

Effects of linear trends on estimation of noise in GNSS position time-series

K. Dmitrieva, P. Segall and A.M. Bradley

Department of Geophysics, Stanford University, Stanford, CA 94305, USA. E-mail: dmitrieva@stanford.edu

Accepted 2016 October 13. Received 2016 September 27; in original form 2016 April 13

SUMMARY

A thorough understanding of time-dependent noise in Global Navigation Satellite System (GNSS) position time-series is necessary for computing uncertainties in any signals found in the data. However, estimation of time-correlated noise is a challenging task and is complicated by the difficulty in separating noise from signal, the features of greatest interest in the time-series. In this paper, we investigate how linear trends affect the estimation of noise in daily GNSS position time-series. We use synthetic time-series to study the relationship between linear trends and estimates of time-correlated noise for the six most commonly cited noise models. We find that the effects of added linear trends, or conversely de-trending, vary depending on the noise model. The commonly adopted model of random walk (RW), flicker noise (FN) and white noise (WN) is the most severely affected by de-trending, with estimates of low-amplitude RW most severely biased. FN plus WN is least affected by adding or removing trends. Non-integer power-law noise estimates are also less affected by de-trending, but are very sensitive to the addition of trend when the spectral index is less than one. We derive an analytical relationship between linear trends and the estimated RW variance for the special case of pure RW noise. Overall, we find that to ascertain the correct noise model for GNSS position time-series and to estimate the correct noise parameters, it is important to have independent constraints on the actual trends in the data.

Key words: Time-series analysis; Transient deformation.

1 INTRODUCTION

We are currently in the third decade of continuous Global Navigation Satellite System (GNSS) recordings of crustal motion. Daily position time-series provide highly precise estimates of GNSS velocities (Prawirodirdjo & Bock 2004; Li *et al.* 2012; Kierulf *et al.* 2014; Mantovani *et al.* 2015). However, the presence of time-correlated (or coloured) noise in the time-series complicates these estimates. First, the estimates of signals, such as linear trends, in the data can trade-off with estimates of time-correlated noise. Secondly, the presence of time-correlated noise drastically increases the velocity uncertainty (Williams 2003; Williams *et al.* 2004; Langbein 2012), yet these noise parameters can be difficult to estimate robustly (Langbein 2012; Dmitrieva *et al.* 2015).

The task of estimating noise becomes easier when the signal is known. Previously, we developed a network method of analysing noise in GNSS time-series from intraplate regions, where we can assume small or well-characterized signals (Dmitrieva *et al.* 2015). Rigid plate rotations are generally well-known *a priori*. In our previous analysis of data from the North American mid-continent, we also corrected for trends due to glacial isostatic adjustment (GIA). We found that the noise estimate for a network of stations was unchanged after the removal of modeled linear trends due to GIA. This prompted a further investigation into the effects of linear trends

on the estimates of time-correlated noise, as the trade-off between noise estimates and trends has been known for some time (Langbein & Johnson 1997).

From a scientific standpoint, our main interest is in estimating signals in the GNSS data, such as site velocities or transient signals on a variety of timescales (Melbourne & Webb 2002; Miyazaki *et al.* 2003). However, we need to quantify the time-correlated noise in the data to calculate the uncertainty in the signal. For example, estimation of a linear trend is more accurate if the noise model and the amplitudes of the various noise components are accurately known. Additionally, there is considerable debate about the type and amount of noise present in GNSS data (Hackl *et al.* 2011; Santamaria-Gómez *et al.* 2011; Amiri-Simkooei 2016; Klos *et al.* 2016), making it difficult to determine the true signal uncertainty. In order to correctly model noise in the data, we would ideally like to have strong *a priori* constraints on any signals present. In this paper, we focus on understanding the relationship between estimated time-correlated noise and linear trends in the GNSS time-series.

Time-correlated noise is usually represented by power-law (PL) forms (Agnew 1992), where noise in the spectral domain is proportional to the inverse of the frequency to a power of n —the spectral index: $p \sim f^{-n}$, where n usually ranges from -1 to 3 (Agnew 1992). Some well-known cases of the PL representation are: white noise (WN, $n = 0$), flicker noise (FN, $n = 1$) and random walk (RW,

$n = 2$). However, n could be non-integer, in which case it is referred to as generic PL.

There is no agreement on which noise model is the most representative of GNSS time-series. Some argue for a sum of FN and WN (Williams *et al.* 2004; Ray *et al.* 2008), while others suggest that the sum of RW, FN and WN should be used (Calais *et al.* 2006; King & Williams 2009; Amiri-Simkooei 2013; Dmitrieva *et al.* 2015). Finally, some suggest a sum of PL and WN (Santamaría-Gómez *et al.* 2011; Devoti *et al.* 2016; Klos *et al.* 2016). Moreover, Langbein (2008) suggests that the optimal model is different for different stations. It has also been shown that a PL model can be approximated as a sum of RW and FN (Langbein 2012). In this paper, we explore the above models with synthetic time-series, since knowledge of the true noise and trend allows us to precisely evaluate the effects of linear trends on estimation of the noise parameters. For every noise model and added trend, we perform 100 realizations and then calculate the mean and the standard deviation of the estimated noise parameters.

There are various methods to estimate noise in GNSS time-series, such as spectral estimation (Langbein & Johnson 1997; Zhang *et al.* 1997; Santamaría-Gómez *et al.* 2011), maximum likelihood estimation (MLE, Langbein & Johnson 1997; Langbein 2004; Williams *et al.* 2004), least-squares variance component estimation (Amiri-Simkooei *et al.* 2007), applying the Allan variance of the rate to the time-series (Hackl *et al.* 2011) and Kalman-filter-based MLE network noise estimation (Dmitrieva *et al.* 2015). We previously showed that when estimating time-correlated noise independently for individual stations, the time-correlated noise, especially RW, can be systematically underestimated (Dmitrieva *et al.* 2015). Estimating noise parameters for a network of stations simultaneously provides more robust estimates of the average RW variance (Dmitrieva *et al.* 2015), although in this approach we estimate only an average set of parameters for the entire network. Since in this paper all data are synthetic, there is no disadvantage to estimating noise parameters for the network rather than for individual time-series, as long as all time-series within the network have the same noise parameters. This way we gain more precision in the estimation of lowest frequency noise (such as RW or high-exponent PL). In order to estimate average noise parameters for a network, we modify the MLE method (Langbein 2004), calculating the likelihood of each time-series having the given noise covariance, and then maximizing the sum of these likelihoods, rather than maximizing each individual likelihood:

$$2 \sum_{i=1}^M \mathcal{L}(x, C) = - \sum_{i=1}^M [\ln(\det(C)) + r_i^T C^{-1} r_i + N \ln(2\pi)], \quad (1)$$

where M is the number of time-series in the network, C is the data covariance matrix, N is the number of observations and r_i are the residuals of the model fit for the i th time-series. To speed up the likelihood calculation, we use Cholesky factorization of the covariance matrix (Langbein & Johnson 1997; Bos *et al.* 2008).

In this paper, we explore the relationship between time-correlated noise estimates and linear trends in the data. First, we present a theoretical derivation of how trends affect the estimate of RW amplitude in case of a simple pure RW noise model. Then we look at how adding linear trends to various noise models affects the estimates of those noise parameters. We also explore how noise could be perceived as trend and how removing an apparent linear trend affects the noise estimates. Finally, we estimate noise and trend simultaneously and compare the results to previous tests. The main goal of

this paper is to develop an understanding of how noise estimates are affected by linear trends.

2 THEORETICAL RELATIONSHIPS BETWEEN TREND AND RANDOM WALK VARIANCE

In this section, we develop a theoretical relationship between trend and the estimated RW variance for the case of pure RW, and show how the estimate of RW scale changes with the addition or subtraction of a linear trend. This case is tractable because the first difference of an RW is simply WN.

Let z_i be an RW, where $i = 0, \dots, n$ is the epoch. If the period between two epochs Δt is constant, then $t_i = i\Delta t$ and $t_n \equiv T = n\Delta t$. A discrete RW process with variance $\tau^2 t$ is a cumulative sum of WN: $z_i = \tau\sqrt{\Delta t} \sum_{j=1}^i r_j$, where τ is the RW scale parameter with units of $\text{mm yr}^{-0.5}$ and r is a random vector with zero mean and unit variance. The difference of the series z is WN: $z_i - z_{i-1} = \tau\sqrt{\Delta t} r_i$. In the following, let $\text{Diff}(\cdot)$ denote the vector of first differences. The expectation of the mean of the differences is 0 because the difference vector $\text{Diff}(z)$ is proportional to r . Thus, the variance of the differences is

$$\text{var}[\text{Diff}(z)] = E \left[\frac{1}{n} \sum_{i=1}^n (z_i - z_{i-1})^2 \right] = \tau^2 \Delta t, \quad (2)$$

where E denotes expected value. We obtain the simple estimator

$$\hat{\tau}^2(z) \equiv \frac{1}{n\Delta t} \sum_{i=1}^n (z_i - z_{i-1})^2 = \frac{1}{T} \sum_{i=1}^n (z_i - z_{i-1})^2, \quad (3)$$

whose expectation for RW z is $E[\hat{\tau}^2(z)] = \tau^2$. Simply put, the estimate of the RW scale parameter is proportional to the variance of the difference series.

We now apply this estimator to the time-series $y_i \equiv y(t_i) = st_i + z_i$, which is a sum of RW z and linear trend st with slope s . The time-series first difference is $\text{Diff}(y) = \tau\sqrt{\Delta t} r + s\Delta t$; that is WN with non-zero mean. The expectation of the estimator with this input is thus

$$E[\hat{\tau}^2(y)] = \frac{1}{\Delta t} [\tau^2 \Delta t + s^2 \Delta t^2] = \tau^2 + s^2 \Delta t. \quad (4)$$

Eq. (4) gives the relationship between the estimated scale parameter for RW and the trend in the data. Surprisingly, the estimate depends on the sampling interval. This can be understood as follows. In the limit of very sparse sampling, it is hard to distinguish between RW and a trend. With finer sampling, RW and trend become more distinct.

In the preceding analysis, the linear trend is independent of the RW, and the expectation of the estimator $\hat{\tau}$ increases. De-trending has the opposite effect. De-trending adds a linear trend that is correlated with the RW. The expectation of $\hat{\tau}$ decreases. This can be understood by a derivation similar to that in eq. (4) for simple de-trending procedures. For example, suppose a linear trend is removed such that a time-series y_i starts and ends at 0. Then, the slope $s = -z_n/T$ (the MLE for pure RW errors), and thus $y_i = z_i - z_n t_i/T$. Proceeding as before, we construct the first difference $\text{Diff}(y)$:

$$\begin{aligned} y_i - y_{i-1} &= \tau\sqrt{\Delta t} r_i - \frac{z_n}{T} \Delta t = \tau\sqrt{\Delta t} \left[r_i - \frac{\sum_{j=1}^n r_j}{T} \Delta t \right] \\ &= \tau\sqrt{\Delta t} \left[\left(1 - \frac{\Delta t}{T}\right) r_i - \frac{\Delta t}{T} \sum_{j=1, j \neq i}^n r_j \right]. \end{aligned} \quad (5)$$

Then, the expectation of the estimator is

$$\begin{aligned}
 E[\hat{\tau}^2] &= \frac{1}{T} E \left[\sum_{i=1}^n (y_i - y_{i-1})^2 \right] \\
 &= \frac{\tau^2 \Delta t}{n \Delta t} \sum_{i=1}^n E \left[\left(1 - \frac{\Delta t}{T} \right) r_i - \frac{\Delta t}{T} \sum_{j=1, j \neq i}^n r_j \right]^2 \\
 &= \frac{\tau^2}{n} \sum_{i=1}^n \left[\left(1 - \frac{\Delta t}{T} \right)^2 E[r_i^2] + \left(\frac{\Delta t}{T} \right)^2 \sum_{j=1, j \neq i}^n E[r_j^2] \right]. \tag{6}
 \end{aligned}$$

In the second line, we used $T = n\Delta t$. In the third line, we used $E[r_i r_j] = 0$ when $i \neq j$ to remove all terms in $r_i r_j$, $i \neq j$. Now we use $E[r_i^2] = 1$ to finish:

$$\begin{aligned}
 E[\hat{\tau}^2] &= \frac{\tau^2}{n} \sum_{i=1}^n \left[\left(1 - \frac{\Delta t}{T} \right)^2 + \left(\frac{\Delta t}{T} \right)^2 (n-1) \right] \\
 &= \tau^2 \left[\left(1 - \frac{\Delta t}{T} \right)^2 + \left(\frac{\Delta t}{T} \right)^2 (n-1) \right] \\
 &= \left(1 - \frac{\Delta t}{T} \right) \tau^2, \tag{7}
 \end{aligned}$$

where we have again made use of $n\Delta t = T$ to obtain the last line. Eq. (7) should be compared with eq. (4). In eq. (4), there is a term $s^2 \Delta t$ in addition to τ^2 ; in eq. (7), there is instead a term $-\tau^2 \Delta t / T$. This makes sense in that the slope removed to obtain (7) is $s = -z_n / T$, but $z_n \sim \tau \sqrt{T}$ so that $s = -\tau / \sqrt{T}$. Again, while the bias is small for pure RW, we show that it can be considerably larger when FN and WN are present.

While it is possible to derive an analytical expression for pure RW, when any other noise component is added to the noise model it appears not to be possible to derive closed-form expressions for the expected value of the noise parameter. Instead, we explore the effects of linear trends on the estimates in the next section using synthetic data.

3 EMPIRICAL RELATIONSHIP BETWEEN TRENDS AND NOISE ESTIMATES

In this section we perform tests of more realistic noise scenarios for GNSS position time-series. We use synthetically generated time-series consisting of a sum of time-correlated and WN. Since there is no general agreement on which noise model is the most appropriate for GNSS time-series, we consider three commonly used noise models, as discussed in the Introduction. The inferred noise parameters depend on the topocentric components analysed, with horizontal components of GNSS positions being more precise than the vertical. We explore a range of noise amplitudes based on estimates reported in the literature. The first model we consider is a sum of RW, FN ($4 \text{ mm yr}^{-0.25}$) and WN (1 mm), with three RW amplitudes: 1, 0.5 and $0.1 \text{ mm yr}^{-0.5}$. Secondly, we consider a sum of PL (amplitude of $3 \text{ mm yr}^{-0.25n}$) and two different spectral indices $n = 1.4$, which lies between RW and FN, and $n = 0.3$, which lies between FN and WN) and WN (1 mm). Finally, we consider a sum of FN of $4 \text{ mm yr}^{-0.25}$ and WN of 1 mm.

First, we investigate how adding various linear trends affects the estimated time-correlated noise. For each scenario, we generate a

network of four time-series each with 10 yr of daily data, fixed noise and different linear trends varying from 0 to 1 mm yr^{-1} with an increment of 0.1 mm yr^{-1} . Then, we estimate noise parameters assuming no trend, and compare the means and standard deviations of the estimates (Fig. 1). The top panel shows the mean and standard deviation of the estimates of the RW scale parameter for the RW + FN + WN model. As expected smaller amplitudes of RW are most affected by the addition of a linear trend. When RW is high ($1 \text{ mm yr}^{-0.5}$), the mean estimate of RW amplitude exceeds the true value by one standard deviation when linear trend is 0.52 mm yr^{-1} (dashed line) and exceeds the true value by 10 per cent ($1.1 \text{ mm yr}^{-0.5}$) when the linear trend exceeds 0.63 mm yr^{-1} . For RW of $0.5 \text{ mm yr}^{-0.5}$, the mean estimate exceeds the true value by one standard deviation when the trend exceeds 0.27 mm yr^{-1} (dashed line) and exceeds the true value by 10 per cent ($0.55 \text{ mm yr}^{-0.5}$) once the linear trend is 0.34 mm yr^{-1} . In the case of low RW of $0.1 \text{ mm yr}^{-0.5}$, adding even 0.13 mm yr^{-1} of trend causes the mean to exceed the true value by one standard deviation. In summary, when the RW variance is large moderate trends do not significantly affect the RW amplitude estimate, while low-level RW can be strongly influenced by the presence of a trend in the data. Note also that the mean of the RW estimates in all three cases approaches a common value when the trends exceed $\sim 1 \text{ mm yr}^{-1}$, suggesting that for sufficiently large trend the site velocity dominates the estimated RW. We do not show the corresponding FN or WN estimates, as they are not greatly affected by the addition of linear trends for the range of parameters tested.

The second panel of Fig. 1 shows the means and standard deviations of the spectral index n for the PL + WN noise model. We consider two cases, first, the PL + WN model with high spectral index $n = 1.4$ and then with very low index $n = 0.3$. We estimate both the spectral index and the amplitude of the PL component, but only show the estimates of n , since it is significantly more affected by the added trend. Fig. 1 shows that adding a linear trend affects noise with $n = 0.3$ much more than noise with $n = 1.4$. For $n = 1.4$, the mean of the estimate exceeds the true value by one standard deviation once the linear trend is 0.37 mm yr^{-1} (dashed line) and it exceeds the true n by over 10 per cent only for trends exceeding 1 mm yr^{-1} . For $n = 0.3$, adding even 0.1 mm yr^{-1} of trend causes the mean estimated n to exceed the true value by over 50 per cent. As with the RW + FN + WN model, the estimate of the spectral index for the case when $n = 0.3$ converges to a higher value when a sufficiently large trend is added.

The bottom panel of Fig. 1 shows how the presence of a linear trend affects the estimates of FN amplitude in a FN + WN model. We find that even with a 1 mm yr^{-1} trend, the mean estimate still does not exceed 10 per cent of the true FN amplitude. The mean of the FN amplitude estimate exceeds the true value by one standard deviation when the linear trend is 0.4 mm yr^{-1} (dashed line), but in this case this results mainly from the small standard deviation in the estimate (there are fewer parameters estimated compared to previous models).

We next consider how de-trending affects the estimates of noise parameters. This is important because long-period noise could be interpreted as a trend. Using synthetic data, we calculate a mean of the absolute values of the estimated apparent trend for all six noise scenarios explored in this paper. We emphasize that for these estimates the time-series consisted only of noise with no trend. The results for 10 yr of daily positions time-series are shown in Table 1. The calculations show that a significant trend could be estimated when there is in fact no underlying linear signal. The apparent linear trend is greater for models with noise with higher spectral

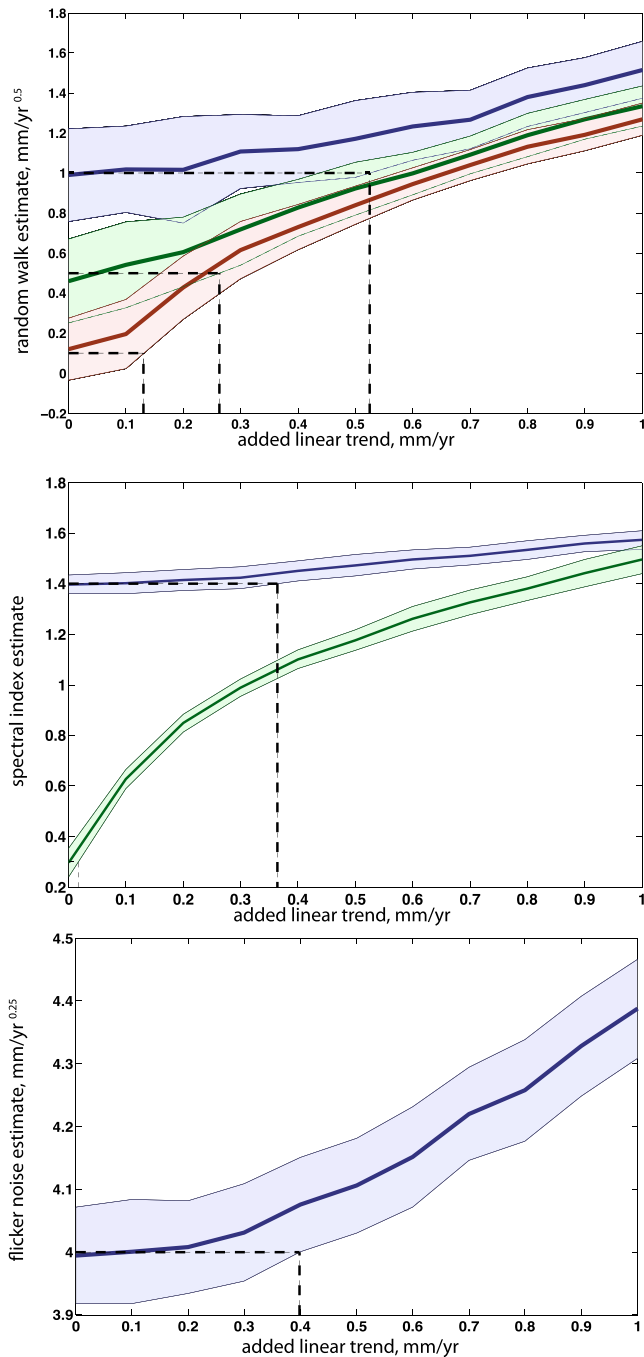


Figure 1. Changes in estimated noise parameters due to the presence of a linear trend in the data. For each combination, we create four time-series containing the same amount of noise sampled daily for 10 yr. A linear trend is added, with values ranging from 0 to 1 mm yr⁻¹, with an increment of 0.1 mm yr⁻¹. The estimated noise parameters are shown assuming that the data contain noise only. Each combination of noise model, noise amplitude and trend, is repeated 100 times. The mean (thick lines) and standard deviation (thin lines) of the estimates are shown. Top: the noise model is RW + FN + WN, the true RW is 1 mm yr^{-0.5} for the purple curves, 0.5 mm yr^{-0.5} for the green curves and 0.1 mm yr^{-0.5} for the pink curves. Middle: PL (3 mm yr^{-0.25n} amplitude and spectral index $n = 1.4$ (blue) and $n = 0.3$ (green)) + WN model. Bottom: FN (4 mm yr^{-0.25}) + WN. The dashed lines show where the mean of the estimates exceeds the true value by one standard deviation.

Table 1. Average (over 1000 estimations for each value) of absolute value of estimated linear trend in the synthetic time-series (10 yr of daily data) that consist only of noise.

Noise model	Mean trend (mm yr ⁻¹)
RW(1 mm yr ^{-0.5}) + FN(4 mm yr ^{-0.25}) + WN(1 mm)	0.30
RW(0.5 mm yr ^{-0.25}) + FN(4 mm yr ^{-0.25}) + WN(1 mm)	0.18
RW(0.1 mm yr ^{-0.25}) + FN(4 mm yr ^{-0.25}) + WN(1 mm)	0.12
PL($n=1.4$, 3 mm yr ^{-0.075}) + WN(1 mm)	0.21
PL($n=0.3$, 3 mm yr ^{-0.35}) + WN(1 mm)	0.02
FN(4 mm yr ^{-0.25}) + WN(1 mm)	0.11

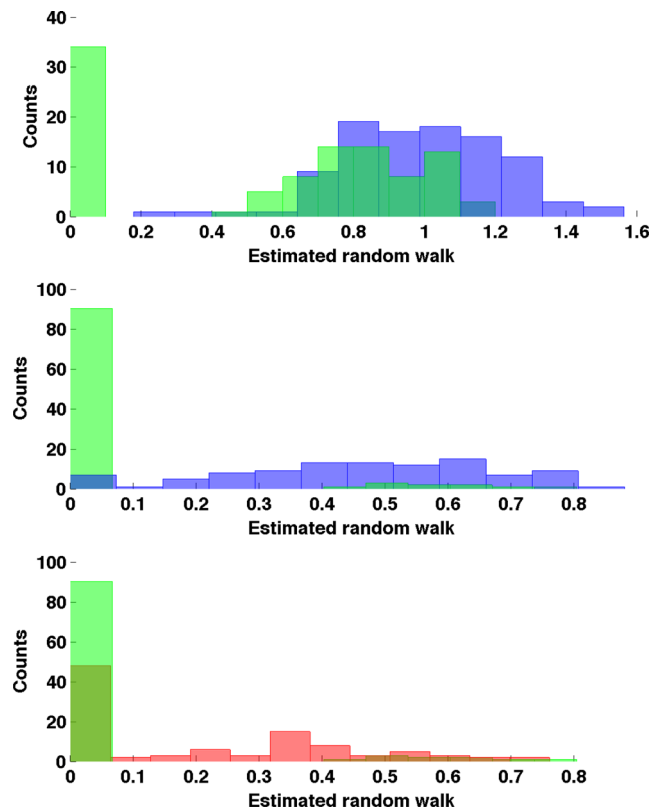


Figure 2. Effects of de-trending on noise estimates (RW + FN + WN model). Synthetic time-series contain RW + FN + WN (all panels have the same FN 4 mm yr^{-0.25} and WN 1 mm), RW from top: 1 mm yr^{-0.5}, mid and bottom: 0.5 mm yr^{-0.5}. The apparent trend is subtracted and then noise parameters are estimated. Histograms show the distribution of estimated RW amplitude for 100 trials. Blue: original, green: de-trended (intercept and slope removed) and red: de-trended (just the slope removed).

indices, such as RW and high n PL, but is still present for FN + WN. This emphasizes how time-correlated noise affects both the velocity estimate as well as the uncertainty in that estimate.

We again use synthetic tests to explore the impact of de-trending on estimates of time-correlated noise parameters. For each scenario, we generate a network of four time-series with 10 yr of daily data, fixed noise parameters, but with no linear trend. We use linear least squares with the appropriate data covariance to estimate and remove apparent trends, then estimate noise parameters from the residuals.

For the RW + FN + WN model (Fig. 2), in the case of high RW (1 mm yr^{-0.5}) removing a linear fit significantly biases the RW estimate. One-third of the tests have zero estimated RW amplitude, while the remaining two-thirds have non-zero estimates but are still

Table 2. Velocity uncertainty with the RW + FN + WN model, where FN is 4 mm yr^{-0.25}, WN is 1 mm and RW amplitude as shown in the table.

Random walk, mm yr ^{-0.5}	1.4	1.0	0.5	0.1	0
Velocity uncertainty, mm yr ⁻¹	0.5	0.35	0.21	0.13	0.13

biased to low values. For moderate RW (0.5 mm yr^{-0.5}), 90 per cent of the estimated scale parameters are zero following trend removal. The mean estimate of RW amplitude is only 0.06 mm yr^{-0.5}. Initially, we estimate and remove both the slope and the intercept of the linear trend, since this is more conventional. However, we found that removing just the slope produces a different result (Fig. 2, bottom panel). Subtracting the admittedly small intercept brings the estimate of RW down. Even when we remove just the apparent trend (without the intercept), the RW amplitude is underestimated, the mean estimate of RW is now 0.2 mm yr^{-0.5} and almost a half of the estimates are now at 0 mm yr^{-0.5} (the true value is 0.5 mm yr^{-0.5}). Thus, for the RW + FN + WN model, removing an apparent linear trend leads to a significant underestimation of the RW amplitude. This bias leads to an underestimation of velocity uncertainty (Table 2). We do not show estimates of FN and WN amplitudes as they are not strongly affected by de-trending for this noise model.

For the case of PL + WN model (Fig. 3), for both low and high spectral index, n is just slightly underestimated after removal of a fitted trend. For $n = 1.4$, the mean estimate before de-trending is 1.40, while it is 1.38 after de-trending. For $n = 0.3$ prior to de-trending, the mean estimate is 0.30 and $n = 0.28$ after the trend is removed. There is no change in the estimate of the amplitude of the PL or WN amplitudes for the PL + WN model. For FN + WN model (Fig. 4), there is almost no change in the FN estimate. Before de-trending the mean estimate is 4.00 mm yr^{-0.25} and it is 3.98 mm yr^{-0.25} after de-trending.

Finally, we test a common approach of estimating noise and linear trend simultaneously. For this, we estimate noise parameters and linear trend by maximizing the likelihood for individual time-series. We focus on WN + FN + RW, as previous tests showed RW estimates are most influenced by trends. Given that the assumed noise and trend model correctly describe the synthetic data, it is not surprising that on average the trend is correctly estimated in both high (1 mm yr^{-0.5}) and low (0.1 mm yr^{-0.5}) RW cases. For the high RW case, 30–40 per cent of RW estimates are zero, while the rest are distributed around the true value. In the low RW case, 90–95 per cent of RW estimates are zero, with the rest significantly higher than the true RW value. The RW variance estimates are generally uncorrelated with the trend estimates. FN and WN estimates are unbiased in all cases. This is consistent with the results of Dmitrieva *et al.* (2015) who showed that single time-series estimates tend to be biased to low values of RW.

4 DISCUSSION

To better understand the dependence of the noise estimate on the linear trend consider the power spectra plotted in Fig. 5, which shows theoretical noise components: RW, FN, WN, their sum and a linear trend. At high frequencies, the noise is mainly affected by WN, in the mid-frequencies FN is dominant, while RW only dominates for a limited band-width at the lowest frequencies. Fig. 5 also shows that a finite linear trend has a slope of -2 , as does RW. (Although both trend and RW have the same slope in the amplitude domain the phasing is very different, which is clear in the time domain). With realistic amounts of FN and WN, RW only dominates at the

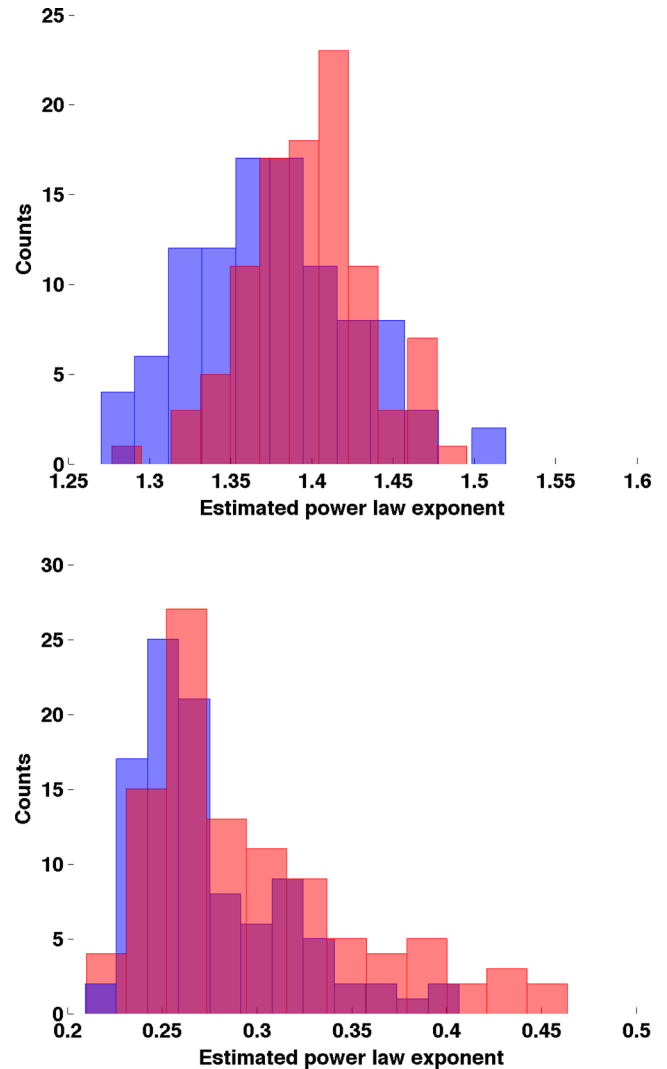


Figure 3. Effects of de-trending on noise estimates (PL + WN model). Synthetic time-series contain PL 3 mm yr^{-0.5n} with $n = 1.4$ (top) and $n = 0.3$ (bottom) and WN 1 mm. The apparent trend is subtracted and then noise parameters are estimated. Here, we show histograms of the distribution of PL spectral index estimates for 100 trials. Red: original and blue: de-trended.

lowest frequencies, making it harder to estimate and more likely to trade-off with trend.

To test this, we compared the effects of adding a linear trend to the RW + FN + WN model for typical and very low amplitude FN. Based on the spectral plots, we expect that with very low FN the added trend would lead to a smaller bias in the RW estimate. Indeed, Fig. 6 shows that with the very low FN adding 1 mm yr⁻¹ linear trend increases the RW estimate by only ~ 50 per cent, whereas with typical FN amplitude, the RW estimate is biased by ~ 260 per cent of the true value.

Perhaps a counterintuitive result is that for pure RW noise: the sampling frequency determines how much a trend biases the noise estimate (eq. 4). Perhaps some insight can be gained by considering the limiting case of two data points at the beginning and end of the time-series. In this limit, the RW time-series is indistinguishable from a linear trend. As the sampling interval decreases, the difference time-series for RW approaches a fixed distribution (in this case Gaussian WN with zero mean), whereas the difference series for the trend plus RW is WN with non-zero mean, $s\Delta t$. Thus, an ML

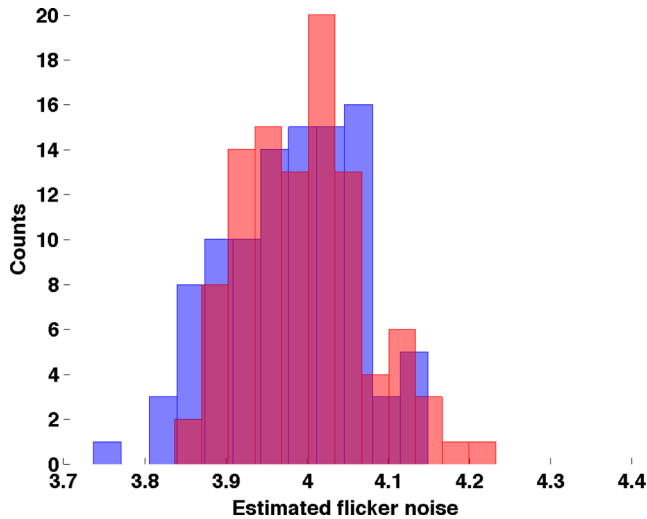


Figure 4. Effects of de-trending on noise estimates (FN + WN model). Synthetic time-series contain FN $4 \text{ mm yr}^{-0.25}$ and WN 1 mm . The apparent trend is subtracted and then noise parameters are estimated. Here, we show histograms of the distribution of FN amplitude estimates for 100 trials. Red: original and blue: de-trended.

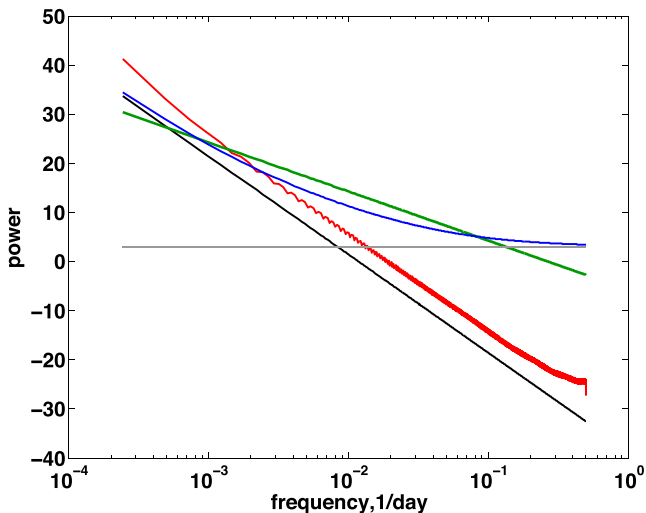


Figure 5. Power spectra showing the relationship between various noise components and a linear trend. Theoretical slope for $1 \text{ mm yr}^{-0.5}$ RW (black), $4 \text{ mm yr}^{-0.25}$ FN (green) and 1 mm WN (grey). The blue line is a sum of RW, FN and WN. The red line is a power spectrum of 10 yr of daily data with slope of 3.8 mm yr^{-1} . We zero-padded the linear trend time-series to avoid artefacts due to the finiteness of the slope.

estimator can better differentiate trend from RW with more data points, even though their amplitude spectra are similar.

Our simulations show that time-correlated noise can be incorrectly perceived as a linear trend. Removing this apparent, but non-existent, trend may bias the noise estimate to low values; RW noise is especially sensitive to de-trending. For the RW + FN + WN model, de-trending can significantly decrease estimates of RW amplitude. Removing a linear trend also decreases the estimate of the spectral index for PL + WN model, but to a lesser extent. Removal of an apparent trend does not affect the estimate of FN for the FN + WN model, in part because of the difference in spectral slope between the trend (-2) and FN (-1). In contrast to the PL model, the spectral index is fixed when estimating noise parameters for the FN + WN model.

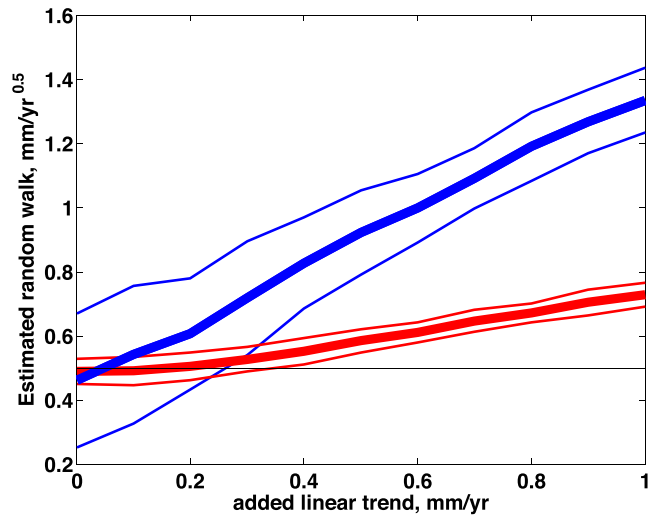


Figure 6. Comparison of the effects of added linear trend on RW estimates, for RW + FN + WN with typical FN ($4 \text{ mm yr}^{-0.25}$, in blue) and very low ($0.1 \text{ mm yr}^{-0.25}$, in red). For both cases, RW is $0.5 \text{ mm yr}^{-0.5}$ and WN is 1 mm . Thick lines indicate the mean of 100 estimates and thinner lines are one standard deviation. Black line shows true RW.

For the RW + FN + WN model, de-trending can result in very low and even null estimates of RW, even when the true RW amplitude is significant. In the first two tests, here we maximized the sum of the likelihoods from multiple time-series. We have previously shown that network approaches are more precise at estimating low levels of RW (Dmitrieva *et al.* 2015). When maximizing the likelihood for each time-series, as typically done, de-trending is even more likely to result in null estimates of RW. We conclude that de-trending the data can lead to biased or even vanishing RW estimates. At the same time, accurately estimating weak RW is difficult in the presence of trends. The estimated RW amplitudes can be significantly larger than the true values, when unaccounted for trends are present. Hence, one has to be very careful about de-trending the time-series, since this could lead to either completely neglecting or significantly overestimating the RW variance.

The FN + WN model is insensitive to both moderate linear trends as well as to de-trending. With strong *a priori* knowledge that FN + WN is the correct noise model, one could be somewhat liberal with removing trends. The same holds for the PL + WN model, as we observe that de-trending only weakly biases estimates of the spectral index. For PL indices that lie between RW and FN, such as $n = 1.4$, adding even moderate trends does not have a significant effect on the estimate of the spectral index. However, for PL noise with lower spectral index, such as $n = 0.3$, even a very small (e.g. 0.05 mm yr^{-1}) trend has a significant effect on the estimate of the spectral index. Removing a linear trend for low index PL noise does not affect the estimate of the index. Hence, if one is confident in the PL + WN model and MLEs of spectral index are low, it is fair to assume that those estimates are accurate. However, higher spectral index estimates could be due to (1) the index actually being high, (2) a residual linear trend or (3) PL + WN not being the correct model to use.

We do emphasize, however, that there is independent evidence that GNSS monument motion contributes RW to geodetic time-series (Wyatt 1989; Johnson & Agnew 1995). We have shown that de-trending prior to noise estimation or estimation noise for individual time-series can lead to null estimates of RW. This can potentially lead to the erroneous conclusion that a simpler FN + WN model,

that requires only two parameters to estimate. The best solution to this problem may be to analyse data from areas where there are *a priori* constraints on trends in the data, as discussed below. In order to ensure that RW is not missed, we recommend using an approach where noise parameters are estimated simultaneously for a group of stations. Analysing time-series in these areas can allow us to determine average noise models to describe GNSS time-series. Such areas could be the interiors of plates far from plate-boundary deformation, and also far from large GIA signals, or where such effects are well modeled. As noted in the Introduction, preliminary work in the North American mid-continent found that removing GIA velocities barely influenced estimates of noise parameters (Dmitrieva *et al.* 2015). In that case, the average GIA signals for the horizontal time-series were low, with a mean of 0.28 mm yr^{-1} , and the estimated RW was relatively high, $1 \text{ mm yr}^{-0.5}$.

One could reasonably argue that estimating noise parameters simultaneously with the linear trends is the correct approach. We tested that treating each time-series individually. The results show that trend estimates are close to the true values, however RW is often underestimated, especially with low-amplitude RW. Dmitrieva *et al.* (2015) adopt the philosophy that local, non-tectonic trends in data should be treated as noise for the purposes of tectonic studies. In these studies, one is primarily interested in deformations that are spatially coherent over length scales appropriate to the processes under consideration. For GIA, the length scales are much longer than the lithospheric thickness, for tectonic studies the appropriate length scales are typically longer than the crustal thickness. Local trends possibly resulting from geomorphic processes (e.g. slumping), localized fluid withdrawal or injection, can contaminate these long-wavelength signals, especially if the GNSS stations were not installed for such applications. Simultaneously estimating a noise model and trend at each site independently may indeed provide a better measure of the intrinsic accuracy of the GNSS system, however a more conservative approach, useful for modeling studies is to treat trends of unknown origin as noise. Thus, we advocate removing only trends due to known processes such as plate motion for sites well removed from plate boundaries, and GIA motions, when they are not the focus of study. Of course there is some uncertainty in our knowledge of these trends that should be propagated through to the noise estimates.

5 CONCLUSIONS

When a small-to-moderate linear trend is added to a pure RW time-series with daily sampling, the effects of the trend on the estimates of RW variance are small. However, for more realistic noise models, the results vary significantly depending on the noise model, and whether the trend is estimated simultaneously with the noise parameters. In the presence of WN and FN, RW is both very sensitive to prior detrending as well as unmodeled residual trends. It can be difficult to either confirm or reject the presence of (especially weak) RW in the data without *a priori* constraints on the signal trend. Additionally, when estimated for individual time-series RW estimates are often biased to zero.

Both FN + WN and PL + WN (with the PL spectral index $1 < n < 2$) models are relatively insensitive to detrending. However, for the PL + WN model with lower spectral indices $0 < n < 1$, an added trend drastically increases the estimate of the spectral index.

In order to know the uncertainties of the estimated linear trends, we need to know the time-correlated noise model and variances. At the same time, estimates of time-correlated noise depend on

knowledge of any linear trends present in the data. To determine the best noise model and conservative variances in the GNSS position time-series for tectonic studies, we recommend focusing on areas where trends are either very small or well-known *a priori*, and only removing these known signals. This approach may lead to conservative estimates of the GNSS system noise, but at least partially accounts for the possible presence of non-tectonic signals in actual time-series. In addition, we suggest that analysing data from groups of stations will provide more robust estimates of RW variance.

ACKNOWLEDGEMENTS

This work was supported by NASA under the NASA Earth and Space Science Fellowship Programme (grant no. 14-EARTH14R-47) and SCEC (grant no. 13057).

REFERENCES

- Agnew, D.C., 1992. The time-domain behavior of power-law noises, *Geophys. Res. Lett.*, **19**(4), 333–336.
- Amiri-Simkooei, A.R., 2013. On the nature of GPS draconitic year periodic pattern in multivariate position time series, *J. geophys. Res.*, **118**(5), 2500–2511.
- Amiri-Simkooei, A.R., 2016. Non-negative least-squares variance component estimation with application to GPS time series, *J. Geod.*, **89**(6), 591–606.
- Amiri-Simkooei, A.R., Tiberius, C.C.J.M. & Teunissen, S.P., 2007. Assessment of noise in GPS coordinate time series: methodology and results, *J. geophys. Res.*, **112**, B07413, doi:10.1029/2006JB004913.
- Bos, M.S., Fernandes, R.M.S., Williams, S.D.P. & Bastos, L., 2008. Fast error analysis of continuous GPS observations, *J. Geod.*, **82**(3), 157–166.
- Calais, E., Han, J.Y., DeMets, C. & Nocquet, J.M., 2006. Deformation of the North American plate interior from a decade of continuous GPS measurements, *J. geophys. Res.*, **111**, B06402, doi:10.1029/2005JB004253.
- Devoti, R., Pietrantonio, G., Pisani, A.R. & Riguzzi, F., 2016. Permanent GPS networks in Italy: analysis of time series noise, in *VIII Hotine-Marussi Symposium on Mathematical Geodesy*, Volume 142 of the series *International Association of Geodesy Symposia*, eds Sneeuw, N. et al., pp. 51–58, Springer International Publishing.
- Dmitrieva, K., Segall, P. & DeMets, C., 2015. Network-based estimation of time-dependent noise in GPS position time series, *J. Geod.*, **89**(6), 591–606.
- Johnson, H.O. & Agnew, D.C., 1995. Monument motion and measurements of crustal velocities, *J. geophys. Res.*, **22**(21), 2905–2908.
- Hackl, M., Malservisi, R., Hugentobler, U. & Wonnacott, R., 2011. Estimation of velocity uncertainties from GPS time series: examples from the analysis of the South African TrigNet network, *J. geophys. Res.*, **116**, B11404, doi:10.1029/2010JB008142.
- Kierulf, H.P., Steffen, H., Simpson, M.J.R., Lidberg, M., Wu, P. & Wang, H., 2014. A GPS velocity field for Fennoscandia and a consistent comparison to glacial isostatic adjustment models, *J. geophys. Res.*, **119**(8), 6613–6629.
- King, M.A. & Williams, S.D.P., 2009. Apparent stability of GPS monumentation from short-baseline time series, *J. geophys. Res.*, **114**, B10403, doi:10.1029/2009JB006319.
- Klos, A., Bogusz, J., Figurski, M. & Gruszczynski, M., 2016. Error analysis for European IGS stations, *Stud. Geophys. Geod.*, **60**(1), 17–34.
- Langbein, J., 2004. Noise in two-color electronic distance meter measurements revisited, *J. geophys. Res.*, **109**, B04406, doi:10.1029/2003JB002819.
- Langbein, J., 2008. Noise in GPS displacement measurements from Southern California and Southern Nevada, *J. geophys. Res.*, **113**, B05405, doi:10.1029/2007JB005247.
- Langbein, J., 2012. Estimating rate uncertainty with maximum likelihood: differences between power-law and flicker/Drandom-walk models, *J. geophys. Res.*, **86**(9), 775–783.

- Langbein, J. & Johnson, H., 1997. Correlated errors in geodetic time series: implications for time-dependent deformation, *J. Geod.*, **102**(B1), 591–603.
- Li, Q., You, X., Yang, S., Du, R., Qiao, X., Zou, R. & Wang, Q., 2012. A precise velocity field of tectonic deformation in China as inferred from intensive GPS observations, *Sci. China Earth Sci.*, **55**(5), 695–698.
- Mantovani, E., Viti, M., Cenni, N., Babbucci, D. & Tamburelli, C., 2015. Present velocity field in the Italian region by GPS data: geodynamic/tectonic implications, *Int. J. Geosci.*, **6**(12), 1285–1316.
- Melbourne, T.I. & Webb, F.H., 2002. Precursory transient slip during the 2001 $M_w = 8.4$ Peru earthquake sequence from continuous GPS, *J. geophys. Res.*, **29**(21), 2032, doi:10.1029/2002GL015533.
- Miyazaki, S.I., McGuire, J.J. & Segall, P., 2003. A transient subduction zone slip episode in southwest Japan observed by the nationwide GPS, *J. geophys. Res.* **108**(B2), 2087, doi:10.1029/2001JB000456.
- Prawirodirdjo, L. & Bock, Y., 2004. Instantaneous global plate motion model from 12 years of continuous GPS observations, *J. geophys. Res.*, **109**, B08405, doi:10.1029/2003JB002944.
- Ray, J., Altamimi, Z., Collilieux, X. & van Dam, T., 2008. Anomalous harmonics in the spectra of GPS position estimates, *GPS Solut.*, **12**(1), 55–64.
- Santamaría-Gómez, A., Bouin, M.N., Collilieux, X. & Wöppelmann, G., 2011. Correlated errors in GPS position time series: implications for velocity estimates, *J. geophys. Res.*, **116**, B01405, doi:10.1029/2010JB007701.
- Williams, S.D.P., 2003. The effect of coloured noise on the uncertainties of rates estimated from geodetic time series, *J. Geod.*, **76**(9–10), 483–494.
- Williams, S.D.P., Bock, Y., Fang, P., Jamason, P., Nikolaidis, R.M., Prawirodirdjo, L., Miller, M. & Johnson, D.J., 2004. Error analysis of continuous GPS position time series, *J. geophys. Res.*, **109**, B03412, doi:10.1029/2003JB002741.
- Wyatt, F., 1989. Displacement of surface monuments—vertical motion, *J. geophys. Res.*, **94**, 1655–1664.
- Zhang, J., Bock, Y., Johnson, H., Fang, P., Williams, S., Genrich, J., Wdowinski, S. & Behr, J., 1997. Southern California permanent GPS geodetic array: error analysis of daily position estimates and site velocities, *J. geophys. Res.*, **109**(B3), 18 035–18 055.