Table 2 indicates, no long-term bias was introduced in any of the transformation parameters by alternating the mapping functions or the meteorological data. To ascertain that further upon, we performed the 14-parameter Helmert transformations from the respective global solutions (results not shown here). The scale rates do not exceed the 0.1 mm/a level. Employing erroneous models will have repercussions mainly in the form of unusually large scatter. For instance, employing meteorological data from the surface fields of ERA-Interim leads to a WRMS in the scale series of 6.7 mm (e.g., Heinkelmann, Balidakis, et al., 2016). For contrast, the transformation between the estimates of the VLBI data analysis and the coordinates of the secular International Terrestrial Reference Frame 2014 (ITRF14) results in a bias of −3.7 mm and a WRMS of 2.5 mm (cf. Table 2).

8. Comparison of IWV Trends From VLBI, GPS, and NWMs

To investigate the IWV trend differences between VLBI, GPS, and ERA-Interim rather than the impact of the sampling rate, all series were synchronized beforehand. As all GPS series and most VLBI series span less than a climate normal of 30 years and are not evenly sampled (mainly VLBI) we only consider the series that are longer than 10 years and experience statistically significant IWV rates. To ensure that trend differences between VLBI

Table 2
The Time Average of the Seven-Parameter Helmert Transformation Time Series and in Parenthesis the WRMS Thereof, Calculated Employing the Least Median of Squares Estimator

	T_x (mm)		T_y (mm)		T_z (mm)		R_x (mm)		R_y (mm)		R_z (mm)		Sc (mm)	
PMF → VMF1	0.0	(0.4)	−0.0	(0.6)	0.0	(0.7)	0.0	(0.6)	−0.0	(0.4)	0.0	(0.3)	−0.0	(1.1)
HmgInSitu → GFZ-PT	−0.0	(0.7)	−0.0	(0.8)	0.0	(1.1)	0.0	(0.7)	−0.0	(0.6)	−0.0	(0.5)	−0.0	(1.8)
HmgInSitu → RawInSitu	−0.0	(0.6)	−0.0	(0.8)	0.0	(0.9)	0.0	(0.7)	0.0	(0.5)	0.0	(0.5)	−0.0	(1.3)
HmgInSitu → ERAinML	0.0	(0.5)	−0.0	(0.7)	0.0	(0.8)	0.0	(0.5)	0.0	(0.4)	−0.0	(0.4)	−0.0	(1.2)
PMF → ITRF14	0.2	(0.3)	−0.1	(0.0)	1.5	(1.0)	0.1	(0.5)	−0.1	(0.0)	0.0	(0.1)	−3.7	(2.5)

Note. The rotation parameters and the scale have been multiplied with the average radius of the Earth to yield deformation of the surface of the Earth. PMF = Potsdam mapping functions.

and GPS do not stem from differences in the sensor position (mainly the height difference), atmospheric ties were applied to the estimated d^z_{nh} from GPS and were referred to the collocated VLBI station. Due to the fact that the d^z_{nh} tie between stations with a height difference of dozens of meters displays an S_a harmonic term as large as 0.5 mm, no empirical function was used but the differences between the ray-traced d^z_{nh} at the VLBI and GPS stations, as applied by Heinkelmann, Willis, et al. (2016).

Synchronizing IWV series of VLBI with GPS in some cases strips VLBI of its main advantage, which condenses to their length and stability. That is, the length of both VLBI and GPS time series reduces to their common period. This means that some very long VLBI time series (e.g., almost 35 years for Wettzell) are significantly reduced by over 10 years. To exploit the full lengths of the VLBI time series, we first compare the VLBI-derived IWV trends to those from ERA-Interim, which are also available from 1979 (section 8.1). Afterward, we compare the trends from the VLBI-derived IWV series with the GPS-derived series (section 8.4). In terms of baseline length repeatability, the best VLBI solution was obtained when PMF together with homogenized in situ meteorological data were used; hence, the IWV series calculated from the related d^z_{nh} are used for the ensuing comparisons. To transform d^z_{nh} series to IWV, we employ water vapor-weighted mean temperature calculated following equation (4).

8.1. Comparison of IWV Rates From VLBI and ERA-Interim

We synchronized ERA-Interim with VLBI by interpolating the ERAinML IWV series (equation (2)) at the epochs of the VLBI-derived zenith delay estimates employing Lagrange polynomials. Having synchronized ERA-Interim with VLBI, 90% of the reliable IWV series extracted from both sets experience statistically significant trends (1σ level). In 24% of the cases the differences are nonnegligible. Some stations shown in Figure 9 display unexpectedly large trend differences. HRAS 085 (Fort Davis) — as well as PIETOWN (USA) — is a station that had experienced unexplained nonlinear motion, thereby does not participate in the No Net Translation/No Net Rotation conditions for the datum definition. Nevertheless, due to the fact that d^z_{nh} was always estimated together with the coordinates in a single adjustment, the IWV trend should not be affected. Moreover, the recorded pressure series has suffered multiple jumps and drifts, and the temperature time series had to be segmented thrice and offsets be introduced in the homogenization process (−3.4, −2.7, and −0.9 °C). The IWV rate at HRAS 085 estimated from the different VLBI solutions carried out herein is around $-0.5\ \mathrm{kg\ m^{-2}\ a^{-1}}$, which is not realistic; hence, it is excluded from our considerations hereafter. Another station that displays unusually large trends is SESHAN25 (Shanghai, China). The in situ series feature three jumps in pressure larger than 1 hPa and a 1 °C jump. The reason that the trends at ZELENCHK (Zelenchukskaya, Russia) are quite large mainly stems from the rather short length and from the 11 hPa and the 2 °C jumps. In the temperature series of KOKEE (Kauai, Hawaii, USA), four offsets were introduced, as well as a single 1.6-hPa offset. Nevertheless, we ascribe the different trend sign at KOKEE (i.e., site on an island) to the finite resolution of the NWM, as well as to the trend being very close to zero.

As discussed in section 8.2, the good agreement shown in Figure 9 is conditional upon the use of appropriate meteorological data to mitigate the tropospheric delay.

The fact that the relative errors between VLBI and ERA-Interim are not even smaller is mainly due to the fact that the tropospheric parameters from NWM provide a "snapshot" of the atmospheric state at the epochs the solutions refer to, which does not hold for the batch least squares processing carried out here. In addition, although it is not expected of climate signals to have steep spatial gradients, microclimate features detectable

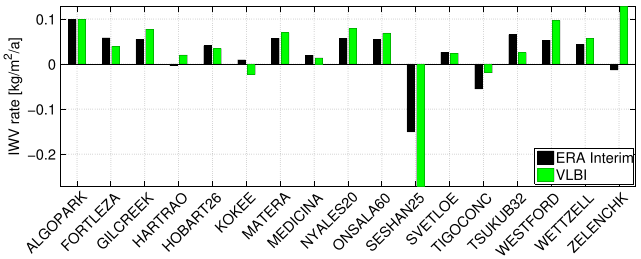

Figure 9. IWV rates from ERA-Interim and VLBI. Shown are the stations with sufficiently long observation record that yields statistically significant trends. IWV = integrated water vapor; VLBI = very long baseline interferometry.

by techniques such as VLBI and GNSS may not be discernible in the fields of a meso-β scale NWM such as ERA-Interim, owing to the finite resolution of the latter. In point of fact, series of residual tropospheric parameters estimated using space geodetic observations describe the water vapor distribution bounded in an inverse cone. For instance, a geodetic midlatitude station at sea level where an elevation mask below 7° is applied, senses water vapor redistribution limited by an inverse cone with a base radius of 100 km. This is the usual case for the analysis of GNSS observations because this technique is subject to multipath scattering and phase center variations. The directional antennas utilized for VLBI observations render the adoption of an elevation mask or an elevation-dependent weighting strategy inessential, in this context. Accounting for the fact that 90% of water vapor resides within 5 km from sea level, the effective radius is limited to less than 40 km, that is well below the native spatial resolution of ERA-Interim. Thus, the relative errors could be partly attributed to the large-scale spatial representativeness of the model.

Comparing the rates of the IWV series from GPS with those from the model levels of ERA-Interim shows a good agreement. Disregarding the stations with relatively short data span or dubious rates, we find IWV rate differences larger than the formal errors thereof in 28% of the stations. Seventy-three percent of the GPS stations show the same trend sign as ERA-Interim, and in 81% of the GPS and 77% of the ERA-Interim series the trends are positive.

We note that should the common approach—that is performing a least squares adjustment accounting neither for autocorrelation nor for seasonal signals—for the calculation of IWV trends be adopted in lieu of the current, the rate errors are deflated and so does the number of statistically insignificant differences. The Theil-Sen variant adopted here (equation (B2)) implicitly accounts for the autocorrelation in the series (Blewitt et al., 2016).

8.2. Impact of Meteorological Data on VLBI-Derived IWV Trends

So far, we have seen that alternating the meteorological data source impacts on the baseline length repeatability (section 7.1) and distorts the underlying reference frame (section 7.2). Figure 10 (top) illustrates the statistically significant IWV rates estimated varying the input meteorological data.

Employing ERAinML instead of HomogInSitu induces significant differences in 22% of the stations. The empirical model GFZ-PT induces significant relative errors with respect to HomogInSitu in only 11% of the stations. Sixty-one percent of the RawInSitu-derived IWV trends are significantly different from those estimated from the HomogInSitu data set.

Nevertheless, despite the fact that the estimated d^z_{nh} differs considerably, their sums with their respective d^z_h (the a priori d_{atm} in the VLBI adjustment) are very similar. This stems from the d^z_{nh} estimates trying to compensate for impurities in the d^z_h, as well as the fact that mf_h and mf_{nh} are similar. For instance, utilizing the in situ pressure at Sejong (South Korea) during 2015 in VLBI data analysis yields negative d^z_{nh}. From a physical viewpoint this is absurd but from a mathematical it is expected; the recorded pressure had a 45-hPa positive offset that led to very high a priori d^z_h and consequently very small estimated d^z_h. Subtracting reasonably accurate d^z_h (e.g., ERAinML) from Sejong's zenith total delays yields meaningful d^z_{nh} and eventually IWV. Of course, using the pressure source, we think is the best already in the VLBI analysis, would give the best result. Then, we would not have to apply any corrections a posteriori. However, in some cases we may have a VLBI solution available but using the wrong a priori pressure. For operational VLBI analysis, the nonhomogenized recorded pressure at the sites is usually used, while GNSS analysis typically applies an empirical model in the fashion of GPT2. If we do not want to reprocess the data (it is not a huge effort for VLBI, but it is for GNSS), we can try to

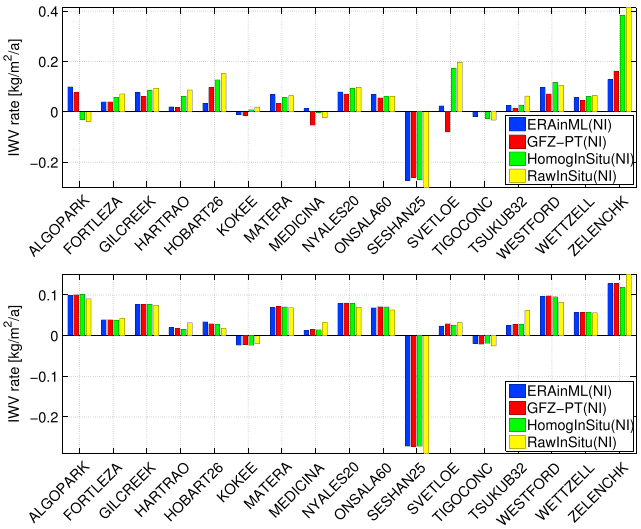

Figure 10. IWV rates from the solutions where the meteorological data were alternated, employing T_m from numerical integration (NI). The rates in the top panel were calculated utilizing the d^z_{nh} estimates from the VLBI analyses directly, whereas the rates in the bottom panel stem from the subtraction of a d^z_h set constant over all solutions from the zenith total delays from the VLBI analyses. Shown are the stations with sufficiently long observation record that yields statistically significant trends. IWV = integrated water vapor; VLBI = very long baseline interferometry.

correct the estimated d^z_{nh} a posteriori, using the described approach. The question, which we investigate by varying the meteorological input data source, is how this approach affects the results. Figure 10 (bottom) illustrates the IWV rates estimated after the same d^z_h (from ERAinML) was subtracted from the zenith total delays of each solution. The differences are mostly reconciled, but we still find differences as large as 0.04 kg m^{-2} a^{-1}. In the geodetic adjustment, the estimated d^z_{nh} is highly anticorrelated with the radial coordinate component and the clock offset. Owing to that and the mf_{nh} being slightly larger than mf_h impurities in the a priori d_{atm} cannot be fully compensated. Although there are no significant differences between the HomogInSitu, ERAinML, and GFZ-PT solutions, at 24% of the stations the IWV rate differences between HomogInSitu and RawInSitu are detectable.

This result highlights the paramount importance of utilizing homogenized pressure series to account for the atmospheric propagation delay.

8.3. Impact of Mapping Functions on VLBI-Derived IWV Trends

Alternating the mapping functions in the VLBI data analysis between PMF, VMF1, and the empirical GFZ-PT yields negligible difference in the estimated IWV rates (cf. Figure 11). The relative errors from the intercomparison are below 4 kg m^{-2} a^{-1}. Of course, we cannot state that this result applies to all mapping functions, since the development of all three included either operational analysis (VMF1) or reanalysis fields (ERA-Interim for both PMF and GFZ-PT and ERA5 for GFZ-PT) from ECMWF. Nevertheless, comparing the PMF in its full expansion (symmetric and asymmetric of first and second orders) against ray-traced delays from GEOS 5.9.1 tavg3_3d_asm_Nv FP-IT model yields relative equivalent height errors an order of magnitude smaller than

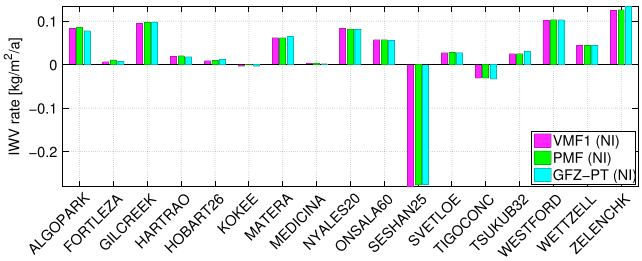

Figure 11. IWV rates from the solutions where the mapping functions were alternated. Shown are the stations with sufficiently long observation record that yields statistically significant trends. IWV = integrated water vapor; PMF = Potsdam mapping functions.

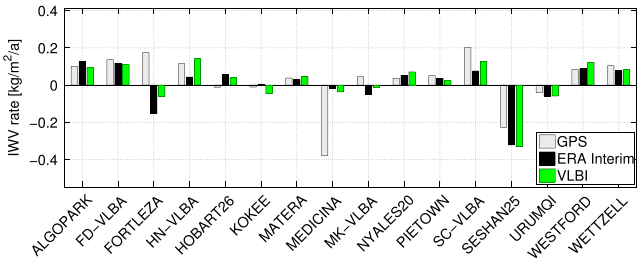

Figure 12. IWV rates from GPS, ERA-Interim, and VLBI. Shown are the stations with sufficiently long observation record that yields statistically significant trends. IWV = integrated water vapor; VLBI = very long baseline interferometry.

those stemming from the comparison of PMF with GFZ-PT. Thus, our conclusions regarding the impact of alternating mapping functions in VLBI data analysis on linear IWV trends are reliable.

8.4. Comparison of IWV Rates From GPS, VLBI, and ERA-Interim

Figure 12 provides a dipiction of the trends estimated from GPS, VLBI, and ERA-Interim after the related IWV series were synchronized. Although most VLBI stations are collocated with more than one GPS stations that could serve the study, we opt for retaining only the one with the longest span of parallel observations with VLBI. Comparing Figures 12 to 9, one notices that IWV trends at the same stations can be very different (e.g., at Fortaleza where a sign different from the GPS series is obtained). This is due to the IWV trends being sensitive to the sampling rate as well as the data span.

Despite the seemingly good agreement, the IWV trends are quite different at some sites. Forty-four percent of the differences between GPS and ERA-Interim, 22% of the differences between GPS and VLBI, and 22% of the differences between VLBI and ERA-Interim are significant. To exploit the complete length of the GPS-derived tropospheric parameters, we interpolated IWV from ERAinML at the epochs of the GPS estimates. The comparison between GPS and ERA-Interim does not reveal an agreement as good as the comparison between VLBI and ERA-Interim, since significant differences were found in more than 29% of the stations.

9. Conclusions

The results show that the long-term IWV trends from VLBI are of good quality. Since some VLBI stations now have been observing for more than 30 years, the data from these stations can be used for monitoring climate change, and to validate and improve climate models. Furthermore, many GPS stations now provide time series longer than 20 years, thus approaching a length interesting for climate studies. Since dense GPS networks exist in many areas, this technique will certainly get more interesting for climate research in the near future. However, since many GPS time series contain frequent jumps due to, for example, antenna changes, it will be crucial to homogenize these series by comparing with other techniques, such as VLBI.

The zenith hydrostatic delay, being mainly a function of pressure, is subject to inhomogeneities introduced by the related observations. We have addressed this issue by homogenizing all in situ meteorological records available by IVS employing the penalized maximal t test and reference series from the model levels of ERA-Interim. We recommend the use of homogenized meteorological records in VLBI data analysis, employing model level data as a reference, because it improves the baseline length repeatability, and employing the raw observations distorts 44% of the estimated IWV rates. Nevertheless, applying more appropriate hydrostatic delays a posteriori remedies for most cases (cf. section 8.2).

The PMF (section 3) and GFZ-PT (section 4) have been introduced here. PMF features a more rigorous parameterization that renders it slightly more accurate than VMF1, in terms of the assembled delays themselves. GFZ-PT differs from previous empirical models because it additionally features high-frequency harmonic terms, robustly estimated rates, and harmonic functions of gradient components estimated rigorously from ray-traced delays. Although alternating the mapping function impacts on the scale of the networks realized by the coordinates estimated from the different VLBI data analysis setups (at the 1-mm level), it yields negligible relative errors in the estimated IWV rates. Hence, PMF, GFZ-PT, or VMF1 may be used interchangeably in this regard.

The five different approaches studied to obtain IWV given a NWM do not have a significant impact on the linear rates, regardless of the method adopted to estimate them (least squares or equation (B2)). Other variables

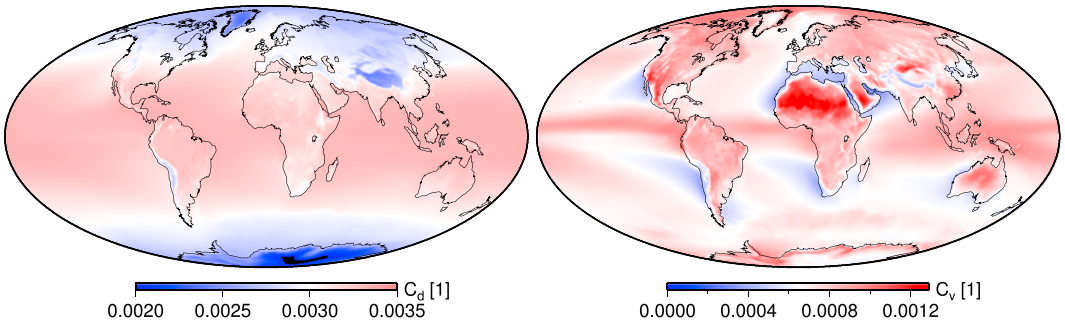

Figure A1. Hydrostatic (left) and nonhydrostatic gradient mapping function coefficients estimated from ray tracing in ERA-Interim. Proximity to white indicates proximity to the respective average values (0.00309 and 0.00075).

such as the time series length and the relaxation time for equation (B2) are more critical to the rate determination. In general, IWV rates are not linear; hence, further investigations should be conducted in this regard.

Comparing IWV rates from VLBI to ERA-Interim reveals a very good agreement when the complete series of the former are employed. The differences between GPS and ERA-Interim are smaller than the formal errors thereof at 72% of the cases. The discrepancies between IWV derived from space geodetic observations and NWM can stem from both issues with the NWM and errors in the analysis of the space geodetic data. For example, there could be issues related to the finite spatiotemporal resolution of the NWM or unmodeled loading effects in the space geodetic data analysis. To assess the validity of these presumptions, we look forward to the release of the hourly ERA5 reanalysis fields prior to 2010 by the fourth quarter of 2018, as well as to employing a JTRF14-like frame based on Kalman filter (e.g., Soja et al., 2016) for the geodetic analysis.

Appendix A: Potsdam Mapping Functions

The mapping functions mf_i and mf_g describe the elevation dependence of the symmetric delays (Herring, 1992) and the asymmetric delays (Chen & Herring, 1997), respectively. We adopted the continued fraction form (Marini, 1972) confined to third and first orders for $\sin(\varepsilon)$ and $\sin(\varepsilon)\tan(\varepsilon)$, respectively:

$$mf_i = \begin{cases} \dfrac{1+\dfrac{a_i}{1+\dfrac{b_i}{1+c_i}}}{\sin(\varepsilon)+\dfrac{a_i}{\sin(\varepsilon)+\dfrac{b_i}{\sin(\varepsilon)+c_i}}}, & \text{for } i = h \vee nh \\ \dfrac{1}{\sin(\varepsilon)\tan(\varepsilon)+C} & \text{for } i = g. \end{cases} \tag{A1}$$

Further expansions yield an insignificant precision improvement (below 1 mm at $\varepsilon = 3°$) at a limited number of extraordinary cases, but the additionally estimated coefficients are not statistically significant.

In this study, we employ ERA-Interim (6-hourly 0.5° × 0.5° fields on 60 model levels) and the in-house ray-trace algorithm developed at GFZ by Zus et al. (2012). We performed ray tracing at $\varepsilon = \left[\,3\ 5\ 7\ 10\ 15\ 20\ 30\ 50\ 70\ 90\,\right]^\circ$ and azimuth granularity $\delta\alpha = 15°$.

For every station (or grid node) at every epoch, we perform least squares fittings that result in the estimation of $\mathbf{x} = \left[\,\mathbf{a}\ \mathbf{b}\ \mathbf{c}\ \mathbf{C}\ \mathbf{G_{NS}}\ \mathbf{G_{EW}}\ \mathbf{G_{NN}}\ \mathbf{G_{NE}}\ \mathbf{G_{EE}}\,\right]^\top$ for the hydrostatic and nonhydrostatic components, separately. To keep the character of the mapping function coefficients in equation (A1) purely symmetric, the symmetric and asymmetric parameters are estimated separately. Estimating the mapping function coefficients rather than calculating them from a single ray at a low elevation (VMF1 concept) is computationally more expensive but is not subject to the systematics described in Zus, Dick, Douša, et al. (2015). To account for the representative errors of the delays at elevation angles close to the horizon, we apply a weighting scheme suggested by Zus, Dick, Heise, et al. (2015), $P_{ij} = \delta_{ij}\sin(\varepsilon_i)\sin(\varepsilon_j)$, where δ_{ij} denotes Kronecker's delta. The estimation of b, c, C, G_{NN}, G_{NE}, and G_{EE} causes a minimal increase in the computational overhead but results in a significantly better precision, expressed as a closer fit to the ray-traced delays. Estimating b and c in addition to a reduces the formal errors of a estimates at least 1000 times. In the presence of inversions, which occur more often to the nonhydrostatic component than to the hydrostatic, absolute constraints are imposed on the estimation of b and c. An alternative treatment assessed is applying an elevation-dependent weighting scheme

(e.g., cosecant of elevation angle); this approach does not require imposing absolute constraints. Due to the fact that the estimation of C is unstable in the absence of tropospheric asymmetry and C does not display notable seasonal variability (maximum relative power spectral density is well below 1%), we estimate them as "global parameters" for each node (cf. Figure A1). Nevertheless, in spite of the gradient components changing when alternating the mf_g, the difference between the assembled delays (employing either a site-specific or a global average C, e.g., 0.003 over land) does not exceed 1 mm even at ε as low as 5°. Hence, a constant C is retained.

Appendix B: GFZ-PT

The approach termed as GFZ-PT consists of the following ansatz:

$$y(t) = A + Bt + \sum_{i=1}^{3} C_i \sin\left(2\pi i \frac{\mathrm{DoY}}{365.25} + D_i\right) + \sum_{j=1}^{3} E_j \sin\left(2\pi j \frac{\mathrm{HoD}}{24} + F_j\right), \tag{B1}$$

where DoY denotes the day of year and HoD denotes the hour of day. All GFZ-PT parameters y (pressure, temperature, relative humidity, zenith delays, mapping function coefficients, gradient vector components of the first and second orders, and water vapor-weighted mean temperature) feature annual (S_a), semiannual (S_{sa}), terannual (S_{ta}), diurnal (S_1), semidiurnal (S_2), and terdiurnal (S_{03}) frequencies. C_i denotes the amplitude of the seasonal signals, and D_i the phase thereof. E_j denotes the amplitude of the high-frequency signals, and F_i the phase thereof. The term A denotes the time average of the signal. Harmonic terms that correspond to the 11-year sunspot cycle yield insignificant differences for the current application; hence, they were dropped. B is the rate of the series that is robustly estimated by modifying the Theil-Sen estimator (e.g., MIDAS, Blewitt et al., 2016) to better account for seasonal variations:

$$\begin{aligned} B &= \mathrm{median}\left(\frac{y_l - y_m}{t_l - t_m}\right), \forall l > m \text{ and} \\ &\left|t_l - t_m - \mathrm{argmax}\left(\mathrm{PSD}(y, \sigma_y)\right)\right| \le a, \end{aligned} \tag{B2}$$

where the signal y_l is sampled at time t_l, PSD is a power spectral density estimator operator, σ_y represents the formal errors of y, and a is the relaxation coefficient.

Acknowledgments
The ECMWF is acknowledged for making publicly available model level data from ERA-Interim (Dee et al., 2011) and ERA5 reanalysis (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim and https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5, respectively). The IVS is acknowledged (Nothnagel et al., 2015) for coordinating the VLBI experiments that were analyzed in the framework of this study. The maps were produced using the GMT (Wessel et al., 2013). We make available the Potsdam mapping functions and gradient components of first and second orders for all VLBI stations that have participated in geodetic experiments for the period 1979–2018 under ftp://ftp.gfz-potsdam.de/pub/home/kg/kyriakos/PMF/VLBI/. K. B. acknowledges financial support from the Helmholtz Association of German Research Centers under grant ZT-0007 (ADVANTAGE, Advanced Technologies for Navigation and Geodesy). The authors would like to thank two anonymous reviewers for their valuable comments on the manuscript.

References

Alshawaf, F., Balidakis, K., Dick, G., Heise, S., & Wickert, J. (2017). Estimating trends in atmospheric water vapor and temperature time series over Germany. *Atmospheric Measurement Techniques*, *10*(9), 3117–3132. https://doi.org/10.5194/amt-10-3117-2017

Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A., & Ware, R. H. (1992). GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System. *Journal of Geophysical Research*, *97*(D14), 15,787–15,801. https://doi.org/10.1029/92JD01517

Blewitt, G., Kreemer, C., Hammond, W. C., & Gazeaux, J. (2016). MIDAS robust trend estimator for accurate GPS station velocities without step detection. *Journal of Geophysical Research: Solid Earth*, *121*, 2054–2068. https://doi.org/10.1002/2015JB012552

Böhm, J. (2004). Troposphärische Laufzeitverzögerungen in der VLBI (PhD thesis), Vermessung und Geoinformation, Technische Universität Wien, Wien, Österreich, Geowissenschaftliche Mitteilungen. 68 (p. 70).

Böhm, J., Möller, G., Schindelegger, M., Pain, G., & Weber, R. (2015). Development of an improved empirical model for slant delays in the troposphere (GPT2w). *GPS Solutions*, *19*(3), 433–441. https://doi.org/10.1007/s10291-014-0403-7

Böhm, J., Urquhart, L., Steigenberger, P., Heinkelmann, R., Nafisi, V., & Schuh, H. (2013). A priori gradients in the analysis of space geodetic observations. In Z. Altamimi & X. Collilieux (Eds.), *Reference frames for applications in geosciences* (pp. 105–109). Berlin: Springer.

Böhm, 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. *Journal of Geophysical Research*, *111*, B02406. https://doi.org/10.1029/2005JB003629

Chen, G., & Herring, T. A. (1997). Effects of atmospheric azimuthal asymmetry on the analysis of space geodetic data. *Journal of Geophysical Research*, *102*(B9), 20,489–20,502. https://doi.org/10.1029/97JB01739

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Deng, Z., Gendt, G., & Schöne, T. (2016). *Status of the IGS-TIGA tide gauge data reprocessing at GFZ* (pp. 33–40). Cham: Springer International Publishing. https://doi.org/10.1007/1345_2015_156

Dill, R., & Dobslaw, H. (2013). Numerical simulations of global-scale high-resolution hydrological crustal deformations. *Journal of Geophysical Research: Solid Earth*, *118*, 5008–5017. https://doi.org/10.1002/jgrb.50353

Eriksson, D., MacMillan, D. S., & Gipson, J. M. (2014). Tropospheric delay ray tracing applied in VLBI analysis. *Journal of Geophysical Research: Solid Earth*, *119*, 9156–9170. https://doi.org/10.1002/2014JB011552

Gelaro, R., McCarty, W., Surez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Heinkelmann, R., Balidakis, K., Phogat, A., Lu, C., Mora-Diaz, J. A., Nilsson, T., & Schuh, H. (2016). *Effects of meteorological input data on the VLBI station coordinates, network scale, and EOP* (pp. 1–8). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/1345_2016_243

Heinkelmann, R., Boehm, J., Schuh, H., Bolotin, S., Engelhardt, G., MacMillan, D. S., et al. (2007). Combination of long time-series of troposphere zenith delays observed by VLBI. *Journal of Geodesy*, *81*(6), 483–501. https://doi.org/10.1007/s00190-007-0147-z

Heinkelmann, R., Willis, P., Deng, Z., Dick, G., Nilsson, T., Soja, B., et al. (2016). Multi-technique comparison of atmospheric parameters at the DORIS co-location sites during CONT14. *Advances in Space Research*, *58*(12), 2758–2773. https://doi.org/10.1016/j.asr.2016.09.023

Herring, T. A. (1992). Modeling atmospheric delays in the analysis of space geodetic data. In *Proceedings of refraction of transatmospheric signals in geodesy* (Vol. 36, pp. 157–164). Netherlands: Geodetic Commission Publications on Geodesy.

Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of change points with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590–1598. https://doi.org/10.1080/01621459.2012.737745

Kotsakis, C., Vatalis, A., & Sansò, F. (2014). On the importance of intra-frame and inter-frame covariances in frame transformation theory. *Journal of Geodesy*, *88*(12), 1187–1201. https://doi.org/10.1007/s00190-014-0753-5

Landskron, D., Hofmeister, A., & Böhm, J. (2017). *Refined tropospheric delay models for CONT11* (pp. 65–69). Cham: Springer International Publishing. https://doi.org/10.1007/1345_2015_56

MacMillan, D. S. (1995). Atmospheric gradients from very long baseline interferometry observations. *Geophysical Research Letters*, *22*(9), 1041–1044. https://doi.org/10.1029/95GL00887

MacMillan, D. S., & Ma, C. (1994). Evaluation of very long baseline interferometry atmospheric modeling improvements. *Journal of Geophysical Research*, *99*(B1), 637–651. https://doi.org/10.1029/93JB02162

MacMillan, D. S., & Ma, C. (1997). Atmospheric gradients and the VLBI terrestrial and celestial reference frames. *Geophysical Research Letters*, *24*(4), 453–456. https://doi.org/10.1029/97GL00143

Marini, J. W. (1972). Correction of satellite tracking data for an arbitrary tropospheric profile. *Radio Science*, *7*(2), 223–231. https://doi.org/10.1029/RS007i002p00223

Nilsson, T., Böhm, J., Wijaya, D. D., Tresch, A., Nafisi, V., & Schuh, H. (2013). *Path delays in the neutral atmosphere* (pp. 73–136). Berlin: Springer. https://doi.org/10.1007/978-3-642-36932-2_3

Nilsson, T., & Elgered, G. (2008). Long-term trends in the atmospheric water vapor content estimated from ground-based GPS data. *Journal of Geophysical Research*, *113*, D19101. https://doi.org/10.1029/2008JD010110

Nilsson, T., Soja, B., Balidakis, K., Karbon, M., Heinkelmann, R., Deng, Z., & Schuh, H. (2017). Improving the modeling of the atmospheric delay in the data analysis of the intensive VLBI sessions and the impact on the UT1 estimates. *Journal of Geodesy*, *91*(7), 857–866. https://doi.org/10.1007/s00190-016-0985-7

Nilsson, T., Soja, B., Karbon, M., Heinkelmann, R., & Schuh, H. (2015). Application of Kalman filtering in VLBI data analysis. *Earth, Planets and Space*, *67*(1), 136. https://doi.org/10.1186/s40623-015-0307-y

Nothnagel, A., Alef, W., Amagai, J., Andersen, P. H., Andreeva, T., Artz, T., et al. (2015). *The IVS data input to ITRF2014*. Potsdam, Germany: International VLBI Service for Geodesy and Astrometry, GFZ Data Services. https://doi.org/10.5880/GFZ.1.1.2015.002

Petit, G., & Luzum, B. (Eds.) (2010). IERS conventions (2010) (*IERS Technical Note 36*). Frankfurt am Main.

Rousseeuw, P. J., & Leroy, A. M. (2005). *Multiple regression* (pp. 75–157). John Wiley and Sons Inc. https://doi.org/10.1002/0471725382.ch3

Schuh, H., Panafidina, N., Boehm, J., & Heinkelmann, R. (2006). Climatic signals observed by VLBI. *Acta Geodaetica et Geophysica Hungarica*, *41*(2), 159–170. https://doi.org/10.1556/AGeod.41.2006.2.2

Soja, B., Nilsson, T., Balidakis, K., Glaser, S., Heinkelmann, R., & Schuh, H. (2016). Determination of a terrestrial reference frame via Kalman filtering of very long baseline interferometry data. *Journal of Geodesy*, *90*(12), 1311–1327. https://doi.org/10.1007/s00190-016-0924-7

Wang, X. L., Wen, Q. H., & Wu, Y. (2007). Penalized maximal t test for detecting undocumented mean change in climate data series. *Journal of Applied Meteorology and Climatology*, *46*(6), 916–931. https://doi.org/10.1175/JAM2504.1

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., & Wobbe, F. (2013). Generic Mapping Tools: Improved version released. *Eos Transactions, American Geophysical Union*, *94*(45), 409–410. https://doi.org/10.1002/2013EO450001

Wresnik, J., Haas, R., Boehm, J., & Schuh, H. (2007). Modeling thermal deformation of VLBI antennas with a new temperature model. *Journal of Geodesy*, *81*(6), 423–431. https://doi.org/10.1007/s00190-006-0120-2

Zus, F., Bender, M., Deng, Z., Dick, G., Heise, S., Shang-Guan, M., & Wickert, J. (2012). A methodology to compute GPS slant total delays in a numerical weather model. *Radio Science*, *47*, RS2018. https://doi.org/10.1029/2011RS004853

Zus, F., Dick, G., Douša, J., & Wickert, J. (2015). Systematic errors of mapping functions which are based on the VMF1 concept. *GPS Solutions*, *19*(2), 277–286. https://doi.org/10.1007/s10291-014-0386-4

Zus, F., Dick, G., Heise, S., & Wickert, J. (2015). A forward operator and its adjoint for GPS slant total delays. *Radio Science*, *50*, 393–405. https://doi.org/10.1002/2014RS005584