An Algorithm for Automatic Inertial Sensors Calibration: Proceedings of the ION GNSS 2013

GUERRIER, Stéphane, et al.

Abstract

We present an algorithm for determining the nature of stochastic processes together with its parameters based on the analysis of time series of inertial errors. The algorithm is suitable mainly (but not only) for situations when several stochastic processes are superposed. In such cases, classical approaches based on the analysis of Allan variance or PSD are likely to fail due to the difficulty of separating the underlying error-processes in the spectral domain. The developed alternative is based on the recently proposed method called the Generalized Method of Wavelet Moments (GMWM), whose resulting estimator was proven to be consistent and asymptotically normally distributed. The principle of this method is to match the empirical and model-based wavelet variances (WV). In this study we propose a goodness-of-fit criterion which can be used to determine the suitability of a candidate model and apply it to low-cost inertial sensors. The suggested approach of model selection relies on an unbiased estimate of the distance between the theoretical WV and the empirical WV which would be obtained on an independent sample issued […]

Reference


Available at:
http://archive-ouverte.unige.ch/unige:30305

Disclaimer: layout of this document may differ from the published version.
An Algorithm for Automatic Inertial Sensors Calibration
Stéphane Guerrier¹, Roberto Molinari¹, Jan Skaloud², and Maria-Pia Victoria-Feser¹

¹University of Geneva, Research Center for Statistics and HEC Genève, CH-1211 Geneva, Switzerland
²École Polytechnique Fédérale de Lausanne (EPFL), Geodetic Engineering Laboratory (TOPO), CH-1015 Lausanne, Switzerland

Abstract—We present an algorithm for determining the nature of stochastic processes together with its parameters based on the analysis of time series of inertial errors. The algorithm is suitable mainly (but not only) for situations when several stochastic processes are superposed. In such cases, classical approaches based on the analysis of Allan variance or PSD are likely to fail due to the difficulty of separating the underlying error-processes in the spectral domain. The developed alternative is based on the recently proposed method called the Generalized Method of Wavelet Moments (GMWM), whose resulting estimator was proven to be consistent and asymptotically normally distributed. The principle of this method is to match the empirical and model-based wavelet variances (WV). In this study we propose a goodness-of-fit criterion which can be used to determine the model selection for two types of MEMS-IMUs, the latter of higher accuracy - is to estimate the parameter set \( \theta \) based on the Allan Variance (AV) [1]. However, in situations where several stochastic processes are superposed, the methods of identifying noise parameters in AV plots or PSD analysis, as discussed for instance in Ref. [2], may be less appropriate. The reason is that AV-based methods work reasonably well only for processes which are clearly identifiable and separable in the spectral domain and not subject to spectral ambiguity [3]. However, in the error behaviour of inertial sensors of lower quality, the existence of spectral ambiguity is likely to occur.

For this purpose, Ref. [4] proposed an estimation method based on Wavelet Variance (WV) for processes (hereinafter composite processes) which are the result of a sum of underlying latent processes. In a nutshell, using a wavelet-based filter, the procedure takes the estimated WV at each scale and associates it to the WV implied by the postulated model, thus estimating the parameters of the latter which minimize a certain distance with the estimated WV. The procedure is called the Generalized Method of Wavelet Moments (GMWM) and can be represented as the solution of the following optimization problem

\[
\hat{\theta} = \arg\min_{\theta \in \Theta} (\hat{\nu} - \nu(\theta))^T \Omega (\hat{\nu} - \nu(\theta))
\]

where \( \theta \in \Theta \subseteq \mathbb{R}^p \) represents the model parameters, \( \hat{\nu} \) the estimated WV from the gathered data, \( \nu(\theta) \) the WV implied by the model as a function of \( \theta \) and \( \Omega \) is an appropriate penalty term.\(^{1}\)

\(^{1}\)\{\( Y_t \)\} can also be non-stationary but with stationary backward differences of order \( d \). The first order backward difference of \( Y_t \) is \( Y_t^{(1)} = Y_t - Y_{t-1} \) and the backward difference of order \( d \) is \( Y_t^{(d)} = Y_t^{(d-1)} - Y_{t-1}^{(d-1)} \).

\(^{2}\)When based on Haar's wavelet these variances differ from AV only by a scale factor.
The expected squared loss between the predicted values (having the realization $\hat{Y}$) where $\| | \hat{Y} - Y | |^2$ is taken from [7] (see Example 2.1). To explain how this result can be applied to the GMWM selection criterion for GMWM estimator which is based on the theoretical results of Ref. [6] (in particular Theorem 1). In [5] we have investigated the application of the GMWM estimator in characterizing the residual inertial error behaviour of a MEMS IMU. There, the GMWM estimated the parameters of the stochastic processes as well as their incertitude (i.e. 95% confidence intervals). The structure of the processes were, however, chosen arbitrarily by the designer who compared the obtained “goodness-of-fit” among models of different complexity. The aim of this contribution is to propose an algorithm that automatizes such a procedure.

To reach this goal we will first introduce (in Sec. II) an objective function that weights the goodness-of-fit between WV against the complexity of the tested model for which we define a suitable (i.e. consistent) estimator. Then, in Sec. III we describe the details of an algorithm that evaluates such an estimator. Later, in Sec. IV and Sec. V we illustrate the algorithm performance on an error signal from two-types of MEMS gyroscopes. We conclude our contribution with summary and perspectives in Sec. VI.

II. MODEL SELECTION FOR SENSOR CALIBRATION

In the context of inertial sensor calibration the task of model selection consists in finding, among a list of candidate models, the one(s) which appear to best describe the stochastic nature of the considered sensor. For this purpose, we propose a model selection criterion for GMWM estimator which is based on the theoretical results of Ref. [6] (in particular Theorem 1). To explain how this result can be applied to the GMWM estimator we shall first present a simple example considering linear regression models which is taken from [7] (see Example 2.1).

Consider the linear model $y = X\theta + \varepsilon$ where $X \in \mathbb{R}^{n \times p}$ is a full-ranked non-random matrix, $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$ and $\theta \in \Theta \subseteq \mathbb{R}^p$. Let $\hat{\theta}_{LS}$ denote the least square estimator of $\theta$, i.e.

$$\hat{\theta}_{LS} = \arg\min_{\hat{\theta} \in \Theta} \| y - X\hat{\theta} \|^2_2.$$ 

where $\| \hat{y} \|^2_2 = \hat{y}^T \hat{y}$. Suppose that we wish to find an estimator for the following criterion:

$$C = E \left[ \mathbb{E}_0 \left[ \| Y^0 - \hat{Y} \|^2_2 \right] \right]$$

where $\hat{Y}$ denotes the prediction vector constructed from $\hat{\theta}_{LS}$ (having the realization $\hat{y} = X\hat{\theta}_{LS}$). The double expectation in the above equation signifies that we wish to compute the expected squared loss between the predicted values $\hat{Y}$ (computed on the sample $Y$) and the observations $Y^0$ which are issued from the same (unknown) probabilistic model as $Y$.

Theorem 1 of Ref. [6] can be stated under some suitable regularity conditions and for the Mahalanobis loss function as:

$$C = E \left[ \mathbb{E}_0 \left[ \delta^T \Phi \delta \right] \right] = E \left[ \hat{\alpha}^T \Phi \hat{\alpha} \right] + 2 \text{tr} \left\{ \text{cov} \left[ f_1(y), \Phi f_2(y, \hat{\theta}) \right] \right\}$$

where $\hat{\alpha} = f_1(Y^0) - f_2(Y, \hat{\theta})$, $\delta = f_1(y) - f_2(y, \hat{\theta})$ and $\Phi$ is a positive definite weighting matrix. Therefore, a “natural” (and consistent) estimator of $C$ is given by:

$$\hat{C} = \hat{\alpha}^T \Phi \hat{\alpha} + 2 \text{tr} \left\{ \text{cov} \left[ f_1(y), \Phi f_2(y, \hat{\theta}) \right] \right\}$$

where $\text{cov}(\cdot)$ is obtained analytically up to a value of $\theta$, the model’s parameters, which is then replaced by $\theta$, or by resampling methods (see e.g. [8]).

The term $\delta^T \Phi \delta$ in Eq. (2) is equivalent to $\| Y^0 - \hat{Y} \|^2_2$ with $\Phi = I$. $f_1(Y^0) = Y^0$ and $f_2(Y, \theta_{LS}) = X\theta_{LS}$. Therefore, by applying Eq. (2) to $C$ we obtain:

$$\hat{C} = \| y - \hat{y} \|^2_2 + 2 \text{tr} \left\{ \text{cov} \left( y, X\hat{\theta}_{LS} \right) \right\}$$

where $H$ denotes the “hat” matrix of $X$, i.e. $H = X(X^T X)^{-1} X^T$. Note that $\hat{C}$ is (unsurprisingly) equivalent to Mallow’s $C_p$ (see Ref. [9]). While the first term $\| y - \hat{y} \|^2_2$ decreases with model complexity (i.e. the number of parameter $p$) the second term penalises the dimension of the model and increases with $p$.

The objective function given in Eq. (1) used for GMWM estimators can be seen as a discrepancy measure which aims at minimizing the error between the “predictions” made by the postulated model (i.e. $\nu(\hat{\theta})$) and the observed WV of the process (i.e. $\nu$). With this in mind we define the Wavelet Variance Information Criterion (WVIC) as:

$$\text{WVIC} = E \left[ \left( \nu^0(\hat{\theta}) - \nu(\hat{\theta}) \right)^T \Omega \left( \nu^0(\hat{\theta}) - \nu(\hat{\theta}) \right) \right]$$

The reason for defining the WVIC as the objective function of the GMWM estimator evaluated at $\hat{\theta}$ is motivated by Corollary 2.1 of Ref. [7]. It can be observed that in the context of linear regression Mallow’s $C_p$ and the Akaikes Information Criterion (see Ref. [10]) are built similarly for, respectively, the least squares and maximum likelihood.

Using Eq. (2) and assuming the matrix $\Omega$ to be constant we can write the following:

$$\text{WVIC} = E \left[ \left( \nu^0 - \nu(\hat{\theta}) \right)^T \Omega \left( \nu^0 - \nu(\hat{\theta}) \right) \right] + 2 \text{tr} \left\{ \text{cov} \left[ \nu, \nu(\hat{\theta}) \right] \Omega^T \right\}$$

Thus, a suitable (consistent) estimator of (5) can be obtained as:

$$\hat{\text{WVIC}} = \left( \nu^0 - \nu(\hat{\theta}) \right)^T \Omega \left( \nu^0 - \nu(\hat{\theta}) \right) + 2 \text{tr} \left\{ \text{cov} \left[ \nu, \nu(\hat{\theta}) \right] \Omega^T \right\}$$

where $\text{cov}(\cdot)$ denotes a consistent estimator of the covariance. Considering the form of expression (7), one can recognize
its similarity with Mallow’s \(C_p\) criterion as presented in Eq. (4). In fact, the observations in the \(C_p\) are replaced by the WV \(\tilde{\nu}\) and these are then weighted by a matrix \(\Omega\). In brief, the criterion WVIC is in some sense a weighted version of Mallow’s \(C_p\), based on a function (the WV) of the observations instead of the observations themselves. We can define the first term in Eq. (7) as the “Apparent Error” (which measures the fit of the model to the signal) and the second term as the “Optimism” (which measures the complexity of the model).

As explained in Ref. [7] we could use the distribution of \(\{Y_t\}\) to analytically obtain an expression for \(\tilde{\text{cov}}(\cdot)\) up to a value of \(\theta\) which could then be replaced by \(\hat{\theta}\). However, such an approach is model-dependent and deriving an expression for \(\tilde{\text{cov}}(\cdot)\) is generally difficult. For these reasons, we propose using resampling methods to estimate this quantity. Conveniently, the bootstrap or, more exactly, the parametric bootstrap, provides a direct way of estimating \(\text{cov}(\cdot)\) (see e.g. [8] for details). The principle is to generate \(H\) simulated samples \(\{y^*_t, \theta^*_k\}_{h=1}^{H}\) from \(F_{\hat{\theta}}\) and to estimate firstly the \(\hat{v}^*\) based on the simulated sample and, secondly, the \(\nu(\hat{\theta}^*)\) implied by \(\hat{\theta}^*\) (based on \(\hat{v}^*\) and \(\Omega\)). Then, we may estimate the quantity of interest by the observed bootstrap covariance, say \(\tilde{\text{cov}}(\cdot)\) which is defined as:

\[
\tilde{\text{cov}}(\hat{v}, \nu(\hat{\theta})) = \frac{1}{H-1} \sum_{h=1}^{H} \left( \hat{v}^*(h) - \nu(\hat{\theta}^*) \right)^T \times \\
\left( \nu(\hat{\theta}^*(h)) - \nu(\hat{\theta}^*) \right) \tag{8}
\]

where \(\hat{v}^*(h)\) and \(\hat{\theta}^*(h)\) denote respectively the estimated values for \(\hat{v}\) and \(\hat{\theta}\) obtained on the \(h\)th simulated sample and where

\[
\hat{v}^* = \frac{1}{H} \sum_{h=1}^{H} \hat{v}^*(h) \quad \text{and} \quad \nu(\hat{\theta}^*) = \frac{1}{H} \sum_{h=1}^{H} \nu(\hat{\theta}^*(h)).
\]

Therefore, WVIC can be computed by replacing \(\tilde{\text{cov}}(\cdot)\) by \(\text{cov}(\cdot)\) as defined in Eq. (8) in Eq. (7). Moreover, if the matrix \(\Omega\) is estimated by \(\tilde{\Omega}\) then we may replace \(\Omega^*\) by \(\tilde{\Omega}^*\) in Eq. (7).

In practice, the WVIC will be employed to select among \(K\) candidate models the one(s) that describes best the behaviour of the process \(\{Y_t\}\). Suppose that these \(K\) candidate models \(F_{\theta_k}, \ k = 1, ..., K\), contain the true model and are such that there exists one model, say \(F_{\theta_k}\), such that it “includes” all other models (i.e. models \(F_{\theta_k}, \ k = 1, ..., K; \ k \neq k^*\) are nested in model \(F_{\theta_k}\)). Then, \(\theta_k\) is consistent regardless of which model among the \(K\) candidates is the true model and should therefore be used to generate the \(H\) simulated samples required to compute \(\tilde{\text{cov}}(\hat{v}, \nu(\hat{\theta}))\) in (8).

III. ALGORITHM FOR MODEL SELECTION

Assuming that we have a list of \(K\) potential models, we need to obtain the values WVIC associated to each candidate model and compare them in order to select the model(s) with the lowest values(s). As already explained in Sec. II, the second term in Eq. (7) is not a directly observable statistic. An alternative solution to estimate this term is the parametric bootstrap (estimating in fact \(\tilde{\text{cov}}(\cdot)\)). This provides a direct way of estimating the covariance \(\tilde{\text{cov}}(\cdot)\) (see Ref. [8] for details). The procedure for computing \(\tilde{\text{cov}}(\cdot)\) as defined in Eq. (7) for a given model \(k\) is the following: assuming one already has an estimate of \(\hat{\theta}_k\) (i.e. as defined in Sec. II \(F_{\hat{\theta}_k}\)), such that it “includes” all other models:

1) Simulate a sample \(\{y^*_t\}_{h=1}^{H}\) from \(F_{\hat{\theta}_k}\), for \(h = 1, ..., H\).
2) Estimate \(\hat{v}^*(h)\) based on \(\{y^*_t\}_{h=1}^{H}\), for \(h = 1, ..., H\).
3) Estimate for the model of interest \(\hat{\theta}^*_k\) based on \(\hat{v}^*(h)\) using Eq. (1), for \(h = 1, ..., H\). Note that \(\hat{\theta}_k\) and \(\hat{\theta}^*_k\) denote, respectively, the parameter vector and parameter space of model \(k\).
4) Compute \(\nu(\hat{\theta}^*_k)\) based on the previously estimated \(\hat{\theta}^*_k\), for \(h = 1, ..., H\).
5) Compute \(\tilde{\text{cov}}(\hat{v}, \nu(\hat{\theta}))\) as defined in Eq. (8).

This procedure can be computationally intensive, especially if the most complex model \(F_{\hat{\theta}_k}\) includes a large number of sub-models to select from. The estimation of the WV \(\hat{\nu}(\cdot)\) for the \(H\) simulated samples (i.e. the step 4) of the above procedure) has by far the largest computational cost. Since these values are computed at \(F_{\hat{\theta}_k}\), they do not depend on the considered model and therefore need to be computed only once to compute \(\tilde{\text{cov}}(\cdot)\) for all candidate models. Hence, instead of running this procedure independently for each model, we propose an alternative approach that estimates WVIC for all models. This approach is presented in Algorithm 1 and requires (approximately) the same computational time as would be needed to compute the estimate of WVIC on a single model.

IV. EXPERIMENTAL RESULTS

A

In this section we illustrate the previously described algorithm for automated selection of a model structure and its parameters via a practical example. For this purpose we utilize data from a MEMS gyroscope (gyro A) that present rather complex error characteristics. The performance of the gyro A is rather poor (i.e. its “stated” bias stability and noise density are \(\approx 1-5\) deg/s and \(\approx 0.1\) deg/s/\(\sqrt{\text{Hz}}\), respectively). These data were observed in static conditions during several hours with sampling frequency of 100 Hz.

After removing the mean, the empirical WV of the series of the gyro A is presented as a black plot in Fig. 1. As the steepness of slope on the left part of this log-log plot is significantly greater than -1 we can conclude that the residual quantisation noise (QN) is negligible [2]. Similar analysis of the slope on the right-side (i.e. smaller than 1) evoke the absence of a drift ramp (e.g. see Fig. 5 in [2]). Hence, a probably reasonable list of possible models includes white noise (WN), random walk (RW) and several Gauss-Markov (GM) models of the first-order. By limiting arbitrarily the number of possible GM models to 3 we define the model of “highest-complexity” \(k^*\) as \([\text{WN} + \text{RW} + \text{3GM}]\). Overall, this structure represents 12 possibilities (i.e. model sub-structures) to be evaluated by the proposed algorithm employing the estimator defined in Eq. (8).

The algorithm classified numerically the “best-model structure” as that of \([\text{M1:WN + 2GM}]\) followed by the models \([\text{M2:WN + RW + 2GM}], \text{M3:3GM}, \text{M4:WN + 3GM}]\),
Algorithm 1: Automatic model selection for inertial sensors

\[ \nu = -6 - 4 - 3 \times 10^{-6} - 10^{-5} - 10^{-4} \times 10^{-3} \]

\[ Q_k = \min_{\theta_k \in \Theta_k} (\hat{\nu} - \nu(\theta_k))^T \Omega (\hat{\nu} - \nu(\theta_k)). \]

\[ \hat{\theta}_k = \arg\min_{\theta_k \in \Theta_k} (\hat{\nu} - \nu(\theta_k))^T \Omega (\hat{\nu} - \nu(\theta_k)). \]

\[ \hat{\theta}_{k^*} = \arg\min_{\theta_k \in \Theta_k} (\hat{\nu} - \nu(\theta_k))^T \Omega (\hat{\nu} - \nu(\theta_k)). \]

\[ \hat{\nu} = \frac{1}{7} \sum_{h=1}^{H} \hat{\nu}^{(h)}. \]

\[ D_k = \frac{1}{7} \sum_{h=1}^{H} (\hat{\nu}^{(h)} - \nu^{(h)}) (\nu(\hat{\theta}_k^{(h)}) - \nu(\hat{\theta}_k^{(h)})). \]

\[ \nu(\hat{\theta}_k^{(h)}) = \frac{1}{7} \sum_{h=1}^{H} \nu(\hat{\theta}_k^{(h)}). \]

\[ \hat{\nu}(\hat{\theta}_k^{(h)}) = \frac{1}{7} \sum_{h=1}^{H} \hat{\nu}(\hat{\theta}_k^{(h)}). \]

Step 1: Compute the empirical WV \( \hat{\nu} \) on the observed time series \( \{y_t\}, t = 1, ..., T \).

Step 2: Compute the objective function of the GMWM at the solution for all candidate models and estimate \( \hat{\theta}_{k^*} \):

for \( k = 1 \to K \) do

\[ Q_k = \min_{\theta_k \in \Theta_k} (\hat{\nu} - \nu(\theta_k))^T \Omega (\hat{\nu} - \nu(\theta_k)). \]

if \( k = k^* \) then

\[ \hat{\theta}_{k^*} = \arg\min_{\theta_k \in \Theta_k} (\hat{\nu} - \nu(\theta_k))^T \Omega (\hat{\nu} - \nu(\theta_k)). \]

end

end

Step 3: Compute the terms \( \hat{\nu}(\hat{\theta}_k^{(h)}) \) and \( \nu(\hat{\theta}_k^{(h)}) \), for \( k = 1, ..., K \) and \( h = 1, ..., H \):

for \( h = 1 \to H \) do

Simulate a sample \( \{y_t^h\} \) of length \( T \) under model \( F_{\hat{\theta}_k^{(h)}} \).

\[ \hat{\theta}_{k^*} = \arg\min_{\theta_k \in \Theta_k} (\hat{\nu} - \nu(\theta_k))^T \Omega (\hat{\nu} - \nu(\theta_k)). \]

\[ \hat{\nu}(\hat{\theta}_k^{(h)}) = \frac{1}{7} \sum_{h=1}^{H} \hat{\nu}(\hat{\theta}_k^{(h)}). \]

end

Step 4: Compute the term \( \hat{\nu}(\nu(\hat{\theta}_k)) \), for \( k = 1, ..., K \):

\[ \hat{\nu}(\nu(\hat{\theta}_k)) = \frac{1}{7} \sum_{h=1}^{H} \hat{\nu}(\nu(\hat{\theta}_k^{(h)})). \]

Step 5: Compute WVIC \( k \), for \( k = 1, ..., K \):

\[ \text{WVIC}_k = Q_k + 2 \text{tr} (D_k \Omega^T). \]

end

Note: WVIC \( k \) denotes the WVIC associated to model \( k \).

Step 6: Select model with the smallest WVIC:

\[ \hat{k} = \arg\min_{k \in \{1, ..., K\}} \text{WVIC}_k. \]

V. EXPERIMENTAL RESULTS B

We repeat the WV analysis and model selection for another MEMS gyro (B), whose stated performance is significantly higher with respect to the previously investigated gyro A (i.e. gyro B’s "stated" bias stability and random walk are \( \approx 10-15 \) deg/h and \( \approx 0.25 \) deg/\( \sqrt{\text{h}} \), respectively). Again, the error signal was observed in static conditions over several hours and the sampling frequency was 250 Hz.

In this case we have considered only 4 models. These models were chosen as candidate models based on the analysis of the plot of the empirically observed WV of the gyro-error signal. These are plotted as a black line in Fig. 2. As the slope of this line at the initial scales is close to -1 it is judicious to consider (not only white but also quantization noise. The behaviour of the WV at the larger scales is attempted to be first modelled by one Gauss-Markov process of the first order.

M5:[RW + 3GM], M6:[WN + RW + 3GM], M7:[WN + GM], M8:[2GM] and so on. Nevertheless, it should be mentioned that the implied fits of wavelet variances for models M2-M6 are nearly identical. They also do not differ significantly from M1. Therefore, the WVIC criterion of the selected model is smallest not only due to fact that the fit of the WV is small, but also because its structure is simpler in comparison to models M2-M6. The graphical comparisons of fitting the empirical WV with those implied by models M1 and the significantly worse models M7 and M8 are depicted in Fig. 1. The visual correlation with the empirical WV shows that the models M7 and M8 are too “simple” to represent the observed stochastic process. The same can be concluded for even simpler models.

Fig. 1. Gyro A error signal: Graphical comparisons (log-log scale) between empirical wavelet variances (WV) and the those implied by the model [WN + 2GM] (selected as optimal) and the significantly worse model structures [WN + GM] and [2GM].

Approximate WV implied by all models which includes model WN + 2GM

WV implied by models 2GM and WN + GM

Empirical WV \( \nu \)

95 % Confidence interval

\[ \nu \approx 10^{-3} \]

\[ \nu \approx 10^{-2} \]

\[ \nu \approx 10^{-1} \]

\[ \nu \approx 10 \]
Hence the Model 1 is defined as \([WN + QN + GM]\). Another Gauss-Markov process was added with the aim to better adapt the curve of the last scales in the Model 2: \([WN + QN + 2*GM]\). Alternatively, a random walk can be considered for the same purpose, which defines Model 3: \([WN + QN + GM + RW]\). As a mean to improve the fit at the very last scale, a drift (DR) was considered in Model 4: \([WN + QN + 2GM + DR]\).

The wavelet variances implied by all four models are plotted in Fig. 2. It can be seen that their respective differences are relatively small at the initial scales, but differ substantially at higher scales. However, the confidence of WV variances at higher scales decreases also. This decrease is related to the duration of the error-signal and alternatively, longer time-series can be considered to improve the confidence level at higher scales of interest. Nevertheless, the WV implied by a “good model” shall stay within the 95\% confidence interval of the empirical WV.

Furthermore, Table I shows the numerical results of the selection search in ascending order with respect to the WVIC criterion. It displays the number of parameters implied by the respective model as well as the terms of Eq. (7) linked to model complexity. According to the proposed selection procedure, the WVIC criterion suggests Model 2 \([WN + QN + 2GM]\) as optimal.

Certainly, the final decision on model complexity shall consider also other factors related to the integrated navigation. These are, for instance, connected to the size of the augmented (error) state-vector, its observability or correlation between its states. These considerations are, however, out of the scope of the investigations presented herein.

**TABLE I**

**Comparisons between the Apparent Error (App), Optimism (Opt) and WVIC values for the considered models in gyro B**

<table>
<thead>
<tr>
<th>N.par</th>
<th>App</th>
<th>Opt</th>
<th>WVIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 2</td>
<td>6</td>
<td>11580.62</td>
<td>7.19</td>
</tr>
<tr>
<td>Model 1</td>
<td>4</td>
<td>11584.05</td>
<td>4.01</td>
</tr>
<tr>
<td>Model 3</td>
<td>5</td>
<td>11582.02</td>
<td>6.46</td>
</tr>
<tr>
<td>Model 4</td>
<td>7</td>
<td>11580.62</td>
<td>11.85</td>
</tr>
</tbody>
</table>

**VI. Conclusions and Perspectives**

In this contribution we have presented an algorithm that analyses a time series of (residual) error signal from an inertial sensor and proposes a model that best describes the underlying stochastic process within a set of candidate models. The algorithm is essentially based on fitting the wavelet variances implied by a model-structure to their empirically observed counterpart. Its core is based on the GMWM estimator that bears certain analogies with the AV approach. However, the GMVM estimator has nice properties, because - as opposed to the previously investigated fits to AV - it is proven to provide consistent estimates also for composite stochastic processes. This property appears to be important especially when characterizing the stochastic behaviour of poorer (i.e. MEMS) inertial sensors.

In simple terms the presented algorithm “scans” through all the possible model combinations up to a model of highest complexity that is defined by a designer. These models are then evaluated by a criterion that weights the goodness of model’s fit against its complexity. The choice on the model of highest complexity stays arbitrary and somehow dictates the processing time. In this sense, the bottle-neck is not necessarily the estimate of the model parameters (this part is rather fast, once the empirical WV are computed) but rather in the bootstrap estimate of the covariances.

Although the presented algorithm is completely general and does not depend on the “nature” of the analysed observations, the experimental part shows its implication on the time series issued from two-types of MEMS gyroscopes. When exercised on all sensors of an IMU (and possibly in varying environmental conditions), the output of the algorithm will provide a stochastic model, whose structure and parameters - up to few exceptions - define the augmented “error-state” of a Kalman filter/smoothers employed for integrated navigation. Among these exceptions we find the random-constant state (because the mean is removed from the time-series before processing) as well as the states modelling any cross-coupling effects (e.g. misalignment between axes). Although already investigated, further publications will present the implication of the estimated stochastic model by this procedure in the scope of the Kalman filtering/smoothing.
REFERENCES


