

Article

Comparison of Enhanced Noise Model Performance Based on Analysis of Civilian GPS Data

Andy W. R. Soundy, Bradley J. Panckhurst, Phillip Brown, Andrew Martin, Timothy C. A. Molteno  and Daniel Schumayer * 

Department of Physics, University of Otago, 730 Cumberland St, Dunedin 9016, New Zealand; souan326@student.otago.ac.nz (A.W.R.S.); panbr907@student.otago.ac.nz (B.J.P.); phillb@elec.ac.nz (P.B.); andrew.martin@otago.ac.nz (A.M.); tim@physics.otago.ac.nz (T.C.A.M.)

* Correspondence: daniel.schumayer@otago.ac.nz; Tel.: +64-479-7815

Received: 1 August 2020; Accepted: 21 October 2020; Published: 24 October 2020



Abstract: We recorded the time series of location data from stationary, single-frequency (L1) GPS positioning systems at a variety of geographic locations. The empirical autocorrelation function of these data shows significant temporal correlations. The Gaussian white noise model, widely used in sensor-fusion algorithms, does not account for the observed autocorrelations and has an artificially large variance. Noise-model analysis—using Akaike’s Information Criterion—favours alternative models, such as an Ornstein–Uhlenbeck or an autoregressive process. We suggest that incorporating a suitable enhanced noise model into applications (e.g., Kalman Filters) that rely on GPS position estimates will improve performance. This provides an alternative to explicitly modelling possible sources of correlation (e.g., multipath, shadowing, or other second-order physical phenomena).

Dataset License: BY-NC-ND

Keywords: noise models; stochastic process; GPS error; Ornstein–Uhlenbeck process; sensor fusion; embedded computing

1. Introduction

Numerous applications are built using location provided by Global Positioning System (GPS) receivers. While centimetre precision has been demonstrated for dual-frequency [1] and differential GPS systems, a very large number of systems rely on single-frequency non-differential units, which report positions with typical precision of many metres [2]. We use the term “noise” to refer to the errors in a sequence of reported positions. The sources of the noise in the positions include atmospheric and ionospheric phenomena [3], multipath effects [4,5], canopy trees [6], and a wide variety of physical phenomena that may be impractical to model. However, other types of errors are also inadvertently added to the signal, e.g., errors present in the modelling of satellite orbits [7], finite precision of the digital components of receivers, variations in the built-in phased-locked loop [8], or the signal processing techniques [9], etc. The common structure and frequent formulations of GPS algorithms are discussed in Xu and Xu [10], including some aspects of the mathematical formulation that carry non-trivial correlation in readings between epochs. In summary, however, it is sufficient for this work to realise that there are noise sources that can be adequately modelled by uncorrelated processes, e.g., white noise, and there are sources that inherently carry temporal correlations.

All GPS signals are encumbered with noise, whose sources are hard to distinguish [11,12]. The noise could be attributed to satellite clock and orbit errors, to the propagation of the GPS signal through media (ionosphere, troposphere, etc.), and to the receiver units. Many of these noise components are time-correlated and caused by stochastic processes external to the useful signal.

It would be ambitious to model and distinguish between all these processes; alternatively, one can consider their cumulative effect on the time series [13,14]. Irrespective of the source, a robust signal-processing algorithm needs some model of noise processes. Such a model unavoidably contains assumptions on the nature of the noise, e.g., amplitude and correlation in the time domain, which must be experimentally validated. Often, the measurement errors are modelled by independent, identically distributed random Gaussian variables (Gaussian white noise). While this model is sufficient for many applications, it is generally accepted that a linear combination of coloured noise (whose power spectrum follows a $P \sim f^{-\kappa}$ law with a characteristic exponent κ) is a more suitable approximation [15]. With the assumption of white noise, the time correlation in GPS datastreams leads to biased and inconsistent parameter estimates [16], especially for sampling rates above 5 Hz. Moreover, it is unrealistic that measurements from different satellites would have the same accuracy; therefore, stochastic models assuming independent and constant variance result in unreliable statistics, and stochastic modelling is needed. Directly estimating statistical parameters of the GPS time series by explicitly taking into account the heteroscedastic space- and time-correlated error structure of the observations could result in improved accuracy at minimal computational cost.

Noise in long GPS time series has been analysed for geodetic purposes by [17,18], who found $-1 \leq \kappa \leq 3$ in the relevant frequency band. Their results are, however, not conclusive, as many spectra would fit this range, e.g., white noise ($\kappa = 0$) or flicker noise ($\kappa = 1$). Many processes that these GPS observations are expected to capture, e.g., earthquakes or the wobble of the Earth around its axis, have similar spectra [19]. Consequently, the physical process being observed by GPS and the inherent noise of a GPS signal are present simultaneously. The noise amplitude changes with time and, for older data, can be a considerable fraction of the signal [20].

In this study, we analyse time series of GPS measurements with duration between a day and a week at 1 Hz sampling frequency. Observations were made in the Southern Hemisphere with stationary units. The recordings show significant autocorrelation incompatible with a Gaussian white noise model. The observed presence of correlations implies that there are some stochastic equations governing the dynamics of the signal. There are two major sources that plausibly introduce correlations: (a) the movement of GPS satellites and, thus, their constellation, and (b) the algorithms coded in the GPS emitters and receivers. Both sources are, to a high degree, deterministic; thus, in a perfect environment, they would produce a strongly correlated signal. However, there are many other phenomena, e.g., weather and surface reflectance, that introduce effects into the signal that can be most effectively modelled by noise, rather than searching for their respective governing equations. We develop mathematical models for the observed noise and suggest how these models can be used for improved positioning and uncertainty quantification in future devices.

2. Description of the Data Collection Process

Data were collected from civilian L1-band GPS units at multiple locations. At each location, data were collected from four units with a Telit Jupiter-F2 integrated GPS module, as well as a single standard USB-connected GPS receiver—GlobalSat BU-353S4. The first- and second-order statistics of the location data showed no significant difference between the two devices. For brevity, we present our findings using the data obtained only with the GlobalSat BU-353S4.

The location of satellites is a major factor of GPS accuracy degradation. Devices function well when the satellites are widely distributed [21]; hence, the altitude errors in measurements are typically much larger than those in latitude or longitude. The GPS satellites operate at an inclination of around 55° ; consequently, for devices at higher latitudes, the GPS satellite positions are biased strongly towards the equator.

To ensure that our analysis was not biased due to specifics of the GPS satellite constellation geometry, we carried out measurements at four locations with latitudes between 45° and 10° South (see Table 1). Data were recorded at each site for approximately 48 h. The decimated latitude and

longitude errors (deviations from time average) are shown in Figure 1. Disregarding the initialisation period (typically ~ 1 min), the data resemble a random walk.

Table 1. Measurement sites and dates of observations.

Measurement Site	Latitude (South)	Longitude (East)	Dates
North Stradbroke Island	27°25'14''	153°30'52''	14 September 2014
Port Douglas	16°29'07''	145°27'59''	17 September 2014
Thursday Island	10°34'79''	142°13'47''	21 September 2014
Dunedin	45° 51'50''	170°30'48''	6 August 2014 7 August 2014 15 August 2014 27 September 2014 5 October 2014 22 October 2014

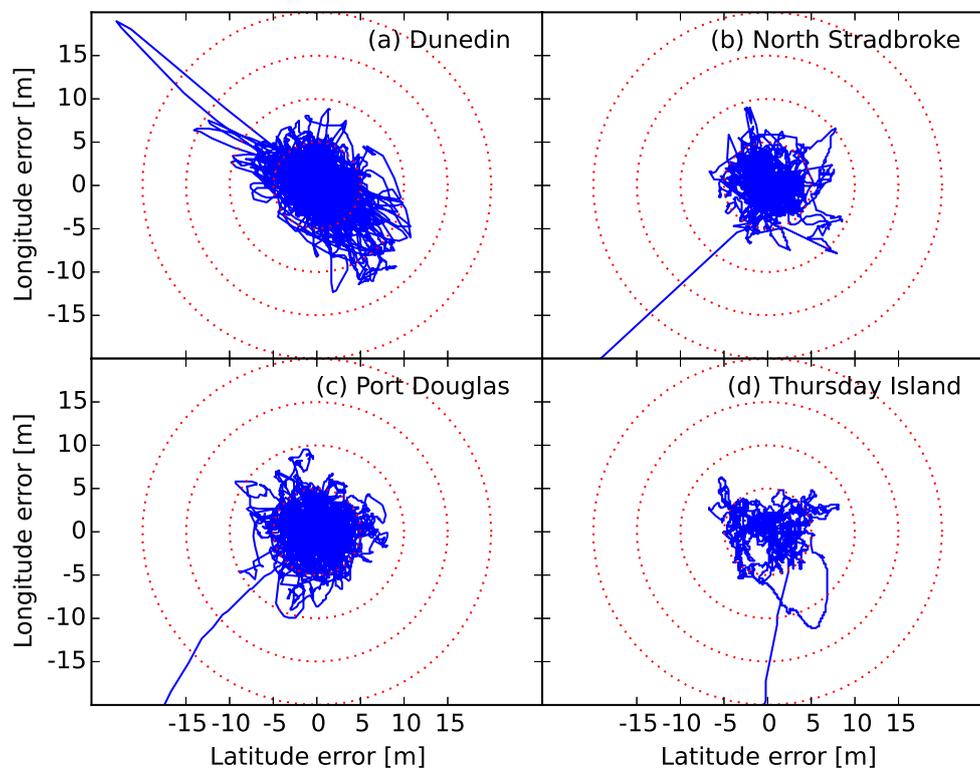


Figure 1. The longitude and latitude errors at each location. Due to the large number of observations, the data have been decimated.

While this cursory look seems suggestive, calculating the residuals defined as $r_t = X_t - \bar{\zeta}_0$ (here, $\bar{\zeta}_0$ denotes the long-time average, or the ‘true’ position) and creating a QQ-plot [22] reveals the non-Gaussian nature of r_t (for technical details see Appendix A).

In Figure 2 the QQ-plot of the latitude data collected in Dunedin is depicted. In order to support the graphical findings, we have also calculated two more normality tests [23], one based on d’Agostino and Pearson’s work [24,25], and another one suggested by Lilliefors [26]. The latter is the most suitable for our use, as it does not assume the knowledge of mean and standard deviation of the normal distribution, but estimates these from the sample. The null hypothesis of both tests is that the data come from a normal distribution. Since both tests return p -values much smaller than 0.01, we can thus conclude at a 99% confidence level that latitude data do not support the assumption of normal distribution. Similar conclusions can be drawn for longitude and altitude data at all locations.

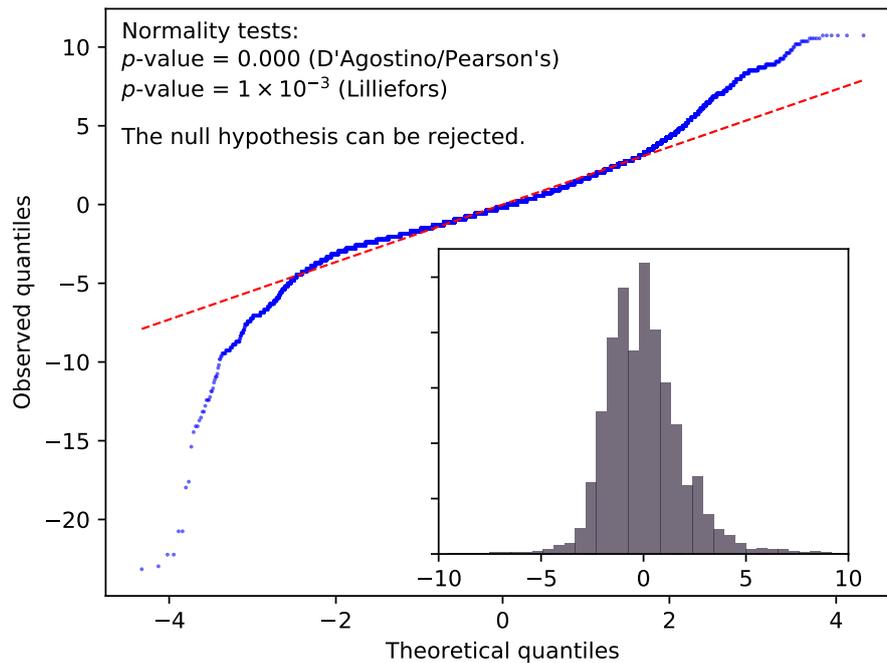


Figure 2. The main figure shows the QQ-plot of the residual of latitude data (measured in metres) collected in Dunedin. The inset depicts the histogram of the same data. The p -values of two normality tests are also given in the top left corner. Based on this evidence, the data do not support the assumption that the residual follows a Gaussian distribution.

3. Analysis of Observations

One may think of the collection of positions—emitted by the GPS chip—as a stochastic process, a sequence of random variables, $X_1, X_2, \dots, X_t, \dots$ labelled by time $t = 1, 2, \dots$. In our study, X_t may represent the latitude, longitude, or altitude returned by a GPS unit. We denote random variables with capital letters, e.g., X_t , and its particular realisation with lower case letter, e.g., x_t . Any estimator of a statistic, e.g., the expectation value μ , is denoted by the same letter but decorated with a $\hat{\cdot}$ symbol, e.g., $\hat{\mu}$.

Stochastic Processes

The simplest non-trivial stochastic process can be formulated as the sum of the true position, ζ_0 , and an additive noise term:

$$X_t = \zeta_0 + E_t. \quad (1)$$

The measurement error, E_t , is often considered to be a Gaussian white noise with zero mean and variance σ^2 . The problems of fitting Gaussian noise models to GPS measurements have been recognised by [27,28], and two models were proposed to describe the measurements more accurately: a moving average process (MA) and an autoregressive model (AR).

We check whether adjacent recordings can be considered independent random variables using the empirical autocorrelation function [29–33], defined as

$$\hat{\rho}(t, s) = \frac{\text{cov}(X_t, X_s)}{\sigma_t \sigma_s} = \frac{\mathbb{E}[(X_t - \mu_t)(X_s - \mu_s)]}{\sigma_t \sigma_s}, \quad (2)$$

where μ_r and σ_r are the expectation value and variance of the random process at step r , respectively. The process $\{X_t\}$ is stationary if μ_t and σ_t are both time-independent; in such a case, the autocorrelation function depends only on the time difference $h = s - t$ (lag) between its arguments and not explicitly on the arguments themselves. Henceforth, we assume stationarity; thus, $\rho(t, s) = \rho(h)$ while $\mu_t = \mu$

and $\sigma_t = \sigma$ both become constants, which are calculated by time-averaging. Furthermore, we consider the empirical partial autocorrelation function calculated using the Yule–Walker method [32,34] and defined as

$$\hat{\rho}_p(h) = \frac{\text{cov}(X_t, X_{t+h} | X_{t+1}, X_{t+2}, \dots, X_{t+h-1})}{\sigma_t \sigma_{t+h}}, \quad (3)$$

which provides the amount of correlation between a variable and a lag of itself that is not explained by correlations at all lower-order lags.

As an example, Figure 3 shows the autocorrelation of a typical time series. The autocorrelation function decays slowly. Fitting an exponential decay to the first half-hour dataset provides a characteristic time estimate of $\tau = 20.98 \pm 0.02$ minutes, i.e., it takes approximately an hour ($\sim 3\tau$) to reach statistical independence. Using the reported position for periods much shorter than τ would lead to a biased estimate. We wish to draw the reader’s attention to the fact that long-term correlations usually imply a power-law decay of correlation, while our fit assumes an exponential correlation. While exponential decay of correlation is relatively fast, compared to power-law decay, it seems to have an observable effect on location time series, and its inclusion in modelling may provide better positioning algorithms in the future.

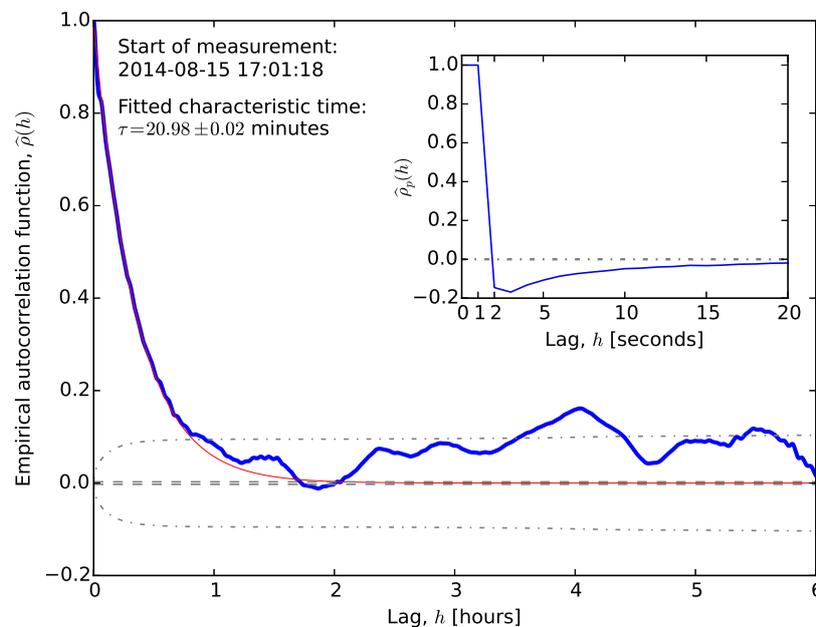


Figure 3. Correlogram of latitude errors of a time series from Dunedin of ~ 250 h. The dashed lines correspond to the classical 95% confidence interval, while the dash-dotted curve corresponds to the Box–Jenkins standard error. The thin solid line represents the exponentially decaying curve fitted to the first half-hour section. The inset depicts the empirical partial autocorrelation function.

The theoretical autocorrelation function of Gaussian white noise is the Dirac-delta, $\rho(t, s) \propto \delta(t - s)$. Assuming this model, one may test for statistical independence of X_t and X_{t+h} by checking whether the autocorrelation function is different from zero, i.e., whether it falls within the confidence interval around $\hat{\mu} = 0$ [35,36].

$$-\frac{z\left(1 - \frac{\alpha}{2}\right)}{\sqrt{n}} \hat{\sigma} \leq \hat{\rho}(h) \leq \frac{z\left(1 - \frac{\alpha}{2}\right)}{\sqrt{n}} \hat{\sigma},$$

where $z(x)$ is the cumulative distribution function of the standard normal probability distribution, α is the required level of significance, $\hat{\sigma}$ denotes the sample standard deviation, and n is the number of measurements. This confidence interval at the $\alpha = 0.05$ level is plotted with dashed lines in Figure 3.

The majority of $\hat{\rho}(h)$ falls outside this interval; therefore, our observations and the assumption of Gaussian white noise are inconsistent with each other. One may take into account that the empirical autocorrelation function estimated from a finite-length time series also carries error [37] and could derive a modified confidence interval. At significance level α , this is given by the Box–Jenkins standard error [38–40]:

$$-\frac{z(1-\frac{\alpha}{2})}{\sqrt{n}} S(h) \leq \hat{\rho}(h) \leq \frac{z(1-\frac{\alpha}{2})}{\sqrt{n}} S(h),$$

where $S(h) = \sqrt{1 + 2 \sum_{k=1}^{h-1} \hat{\rho}^2(k)}$. Figure 3 also depicts the Box–Jenkins standard error with dash-dotted lines, and it is apparent that the two confidence intervals differ from each other substantially, thus confirming the statistical dependence in the measured location readings. It is worth noting here that the exponential fit and the first crossing of $\hat{\rho}(h)$ with the Box–Jenkins confidence interval agree with each other. The inset in Figure 3 shows $\hat{\rho}_p(h)$ for $h \leq 20$ seconds with its standard error at a 0.95 significance level. This graph may suggest the application of a high-order autoregressive noise model (see next section), since $\hat{\rho}_p(h)$ vanishes only for lags larger than approximately 30 s.

However, we find that other datasets exhibit much shorter ‘memory’, so lower-order processes might also be sufficient.

Datasets from other locations exhibited different behaviour (see Figure 4). The autocorrelation still takes less than an hour to reach the Box–Jenkins standard error, but the initial drop is much sharper. This drop may explain the different behaviour of the partial autocorrelation function, which vanishes for lags larger than unity. Figures 3 and 4 illustrate the extreme cases.

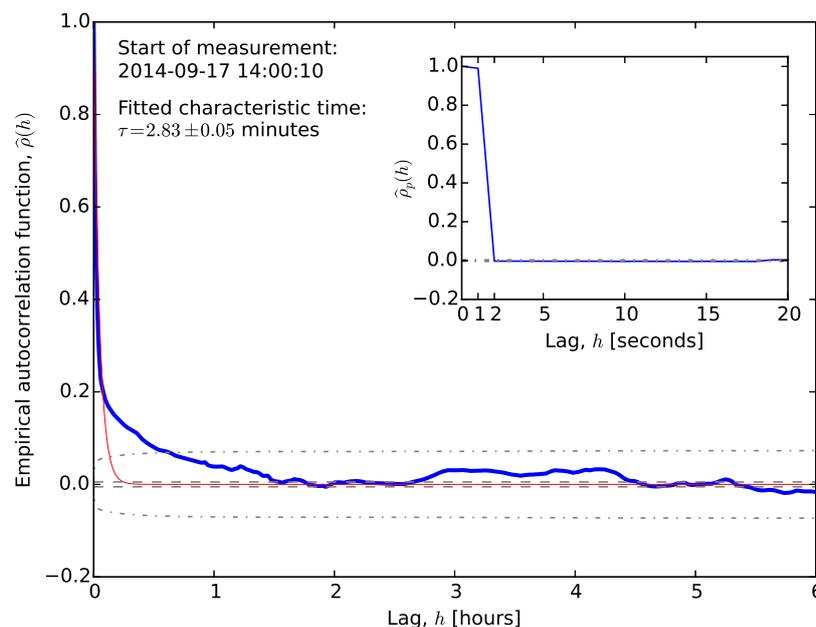


Figure 4. Correlogram of longitude errors for the GPS data taken in Port Douglas. The line styles are identical to those in Figure 3.

The autocorrelation of the Thursday Island dataset, see Figure 5, displays persistent oscillations not observed at other locations. This could be because of a temporary loss of power experienced on Thursday Island, which forced the data to be split into two sets, each approximately 24 h long, or an artefact of the local geography contributing to long-term multipath effects.

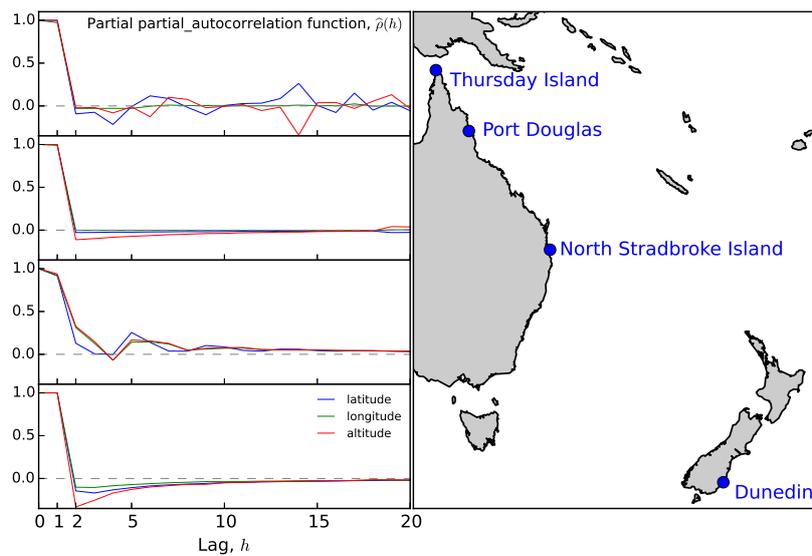


Figure 5. The partial autocorrelation function, $\hat{\rho}_p(h)$, for longitude, latitude, and altitude at all geographic locations visited. The lag is measured in seconds.

4. Modelling the Noise

In this section, we develop some statistical models for the observed noise. A model is a relationship between random variables, $\{X_{t+k}\}_{k \in \mathbb{K}}$, in the following form:

$$h(\{X_{t+k}\}_{k \in \mathbb{K}}) = g(\{E_{t+\ell}\}_{\ell \in \mathbb{L}}), \quad (4)$$

where \mathbb{K} and \mathbb{L} are sets of integers, h and g are fixed functions, and $E_t \sim \mathcal{N}(0, \sigma^2)$. Here, the quartet $(h, g, \mathbb{K}, \mathbb{L})$ constitutes the model to be inferred. To simplify the task, we presume that h and g are linear in all their variables (so-called linear models). The stochastic model is then

$$\sum_{k=0}^{\infty} h_k X_{t-k} = \sum_{\ell=0}^{\infty} g_\ell E_{t-\ell}, \quad (5)$$

where h_k and g_ℓ are all constants with $h_0 = g_0 = 1$ (convenience), and causality restricts summations to the current or past values of X_t or E_t [41]. This model requires an infinite number of coefficients, which cannot be estimated using a finite number of observations. Therefore, we focus on finite-parameter linear models and estimate the unknowns via the maximum likelihood method, in which the likelihood function below is constructed for Gaussian processes.

The likelihood function, L , is the joint probability density of obtaining the actually observed data $\mathbf{x} = (x_1, x_2, \dots, x_N)$, according to the model. It is a function of the unknown parameters $\boldsymbol{\beta} = (h_0, h_1, \dots, g_0, g_1, \dots)$, and σ^2 . Under general assumptions, the log-likelihood function of a finite parameter linear model is:

$$\log(L(\boldsymbol{\beta}, \sigma^2)) = -\frac{1}{2} \left[N \log(2\pi) + \log(\det(\sigma^2 \boldsymbol{\Sigma}(\boldsymbol{\beta}))) + \frac{1}{\sigma^2} \mathbf{x}^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}) \mathbf{x} \right]$$

where the (m, n) -th component of the covariance matrix $\sigma^2 \boldsymbol{\Sigma}(\boldsymbol{\beta})$ is calculated from

$$\sigma^2 \Sigma_{m,n} = \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} e^{i(m-n)\omega} \left| \frac{g_0 + g_1 z + g_2 z^2 + \dots}{h_0 + h_1 z + h_2 z^2 + \dots} \right|_{z=e^{-i\omega}}^2 d\omega.$$

Although the likelihood function, L , is derived under Gaussian processes, if the time series is causal and invertible, then maximising this likelihood leads to asymptotically consistent

estimators [42,43]. This approximation is convenient, as the asymptotic likelihood function has a unique stationary point, the global maximum [44].

Below, we introduce and analyse finite-parameter linear noise models and compare their performance, including the Gaussian white noise model for a benchmark.

4.1. Gaussian White Noise Model

The simplest noise model treats measurements as statistically independent, normally distributed random variables: $X_t - \mu = E_t$ for fixed μ . Thus, $X_t \sim \mathcal{N}(\mu, \sigma^2)$, with yet unknown parameter values μ and σ . Using the maximum likelihood approach, these parameters are estimated from the observations as

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N x_i \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{\mu})^2.$$

4.2. Autoregressive Models

An autoregressive process of order p , abbreviated as AR(p), is a stochastic process where h is a linear combination of p terms:

$$X_t = - \sum_{k=1}^p h_k X_{t-k} + E_t,$$

where h_k are constants. The process has $p + 1$ parameters: $\{h_k\}_{k=1}^p$ and σ . The estimators \hat{h}_i and $\hat{\sigma}^2 = \hat{\rho}(0) + \sum_{k=1}^p \hat{\rho}(k) \hat{h}_k$ are found by linear algebra [37]. Alternatively, the estimators can be found by least squares regression [45]. The parameter estimation procedure costs, at most, $O((p + 1)^2 + np)$ operations, where n is the number of observations.

As the Laplace transformation of X_t suggests the autocorrelation function of an AR(p) process is formally the sum of p exponentially decaying terms, thus, $\hat{\rho}(h)$ exhibits exponential decay or oscillatory behaviour [32], while $\hat{\rho}_p(h)$ vanishes for lags $p < h$.

4.3. Moving Average Model

Rather than using the past values of the random variable, a moving average model of order q , abbreviated as MA(q), regresses on past errors, $E_{t-k} = X_{t-k} - \mathbb{E}(X_t)$ for $k = 0, \dots, q$, and is defined as

$$X_t = \mu + \sum_{\ell=0}^q g_\ell E_{t-\ell}.$$

The process has $q + 2$ parameters. One may cast any stationary AR(p) process as an MA(∞) model and an invertible MA(q) model as an AR(∞) process. Nevertheless, MA(q) models are useful, as one cannot fit an AR(∞) model on a finite number of observations.

Unlike autoregressive processes, for MA processes, any estimator based on a finite amount of observations inherently carries some loss of information. Resultantly, parameter estimation is harder than for AR processes of comparable order.

The qualitative behaviours of the autocorrelation and partial autocorrelation functions, $\hat{\rho}(h)$ and $\hat{\rho}_p(h)$, are similar to those in autoregressive processes, but with roles interchanged. The autocorrelation function vanishes for $h > q$, while the partial autocorrelation function is dominated by damped exponential and/or periodic functions. In the examples shown in Figures 3 and 4, the autocorrelation function vanishes too slowly to be described by a low-order MA process.

4.4. Autoregressive Moving Average Model

The combination of AR(p) and MA(q) processes can be formulated as

$$X_t = \mu - \sum_{k=1}^p h_k X_{t-k} + \sum_{\ell=0}^q g_\ell E_{t-\ell}.$$

The combination inherits most of its properties from these simpler processes. Neither the autocorrelation nor the partial autocorrelation function vanishes after a threshold lag. The autocorrelation function exhibits damped exponential and/or oscillatory behaviour for lags $h > q - p$, while $\hat{\rho}_p(h)$ is dominated by exponential and periodic terms for lag values $h > p - q$. The variance of the sample autocorrelation function can exceed the Box–Jenkins error (see Figure 3). As a result, it is often difficult to determine the model orders; model verification and iterative parameter estimation are needed.

For normally and independently distributed E_t with zero mean and constant variance, and for stationary $\{X_t\}$, Melard [46] provided an efficient algorithm for calculating the exact likelihood function of an ARMA(p, q) model. This algorithm requires $O(p^2 + q^2 + n(p + 3q))$ operations for n observations, which—in practice—can be reduced by considering the observations in a finite window.

4.5. Ornstein–Uhlenbeck Model

In this section, we consider a stochastic differential equation (SDE) rather than a discrete stochastic process. Stochastic partial differential equations are a natural choice to model dynamical systems subject to random fluctuation. For an introduction to SDEs, we point to Evans' [47] and Øksendal's [48] excellent monographs.

The scatter-plots in Figure 1 resemble a random walk on a phenomenological level. However, in a Brownian motion, the variance of the position increases with time, which we do not observe here. On the contrary, we observe (and expect) that the GPS reading will return to the true location even if it deviates from it momentarily, i.e., it is mean-reverting. Furthermore, Figure 3 shows that the correlation function decreases exponentially. These properties suggest an Ornstein–Uhlenbeck (OU) process [49], defined by the SDE

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t, \quad (6)$$

where θ , μ , and σ are real parameters, with $\theta, \sigma > 0$. For historical and physical reasons, the parameters θ and σ are called the drift and diffusion coefficients, respectively. The fluctuation part, $W_t \sim \mathcal{N}(0, t)$, is a Wiener process. The same process can also be represented as

$$X_t = X_0 e^{-\theta t} + \mu (1 - e^{-\theta t}) + \sigma e^{-\theta t} W_{(e^{-2\theta t} - 1)/2\theta}.$$

The first term corresponds to the long-term but diminishing influence of the initial observation. The second reaches the steady value μ exponentially fast, while the third describes a scaled and time-transformed random fluctuation. The expectation value at time t is $\mathbb{E}(X_t) = \mu (1 - e^{-\theta t}) + e^{-\theta t} \mathbb{E}(X_0) \approx \mu$.

One may simplify (6) by eliminating μ with the $X_t \mapsto X_t - \mu$ transformation, but we keep μ explicit, as it is the 'true' position of the GPS device. The parameter θ sets the scale of the autocorrelation function $\rho(h)$, while σ determines the noise between measurements. The OU process is the continuous-time generalisation of an AR(1) process. It has several desirable properties: (a) It is mean-reverting for $\theta > 0$, (b) it has finite variance for all $t \geq 0$ contrary to Brownian motion, (c) its analytic covariance $\text{cov}(X_t, X_u) \sim e^{-\theta|t-u|}$ permits a long autocorrelation time, (d) the autocorrelation function is insensitive to the sampling rate, and (e) it approaches Gaussian white noise as $\theta \rightarrow \infty$. More precisely, the OU process is the only non-trivial, mean-reverting, Gauss–Markov process, i.e., whose transitions are memory-less, and it thus allows predictions of the future state based on the

current state [50]. This does not prohibit more sophisticated models for GPS time series, but these models will lack at least one of the properties listed above.

To model our observations as an Ornstein–Uhlenbeck process, we must estimate θ , μ , and σ using maximum likelihood estimation [51]. Tang and Chen [52] provided a detailed derivation of estimators for these parameters, and Dacunha-Castelle and Florens-Zmirou [53] showed that for time-equidistant observations, the maximum likelihood estimators are consistent and asymptotically normal, irrespective of the sampling time-step, Δt .

The estimators for θ , μ , and σ require $\sim 2n$ operations each. Therefore, the overall parameter estimation takes $O(6n)$ operations. However, Tang also gives explicit estimates for the asymptotic form of these estimators, e.g.,

$$\mathbb{E}\left(e^{-\hat{\theta}\Delta t}\right) = e^{-\theta\Delta t} - \frac{4 - 3\theta\Delta t}{n} + O\left(\frac{1}{n^2}\right).$$

Thus, one needs only a moderate number of observations, n , or short sampling time Δt to estimate the θ parameter, especially compared with the estimator of the mean for an uncorrelated process, which would have $O(1/\sqrt{n})$ (much slower) convergence to the true mean.

Parameter estimation for SDEs is a rapidly developing field, especially the sub-field of online estimators that update the parameter estimates as newer observations become available. Such estimators are usually associated with constant computational cost, as their performance is independent of n . A method of online estimators for a sampled, continuous-time Ornstein–Uhlenbeck is the subject of future publication.

After estimating μ , θ , and σ , we simulated an ensemble of Ornstein–Uhlenbeck processes using the Euler–Maruyama method [54]. We calculated and plotted the autocorrelation function for each run (see Figure 6) as well as the range covered by these functions. Clearly, the OU noise model represents the autocorrelation in the data up to at least two hours. However, it is unable to capture the undulation of $\hat{\rho}(h)$. Therefore, an Ornstein–Uhlenbeck process should be complemented with some other process for longer-term predictability. We believe that a more complicated noise model is unwarranted unless the sampling period exceeds an hour.

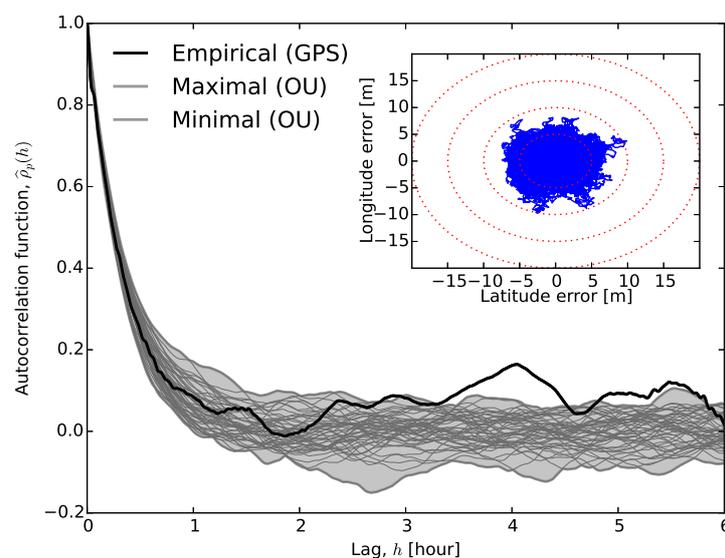


Figure 6. The thick black solid line depicts $\hat{\rho}(h)$ for the GPS recordings on 15 August, 2014 in Dunedin, as in Figure 3. All grey lines represent the autocorrelation function of a simulated Ornstein–Uhlenbeck process with input parameters $\mu = \hat{\mu} \cong 45.864$, $\theta = \hat{\theta} \cong 1.329 \times 10^{-4}$, and $\sigma = \hat{\sigma} \cong 2.745 \times 10^{-7}$. The lines labelled as ‘maximal’ and ‘minimal’ are connecting the extremal values of simulated autocorrelation functions at all lags.

5. Quantitative Analysis of Noise Models

In this section, we present and discuss the quantitative performance of the previously described noise models. For each model, the parameters were chosen to best fit the location data. The process of selecting which of these models best describes the data, called model selection, is a large and complex field in statistics; thus, we leave detailed explanation to the technical literature, e.g., [55] and references therein.

There are three properties of model selection criteria: consistency, efficiency, and parsimony. Consistency means that the criterion selects the true model with probability approaching unity with increasing sample size provided the true model is among those considered. However, nothing guarantees that the true model is simple and tractable; thus, it is also necessary to measure how much error one might make by accepting a model knowing that it is not the true model. Efficiency is a measure of this property; as the sample size becomes large, the error in predictions using the selected model becomes indistinguishable from the error obtained using the best model among the candidates. Parsimony dictates that if all measures are equal for two competing statistical models, the preferred one should be the simpler model.

The most common model selection criteria use the likelihood function, L , and introduce some penalty proportional to the model complexity. Akaike's Information Criterion (AIC) is defined as $AIC = -2 \ln(L) + 2k$, where k is the number of fit parameters in the model [56,57]. The best model has the smallest AIC value—it encodes the most information using the least number of fitted parameters. The AIC has a definite minimum value with respect to k , and the appropriate order of the model is determined by the value of k at this minimum. The AIC is biased for small samples, and a bias-corrected version (AICc) is gaining popularity [43]. Here, we analyse time series with $n > 10^5$ and $k \approx 10$, so we can ignore this correction. Other criteria have been developed, e.g., Bayesian Information Criterion (BIC), which more severely penalises the introduction of extra model parameters by replacing the $2k$ term with $k + k \ln(n)$.

Regarding the model selection criteria, AIC and AICc are both inconsistent, while BIC is only asymptotically consistent. Meanwhile, AIC is efficient, while BIC is not. Therefore there is no clear choice between AIC and BIC. The latter is asymptotically consistent as the sample size approaches infinity, while the former selects models that have acceptable predictive power [58].

We use AIC, as it is almost universally accepted. While it does not validate a statistical model, it determines relative probabilities between alternative models. We compare AIC values for a selection of models. Table 2 shows AIC scores for model fits of latitude, longitude, and altitude at four geographic locations. For each location, we individually fitted different noise models using the same observation datasets. For the AR(p), MA(q), and the ARMA(p, q), we increased the number of parameters involved. Occasionally the statistics routine was unable to converge for the MA(q) processes. In these cases, the corresponding entry in Table 2 is left empty.

Table 2 clearly shows that the Gaussian white noise model, irrespective of the geographic location or the coordinate, is the worst-performing noise model. All of its AIC values are positive, indicating a very low likelihood that the observations are described by the model. The MA models also seem a poor fit to the data, although one may increase the order of the model, q , to achieve lower AIC scores. Our standard numerical routine, based on the Yule–Walker method [37], could not converge. In contrast, an AR model with increasing order fit best among the noise models examined. Although both the ARMA(p, q) and Ornstein–Uhlenbeck models could achieve comparable AIC values (with the exception of the North Stradbroke Island dataset), the quantitative comparison of AIC values always favoured the high-order AR model. We note here that slightly lower AIC values could be achieved for autoregressive models by increasing their order, but the gain is becoming negligible around $p \approx 10$ (see Figure 7).

Figure 7 also captures two typical phenomena: For low values of p , the AR(p) models do improve significantly with increasing values of p , but there is a 'kink' in the AIC curve around approximately $p = 10$, where some improvement can be still achieved by increasing p , but these may

not be worthwhile considering the increasing computational workload. In many GPS applications, the sampling period is often shorter than a few minutes or an hour, so the aim of developing a noise model is to focus on the short-term behaviour of the time series. For this reason, the choice of $p = 9$ seems a good compromise between modelling performance and computational load (cf. Table 2).

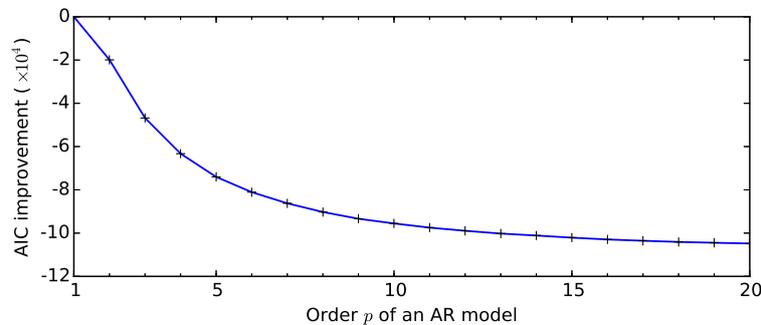


Figure 7. The Akaike’s Information Criterion (AIC) score for autoregressive (AR) models of increasing order using the Dunedin GPS altitude dataset recorded on 15 August 2014, scaled relative to the AIC score for $p = 1$.

Table 2. AIC values (rounded) of noise models for GPS time series.

Thursday Island	Gauss	AR(9)	MA(1)	MA(3)	ARMA(1, 1)	OU
Latitude	368,278	−419,119	240,593		−391,014	−390,082
Longitude	490,408	−358,983	369,403		−201,441	−212,528
Altitude	616,549	−151,204	488,436		−106,982	−107,027
Port Douglas					ARMA(1, 1)	
Latitude	1,080,721	−1,104,629	749,275		−1,048,908	−1,031,940
Longitude	1,254,838	−415,461	927,860		−887,968	−1,016,929
Altitude	1,445,244	−844,480	1,112,076		−924,883	−659,902
North Stradbroke					ARMA(2,2)	
Latitude	616,991	−846,881	417,772	160,425	−54,584	4099
Longitude	552,469	−847,589	369,864	138,410	48,340	5140
Altitude	991,469	−638,245	783,784	549,365	280,719	401,359
Dunedin					ARMA(2,2)	OU
Latitude	2,827,919	−4,855,973				−4,762,660
Longitude	2,202,644	−5,333,179	945,600		−5,343,304	−5,292,592
Altitude	4,380,261	−4,351,240	3,112,342		−4,360,823	−4,114,385

In summary, the Gaussian white noise model is the worst fit to the observation. The next worst model, according to the AIC, is a low-order MA process, followed by the Ornstein–Uhlenbeck model. A higher-order AR scores better, with the higher-order AR processes scoring best. The AR process has the further advantage of robust parameter estimation.

6. Discussion

The use of Gaussian white noise models is very common in sensor fusion algorithms that incorporate GPS data (Hamilton 1994). We have shown here that the assumption of uncorrelated noise does not capture important features of real-world GPS data. Fitting a Gaussian white noise model results in an artificially large variance and does not account for the observed autocorrelations. Analysis of location data shows that the error between the position measurement and the ‘true’ location shows strong autocorrelation up to approximately an hour. We proposed stochastic noise models that take into account such autocorrelation. Estimating the model parameters for the Ornstein–Uhlenbeck and

AR models based on likelihood maximisation is straightforward where either the analytic solution is known or the numerical algorithm is well known and can be performed efficiently. Conversely, estimating the parameters for the MA or ARMA models is more involved and computationally expensive. If the stochastic process is known to be non-stationary, then the MA model is non-invertible, and further measures are required to estimate the model parameters. Here, we mention an alternative approach for supporting the suitability of novel noise models. One can calculate the second moment of the residuals, $\langle E_f^2 \rangle$, using single-path or multipath ensemble averaging for the observed time series and the one generated by the model. However, this approach is left for future analysis.

Autoregressive processes of moderate order provide a sufficient noise model for the signal returned by the commercial GPS device, and the extra computational cost compared to that of a naive Gaussian white noise model is negligible. Increasing the order of the autoregressive process results in lower AIC values, indicating that, from the purely statistical perspective, the likelihood maximisation outweighs the increase in complexity. However, for the less accurate GPS units, the higher-order AR processes are only slightly more likely than the first-order AR(1) process. For an AR(1) process, we would advise a discretised Ornstein–Uhlenbeck process, which brings several advantages. First, the sampling period of the device is explicit in its formulation; thus, variable sampling periods can easily be fit into this description. Second, the other parameters of an Ornstein–Uhlenbeck process give clear insight into the noise characteristics: θ describes how the rate measurements revert towards the mean μ , while σ indicates the amount of noise present between each measurement.

Wendel and Trommer [59] analysed the effect of time-correlated noise on the applicability of classical Kalman Filtering, and proposed new filter equations. The noise models we proposed here can thus be incorporated into a Kalman Filter or Bayesian inference algorithms and provide both an estimate for the location along with quantified uncertainty. In addition, recent work [60] has shown that there are practical algorithms for sensor fusion in multi-dimensional systems that allow arbitrary noise distributions and approach the computational efficiency of Kalman Filtering.

Author Contributions: Conceptualisation: A.W.R.S., T.M., P.B., and D.S.; Methodology: A.W.R.S., A.M., and D.S.; Software and validation: A.W.R.S., A.M., and T.M.; Formal analysis: A.W.R.S., A.M., and D.S.; Resources, B.J.P., P.B., and T.C.A.M.; Writing: A.W.R.S., A.M., and D.S.; Review and editing: A.M., T.M., and D.S.; Visualisation: D.S.; Supervision: D.S. and T.C.A.M.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the New Zealand Ministry of Business, Innovation, and Employment under grant number UOOX-1208. D. Schumayer is grateful for the financial support from the Dodd-Walls Centre for Photonic and Quantum Technologies.

Acknowledgments: The authors are grateful for the reviewers for their recommendations, which helped to clarify some of our earlier discussions. We are also indebted for one of their suggestions, which will provide an alternative measure of goodness for the noise models presented in this study.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. Short Description of the QQ-Plot Method for Testing Normality

Before constructing non-Gaussian noise models, it is desirable to show that the distribution of noise is significantly different from a Gaussian distribution [23]. We thus carried out three normality tests for each coordinate (latitude, longitude, and altitude): the d’Agostino–Pearson test [25], the Lilliefors test [26], and the graphical QQ-plot method [61]. In Section 2, we demonstrated the deviation of the residuals from a Gaussian distribution. Here, a short description of the QQ-plot method is provided.

The central concept of the QQ-plot method is that of the quantiles, a measurable functional of the empirical distribution, such as the sample mean and sample median. The empirical distribution function, F_n , of a sample of size n , is defined as $F_n(x) = P_n^*((-\infty, x))$, or, in words, the ratio of elements of the sample that are less than x . The symbol $*$ means that the probability is estimated by the observed frequency. Assuming that $F_n(x)$ converges in probability to a distribution function $F(x)$

as $n \rightarrow \infty$, we can define the quantile, ζ_p , of order p as $\zeta(p) = \sup\{x | F(x) \leq p\}$. Here, the usage of sup guarantees that this definition remains meaningful even if $F(x)$ has points of discontinuity. As a function of p , the quantile $\zeta(p)$ is the inverse function of $F(x)$. A widely used example of a quantile is the median, which is simply $\zeta(\frac{1}{2})$, i.e., the value \hat{x} of the sample for which half of the observations are smaller than \hat{x} , and half of them are higher than \hat{x} . Similarly, we may divide up the entire range of the random variable into segments in a way that the number of samples belonging into each segment are the same. In the sample quantile $\zeta_n^*(p) = x_\ell$, where $\ell = [np] + 1$, x_k is equal to the k -th smallest value of the sample.

The sample quantiles provide us a graphical approach to check whether the sample comes from a given probability distribution, $G(x)$. Let us assume that X_1, X_2, \dots are independent, identically distributed random variables, with $E[X] = 0$ and $\text{Var}[X] = \sigma^2 < \infty$. Even if the observed values do not satisfy these mild conditions, after normalisation, the two conditions can be fulfilled. Let G denote the distribution function of the standard normal distribution, while F_n stands for the empirical distribution function of S_n^* , where

$$S_n^* = \frac{1}{\sqrt{n\sigma^2}} (X_1 + X_2 + \dots).$$

One measure of whether the sample $\{x_1, x_2, \dots\}$ comes from a standard normal distribution is the ‘closeness’ of F_n and G . The closeness is visually represented in a two-dimensional plot, where the abscissa is the t -th quantile, $\zeta(t)$, while the ordinate is the t -th sample quantile, $\zeta^*(t)$. The closer they are, the closer the points are to the diagonal. The plot of points $(\zeta(t), \zeta^*(t))$ for $t \in \mathbb{R}$ is called the quantile–quantile plot, or, in short, the QQ-plot.

References

- Zumberge, J.F.; Heflin, M.B.; Jefferson, D.C.; Watkins, M.M.; Webb, F.H. Precise point positioning for the efficient and robust analysis of GPS data from large networks. *J. Geophys. Res.* **1997**, *102*, 5005–5017. [[CrossRef](#)]
- Zandbergen, P.A.; Barbeau, S.J. Positional accuracy of assisted GPS data from high-sensitivity GPS-enabled mobile phones. *J. Navig.* **2011**, *64*, 381–399. [[CrossRef](#)]
- Olynik, M.C. *Temporal Characteristics of GPS Error Sources and Their Impact on Relative Positioning*; University of Calgary: Calgary, AB, Canada, 2002.
- Van Nee, R.D. Multipath effects on GPS code phase measurements. *Navigation* **1992**, *39*, 177–190. [[CrossRef](#)]
- Georgiadou, Y.; Kleusberg, A. On carrier signal multipath effects in relative GPS positioning. *Manuscripta Geod.* **1988**, *13*, 172–179.
- Sigrist, P.; Coppin, P.; Hermy, M. Impact of forest canopy on quality and accuracy of GPS measurements. *Int. J. Remote Sens.* **1999**, *20*, 3595–3610. [[CrossRef](#)]
- Griffiths, J.; Ray, J.R. On the precision and accuracy of IGS orbits. *J. Geod.* **2009**, *83*, 277–287. [[CrossRef](#)]
- Forte, B. Analysis of the PLL phase error in presence of simulated ionospheric scintillation events. *Radio Sci.* **2012**, *47*, 1–21. [[CrossRef](#)]
- Tiberius, C.; Jonkman, N.; Kenselaar, F. The stochastics of GPS observables. *GPS World* **1999**, *10*, 49–54.
- Xu, G.; Xu, Y. *GPS: Theory, Algorithms and Applications*; Springer: Berlin/Heidelberg, Germany, 2016.
- Santamaría-Gómez, A.; Bouin, M.N.; Collilieux, X.; Wöppelmann, G. Correlated errors in GPS position time series: Implications for velocity estimates. *J. Geophys. Res. Solid Earth* **2011**, *116*. [[CrossRef](#)]
- Zhang, J.; Bock, Y.; Johnson, H.; Fang, P.; Williams, S.; Genrich, J.; Wdowinski, S.; Behr, J. Southern California permanent GPS geodetic array: Error analysis of daily position estimates and site velocities. *J. Geophys. Res. Solid Earth* **1997**, *102*, 18035–18055. [[CrossRef](#)]
- Bos, M.; Fernandes, R.; Williams, S.; Bastos, L. Fast error analysis of continuous GPS observations. *J. Geod.* **2008**, *82*, 157–166. [[CrossRef](#)]
- Olivares, G.; Teferle, F. A Bayesian Monte Carlo Markov Chain Method for Parameter Estimation of Fractional Differenced Gaussian Processes. *IEEE Trans. Signal Process.* **2013**, *61*, 2405–2412. [[CrossRef](#)]

15. Jin, S.; Wang, J.; Park, P.H. An improvement of GPS height estimations: Stochastic modeling. *Earth Planets Space* **2005**, *57*, 253–259. [[CrossRef](#)]
16. Borre, K.; Tiberius, C. Time Series Analysis of GPS Observables. In Proceedings of the 13th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS 2000), Salt Lake City, UT, USA, 19–22 September 2000; pp. 1885–1894.
17. Segall, P.; Davis, J.L. GPS applications for geodynamics and earthquake studies. *Annu. Rev. Earth Planet. Sci.* **1997**, *25*, 301–336. [[CrossRef](#)]
18. Amiri-Simkooei, A.R. Noise in multivariate GPS position time-series. *J. Geod.* **2009**, *83*, 175–187. [[CrossRef](#)]
19. Williams, S.D.P.; Bock, Y.; Fang, P.; Jamason, P.; Nikolaidis, R.M.; Prawirodirdjo, L.; Miller, M.; Johnson, D.J. Error analysis of continuous GPS position time series. *J. Geophys. Res. Solid Earth* **2004**, *109*, B03412. [[CrossRef](#)]
20. Santamaría-Gómez, A.; Bouin, M.N.; Collilieux, X.; Wöppelmann, G. Chapter Time-Correlated GPS Noise Dependency on Data Time Period. In *Reference Frames for Applications in Geosciences*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 119–124. [[CrossRef](#)]
21. Kaplan, E.D.; Hegarty, C.J. *Understanding GPS Principles and Applications*, 2nd ed.; Artech House: Norwood, MA, USA, 2006.
22. Thode, H. *Testing for Normality*; Taylor & Francis Group: Abingdon, UK, 2019.
23. Yap, B.W.; Sim, C.H. Comparisons of various types of normality tests. *J. Stat. Comput. Simul.* **2011**, *81*, 2141–2155. [[CrossRef](#)]
24. Pearson, E.S.I. Note on Tests for Normality. *Biometrika* **1931**, *22*, 423–424. [[CrossRef](#)]
25. D’Agostino, R.B.; Belanger, A. A Suggestion for Using Powerful and Informative Tests of Normality. *Am. Stat.* **1990**, *44*, 316–321.
26. Lilliefors, H.W. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *J. Am. Stat. Assoc.* **1967**, *62*, 399–402. [[CrossRef](#)]
27. Abdel-Hafez, M.F. The autocovariance least-squares technique for GPS measurement noise estimation. *IEEE Trans. Veh. Technol.* **2010**, *59*, 574–588. [[CrossRef](#)]
28. Fallahi, K.; Cheng, C.T.; Fattouche, M. Robust positioning systems in the presence of outliers under weak GPS signal conditions. *IEEE Syst. J.* **2012**, *6*, 401–413. [[CrossRef](#)]
29. Durbin, J. The Fitting of Time-Series Models. *Rev. Int. Stat. Inst.* **1960**, *28*, 233–244. [[CrossRef](#)]
30. Durbin, J. *Efficient Fitting of Linear Models for Continuous Stationary Time Series from Discrete Data*; Technical Report; North Carolina State University, Department of Statistics: Raleigh, NC, USA, 1960.
31. Durbin, J. Efficient fitting of linear models for continuous stationary time series from discrete data. *Bull. Int. Stat. Inst.* **1961**, *38*, 273–281.
32. Cryer, J.D.; Chan, K.S. *Time Series Analysis: With Applications in R*; Springer: New York, NY, USA, 2008.
33. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*, 5th ed.; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2016.
34. Walker, A.M. Large-Sample Estimation of Parameters for Autoregressive Processes with Moving-Average Residuals. *Biometrika* **1962**, *49*, 117–131. [[CrossRef](#)]
35. Box, G.E.P.; Pierce, D.A. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *J. Am. Stat. Assoc.* **1970**, *65*, 1509–1526. [[CrossRef](#)]
36. Ljung, G.M.; Box, G.E.P. On a measure of lack of fit in time series models. *Biometrika* **1978**, *65*, 297–303. [[CrossRef](#)]
37. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C. *Time Series Analysis, Forecasting and Control*, 3rd ed.; Prentice Hall: Englewood Cliffs, NJ, USA, 1994.
38. Bartlett, M.S. On the theoretical specification of sampling properties of an autocorrelated process. *J. R. Stat. Soc. B* **1946**, *8*, 27–41.
39. Anderson, O. *Time Series Analysis and Forecasting: The Box-Jenkins Approach*; Butterworths: Carlisle City, UK, 1976.
40. Fishman, G. *Principles of Discrete Event Simulation*; Wiley Series on Systems Engineering and Analysis; John Wiley & Sons: Hoboken, NJ, USA, 1978.
41. Porat, B. *Digital Processing of Random Signals: Theory and Methods*; Prentice Hall Information and System Sciences Series; Prentice Hall: Englewood Cliffs, NJ, USA, 1994.

42. Akaike, H. Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika* **1973**, *60*, 255–265. [[CrossRef](#)]
43. Brockwell, P.J.; Davis, R.A. *Time Series: Theory and Methods*; Springer Series in Statistics; Springer: New York, NY, USA, 1991.
44. Astrom, K.; Soderstrom, T. Uniqueness of the maximum likelihood estimates of the parameters of an ARMA model. *IEEE Trans. Autom. Control* **1974**, *19*, 769–773. [[CrossRef](#)]
45. Mann, H.B.; Wald, A. On the Statistical Treatment of Linear Stochastic Difference Equations. *Econometrica* **1943**, *11*, 173–220. [[CrossRef](#)]
46. Melard, G. Algorithm AS197: A Fast Algorithm for the Exact Likelihood of Autoregressive-Moving Average Models. *J. R. Stat. Society. Ser. C Appl. Stat.* **1984**, *33*, 104–114. [[CrossRef](#)]
47. Evans, L.C. *An Introduction to Stochastic Differential Equations*; American Mathematical Society: Berkeley, Berkeley, CA, 2014.
48. Øksendal, B. *Stochastic Differential Equations: An Introduction with Applications*; Universitext; Springer: Berlin/Heidelberg, Germany, 2010.
49. Uhlenbeck, G.E.; Ornstein, L.S. On the Theory of the Brownian Motion. *Phys. Rev.* **1930**, *36*, 823–841. [[CrossRef](#)]
50. Doob, J.L. The Brownian Movement and Stochastic Equations. *Ann. Math.* **1942**, *43*, 351–369. [[CrossRef](#)]
51. Pedersen, A.R. A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations. *Scand. J. Stat.* **1995**, *22*, 55–71.
52. Tang, C.Y.; Chen, S.X. Parameter estimation and bias correction for diffusion processes. *J. Econom.* **2009**, *149*, 65–81. [[CrossRef](#)]
53. Dacunha-Castelle, D.; Florens-Zmirou, D. Estimation of the coefficients of a diffusion from discrete observations. *Stochastics* **1986**, *19*, 263–284. [[CrossRef](#)]
54. Kloeden, P.E.; Platen, E. *Numerical Solution of Stochastic Differential Equations*; Springer: Berlin/Heidelberg, Germany, 1992; Volume 23.
55. Claeskens, G.; Hjort, N.L. *Model Selection and Model Averaging*; Cambridge Series in Statistical and Probabilistic Mathematics; Cambridge University Press: Cambridge, UK, 2008; p. 332.
56. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [[CrossRef](#)]
57. Schwarz, G. Estimating the Dimension of a Model. *Ann. Stat.* **1978**, *6*, 461–464. [[CrossRef](#)]
58. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer Series in Statistics; Springer: Berlin/Heidelberg, Germany, 2003.
59. Wendel, J.; Trommer, G.F. An Efficient Method for Considering Time Correlated Noise in GPS/INS Integration. In Proceedings of the 2004 National Technical Meeting of The Institute of Navigation, San Diego, CA, USA, 26–28 January 2004; pp. 903–911.
60. Fox, C.; Norton, R.A.; Morrison, M.E.; Molteno, T.C. Sequential Bayesian Inference for Dynamical Systems Using the Finite Volume Method. In *2017 MATRIX Annals*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 13–23.
61. Thode, H. *Testing for Normality*; Marcel Dekker: New York, NY, USA, 2002.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).