Analysis of the correlations of TLS range observations from plane fitting residuals

Authors: Gaël Kermarrec[1] *, Michael Lösler[2], Jens Hartmann[1]

[1]Geodetic Institute, Leibniz Universität Hannover, Nienburger Str. 1, 30167 Hannover, Germany
[2]Faculty 1: Architecture – Civil engineering – Geomatics, Laboratory for Industrial Metrology, Frankfurt University of Applied Sciences, Nibelungenplatz 1, 60318 Frankfurt am Main, Germany

 * Correspondence: kermarrec@gih.uni-hannover.de; Tel.: +49-511-7621-4736

Abstract

Terrestrial laser scanners (TLS) record a large number of points within a short time. Correlations between observations are unavoidable but often neglected in stochastic modelling. The main consequences are an overestimated precision of the point clouds and potential wrong test decisions when used for deformation analysis with rigorous statistical procedures. Regarding physical considerations, a fractional Gaussian noise, defined by a so-called Hurst exponent, or a combination of fractional Gaussian noises could be used to model the noise of range measurements: temporal correlations are expected to have a long range dependency due to the high recording rate. Alternatively, a Matérn process can be used, where the power spectral density is damped at low frequencies. The corresponding parameters of both noise models can be estimated from the residuals of a least-squares approximation of the TLS point cloud: these residuals yield precious information about the correlation structure of the raw polar observations. Unfortunately, the correlation structure of the residuals may be affected by the scanning configuration and settings. In this contribution, we propose to quantify the impact of these effects on the estimation of a global Hurst exponent. Based on the results from simulations of a plane, real data analysis from indoor and outdoor experiments can be better understood, which allows one to identify the dominant noise structure and its parameters. Our methodology makes use of a combination of two estimators: the Whittle maximum likelihood and the generalised Hurst estimator; both of them provide information on the noise, paving the way for a simple and global model for TLS range correlations, usable in point clouds analysis independently of the object under consideration.

Keywords: Terrestrial Laser Scanner, correlations, Hurst parameter, residuals

1. Introduction

The latest generation of terrestrial laser scanners (TLS) is able to record millions of points during a short time period with a measurement accuracy of a few millimetres or less. Loaded in dedicated software, these point clouds can be used to compute, for example, a cloud to cloud or cloud to mesh distance (Holst et al. 2016; Lague et al. 2013). Promising applications include the monitoring of bridges, dams, tunnels and towers (Idrees and Pradhan 2018; Laefer and Truong-Hong 2017; Mukupa et al. 2016; Neuner et al. 2016; Suchocki 2020 and the references within). Statistical tests for deformation, such as the congruence test, can only be performed based on a parametric representation of the point cloud (Niemeier 2002); B-spline surfaces are a promising way to compute such an approximation (Kermarrec et al. 2020; Koch 2010). Unfortunately, tests based on surface approximation are the most powerful only if a reliable stochastic model is specified (Pelzer 1971): this requisite includes both the description of the variance and the correlation structure of the observations. Whereas the variance of

the TLS angles can be taken from the manufacturer's specifications, the variance of the range was shown to follow a power law function regarding the intensity values (Wujanz et al. 2017). This empiric all-embracing stochastic model depends on the TLS under consideration and can be determined from calibration. This proposal is physically justified by the radar range equation and accounts for different influencing factors, such as the property of the object scanned or the distance to the TLS (Soudarissanane et al. 2011). Using this model, a diagonal variance covariance matrix (VCM) of the raw observations can be set up. The latter, unfortunately, does not account for correlations between the measurements:

- Statistical tests may be biased (Kermarrec et al. 2020) so that an incorrect rejection of the null hypothesis that a deformation occurred can happen or
- the precision of the parameters adjusted may be underestimated (cf. Jäger et al. 2005, Ch. 5.5).

The detection of correlations is nearly impossible by visual inspection from a time series, such as the residuals of a mathematical approximation from a TLS point cloud, as illustrated in Figure 1 (top). Fortunately, the non-null slope of their power spectral densities (psds) provides some indication of the presence of correlations and their structure: the challenge of modelling range correlation can be addressed, with the aim of deriving a practically usable correlation function. In this contribution – and our work in general – we follow the goal that the latter should be usable by a TLS practitioner and implemented in free software to judge the point cloud quality with a higher trustworthiness.
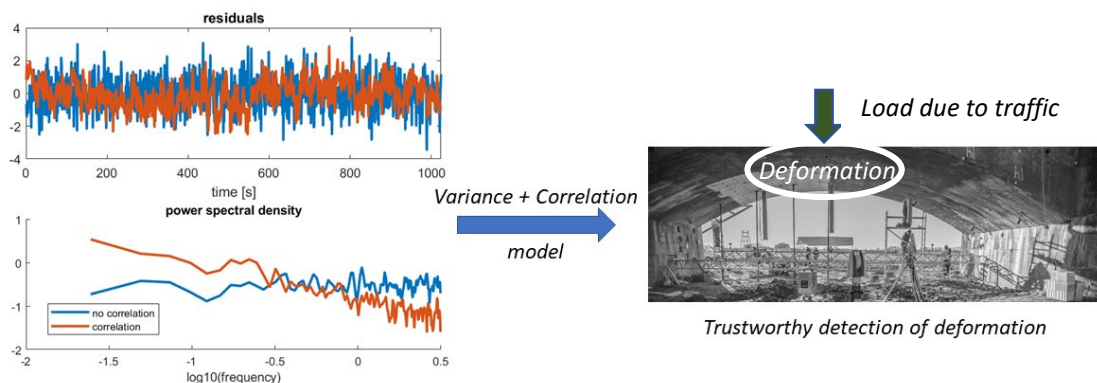

Fig 1: left: the residuals of a simulated point cloud approximation (top) and their psd in the case of correlated (bottom, red line) and non-correlated (bottom, blue line) observations. Right: Including information about the correlation structure leads to a trustworthy detection of deformation

Kauker and Schwieger (2017) and Schmitz et al. (2020) started investigating the correlations of TLS range measurements using an elementary model and the resolution capability of the scanner, respectively. They both based their modelling on spatial correlations, concluding that correlations should not be undertaken. They used, for example, the concept of the effective number of observations. Kermarrec et al. (2018) started alternatively from the raw observations of the sensors, making a parallel between TLS and the Global Navigation Satellite System (GNSS) phase observations: they proposed to model their correlation structures as temporal, i.e. time-dependent. Consequently, they chose the Matérn covariance model, a popular and widely used function in geostatistics (Gelfand et al. 2010) and GNSS co-ordinate time series analysis unter the namegeneralised Gauss-Markov noise model (Bos et al. 2013).

This contribution aims to validate the proposed temporal model for TLS range correlations. We developed a new methodology to estimate the corresponding correlation parameters from the residuals of a least-squares (LS) adjustment, starting from the raw point clouds. Both the simulations and real data of a plane will be used to show the potential of our proposal. Indoor and outdoor experiments should highlight that the correlation structure of single-mode laser light as a combination of a white and a flicker part can be found in the residuals (the so-called Voigt profile, see Di Domenico et al. 2010; Stéphan et al. 2005). Moreover, the expected dependency of the correlation regarding the distance to the object, the data quality and the data rate will be analysed.

More specifically:

- We will investigate how and when the residuals of a plane adjustment can be used to derive trustworthy correlation parameters. We will introduce two dedicated correlation estimators and simulate different scanning scenarios to reach that goal.
- We will apply the methodology and knowledge gained from the simulations to estimate the global correlation parameters of TLS range measurements using real data sets from a plane scanned with a Z+F IMAGER 5016. Both indoor and outdoor observations will be used for the sake of comparison.
- We will show the extent to which the correlation structure varies depending on the scanning settings and the distance from the TLS to the scanned object. We aim to provide physically plausible explanations of the results by additionally analysing their spectral content.

These investigations are performed using recent developments from time series analysis to estimate trustworthy correlation parameters (Sykulski et al. 2019). They pave the way:

- for the analysis of the noise structure of TLS range observations,
- and, thus, for the determination of a global model for TLS range correlations without using any iterative approaches, such as variance covariance estimation methods (Teunissen and Amiri-Simkooei 2008).

The methodology developed is independent of the scanner under consideration and can be used with freely available functions. It can be easily implemented and extended to more challenging surface modelling using, for example, B-spline surfaces, without any limitation of the shape of the object scanned.

2. Material and Methods

A plane is a simple object widely used to calibrate and estimate the uncertainties of sensors. For that reason, we will make use of both simulated and real data corresponding to a plane in this contribution. This section gives a brief mathematical background about the estimation of its parameters using a non-linear Gauss-Helmert model (Lenzmann and Lenzmann 2004). We will put the stress on how to obtain the range residuals of the approximation, from which we will estimate the correlation parameter of the original observations. We also introduce the noise model chosen by briefly defining the fractional Gaussian noise (fGn) and the way to estimate its parameter: the Hurst exponent (Koutsoyiannis 2002 and the references within).

2.1 Plane fitting with non-linear LS from TLS observations

A plane is an infinitely extended surface defined by a normal vector $\mathbf{n}^{\mathrm{T}} = \begin{pmatrix} n_x & n_y & n_z \end{pmatrix}$ and a surface parameter d, i.e.,

$$\mathbf{n}^{\mathrm{T}} \mathbf{P}_i = d, \tag{1}$$

where $\mathbf{P}_i$ is an arbitrary Cartesian point lying on the plane (Bronshtein et al. 2007, p. 214f). The normal vector $\mathbf{n}$ and the surface parameter d depend on each other. A change in the vector length of $\mathbf{n}$ yields a corresponding change in d but represents the identical plane. An additional restriction is needed to overcome the ambiguity of the plane parameters. A common approach is to normalize the normal vector, i.e.,

$$\mathbf{n}^T\mathbf{n} = 1. \tag{2}$$

In this case, Eq. (1) becomes the so-called Hesse normal form and the surface parameter d represents the shortest distance between the origin of the co-ordinate system and the plane.

The parameters of the plane can be estimated for a given set of points not lying on a straight line using LS techniques. By treating the points as observations and introducing corresponding residuals $\mathbf{v}_i^T = \begin{pmatrix} v_x & v_y & v_z \end{pmatrix}$, which we considered as being distributed as $\mathbf{v} \sim N(0, \Sigma)$, the functional model of the plane reads:

$$\mathbf{f}(\mathbf{x}, \mathbf{v}) = \mathbf{n}^T(\mathbf{P}_i + \mathbf{v}_i) - d = 0. \tag{3}$$

Here, $\mathbf{x}^T = \begin{pmatrix} n_x & n_y & n_z & d \end{pmatrix}$ is the vector of parameters to be estimated.

The well-known objective function of the LS adjustment is given by

$$\Omega(\mathbf{v}) = \mathbf{v}^T\Sigma^{-1}\mathbf{v} = \min \tag{4}$$

and minimises the weighted sum of squared residuals. The objective function, the condition equations and the side condition are combined by the Lagrangian function to estimate the parameters. By introducing appropriate approximation values $\mathbf{x}_0$ and $\mathbf{v}_0$, the linearized normal equation of the Lagrangian function yields (Kupferer 2005, p. 39)

$$\begin{bmatrix} \mathbf{B\Sigma B}^T & \mathbf{A} & \mathbf{0} \\ \mathbf{A}^T & \mathbf{0} & \mathbf{R}^T \\ \mathbf{0} & \mathbf{R} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{k}_1 \\ \mathbf{d\hat{x}} \\ \mathbf{k}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{-w} \\ \mathbf{0} \\ \mathbf{-r} \end{bmatrix}. \tag{5}$$

Here, the vectors of Lagrangian multipliers are denoted by $\mathbf{k}_1$ and $\mathbf{k}_2$. The matrices $\mathbf{A}$ and $\mathbf{B}$ contain the partial derivatives of the linearized functional model $\mathbf{f}$ of the plane regarding the unknown parameters and the residuals, respectively. The linearized restriction of the side condition, cf. Eq. (2), is given by $\mathbf{R d\hat{x}} + \mathbf{r} = \mathbf{0}$ and $\mathbf{w} = -\mathbf{B v}_0 + \mathbf{f}(\mathbf{x}_0, \mathbf{v}_0)$ is the vector of misclosures. The iterates $\hat{\mathbf{x}} = \mathbf{x}_0 + \mathbf{d\hat{x}}$ and $\mathbf{v} = \Sigma\mathbf{B}^T\mathbf{k}_1$ are introduced as new approximations to the next iteration step until the system of equations converges. Eq. (5) describes a general LS method, which is known as the Gauss-Helmert model. For the detailed derivation of the Gauss-Helmert model, the interested reader is referred to the contributions by Koch (2014) and Neitzel (2010).

The vector $\mathbf{P}$ in Eq. (1) represents Cartesian co-ordinates: TLS raw observations are recorded as polar co-ordinates so that a co-ordinate transformation is needed (see Fig. 2). Even though Lösler (2020, to be published) showed that the LS results are independent of the choice of the co-ordinate representation if (i) the functional model, (ii) the stochastic model and (iii) the objective function are rigorously transformed, the condition number of the normal equation system depends on the co-

ordinate representation. Numerically stable equation systems are more likely for models using Cartesian co-ordinates. For that reason, the raw observations and the related dispersion matrix are transformed into their Cartesian representations by

$$\mathbf{P}_i = \begin{pmatrix} x \\ y \\ z \end{pmatrix}_i = \begin{pmatrix} r\sin\phi\cos\lambda \\ r\sin\phi\sin\lambda \\ r\cos\phi \end{pmatrix}_i , \tag{6a}$$

$$\mathbf{\Sigma}_{\mathbf{cart}} = \mathbf{F}\mathbf{\Sigma}_{\mathbf{pol}}\mathbf{F}^{\mathrm{T}}, \tag{6b}$$

where $\mathbf{F}$ denotes the matrix of the linear transformation between polar and Cartesian co-ordinates.
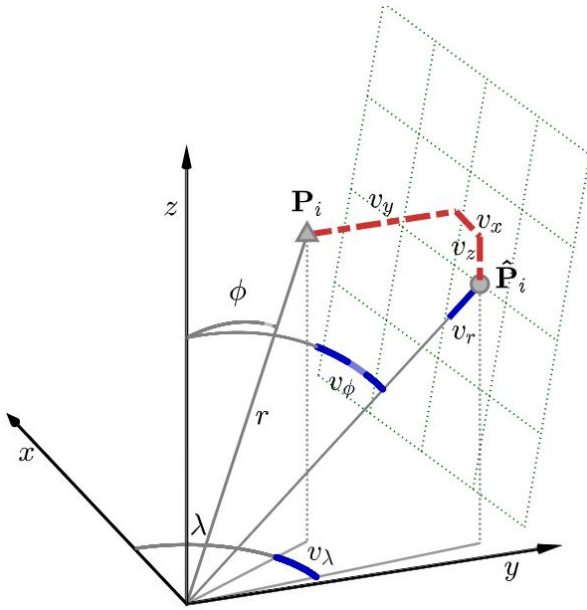

Fig 2: Schematic representation of the transformation between polar and Cartesian co-ordinates

2.1. Backward transformation of the residuals

In this contribution, we aim to estimate the correlation parameter of the TLS range measurements from the residuals of the adjustment, i.e. $\mathbf{v} \sim N(0,\mathbf{\Sigma})$. Consequently, the residuals of the plane adjustment (Eq. (3)) have to be backwards transformed to polar co-ordinates.

According to Lösler (2020), the pointwise polar residuals are expressed by means of the co-ordinates and the cartesian residuals (see Fig, 2), as:

$$v_{\lambda i} = \sqrt{\left(x_i + v_{x_i}\right)^2 + \left(y_i + v_{y_i}\right)^2 + \left(z_i + v_{z_i}\right)^2} - \sqrt{x_i^2 + y_i^2 + z_i^2} \tag{7a}$$

$$v_{\phi i} = a\tan\left(\frac{\sqrt{\left(x_i + v_{x_i}\right)^2 + \left(y_i + v_{y_i}\right)^2}}{\left(z_i + v_{z_i}\right)}\right) - a\tan\left(\frac{\sqrt{\left(x_i\right)^2 + \left(y_i\right)^2}}{\left(z_i\right)}\right)$$

$$v_{\phi i} = a\tan\left(\frac{\sqrt{\left(x_i + v_{x_i}\right)^2 + \left(y_i + v_{y_i}\right)^2}}{\left(z_i + v_{z_i}\right)}\right) - a\tan\left(\frac{\sqrt{\left(x_i\right)^2 + \left(y_i\right)^2}}{\left(z_i\right)}\right) \tag{7c}$$

Henceforth, we call the range residuals the vector $\mathbf{v}_r = \left[v_{r1}, \ldots, v_{rn}\right]$. This vector is assumed to be temporally sorted: $v_{ri} = \mathbf{v}_r\left(t_i\right)$, t_i being the epoch at which the corresponding observation was made. We call $C_{\mathbf{v}_r}$ the covariance between two range measurements taken at t_i and $t_i + \tau$, τ being the time increment:

$$C_{\mathbf{v}_r}\left(t_i, t_i + \tau\right) = E\left(\left(\mathbf{v}_r\left(t_i, t_i + \tau\right) - E\left(\mathbf{v}_r\right)\right)\left(\mathbf{v}_r\left(t_i\right) - E\left(\mathbf{v}_r\right)\right)\right) \tag{8}$$

$E\left(\bullet\right)$ is the expectation operator.

In a first approach, we assume that the noise angles are uncorrelated. In this contribution, only the TLS range measurements are considered to be temporally correlated.

2.2. Noise model

Rather than estimating the covariance function from the residuals empirically using, for example, LS variance covariance estimation (Teunissen and Amiri-Simkooei 2007), we propose to model the correlated noise as a fGn (Granger and Joyeux 1980; Mandelbrot and Ness 1968). This choice is based on strong physical assumptions: it combines sensor knowledge with the propagation of electromagnetic waves in random medium and follows the previous work of Kermarrec and Schön (2014) for GPS (see also Wheelon 2001).

Modelling of the noise as a fGn has the advantage that only one correlation parameter has to be estimated: the Hurst parameter. Many estimators for this parameter are available either in the time or frequency domain, allowing a great flexibility in case of potential additional effects due to a functional mismodeling, outliers, a small number of observations or additional white noise (WN). Interested readers can refer exemplarily to Bardet et al. (2003).

Prior to introducing this noise, we propose to briefly explain the reason why a temporal modelling was chosen.

Why temporal correlations?

We model the correlation of the TLS range observations as temporal, i.e. time-dependent. This is justified by the fact that range measurements – whether from phase or pulse lasers – are a measure of time (Rueger 1996). Figure 3 shows exemplarily the laser signal, emitted at time t and received after reflection and a double atmospheric propagation at $t + dt$ (Andrew and Phillips 2005). The range is computed from the time difference dt: it is, thus, indirectly a measure of time (Pfeifer and Briese 2007). Consequently, the signal emitted will have some affinity with the signals received at different epochs:

- They travelled the same random medium, which acts to correlating them (atmospheric noise), and

- they were emitted by the same instrument and still carry some information about the noise of the original signal. The noise structure of a laser was shown, both empirically and theoretically, to be a combination of WN at high frequencies and flicker noise (FN) at low

frequencies (see, e.g., Di Domenico et al. 2010). We expect, thus, to find this structure in the residuals of the adjustment, which reflects the laser measurement noise.

Due to the large number of observations recorded by the TLS in a short amount of time, the degree of affinity between the reference signal and all other ones recorded with an increasing time interval will stay high: the temporal correlation is expected to be long-range dependant. Any spatial effects stemming from the reflected surface affect the intensity of the signal received and are modelled by the variance factor, following, for example, Wujanz et al. (2017). By means of this modelling, the correlation function is said to be "separated" (Gelfand et al. 2010). The intensity factor for objects with strong intensity variations could take a similar form to the elevation dependency for GNSS (see Kermarrec and Schön 2014, Eq. 10). The analysis of the variance is beyond the scope of this paper, which concentrates on the correlation structure of range measurements.
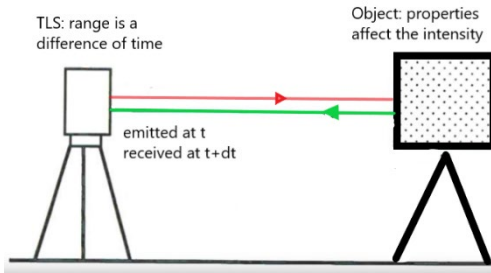

Fig. 3: Motivation of modelling correlations as temporal

2.2.1. Fractional Gaussian noise (fGn)

We postulate that the high rate of measurements induces a long dependency between the observations. The temporal correlations between the first observations may decay quickly but will stay much higher than zero (Koutsoyiannis 2002). Processes with a long-range dependence possess this property, which is linked with the slow decay of their autocorrelation C_{LRD} to zero, so that their sum does not converge (Beran 2013). More precisely,

$$C_{LRD}(\tau) \approx c\tau^{-\delta} \tag{9}$$

where τ is the time lag, c a positive constant and $0 < \delta < 1$. As τ increases, the dependence between the observations stays strong, which implies a fat-tailed autocorrelation function.
The long-range dependence of the fGn for a stationary process is contained in the Hurst exponent. This exponent is related to the self-similarity of the process: the latter appears statistically identical under rescaling the time axis. The fGn is fully characterized by the Hurst exponent H_{fGn} and the variance $\sigma^2_{H_{fGn}}$, which has the nice property to be bounded; this property is more likely to happen in a real case scenario. Its autocorrelation $C_H(\tau)$ of a fGn is given by

$$C_H(\tau) = \frac{1}{2}\left(|\tau+1|^{2H} - 2|\tau|^{2H} + |\tau-1|^{2H}\right). \tag{10}$$

The fGn has a power law spectrum, i.e. $W(f) \propto \frac{1}{f^\beta}$, with f, β being the frequency and the power law of the process, respectively. This model is popular for many kinds of physical processes, particularly atmospheric ones (Vyushin and Kushner 2009). The psd appears on a log-log plot as a simple straight line.

Regarding an fGn, we have the relationship $H_{fGn} = \dfrac{\beta+1}{2}$ with $-1 < \beta < 1$. A noise with $H_{fGn} = \frac{1}{2}$ is called WN, whereas $H_{fGn} = 1$ is FN.

Generation of an fGn

In this contribution, we generate the fGn noise with the Matlab function called ffG, freely available in the Matlab file exchange section (Stoev 2020). It is based on the circulant embedding method for persistent noise, resulting in a reproduction of its exact autocovariance. This noise is simulated: ee will never get the exact value of the Hurst parameter. Monte Carlo (MC) simulations are unavoidable when studying the performance of the parameter estimation for the residuals of the plane fitting.

2.2.2. Matérn model

The fractional Brownian motion (fBm) is a generalisation of a Brownian motion (Lilly et al. 2017). It is a non-stationary process with stationary increments and possesses the long-term memory, also called persistency or positive correlations when $H_{fGm} > \frac{1}{2}$.

A Matérn process is an extension of the fBm. The power spectrum of a Matérn process is still fractional but includes an additional damping parameter α, i.e.

$$W(\omega) = \frac{\alpha^{2\nu-1}}{c_{matern}} \frac{\sigma^2_{matern}}{\left(\omega^2 + \alpha^2\right)^\nu} \tag{11}$$

where ω is the frequency. Moreover, σ^2_{matern} is the variance of the process, c_{matern} a normalizing constant and ν the smoothness parameter. ν is linked with the slope of the psd at high frequencies and is defined as half $\frac{\beta}{2}$ (Voss 1991). Through the damping of $W(\omega)$ at low frequencies, an unlimited growth of the process is avoided: this property may be considered as hardly explainable for residuals of a plane adjustment, making the Matérn process an alternative to the modelling of the range correlation with an fBm.

When the damping factor is close to zero, the Matérn process becomes a fGn (Stein 1999), which reduces the number of parameters to estimate. The smoothness is linked with the Hurst parameter by the relationship $H_{fGn} = \nu - \frac{1}{2}$. A prior knowledge of the expected noise is mandatory for a correct use of the estimators, i.e. one should always ensure that the process used to estimate the smoothness is a fBm, using, for example, the cumulative sum of the original process when this latter is a fGn.

2.2.3. Noise modelling for range measurements

In the following, we adopt the notation $N_r = WN + fGn$, where N_r is the TLS range noise, expressed as a combination of two fGn: the WN and a correlated noise for which $H_{fGn} > \frac{1}{2}$. In the following Sections 2.2.1. and 2.2.2, fGn can be a pure fGn, a combination of fGn with different Hurst parameters (see Paschotta 2008 for the specific case of laser noise) or a damped fBm, i.e. a Matérn process. We do not combine the Matérn noise with a fGn for the sake of simplification and due to the challenging parameter estimation.

We define the fraction of WN as $R_{WN} = \dfrac{\sigma_{WN}^2}{\sigma_{H_{fGn}}^2}$, where σ_{WN}^2 is the variance of the WN. We note that

when $R_{WN} = 0$, the estimation of the Hurst parameter H_{N_r} from N_r corresponds to H_{fGn} .

When $R_{WN} \neq 0$, two modelling strategies of the range correlation can be proposed:

- We filter the range residuals from their WN component, using, for example, maximum likelihood, as in the Hector software (Bos et al. 2003). However, from the experience of the authors with simulated fGn time series, no WN could be detected in more than 50 % of the cases, although one was added to the fGn. Similar experiences were made by using the EM algorithm (using an autoregressive noise model; see Kargoll et al. 2018). Such untrustworthy estimation biases the Hurst exponent estimation accordingly, making a separation fGn/WN challenging.

- We estimate a biased and underestimated Hurst H_{N_r} (or, alternatively, a smoothness parameter v_{N_r} for a Matérn model). Using this strategy, we consider indirectly that $R_{WN} = 0$. The amount of WN is, thus, included in H_{N_r} without having to estimate any ratio. The price to pay with this approximation is a loss of physical meaning for the Hurst parameter, although variations of H_{N_r} give precious information about the dominant noise sources, as shown in Section 4.

In this contribution, we will make use of the latter strategy. We justify our choice by the challenging extraction of multiple fGn from a time series, which necessitates previous knowledge about the parameters expected. Consequently, we propose a global model and estimate only one Hurst parameter. Random walk ($\beta = 2$) is present at low frequencies and mostly associated with functional misspecifications which are not considered in this contribution. We will simulate, measure and estimate planes, eliminating the risk of such challenging situations. However, an indication of the presence of random walk is given by the strong increase of H_{N_r} , which can even become higher than 1 with some estimators, as shown in Section 2.2.5.

The accurate and physically meaningful estimation of H_{fGn} only necessitates specific filtering procedures: they are beyond the scope of this paper and lead to the next contribution where specific methods from time series analysis will be used. We here, thus, focus on a global estimation of the noise parameters.

2.2.4. Estimation of the Hurst parameter

A power law noise can be considered as a Matérn process with α close to zero, provided that the noise is transformed into a fBm: it is, thus, possible to use the same estimator for both noise models mentioned previously. We will compare two estimators for the sake of completeness: (i) the Whittle likelihood estimator (WhiE), which is specifically designed for the Matérn process, and (ii) the generalised Hurst estimator (GHE) used here in the time domain for the estimation of the Hurst parameter.

Debiased WhiE method
The maximum likelihood estimator is a purely numerical method which is often considered to be the best estimator obtainable, provided that the process is consistent with the model. Although the estimates are asymptotically unbiased, the estimator is asymptotically efficient and fast to compute. It may perform poorly if the assumption is incorrect or for short samples: the use of information criteria

(Section 2.3.5) is, thus, unavoidable to check the model assumption. We can evaluate the log-likelihood for Gaussian data as

$$l(H) = -\log\left(\left|\mathbf{\Sigma_H}\right|\right) - \mathbf{X_H^T}\mathbf{\Sigma_H^{-1}}\mathbf{X_H} \tag{12}$$

where $\mathbf{X_H}$ denotes the column vector of length n_{obs} and $\mathbf{\Sigma_H}$ is a fully populated VCM, whose components are given using the covariance function derived from Eq. (10). $|\bullet|$ states for the determinant of the matrix. Matrix inversions can be avoided using the WhiE, which aims to provide faster estimation with only a slight inaccuracy. In that case, the Whittle likelihood in its discretized form is given by

$$l_W(H) = -\sum_{\omega \in \Omega}\left[\log\left(\tilde{f}(\omega, H)\right) + \frac{I(\omega)}{\tilde{f}(\omega, H)}\right] \tag{13}$$

with Ω as the set of discrete Fourier frequencies, $\tilde{f}(\omega, H)$ the continuous-time process spectral density and $I(\omega)$ the periodogram $I(\omega) \infty \sum_{j=1}^{N}\left|X_{H,j}e^{-ij\omega}\right|^2$.

In this contribution, we use the WhiE as implemented in Matlab by Lilly (2020). The cut-off or the smoothness parameter can be fixed to optimise their determination (see Kaufman and Shaby 2013). Exemplarily, if the noise of the process corresponds to a power law noise, we can fix $\alpha = 1e^{-4}$ and estimate H_{fGn}.

An alternative to the WhiE is provided by the GHE. This estimator could be shown to be optimal in estimating the Hurst exponent in the case of additional low frequencies due to a suboptimal functional model (Kermarrec 2020). We propose to introduce this estimator shortly.

The generalised Hurst estimator

The GHE was introduced by Barabasi and Vicsek (1991). It is used in medical or financial domains (Di Matteo et al. 2003). We make use of the GHE in the time domain, which we define by using the first-order moments of the distribution of increments: $K_1(t) = \dfrac{\left\langle\left|X_H(t+\tau) - X_H(t)\right|\right\rangle}{\left\langle\left|X_H(t)\right|\right\rangle}$. τ is varied between 1 and $\tau_{\max}$, which is usually taken to 20. It can be shown that H_1 is linked to $K_1(t)$ by the relationship

$$H_{fGn} = H_1 \sim \frac{\log\left(K_1(\tau)\right)}{\log(\tau)}. \tag{14}$$

We use the Matlab function genhurst to estimate the Hurst exponent, for which the residuals are to be forced to be fBm by computing their cumulative sum.

From now on, we will call $\hat{H}_{fGn}$ the estimated H_{fGn}, using either the WhiE or GHE. Similarly, we define $\hat{H}_{N_r}$, the estimated H_{N_r} as well as $\hat{H}_{res}$ the global Hurst parameter estimated from the residuals.

2.2.5. Choice of the noise model using information criteria

The noise model of the range measurements in a simulated framework is fixed in advance. The situation is slightly more complicated with real data. We mentioned for the specific case of TLS observations that the noise could be an fGn or a damped fBm, eventually overlapped with WN: these assumptions need to be validated prior to general conclusions about the correlation structure.

Although improvable in the case of correlations, we perform the validation of the noise model using both the Akaike information criterion (AIC) and Bayesian information criterion (BIC), following He et al. (2019). Such criteria provide a first indication about the best noise structure. We chose to test two competitive noise models: the power law/WN model and the Matérn model.

2.3. Summary of the methodology

The following five steps are performed to extract the correlation structure of the TLS range residuals:

- Using simulated or real observations from a plane
- Estimation of the parameters of a plane using the GHM framework
- Computation of the residuals in Cartesian co-ordinates
- Backwards transformation to obtain the range residuals
- Estimation of the Hurst or smoothness parameter with WhiE and/or GHE using a global model

In the case of simulated observations, a validation of the Hurst parameter estimation can be added by computing the root mean squared error regarding the true value of the parameter; the validation for real data is performed using the information criterion approaches. The methodology applied is summarised in Fig 4.
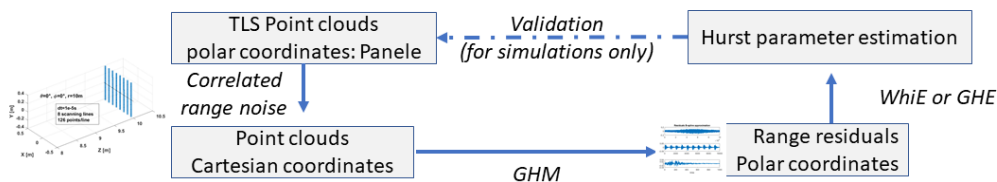

Fig 4: Summary of the methodology proposed to estimate the correlation parameter from the range residuals of a non-linear GHM plane parameter estimation.

3. Calculations: simulations

In this section, we apply the methodology proposed. We aim:

- to determine the extent to which a trustworthy estimation of the correlation parameter can be performed, depending on the scanning configuration or scanner settings, and
- to translate this knowledge to real data analysis in Section 4 so that a deeper understanding of the results is possible.

We recall that the results of the plane fitting, as defined in Eq. (1), are not the topic of this contribution. We focus on the determination of the correlation structure of the raw observations from the residuals of the fitting, i.e. the Hurst parameter or the smoothness.

We will describe the simulated observations in the first subsection. A second part is devoted to the ratio used to interpret the results, which are presented in the third part. We will provide some indications about which configuration should be chosen for optimal results from real data analysis.

3.1. Simulation of the observations

We simulate TLS raw observations from a plane of size 1 x 1 m. We further consider that the simulated plane is aligned centrically at the height of the tilting axis of the TLS. This ensures a so-called reference configuration. The arrangement of the plane can be changed freely to simulate different set-ups. In

such cases, the points may no longer lie on a rectangular grid, similar to what would happen in a real case. As in Eq. (1), we call d the distance between the origin of the TLS and the centre of the panel.

We check that the number of observations n_{obs} is at least 400 for each point cloud simulated to ensure a trustworthy estimation of the Hurst parameter.

The scanning rate, Δt , is defined as the temporal spacing between two consecutively emitted signals. It is linked with the maximum number of pixel/360° regarding the scanning time.

The polar co-ordinates simulated are noised component wise. The standard deviations for range and angle are chosen to correspond to the TLS used in Section 4.

- We add a Gaussian noise with a standard deviation of 7e^{-5} rad – generated with the Matlab function randn – to the vertical and horizontal angles, respectively.
- We add an fGn or a combination of fGn and WN to the range, as mentioned in Section 2. We choose a standard deviation for N_r equal to 0.25e^{-3} m: according to the manufacturer's specifications, this value is specified as a 1 sigma noise corresponding to a distance to the object of d = 10 m, a white percentage of the object of 37 % and a data rate of 136.719 pixel/s by the TLS. These parameters are suitable for a lot of applications in TLS and close to the choice of the practical measurements in Section 4. We note that increasing the standard deviation (i.e. increasing the distance or changing the colour to black) will have a favourable effect on the Hurst estimation: the higher the standard deviation of the range noise, the less WN coming from the angles and functional misspecification will be present in the range residuals.

We put the stress on the fact that the simulations should validate the extent to which the noise structure of the residuals of the adjustment corresponds to the one of the raw observations in a *global* model: we do not aim to estimate the exact value of H_{fGn} .

3.2. Settings

Range noise vector

We generated noise vectors with three different Hurst exponents: 0.7, 0.8 and 0.9, the latter corresponding nearly to FN. A WN with $R_{WN} = [0, 0.2, 0.5]$ is added to the fGn:

- $R_{WN} = 0$ corresponds only to an fGn. This case allows us to identify the extent to which WN can be found in the residuals without correspondence to any WN in the observations. This may occur, for example, for a suboptimal scanning configuration and is detected by an underestimation of the $\hat{H}_{N_r}$ regarding H_{N_r} .

- We recall that if $R_{WN} \neq 0$, $H_{N_r} < H_{fGn}$. In this contribution, we only estimate H_{N_r} .

For the sake of simplicity and without a lack of generality due to the self-similarity property of the fGn, we generate one noise vector of size n_{obs} for the all the observations simulated. We use a MC approach because the Hurst parameter may not correspond exactly to the true value and simulate 2000 different noise vectors for each scenario.

Scanning configuration

The TLS point clouds corresponding to different configurations are generated to analyse the impact of the scanning configuration on the estimation of the correlation parameter. The orientation of the plane is changed by acting on its elevation El and azimuth Az. Because of the symmetry of the problem, we only vary the azimuth in steps of 5° between 0° (the so-called "reference configuration") and 30°.

Data rate

The data rate is varied to mimic a real data point cloud. Two values are considered: $\Delta t = \left[1.5e^{-5}, 1.8e^{-6} \right]$, which corresponds approximately to a setting "premium, high" or "preview, normal" for the first value, and "superhigh, normal" or "premium, extremely high" for the second one (see Appendix 1 for more details on the settings). The first value is chosen as the "worthiest" case.

Distance

The distance from the TLS to the plane affects the number of points per scanning line, i.e. the temporal distance between the emitted and the received signal. Increasing the distance for a given scanning rate leads to a more challenging LS estimation, which may affect the residuals by introducing additional WN. However, in a real case, this is linked with an increase of the standard deviation of the range, so that this effect should be damped. We consider two distances: 10 and 20 m.

Table 1 summarised the settings chosen for the simulations.

Table 1: Simulation settings

El [°]	Az [°]	H_{fGn}	R_{WN}	Δt [s]	d [m]
[0]	[0 − 50°]	0.7, 0.8, 0.9	0, 0.2, 0.5	$1.5e^{-5}, 1.8e^{-6} s^{-1}$	10, 20

3.3. Evaluation of the results

We estimate the parameters for each simulated plane using a non-linear GHM, following Section 2.1. The range residuals of the approximation are transformed into polar co-ordinates and the Hurst parameter is estimated using a WhiE or a GHE. The AIC and BIC are computed to ensure that the power law assumption is valid. We fix the range parameter α to 0.001 for the WhiE and following Stein (1999) in order to force the estimator to estimate the slope of the psd at a high frequency. We compare the results with the one given with the GHE with $\tau_{max} = 20$. The Hurst parameter of the simulated noise may be a combination of WN and fGn.

We make use of MC simulations to assess the trustworthiness with which the global Hurst parameter is estimated. We compute $\hat{H}_{res}$ from the residuals for each run and define the following ratio, which we express in %:

$$R_{N_r} = 100 \frac{\hat{H}_{res} - \hat{H}_{N_r}}{\hat{H}_{N_r}} \tag{15}$$

We compute this ratio from the residuals of a plane estimation using a GHM without prior stochastic information. This approximation is intentional: integrating VCM of the raw polar observations in a LS adjustment necessitates the use of the propagation law for the transformation in Cartesian co-

ordinates. These mathematical correlations lead to a fully populated VCM, which increase the computational effort significantly: this is not our aim here as the estimation should be performable by a TLS practitioner.

This ratio allows one to draw conclusions about the trustworthiness of the estimator without being affected by the slight discrepancy between the Hurst exponent simulated regarding the true value. We compute R_{N_r} for each MC run and, finally, its mean over all runs called $R_{N_r,MC}$.

We consider a ratio below 2 % as an acceptable estimation of the Hurst parameter. For $H_{N_r} = 0.8$, this corresponds to $\hat{H}_{N_r} \approx 0.81$.

3.4. Results

The results are divided into two parts: the influence of the scanning configuration by means of the tilt angle, followed by the influence of the scanning settings and the distance. We focus on the performance of both estimators.

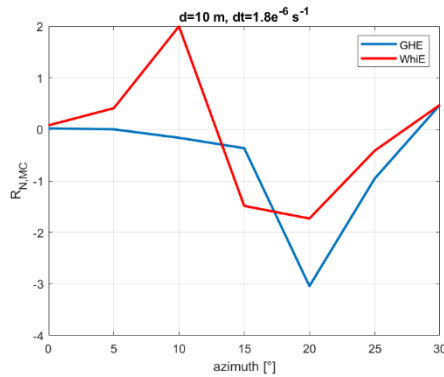#### 3.4.1. Varying the tilt angle of the plane

Fig. 5: $R_{N_r,MC}$ computed with the GHE and the WhiE by varying the azimuth from 0 to 30°. The distance simulated is 10 m and the scanning rate $1.8e^{-6}\ s^{-1}$. The noise parameters were $H_{fGn} = 0.7$ and $R_{WN} = 0$.

We see from Figure 5 that the optimal configuration is – as intuitively expected – $[El, Az] = [0,0]°$. In that case, both estimators perform ideally and $R_{N_r,MC}$ is close to zero. When the azimuth angle increases, the WhiE is less trustworthy than the GHE, for which $R_{N_r,MC}$ does not exceed -1 %, even for $[El, Az] = [0,15]°$, which should be compared to the 2 % obtained with WhiE for the same configuration. Interestingly, increasing the azimuth further from 20° is not synonymous with an increase of $R_{N_r,MC}$. We can conclude that the scanning configuration does not greatly impact the Hurst parameter estimation; the maximum difference reaches only 3 %. Other settings were tested without changing this conclusion. They are not presented for the sake of brevity.

3.4.2. Varying the distance and the scanning rate

The optimal setting $[El, Az] = [0,0]°$ in a real scenario is often linked with an optimal angle of incidence of the TLS laser beam to the object. Since we want to be able to interpret the results of a real experiment in the light of the simulations, we will only consider cases where $[El, Az] = [0,5]°$.

Following Figure 5, this slight discrepancy regarding the optimal configuration does not lead to a significant increase of $R_{N_r,MC}$.

Table 2: $R_{N_r,MC}$ for the scanning configuration $[El, Az] = [0,5]°$ by varying $[H_{fGn}, R_{WN}]$, as described in Table 1. The results are given for both WhiE and GHE. The simulated distances TLS-plane are 10 and 20m. A total of 2000 MC simulations were performed to compute the mean values of the ratio.

	H_{fGn} /WN	0	0.2	0.5
d=10 / 20 m				
$\Delta t = 1.8e^{-6}s^{-1}$				
WhiE	0.6	-1.40/-1.71	-2.46/2.82	1.73/-2.77
GHE		-2.11/-1.95	-2.88/-0.95	7.57/-2.26
WhiE	0.7	0.41/-0.01	0.99/-0.66	-0.52/-2.45
GHE		-0.00/-1.81	0.29/1.72	1.39/-0.29
WhiE	0.8	-3.04/-0.06	0.31/-1.77	-6.17/-5.10
GHE		2.52/0.95	1.22/-0.29	-1.95/-1.07
$\Delta t = 3.5e^{-5}$	Identical results			

The ratios $R_{N_r,MC}$ obtained by varying the scanning rate and the distance are presented in Table 2. The adequation of the 2000 MC values with the Gaussian distribution was checked by plotting the corresponding histogram and performing a one-sample Kolmogorov-Smirnov test.

We, firstly, note that the increase of the scanning rate only changed the results in a sub-decimal range: these differences can be considered insignificant; they are not presented for the sake of brevity. We, therefore, do not expect the scanning rate chosen for a real case to affect the estimation of the Hurst parameter greatly. Please note that this result does not mean that the correlation *structure* will be independent of the scanning setting, as shown in Section 4.

The two estimators do not perform similarly when the true Hurst exponent and the amount of WN is varied. The GHE has a better performance for estimating the global Hurst parameter than the WhiE as H_{fGn} increases, provided that $R_{WN} < 0.7$. For $R_{WN} = 0$, we note that a Hurst exponent of 0.7 is favourable to the GHE, as $R_{N_r,MC} \approx 0$. Globally, the GHE is trustworthy: $R_{N_r,MC}$ does not exceed 3 %, except for $R_{WN} = 0.5$ and $H_{fGn} = 0.6$. In those cases, the GHE performs poorly, although increasing the distance to 20 m decreased the $R_{N_r,MC}$ from 7 to 2 %. This effect is linked with the fact that the GHE acts on the time domain: when the Hurst exponent is close to 0.5, the GHE is not able to distinguish WN and fGn successfully. In that particular case, the WhiE – a frequency-based estimator – has an opposite performance: for $R_{WN} = 0.5$, $R_{N_r,MC}$ increases with the Hurst exponent from 1.7 to 6 %. Increasing the distance decreases the $R_{N_r,MC}$ to 5 % for the less optimal configuration $H_{fGn} = 0.8$ and $R_{WN} = 0.5$. Consequently, the WhiE has a better performance than the GHE for low $H_{fGn} = 0.6$ and $R_{WN} = 0.5$. As the Hurst parameter increases, $R_{N_r,MC}$ stays under 2 % for both estimator and $R_{WN} \leq 0.2$, although we note a slightly better performance with the GHE for $H_{fGn} = 0.8$.

We notice further that increasing the distance does not affect the global Hurst parameter estimation unfavourably, provided that enough MC samples are generated as the standard deviation of the ratio

increased from 3 % to approximately 6 % with the distance. This necessary increase of samples may, however, not be given in a real case.

3.5. Summary of the simulations

We summarised the results of the simulations as follows:

- The optimal scanning configuration corresponds to $[El, Az] = [0, 0-5]°$, independent of the estimator chosen.
- Neither the distance nor the scanning rate affects the trustworthiness of the global Hurst parameter determination. This result depends, however, on the number of samples generated: it may not be transferable in a real case analysis.
- The GHE has a better performance than the WhiE for a low level of WN contamination and a true $H_{fGn} \geq 0.7$. The contrary holds true for a high level of WN and $H_{fGn} \leq 0.7$, where the WhiE should be preferred.

Globally, the two estimators perform similarly and the ratio of difference stays below 2 %, except for the challenging cases mentioned previously. We recall that 2 % implies a difference of less than 0.015 for a Hurst parameter of 0.7. Considering that one never knows the true correlation parameter in a real case, the estimation can be considered trustworthy. The difference is even below 1 % for some Hurst parameters. It is, thus, possible to extract the global Hurst parameter from the residuals of an LS adjustment.

4. Results of real case analysis

Having determined the optimal settings, we are able to interpret the results of a real data analysis with higher confidence. After a short description of the indoor and outdoor experiments, we extract the correlation parameter for the different settings using the methodology described previously.

4.1. Description of the experiments

A planar ALUCORE® panel with the dimension of 1 x 1m was used in the measurements performed. The colour of the panel is 43 % white. The planarity of the panel was determined to be ≤ 0.5 mm. The measurements were carried out by a laser tracker (Leica AT960) and a corner cube reflector. The accuracy is specified with Ux,y,z = +/-15 µm + 6 µm/m as a maximum permissible error (Hexagon 2015). The 3D object capturing was performed by the Z+F IMAGER 5016 (Zoller & Fröhlich GmbH, Wangen im Allgäu, Germany). Outliers were eliminated using the methodology presented in Appendix 2.

4.1.1. Indoor experiment

The measurements took place at the measuring laboratory of the Geodetic Institute in Hannover.

The panel was observed from a total of five standpoints with distances between 2 and 50 m. We ensured that the incidence angles for the reference configuration were nearly parallel to the face normal of a panel. It should be noted that over-radiation occurs when using ALUCORE® panels and the Z+F IMAGER 5016 at an optimal angle of incidence. A visualisation of a point cloud clearly shows this unwanted effect in Figure 6. These artefacts in the distance measurement lead to outliers and strongly affect the adjustment of the plane: similar to a snowball effect, the residuals – from which we wish to determine the correlation parameters – show suspicious variance increases that are likely to mislead the WhiE or GHE. We turned the plane slightly from 5° in azimuth to avoid that drawback. Simulations have highlighted that this setting did not impact the estimation of the Hurst parameter more than 2 % at a distance of 10 m. Thus, no additional sensors were used to control the alignment.
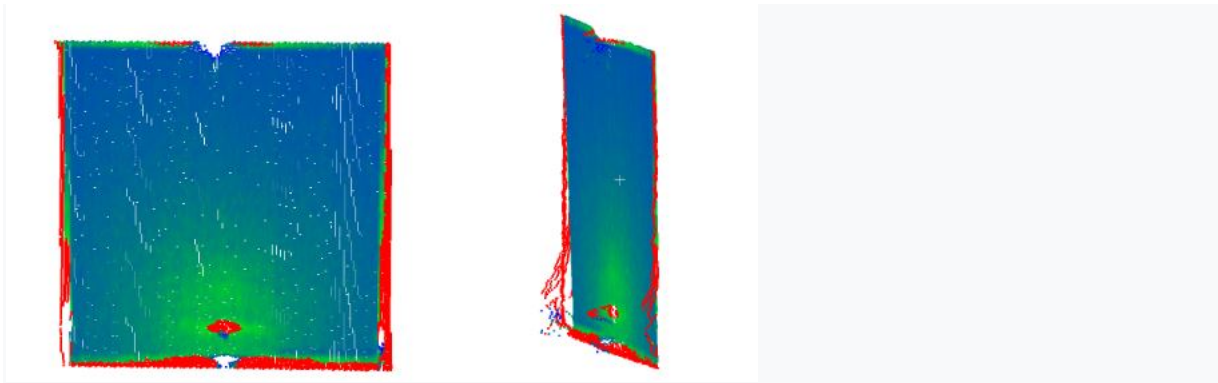
Fig 6: Left: example of strong reflection occurring in the middle of the panel (57 %) at a distance of 20 m leading to a variance increase in the residuals. Right: outliers due to edge effects have to be eliminated prior to the correlation analysis (Jüngerink 2019)

The measurements were made indoor in a laboratory. We do not expect a turbulent variation of refractive index due, for example, to temperature fluctuations that affects the results. The data acquisition was performed under fixed settings with the resolution setting "superhigh" and "high" and the quality setting "high".

The ratio of the standard deviation of the angle to the range is below 100, according to the manufacturer's datasheet, so that the WN of the range residuals cannot be assessed to the angle noise. The adjustments were computed without stochastic information, i.e. the VCM of the raw observations was the identity matrix, following the simulations.

4.1.2. Outdoor data set

Fig 7: Outdoor experiment, Herrenhäuser Allee, Hannover, Germany. The panel is 43 % white. The ranges were varied between 75 and 200 m (Jüngerink 2019).

The measurements from the panel (Fig. 7, left) were carried out on 16 May 2019 at EDM calibration base in Hannover, Germany. The laser scanner was set up at the distances of 75, 125 and 200 m towards the ALUCORE® panel (see Fig. 7). Although shorter distances were measured during this experiment (20 and 50 m), they are not considered in this contribution for the sake of conciseness: such distances have already been investigated with indoor measurements, leading to similar results.

A Leica TS30 was used to set up the reference constellation of the panel regarding the TLS. The distance of both sides of the panel was measured here to correct the arrangement of the panel. The weather was dry but cloudy. The temperature rose steadily from 16.5 to 21 °C within the five-hour measurement period with smaller fluctuations in the range of 0.5 °C. The slow rise in temperature should not have any negative effects on the distance measurement, as the temperature has been updated for each standpoint of the laser scanner. The temperature-related correction was performed by the laser scanner. This does not mean that turbulent variations of the air's refractive index will not

affect the correlation structure (Wheelon 2001), as an atmospheric noise is more likely to be present as the distance between the scanner and the object increases.

Only the part of the panel selected was scanned to reduce the measurement duration and the amount of data. The laser scanning was performed with the resolutions middle, high und superhigh, and the quality was chosen as normal. The resolution middle is not considered in this contribution as not enough points were available for a trustworthy Hurst estimation.

The correct Hurst parameter is unknown in real cases and without further physical investigations can only be estimated. We apply our methodology and adjust the point cloud with a non-linear Gauss-Helmert model. In a second step, we estimate the correlation parameters and choose the correlation model by means of information criteria. We adopt and compare two approaches, GHE and WhiE, because of the property of these estimators; any difference provides additional information about the correlation structure or dominant noise sources. We do not investigate the fraction of WN explicitly as we search for a global model. We take for granted that the physically related Hurst parameter (exemplarily FN or atmospheric noise or a combination of both) may be slightly underestimated in the presence of WN. Physical noise investigations necessitate an adequate filtering and are left to further dedicated contributions.

We chose to estimate the Hurst parameter "batch-wise", i.e. we cut the residuals into batches of 1000 observations and compute the mean of the Hurst parameter over all batches. Following Sykulski et al. (2019), we, thus, ensure a trustworthy estimation of the *mean* global correlation parameter.

4.2. Results

The results for the indoor and outdoor measurements are presented in Table 3 and 4, respectively. We highlight with cases when the Matérn model was the most optimal model from the AIC/BIC in red, otherwise a power law model, i.e. a fGn, was considered. We adopted the 3σ threshold for the outlier elimination.

4.2.1. Indoor experiment

Table 3 gives the mean values for all batches of 1000 observations over the whole residual vector. We decreased the batch size to 500 observations for the setting "high" from 20 m to ensure that at least three batches could be made: the estimation of the parameter may be slightly less reliable for these cases (see Section 3). The same approach was used for indoor and outdoor measurements.

Table 3: Results of the indoor experiment. The setting is varied from ultrahigh to high for distances from 2 to 50 m. The WhiE and GHE are compared. The values in red correspond to a Matérn model being considered as optimal.

2m	Ultrahigh	Superhigh	high
WhiE	0.83 (0.04)	0.81 (0.04)	0.72 (0.02)
GHE	*0.68 (0.05)*	*0.78 (0.08)*	*0.74 (0.02)*
5m			
WhiE	0.92 (0.03) Cut-off 0.3	0.87 (0.16) Cut-off 0.5	0.59 (0.04)
GHE	*0.63 (0.11)*	*0.63 (0.08)*	*0.56 (0.04)*
10m			
	0.92 (0.02)	0.85 (0.15)	0.66 (0.06)
	0.72 (0.02)	0.57 (0.03)	0.58 (0.03)
20m		500 obs	500 obs

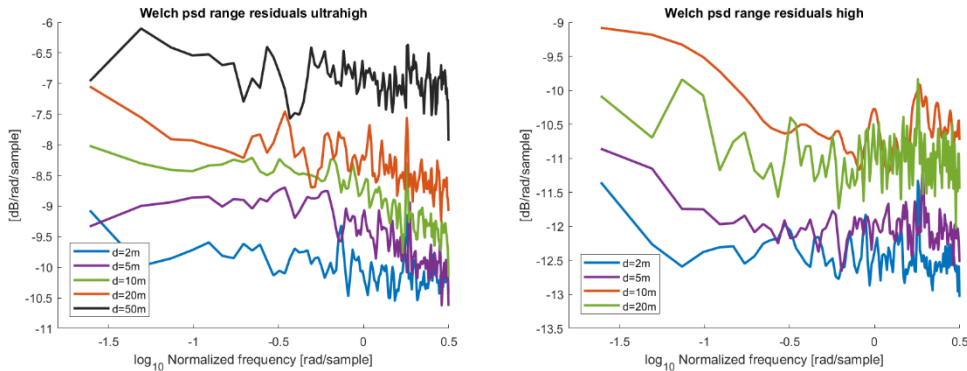|			
	0.80 (0.03)	0.71 (0.03)	0.61 (0.02)
	0.71 (0.04)	*0.68 (0.03)*	*0.61 (0.02)*
50m		500 obs	300 obs
	0.78 (0.1)	0.62 (0.06)	0.60 (0.04)
	0.62 (0.04)	*0.58 (0.08)*	*0.55 (0.05)*

Fig. 8: psd of the residuals, indoor experiment. Left: by varying the distance between the TLS and the panel for a distance from 2 m (blue line) to 50 m (black line) for the setting ultrahigh. Right: for the setting high. The x-axis represents the log10 for the normalized frequency. The psd is obtained from the Matlab function pwelch and is given in dB/rad/samples.

Varying the distance for a given setting

The distance between the TLS and the panels for the indoor experiment was varied from 2, 5, 10 and 20 up to 50 m. The settings ultrahigh and high lead to similar behaviours of the Hurst parameter versus distance: the parameter slightly increases up to the range of 10 m and decreases by remaining at a high level over 0.6 for the range 50 m. At that range, the standard deviation of the Hurst parameter reaches nearly 0.08 due to the low number of observations per batch combined with the decrease of the number of batches.

Settings ultrahigh and superhigh

The noise at short range is a combination of WN and FN, which can be seen in Figure 8 (left, blue line). The FN comes typically from the laser (Van der Ziel 1970) and is particularly visible in the low frequency domain of the spectrum. The discrepancy between the estimation from WhiE and GHE (0.83 versus 0.68, respectively) comes from that particularity since the GHE focuses on the estimation in the middle of the spectrum, where WN predominates.

The noise model switches clearly at distance 5 and 10 m from a fGn, i.e. a combination of power law noises, to a Matérn noise. Both the AIC and BIC confirm what the psd highlights (magenta line): a damping of the power law noise at low frequencies. In such cases, the GHE is not adequate as the damping at low frequency affects the slope and, thus, decreases the Hurst parameter estimated. In the point of view of the authors, this model change is due to the increase of observations, leading to a flattening of the spectrum: a pressure to grow combined with some drag or resistance on that growth. This effect is known in GNSS co-ordinates time series analysis (see He et al. 2019). Without the high increase of FN at low frequencies, the psd at the range of 2 m would also have flattened similarly. The Hurst parameter found for the ranges 5 and 10 m is a further indication of the presence of FN combined with WN and potentially an increasing atmospheric noise.

The psd at a distance of 20 m has a linear slope: the noise is an fGn with a Hurst parameter of approximately 0.7 – 0.75, the value of the superhigh setting having been computed for a lower number

of batches than for ultrahigh. The decrease of the FN can be interpreted as a simultaneous increase of the atmospheric noise, which is expected to have a Hurst parameter of approximately 0.8 (Wheelon 2001, Ch. 6): the slightly smaller values come from the WN component, which biased the Hurst parameter from the atmospheric noise towards 0.5. At the distance of 50 m, we still clearly identify a similar slope between an angular frequency of 0.3 and 1.7 rad/samples to lower distances. However, a decrease of the power of the low frequency component decreases the Hurst parameter artificially, particularly with the GHE: a value close to 0.6 is obtained for the ultrahigh setting compared to 0.65 and 0.75 with the WhiE. This lower power from the simulations, shown in Figure 8 (black line), could come from the challenging approximation due to a lower data quality.

Setting high

No Matérn model can be identified as optimal for the setting high. This is a strong confirmation that this model is best fitted to the residuals as the temporal distance between emitted and received signal is short. We identify a combination of WN and FN for the distance of 2 m (in Fig. 8, right), leading to a Hurst parameter of 0.72.

Summary

More generally, the amount of WN is higher for the setting high than for ultrahigh and superhigh: this effect can be linked to the data rate, i.e. the correlation between the measurements is smaller than for higher number of observations. If the WN leads to a flattening of the spectrum at high frequencies, at low frequencies, we clearly identify a power law noise with a slope close to the one found for the other setting. As this slope is becoming more pronounced as the distance increases (green and red line, Fig 8 right), we interpret it as an atmospheric noise: the power of the latter increases as the distance increases. However, one should not underestimate that the determination of the Hurst parameter is made challenging due to the additional sinus-like pattern of the psd, which needs further investigation which is beyond the scope of this present paper.

Varying the data rate for a given distance

If we consider the results of Table 3 for a given distance (one line), by varying the resolution from ultrahigh to high, we observe a decrease of the correlation parameter between 0.2 for the 20 m case and 0.1 for the 2 m case for an estimation with the WhiE. These variations are less pronounced with GHE and stay within a range of 0.1, except for the 10 m case. We interpret this effect as coming from the amount of WN, which leads to an underestimation of the slope, following the results of simulations. This holds particularly true in the presence of higher low frequencies because the GHE is less sensible to that part of the spectrum than the WhiE. There is, thus, a decrease of the correlation with the decrease of the scanning rate for a given quality (here high). This effect is expected when dealing with temporal correlations. However, it does not mean that the observations become fully uncorrelated: long-range correlations corresponding to an fGn with a Hurst parameter of 0.6 are still detected for the setting "high" at a distance of 50 m, which is far from negligible. This leads to a VCM that stays fully populated, and, therefore, its inverse impacts the test statistics for detection deformation and their probability distribution (Kermarrec et al. 2020)

4.2.2. Outdoor experiment

The results for the outdoor experiment are presented in Table 4.

Table 4: Results of the outdoor experiment. The resolution is varied from superhigh to high for distances from 75 to 200 m. The WhiE and GHE are compared.

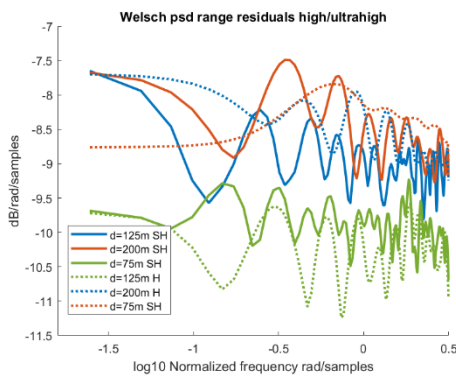| 75m	Superhigh 400 obs	High 300 obs 1 batch only
	0.68 (0.007)	0.62
	0.66 (0.02)	0.63
125m	400 obs	125 obs 1 batch only
	0.72	0.72
	0.71	0.66
200m	300 obs	100 obs 1 batch only
	0.79	0.68
	0.73	0.63

Fig. 9: psd of the residuals, outdoor experiment. The distance is varied from 75 m (green line) to 200 m (red line) for the setting superhigh. The x-axis represents the log10 for the normalized frequency. The psd is obtained from the Matlab function pwelch and is given in dB/rad/samples.

The outdoor observations confirmed the fact that the atmospheric noise predominates with increasing distance: although the time between emitted and receiver signal increases with the distance, the Hurst parameter for the resolution high remains constant or even increases, reaching 0.68 with the WhiE for the range of 200 m. It remains similar for all three distances, independent of the setting used. This finding is coherent with the constant slopes found in the psd (see Fig. 9). The same behaviour is observed for the resolution superhigh. We justify this finding by the correlations due to turbulent variations of the refractivity index, which may have been strong due to the atmospheric conditions (dry). These correlations are more pronounced regarding the other noise (FN, WN) as the distance increases. Clearly, more investigations are needed to validate this phenomenon, combined with a strategy to extract that noise from the WN and FN of the laser. One should, furthermore, keep in mind that the estimation is made challenging due to the presence of outliers and the small number of observations per batch. Thus, the importance of using an unbiased estimator, such as the WhiE, for small samples is confirmed.

The high correspondence between the values obtained with GHE and WhiE is a strong indication that the noise corresponds to an fGn and that no additional WN at high frequencies or higher order power law noise at low frequencies are present. The sinus pattern visible in the psd (Fig. 9) does not affect the determination of the Hurst parameter with the WhiE in the frequency domain: the results obtained with GHE, which operates in the time domain, are similar, which should be expected from the simulations.

4.3. Summary of the real experiments

The following conclusions can be drawn from the real data analysis both indoor and outdoor:

- The combination of two estimators for the Hurst parameter gives a strong indication about the dominant noise sources.
- The noise is a combination of WN and FN at short distance.
- We note an increase of WN with a decrease of the data rate.
- The Matérn model is more adequate than the fGn as time between the emitted and received signal decreases, i.e. at short distance and for a high data rate. In such cases, the slope of the psd at high frequency highlights the presence of a strong FN.
- We could identify an increase of the atmospheric noise with the distance, particularly visible in the outdoor experiment. This noise balances the increase of WN for the resolution high and additional noise due to the potential functional misspecification coming from a lower data quality.

Based on these conclusions, we propose either

(i) to estimate the global Hurst parameter from the residuals of an approximation for an accurate correlation structure (Matérn model or fGn) or

(ii) to fix the Hurst parameter following the previous results to a mean value of 0.7. Such a simplification does not account for slight variations occurring with the distance or the atmospheric condition and should remain a first approximation. The variance model can be taken from the intensity model of Wujanz et al. (2017) or from the manufacturer's datasheet.

5. Conclusions

This study proposes and validates a strategy to estimate the correlation parameter of the TLS distance measurement. Consequently, we used a parametric modelling of the point clouds and analysed the residuals of the LS approximation. We also chose to model the correlation structure of the range as an fGn: this noise has the properties of being

- physically plausible due to the power law dependency of its psd,
- simple, as only one parameter, called the Hurst exponent, must be estimated from the residuals, and
- linked, to some extent, to the Matérn process, as a damped version of an fBm at low frequencies.

We proposed two estimators for the Hurst exponent in this contribution: one acting in the time domain and the other called the debiased WhiE, recently developed and known to be trustworthy for small samples. Although the latter was developed for the parameter of the Matérn model, we showed that it can be used to estimate the parameters of the fGn, provided that a small cut-off frequency is chosen.

In a simulated framework, we generated TLS point clouds from a plane of size 1 x 1 m, (i) scanned at a given distance (ii) with different data rates, and (iii) under different scanning configurations. We noised the range observations by adding an fGn generated with a known Hurst parameter and additionally combined with WN. We estimated the parameters of the plane from the simulated point clouds with a non-linear LS approximation known as the Gauss-Helmert model. Rather than concentrating on the parameters of the plane, we focused on the correct estimation of the Hurst exponent from the residuals of the fitting. For the sake of simplicity of our proposal and its wider use in the community, we modelled the noise as global, i.e. we did not separate the WN from the correlated noise specifically. The simulations highlighted that, depending on the level of WN and the distance from the TLS to the plane, the GHE or WhiE have a different performance. This finding was particularly true for suboptimal

scanning configurations and allowed us to define optimal configurations and settings: the difference between a true and estimated Hurst exponent was beyond 2 %. The difference in the behaviour of the two estimators gave some indication concerning the dominant noise source (WN of fGn).

The valuable conclusions of the simulations were used to analyse results obtained with real data from indoor and outdoor experiments, recorded with different resolution from high to ultrahigh. We estimated the noise parameters – modelled as a fGn or Matérn model – and analysed the psd of the residuals of the plane approximation. Combining these results, we showed that the correlated noise at very short distances can be described as a combination of FN and WN. As the distance increases, the noise structure evolved from a combination of two noises to a Matérn model with a Hurst parameter close to the FN for the resolutions ultrahigh and high. We interpreted this finding by the high number of observations or as a transition regime. As the distance was further increased to 20 m, the noise became an fGn once more, with a lower Hurst exponent than at a short range. This highlighted the shift in the predominant noise, driven by the atmospheric correlation as the distance increases. The outdoor experiment confirmed this result, although the estimation of the Hurst parameter was made challenging because of the distance and the data quality. Using the resolution high, a high level of long-range dependency could be found. This finding goes, to some extent, against the sometimes assumed lack of correlations for lower data rate and high distance object TLS.

Our results pave the way for the determination of a correlation model for TLS range observations based on the simple fGn without having to estimate the corresponding parameters empirically. In this contribution, we proposed a global model for estimating the correlation parameter: we did not estimate the variance of the WN or fGn separately. The accurate and physically based modelling of the atmospheric correlations and their quantification regarding other noise sources, such as FN or WN, remain the topic of a specific contribution. Although tested for the specific case of a Z+F IMAGER 2016F, our method is easily transferable to all kinds of laser scanners, including time of flight laser. It is, moreover, general enough to be applied for the study of the correlation structure of TLS angles.

Appendix 1: Settings

In this appendix, we shortly summarise the settings of a Z+F IMAGER 2016F, taken from the manufacturer's datasheet.

Table 4: resolution settings for a Z+F IMAGER 2016F. The green cells correspond to typical settings for a 360° scan

Quality

Resolution	Pixel/360°	Time between measurements (s)			
		less	normal	high	premium
preview	1250	-	1.5E-05	-	-
low	2500	7.3E-06	1.5E-05	2.9E-05	-
middle	5000	3.7E-06	7.3E-06	1.5E-05	2.9E-05
high	10000	1.8E-06	3.7E-06	7.3E-06	1.5E-05
super high	20000	9.1E-07	1.8E-06	3.7E-06	7.3E-06
ultra-high	40000	-	9.1E-07	1.8E-06	3.7E-06

| extremely high | 80000 | - | 9.1E-07 | 9.1E-07 | 1.8E-06 |

max. number of pixels per profile	200000

Appendix 2: Treatment of outliers

In this contribution, outliers were eliminated using the concept of the least median of squared (LMS) residuals, introduced by Rousseeuw (1984). In this appendix, we propose to briefly explain the corresponding procedure. The advantage of the LMS is the high breakdown point of up to 50 %, i.e., the largest possible value. The objective function of the LMS reads

$$\min_{} \operatorname*{median}_{i} v_i^2. \tag{16}$$

To derive the robust solution, m u-tuples or subsamples are drawn from the set of all n observations. Each subsample contains the minimal number of observations required u to get an exact solution. In the case of the plane, $u = 3$ is the number of points required. Referring to these m subsample solutions, the residuals of the whole set of observations are derived. The solution, which fulfils the LMS objective function in Eq. (16), is the robust solution of the LMS. Outliers are identified by $|v_i| > k\hat{\sigma}_{\text{MAD}}$, where k is a sensibly selected threshold value and $\hat{\sigma}_{\text{MAD}} = 1.483 \sqrt{\operatorname*{median}_{i} v_i^2}$ is a robust estimation of σ_0 (Rousseeuw 1984), the a priori standard deviation of the observations.

Since the results are derived by combinatorial analysis, the disadvantage of the LMS is the computational effort if the sample size n becomes large regarding u. The total number of subsamples is (e.g. Lösler 2011) $m_{\max} = \dfrac{n!}{u!(n-u)!}$.

To reduce the computational effort, the minimal number of subsamples required can be obtained by

$$m_{\min} = \frac{\ln(1-P)}{\ln\left(1-(1-\varepsilon)^u\right)}, \tag{17}$$

where P is the probability of drawing at least one outlier-free subsample, ε is the fraction of expected outliers and m is the number of subsamples drawn (Rousseeuw and Leroy 1987, p. 198).

Bibliography

Andrews, L. C., Phillips, R.L., 2005. Laser beam propagation through random media. 2nd ed. Bellingham, Wash. : SPIE (SPIE Press monograph, PM152).

Barabási, A.L., Vicsek, T., 1991. Multifractality of self-affine fractals. Physical Review A 44 (4), 2730-2733. DOI: 10.1103/physreva.44.2730.

Bardet, J.M., Lang, G., Oppenheim, G., Philippe, A., Taqqu, M.T.S, 2003. Generators of long-range dependent processes: A survey. In P. Doukhan, G.Oppenheim, and M. S.Taqqu (Eds.),Theory and Applications of Long-Range Dependence, pp. 579–623. Boston:Birkhäuser

Beran, J., 2013. Long memory processes: probabilistic properties and statistical methods. Springer Verlag Berlin and Heidelberg GmbH & Co. K

Bos, M.S., Fernandes, R.M.S., Williams, S.D.P., 2013. Fast error analysis of continuous GNSS observations with missing data. Journal of Geodesy 87, 351-360.

Bronshtein, I.N., Muehlig, H., Musiol, G., Semendiaev, K.A., 2007. Handbook of Mathematics, fifth ed. Springer, Berlin, Heidelberg.

Di Domenico, G., Schilt, S., Thomann, P., 2010. Simple approach to the relation between laser frequency noise and laser line shape. Applied Optics 49 (25), 4801-4807. DOI: 10.1364/AO.49.004801.

Di Matteo, T., Aste, T., Dacorogna, M.M., 2003. Scaling behaviors in differently developed markets. Physica A: Statistical Mechanics and Its Applications 324 (1-2), 183-188. DOI: 10.1016/S0378-4371(02)01996-9.

Gelfand, A.E., Diggle, P., Fuentes, M., Guttorp, P., 2010. Handbook of Spatial Statistics. Chapman & Hall/CRL, Boca Raton.

GeoSiberia, 2007. International Exhibition and Scientific Congress. Novosibirsk, Russia, 25 April 2007. European Association of Geoscientists & Engineers.

Granger, C.W.J., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. Journal of Time Series Analysis 1 (1), 15-29. DOI: 10.1111/j.1467-9892.1980.tb00297.x.

He, X., Bos, M.S., Montillet, J.P., 2019. Investigation of the noise properties at low frequencies in long GNSS time series. Journal of Geodesy 93, 1271-1282. DOI: 10.1007/s00190-019-01244-y.

Holst, C., Kuhlmann, H., 2016. Challenges and present fields of action at laser scanner based deformation analysis. Journal of Applied Geodesy 10 (1), 17-25.

Idrees, M., Pradhan, B., 2018. Geostructural stability assessment of cave using rock surface discontinuity extracted from terrestrial laser scanning point cloud. Journal of Rock Mechanics and Geotechnical Engineering, 10 (3) 534–544.

Jäger, R., Müller, T., Saler, H., Schwäble, R., 2005. Klassische und robuste Ausgleichungsverfahren – Ein Leitfaden für Ausbildung und Praxis von Geodäten und Geoinformatikern. Wichmann, Heidelberg. ISBN: 3-87907-370-8

Jüngerink J., 2019. Parametrisierung des intensitäts-basierten stochastischen Modells für den Z+F IMAGER 5016. Bachelor Thesis, Leibniz Universität Hannover, Fakultät für Bauingenieurwesen und Geodäsie, Geodätisches Institut Hannover.

Kargoll, B., Omidalizarandi, M., Loth, I., Paffenholz, J-A., Alkhatib, H., 2018. An iteratively reweighted least-squares approach to adaptive robust adjustment of parameters in linear regression models with autoregressive and t-distributed deviations. Journal of Geodesy 92 (3), 271-297. DOI: 10.1007/s00190-017-1062-6.

Kauker, S., Schwieger, V., 2017. A synthetic covariance matrix for monitoring by terrestrial laser scanning. Journal of Applied Geodesy 11 (2), 77-87.

Kaufman, C. G. and Shaby, B. A., 2013. The role of the range parameter for estimation and prediction in geostatistics. Biometrika, 100,473–484.

Kermarrec, G., Schön, S., 2014. On the Ma tern covariance family: a proposal for modeling temporal correlations based on turbulence theory. Journal of Geodesy 88 (11) 1061–1079.

Kermarrec, G., 2020. On estimating the Hurst parameter from least-squares residuals. Case study: Correlated terrestrial laser scanner range noise. Mathematics 8 (5), 674. DOI: 10.3390/math8050674.

Kermarrec, G., Kargoll, B., Alkhatib, H., 2020. Deformation analysis using B-spline surface with correlated terrestrial laser scanner observations – A bridge under load. Remote Sensing 12 (5), 829. DOI: 10.3390/rs12050829.

Kermarrec, G., Neumann, I., Alkhatib, H., Schön, S., 2018. The stochastic model for Global Navigation Satellite Systems and terrestrial laser scanning observations: A proposal to account for correlations in least-squares adjustment. Journal of Applied Geodesy 13 (2).

Koch, K.R., 2010. Nurbs surface with changing shape. Allgemeine Vermessungs-Nachrichten 3, 83-89.

Koch, K.R., 2014. Outlier detection for the nonlinear Gauss Helmert model with variance components by the expectation maximization algorithm. Journal of Applied Geodesy 8 (3). DOI: 10.1515/jag-2014-0004.

Koutsoyiannis, D., 2002. The Hurst phenomenon and fractional Gaussian noise made easy. Hydrological Sciences Journal 47 (4), 573-595. DOI: 10.1080/02626660209492961.

Kupferer, S., 2005. Anwendung der Total-Least-Squares-Technik bei geodätischen Problemstellungen. Dissertation. Karlsruhe, University http://digbib.ubka.uni-karlsruhe.de/volltexte/1000003449.

Laefer, D.F., Truong-Hong, L., 2017., Toward automatic generation of 3D steel structures for building information modelling. Automation in Construction 74, 66-77. DOI: 10.1016/j.autcon.2016.11.011.

Lague, D., Brodu, N., Leroux, J. 2013 Accurate 3D comparison of complex topography with terrestrial laserscanner: Application to the Rangitikei canyon (N-Z).ISPRS J. Photogramm. Remote Sens. 82, 10–26.

Lenzmann, L., Lenzmann, E., 2004. Strenge Auswertung des nichtlinearen GaußHelmert-Modells. Allgemeine Vermessungs-Nachrichten 111, 68-73.

Lilly, J.M., Sykulski, A.M., Early, J.J., Olhede, S.C., 2017. Fractional Brownian motion, the Matérn process, and stochastic modeling of turbulent dispersion. Nonlinear Processes Geophysics 24, 481-514.

Lösler, M., 2011. Robust parameter estimation of the spatial. Helmert-transformation. Allgemeine Vermessungs-Nachrichten 118 (5).

Lösler, M., Lehmann, R., Neitzel, F., Eschelbach, C., 2020. Bias in least-squares adjustment of implicit functional models. Survey Review DOI: 10.1080/00396265.2020.1715680.

Lösler, M., 2020. Zfv to be published

Mandelbrot, B., Ness, J.W., 1968. Fractional Brownian motions, fractional noises and applications. SIAM Review 10, 422-437.

Mukupa, W., Roberts, G.W., Hancock, C.M., Al-Manasir, K., 2016. A review of the use of terrestrial laser scanning application for change detection and deformation monitoring of structures. Survey Review 49 (353), 1-18. DOI: 10.1080/00396265.2015.1133039.

Neitzel, F., 2010. Generalization of total least-squares on example of unweighted and weighted 2D similarity transformation. Journal of Geodesy 84 (12), 751-762. DOI: 10.1007/s00190-010-0408-0.

Neuner, H., Holst, C., Kuhlmann, H., 2016. Overview on Actual Modelling Strategies of Point Clouds for Deformation Monitoring. In: allgemeine vermessungs nachrichten (avn), 123 (11-12) 328 – 339.

Niemeier, W., 2002. Ausgleichungsrechnung. Eine Einführung für Studierende und Praktiker des Vermessungs- und Geoinformationswesens. Walter de Gruyter, Berlin, New York.

Paschotta, R., 2008 Encyclopedia of Laser Physics and Technology. Wiley-VCH, Weinheim.

Pelzer, H., 1971. Zur Analyse geodätischer Deformationsmessungen, Dtsch. Geodät. Komm., Ser. C. 164, p. 86

Pfeifer, N., Briese, C., 2007. Laser scanning – principles and applications: GeoSiberia 2007 – International Exhibition and Scientific Congress. Novosibirsk, Russia, 25 April. European Association of Geoscientists & Engineers.

Rousseeuw, P.J., 1984. Least median of squares regression. Journal of the American Statistical Association 79 (388), 871-880. DOI: 10.1080/01621459.1984.10477105.

Rousseeuw, P.J., Leroy, A.M., 1987. Robust Regression and Outlier Detection. J. Wiley, New York.

Rueger, J.M., 1996. Electronic Distance Measurement. Springer-Verlag, Berlin, Heidelberg.

Schmitz, B., Holst, C., Medic, T.D., Lichti, D., Kuhlmann, H., 2019. How to efficiently determine the range precision of 3D terrestrial laser scanners. Sensors 19 (6), 1466. DOI: 10.3390/s19061466.

Schmitz, B., Kuhlmann, H., Holst, C., 2020. Investigating the resolution capability of terrestrial laser scanners and its impact on the effective number of measurements. ISPRS Journal of Photogrammetry and Remote Sensing 159, 41-52.

Soudarissanane, S., Lindenbergh, R., Menenti, M., Teunissen, P., 2011. Scanning geometry: Influencing factor on the quality of terrestrial laser scanning points. ISPRS Journal of Photogrammetry and Remote Sensing 66, 389-399.

Stein, M.L., 1999, Interpolation of spatial data: some theory for kriging. Bautechnik 94 (2).

Stéphan, G.M., Tam, T.T., Blin, S., Besnard, P., Têtu, M., 2005. Laser line shape and spectral density of frequency noise. Physical Review A 71 (4). DOI: 10.1103/PhysRevA.71.043809.

Stoev, S., 2020. fftfgn (https://www.mathworks.com/matlabcentral/fileexchange/5702-fftfgn), MATLAB Central File Exchange. Retrieved August 10, 2020.

Suchocki, C., 2020. Comparison of time-of-flight and phase-shift TLS intensity data for the diagnostics measurements of buildings. Materials (Basel) 13 (2). DOI: 10.3390/ma13020353.

Sykulski, A.M., Olhede, S.C., Guillaumin, A.P., Lilly, J.M., Early, J., 2019. The debiased Whittle likelihood. Biometrika 106 (2).

Teunissen, P.J.G., Amiri-Simkooei, A.R., 2008. Least-squares variance component estimation. Journal of Geodesy 82, 65-82.

van der Ziel, A., 1970. Noise in solid-state devices and lasers. Proceedings of the IEEE 58 (8), 1178-1206. DOI: 10.1109/PROC.1970.7896.

Voss, R.F., 1991. Random fractals: characterization and measurement. In: Pynn, R., Skjeltorp, A. (Eds.), Scaling Phenomena in Disordered Systems. Springer Science+Business Media, New York, pp. 1–11.

Vyushin, D.I., Kushner, P.J., 2009. Power-law and long-memory characteristics of the atmospheric general circulation. Journal of Climate 22 (11) 2890-2904. DOI: 10.1175/2008JCLI2528.1.

Wheelon, A.D., 2001. Electromagnetic Scintillation Part I Geometrical Optics. Cambridge University Press, Cambridge, UK.

Wujanz, D., Burger, M., Mettenleiter, M., Neitzel, F., 2017. An intensity-based stochastic model for terrestrial laser scanners. ISPRS Journal of Photogrammetry and Remote Sensing 125, 146-155.