Efficient error model construction

Zeev Berman*

Berman Consulting and Training Ltd, Kinneret, 15105, Israel

This paper describes a new method for error model construction. Instead of the standard local slope analysis of the Allan variance, two major modifications are proposed: (1) the Direct Bound principle, i.e. finding an entire error model that generates analyzing tool values that tightly bound the analyzing tool values generated by the real data; (2) instead of using Allan variance as a unique analyzing tool, a variety of analyzing tools termed Direct Predictor (DP) types 0, 1, 2, and 3 are introduced. In the paper, a uniform structure of DPs is developed and their parameterization is extended. For a nominal model that consists of a Markov process with additive white noise, the analytical functions for DPs are presented (for infinite data length). The errors due to the final data length are analyzed. Using these results, a reliable optimization problem is presented to implement the Direct Bound approach. The flexibility of working with hard and soft bounds is introduced. The presented simulation results show that the proposed method is indeed efficient and provides satisfactory results for model parameter estimations. The paper concludes with a description of an entire engineering process to cover test design and its analysis.

I. Introduction

This paper is a part of ongoing research on inertial sensor calibration under changing temperature. In it, we address the question of how to find a model (time invariant) that can be used to bound the performance of the underlying navigation system, even if at the sensor level, the random (residual after calibration) errors may follow time-variant dynamics. This is the precise reason why system identification models are not popular in the navigation community. The standard approach is to use Allan variance (AVAR) analysis and construct the error model from different local slopes of the Allan plot. A survey of the literature on inertial sensor calibration and error model analysis reveals many papers on the subject (see for example [1-3]). In this context, the IEEE Standard group's attempt to create a common terminology and framework for gyro modeling is very promising. The standard approach [4] divides the error sources into two groups: stochastic and environmental. For stochastic errors, significant effort has been invested in creating a suitable stochastic model, using the AVAR and related power spectral densities (PSDs). The attempt to present methodologies for the development of models for post-calibration residual errors [5] marks important progress. The models related to PSD, for example flicker noise—termed bias instability, are well adapted for time-invariant, stationary systems and work well at room temperature, but they are not well-suited to dealing with environmental (mainly thermal) sensitivity. Our goal is to propose an efficient method of constructing an error model under environmental sensitivity, with relatively short data-collection time and satisfactory accuracy. This new approach [6, 7] is based on two principles:

*CEO, email: berman.consult@gmail.com, web-site: http://www.bermanconsult.com

- Instead of looking for local slopes of AVAR (or other analyzing functions), we propose finding an entire model that generates analyzing tool values that tightly bound those generated by the real data. This is called the Direct Bound approach.
- Instead of using AVAR as a unique analyzing tool, a variety of analyzing tools, termed Direct Predictor (DP) types 0, 1, 2, and 3, are proposed.

The applicability of the proposed approach was verified with real-life data taken from MEM's gyro calibration. For further details see [6].

The main drawback in the proposed approach [6] was that the analysis and application were carried out using simulations and trial and error search. The purpose of this paper is to convert the ideas presented and verified in [6] into a concrete engineering procedure that can be easily implemented by the navigation community.

The paper is structured as follows. In section II, the DPs are defined as linear operators acting on the data. It is shown that all DPs have the same unified structure with different parameters for types 0, 1, 2, and 3. The nominal model in this paper is a first-order Gaussian Markov process with additive white noise. In Section III, a unified analytical function to describe the DP values for the nominal model and infinite data length is developed. In section IV, an approximation for DP errors due to finite length is proposed and verified for a wide range of cases. Section V presents an optimization problem that implements the Direct Bound approach. It turns out that in order to get reliable solutions from optimization algorithms; one must carefully define the underlying optimization problem, with proper scaling and normalization. After doing so, we present the error model accuracy for instrumental gyro model, and for all four types of DPs. In this case, the DP type 3 outperforms the Allan Variance. Although the Direct Bound approach defines only hard bounds over the test data, in this section hard bound and soft bound applications are discussed and presented. Section VI is devoted to a discussion of the entire engineering procedure, error model construction, evaluation of its accuracy, and the selection of design parameters—mainly duration of data collection.

II. The General Structure of Direct Predictors

The AVAR ${\sigma_y}^2(\tau)$ for infinite data length is defined as follows:

$$ {\sigma_y}^2(\tau) = \frac{1}{2} E\left(\Delta(\tau, y)\right)^2 \tag{1} $$

Where E is the expectation operator and $\Delta(\tau, y)$ is defined as

$$ \Delta(\tau, y) = \frac{1}{n}\sum_{i=n+1}^{2n} y(i) - \frac{1}{n}\sum_{i=1}^{n} y(i) \qquad n = \frac{\tau}{dt} \tag{2} $$

dt is the sampling time. The Allan deviation (ADEV) is the square root of AVAR and is denoted $\sigma_y(\tau)$.

The first interpretation for $\Delta(\tau, y)$ is the difference between successive mean values over two sampling periods. We will provide an additional interpretation for $\Delta(\tau, y)$ in the context of mean prediction error. Refer to $i = n$ as the current time; then $D^E = \frac{1}{n}\sum_{i=1}^{n} y(i)$ is the estimated mean data value (drift in gyro nomenclature) over the past. We can interpret $\Delta(\tau, y)$ using the following observation:

$$\Delta\left(\tau, y\right) = \frac{1}{n}\sum_{i=n+1}^{2n}\left(y(i) - D^E\right) \tag{3}$$

$\left\{y(i) - D^E\right\}_{i=n+1}^{2n}$ is the sequence of prediction errors for "future" indices $i = n+1, \cdots$.

So $\Delta(\tau, y)$ is the mean prediction error (over the sampling period with length τ) and AVAR (for infinite data length) is half the covariance of the mean prediction error. The DPs introduced in [6] are generalizations of this structure; all of them are half the covariance of the mean prediction error over the sampling interval, but they are distinct in terms of estimation and prediction methods.

DP type 0 is simply the AVAR, for which the estimator is given by:

$$D_0^E = \frac{1}{n}\sum_{i=1}^{n} y(i) \tag{4}$$

and the mean prediction error is denoted by:

$$\Delta\left(\tau, y, 0\right) = \frac{1}{n}\sum_{i=n+1}^{2n} y(i) - \frac{1}{n}\sum_{i=1}^{n} y(i) \tag{5}$$

The third argument in Δ is the predictor type.

DP type 1 is based on estimations calculated over a fixed interval (independent of τ).

$$D_1^E = \frac{1}{m}\sum_{i=1}^{m} y(i) \tag{6}$$

$$\Delta\left(\tau, y, 1\right) = \frac{1}{n}\left(\sum_{i=m+1}^{m+n} y(i) - \frac{1}{m}\sum_{i=1}^{m} y(i)\right) \tag{7}$$

DP type 2 is based on the Kalman filter estimation; this calculation requires definition of a nominal model. The nominal model assumed in this paper is a first-order Gaussian Markov process with additive white noise. To describe D_2^E as a function of data samples, some analysis is needed. For Kalman filter implementation, the measurement is defined by $y(i) = x(i) + v(i)$; $x(i)$ is the first-order Gaussian Markov process, given by $x(i) = \alpha_f x(i-1) + w(i-1)$, and $v(i), w(i)$ are measurement and process random noise processes, respectively. The Kalman filter iterations (see [8] for notations) are:

$$\begin{aligned} x_e^-(i) &= a_f x_e(i-1) \\ x_e(i) &= x_e^-(i) + k\left(y(i) - x_e^-(i)\right) \end{aligned} \tag{8}$$

$x_e(i)$ is the estimate, k is the Kalman gain.
Combining them:

$$x_e(i) = (1-k)\alpha_f x_e(i-1) + ky(i) = \mu x_e(i-1) + ky(i) \tag{9}$$

with $\mu = (1-k)\alpha_f$. In the context of this paper, the Kalman filter gain k is taken as its steady-state value. See the Appendix for the steady-state Kalman filter gain calculation in this case. By inserting the above iteration for $i = 1,2,..m$, using $x_e(0) = 0$, we get:

$$D_2{}^E = x_e(m) = k\sum_{i=1}^{m} \mu^{m-i} y(i) \tag{10}$$

and therefore

$$\Delta(\tau, y, 2) = \frac{1}{n}\left(\sum_{i=m+1}^{m+n} y(i) - nk\sum_{i=1}^{m} \mu^{m-i} y(i)\right) \tag{11}$$

DP type 3 applies both the optimal estimator and the optimal predictor (for a nominal model). Consider again the prediction error series $\left\{y(i) - D^E\right\}_{i=m+1}^{m+n}$. For an optimal predictor that applies the underlying system model (first-order Gaussian Markov process with additive white noise), we can calculate the prediction error series as $\left\{y(i) - \alpha_f{}^{i-m} D^E\right\}_{i=m+1}^{m+n}$ (see [8] for details). In this case, DP type 3 can be calculated as:

$$\Delta(\tau, y, 3) = \frac{1}{n}\left(\sum_{i=m+1}^{m+n} y(i) - k\left(\alpha_f + \alpha_f{}^2 + \cdots \qquad \sum_{i=1}^{m} \mu^{m-i} y(i)\right)\right. \tag{12}$$

$$\Delta(\tau, y, 3) = \frac{1}{n}\left(\sum_{i=m+1}^{m+n} y(i) - k\alpha_f \frac{1-\alpha_f{}^n}{1-\alpha_f}\sum_{i=1}^{m} \mu^{m-i} y(i)\right) \tag{13}$$

After describing the four types of predictors, we can deduce that all of them obey the same general structure:

$$\Delta(\tau, y) = \frac{1}{n}\left(\sum_{i=m+1}^{m+n} y(i) - \beta\sum_{i=1}^{m} \mu^{m-i} y(i)\right) \tag{14}$$

with the following parameter settings:
DP type 0: $m = n, \mu = 1, \beta = 1$

DP type 1: $\mu = 1, \beta = \dfrac{n}{m}$

DP type 2: $\mu = (1-k)\alpha_f, \beta = nk$

DP type 3: $\mu = (1-k)\alpha_f, \beta = k\alpha_f \dfrac{1-\alpha_f{}^n}{1-\alpha_f}$

Note that since all DPs satisfy the same structure, they are equivalent in terms of computational complexity.

III. Direct Predictor Analysis

In this section we take the general structure of the mean prediction error described in Eq. (14) and insert the structure of the signal generated by the nominal system model that consists of the first-order Gaussian Markov process with additive white noise:

$$\begin{aligned} y(i) &= x(i) + v(i) \\ x(i) &= \alpha x(i-1) + w(i-1) \end{aligned} \tag{15}$$

where the sensor additive $v(i)$ is white noise (which after integration causes random walk) and $w(i)$ is process noise related to the Markov process. First we will develop an equation for the mean prediction error $\Delta(\tau, y)$ for this particular signal. Then its covariance will be calculated to provide DP values (for infinite data length).

Since the predictor is a linear operator and our signal $y(i)$ can be considered the sum of two signals—$x(i), v(i)$—we can analyze them separately. We start with the Markov process. One can solve the recursion defined in Eq. (15) to get:

$$x(i) = \alpha^{i-1}x(1) + \alpha^{i-2}w(1) + \alpha^{i-3}w(2) + ... + \alpha w(i-2) + w(i-1) \tag{16}$$

Subjecting Eq. (16) to a straightforward but relatively long and involved algebraic manipulation, the following expression is obtained:

$$\sum_{i=1}^{m} \mu^{m-i} x(i) = \mu^{m-1} S\left(\frac{\alpha}{\mu}, m\right) x(1) + \sum_{i=1}^{m-1} \mu^{m-1-i} S\left(\frac{\alpha}{\mu}, m-i\right) w(i) \tag{17}$$

where $S(q,n)$ is a sum of geometric series with ratio q, n elements and 1 as the first element.

$$S(q,n) = 1 + q + \cdots = \frac{1-q^n}{1-q} \tag{18}$$

Note that $S(q,1) = 1$.

Similarly, one can show that

$$\sum_{i=m+1}^{m+n} x(i) = S(\alpha, n) x(m+1) + \sum_{i=m+1}^{m+n-1} S(\alpha, m+n-i) w(i) \tag{19}$$

Now we need to find a substitution for $x(m+1)$, which is

$$x(m+1) = \alpha^{m-1}x(1) + w(m) + \sum_{i=1}^{m-1} \alpha^{m-i} w(i) \tag{20}$$

Finally, after omitting some details which are straightforward but quite involved, we get the following general structure for the mean prediction error of the Markov process for all types of DPs:

$$\begin{aligned} \Delta(\tau, x) = &\frac{1}{n}\left\{ S(\alpha, n)\alpha^m - \beta\mu^{m-1} S\left(\frac{\alpha}{\mu}, m\right) \right\} x(1) + \\ &+ \frac{1}{n}\sum_{i=1}^{m-1}\left[S(\alpha, n)\alpha^{m-i} - \beta\mu^{m-1-i} S\left(\frac{\alpha}{\mu}, m-2-i\right) \right] w(i) + \\ &+ \frac{1}{n}\sum_{i=m}^{m+n-1} S(\alpha, m+n-i) w(i) \end{aligned} \tag{21}$$

From Eq. (21), we can calculate half the variation of the mean prediction error; its root square is the DP value for infinite data length, and in particular, for type 0 we will have the exact AVAR expression for the Markov process.

Recall that $x(1), w(i)\ i = 1,2,\cdots$ -1 are independent. $x(1)$ has variance $S_m{}^2$ and $w(i)\ i = 1,2,...,N$ has covariance $S_w{}^2 = S_m{}^2(1-\alpha^2)$. Using this information, the covariance calculation is direct.

$$\sigma_x{}^2(\tau) = \frac{1}{2}Var\left(\Delta(\tau,x)\right) = \frac{1}{2n^2}\left\{S(\alpha,n)\alpha^m - \beta\mu^{m-1}S\left(\frac{\alpha}{\mu},m\right)\right\}^2 S_m{}^2 + \\ + \frac{1}{2n^2}\sum_{i=1}^{m-1}\left[S(\alpha,n)\alpha^{m-i} - \beta\mu^{m-1-i}S\left(\frac{\alpha}{\mu},m-2-i\right)\right]^2 S_w{}^2 + \\ + \frac{1}{2n^2}\sum_{i=m}^{m+n-1} S(\alpha,m+n-i)^2\, S_w{}^2 \tag{22}$$

Using $S_w{}^2 = S_m{}^2\left(1-\alpha^2\right)$ we get the general formula for DP values (for infinite data length) for the first-order Gaussian Markov process:

$$\sigma_x{}^2(\tau) = \frac{S_m{}^2}{2n^2}\left[\begin{array}{l}\left(S(\alpha,n)\alpha^m - \beta\mu^{m-1}S\left(\frac{\alpha}{\mu},m\right)\right)^2 + \\ +\left(1-\alpha^2\right)\sum_{i=1}^{m-1}\left[S(\alpha,n)\alpha^{m-i} - \beta\mu^{m-1-i}S\left(\frac{\alpha}{\mu},m-2-i\right)\right]^2 + \\ +\left(1-\alpha^2\right)\sum_{i=m}^{m+n-1} S(\alpha,m+n-i)^2\end{array}\right] \tag{23}$$

To complete the calculation for our process, we need to calculate the DP for v, the white noise component. We start with Eq. (14) for this case:

$$\Delta(\tau,v) = \frac{1}{n}\left(\sum_{i=m+1}^{m+n} v(i) - \beta\sum_{i=1}^{m}\mu^{m-i}v(i)\right) \tag{24}$$

Recall that $v(i)$ is independent with variance $S_v{}^2$; the square of the DP values (half covariance) for white noise is given as:

$$\sigma_v{}^2(\tau) = \frac{S_v{}^2}{2n^2}\left(\sum_{i=m+1}^{m+n} 1 + \beta^2\sum_{i=1}^{m}\mu^{2(m-i)}\right) = \frac{S_v{}^2}{2n^2}\left(n + \beta^2 S(\mu^2,m)\right) \tag{25}$$

Now, since the Markov process and white noise components are independent, we can combine Eqs. (23) and (25) to obtain:

$$\sigma_y(\tau) = \sqrt{\sigma_x{}^2(\tau) + \sigma_v{}^2(\tau)} \tag{26}$$

Eq. (26), together with Eqs. (23) and (25), provides a closed-form solution for $\sigma_y(\tau)$ of the first-order Gaussian Markov process with additive white noise. The model parameters are S_m, S_v - the standard deviations of the Markov process and white noise components, respectively; the third model parameter is α, related to the Markov process time constant t_m by $\alpha = \exp(-dt/t_m)$, where dt is sampling time. The other parameters, $m.\mu, \beta$, are related to DP structure (type and parameters). The parameter n is related to the argument τ by $n = \tau / dt$.

Note that the notation $\sigma_y(\tau,...)$ is used in this paper for DPs calculated for infinite data length by Eqs. (26), (23), and (24). The DP value calculated for final data length will be denoted $DP(\tau,...)$.

IV. Direct Predictor Errors Due to Finite Data Length

We begin by describing how the DP for finite data length is calculated, using a non-overlapping approach. The total number of samples is denoted by N; for any τ we divide the data sample into N_τ groups (windows), every window with $m+n$ elements, and every group starting with the index n_k, such that $n_1 = 1, n_2 = m+n+1, n_3 = 2(m+n)+1, \ldots n_{N_\tau} = (N_\tau - 1)(m+n) - 1$. The mean prediction error, calculated for window k, is defined by:

$$\Delta(\tau, y, k) = \frac{1}{n}\left(\sum_{i=n_k+m}^{n_k+m+n-1} y(i) - \beta \sum_{i=n_k}^{n_k+m-1} \mu^{m-i} y(i) \right) \tag{27}$$

The square of DP, denoted DQ, is estimated as follows:

$$DQ(\tau, y, N) = DP(\tau, y, N)^2 = \frac{1}{2N_\tau} \sum_{k=1}^{N_\tau} \Delta(\tau, y, k)^2 \tag{28}$$

To compare DQ with a Chi-square distribution, we make the following normalization

$$DQ(\tau, y, N) = \frac{2\sigma_y(\tau)^2}{2N_\tau} \sum_{k=1}^{N_\tau} \left(\frac{\Delta(\tau, y, k)}{\sqrt{2}\sigma_y(\tau)} \right)^2 \tag{29}$$

If $\Delta(\tau, y, k)$ $k = 1, 2, ..., N_\tau$ is independent, $\sum_{k=1}^{N_\tau} \left(\frac{\Delta(\tau, y, k)}{2\sigma_y(\tau)} \right)^2$ has a Chi-squared distribution with N_τ degrees of freedom. In this case, the mean is N_τ and variation is $2N_\tau$ (see [9]). Therefore, the mean of $DQ(\tau, y, N)$ is $\sigma_y(\tau)^2$ and its variance is $\frac{2\sigma_y(\tau)^4}{N_\tau}$. Moreover, for large N_τ, a Gaussian distribution for $DQ(\tau, N, y)$ can be assumed.

Similarly DP, which is the square root of DQ, can be described using the Chi distribution of $\sqrt{\sum_{k=1}^{N_\tau} \left(\frac{\Delta(\tau, y, k)}{\sqrt{2}\sigma_y(\tau)} \right)^2}$ with mean $m_p = \sqrt{2}\frac{\Gamma(0.5(N_\tau + 1))}{\Gamma(0.5N_\tau)}$ and variance $\sigma_p{}^2 = N_\tau - m_p{}^2$.

$\Gamma(x)$ denotes the gamma function. For the purpose of the analysis required here, we will avoid the use of gamma functions, mainly because we found a that linear approximation provides a simpler and satisfactory result. The following analysis is then performed.

$$y_0 = \sqrt{x_0}, \quad y_0 + \delta y \approx \sqrt{x_0} + \frac{1}{2\sqrt{x_0}} \delta x \tag{30}$$

In our case $x_0 = \sigma_y(\tau)^2, y_0 = \sigma_y(\tau)$ are the mean values of DQ, DP, and $\delta x, \delta y$ are the standard deviations of DQ, DP, respectively. To conclude the above discussion:

$$mean\left(DP\left(\tau, y, N\right)\right) = \sigma_y(\tau) \tag{31}$$

$$std\left(DP\left(\tau, y, N\right)\right) = \frac{1}{\sqrt{2N_\tau}} \sigma_y(\tau) \tag{32}$$

Eqs. (31) and (32) provide us with the error analysis due to the final data length that we were looking for. However, before we can accept these results we need to examine our assumptions:

- The sequence $\Delta(\tau, y, k)$ is assumed to be independent, due to the underlying Markov process; in general, this is not true. However, the mean dependence is expected to be weak because the dependence exists only for adjacent windows and the number of windows is large.
- The linear approximation for square root should be valid; this means that the standard deviation of DQ should be much smaller than its mean value, and indeed, it is valid for large N_τ.

The final verification of this assumption was carried out by simulation. We considered a variety of systems and our conclusion was that for a collection time of one hour, $dt = 1, 1 \le \tau \le 300$, this approximation provides poor results. If the collection time is on the order of two hours or more and the quality of the sensors is in the range of $0.1 \div 100\ °/h$, Eqs. (31) and (32) are certainly valid. The following plot compares the values calculated by Eqs. (31) and (32) (precise) with the results of 500 Monte Carlo runs, each describing a two-hour data collection. The presented case is for DP type 3, and instrument model $S_m = 100°/h, S_v = 1°/\sqrt{h}, t_m = 25\sec$

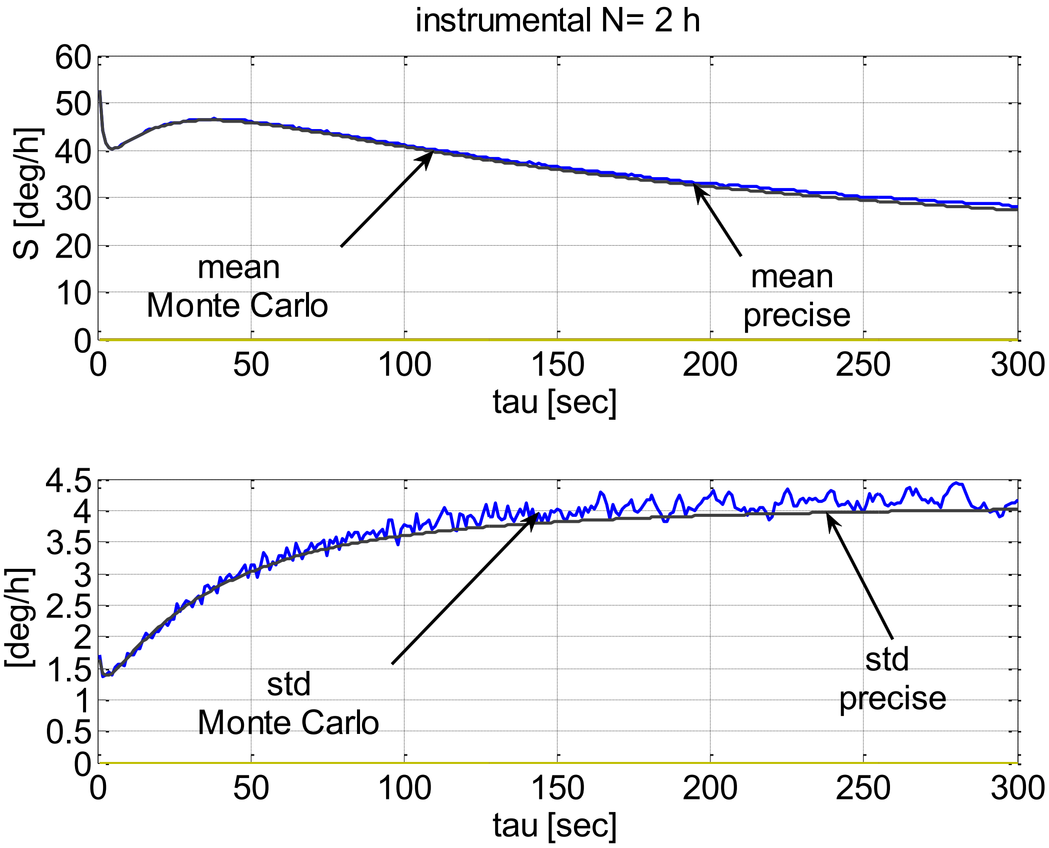

Figure 1. Errors due to finite length: instrumental system

Figure 2 describes the same results for tactical level system $S_m = 1°/h, S_v = 0.05°/\sqrt{h}, t_m = 100\,\text{sec}$.

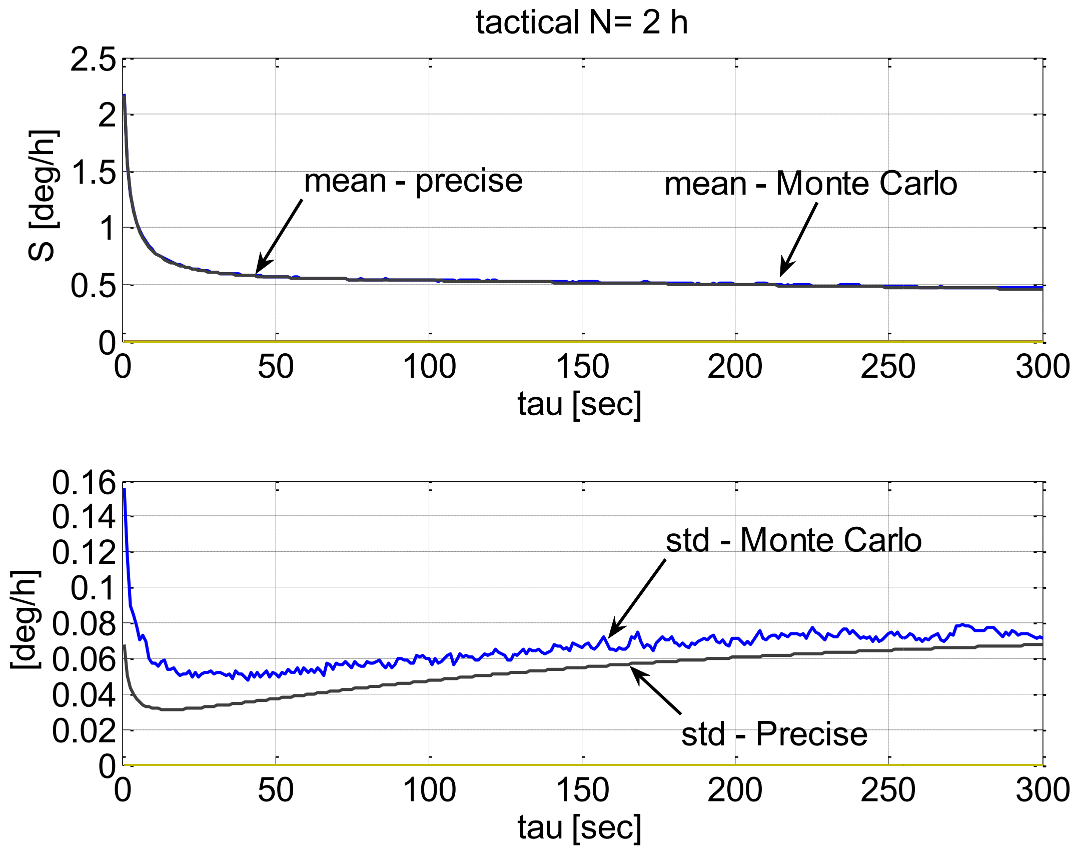

Figure 2. Errors due to finite length: tactical system

V. Optimization Techniques for Direct Bound Implementation

Given DP values calculated from the real data, $DP(\tau)$ (arguments of the signal y and data length N are omitted), we are looking for parameters S_m, S_v, t_m such that the model $M_D(S_m, S_v, t_m)$ will generate $\sigma(\tau, M_D, T)$ to be a tight bound for $DP(\tau)$. Observe that the notation for $\sigma_y(\tau)$ was modified to $\sigma(\tau, M_D, T)$, with T representing the DP type (with its parameters). Where the DP type is not important, the notation $\sigma(\tau, M_D)$ will be used as well.

Let us write the formal problem definition. We consider a discrete sequence of $\{\tau_i\}_{i=1}^{K_\tau} = \tau_{\min}, \tau_{\min} + d\tau, \tau_{\min} + 2d\tau, ..., \tau_{\max}$. $DP(\tau_i)$ and we need to solve the following optimization problem:

$$\min_{S_m, S_v, t_m} \left\{ \sum_{i=1}^{K_\tau} \left(\sigma\left(\tau_i, M_D\left(S_m, S_v, t_m\right)\right) - DP(\tau_i) \right)^2 \right\} \quad (33)$$

$$\textit{subject to } \sigma\left(\tau_i, M_D\left(S_m, S_v, t_m\right)\right) \geq DP(\tau_i) \;\; i = 1, 2, ..., K_\tau \quad (34)$$

For this problem, we were surprised to get poor, unstable results with a significant number of outlier solutions, even with different optimization methods. We thus understood that some modifications in the above definition are required. In the following we describe the modification required to get stable solutions.

The physical values of S_m, S_v are very small. To get a reliable optimization method we need to introduce proper scaling. The choice is to work with a vector of engineering units: $^\circ/h, ^\circ/\sqrt{h}, \sec$ and to introduce the following notation:

$M_D(S_m, S_v, t_m) = M_D{}^S(X)$ with $X(1) = s_1 S_m, X(2) = s_2 S_v, X(3) = t_m$ such that $s_1 = \frac{180}{\pi}3600, s_2 = \frac{180}{\pi}\sqrt{3600}$.

Now observe that the minimization defined in Eq. (33) is related to the norm of residuals for the following equation:

$$DP(\tau_i) = \sigma_y\left(\tau_i, M_D{}^S\left(X^*\right)\right) + v(\tau_i) \tag{35}$$

where X^* represents the true model parameters and $v(\tau_i)$ is the error due to finite data length. We know that standard deviation of this error is given by $std\left(DP(\tau)\right) = \frac{1}{\sqrt{2N_\tau}}\sigma_y(\tau)$. Therefore, to deal with normalized residuals we need to consider

$$\min_{S_m, S_v, t_m}\left\{\sum_{i=1}^{K_\tau}(r_i)^2\right\} \tag{36}$$

$$r_i = \sqrt{2N_\tau}\,\frac{\sigma_y\left(\tau_i, M_D{}^S(X)\right) - DP(\tau_i)}{\sigma_y\left(\tau_i, M_D{}^S(X)\right)} \tag{37}$$

We can define a penalty factor ρ and combine the constraints defined in Eq. (34) into the minimization defined in Eq. (36) by replacing r_i by its bounded version $r_i{}^B$:

$$r_i{}^B = \begin{cases} r_i & if \;\; r_i > 0 \\ \rho r_i & else \end{cases} \tag{38}$$

The penalty factor ρ is selected to ensure the condition:

$$\sigma_y\left(\tau_i, M_D{}^S\left(X^*\right)\right) \geq DP(\tau_i) \tag{39}$$

Moreover, ρ can provide flexibility from hard bound (high ρ), via soft bound (moderate ρ) to best match (standard least square: $\rho = 1$). The optimization problem defined by the cost function described in Eqs. (36–39) was tested for a variety of cases and reliable, stable results were obtained.

Having developed a robust optimization method, we approached the problem of DP selection and the accuracy of the estimated model parameters. Our first test was for the instrumental system, with parameters: $S_m = 100^\circ/h, S_v = 1^\circ/\sqrt{h}, t_m = 25\sec$. The test duration was two hours, sampling time $dt = 1\sec$ and the range of prediction time τ was $1 \leq \tau \leq 300$, termed the sequel tau range. The following plot provides insight into how the optimization works. We present four plots:

- DP – a sample of DPs calculated from sample data—actual

- $\sigma_y\left(\tau, M_D{}^*\right)$ – DP calculation based on the true model—precise
- $\sigma_y\left(\tau, M_D{}^{EH}\right)$ – DP calculated using the model found by hard bound optimization, $\rho = 100$ —hard bound.
- $\sigma_y\left(\tau, M_D{}^{ES}\right)$ – DP calculated using the model found by soft bound optimization, $\rho = 10$ —soft bound.

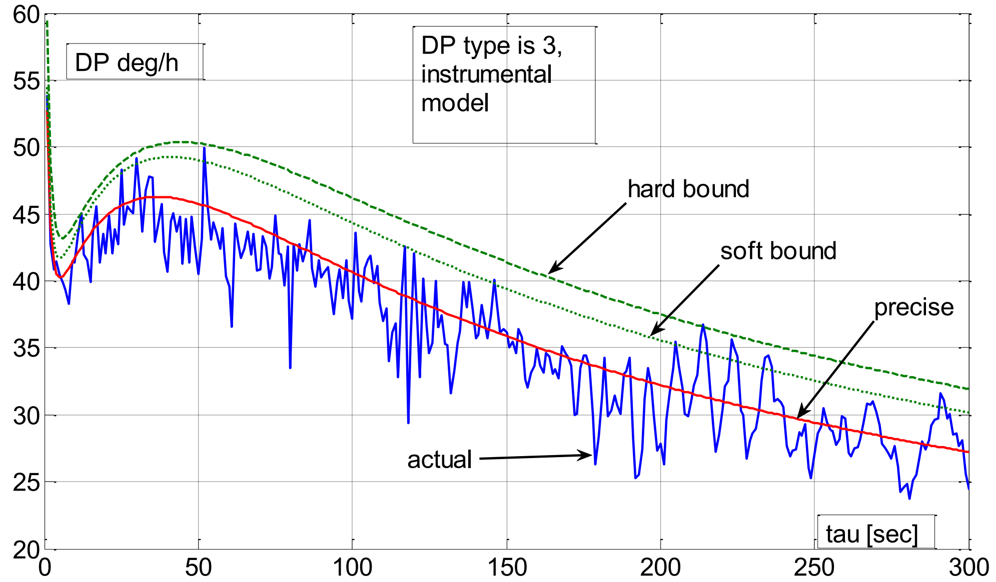

Figure 3. DPs found by the optimization algorithm

The following table summarizes the results of 100 Monte Carlo runs. For every run, we calculated the estimated parameters for all four types of DP. After all runs, mean value and standard deviation (std) were calculated for each type. The value "total rel" is rss (root sum square) of the mean and standard deviation, divided by the nominal parameter. Using this table we can evaluate performance and compare the accuracy of each DP type.

Table 1. The errors of estimated model parameters: instrumental system.

The estimated model parameters errors				DP type 0	1	2	3
Hard bound	S_m	$°/h$	mean	22	27.8	45.2	12.6
			std	6.1	8.7	51.0	4.8
			total rel	*22%*	*29%*	*68%*	*13%*
	S_v	$°/\sqrt{h}$	mean	0.02	0.34	0.10	0.07
			std	0.04	0.4	0.2	0.1
			total rel	*4%*	*52%*	*22%*	*12%*
	t_m	sec	mean	10.3	22.5	30.8	5.7
			Std	5.7	12.8	36.5	5.6
			total rel	*47%*	*103%*	*191%*	*32%*
Soft bound	S_m	$°/h$	mean	15.1	17.1	23.4	8.7
			std	5.2	6.6	6.7	4.2
			total rel	*16%*	*18%*	*24%*	*9.6%*
	S_v	$°/\sqrt{h}$	mean	0.01	0.21	0.08	0.05
			std	0.03	0.24	0.08	0.08
			total rel	*3%*	*31%*	*11%*	*9%*
	t_m	sec	mean	6.8	12.5	14.4	3.6
			std	4.6	8.0	6.8	4.8
			total rel	*33%*	*59%*	*63%*	*24%*

The following observations can be made from the table:

- The mean of estimated errors shows that the estimators are strongly biased. After some consideration, this should not be a surprise. The nature of bounding is that it provides higher DP plots than the actual one. The bias estimator is the price for our attempt to bound the performance of a system that is not necessarily time-invariant. Of course, we can reduce this price by allowing some crossing of the actual DP (soft bound).

- In principle, we have three different quality criteria: the accuracies of S_m, S_v, t_m. The best DP in one category is not necessarily the best one in another.

- In the case presented here, DP type 3 is the best for the Markov process. Its errors for Markov process parameters are about 50% better than the second-best DP, which is type 0, the AVAR. For white noise the AVAR is better, but DP type 3 can be considered satisfactory with 10% accuracy.

- DP types 1 and 2 show relatively poor performance in this case.

- In general, for soft bound optimization, two hours of data collection to get 10% accuracy in S_m, S_v and 25% in the time constants seems to be a very efficient and precise approach to error model construction.

VI. Recommendations for Practical Applications and Summary

Good engineering practice requires answering the following questions:

#1. How should the error model parameters be calculated?

#2. How should the test be designed to obtain satisfactory accuracy of these parameters?

#3. How should the test be monitored and sensor malfunctions detected?

In this paper, tools to deal with all of these questions were developed and presented.

With respect to #1, our position is clear: *never use local slopes*, use matching techniques, such as Eqs. (36–38). The decision of whether to use hard bound, soft bound or even best matching depends on the application tradeoff between the estimation accuracy and sensitivity to detecting outliers with respect to the time-invariant model. In the preparation phase, one can analyze what kind of accuracy degradation is related to higher outlier detection. The preparation phase, which is based on simulations, only provides the right answers for #2. In this phase, we need to select the proper collection time, range of tau, kind of matching used by the estimation (hard bound, soft bound, best matching), and type of DP to be used for real data. Our recommendation is to select the minimum collection time, the maximum tau range, and the hardest bound that provides satisfactory estimation accuracy. The minimum collection time saves costs, the maximum tau range and hardest bound provide good detectability for outliers. Of course we will select the best DP for the case. Recall that beside freedom of selection from four different types, there also exists freedom in DP parameter selection. Figure 4 describes the flow of the simulation, applied for test design.

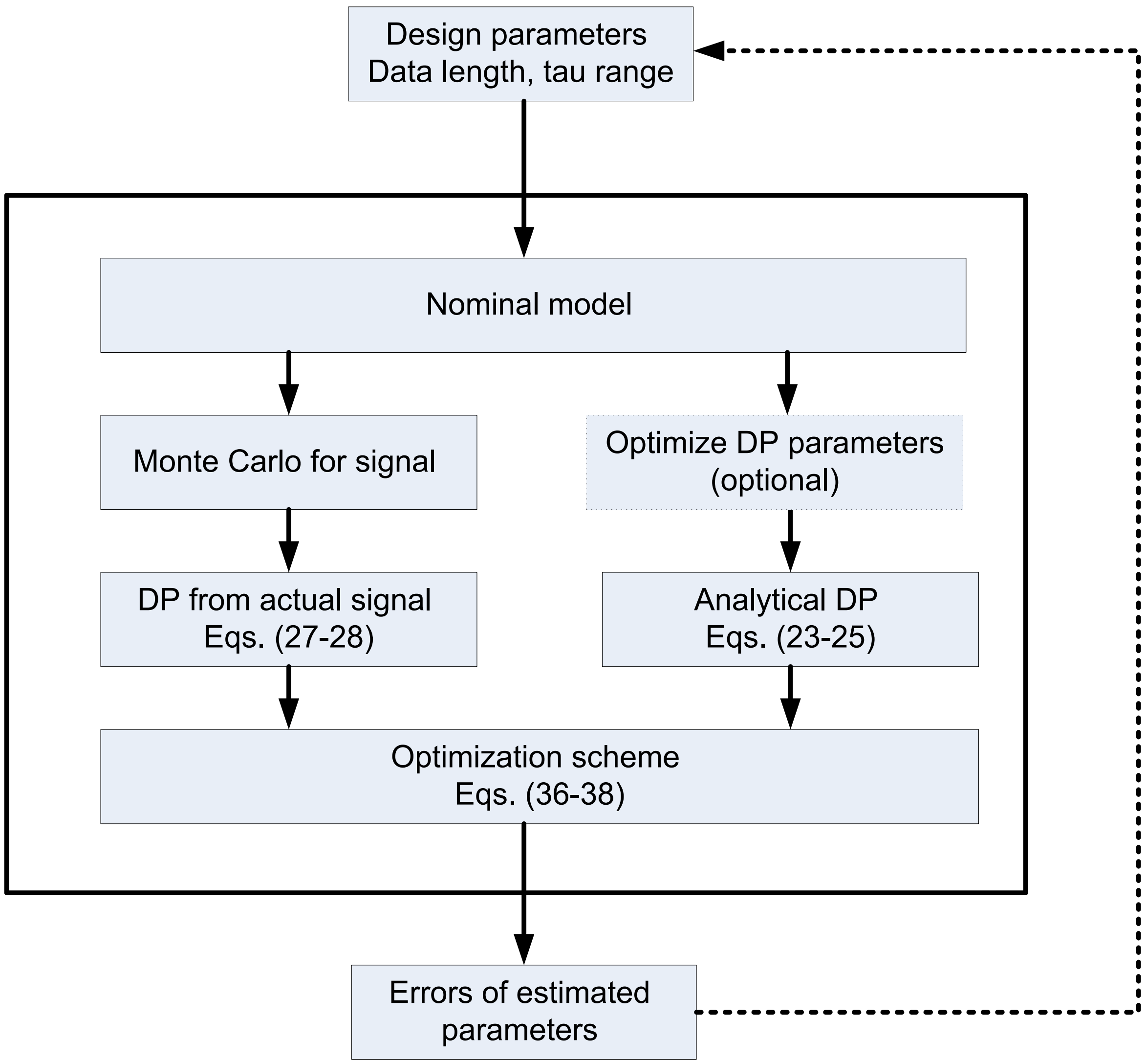

Figure 4. Diagram block for simulation phase

After defining the test parameters, the real-data test appears straightforward; it is such in the path of model parameter estimations. An additional path of quality of matching, which measures the residuals of the optimization function and compares them with the statistics of errors due to finite length (Eq. 32) provides some, perhaps partial, answers to problem #3.

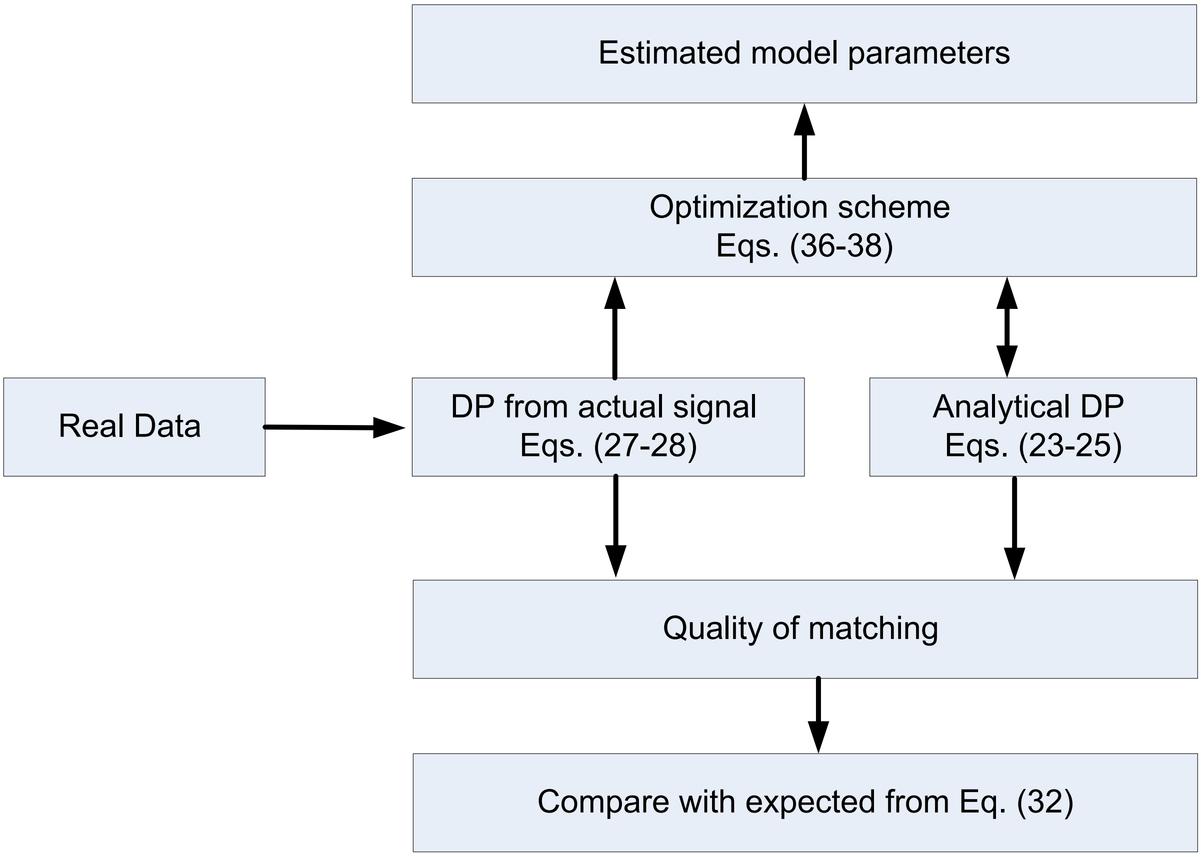

Figure 5. Block diagram for real-data test

Appendix

Calculation of the Steady-State Kalman Gain

The development here is based on notations and the equation presented in [8].
The measurement matrix is scalar equal to 1:

$$z(i) = y(i) + v(i) \Rightarrow H = 1 \tag{40}$$

The propagation system matrix is scalar equal to α

$$y(i) = \alpha y(i-1) + w(i-1) \Rightarrow \Phi = a \tag{41}$$

The covariance of measurement noise error $v(i)$ is r and the covariance of process noise $w(i)$ is q.

$$K_k = P_k^- H^T \left(H P_k^- H^T + R \right)^{-1} \Rightarrow k_k = \frac{p_k^-}{p_k^- + r} \tag{42}$$

$$P_k^+ = \left(I - K_k H \right) P_k^- \Rightarrow p_k^+ = (1 - k_k) p_k^- \tag{43}$$

$$P_{k+1}^{-} = \Phi P_k^{+} \Phi^T + Q \Rightarrow p_{k+1}^{-} = \alpha^2 p_k^{+} + q \tag{44}$$

The steady-state gain is denoted by k. Under steady-state conditions $p_{k+1}^{-} = p_k^{-} = p$.

$$p = \alpha^2 (1-k) p + q \tag{45}$$

Using

$$k = \frac{p}{p+r} \Rightarrow (1-k) = \frac{r}{p+r} \tag{46}$$

one gets

$$p^2 + \left(r - \alpha^2 r - q\right) p - qr = 0 \tag{47}$$

The positive solution of Eq. (47) is given by

$$p^* = \frac{r - \alpha^2 r - q}{2} + \sqrt{\left(\frac{\left(r - \alpha^2 r - q\right)^2}{4} + qr\right)} \tag{48}$$

The final solution is obtained by inserting Eq. (48) into Eq. (46)

$$k = \frac{p^*}{p^* + r} \tag{49}$$

References

[1] Aggarwal, P., Syed, Z., Niu, X., and El-Sheimy, N., "A Standard Testing and Calibration Procedure for Low Cost MEMs Inertial Sensors and Units," *Journal of Navigation*, Vol. 61, 2008, pp. 323-336.

[2] Titterton, D. H. and Weston, J. L., *Strapdown Inertial Navigation Technology*, Peter Peregrinus Ltd., UK, 1997.

[3] Flenniken, W. S. IV, Wall, J. H., and Bevly, D. M., "Characterization of Various IMU Error Sources and the Effect on Navigation Performance," *ION GNSS*, 2005.

[4] "IEEE Standard Specification Format Guide and Test Procedure for Single-Axis Interferometric Fiber Optic Gyros," *IEEE Std* 952-1997 (R2003).

[5] Xing, Z. and Gebre-Egziabher, D., "Modeling and Bounding Low Cost Inertial Sensor Errors," *Proceedings of IEEE/ION PLANS* 2008, May 6–8, 2008.

[6] Berman, Z., "Inertial Sensors—Further Developments in Low Cost Calibration and Testing," *PLANS* 2012, April 2012.

[7] Berman, Z.,"Inertial sensors—A New Approach for Low Cost Calibration and Testing," *Inertial Sensors and Systems Gyro Technology Symposium* , Karlsruhe, 2011.

[8] Gelb, A., *Applied Optimal Estimation*, M.I.T. Press, Cambridge, Massachusetts, and London, England 1974.

[9] Papoulis, A., *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, USA, 2nd ed., 1984.