Chameleonic noise in GPS position time series

Alvaro Santamaría-Gómez[1,1] and Jim R Ray[2,2]

[1]GET, Université de Toulouse, CNES, CNRS, IRD, UPS, Toulouse, France
[2]National Geodetic Survey (retired), Silver Spring, MD 20910, USA

November 29, 2022

Abstract

It has been standard practice for about two decades to compute GPS-based station velocity uncertainties using the apparent noise statistics of the non-linear position residuals rather than assume white noise (WN) behavior. The latter choice would yield unrealistic velocity uncertainties. The most common noise types used are power-law, usually close to flicker noise (FN), over most frequencies mixed with WN at the shortest periods. The complicating impact of offsets in the position time series, mostly caused by equipment changes or tectonic events, has not been fully appreciated. These are far less benign than recently suggested. In addition to contributing a pseudo-random walk noise (RW) component to the velocity errors, estimating offset parameters changes the apparent noise color towards whiter. Spectral power is effectively drained by offsets at periods longer than roughly the mean span between them. This consequently promotes a Gauss-Markov process as the apparently preferred noise model and, importantly, obscures the presence of RW and long-period Earth deformation in the series. Both effects can lead to potentially under-estimated velocity uncertainties. The full value of decadal-long GPS time series for geodynamical applications is thereby greatly eroded by recurring offsets, especially when they occur quasi-regularly. In addition, contrary to common assumption, the noise color is generally not fixed with time, but clearly becomes whiter in more recent data. The origin of the colored noise and its whitening over time remain elusive.

Posted on 29 Nov 2022 — CC-BY 4.0 — https://doi.org/10.1002/essoar.10503150/v2 — This a preprint and has not been peer reviewed. Data may be preliminary.

Chameleonic noise in GPS position time series

Alvaro Santamaría-Gómez[1] and Jim R. Ray[2]

[1] GET, Université de Toulouse, CNES, CNRS, IRD, UPS, Toulouse, France.

[2] National Geodetic Survey (retired), Silver Spring, MD 20910, USA.

Corresponding author: Alvaro Santamaría-Gómez (alvaro.santamaria@get.omp.eu)

Key points:

- Position offsets attenuate most of the power spectrum at long periods and severely impact velocity estimates and their uncertainties.
- If the impact of offsets is not considered, the best-fitting noise color changes from flicker to Gauss-Markov.
- Position offsets reduce the chances of detecting random walk noise and long-period Earth deformation signals in long GPS series.

Abstract

It has been standard practice for about two decades to compute GPS-based station velocity uncertainties using the apparent noise statistics of the non-linear position residuals rather than assume white noise (WN) behavior. The latter choice would yield unrealistic velocity uncertainties. The most common noise types used are power-law, usually close to flicker noise (FN), over most frequencies mixed with WN at the shortest periods. The complicating impact of offsets in the position time series, mostly caused by equipment changes or tectonic events, has not been fully appreciated. These are far less benign than recently suggested. In addition to contributing a pseudo-random walk noise (RW) component to the velocity errors, estimating offset parameters changes the apparent noise color towards whiter. Spectral power is effectively drained by offsets at periods longer than roughly the mean span between them. This consequently promotes a Gauss-Markov process as the apparently preferred noise model and, importantly, obscures the presence of RW and long-period Earth deformation in the series. Both effects can lead to potentially under-estimated velocity uncertainties. The full value of decadal-long GPS time series for geodynamical applications is thereby greatly eroded by recurring offsets, especially when they occur quasi-regularly. In addition, contrary to common assumption, the noise color is generally not fixed with time, but clearly becomes whiter in more recent data. The origin of the colored noise and its whitening over time remain elusive.

1. Introduction and motivation

This paper addresses the central problem of GPS time series analysis: the separation of what is considered signal, i.e., the systematic variations, from what is considered to be noise, i.e., the random variations of usually unknown origin. A functional model is commonly fit to a GPS position time series to represent the signal, consisting of a linear trend, periodic variations, and irregularly timed position offsets as needed. It was realized in the 1990s that weighted least-squares fitting for the functional model with formal error propagation yields highly optimistic uncertainties for geodetic parameters. Then the discovery that the GPS-based errors are correlated both spatially (Wdowinski et al., 1997) and temporally (Zhang et al., 1997) led to more robust methods to assess and quantify the levels and types of noise in time series of GPS station positions (Mao et al., 1999). The main goal was to obtain more reliable uncertainties for GPS-based velocities, which are needed for many geodynamical studies such as measuring intra-plate tectonic stability and correcting vertical land motion from tide gauge records. A general consensus quickly emerged that GPS time series errors combine mostly white noise (WN) at shorter periods, typically less than a month, with flicker noise (FN) or similar power-law noise (PL) over longer spans, from monthly up to decadal periods (Mao et al., 1999; Zhang et al., 1997), although some level of background random walk (RW) error cannot be excluded. Consequently, velocity uncertainties can be under-estimated by as much as an order of magnitude if WN alone is assumed (Mao et al., 1999).

Development quickly followed of several mathematical tools to enable users to evaluate models of alternative noise types and amplitudes in observed GPS time series (Amiri-Simkooei et al., 2007; Bos et al., 2013; Langbein, 2004; Williams, 2008), which have been widely adopted for published noise analyses. One of the most common mathematical approaches found in the literature is the maximum likelihood estimator (MLE) which is well adapted to fit any noise model or combination of noise models to univariate time series. In order to discriminate the best-fitting noise model, the estimated likelihood differences (dML) are often used after accounting for the different degrees of freedom of each noise model. Very recently, this approach has been extended to analyze spatial common-mode noise models from regional GPS networks (Dmitrieva et al., 2015).

In this study, we analyze the impact the functional model, particularly position offsets, has on the apparent color of the noise model fitted to the position time series. Two very recent papers have raised questions about the orthodox interpretation of the long time series that are increasingly available nowadays. He et al. (2019) conclude that GPS noise does not flatten at low frequencies while assuming that the noise properties are stable over time, which we will reiterate below is not justified, in part due to the effect of occasional position offsets. Meanwhile, Wang and Herring (2019) attempt to show that the impact of position offsets is less serious than previously thought. These authors base their conclusions on serious misrepresentations. It is the objective of this contribution to clarify these points and to offer new perspectives on the information obtained from the colored noise in GPS position time series.

2. Background

2.1 Basics of time series computation

Most studies of GPS noise start with some particular set of time series with little or no regard for the process that was used to generate them or how alternative results might compare. It is, however, instructive to first consider the processing steps that produce observed GPS time series as this affects

their noise content. “GPS” in the following can increasingly be regarded generically as “GNSS”, but in actual fact, for the long time series considered here, GPS observations have been used exclusively.

Three types of GPS data processing can be distinguished that produce time series of station positions: 1) full global network solutions free of any a priori over-constraints; 2) precise point positioning (PPP) of several sorts; and 3) regional network solutions. The International GNSS Service (IGS) and its Analysis Centers (ACs) are the prime examples of global network providers. Their occasional IGS reprocessing campaigns generate the GNSS inputs for the International Terrestrial Reference Frame (ITRF; Altamimi et al. 2016) and for many PPP solutions. PPP allows a single GNSS station to be positioned by assuming a priori satellite orbits and clocks from some given global solution provider (Zumberge et al., 1997). This is an efficient and autonomous method to densify a global solution that is usually limited in size by CPU constraints. The quality of PPP solutions is poorer when integer phase ambiguities are not fixed, but improves notably, especially in the east component, when double-differenced or one-way ambiguities are resolved, which is the usual case today. Regional networks are processed for many local or national applications requiring dense coverage, and can yield highly precise differential positions due to the natural rejection of common-mode errors.

The assignment of absolute geocentric coordinates to estimated station positions depends on the type of processing used. The IGS global network positions are realized, after combining independent AC solutions, by aligning each unconstrained daily frame via no-net-rotation (NNR) to the current linear ITRF coordinates using a "core" subset of well-distributed and reliable reference stations, typically numbering between about 40 and 80 (Rebischung et al., 2016b). This alignment process unavoidably redistributes non-linear displacements, both genuinely physical as well as technique artifacts, among the reference stations to all other network members (Collilieux et al., 2012). By including as many reference stations as possible, as uniformly spaced as possible, the net alignment-induced scatter should be much less than if only a few fiducials were used. Ray et al. (2017) showed that the long-term weighted root-mean-square (WRMS) of this rotational scatter is about 25 to 30 μas, equivalent to about 0.8 to 0.9 mm of equatorial motion, which corresponds to the estimated accuracy of daily IGS polar motion measurements. This same level of scatter also matches the observed WRMS error floor for daily IGS horizontal GPS positions after removing periodic variations, actually, about 0.9 and 0.8 mm for North and East components, respectively. Individual stations often display scatter that is considerably larger than these floors due to a variety of local effects. Ray et al. (2017) also found a WRMS error floor for non-periodic GPS station heights of 3.0 mm, which is commensurate with the usual ratio of GPS vertical to horizontal errors. The spectral qualities of the net frame alignment scatter were not studied by Ray et al. (2017) except to describe them as broadband and with more power at longer periods, i.e., reddish noise.

At a minimum, time series of PPP estimates inherit the net noise properties of their reference global network solution, and its ITRF alignment, that have been fixed in the processing. In fact, contrary to assertions by Zumberge et al. (1997), the actual performance of a PPP solution must always be degraded somewhat by loss of the full solution covariance information and by any potential deviation from identical modeling applied in the global network versus the PPP data reductions. (See further points in section 4.1).

Regional network solutions are particularly effective in estimating very precise horizontal positions and velocities by innately suppressing spatially coherent errors and by being generally less sensitive to large-scale effects, e.g., orbit errors. But they do not provide reliable absolute geocentric heights for the same reason and long-term alignments of regional frames to ITRF are subject to significant biases (Legrand et al., 2010).

2.2 Analysis effects in GPS noise

There have been few quantitative comparisons between the GPS noise characteristics from different IGS ACs or from different types of computation strategy (global, PPP, regional). Using the AC-specific station residuals from continuously updated operational IGS combined products might be instructive to some extent, but the interpretation is confused by comparison to the weighted AC mean. To address this concern, Rebischung et al. (2016a) formed individual long-term solutions for each AC in the second IGS reprocessing campaign (i.e., Repro2) and examined the station residuals. There was a large variation among ACs in the number of stations processed, but the quality of solutions was generally rather similar. The background non-periodic power is well described, except for JPL, as high-frequency WN plus a PL with spectral index near -1 for North and East and near -0.8 for up. The JPL behavior differs in having less WN, which is undetectable for heights, and having smaller PL spectral indices of -0.72 to -0.74 for North and East and -0.52 for up.

On the surface this result seems to imply that WN is largely a product of data analysis and not derived from GPS observational noise. But that conclusion could be undermined by the facts that the JPL Kalman filter processor might simply absorb WN into its parameter estimates or that the JPL use of data arcs that overlap by 3 hr with adjoining days before and after could exert a strong smoothing effect. EMR/NRCan is the only other IGS AC using the JPL Kalman filter but they fit strict 24-hr arcs and have noise properties similar to the other non-JPL ACs, which supports the second possibility of implicit smoothing.

The attenuation of WN is at least partly produced by the implicit smoothing effect of including 3 hr overlaps from the previous and the following daily observations. This is the case in the JPL processing, but it is less clear how any difference of the JPL analysis scheme can reduce their spectral indices by ~0.3. After all, the geophysical processes that affect GPS data and positions, such as surface pressure loading, contribute effects that have distinctly red spectral variations, at least up to periods of a year or longer. One guess is that the critical element might involve JPL's estimation of satellite solar radiation pressure (SRP) variations as 3D stochastic offsets rather than as harmonic parameters, as most ACs do, but EMR also uses 3D stochastic SRP estimation.

Concerning position offsets, Williams (2003) showed that the presence of undetected and unmodeled position offsets in GPS time series contributes a RW noise component. He also evaluated the impact of adding offset parameters to the functional model, for various background noise types. Griffiths and Ray (2016) quantified the magnitude of the RW-like noise due to adding position offsets in an ITRF-type global network solution. Under these circumstances and because some station hardware changes cause discontinuities while others do not, different ACs are expected to introduce varying amounts of RW noise when they use incomplete metadata for the station equipment changes. This effect is probably not dominant among AC noise differences, however.

2.3 Assumptions of noise analysis and the evolution of colored noise

In constructing analytical methods to evaluate GPS positional noise, a number of assumptions must be made, some of them by implication only. Basic is the natural notion that the total observed GPS noise is a superposition of different physical processes each with its own spectral characteristics. Some processes might be so slight, compared to others, that their noise effects cannot be detected directly or the data span that would be needed is very long, e.g., RW. Under these circumstances, it is not generally possible to invert a GPS spectrum to infer its constituent noise components. Instead, forward modeling of a range of possible noise types and amplitudes is usually performed and tested by best-fitting against the observations, which is of course non-unique and prone to subjectivity.

Another key assumption almost always made is that the mix of noise characteristics for any given time series is constant over time, despite observations of Williams et al. (2004) and Santamaría-Gómez et al. (2011) to the contrary. Time series, especially long ones, often display non-stationary noise properties that are evident to the eye. For some examples, please consider the residual JPL PPP series for these long-running stations: ALIC, CAS1, CHUM, COCO, DAV1, GENO, KIR0, KIRU, MAS1, MATE, MEDI, NRC1, ONSA, PIN1, RAMO, SFER, TOW2, UCLU, ZIMM. There has been an overall trend towards improved GPS tracking hardware and more robust ground networks on the one hand, but degradations and gradual failures at individual stations can happen at any time. This is why visual inspection of the series, as done in this study, is of uttermost importance to support the conclusions of a given analysis. The JPL plots also illustrate the pervasiveness of position offsets in any GPS solution, which average about 0.9 per station per decade in the IGS Repro2 global analysis (Rebischung et al., 2016b), the dataset used by Griffiths and Ray (2016).

2.4 Comments on Wang and Herring (2019)

Partly to address objections raised by an anonymous reviewer #1 (AnR1), we review here results from Wang and Herring (2019) [WH19]. This is done reluctantly for fear of appearing to limit the scope of this contribution to a mere rebuttal, which is not our intent. This does give us a chance, however, to introduce several important themes that recur later in our own results.

Amidst the ponderous formalism and numerous noise models considered by WH19, one could easily overlook the simple fact that any connection between their simulations and real-world GNSS analysis is not obvious. Each of their 22 different noise models, mostly quite unrealistic, is simulated with known covariances and then the impact of position offset parameters on formal velocity uncertainties is assessed by least-squares solutions using the same known covariances. Such a circumstance never arises with real data where the contributing errors are unknown in type and magnitude. As mentioned in the previous section, a range of error models is normally tested while estimating the deterministic parameters and their uncertainties, in search of a maximally likely outcome.

Meanwhile, adding offset parameters to the functional model itself, while commonly assuming WN series, changes the noise color of the residual series by removing progressively more low-frequency power, as we show in section 4.2. So the true power spectrum becomes more obscured and distorted as the number of offsets per unit time increases, putting the ideal world simulated by WH19 ever farther out of reach as offsets accumulate. Nonetheless, AnR1 asserts that this, our core finding, "is not anything new" even though this phenomenon is never mentioned by WH19 and it rules out any practical utilization of the WH19 approach using real data with offsets. We found only one prior study of the noise "whitening" effect of adding position offset parameters, the simulation by Chen et al. (2018). Position offsets were one of several analysis options they examined, obtaining median impacts consistent with ours but with more limited insight into the underlying dynamics of the problem, which is our prime interest here.

WH19 take pains to aggressively dismiss the earlier study by Griffiths and Ray (2016) [GR16] of the effects of position offset parameters, while repeatedly misstating their analysis, such as by equating "adding one additional offset parameter" to the doubling of offsets per unit time discussed by the latter authors. GR16 never considered the estimated velocity uncertainties whatsoever, contrary to multiple WH19 assertions, but instead they examined the increase in dispersion of actual rate estimates, not simulated data series, as offset parameters were added artificially to a long-term global reference frame solution. WH19 and AnR1 both claim that the rate error scatters obtained by GR16 were biased by not having included full temporal noise correlations. The impact of a given

process noise on velocity estimates has been simulated by WH19 (their Fig. 7). However, they assumed an unrealistic RW process noise for the actual GPS data. If one instead considers realistic PL noises with spectral indices close to FN, then the scatter of velocity estimates is negligibly impacted, consistent with GR16. These points are reiterated in more detail in section 5.2.

Most offensive of all is the cavalier repudiation by WH19 of pleas by Williams (2003a) and GR16, which we also support in this contribution, for station operators to maintain the strictest possible configuration control of GNSS reference stations to minimize the occurrence of position offsets and to maximize the geodynamical value of long GNSS time series.

3. Data and methods

In this study we used residual daily GPS position time series from the JPL online service (https://sideshow.jpl.nasa.gov/post/series.html; accessed in May 2019). These series are estimated using the PPP technique with phase ambiguity resolution and JPL final orbits and clocks products. The daily PPP frames were NNR-aligned to the IGS14 realization of ITRF2014 (Rebischung et al., 2016b) using transformation parameters estimated at the same time as the orbit/clock products in a full global network solution. JPL data reduction models and procedures were updated following Repro2 (IGS Mail 7637, 5 June 2018) and applied in the network solution and in the PPP processing, including antenna calibrations consistent with the IGS14 frame. This latest JPL PPP reprocessing ensures the temporal coherence of the historic series with their current operational products.

The residual series were obtained by JPL after removing linear trends, seasonal variations, position offsets, and outliers. JPL uses an automatic offset detector that Gazeaux et al. (2013) ranked as the second-best automatic solution in their study.

From the initial set of 2567 stations available, we retained 137 stations having at least 20 years of data since 1999, less than 10% of missing points, and gaps not longer than 5 months. These 137 stations were screened for quality using SARI (Santamaría-Gómez, 2019) in which we removed several outliers most probably caused by snow/ice on the antennas. In addition, we added position offsets for 8 stations that were visually identified and judged to be significant against the colored noise content in the series. After this check, we rejected 10 more stations having unusual residual distributions or strong changes in their variability, leaving 127 long, high-quality stations for this study (see all the information in the supplemental material). We observed a very irregular temporal sampling for some of these stations with, occasionally, multiple points estimated on the same day separated by only a few hours. This might be caused by JPL merging their global network and PPP estimates for the same station. To avoid this, we removed the daily estimates with epochs before 6 h UTC or after 18 h UTC. We also note there are more than 3 offsets per station per decade on average among the selected 127 stations, which is particularly large compared to the value for the IGS Repro2 series.

We acknowledge the JPL PPP series have distinct noise characteristics compared to other GPS position series, especially if they are derived from a filtered regional solution. However, the JPL series provide the longest homogeneously reprocessed residual series to date, before the next IGS reprocessing campaign, and for an extensive number of global stations. Judging from IGS combination reports (see weekly orbit and frame combination summaries at https://lists.igs.org/pipermail/igsreport/), the JPL analysis quality is outstanding also. This dataset is therefore the best available for addressing several points concerning the noise content and the impact of position offsets.

In the following, we used the Lomb-Scargle periodogram (Scargle, 1982) to represent the error power spectra and the Create and Analyze Time Series (CATS) software (Williams, 2008), which implements MLE, for the estimation of the power spectra of a noise model. Here, we define, on one hand, the noise power spectrum as the power spectrum that is obtained by fitting a pre-defined noise model that is conveniently used to assess the formal errors of a functional model, especially linear trends. On the other hand, we define the error power spectrum as the power spectrum that is actually observed in residual GPS series. In other words, the error spectrum represents the spectral distribution of the variance that is not explained by the reduction of the raw GPS observations nor by the fitted model to the series, and it includes random and systematic variations, both physical and spurious. Generally speaking, the noise spectrum is obtained by fitting the error spectrum with a particular stochastic process or mix of processes. Ideally, the noise spectrum and the error spectrum should be very similar, but we show in the next section that this is not always the case and that a distinction between them needs to be recognized.

JPL obtained their position residual series by assuming WN series, as is common practice. Therefore, in order to avoid biasing the estimates of the noise spectrum, we fitted the residuals with their corresponding functional model in CATS. This provides the same noise parameters as if the original series were used, i.e. before removing the functional model, while allowing us to use exactly the same input residual series to compute a consistent error power spectrum with the periodogram. In addition, we also included up to the 10^{th} harmonic of the 1.04 cpy draconitic oscillation that was not removed by JPL.

4. Results

4.1 Noise types in the JPL PPP series

Before testing the validity of candidate noise models for the 127 residual JPL series, we shall evaluate first the GPS error power spectra, which are shown in Fig. 1. Overall, the error spectra have a

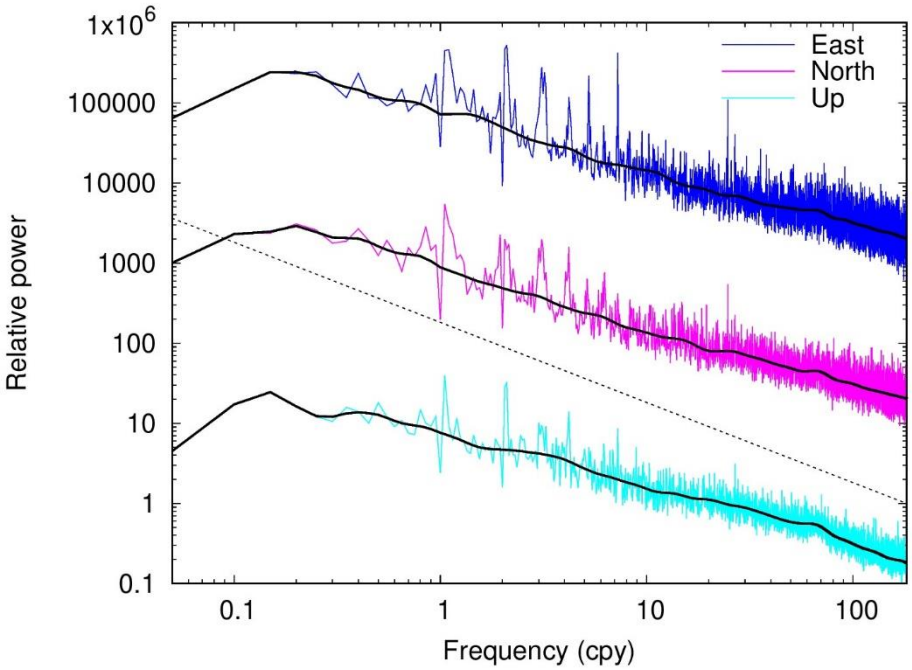


Figure 1. Stacked error spectra of the 127 selected JPL series for the three coordinate components. The spectra are shifted for the sake of visibility. The thick solid black lines represent the smoothed spectra. The thin dotted black line represents a pure flicker noise process.

crowbar shape that could be described by a power-law process up to periods of ~5 years. The slope of the power-law is clearly lower than FN, especially in the vertical component. Beyond 5 years, the power spectra decay significantly in all the components. The loss of power at long periods is further examined in Section 4.2. At the shortest periods, we observe that WN in not visible in any coordinate component. The PL clearly dominates the stacked spectra even at high frequencies of at least 100 cycles-per-year (cpy) or more. To quantify an upper bound of the WN amplitude, we assume the PL is actually dominated by FN, which is likely for the horizontal components, but not so much for the vertical component as we discuss later on. Considering a typical FN amplitude for the daily JPL series of 4 mm $yr^{-0.25}$ for the horizontal and 11 mm $yr^{-0.25}$ for the vertical component (values estimated with CATS), and a cross-over period from FN to WN of less than 4 days, we bound the maximum white noise level to be 0.4 and 1.1 mm for the horizontal and vertical components, respectively. This is much lower than the WN reported for the other AC solutions of the last IGS Repro2 (Rebischung et al., 2016a) and reflects the smoothing inherent in the JPL global processing used for their PPP series. In addition, Ray et al. (2013) and Amiri-Simkooei et al. (2017) have found spurious ~5 d signals in the preceding JPL PPP time series that are not present in their corresponding global solutions, the source of which was unknown. The ~5-day signals reported in those past studies are not visible in this latest JPL solution, but nevertheless we still note a bump around 5.5 days (~65 cpy) common to all the components and of unknown origin, but possibly related to the overlapped 3 + 24 +3 hr data arcs used by JPL. The power spectrum decays faster for periods shorter than 5.5 days in the vertical component only.

Figure S1 in the Supporting Information shows the distribution of the 127 stations used in this study. Most of the stations are concentrated in Europe, Australasia and, especially, Western USA. To explore to what extent the stacked error spectra of Fig. 1 could be biased by the geographic distribution of the 127 sites, we computed two additional smoothed stacked spectra corresponding to the stations located in Western USA (62) and Europe (36). Figure S2 in the Supporting Information shows that, while the spectra from the regional subsets are more variable due to the smaller number of sites, their general shapes are consistent.

Because white noise obviously does not play a part in the background error spectrum, in the following noise analysis we average the daily series into weekly series in order to speed up the processing without losing information on the colored noise content.

Using CATS, we estimated the type and amplitude of the colored noise that best describes the error spectra of the individual series. To assess the impact of offsets in the noise analysis, we assume the background noise content of the series is constant over time and can be described by a unique noise model. We will demonstrate later that this hypothesis, commonly applied in most noise analyses, is not valid. At least we can assess how well a single model describes the background noise because all our series are available for the same period of time. Four noise models were tested: FN, PL, flicker plus random walk noise (FNRW) and a generalized Gauss-Markov noise model (GM), which includes a PL process at high frequency and a WN process at low frequency (Langbein, 2004).

In order to discriminate the goodness of fit for each model, we used the differences of the estimated maximum likelihood (dML) plus an empirical threshold that accounts for the different degrees of freedom of each noise model with respect to FN, which is selected as the null hypothesis for the noise model. Using 1000 20-year-long synthetic weekly FN series, the empirical dML thresholds were set as the 95 % percentile of the dML values between FN and the other three model candidates. Together with the noise model, the synthetic series were fitted by a model including a linear trend, annual and semi-annual variations plus draconitic harmonics up to the tenth. In addition, several dML thresholds were estimated for each candidate noise model corresponding to 6 different

numbers of estimated offsets, from 0 up to 9 equally-spaced offsets, as suggested by John Langbein (personal communication). Based on the dML values obtained for different numbers of offsets, an empirical model was fitted and presented in Table 1 as a function of the offset rate (*r*), i.e., the number of offsets per year. For the PL and the GM models, the dML values increase following an exponential function with the addition of position offsets, with a larger increase for GM than PL. Conversely, the dML values for the FNRW model are significantly smaller than that of PL and GM and, importantly, they are insensitive to the rate of offsets in the series. We also note that, as the number of offsets increases in the series, the number of synthetic series for which a false positive RW was found by chance decreases significantly; from around 10 % in case of no offsets down to 0.5 % in case of one offset every 2 years. The reduction in the number of RW false positives with the number of offsets indicates that offsets have a negative impact on our ability to detect RW in the series. We further develop this issue in Section 4.3. Since the dML values have a long-tail distribution, in order to have enough dML values to compute a reliable 95 % threshold, we increased the number of synthetic series to 10,000 for the FNRW model only. The dML thresholds at the 99 % percentile are also provided in Table 1. The extreme 99 % dML value could be used for instance to account for an extremely irregular distribution of offsets in the series.

Table 1. Empirical maximum likelihood differences (dML) of the power-law (PL), general Gauss-Markov (GM) and flicker plus random walk (FNRW) noise models with respect to a flicker noise null hypothesis. The expressions correspond to the model fitted to different dML values, at the 95 % or 99 % percentile, as a function of the offset rate (r) in the series.

	95 %	99 %
PL	$4.0\ e^{0.9r}$	$5.8\ e^{1.0r}$
GM	$5.8\ e^{1.5r}$	$8.2\ e^{1.1r}$
FNRW	2	3

Taking into account the 95 % dML thresholds of Table 1, the distribution of the best fitted noise model per component (labelled “Full series”), the noise model parameters and its degrees of freedom are given in Table 2. The FN model is selected for most of the series in both horizontal components with amplitudes of around 2 mm yr $^{-0.25}$, followed by the GM model. In the vertical component there is no single noise model clearly dominating among the series, though GM is more common than PL or FN. The PL model, being selected two times more often in the vertical than in the horizontal, exhibits values of the spectral index significantly whiter than FN, consistent with the qualitative assessment from the stacked periodogram (Fig. 1). The FNRW noise model (2 degrees of freedom) is not included in Table 2 because it was not selected as the preferred noise model in any of the series. We observe that the spectral index of the GM process is much closer to FN than RW, which is also consistent with the error spectra shown in Fig. 1, and indicates that the first-order GM (FOGM) noise process used by WH19, which combines WN at long periods with RW at shorter periods, may not be the best model to describe the noise in most of the series. We note however that WH19 used the series issued from a regional solution (Herring et al., 2016), which are certainly different from the JPL series and probably even less consistent with a FOGM model. A similar observation was recently made by Langbein (2020) .The selected noise model for each series and coordinate component are included in Table S1 of the Supporting Information. Together with the selected noise model, we provide the dML value obtained with respect to the FN null hypothesis. Generally, the GM has a typical dML value clustered around 3 and 5 for the horizontal and vertical components, respectively. The PL model has median dML values around 4 and 10 for both components, respectively. For completeness, many of the obtained dML values were close or less than 1. If we use the dML thresholds at 99 % from Table 1, the percentage of selected FN series in Table 2 increases by an average of 10 % for the three components.

Table 2. Percentage of the best fitted noise model per component. The percentage in parenthesis corresponds to the use of dML thresholds without considering the impact of offsets. The numbers in parenthesis after the noise model name in the first column provide the degrees of freedom. The values for relevant noise model parameters for the selected fits (σ for variance in mm, k for spectral index and β for cross-over period in years) are provided at the 5%/50%/95% percentiles. Note that the variance corresponds to weekly sampled series.

		East	North	Up
FN (1)	Full series	60% (39%) σ=1.9/2.2/3.0	58% (23%) σ=1.9/2.5/3.4	20% (4%) σ=6.3/7.2/9.4
	Segments	41% σ=1.7/2.0/2.8	43% σ=1.7/2.1/2.8	11% σ=5.8/7.1/13.2
PL (2)	Full series	16% (13%) σ=1.9/2.0/4.0 k=-0.56/-0.71/-1.37	14% (8%) σ=1.8/2.4/4.6 k=-0.54/-0.67/-0.77	37% (20%) σ=4.7/6.0/7.7 k=-0.56/-0.66/-0.78
	Segments	42% σ=1.2/1.5/2.1 k=-0.43/-0.66/-0.77	39% σ=1.2/1.7/2.5 k=-0.42/-0.66/-0.74	68% σ=3.8/4.9/6.8 k=-0.44/-0.63/-0.75
GM (3)	Full series	24% (48%) k=-0.95/-1.25/-2.57 β=0.2/0.5/1.6	28% (69%) k=-1.12/-1.17/-2.60 β=0.2/0.6/1.2	43% (76%) k=-0.81/-1.01/-1.96 β=0.3/0.6/1.4
	Segments	17% k=-0.90/-1.58/-3.63 β=0.1/0.3/0.8	18% k=-0.86/-1.25/-3.64 β=0.1/0.7/1.5	21% k=-0.70/-1.03/-3.59 β=0.1/0.3/2.3

4.2 Loss of low-frequency power

Table 2 also shows in parenthesis the distribution of the selected noise models that would result if the impact of offsets were not taken into account in the estimation of the dML thresholds, i.e. *r = 0* in Table 1. In that case, the apparent noise type that clearly dominates is a GM, especially in the North and Up components. The effect of including the offsets in the dML thresholds effectively transfers the selected noise model from GM to FN, with a reduction/increase of the GM/FN ratios by a factor ~2, respectively. This would be consistent with the estimated error spectra in Fig. 1, i.e., among the tested noise models and in the absence of additional constraints, GM would be the one that better approximates the observed crowbar shape of the power spectrum. However, the GM noise power spectrum at long periods is given by a frequency-independent WN process, which does not explain the sharp loss of low-frequency power observed in the error spectra of Figure 1. It is known that the trend estimation absorbs part of the colored noise at the longest period and that the estimated colored noise from the residual series is biased low for short series (Langbein & Johnson, 1997). Here we argue that when the functional model is fitted assuming WN series, which is standard practice, a major amount of the low-frequency power spectrum is also absorbed by the estimated position offsets.

To demonstrate this, we first reran the noise analysis, but using this time the continuous segments of the series between each pair of estimated offsets. From our initial 127 series of 20 years, we extracted 162 segments longer than 5 years with a median length of almost 9 years. We used the dML thresholds in Table 1 with *r = 0* and the results are shown in Table 2 (labelled "Segments"). Compared to the results from the full series with *r = 0*, the GM is not the dominant selected noise model anymore; in the vertical component it is now PL, and in both horizontal components GM has been replaced by FN and PL with similar scores. We note however that it is difficult to reach firm conclusions about the general noise model that best describes the error spectrum when the series do not have the same length nor cover the same period of time. For instance, by comparing the ratio of the selected GM models between the “Full series”(with *r = 0*) and the “Segments” runs in Table 2, one would be tempted to state that the effect of offsets is to triple the chances to select a GM

model, while we already show this factor is actually closer to 2. This difference could be attributed to the popular approach of comparing the noise content between series having different lengths, or with similar lengths but not covering the same period. This exercise at least confirms that the GM model is apparently promoted by the effect of estimated position offsets in long discontinuous series.

A more rigorous approach that could be applied, especially when comparing noise levels from different sets of series, is to use continuous segments of the same length and period, at the expense of reducing the number of exploitable series. Since that was not possible in this case, we carried out another test to demonstrate the impact of the offsets in the error spectrum by creating synthetic series with the same sampling as our 127 original series and different amplitudes of FN from the ranges in Table 2. These series were then fitted with a linear trend and the same offsets as in the original JPL series. Figure 2 (black curves) shows how the FN power spectrum is absorbed at long periods slightly by the trend and mostly by the offsets. The simulated FN power spectrum starts flattening around a period of 3 years and the power starts dropping around a period of 7 years, which lies close to the median offset separation of the dataset of 8.2 years. The decay follows the crowbar shape of the real error spectra reasonably well. This indicates that, very likely, the error spectra of the GPS series actually would not decay or flatten if the series were continuous, which partly supports the use of the FN as the null hypothesis for testing alternative candidate noise models. In practice however, the lack of observations in that part of the spectrum restrains us to better constraint the null hypothesis to be used.

Using these synthetic FN series, we estimated the sensitivity of estimated noise parameters of FN and PL noise models to the addition of the JPL offsets. We observe a small whitening of the estimated PL spectral index of about 5% in average. The estimated FN amplitude also reduces marginally with the offsets by around 0.1 mm $yr^{-0.25}$.

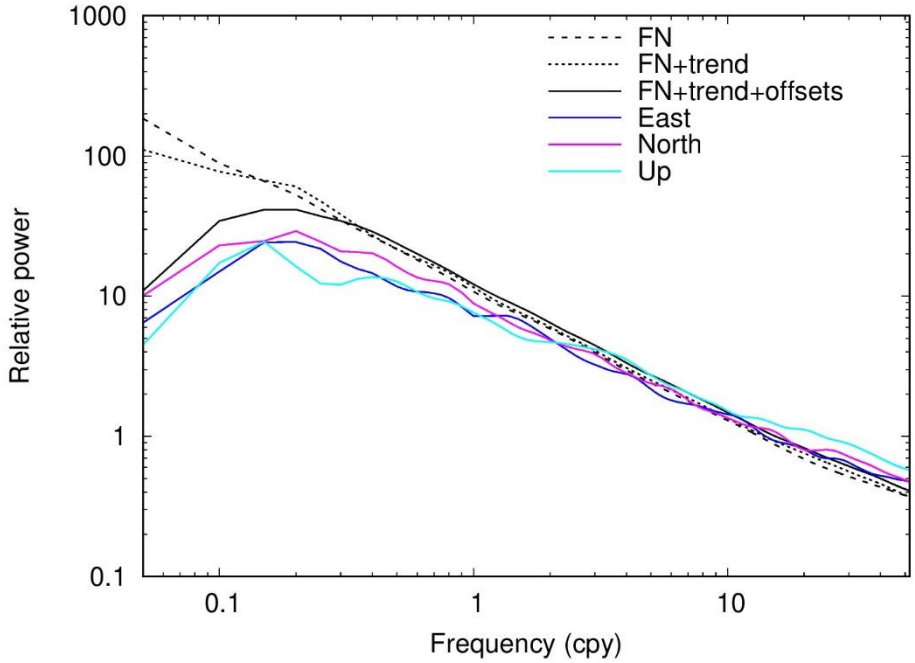


Figure 2. Comparison of simulated (in black color) and real (in colors) smoothed error power spectra. The real spectra for East, North and Up are the same as in Figure 1. The simulated spectra includes flicker noise (dashed), flicker noise with trend estimation (dotted) and flicker noise with trend and offsets estimation (solid).

4.3 Detectability of RW noise

The loss of power at long periods caused by estimating offsets hinders the analysis of the error sources that may contribute in that part of the spectrum, for instance, the RW noise. In addition to a RW contribution due to remaining offsets in the series (Williams, 2003a), it has been suggested that the instability of the antenna monument introduces some level of RW process noise in all series (Johnson & Agnew, 1995). For the best geodetic-class monuments, the amplitude of RW is expected to be several times smaller than the amplitude of FN and, therefore, it can only be detected in very long series or when the spatially-correlated FN is reduced, for instance in regional solutions and in short-baseline solutions (Dmitrieva et al., 2015; Hill et al., 2009; King & Williams, 2009; Langbein & Svarc, 2019). In our dataset of 127 stations of 20 years length we did not detect RW being significantly present in any of the series. Recent studies using network or PPP global solutions also report that RW does not play a significant role in the error spectrum (e.g., He et al. 2019).

The low proportion of detected RW in global solutions has been attributed to the dominance of FN and the series not being long enough (Williams et al., 2004). For instance, in order to detect the RW component over WN in the series, Langbein and Johnson (1997) determined that a series should be at least 5 times longer than the RW/WN cross-over period. The cross-over period between two noise processes depends on the series sampling and on the relative variance of each process. Following the equation of the PL noise power spectrum (Williams, 2003b), the cross-over period for weekly time series having FN amplitude of 2 mm yr $^{-0.25}$ and RW amplitude of 0.5 mm yr $^{-0.5}$ is around 14 years. Using the same equation, we obtain that for a 20-year-long weekly series with 2 mm yr $^{-0.25}$ FN, the theoretical maximum RW amplitude detectable would be around 0.4 mm yr $^{-0.5}$, which represents a FN/RW amplitude ratio of near 5.

Since the estimated position offsets absorb significant amounts of power in the band of the error spectrum where we expect RW to emerge, our chances to detect RW in long series are reduced. A lower probability to detect RW implies that the amount of RW potentially being hidden in the series could be quite large relative to the amount of FN. To quantify this, we ran a new noise analysis using 1000 synthetic weekly series of 20 years having FN+RW in different ratios, from 1 up to 6, and including different offset rates, from none up to one every two years. For different combinations of FN/RW ratios and offset rates, we computed the percentage of series for which the RW component was detected over the FN (true positives), which was still considered as the null hypothesis, using the dML threshold at 95 %. The obtained success rates were then interpolated to create a continuous representation in Figure 3. The sampled success rates are shown by cyan circles in Fig. 3. This figure shows that the success rate of detecting RW is strongly affected by the amplitude ratio FN/RW, as expected due to the spectral limitation imposed by the series length. The RW success rate estimated by MLE is close to zero for FN/RW ratios of 5.5, very close to the FN/RW theoretical limit of our synthetic series (cyan line in Fig. 3). This upper FN/RW bound decreases towards 4 as the offset rate increases. The offsets play a significant role in reducing the chances of RW detection, especially when the FN dominates over the RW with FN/RW ratios of 2 or higher. As an example, for a series having an offset rate between 0.3 and 0.4, which is a common scenario in the JPL series, a weekly series could accommodate a RW amplitude equal to 1/3 of the FN amplitude, i.e., typically ~0.7mm yr $^{-0.5}$ for the horizontal components and more than 2 mm yr $^{-0.5}$ for the vertical component, and yet, we will have around 20 % chance to detect this RW component. The success rate would rise up to more than 40 % in case *r* were 0, which represents an average probability reduction of around 4 % per offset. Lower RW amplitudes than these values have been reported in past studies (King & Williams, 2009), which implies even less chances to detect it; for instance, less than 5 % for a RW amplitude of 0.5 mm yr $^{-0.5}$ and 0.3 offsets yr $^{-1}$. Based on these results with synthetic series, we determined that

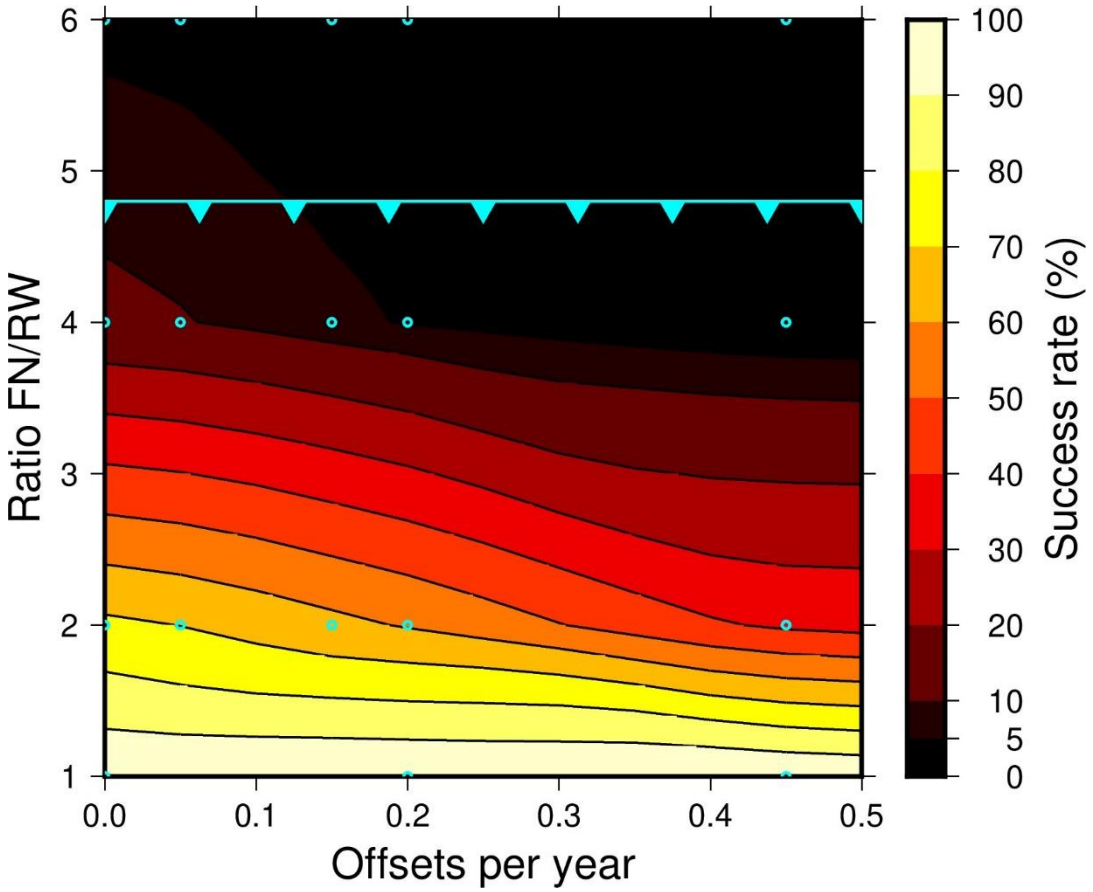


Figure 3. Success rate of random walk detection (RW) over flicker noise (FN) as a function of the amplitude ratio FN/RW (y-axis) and as a function of the rate of offsets per year (x-axis). The cyan line represents the maximum theoretical FN/RW ratio detectable in the synthetic series. The cyan circles represent the locations at which the RW success rate was sampled

the estimated formal uncertainty of the trend would increase by a factor between ~2 and ~3, with *r = 0.45* and *r = 0*, respectively, if a RW amplitude of 0.5 mm yr $^{-0.5}$ were to be detected over a FN amplitude of 2 mm yr $^{-0.25}$. Finally, we observed that there were up to 15 % of the series, mostly with a FN/RW ratio between 2 and 4, for which a significant non-zero RW amplitude was estimated. These results were however rejected (false negatives) because the penalized dML was not high enough to select the FN+RW model over the favored FN null hypothesis.

4.4 Impact of frame alignment on error spectra

It is clear that the functional model can severely alter the power spectrum of the residual series. Therefore, a natural question arises as to what would be the effect of the frame alignment on the colored noise of the series. The frame alignment is part of the functional model in any network solution and is also implicitly applied to any PPP solution via the fixed orbit and clock products from the network solution. To answer this question, we used 20 years of weekly IGS SINEX files starting in 1999. We modified the SINEX files by extracting the IGS14 core stations and replacing their weekly station coordinates by their linearized positions, extracted from the mean position and velocity of the IGS long-term frame, plus a synthetic FN series for each station. The idea is to quantify how much the FN series change after the frame alignment process.

The variance-covariance of the weekly network solutions was left intact in the SINEX files. By doing this, we assume the spatial correlation of position errors and its changes through time are independent of the temporal correlation of each individual station. This hypothesis was necessary as it is not practically possible to synthetize a covariance matrix that accounts for FN in both the

temporal and spatial domains at the same time in a least-squares network solution (Benoist et al., 2020). In addition, this assumption is commonly applied in most studies when the colored noise content in a given series is estimated with no regard for its spatial correlation.

We performed two cumulated solutions using CATREF (Altamimi et al., 2016) on these modified SINEX files, with and without the estimation of the transformation parameters of the frame alignment. In addition, we examined the impact of including/excluding a weekly scale offset and also the impact of the full/diagonal covariance matrix of the SINEX files, i.e., the spatial correlation. Our results indicate that neither the frame alignment nor the consideration of the spatial correlation have a significant impact on the retrieved FN series. The velocity estimation absorbs a small amount of the FN variance at the longest period, as expected, but the estimation of the transformation parameters does not alter significantly the background error spectrum of the residual series after the alignment (Fig. S3). This result also validates the hypothesis that the temporal and spatial correlations can be considered independently when analyzing the noise content of individual series. On a series-by-series basis, the FN amplitude changes typically less than 0.1 mm $yr^{-0.25}$ for both horizontal and vertical components, even if the added FN amplitude was ~3 times larger in the vertical component. The FN change is marginally larger in the vertical component if the scale parameter is estimated. For most of the stations, the FN amplitude increases in the horizontal component and decreases in the vertical component.

These results are considerably different to those recently obtained by Chen et al. (2018) who reported PL amplitude changes at least 2 times larger, while still considering they used daily series instead of weekly as in our case. They noted a clear spatial dependence of the PL changes, with the largest PL reduction in areas of high density of stations and a PL increase elsewhere. Since Chen et al. (2018) used the full IGS network to compute the transformation parameters, their result reflects a network effect, i.e., the transfer of common-mode PL noise between different areas of the network. Such a strong network effect was demonstrated by Ray et al. (2017) for the frame rotations. Our assessment of the alignment impact should be considered more robust as we used the well-distributed IGS14 core network instead of the full IGS network specifically for this reason.

4.5 Evolution of noise types over time

Concerning changes of the colored noise content with time, Williams et al. (2004) and Santamaría-Gómez et al. (2011) showed that the amplitude of the PL noise in the series is generally getting smaller with time. Bos et al. (2010) also questioned the hypothesis that noise properties are constant over time. In that case, the estimated noise content of a given series will depend on the period of time chosen. This dependency can void the conclusions from separate studies about the average noise level in a particular set of series if they do not cover the same time period. It could also have a significant impact on comparisons of noise levels between different sets of series (Amiri-Simkooei et al., 2017; Langbein & Svarc, 2019).

Here, we revisit this issue by using our selection of 127 series from the last JPL PPP solution having the same length, minimum gaps, homogeneous good-quality data from 1999 onwards, and almost twice as long as those in Santamaría-Gómez et al. (2011). We split the original 20-year series into 4 segments of 5 years each and ran separate noise analyses for each segment including the same functional and noise models as for the full JPL series. The selection of the noise model was done using the estimated dML thresholds at 95 % of Table 1.

We observe a change in the noise color throughout the last 20 years of GPS observations (Fig. 4). The proportion of selected FN series is progressively falling in favor of a PL with a spectral index considerably close to WN (Table 3). While it is difficult to extract a clear trend in the estimated noise parameters between the four segments, it is clear that the colored noise is whiter in the second half of the total period. Offsets are more frequent in the first half of the series than in the second, particularly before 2005. For instance, in 2002 and 2004, around 60% of the stations had at least one estimated offset. After 2005, the number of removed offsets stabilizes. If the noise amplitude and type were constant, we should expect the noise to get redder if less offsets were removed in the second half of the series, but we observe the contrary. This result confirms earlier findings that colored noise is becoming whiter for newer observations, but in a more general framework, i.e., considering also changes in the preferred noise model from FN towards PL. The whitening of the noise color can also be seen in the reduction of the median formal velocity uncertainties obtained exclusively from the parameters of the fitted noise models in each segment and component, i.e., not considering here the contribution of offsets to the uncertainty (Fig. 4d). We also note that in the segmented series, the GM model is much less present than the FN and PL models.

Table 3. Percentage of selected PL and FN models, with the median noise amplitude in mm (σ) and spectral index (k), for the four segments of 5 year length as in Figure 4. Values in red are computed from less than 30 series.

		East			North			Up		
		%	σ	k	%	σ	k	%	σ	k
PL	1999-2004	28	1.7	-0.47	17	1.6	-0.45	35	4.6	-0.47
	2004-2009	42	1.3	-0.55	44	1.3	-0.53	55	4.3	-0.47
	2009-2014	53	1.1	-0.43	53	1.2	-0.50	54	4.0	-0.47
	2014-2019	44	1.3	-0.54	56	1.3	-0.54	85	3.9	-0.43
FN	1999-2004	65	2.5	-1	75	2.7	-1	57	7.7	-1
	2004-2009	50	2.0	-1	44	2.0	-1	31	6.5	-1
	2009-2014	37	2.0	-1	38	2.1	-1	35	6.7	-1
	2014-2019	50	1.9	-1	37	2.0	-1	13	7.7	-1

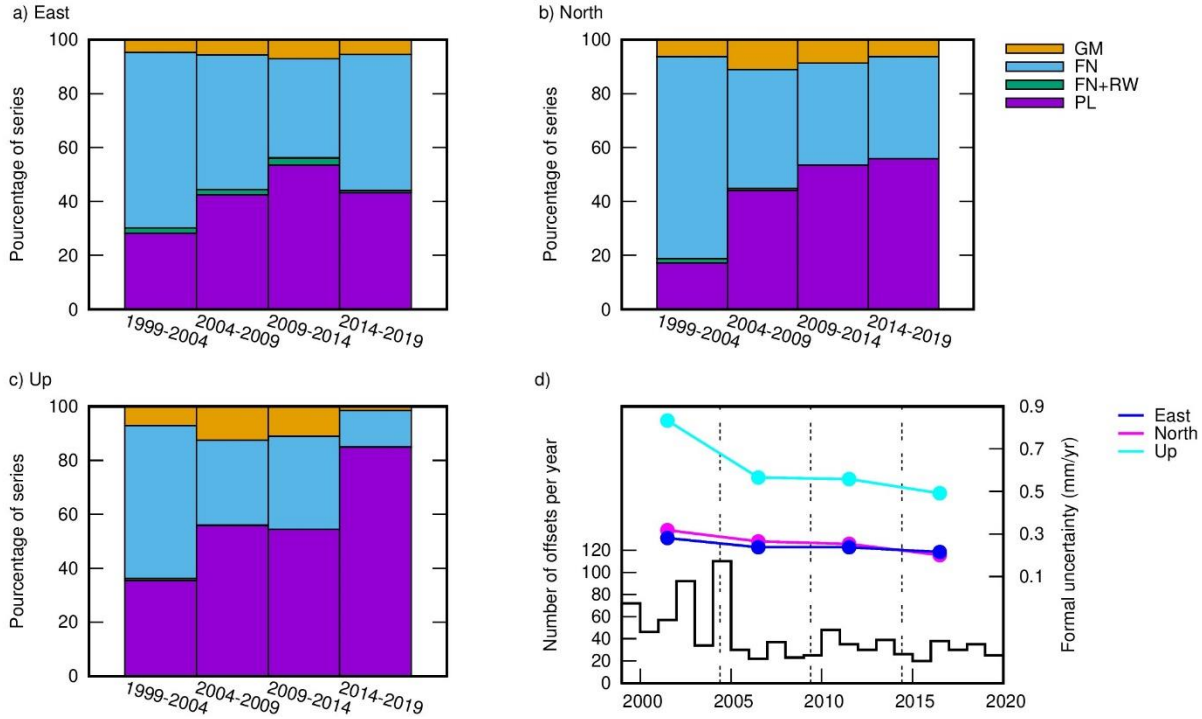


Figure 4. Distribution of the preferred noise model in the JPL series for 4 periods of 5 years each (x-axis) for the a) East, b) North and c) Vertical components, together with the d) distribution of offsets per year and the median formal uncertainty of velocities from the noise estimates for each period and component.

5. Discussion

5.1 Impact of offsets on the noise color

It is clear that offsets should be, first, avoided as much as possible and, second, removed from the series in order to get the least biased estimate of the velocity. The main limitation for reaching the second goal is that offsets are usually mixed and hidden by other features in the series, including colored noise. Manual offset detection by expert analysts generally outperforms automatic algorithms which tend to over-fit the series with spurious offsets (Gazeaux et al., 2013). Despite this finding, with the increasing number of series available and also their length, automatic offset detection is being used nowadays by many scientists as a purely practical matter and it is very likely that this trend will continue.

Leaving aside the question whether the removed offsets are correctly timed or not, we demonstrate how the offset estimation absorbs most of the low-frequency colored noise in the series, depending on the separation between the offsets. This translates into a different shape of the error spectrum and, if the effect is not accounted for, it would bias estimates of the colored noise content and hence velocity uncertainty. The impact is most dramatic in the case of a series with regularly spaced offsets every N years. In that case, the colored noise content of the series becomes unobservable beyond a period of approximately N years. For an irregular distribution of offsets, if their impact is not taken into account, the fit of a noise model will provide, at best, a biased lower noise amplitude and, at worst, it will indicate that the best fitted model is much whiter, for instance a PL model with spectral index closer to zero or even a GM model where the spectral index is zero at the longest periods.

These results confirm the more limited observations recently made by Chen et al. (2018). These authors considered changes in the estimated PL noise amplitude and spectral index as offsets were added to synthetic series of different length. Our results indicate an increase of the estimated PL spectral index, i.e. closer to zero, and a reduction of the colored noise amplitude of the same order as theirs. However, they did not investigate how the offsets affect the selection of the noise model. We also note that PL is not ranked as the best noise model describing our series, especially in the horizontal components, even if the impact of offsets in the dML values is not taken into account.

We observed that, even after taking into account the effect of offsets in the dML thresholds, about a quarter of the horizontal series, and half the vertical series, prefer a GM model (Table 2). We recognized in Section 2 that noise levels and types may vary among the series, which could partially explain the selection of the GM model in some series. However, this result may also be partly due to the limitation of the MLE method itself. Between the FN, PL or FN+RW noise models and a GM noise model, the MLE tends to favor the GM model, which is the one that provides the lowest noise power spectrum at long periods. A lower noise power spectrum at long periods produces a narrower variety of possible outcomes, and therefore, a higher probability of fitting the observed series. This situation is worsened by the presence of position offsets, which essentially reduce the longest periods observable in the series, increasingly favoring the GM model. Offsets reduce our ability to determine precisely the goodness-of-fit of competing noise models as reflected by the increased dML thresholds with the offset rate in Table 1. An additional limitation of the MLE method concerning the GM model is the unrealistically low velocity uncertainty it provides, which is likely caused by unaccounted correlation between the functional and the noise model parameters, especially the cross-over period (Langbein, 2020).

In their recent study, He et al. (2019) analyzed 110 series of at least 12 years length to investigate whether the power spectrum flattens at the longest periods. They found that for 90 % of the series,

the noise could be described by a GM process. This figure dropped to 5 % of the horizontal and 13 % of the vertical series when excluding the GM fits having a cross-over period between the Gaussian and the Markov-like processes smaller than 1 year. From the JPL series in our study, we find that, depending on the coordinate component, between 48 % and 76 % of the series would be better described by a GM process if the impact of offsets in the dML thresholds were not accounted for (Table 2). These numbers are reduced to 27 % and 35 %, respectively if we exclude the GM fits having a cross-over period smaller than 1 year. These results are comparable to those we obtained with the dML thresholds at 95 % including the offset effect (Table 2), which would support the arbitrary threshold on the cross-over period applied in past publications (He et al., 2019; Santamaría-Gómez et al., 2011). However, the percentage of series preferring a GM noise model is still 3 to 5 times larger in our results compared to those from He et al. (2019). To choose the best-fit model, He et al. (2019) used a modified Bayesian Information Criterion penalty function that was added to the logarithm of the maximum likelihood estimate. This penalty function is slightly different than the empirical log-likelihood thresholds that we used when the offset impact is not accounted for (*r = 0* in Table 1), which would not explain these differences.

We may instead explain the difference of our results with those from He et al. (2019) due to a combination of the following reasons. First, they considered series of different lengths and only 20 out of their 110 stations had the maximum length of 20 years, compared to all 127 in our study. If we consider that the GM model was mainly retained in their longest 20 stations, then their GM ratios would rise closer to our results. Second, it is very likely that their series had a different set of estimated offsets than those in our series. Since almost all of their fitted GM models had a cross-over period shorter than a year, it would indicate that the power in many of their error spectra was decaying at shorter periods than our spectra, i.e., their series may have been fitted by more offsets than ours. Third, they used a different software for the noise content analysis that provides slightly lower spectral indices for short series (Bos et al., 2013). It is unknown to us if their software would also provide shorter cross-over periods for the GM model. Fourth, they analyzed PPP series from an older version of the same GNSS software with, importantly, obsolete orbits and clock products consistent with the IGS 1st reprocessing campaign. Those products may possess higher levels of FN (Amiri-Simkooei et al., 2017).

Since offsets can potentially bias the observed noise amplitude and even the noise color, if not accounted for, they might negate the conclusions drawn from the comparison of noise levels between different solutions (Amiri-Simkooei et al., 2017) or different monument types (Langbein & Svarc, 2019). In order to obtain comparable noise levels between different series, we recommend using similar sampling, including the series length and gaps, and also, if possible, the same set of offsets. The latter may not be possible if an offset is generated by incorrect metadata during the reduction of the GPS observations or if the series belong to different instruments with different antenna monuments and tracking histories.

We foresee that, if the current frequency of estimated position offsets does not change or even increases in the future, as the series get longer the error power spectrum will continue to drop at long periods as we have shown in Fig. 2, i.e., the longest period available to estimate the noise variance will not follow the increasing series length. This will prevent us from obtaining further information on the interannual to decadal error band compared to the information we already have today. Like a chameleon, we will never see the true noise color of the GPS position errors at such long periods. For instance, the power spectrum drop caused by the estimated offsets seriously limits our ability to address relevant scientific questions such as the expected periodic Earth deformation at long periods recently modelled by Ding et al. (2020). A signal with a period of ~5.9 years lies very

close to the period where the power spectrum is mostly flat and starts decaying in the best and longest JPL series available today. We have not been able to find any sign of a significant spectral peak emerging out of the noise in this band in any coordinate component (Fig. 1), which confirms a recent modeling of the mechanism of this signal (Gillet et al., 2020).

In addition, due to the presence of offsets, the resulting error spectrum significantly deviates from the expected null hypothesis of a FN. The consequence is that since the error spectrum is not observable at the longer periods, the null hypothesis for the background noise cannot be well defined. We lack enough offset-free series that would indicate whether the GPS error spectra deviate from a FN process at the longest periods: the error spectrum could continue unabated, still following a FN process; it could flatten, indicating that the process generating the FN ceased being dominant; or it could become steeper, indicating that another process noise overcomes the FN, for instance RW noise caused by the monument instability.

A consequence of this is that it will be necessary to use a heuristic null hypothesis for the noise model, not entirely supported by the observations, in order to obtain less biased velocity uncertainty estimates. Somehow, we are already seeing this scenario today when the impact of offsets is not fully accounted for and the best fitting model is a GM process that is arbitrarily rejected because of the unrealistically low formal velocity uncertainty it provides. The estimated position offsets also contribute to hide relatively small RW amplitudes in global GPS solutions as shown in Fig. 3. To avoid this situation, a more conservative approach to estimate the velocity uncertainty would be to fix the noise model to FN and also adding a relatively small amount of RW, compared to the estimated FN, to compensate to some extent for both the unobserved RW noise and also the small reduction of the estimated FN amplitude by the offsets.

5.2 Impact of offsets on velocities

Beyond the impact on the estimated noise level and color, the offsets also have a significant impact on the estimated velocity. Indeed, there exists a threshold on the offset size below which correcting the offset may produce a larger velocity error than if the offset was left uncorrected (Gazeaux et al., 2013; Griffiths & Ray, 2016). Williams (2003) demonstrated that an offset in the middle of the series introduces the largest velocity error compared to a different location in the series. If the offset is located near the beginning or end of the series, the velocity error being added becomes negligible compared to the velocity error without the offset. A similar scenario results if more than one offset is added very close to the beginning or end of the series, provided the length of the series before or after the offsets does not change significantly. In other words, the velocity error obtained from the fitted functional model is inversely proportional to the length of the continuous series between the offsets and not proportional to the number of offsets in the series. Certainly, when the offset is in the middle, the longest continuous segment of the series is the shortest of any other offset location possible, and therefore the velocity error is the largest possible.

In the case of a series free of offsets for N years, adding a new offset at the end will barely affect the present velocity error. However, after adding an offset, as the series length continues to grow, the velocity error will decrease much slower than if the offset had not been added. The maximum impact of the added offset on the velocity error will occur approximately when the series reaches a length of twice the current length (2N years), or slightly earlier in case the series is dominated by FN according to Williams (2003). The actual velocity error added by the offset will be indirectly proportional to the length N. From the epoch of maximum impact around 2N, as the series gets longer, the velocity error will decrease until the impact of the added offset vanishes more than 4 or 5 times N years in the future. Therefore, when planning a configuration change that would likely introduce a new offset

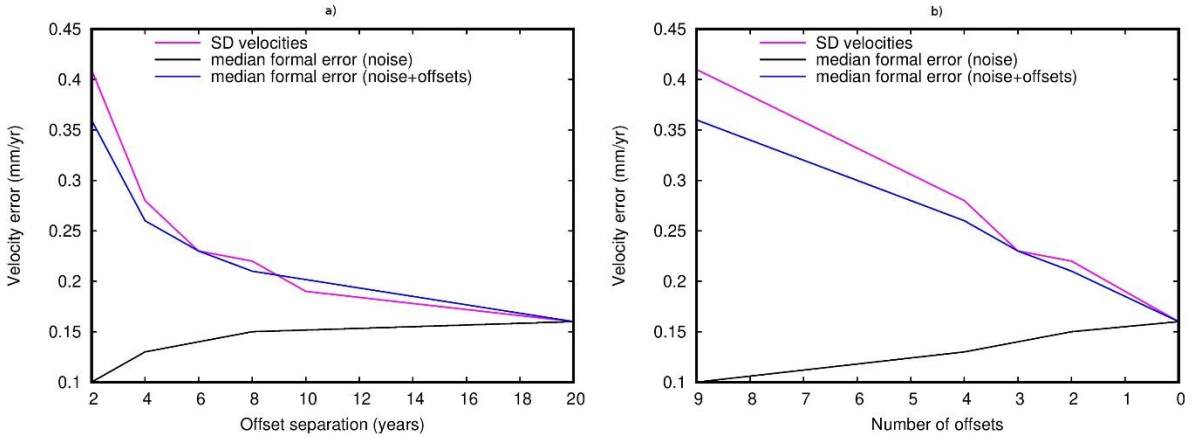


Figure 5. Velocity errors from 1000 synthetic series of 20 years having 6 mm yr -0.25 of FN as a function of a) the separation between offsets of 2, 4, 6, 8, 10 and 20 years, and b) their equivalent number of total offsets in the series, i.e., 9, 4, 3, 2, 1 and 0 offsets (right). Velocity errors are provided from the scatter of the velocity estimates (SD velocities) or from the formal velocity uncertainties from a FN model only (noise) and the FN model plus the functional model (trend+offsets).

into a series, the station manager should consider the increased velocity error that will result in the long-term, especially when the series is approximately twice the current length. For series having several offsets, the velocity error will accumulate the effects of the different offsets.

When the offsets are equally spaced in the series, as typically in simulation studies, both the separation between the offsets and number of offsets provide equivalent assessments of the velocity errors created by offsets, though their representation is still different and may lead to confusion. Figure 5 shows different metrics for the velocity errors from zero-trend synthetic FN series having different numbers of estimated offsets. This figure shows that in the case of considering the total number of equally-spaced offsets (right panel), the velocity error increases linearly. However, in case of considering the equivalent separation between equally-spaced offsets (left panel) the velocity error increases similarly to a RW process generated exclusively by the functional model being fitted. We note that the same velocity errors from the same set of offsets are shown in both panels of Fig. 5, only the independent variable in the X axis of the panels is different. The functions used to predict the increase of the velocity error with offsets are different depending on the chosen independent variable, i.e., quadratic or linear. This reconciles the apparently contradictory results from WH19, who considered the total number of offsets in the series, with those obtained earlier by GR16, who consecutively split the available series in half and considered the rate of offset occurrence, i.e., the inverse of their separation.

Another major difference between the findings by GR16 and those by WH19 is that while the former discussed the impact of offsets on the reference frame realization via a full network inversion assuming WN data, the latter discussed the impact of offsets on the formal velocity errors while using different hypotheses for the noise model. To explain their different results, WH19 emphasized the importance of considering colored noise against WN when estimating the formal velocity error and its change with offsets. They created synthetic series with different types of noise processes, some of which, for instance WN or RW series, are not reasonable examples of actual GPS series. They concluded that, depending on the noise type of the series, the impact of adding a single offset on the formal velocity errors can be quite different. The actual velocity errors are dependent on both the number of offsets and the type and amplitude of the colored noise, which should be generally consistent with a power-law noise spectrum close to FN. In their study, GR16 employed actual IGS series, which contain FN, to observe how the estimated velocity changes by increasingly doubling the rate of artificially inserted position offsets. When the number of offsets in the series is relatively high the scatter of the velocity estimates depends little on the type and amplitude of the colored noise

used to estimate the velocity and its formal error, provided the noise type is consistent with FN. This is due to the larger contribution to the velocity error by the offsets than by the colored noise. This fact is not observed if the noise in the series is consistent with RW, which, we insist, is not a realistic noise model to describe real GPS series. In addition, in their analysis with the simulated series, WH19 estimated the impact of offsets on the formal velocity errors without re-estimating the apparent change of the noise type and its parameter values after adding the offsets; they used the same covariance matrix as the one used for creating the synthetic series in the first place. Therefore, their estimated formal velocity errors do not account for the fact that the error power spectrum at long periods is severely reduced by the added offsets as we have demonstrated in this study. For instance, we show in Fig. 5 that, while the scatter of the estimated velocities assuming WN increases with the number of offsets, the contribution of the colored noise to the formal velocity error decreases because the fitted noise model is progressively whiter. This means that the findings by WH19, with respect to the synthetic series, do not describe the complete impact offsets have on GPS velocities as we do in our study.

By adding equally-spaced artificial offsets into real IGS position series, GR16 found that the velocity error due to the actual rate of offsets removed from the series, i.e., their average span separation, is at the level of 0.09 and 0.34 mm yr^{-1} for the horizontal and vertical components, respectively. These values correspond to the average IGS rate of offset occurrence at that time. In our study, using synthetic FN series that reasonably match the JPL PPP series (Fig. 2), we find velocity errors at the level of 0.08 and 0.21 mm yr^{-1} due to the offsets that JPL removed from the series. These values were obtained from the scatter of the fitted trends with and without the JPL offsets while also fitting the parameters of a FN process. If we consider WN instead of FN, as GR16 did, the scatter of velocities changes by less than 0.02 mm yr^{-1}, supporting the quantification of velocity errors by GR16. Our numbers are smaller most probably due to the fact that colored noise in the JPL series is lower than in the IGS series, and also due to the fact that we considered synthetic series with real irregularly-distributed offsets, while GR16 considered real series with artificial equally-spaced offsets, which tends to increase the error. For completeness, the median formal velocity uncertainty of the real JPL series used in this study, taking into account the offsets and also a FN model, is at the level of 0.09 and 0.29 mm yr^{-1}, respectively. These number were obtained from the real series instead of the synthetic FN series, but very likely, most of the estimated formal error could be explained by the above-mentioned contribution of the offsets alone and marginally by the contribution of the noise model, contrary to the conclusions by Wang and Herring (2019).

Also importantly for velocity estimates is that, due to the removed trend and offsets, it is very difficult to say how much RW truly contributes to the series variance. At most, given the separation between offsets in the series, we provided in Fig. 3 the probability that the RW does not exceed a given threshold. Inversely, by setting a detection threshold, we can assess the maximum amount of RW that could be present in the series. An alternative approach was followed by Langbein (2012) who estimated the upper bound of the RW variance by sequentially increasing it until the dML value of a FN+RW model with respect to a FN model exceeded the value of 2.6. This dML threshold agrees well with our values in Table 1. At the end, the estimated formal uncertainty of the velocity becomes a subjective choice, i.e., depending on how conservative velocity uncertainties one prefers to obtain, a different amount of RW noise could be assumed up to a reasonable limit. Despite RW noise having been detected in several regional or very-short baseline solutions, most of the studies using global GPS solutions have chosen to not consider RW noise in the velocity uncertainty budget, certainly because it has not been detected as a main contributor of the error spectrum. A remarkable exception is Langbein (2012) who proposed using a noise model composed of FN and RW noises instead of the typical PL noise model to obtain a more conservative velocity error budget. We argue

that, with the state-of-the-art GPS series in our study, we cannot rule out the possibility that the RW noise is indeed affecting the velocity estimates even if we were not able to detect it. The GPS velocity errors of a global solution would be much larger if RW noise were to be considered, especially in the vertical component where the larger series variability allows for a larger RW noise to pass undetected. The potential geophysical consequences include, for example, the long-term stability of the terrestrial reference frame (Altamimi et al., 2016), assessing recent sea-level changes (Santamaría-Gómez et al., 2017) or our understanding of the Earth rheology and the glacial history (Schumacher et al., 2018) rely on accurate GPS velocity errors.

In light of the broad impact that offsets have on the GPS velocities as shown for the first time in this study, our recommendation is that if the station manager wants to change an already running GPS antenna, for instance to augment the tracking capabilities of the station, the antenna change should be made immediately after an unplanned offset occurred, e.g., an earthquake that displaced the antenna. Otherwise, we recommend installing a new station nearby to benefit from the already existing infrastructure like access, security, power, communications, etc. As a general rule, a functioning antenna must never be touched for the interest of science, contrary to the advice of WH19.

5.3 Origin of colored noise

The underlying physical origin of colored GPS noise and its change in time is still an open question. Santamaría-Gómez et al. (2011, 2013) investigated the impact of the evolving tracking network, the evolving satellite constellation and the evolving rate of fixed phase ambiguities. None of these effects could explain the observed decrease in colored noise amplitude with time. Here we have demonstrated that the frame alignment process also cannot explain either the observed colored noise or its change with time.

Using synthetic series with different proportions of FN, we were unable to explain the stacked error spectra resulting from the JPL residual series. The real series exhibit spectra that are significantly lower in power at the longest periods even if we account for the damping of the colored noise caused by the estimation of position offsets. The mismatch between the real stacked spectra and the simulated FN spectrum is more evident for the vertical component than for the horizontal (Fig. 2). Spectral indices of the horizontal components have been reported to be closer to flicker noise than the vertical component (Rebischung et al., 2016b; Santamaría-Gómez et al., 2011; Williams et al., 2004). Although surface load mass variations can introduce significant colored noise in the vertical component (Santamaría-Gómez & Mémin, 2015), the fact that flicker noise is visible more clearly in the horizontal component argues against the hypothesis of flicker noise being mainly created by surface mass load variations, which are significantly more important on the vertical component than in the horizontal. Another concurrent indication is the fact that the colored noise content is getting whiter with time even if non-tidal loading corrections are not currently included in the conventional corrections applied to the series and their effects are most likely not decaying with time.

What causes a spectral index that is whiter than flicker is as intriguing as what causes the FN itself. This situation can probably only be explained by a combination of different error sources, i.e., we are looking for a process that dampens FN in the vertical component while having minimal impact on the horizontal components. We demonstrated that the frame alignment process certainly contributes in the same direction, but its magnitude is rather negligible. Actually, surface mass load variations are a good candidate for such a process. Santamaría-Gómez and Mémin (2015) showed that the combination of modeled atmospheric, oceanic and hydrologic load variations results in colored noise that can be described by a PL with a spectral index varying spatially from WN to RW (see map of

figure S1 in Santamaría-Gómez and Mémin (2015)). The median spectral index of the modeled load variations for the vertical component at the 127 JPL series used in this study is −0.7, i.e. significantly whiter than a flicker process. In addition, the median amplitude of the colored noise obtained from the modeled surface loads at the 127 stations and scaled to weekly sampling is ~4 mm. This equals almost two-thirds of the observed PL amplitude of ~7 mm in the residual weekly vertical JPL series, which implies that the surface mass load variations add noise to the series, mostly on the vertical component, but the noise added is whiter than FN. One may conclude that the surface mass loads, contrary to causing the FN observed in the series, probably contribute to hide it, making the vertical series appear whiter than the horizontal components.

6. Conclusions

Offsets in GPS position time series are one of the biggest limitations concerning the precision of the estimated GPS velocities. Yet, the best GPS series are populated with offsets, many of which are related to equipment testing and upgrading which are avoidable. The use of algorithms for automatic offset detection may even increase the number of spurious offsets that are included in the model fitted to the GPS series. Here, we address for the first time the complete impact that estimated position offsets have on the GPS velocities, their uncertainty and the colored noise of the residual series.

We have shown how the estimated offsets absorb most of the series variability at periods longer than their average spacing. The result is that at interannual to decadal periods, the noise power spectra of the GPS position time series changes color, like a chameleon, and becomes uninterpretable as a quantitative measure of the contributing noise processes. The more offsets introduced into the series, the less information is obtained from the error spectrum. This could severely affect our ability to objectively provide a noise model that fully describes the temporal correlation of the GPS series. We have shown how the offsets effectively increase the degrees of freedom of the model fitted to the series, resulting in an inflated dML threshold necessary to reject the flicker noise null hypothesis. If the correct dML threshold is not used, any colored noise model fitted to GPS series having offsets will be biased low, i.e. towards WN, and the preferred noise model may even be wrongly selected, e.g. a Gauss-Markov process instead of a flicker or power-law process. As a result, the estimated formal velocity uncertainty will be biased low. If one is interested in obtaining conservative formal velocity uncertainties, it is necessary to subjectively choose a noise model that is not clearly supported by the observed error spectra, for example a FN model topped with a relatively small amount of RW.

The subjectivity to choose a more accurately representative noise model also concerns the amount of RW that the analyst is willing to add to the velocity uncertainty budget. On most past studies using global network or PPP solutions, the RW has never been included in the velocity error budget because it has not been detected over other noise types. We have shown how the inclusion of offsets in the series renders the detection of RW even more difficult in long series that actually contain RW. We conclude that the common assumption that RW does not contribute to the velocity error is questionable, especially considering that RW has been detected in regional or short-baseline solutions with greater sensitivity due to spatial filtering. The addition of a small quantity of RW, even if not observed directly, will significantly raise the GPS velocity uncertainty and especially in long series with frequent offsets where the amount of RW that could be allocated in the error budget is larger.

The unobservability of the noise power spectra at long periods caused by the estimated position offsets also has severe consequences for our ability to observe and discover long-period Earth deformation signals. A periodic Earth deformation signal with a period of ~5.9 years has been suggested in recent literature. This period is exactly located where the error spectrum is obscured by the estimated position offsets in the best and longest JPL series to date used in this study. Despite recent claims for the detection of this signal, we were unable to find any sign of a significant signal emerging out of the noise in this band in any coordinate component.

The origin of colored noise in GPS position time series and its change with time is still an open question. We have demonstrated how the frame alignment process, that is inherent to the construction of any GPS series, does not play a significant role if it is realized in an optimal way by minimizing the network effect. Surface mass loads could at least partly explain why the noise is whiter in the vertical component than in the horizontal, but they cannot explain why the noise color is becoming whiter in the most recent data. In addition, the offsets already present in the GPS series prevent us from observing the properties of the colored noise at long periods. If offsets continue to appear in the series at the current rate, they will represent a major limitation for our understanding of the GPS error spectrum reducing the value of longer series in the future.

References

Altamimi, Z., Rebischung, P., Métivier, L., & Collilieux, X. (2016). ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions. *Journal of Geophysical Research: Solid Earth*, *121*, 6109–6131. https://doi.org/10.1002/2016JB013098

Amiri-Simkooei, A. R., Tiberius, C. C. J. M., & Teunissen, P. J. G. (2007). Assessment of noise in GPS coordinate time series: Methodology and results. *Journal of Geophysical Research: Solid Earth*, *112*(B7). https://doi.org/10.1029/2006JB004913

Amiri-Simkooei, A. R., Mohammadloo, T. H., & Argus, D. F. (2017). Multivariate analysis of GPS position time series of JPL second reprocessing campaign. *Journal of Geodesy*, *91*(6), 685–704. https://doi.org/10.1007/s00190-016-0991-9

Benoist, C., Collilieux, X., Rebischung, P., Altamimi, Z., Jamet, O., Métivier, L., et al. (2020). Accounting for spatiotemporal correlations of GNSS coordinate time series to estimate station velocities. *Journal of Geodynamics*, 101693. https://doi.org/10.1016/j.jog.2020.101693

Bos, M. S., Bastos, L., & Fernandes, R. M. S. (2010). The influence of seasonal signals on the estimation of the tectonic motion in short continuous GPS time-series. *WEGENER 2008* -

Proceedings of the 14th General Assembly of Wegener, *49*(3), 205–209. https://doi.org/10.1016/j.jog.2009.10.005

Bos, M. S., Fernandes, R. M. S., Williams, S. D. P., & Bastos, L. (2013). Fast error analysis of continuous GNSS observations with missing data. *Journal of Geodesy*, *87*(4), 351–360. https://doi.org/10.1007/s00190-012-0605-0

Chen, G., Zhao, Q., Wei, N., & Liu, J. (2018). Impacts on Noise Analyses of GNSS Position Time Series Caused by Seasonal Signal, Weight Matrix, Offset, and Helmert Transformation Parameters. *Remote Sensing*, *10*(10). https://doi.org/10.3390/rs10101584

Collilieux, X., van Dam, T., Ray, J., Coulot, D., Métivier, L., & Altamimi, Z. (2012). Strategies to mitigate aliasing of loading signals while estimating GPS frame parameters. *Journal of Geodesy*, *86*(1), 1–14. https://doi.org/10.1007/s00190-011-0487-6

Ding, H., Xu, X., Pan, Y., Jiang, W., & Van Dam, T. (2020). A Time-Varying 3-D Displacement Model of the ~5.9-Year Westward Motion and its Applications for the Global Navigation Satellite System Positions and Velocities. *Journal of Geophysical Research: Solid Earth*, *125*(4), e2019JB018804. https://doi.org/10.1029/2019JB018804

Dmitrieva, K., Segall, P., & DeMets, C. (2015). Network-based estimation of time-dependent noise in GPS position time series. *Journal of Geodesy*, *89*(6), 591–606. https://doi.org/10.1007/s00190-015-0801-9

Gazeaux, J., Williams, S., King, M., Bos, M., Dach, R., Deo, M., et al. (2013). Detecting offsets in GPS time series: First results from the detection of offsets in GPS experiment. *Journal of Geophysical Research: Solid Earth*, *118*, 2397–2407. https://doi.org/10.1002/jgrb.50152

Gillet, N., Dumberry, M., & Rosat, S. (2020). The limited contribution from outer core dynamics to global deformations at the Earth's surface. *Geophysical Journal International*, (ggaa448). https://doi.org/10.1093/gji/ggaa448

Griffiths, J., & Ray, J. (2016). Impacts of GNSS position offsets on global frame stability. *Geophysical Journal International*, *204*, 480–487. https://doi.org/10.1093/gji/ggv455

He, X., Bos, M. S., Montillet, J. P., & Fernandes, R. M. S. (2019). Investigation of the noise properties at low frequencies in long GNSS time series. *Journal of Geodesy*, *93*(9), 1271–1282. https://doi.org/10.1007/s00190-019-01244-y

Herring, T. A., Melbourne, T. I., Murray, M. H., Floyd, M. A., Szeliga, W. M., King, R. W., et al. (2016). Plate Boundary Observatory and related networks: GPS data analysis methods and geodetic products. *Reviews of Geophysics*, *54*(4), 759–808. https://doi.org/10.1002/2016RG000529

Hill, E. M., Davis, J. L., Elósegui, P., Wernicke, B. P., Malikowski, E., & Niemi, N. A. (2009). Characterization of site-specific GPS errors using a short-baseline network of braced monuments at Yucca Mountain, southern Nevada. *Journal of Geophysical Research: Solid Earth*, *114*(B11). https://doi.org/10.1029/2008JB006027

Johnson, H. O., & Agnew, D. C. (1995). Monument motion and measurements of crustal velocities. *Geophysical Research Letters*, *22*(21), 2905–2908. https://doi.org/10.1029/95GL02661

King, M. A., & Williams, S. D. P. (2009). Apparent stability of GPS monumentation from short-baseline time series. *Journal of Geophysical Research: Solid Earth*, *114*(B10). https://doi.org/10.1029/2009JB006319

Langbein, J. (2004). Noise in two-color electronic distance meter measurements revisited. *Journal of Geophysical Research: Solid Earth*, *109*(B4). https://doi.org/10.1029/2003JB002819

Langbein, J. (2012). Estimating rate uncertainty with maximum likelihood: differences between power-law and flicker–random-walk models. *Journal of Geodesy*, *86*(9), 775–783. https://doi.org/10.1007/s00190-012-0556-5

Langbein, J. (2020). Methods for Rapidly Estimating Velocity Precision from GNSS Time Series in the Presence of Temporal Correlation: A New Method and Comparison of Existing Methods. *Journal of Geophysical Research: Solid Earth*, *125*(7), e2019JB019132. https://doi.org/10.1029/2019JB019132

Langbein, J., & Johnson, H. (1997). Correlated errors in geodetic time series: Implications for time-dependent deformation. *Journal of Geophysical Research: Solid Earth*, *102*(B1), 591–603. https://doi.org/10.1029/96JB02945

Langbein, J., & Svarc, J. L. (2019). Evaluation of Temporally Correlated Noise in Global Navigation Satellite System Time Series: Geodetic Monument Performance. *Journal of Geophysical Research: Solid Earth*, *124*(1), 925–942. https://doi.org/10.1029/2018JB016783

Legrand, J., Bergeot, N., Bruyninx, C., Wöppelmann, G., Bouin, M.-N., & Altamimi, Z. (2010). Impact of regional reference frame definition on geodynamic interpretations. *WEGENER 2008 - Proceedings of the 14th General Assembly of Wegener*, *49*(3), 116–122. https://doi.org/10.1016/j.jog.2009.10.002

Mao, A., Harrison, C. G. A., & Dixon, T. H. (1999). Noise in GPS coordinate time series. *Journal of Geophysical Research: Solid Earth*, *104*(B2), 2797–2816. https://doi.org/10.1029/1998JB900033

Ray, J., Griffiths, J., Collilieux, X., & Rebischung, P. (2013). Subseasonal GNSS positioning errors. *Geophysical Research Letters*, *40*(22), 5854–5860. https://doi.org/10.1002/2013GL058160

Ray, J., Rebischung, P., & Griffiths, J. (2017). IGS polar motion measurement accuracy. *Geodesy, Astronomy and Geophysics in Earth Rotation*, *8*(6), 413–420. https://doi.org/10.1016/j.geog.2017.01.008

Rebischung, P., Altamimi, Z., Ray, J., & Garayt, B. (2016a). Error Analysis of the IGS repro2 Station Position Time Series (Vol. 2016). Presented at the IGS Workshop. Retrieved from http://www.igs.org/presents/workshop2016

Rebischung, P., Altamimi, Z., Ray, J., & Garayt, B. (2016b). The IGS contribution to ITRF2014. *Journal of Geodesy*, *90*, 611–630. https://doi.org/10.1007/s00190-016-0897-6

Santamaría-Gómez, A. (2019). SARI: interactive GNSS position time series analysis software. *GPS Solutions*, *23*(2), 52. https://doi.org/10.1007/s10291-019-0846-y

Santamaría-Gómez, A., & Mémin, A. (2015). Geodetic secular velocity errors due to interannual surface loading deformation. *GEOPHYSICAL JOURNAL INTERNATIONAL*, *202*(2), 763–767. https://doi.org/10.1093/gji/ggv190

Santamaría-Gómez, A., Bouin, M. N., Collilieux, X., & Wöppelmann, G. (2011). Correlated errors in GPS position time series: Implications for velocity estimates. *JOURNAL OF GEOPHYSICAL RESEARCH-SOLID EARTH*, *116*. https://doi.org/10.1029/2010JB007701

Santamaría-Gómez, A., Bouin, M. N., Collilieux, X., & Wöppelmann, G. (2013). Time-Correlated GPS Noise Dependency on Data Time Period. In Altamimi, Z and Collilieux, X (Ed.), *REFERENCE FRAMES FOR APPLICATIONS IN GEOSCIENCES* (Vol. 138, pp. 119–124). HEIDELBERGER PLATZ 3, D-14197 BERLIN, GERMANY: SPRINGER-VERLAG BERLIN. https://doi.org/10.1007/978-3-642-32998-2_19

Santamaría-Gómez, A., Gravelle, M., Dangendorf, S., Marcos, M., Spada, G., & Wöppelmann, G. (2017). Uncertainty of the 20th century sea-level rise due to vertical land motion errors. *Earth and Planetary Science Letters*, *473*, 24–32. https://doi.org/10.1016/j.epsl.2017.05.038

Scargle, J. D. (1982). Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data. *Astrophysical Journal, Vol. 263, p. 835-853 (1982)*, *263*, 835–853. https://doi.org/10.1086/160554

Schumacher, M., King, M. A., Rougier, J., Sha, Z., Khan, S. A., & Bamber, J. L. (2018). A new global GPS data set for testing and improving modelled GIA uplift rates. *Geophysical Journal International*, *214*(3), 2164–2176. https://doi.org/10.1093/gji/ggy235

Wang, L., & Herring, T. (2019). Impact of Estimating Position Offsets on the Uncertainties of GNSS Site Velocity Estimates. *Journal of Geophysical Research: Solid Earth*, *124*(12), 13452–13467. https://doi.org/10.1029/2019JB017705

Wdowinski, S., Bock, Y., Zhang, J., Fang, P., & Genrich, J. (1997). Southern California permanent GPS geodetic array: Spatial filtering of daily positions for estimating coseismic and postseismic

displacements induced by the 1992 Landers earthquake. *Journal of Geophysical Research: Solid Earth*, *102*(B8), 18057–18070. https://doi.org/10.1029/97JB01378

Williams, S. D. P. (2003a). Offsets in Global Positioning System time series. *Journal of Geophysical Research: Solid Earth*, *108*(B6). https://doi.org/10.1029/2002JB002156

Williams, S. D. P. (2003b). The effect of coloured noise on the uncertainties of rates estimated from geodetic time series. *Journal of Geodesy*, *76*(9), 483–494. https://doi.org/10.1007/s00190-002-0283-4

Williams, S. D. P. (2008). CATS: GPS coordinate time series analysis software. *GPS Solutions*, *12*, 147–153. https://doi.org/10.1007/s10291-007-0086-4

Williams, S. D. P., Bock, Y., Fang, P., Jamason, P., Nikolaidis, R. M., Prawirodirdjo, L., et al. (2004). Error analysis of continuous GPS position time series. *Journal of Geophysical Research: Solid Earth*, *109*. https://doi.org/10.1029/2003JB002741

Zhang, J., Bock, Y., Johnson, H., Fang, P., Williams, S., Genrich, J., et al. (1997). Southern California permanent GPS geodetic array: Error analysis of daily position estimates and site velocities. *Journal of Geophysical Research: Solid Earth*, *102*(B8), 18035–18055. https://doi.org/10.1029/97JB01380

Zumberge, J. F., Heflin, M. B., Jefferson, D. C., Watkins, M. M., & Webb, F. H. (1997). Precise point positioning for the efficient and robust analysis of GPS data from large networks. *Journal of Geophysical Research: Solid Earth*, *102*(B3), 5005–5017. https://doi.org/10.1029/96JB03860

Journal of Geophysical Research: Solid Earth

Supporting Information for

Chameleonic noise in GPS position time series

Alvaro Santamaría-Gómez

GET, Université de Toulouse, CNES, CNRS, IRD, UPS, Toulouse, France.

Jim R. Ray

National Geodetic Survey (retired), Silver Spring, MD 20910, USA.

Contents of this file

Table S1.
Figure S1.
Figure S2.
Figure S3.

Introduction

This Supporting Information file contains the list (Table S1) of 129 GPS stations selected from the online JPL PPP service at https://sideshow.jpl.nasa.gov/post/series.html (accessed in May 2019) together with a statistical description of their quality. Stations that passed the minimum data requirements for this study, but were later rejected because of apparent noise changes in the series are: USUD, MEDI, AVRY, GENO, LEWI, NEAH, RAMO, SFER, CAND and POMM.

In addition to the position offsets removed by JPL, we removed offsets at the following epochs. Some of these offsets are related to known equipment changes while others are unknown.

ALIC 2010.89829311503, 2011.54451086877
DAEJ 1999.72756402614, 2002.71215373097, 2004.15718650034
GRAZ 2004.38109882606
HOB2 2009.34226097444
MPWD 2002.75391153433, 2004.39076167378
REYK 2007.70890954772, 2008.45415801505

TORI 2019.2511936635
YSSK 2000.07480112907

Tables

Table S1. List of the 127 stations used in this study and their corresponding series start, end, length and completeness. The preferred noise model amongst flicker noise (FN), flicker noise plus random walk (FNRW), power-law (PL) and Gauss-Markov (GM) is given for each component together with the maximum likelihood difference (dML) with respect to FN.

See separate file TableS1.txt

Figures

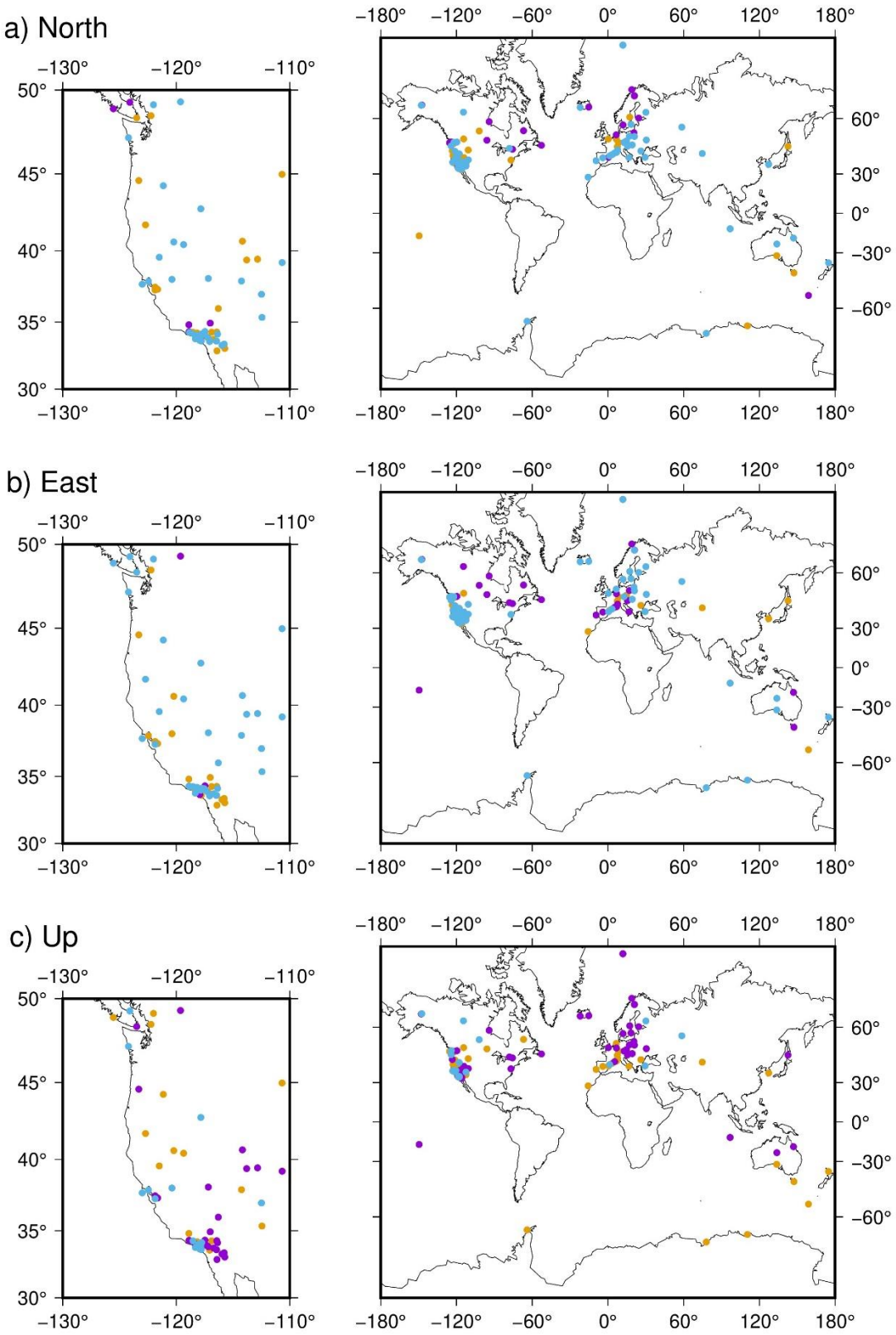

a) North
b) East
c) Up
−130°
−120°
−110°
50°
45°
40°
35°
30°
−180°
−120°
−60°
0°
60°
120°
180°
60°
30°
0°
−30°
−60°

Figure S1. Geographic distribution of the 127 JPL stations used in this study and their preferred noise model for each coordinate component. The colors of the noise models correspond to those of Fig. 4.

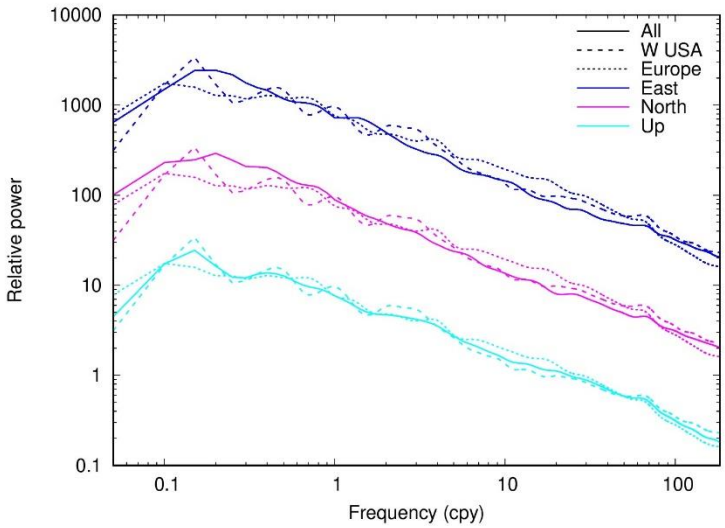


Figure S2. Smoothed stacked spectra for the three coordinate components of the 127 JPL stations used in this study (all, corresponding to Fig. 1), and for the 62 stations located in Western USA and the 36 located in Europe.

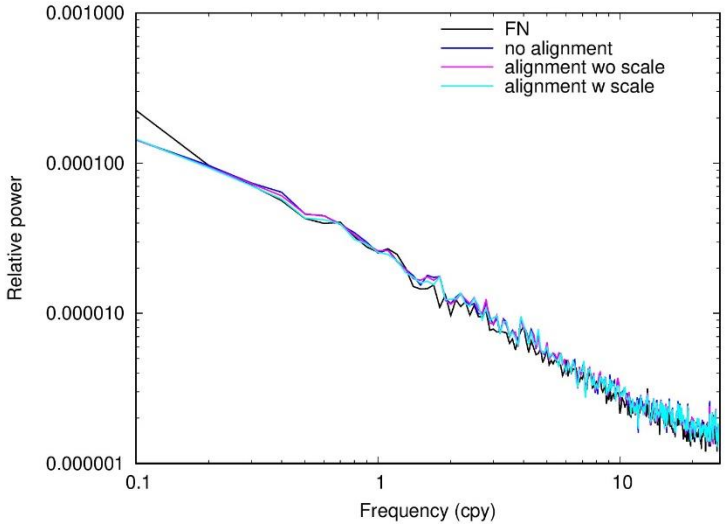


Figure S3. Stacked power spectra of the synthetic weekly series of the IGS14 core stations without alignment (blue), with alignment excluding the scale (pink) and with alignment including the scale (cyan). The black curve represents the stacked power spectra of the weekly flicker noise (FN) series that were added to the weekly SINEX files.