Inertial Sensors: Further Developments in Low-Cost Calibration and Testing

Zeev Berman
Berman Consulting and Training Ltd.
email: berman.consult@gmail.com

***Abstract*—Inertial sensor static calibration under changing temperature is proposed and analyzed. Due to relatively short test duration and non-stationary effects, a modification of the Allan variance, called Direct Predictor (DP), is proposed to describe the residual error model. With a reliable method of acquiring an error model, the calibration procedure is analyzed and the trade-off between temperature profile and calibration errors is shown. Results of an actual test and an in-depth discussion about models tightly bounding the residual errors complete the paper.**

Keywords-*calibration, inertial sensor, error model, Allan variance, Direct Predictor*

I. INTRODUCTION

In this paper, inertial sensor calibration under changing temperature is proposed. This scheme has several crucial advantages over the standard calibration, performed at several stabilized temperatures. The first and in the author's opinion, most important advantage is the system's exposure to real-life conditions, as opposed to stabilized temperature settings that do not characterize the actual system operating cycle. The second advantage is that it saves time, which is a significant expense-reducing factor, even for an automatic test scenario due to high equipment cost. The third advantage, as shown in [3], is the better calibration accuracy, which is achieved mainly due to collection and analysis of data from the entire test, in contrast to the temperature-stabilized approach, where most of the time is wasted waiting for temperatures to stabilize. The fourth advantage stems from the good control achieved in the error model, which is necessary for this approach and allows successful calibration with relatively low-cost equipment (for details see [3]).

A survey of the literature on inertial sensor calibration and error model analysis reveals many papers on the subject (see [4], [6], [7], [8], [9], [10], [11], [12]). In these papers, the issues of calibration method and error model analysis are usually treated separately. In this context, the IEEE standard group's attempt to create a common terminology and framework for gyro calibration and modeling is very promising. The standard approach (see [5]) divides the error sources into two groups: stochastic and environmental. For stochastic errors, there has been a significant effort to create a suitable stochastic model, using the Allan variance and related power spectrum densities (PSDs). This type of model, for example flicker noise, termed bias instability, is well adapted for time-invariant, stationary systems and works well at room temperature, but is not well-suited to deal with environmental (mainly thermal) sensitivities.

The general calibration methods can be found in [4] and [6]. An attempt to identify the Markov process that bounds sensor instability using the Allan variance and frequency domain analysis was proposed in [7]. We use a similar idea to find a Markov process that bounds sensor instability but instead of the Allan variance, our calculation is based on an operator called the direct predictor (DP) and is carried out in the time domain. The DP is a generalization of the Allan variance. Moreover, instead of searching for local properties such as slopes and local minima (as is often done with the Allan variance), we propose a global methodology, in which a DP based on the actual data and a second DP calculated by simulation of the entire error model are compared.

Despite the rich literature on sensor calibration and modeling, two aspects are missing: detailed discussions of considerations related to calibration design and analysis, and detailed reports from well-established navigation companies. The latter is presumably related to the fact that those companies consider the details of their calibration procedure to be intellectual property. Having gained some experience in inertial sensor calibration, worked with several producers and customers, and learned how different details come together into a complete framework, we

decided to share the conclusions and lessons learned with the community. Perhaps this will encourage others to take part in such discussions.

This paper describes the design of a static calibration process, namely a trade-off between the actual temperature profile and calibration accuracy. The main part of the paper is devoted to analysis of residual errors, explicitly defining an error model that well represents the actual residual errors. During the design and analysis of calibration under changing temperatures, it became evident that existing analysis and modeling methods do not support this approach. There are two essential topics that must be clarified:

- Deterministic versus stochastic part of the error model
- Time-invariant versus time-variant analysis approach

To proceed, a general form for a pre-calibration error model is introduced. Such a model can be used with any inertial sensor; for the sake of simplicity, a gyro-type notation is used.

$$\begin{aligned}\delta\omega(T,\omega,t) &= B(T)+\varepsilon_{d2d}+\varepsilon_{ins}(T,t)+\\ &+M(T,\omega)\omega+\varepsilon_{res}(T,\omega,t)\omega+\varepsilon_{RW}(t)\end{aligned} \quad (1)$$

The total sensor error is $\delta\omega(T,t)$, where T is temperature, ω is rate and t is time. The deterministic function $B(T)$ describes the deterministic thermal sensitivity of the drift (bias). The nonlinear (scale factor) deterministic thermal sensitivity error is given by $M(T,\omega)$. The residual errors, due to the no repeatable and instability portions of these sensitivities, are denoted $\varepsilon_{ins}(T,t),\varepsilon_{res}(T,\omega,t)$, respectively. Note that $\varepsilon_{ins}(T,t)$ combines the bias instabilities at constant and changing temperature. The constant (during continuous operation) day-to-day bias is ε_{d2d} and $\varepsilon_{RW}(t)$ is the random walk - for sampled data this is simply white noise.

During the calibration, estimations for $B(T),M(T,\omega)$ are calculated. Based on this estimation, the compensation is performed to reduce sensor error. The post-compensation error model is described in (2).

$$\begin{aligned}\delta\omega(T,\omega,t) &= \varepsilon_{d2d}+\varepsilon_{ins}(T,t)+\\ &+\varepsilon_{res}(T,\omega,t)\omega+\varepsilon_{RW}(t)\end{aligned} \quad (2)$$

Day-to-day error, ε_{d2d}, is defined as a constant; random walk is defined as white noise, and all other errors are collected into two residual terms: $\varepsilon_{ins}(T,t)$ observed at zero input rate, and $\varepsilon_{res}(T,\omega,t)\omega$, which are additional residual errors that appear only for non-zero rate. One can claim that these two residual terms combine the stochastic and deterministic parts one example of the deterministic part being errors in the calibration process.

Nevertheless, after a well-designed calibration process, the dominant part of the residual errors should be random. In the following, all residual terms will be treated as stochastic terms. The characterization of residual bias error, hereafter termed instability, $\varepsilon_{ins}(T,t)$, is a key element in both static and dynamic calibration design and evaluation. Moreover, this term is key for sensor integration algorithm design and for the integrated system's performance evaluation.

- The standard stabilized temperature approach yields a time-invariant error model, where frequency domain analyses such as PSD or the related Allan variance are well justified. However, when working with changing temperatures, a time-variant error model should be allowed, at least in the analytical phase. Our analytical approach is based on the following assumptions:
- The error model can change (slowly) with temperature. Therefore, during the analysis, the summation (integrals, expectations) calculations are limited to a narrow range of time (and temperature).
- A Gaussian distribution cannot be assumed, it must be verified. Deviations from a Gaussian distribution may occur with slow model changes due to temperature or sharp, singular-like thermal sensitivities.
- A time-invariant model should be found that bounds all underlying time-variant error models. Start with the basic model, the sum of first-order Gaussian-Markov processes and random walk. Try to optimize the analytical method with respect to the accuracy of the resulting error model.

The outcome is a reliable method of acquiring an error model. With this reliable method, the proposed calibration design, analysis and optimization all are model-based.

The ideas presented in this paper are general and can be applied to the whole range of sensor accuracies. However, some parts of the algorithms should be tuned to the actual error model. To describe the details of the approach, a concrete example must be presented: a high-end MEMS gyroscope was selected for this purpose. At the end, results of the actual calibration and residual error analysis will be presented. The tested gyro has, roughly speaking: in-run instability of 20º/h with 100 s time constant, and a random walk of 0.6º/sqrt(h). While working with changing temperatures, the temperature (typical, maximal) gradient used to characterize the gyro instability should be defined, based on the system's operational cycle. Virtually every system, over its entire operational life, encounters a variety of temperature gradients. Therefore, defining the temperature gradient for system calibration and testing may be tricky, sometimes necessitating the definition of more than one gradient. In any case, in the example described here, the temperature gradient is 0.6ºC/min. The temperature range is 0ºC–60ºC (roughly speaking, at gyro level).

The paper is structured as follows.

Section II is devoted to finding the proper analytical method. We start with the Allan variance and show that it is equivalent to prediction error calculations. As a result, several generalizations, based on different prediction and estimation schemes, are proposed. Then, the proposed schemes are compared for a particular error model with respect to their sensitivity to error model parameters. The selected method is based on non-model prediction and on model-based estimation; it is named DP type 2. Section III deals with static calibration design and optimization. Different temperature profiles are compared toward better calibration accuracy. Section IV shows the results of the actual test, and presents an in-depth discussion on error model bounding of actual residual errors, and Section V concludes.

II. ALLAN VARIANCE GENERALIZATIONS

In this section, the question of how to select a method to analyze residual (after thermal calibration) static errors is discussed. In the context of this paper, those errors are called instability, so we are seeking a method to obtain an instability model. The critical assumption is that the test is being carried out under changing temperatures. Our requirements from the desired analytical method are as follows:

1. Provide a method to define the underlying error model.
2. Provide good sensitivity properties with respect to error model parameters to ensure good error model accuracy.
3. Provide good sensitivity with respect to changes in the error model due to temperature changes, and in particular provide good ability to detect outliers (for example, singular-like behavior of sensor errors due to thermal changes).

The Allan variance, if implemented correctly, satisfies requirement 1. By proper implementation we do not mean analysis of local properties, such as slopes or minima. In these relatively short tests, the local-type analysis may be misleading. In addition, it is recommended that different error sources not be analyzed separately: in our experience, this error-separation effort, for cases of changing temperature, may produce poor results. By proper implementation we mean to assume an entire error model, to carry out its simulation and to calculate the Allan variance of the simulated process. If the Allan variance of the simulated process matches (bounds) the Allan variance of the actual data, the error model has been found.

We will show that some modification to the standard Allan variance can improve the fulfillment of requirements 2 and 3. For this purpose, the following notation is used. The sensor output is given by:

$$\{x(i) \quad 0 \le idt \le t_{test}\}$$

The sampling time is dt and data are collected during t_{test} seconds. The Allan variance formula is given by:

$$ARAV(\tau) = \frac{1}{\sqrt{2}} \underset{t_0}{rms}\{\Delta(t_0,\tau)\}$$

$$\Delta(t_0,\tau) = \frac{dt}{\tau} \sum_{t_0 \le idt < (t_0+\tau)} x(i) - \frac{dt}{\tau} \sum_{(t_0-\tau) \le jdt < t_0} x(j)$$

To improve the sensitivity properties (requirement 2), alternate forms of $\Delta(t_0,\tau)$ are proposed. They follow from the subsequent interpretation. Consider that $\Delta(t_0,\tau)$ is calculated at time t_0. Then the second sum describes the mean rate estimation over the past interval $[t_0-\tau_E, t_0)$, it is denoted $D^E(t_0,\tau_E)$, (for the time being $\tau = \tau_E$).

$$\Delta(t_0,\tau)=\frac{dt}{\tau}\sum_{t_0\le idt<(t_0+\tau)}x(i)-D^E(t_0,\tau_E)=$$

$$=\frac{1}{\tau}\left\{\sum_{t_0\le idt<(t_0+\tau)}x(i)dt-\tau D^E(t_0,\tau)\right\}$$

Letting $\tau_E\neq\tau$, the subsequent notation is introduced:

$$\Delta(t_0,\tau_P,\tau_E)=$$

$$=\frac{1}{\tau_P}\left\{\sum_{t_0\le idt<(t_0+\tau)}x(i)dt-\tau_P D^E(t_0,\tau_E)\right\}$$

Now, observe that $\tau_P D^E(t_0,\tau_E)$ is an attempt to predict the sensor output integral over the interval $[t_0,t_0+\tau_P)$ while the sum $\sum_{t_0\le idt<(t_0+\tau)}x(i)dt$ represents the sensor output integral over the same interval. Therefore, $\Delta(t_0,\tau_P,\tau_E)$ represents the prediction error of the sensor output integral, normalized by τ_P, or $\Delta(t_0,\tau_P,\tau_E)$ can be defined as the mean rate of prediction error. Summarizing this interim result, The Allan variance can be described by the following algorithm:

$$ARAV(\tau)=\frac{1}{\sqrt{2}}\underset{t_0}{rms}\left\{\Delta(t_0,\tau_P,\tau_E)\right\}\quad\tau=\tau_P=\tau_E$$

The mean rate of the prediction error of the sensor output integral is denoted $\Delta(t_0,\tau_P,\tau_E)$. The fact that the Allan variance is based on prediction error allows us to compare (and modify) this analytical method with well-known algorithms from linear dynamics systems, driven by random noise and used for system identification, modeling and parameter estimation (see, for example, [1] and [2]). In particular, this interpretation is essential for time-domain analyses.

In the context of system identification and parameter estimation, three important questions arise:

- Why work with equal estimation and prediction intervals?
- Why estimate using a simple mean value?
- Why predict using a simple multiplication?

In addition, when dealing with changing temperatures where singular thermal sensitivities might occur, a fourth significant question arises with respect to the use of the rms operator. The structure of the proposed Allan variance generalization, called DP, is as follows:

$$DP(\tau)=\frac{1}{\sqrt{2}}\sigma_{EST}\left\{DP_{sample}(\tau)\right\}$$

The statistical function σ_{EST} is designed to estimate the standard deviation of the sequence $DP_{sample}(\tau,t_0)$ with respect to observation time t_0. In the following, three statistical functions σ_{EST} are used:

$$\sigma_{EST}{}^{rms}\left(DP_{sample}(\tau)\right)=rms\left(DP_{sample}(\tau,t_0)\right)$$

$$\sigma_{EST}{}^{\max}\left(DP_{sample}(\tau)\right)=\frac{1}{3}\max\left(\left|DP_{sample}(\tau,t_0)\right|\right)$$

$$\sigma_{EST}{}^{95}\left(DP_{sample}(\tau)\right)=$$

$$=\frac{1}{2}percentile^{95.4\%}\left(\left|DP_{sample}(\tau,t_0)\right|\right)$$

To comply with requirement 3, i.e. to detect outliers and singularities, we propose considering the distance between 1 and the following ratios.

$$\rho_{rms}^{\max}(\tau)=\frac{\sigma_{EST}{}^{\max}\left(DP_{sample}(\tau)\right)}{\sigma_{EST}{}^{rms}\left(DP_{sample}(\tau)\right)}$$

$$\rho_{rms}^{95}(\tau)=\frac{\sigma_{EST}{}^{95}\left(DP_{sample}(\tau)\right)}{\sigma_{EST}{}^{rms}\left(DP_{sample}(\tau)\right)}\qquad(3)$$

As long as they are close to 1, the distribution is close to Gaussian and there are no significant outliers or singularities. Moreover, the residual thermal sensitivity errors can be well described by a time-invariant system. In this case, the operator $\sigma_{EST}{}^{rms}$ can be used to provide generally smooth results. If the ratio test shows values far from 1, our recommendation is to use the less smooth $\sigma_{EST}{}^{95}$ operator while monitoring $\sigma_{EST}{}^{\max}$, especially against failures. In the past, this type of test has been very powerful for finding problems in sensors (design and production) and in test set-ups, for example, electromagnetic interference or environmental disturbance.

The next step is to propose several alternatives to calculate the DP sample, $DP_{sample}(\tau)$, in order to optimize the sensitivity with respect to error model parameters (requirement 2). The considered alternatives are presented in the following list:

- Allan variance – in this case the estimation is based on mean values, the prediction is a simple multiplication of the estimated drift by the prediction time, and the estimation time is equal to the prediction time.

- DP (Direct Predictor) type 1 – the prediction and estimation schemes are the same as for the Allan variance, but the estimation time is constant.

- DP type 2 – the estimation is optimal for the nominal model; one can show that it is equivalent to a weighted mean with exponential forgetting factor, namely the "fresh" data have higher weight that the "old" data. The effective estimation time is related to the nominal model. The prediction is a simple multiplication as for the above methods.

- DP type 3 – the estimation is as for DP type 2 but the prediction is optimal (for the nominal model). In this case, the prediction is calculated using a weighted sum, with the weight decreasing for increasing times. One can show that this is equivalent to multiplication of the estimated drift by a factor related to the nominal model.

For all methods the prediction time is a free parameter.

While the Allan variance and DP type 1 can be implemented without any assumptions about the system model, DP types 2 and 3 require the use of a nominal error model for optimal estimation and prediction. Note that at this stage, the optimal estimation and prediction can be based on the nominal error model only, while the actual error model is unknown.

Before we select the most sensitive method, the method of error model selection is described. Let us denote $\{x_{data}\}$ as the actual data sequence (the residual errors) and $\{M_i\}_{i\in I}$ as the set of feasible models; the simulated data sequence, based on model M_i, is denoted $\{x_{M_i}\}$. The error model M_k is selected as the system error model if it satisfies two conditions:

1. $DP(x_{M_k})(\tau) \geq DP(x)(\tau) \ \forall \tau$
2. The bound is tight, namely $DP(x_{M_k})(\tau)$ is relatively close to $DP(x)(\tau)$ for all τ.

These conditions can be easily transformed into an optimization problem. However, we prefer to leave the informal definition, because at this stage we are only using a simply trial and error approach to find the selected error model.

The algorithm with the best sensitivity around the nominal model will be selected for subsequent analysis. The nominal error model is: random walk 0.6^o/sqrt(h) and the first-order Markov process has 20^o/h standard deviation and 100 s time constant. The test sensitivity is defined with respect to the process time constant, which will vary from 100 s to 50 s.

For this system, the applied temperature gradient is $0.6°C/\min$, the temperature range is $60°C$, so the test lasts 100 min. We will consider the DP results in the range of $2 \leq \tau \leq 600$. The lower limit is because the random walk in the high frequency analysis is omitted in this work; the upper limit is due to the limited number of samples.

First let us see the results of the Monte Carlo simulation (one test—6000s) for nominal error models.

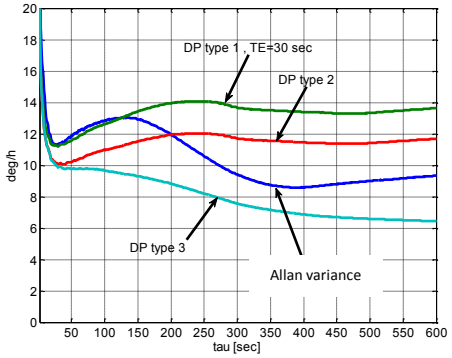

Fig. 1. Direct Predictor and Allan variance comparison for nominal model (Monte Carlo run #1)

Since all of the curves presented in Fig. 1 are interpreted as prediction errors, a method appears to be better if it achieves lower values. From this preliminary observation, one might conclude that DP type 3 is the best method. The Allan variance method is good for short and long times, but becomes poor for intermediate times (which are the most important for instability analyses!). Note that due to the relatively low dynamic range of τ, we do not use a logarithmic scale: for this case a linear scale is much more sensitive.

This is a good opportunity to discuss the common conjecture that one can easily find the instability of a system from the minimum of the Allan variance plot. In Fig. 1, one can find one minimum close to 20 s, and the second at around 350 s. It turns out that the second minimum is just a result of the particular

Monte Carlo run: see Fig. 2, which describes a different Monte Carlo run that has no minimum above 100 s.

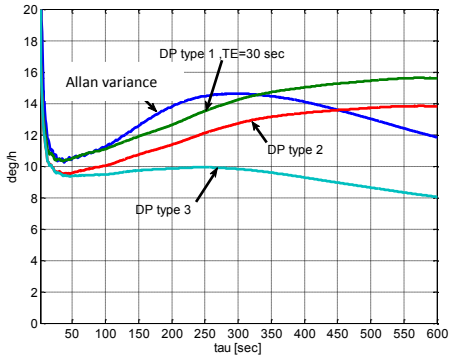

Fig. 2. Direct Predictor and Allan variance comparison for nominal model (Monte Carlo run #2

This observation—that some of the features observed in these plots are simply the result of a particular Monte Carlo run—leads to the conclusion that for a sensitivity study, we must distinguish between changes due to different error models and those due to a new random data sampling. For this purpose, let us consider 10 Monte Carlo runs, 5 with a nominal model and 5 with a different time constant (50 s instead of 100 s). Three different methods are compared: Allan variance, DP type 2 and DP type3. We are looking for the tau segment where the 5 Monte Carlo runs with the 100 s time constant are separated from the 5 runs with the 50 s time constant. This is called the separation segment.

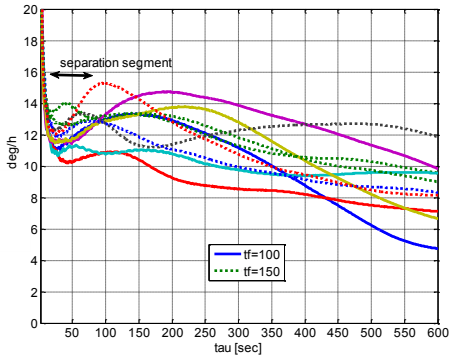

Fig.3. Allan variance separation.

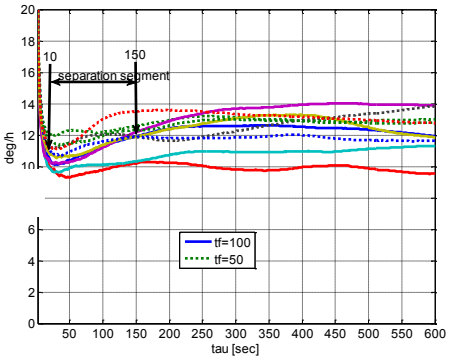

Fig. 4. Direct predictor type 2 separation.

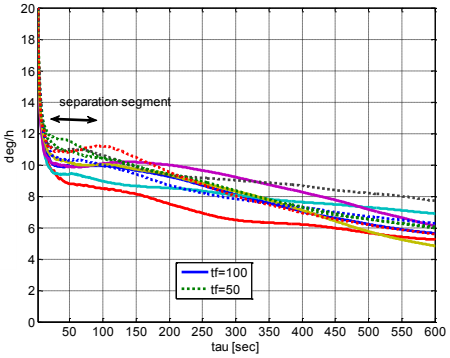

Fig. 5. Direct Predictor type 3 separation.

Note that most of the time, the DP values for time constants 100 and 50 s are mixed together. The length of the separation segment for the Allan variance is 80 s, for DP type 2 it is 140 s, and for DP type 3 it is 75 s. This shows that DP type 2 has significantly better sensitivity properties, at least in this case. From the above three sensitivity plots, one can reach the following conclusions:

- No method provides good separation above 200 s, and thus data above 200 s may be very misleading.
- The best separation is attained by DP type 2 in the range of 10–150 s, so this is the tau range within which conclusions about the error model should be made.

Another way of comparing different calculation schemes is by using the relative sensitivity measure

$$Sen(\tau) = \left| \frac{DP_{model1}(\tau) - DP_{model0}(\tau)}{DP_{model0}(\tau)} \right|$$

Figure 6 compares the relative sensitivities of the Allan variance, DP type 2, and DP type 3 in the range of 10–200 s. The calculations are for a long Monte Carlo (300,000 samples) run.

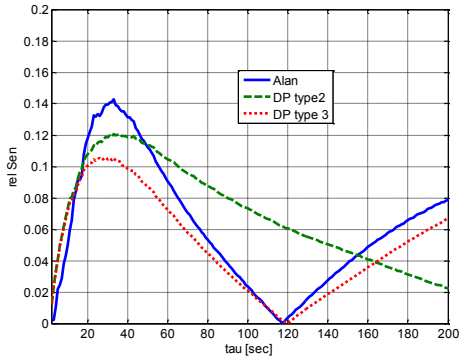

Fig. 6 Relative sensitivity to model time constant.

From Fig. 6 we see that DP type 2 significantly outperforms the Allan variance in the range of 50–200 s, while the Allan variance performs well in the short time range (15–50 s). Since the tau range within which the DP type has better sensitivity properties is wider than that for the Allan variance (the norm of selected sensitivity is larger), and DP type 2 has a significantly longer separation segment, DP type 2 is selected here as the major tool to analyze the error model. Surprisingly, despite being based on optimal calculations, DP type 3 has the least sensitivity with respect to the model time constant.

Although for the problems described in this paper, DP type 2 provides satisfactory results, in more advanced cases, a mix of DPs might be advantageous.

III. STATIC TEST DESIGN

In this section, the static calibration error is analyzed. Let us consider the static part of the sensor error model. By substituting $\omega = 0$ in (1) one gets:

$$\delta\omega(T,t) = B(T) + \varepsilon_{d2d} + \varepsilon_{ins}(t) + \varepsilon_{RW}(t)$$

The day-to-day element, ε_{d2d}, can be easily eliminated by removing the mean value from the data. For the problem discussed in this section, a third-order polynomial for thermal sensitivity $B(T)$ is assumed. In this case:

$$\delta\omega(T,t) = b_0 + b_1 T_n + b_2 T_n^{\,2} + \\ + b_3 T_n^{\,3} + \varepsilon_{ins}(t) + \varepsilon_{RW}(t)$$

The normalized temperature T_n is given by $T_n = \dfrac{T - T_0}{R_T}$. The normalizing parameters T_0, R_T should be selected such that $-1 \le T_n \le 1$. Since the error model is assumed to be known, and it fits into the structure of the linear Gaussian estimation problem, a Kalman filter will provide the optimal calibration results.

The proposed calibration scheme is based on a fifth-order Kalman filter with the state vector $\left[b_0 \quad b_1 \quad b_2 \quad b_3 \quad \varepsilon_{ins}(t)\right]^T$, and the measurement is the sensor output $\delta\omega(T,t)$. During the calibration, $b_0^{\,E}, b_1^{\,E}, b_2^{\,E}, b_3^{\,E}$ are calculated. The calibration error is defined as:

$$\delta c(T) = \sum_{i=0}^{3} \left(b_i - b_i^{\,E}\right) T_n^{\,i}$$

Observe that calibration error describes the drift error due to imperfect calibration. After the calibration, the calibration error is determinist in nature and will always appear in the operational cycle. This is the reason that they should be significantly smaller that the instability term.

After running the Kalman filter simulation for the selected scenario, the standard deviation of the calibration error (as a function of temperature T) is presented. To examine mismatch between the assumed and actual error models, a calculation-based truth-covariance analysis (see [2]) is introduced. Figure 7 describes the standard deviation of the calibration error for the temperature range 0ºC–60ºC and a temperature gradient of 0.6ºC/min. The error model is a random walk of 0.6°/sqrt(h) and the Markov process is 20°/h at 100 s.

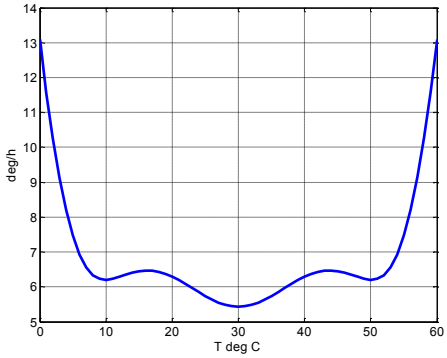

Fig. 7. Calibration error for the basic temperature range.

From Fig. 7, one can learn that in the range of 10ºC–50ºC, the error standard deviation is below 6.5º/h, while for the edges (0ºC and 60ºC), the error is close to 13º/h. These values should be compared with the 20º/h in-run instability. The question of whether these results are satisfactory is worth discussing. The RSS (root sum square) of in-run instability and calibration error is below 24º/h so on the one hand, the increase in total error is not severe. On the other hand, however, one must be aware that the calibration error is deterministic and cannot be filtered out. For those who do not like a 20%

increase in in-run instability or who plan to estimate and remove the in-run instability, let us discuss how to improve this calibration.

The first and simplest proposition is to wait at the extreme temperatures for some time. The proposed profile consists of staying in a "waiting period" at 0ºC, then applying 0.6ºC/min up to 60ºC, then again "waiting" at 60ºC. Figure 8 presents the calibration errors of such calibration designs, where the free parameter is the waiting time for each period.

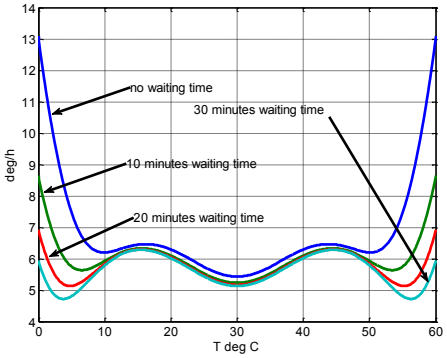

Fig. 8. Calibration errors for different waiting times at extreme temperatures.

From the plot, we see that indeed waiting periods reduce the calibration errors, especially near the extreme temperatures. To get close to a flat plot, a 30-min wait at every edge is required, while a waiting period of 20 min provides calibration errors below 7º/h. How to optimize the temperature profile to get minimum calibration error with respect to test duration appears to be a very interesting research question. At this time, we are not aware of any solution to this problem. From several attempts, the strategy of constant gradient inside the temperature range and waiting at extreme temperatures seems to be a good preliminary candidate for the optimum.

So let us define a profile with 0.6ºC/min from 0ºC–60ºC and 20 min waiting at the edges (140 min, all together) as a basic profile. Now we can apply it several times and see how performance improves.

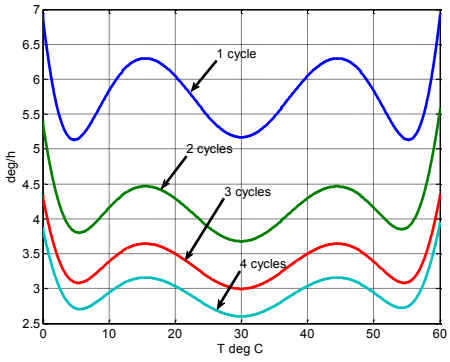

Fig. 9 Calibration error with respect to the number of calibration cycles.

From Fig. 9, one sees that the calibration error can be significantly reduced, at the cost of calibration time. The question of how many cycles to apply should be considered from the perspective of the required performance and additional errors, mainly errors in thermal modeling (thermal sensitivity which is described by a polynomial), repeatability and aging. In this context, the basic cycle, even without a waiting time, is usually satisfactory.

So far, we have presented the calibration error based on perfect knowledge of the instability error model. In real-life, this model is only known approximately. So a very important phase in model-based analysis is to check the algorithm sensitivity with respect to error model parameters. The following plot presents a truth covariance analysis for which the true model is the nominal one and the Kalman filter implemented during the calibration assumes different parameters: the time constant is changed to 30 and 250 s (instead of 100 s) and the standard deviation is changed to 5º/h and 40º/h (instead of 10º/h).

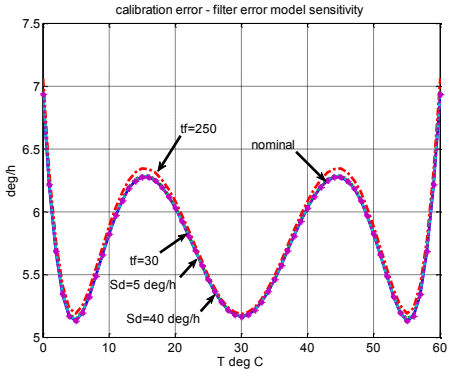

Fig. 10. Calibration errors with respect to filter model parameters.

The low sensitivity of the calibration error with respect to filter model parameters, as

presented in Fig. 10, is a solid justification for using a model-based approach to design and analyze the calibration procedure.

IV. A REAL-LIFE EXAMPLE

In this section, an example of an actual gyro calibration is described. The focus is on the post-calibration error model analysis. Nine MEMS gyro axes were exposed six times to similar temperature profiles. So as not to overwhelm the reader with excessive data, one gyro axis test was selected for detailed presentation.

The gyro temperature was changed within the range –10ºC to 60ºC; one test took almost 1.5 h, and the mean temperature gradient was about 0.6ºC/min. Figure 11 describes the temperature profile, as measured by the gyro temperature sensor.

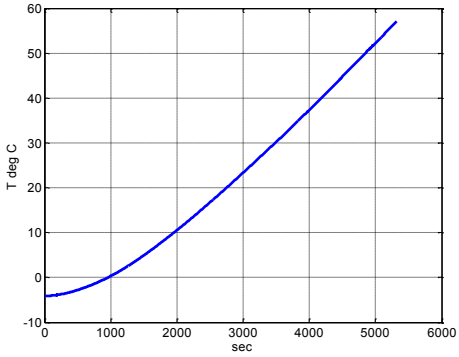

Fig. 11. The temperature profile.

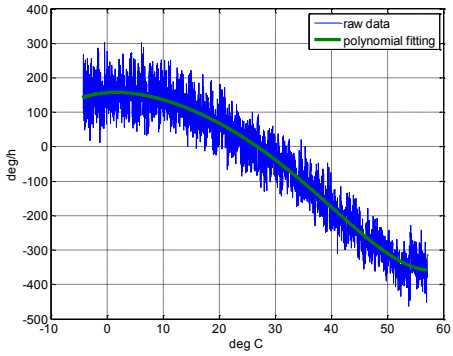

Fig. 12. The raw data and the calibration function

For every test, the preliminary data processing consisted of building 1-s output sums (to reduce the computation times) and removing the mean value to eliminate the influence of day-to-day gyro drift. As a preliminary calibration, a fifth-order polynomial fitting was performed. Figure 12 describes the raw data and the polynomial fitting as a function of temperature.

This preliminary polynomial fitting can be used for preliminary random walk and instability evaluations. A very good view into the system properties can be obtained looking at the smoothing of residual data (the difference between the raw data and the polynomial fit). This is shown in Fig. 13, versus both temperature and time.

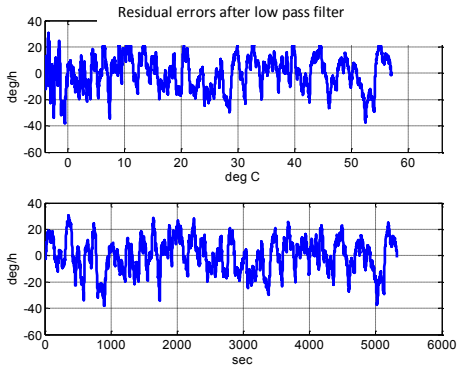

Fig.13. Gyro residual errors after passing low pass filter with respect to temperature and time.

From the above data, one can deduce that the instability is on the order of 20°/h with a typical time constant of 100 s. Note that this 20°/h is out of 600°/h that the thermal calibration corrected. The random walk can be calculated as a standard deviation of the difference between the raw and filtered residual data. The result is 0.66°/sqrt(h). In many cases, this type of analysis will conclude the required analysis. The goal of the work presented in the remaining part of this section is to detail, verify and justify calibration results and the model of residual error.

Two calibration methods: polynomial fitting and Kalman filter (based on the preliminary error model of 20°/h at 100 s and random walk of 0.6°/sqrt(h)) are compared. Both of them assume a fifth-order polynomial for drift temperature sensitivity. Figure 14 describes the difference in the compensated drift with respect to temperature. The difference is less than 3°/h, which, in this case, is negligible. So in this case, both methods are valid calibration methods.

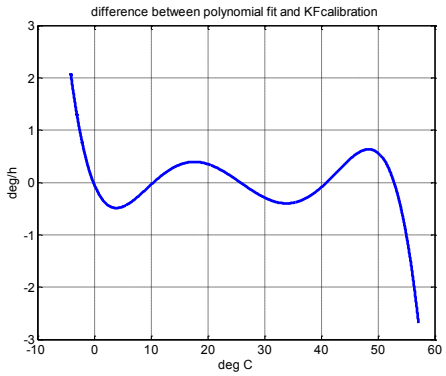

Fig. 14. Compensation differences between polynomial fitting and Kalman filter calibration

To proceed with the error model for residual post-calibration error, DP type 2 samples are evaluated by three different statistical operators: rms, maximum and 94.5th percentile. Figure 15 presents their ratios as defined by (3).

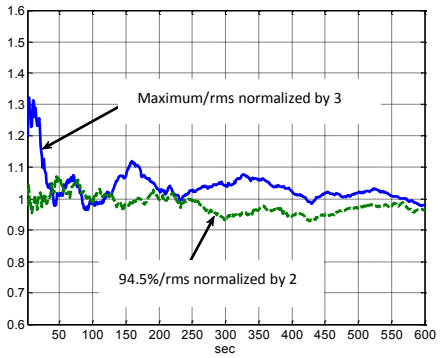

Fig. 15. The ratio between nomalized percentile and standard deviation.

For tau above 50 s, the ratios are in the range of 0.9–1.1, which justifies the Gaussian distribution application. The fact that $\rho_{rms}^{\max}(\tau)$ achieves 1.3 for τ around 20 s calls for some questions and clarifications. However at this stage, investigation of this problem (which is carried out currently) is beyond the scope of this paper. All subsequent analyses are based on rms calculations.

To gain proper insight into the problem, we compare the DP type 2 calculations for several cases: the actual residual error and three simulated calculations that differ in the time constant of the Markov process: 20, 60, and 100 s. All simulated processes have a random walk component of 0.66°/sqrt(h) and a Markov process with a standard deviation of 20°/h. Figure 16 depicts the results.

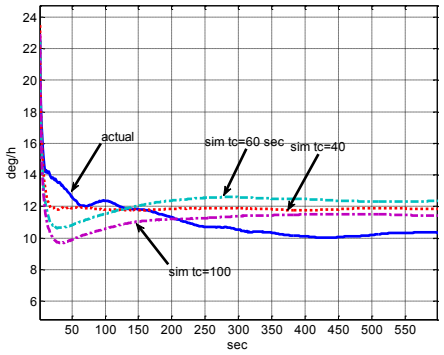

Fig. 16. DP type 2: actual and simulated.

At this stage, it is clear that none of the simulated processes bounds the entire DP type 2 function. The most significant gap is for low times, below 40 s. None of the presented Markov processes can close this gap without losing tightness for large times; however, this gap can be closed by increasing the random walk component. So, the next try will be with a random walk of 1.1°/sqrt(h) and we will compare four simulated processes with time constants of 20, 40, 60 and 100 s.

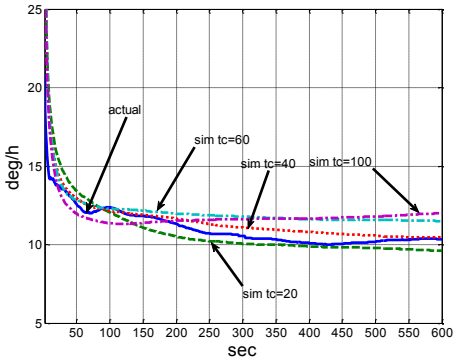

Fig.17. DP: actual and simulated (with random walk of1.1deg/sqrt(h)).

In Fig. 17 it is evident that two simulated signals with Markov processes based on time constants of 40 and 60 s bound the actual DP, while those based on time constants of 20 and 100 s certainly do not bound the actual DP: the first fails for long times and the second does not provide a bound for short times. Therefore, using the DP bound, a relatively good estimation of time constant was achieved.
Although the bounding error model has been found, the bound is not tight, especially for times above 200 s. Perhaps a mix of two Markov processes can provide a tighter bound.

After a few tries, the process with random walk of 1.1°/sqrt(h) and two Markov processes, both of them with standard deviation of 14°/h,

one with a time constant of 20 s and the second with a time constant of 70 s, provide a relatively tight bound, as shown in Fig. 18.

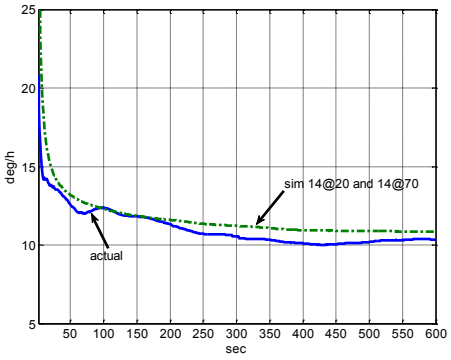

Fig.18. Actual and simulated DP: two Markov processes.

So far, the results of one gyro axis have been presented. Of course, the error model should be based on a large number of samples and should take into account phenomena such as non-repeatability and aging. This type of work forces us to present an error model for multiple test results. As an example of such a work, results of nine axis tests (taken from three IMU's that passed similar temperature profiles) are presented. Figure 19 describes DP (type 2) operating on residual error for nine axes. In addition, maximal and rms values (per tau) are depicted.

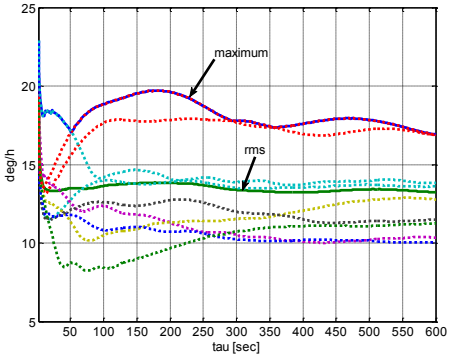

Fig.19. DP of nine residual errors

The next phase is to find one model to cover all nine axis tests; the solution, at this stage, is straightforward. The only remaining question is with respect to the most appropriate statistics (maximum of rms) to integrate all nine DPs. The answer is related to the explicit customer requirements and/or producer attitude (conservative versus risk-taking). In this paper, the error to bound the rms DP is presented. The selected error model consists of a random walk of 1.1°/sqrt(h) and in-run instability of 24°/h at 60 s. Figure 20 compares DPs: the actual one and the one obtained by simulation of the above-described process.

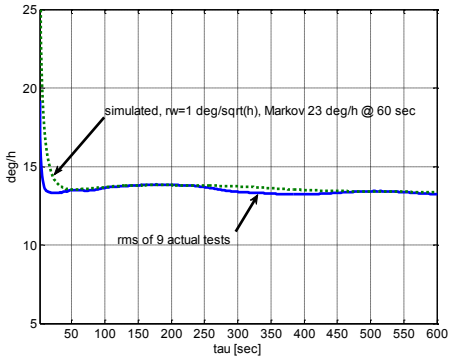

Fig.20. Rms of 9 test DPs and one simulated Markov process DP.

Observe that for times above 50 s, the proposed model provides an extremely tight bound for the actual data. For times below 50 s, our model is not yet perfect; presumably the proposed random walk should be modified by another random walk and fast Markov process.

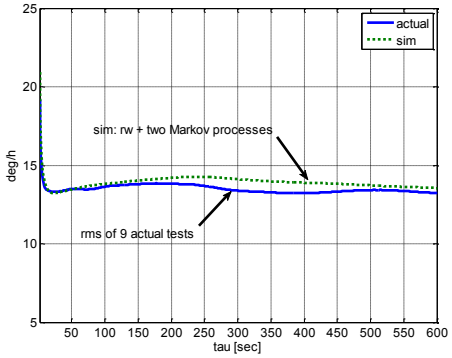

Fig.21. Rms of 9 test DPs and two simulated Markov process DPs.

In this case, the error model consists of a 0.6°/sqrt(h) random walk, one Markov process of 18°/h with a 25 s time constant and the second Markov process with standard deviation of 17°/h and a time constant of 60 s. It is evident that for low tau, the bound is very tight, and in the range of 200–400 it is less tight, but the difference is less than 1°/h, which can be accepted.

V. CONCLUSIONS

In this paper, static calibration with changing temperatures was discussed. Calibration error was described and ways to reduce it were proposed. Attention was directed toward finding a model for residual error. After describing the Allan variance as a type of prediction error calculation, several generalizations were proposed, and among these DP type 2 was selected for its highest

sensitivity. Thus the proposed analysis tool was DP type 2 with two enhancements:

- Comparison between three different statistical functions to estimate standard deviation (via rms, maximum, and 94.5th percentile).
- Comparison of DP values between actual data and data simulated using the proposed error model.

Then an actual example of a MEMS gyro was described and analyzed, with error model derivation for two cases: one for a single test and the second for an assembly of nine tests. The tightness of the models was improved using two Markov processes instead of one. Of course, one might ask "why bother with such subtle differences in the error model?" The answer is not related to algorithm design, where a rough error model is usually acceptable. However, such a precise error model, if implemented in performance evaluation simulation (generally called "truth model"), may cause differences in the evaluation results amounting to on the order of several tens of percentage points.

This paper gives rise to two interesting research questions which were circumvented here by using a trial and error approach. Such an approach can provide good insight into the system properties but it is time-consuming. The research questions that arose in the context of this work are:

- How to find the best error model using an optimization algorithm?
- How to find the best temperature profile for a given test duration? What is the performance limitation for a given test duration?

Such questions warrant further research.

A totally different issue is the question of industry standards: to what degree should the calibration procedures and analytical tools be published, discussed and eventually agreed upon for common, standard use?

REFERENCES

[1] L. Ljung, System Identification Theory for the User.: Prentice Hall Inc., 1987.

[2] A. Gelb, Applied Optimal Estimation: M.I.T. Press, 1974.

[3] Z. Berman, "Inertial sensors—a new approach for low cost calibration and testing," Inertial Sensors and Systems Symposium on Gyro Technology. Karlsruhe, Germany, 2011.

[4] P. Aggarwal, Z. Syed, X. Niu, and N. El-Sheimy, “A standard testing and calibration procedure for low cost MEMs inertial sensors and units,” J. Navigation, vol. 61, pp. 323–336, 2008.

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

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

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

[8] Z. Xing and D. Gebre-Egziabher, "Modeling and bounding low cost inertial sensor errors," Proceedings of IEEE/ION PLANS 2008, May 6–8, 2008.

[9] D. Gebre-Eqziabher, Design and Performance Analysis of a Low-Cost Aided Dead Reckoning Navigator. Dissertation, Stanford University, 2003.

[10] W. T. Fong, S. K. Ong, and A. Y. C. Nee, "Methods for in-field user calibration of an Inertial measurement unit without external equipment," Meas. Sci. Technol., vol. 19, 085202, 2008.

[11] A. G. Hayal, Static Calibration of the Tactical Grade Inertial Measurement Units. M Sc Thesis, Ohio State University, 2010.

[12] E. Nebot, and H. Durrant-Whyte, "Initial calibration and alignment of low-cost inertial navigation units for land vehicle applications," J. Robotic Sys., vol. 16, no. 2, pp. 81–92, 1999.