Cascaded Calibration of Mechatronic Systems via Bayesian Inference

Citation for published version (APA):

Document license:
CC BY-NC-ND

DOI:
10.1016/j.ifacol.2023.10.1489

Document status and date:
Published: 01/01/2023

Document Version:
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Cascaded Calibration of Mechatronic Systems via Bayesian Inference

Max van Meer* Emre Deniz* Gert Witvoet* Tom Oomen* TNO, Dept. of Optomechatronics, Delft, the Netherlands.

Abstract: Sensors in high-precision mechatronic systems require accurate calibration, which is achieved using test beds that, in turn, require even more accurate calibration. The aim of this paper is to develop a cascaded calibration method for position sensors of mechatronic systems while taking into account the variance of the calibration model of the test bed. The developed calibration method employs Gaussian Process regression to obtain a model of the position-dependent sensor inaccuracies by combining prior knowledge of the sensor with data using Bayesian inference. Monte Carlo simulations show that the developed calibration approach leads to significantly higher calibration accuracy when compared to alternative regression techniques, especially when the number of available calibration points is limited. The results indicate that more accurate calibration of position sensors is possible with fewer resources.

Keywords: Mechatronic systems, Bayesian methods, Calibration, Gaussian Process regression

1. INTRODUCTION

High-precision mechatronic systems rely on accurate position measurements to achieve high performance. At the same time, an increasing number of applications requires highly accurate position measurements in mass-produced systems, e.g., satellite swarms for optical communication (Gregory et al., 2010; Kramer et al., 2020), or segmented mirror telescopes (Nelson, 2006).

The accuracy of position measurements relates to their proximity to the actual positions, and precision refers to repeatability. This paper considers sensors that consistently exhibit position-dependent inaccuracies. These repeatable sensor inaccuracies can be measured using a test bed with a more accurate sensor, e.g., coordinate measurement machines (Takamasu et al., 1996) or optical-based test beds (Dresscher et al., 2019). When sensor inaccuracies are measured, a model is fitted to compensate for these inaccuracies through the process of calibration.

The test beds used to calibrate the position sensors of mechatronic systems require calibration themselves, to a standard regarded as an absolute measure of accuracy. This is done by a third party such as a metrology institute, e.g., satellite telescopes (Gregory et al., 2010; Kramer et al., 2020), or segmented mirror telescopes (Nelson, 2006).

Due to this cascade of calibration steps, depicted schematically in Fig. 1, modeling errors in individual calibration steps can stack and limit the achieved accuracy of the sensor calibrated last. Two leading causes of modeling errors are as follows. First, calibration on manual, external calibration instruments is time-consuming. While efforts have been made to partially automate the comparison of test bed sensor readings to accurate external readings (see, e.g., Wu and Wang (2013)), these methods introduce additional development cost and complexity. Consequently, the number of positions at which sensor readings are compared with those of more accurate sensors is limited. Second, some test bed locations may be unreachable to external instruments due to geometry constraints, further limiting the number of available calibration points.

In a parallel line of developments, the application of Gaussian Process regression to mechatronic systems has gained increased attention, see Poot et al. (2022); Rasmussen and Williams (2006), since it admits a highly flexible model structure while taking uncertainty into account using Bayes’ theorem. By specifying a prior that imposes properties such as smoothness and learning hyper-parameters from the data, a model is obtained that yields information not only of the expected function but also the variance of this function space.

Although regression techniques such as lookup tables can model individual functions well if the number of calibration points is large, it is shown in this paper that by taking into account the variance of individual calibration models using Bayesian inference, a significantly more accurate model is obtained, even if the number of calibration points is limited. Moreover, while sequential calibration using Bayesian inference has attained attention in the context of computer models (Jiang et al., 2020), the literature on the cascaded calibration of position sensors is sparse.

Therefore, the aim of this paper is to find an accurate mapping of position sensor readings to ‘true’ position values, i.e., sensor n in Fig. 1, while taking the uncertainty of the intermediate calibration model into account. The contributions of this paper are as follows:

* This work is part of the research programme VIDI with project number 19698, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO). In addition, this research has received funding from the ECSEL Joint Undertaking under grant agreement 101007311 (IMOCO4.E). The Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme.
Fig. 1. Schematic overview of the cascaded calibration problem. An array of sensors is calibrated on each other. Consequently, any imperfection in the calibration of the more accurate sensors is propagated down the chain to the less accurate sensors.

C1: A method for cascaded calibration of position sensors for mechatronic systems is developed. The approach takes the model uncertainty of the first calibration step into account to arrive at a more accurate estimation in the subsequent calibration step.

C2: The effectiveness of the approach is demonstrated through Monte Carlo simulations on a reproducible case study, and it is shown that the developed calibration method yields significantly more accurate models of the sensor offsets than alternatives such as lookup tables. The results indicate that more accurate calibration of mass-produced mechatronic systems is possible with fewer resources.

This paper is structured as follows. First, the problem description is given in Section 2. Next, the developed approach to cascaded calibration is explained in Section 3. Subsequently, simulation results are presented in Section 4, and finally, conclusions are drawn in Section 5.

2. PROBLEM DESCRIPTION

In this section, the problem description is given. First, a motivating example is given. Subsequently, the calibration problem is described, and finally, the problem is formalized in terms of two regression problems.

2.1 Motivating example

A motivating example of cascaded encoder calibration is shown in Fig. 2. The angular position sensor $S_1$ of a mechatronic system requires calibration, but the system is too compact to be accessible by the manual calibration instrument $S_3$, e.g., a theodolite or autocollimator (Gao et al., 2011). Hence, it is calibrated using a test bed with its own sensor $S_2$. Since the systems are aligned during calibration, more accurate sensors can observe the sensor inaccuracies of less accurate sensors.

The manual calibration instrument $S_3$ cannot measure all locations of the test bed because its frame physically obstructs access. Moreover, accurate manual calibration is labor-intensive, especially if high accuracy is required over the entire $360^\circ$ range of motion. Hence, the number of available calibration points of $S_3$ is limited.

Any imperfection in the calibration of the test bed to the manual instrument decreases the accuracy of the mechatronic system when it is calibrated on the test bed. This propagation of modeling errors motivates the need to take the uncertainty of the calibration model of the test bed into account when calibrating the mechatronic system.

2.2 Notation

The following notation is used. Sensor $S_i$, $i \in \{1, 2, 3\}$, is fixed to system $i$, where system 1 is the mechatronic system, system 2 is the test bed, and system 3 is the manual calibration instrument. Sensor $S_i$ yields measurement $y_i \in \mathbb{R}$. All positions are defined w.r.t. the same fixed reference frame. The true position of system $i$ is denoted by $y_i^*$. When systems $i$ and $j$ are aligned for calibration (as detailed in Assumption 2), it holds that $y_i^* = y_j^*$, and hence, measurements $y_i$ and $y_j$ are two different measurements of the same true position.

2.3 Cascaded calibration of sensors

Sensor $S_i$ is generally not perfectly accurate, i.e., $y_i \neq y_i^*$. Sensor inaccuracies can have multiple causes, e.g., manufacturing tolerances, mechanical wear, or imperfect assembly. These inaccuracies to the true position are position-dependent, expressed as follows:

$$y_i = f_{i \rightarrow 1}(y_i),$$

where $f_{i \rightarrow 1} : \mathbb{R} \rightarrow \mathbb{R}$ is a possibly nonlinear function, mapping inaccurate position measurements to true positions. Assumption 1. Measurements $y_i$ of $y_i^*$ are corrupted by zero-mean additive Gaussian white noise $\varepsilon$, with variance $\sigma_{n,i}^2$. Assumption 2. Assume $S_1$ is optically linked for different measurements of the same actual position when they are aligned for calibration.

Fig. 2. Motivating example. A mechatronic system with angular position sensor $S_1$ is optically linked for calibration with a test bed, with its own sensor $S_2$. The test bed itself is calibrated using a highly accurate manual measuring instrument $S_3$. Sensors $S_2$ yield different measurements $y_i$ of the same actual position when they are aligned for calibration.

Fig. 3. Sensor $S_1$ (•) is calibrated on a test bed with sensor $S_2$ (—), which, in turn, is calibrated on $S_3$ (—). Since the systems are aligned during calibration, more accurate sensors can observe the sensor inaccuracies of less accurate sensors.

If sensor $S_i$ of system $i$ is not perfectly accurate, then these inaccuracies are measured by a second system $j$, provided...
Procedure 1 Cascaded calibration of $S_1$ on $S_3$ via $S_2$

1: Align systems 2 and 3 and obtain a data-set $D_2 = \{y_2,k,y_3,k\}_{k=1}^{N_2}$. Use these observations of $y_3 = f_{2\rightarrow 3}(y_2)$ to fit a function $y_3 = f_{2\rightarrow 3}(y_2)$.

2: Align systems 1 and 2 and obtain a data-set $D_1 = \{y_1,k,y_2,k\}_{k=1}^{N_1}$. Then construct $D'_1 = \{y_1,k,f_{2\rightarrow 3}(y_2)\}_{k=1}^{N_1}$ using the model of Step 1, and use these ‘observations’ of $f_{2\rightarrow 3}(y_2)$ as a proxy for $S_i$ accessible by a manual calibration instrument. Therefore, the following assumption is made:

Assumption 2. When a pair $(y_i, y_j)$ of sensor readings is measured at a fixed point in time, it is assumed that misalignment errors are negligible w.r.t. sensor inaccuracies: $|y_i^j - y_j^i| \ll |y_i^j| - y_i^j$. (3)

Hence, both measurements $(y_i, y_j)$ are assumed to describe the same true position $y_i^j \approx y_j^i$ during calibration.

The manual calibration instrument $S_j$ is the most accurate sensor available, and therefore, $y_j$ can effectively be used as a proxy for $y_j^i$. The following function is then defined, assuming systems 3 and 1 are aligned:

$$y_3^i = f_{1\rightarrow 3}(y_1),$$

i.e., $f_{1\rightarrow 3}$ describes the relationship between a sensor reading $y_1$ and its ‘true’ position $y_3^i \approx y_3$.

Assumption 3. Each $f_{1\rightarrow 3}(y_1)$, $i \in \{1, 2\}$, is bijective, i.e., any value of $y_i$ maps to one value of $y_3^i$ and vice versa.

The aim is to obtain a model $\hat{f}_{1\rightarrow 3}$ of $f_{1\rightarrow 3}$, describing the sensor inaccuracy of $S_1$, but importantly, $f_{1\rightarrow 3}$ is never observed directly. In other words, the manual calibration instrument is not aligned with the mechatronic system for calibration, for two reasons: (i) it is economically not viable to perform manual calibration on many different mechatronic systems with sensor $S_1$, and (ii) the mechatronic system may be built too compactly to be physically accessible by a manual calibration instrument.

To prevent the need to calibrate $S_1$ on $S_3$ directly, the test bed with $S_2$ is first calibrated on $S_3$. Subsequently, $S_2$ can calibrate many different mechatronic systems with its own sensor $S_2$. These two steps are detailed in Procedure 1. When $\hat{f}_{1\rightarrow 3}$ is modeled offline through Procedure 1, it is used online to transform inaccurate position measurements $y_1$ to corrected measurements $\hat{f}_{1\rightarrow 3}(y_1)$.

Importantly, cascaded calibration requires making a fit on another fit. Since measurements of $S_2$ are labor-intensive and time-consuming, the first model $\hat{f}_{2\rightarrow 3}$ might be based on a limited amount of data ($N_2 << N_1$), and consequently, it may have a large variance. This potentially deteriorates the accuracy of $\hat{f}_{1\rightarrow 3}$ w.r.t. the true $f_{1\rightarrow 3}$.

In the next section, it is explained how the construction of $\hat{f}_{1\rightarrow 3}(y_1)$ from data through Steps 1-2 of Procedure 1 is framed as a series of regression problems.

2.4 Cascaded calibration through regression

To obtain a model $\hat{f}_{1\rightarrow 3}$ of the sensor inaccuracies of $S_1$ that can be used for calibration, the following cascade of regression problems is defined.

Problem 1. Consider Procedure 1, and suppose that data-sets $D_1$ and $D_2$ are available. Parameterize the models as $y_3 = f_{3\rightarrow 3}(y_1, \alpha)$ and $y_2 = f_{2\rightarrow 3}(y_2, \beta)$, respectively, and let their structures be fully determined by Hilbert spaces $K_1$ and $K_2$ (Wegman, 2006). The aim is to obtain the best possible fit of $\hat{f}_{1\rightarrow 3}$, even though $f_{1\rightarrow 3}$ is not measured directly, but instead by solving two sub-problems:

$$\min_{\beta^*, \alpha^*} J = \left(\frac{\int_{y_1}^{y_1^{\max}} (\hat{f}_{1\rightarrow 3}(y_1, \alpha^*) - f_{1\rightarrow 3}(y_1))^2 dy_1}{y_1^{\max} - y_1^{\min}}\right)^{\frac{1}{2}}$$

subject to

$$\beta^* = \arg\min \|\hat{f}_{2\rightarrow 3}(y_2, \beta) - f_{2\rightarrow 3}(y_2)\|_2^2 + \|\hat{f}_{2\rightarrow 3} - f_{2\rightarrow 3}\|_{K_2}^2,$$

$$\alpha^* = \arg\min \|\hat{f}_{3\rightarrow 3}(y_1, \alpha) - f_{3\rightarrow 3}(y_2, \beta^*)\|_2^2 + \|\hat{f}_{3\rightarrow 3}(y_1, \alpha) - f_{3\rightarrow 3}(y_2, \beta^*)\|_{K_1}^2,$$

where $y_1^{\min}$ and $y_1^{\max}$ specify a range of positions where a good model of the sensor inaccuracies of $S_1$ is desired. Note that solving for $\beta^*$ and for $\alpha^*$ amounts to constructing the fits of Steps 1 and 2 in Procedure 1, respectively.

The cost $J$ cannot be evaluated in practice because $S_1$ can only be compared to $S_2$, and $S_2$ to $S_3$, but not $S_1$ to $S_3$. On the other hand, $J$ can be evaluated in simulation, when $f_{1\rightarrow 3}$ is known. In the following sections, it is shown that for a specific choice of the structure of $f$ through $K_1$ and $K_2$, the cost $J$ is significantly reduced, when compared to conventional regression methods, indicating that more accurate calibration is achievable.

3. CASCADED CALIBRATION VIA BAYESIAN INFERENCE

In this section, the developed solution to Problem 1 is explained. The key idea is to recognize that the two sub-problems in (5) need to be posed in a coupled fashion. If these problems were solved independently, then any inaccuracy in $f_{2\rightarrow 3}(y_2)$ that follows from the fact that $y_3$ is only available at a limited number of positions is over-confidently carried over to $f_{1\rightarrow 2}(y_1)$.

Instead, the developed approach takes into account the uncertainty of $\hat{f}_{2\rightarrow 3}(y_2)$ at locations $y_2$ where no data of $y_3$ is available, through Bayes’ rule.

3.1 Calibration of $S_2$

First, sensor $S_2$ needs to be calibrated on sensor $S_3$; see the first sub-problem in (5). Given a limited number of observations of pairs $(y_2, y_3)$ showing the relation $y_3 = f_{2\rightarrow 3}(y_2)$, there is uncertainty in $\hat{f}_{2\rightarrow 3}(y_2, \beta)$ for values of $y_2$ that are far from calibrated locations. It is explained next how this model uncertainty, or variance, is computed explicitly so that it can be used for more accurate regression in the next section. To this end, a probabilistic viewpoint is adopted.

The model structure of $\hat{f}_{2\rightarrow 3}(y_2, \beta)$ is assumed to be

$$\hat{f}_{2\rightarrow 3}(y_2, \beta) = \phi_2(y_2)^\top \beta,$$
where $\phi_3 : \mathbb{R} \rightarrow \mathbb{R}^D$ maps any $y_2$ into some $D$-dimensional feature space, with weights $\beta$. A Gaussian prior is assumed on $\beta$, i.e.,

$$p(\beta) = \mathcal{N}(\beta_0, \Sigma_{\beta,\beta}),$$

with mean $\beta_0$ and prior variance $\Sigma_{\beta,\beta} \in \mathbb{R}^{D \times D}$. With this model, the likelihood $p(Y_3 | Y_2, \beta)$, or the probability density of the observations given the parameters, is given by

$$p(Y_3 | Y_2, \beta) = \mathcal{N}(\Phi_2(Y_2)^T \beta, \Sigma_{Y_3}),$$

where $\Phi_2(Y_i) := [\phi_i(\bar{y}_i, \ldots, \phi_i(\bar{y}_{N_i})]^T$, $\bar{y}_i \in [1, 2]$, and the variance of the observations is given by

$$\Sigma_{Y_3} = \sigma_{Y_3,n}^2 I.$$  

The prior on $\beta$ is conditioned on the data $D_2$ to obtain the posterior distribution, i.e., the probability of the parameters given the data. From Bayes’ rule, it is known that

$$p(\beta | Y_3, Y_2) = \frac{p(Y_3 | Y_2, \beta) p(\beta)}{p(Y_3 | Y_2)}.$$  

The expression (12) describes the posterior distribution of the parameters. For regression, the predictive distribution $p(Y_3 | Y_2, Y_3) = p(\hat{f}_{2 \rightarrow 3}(Y_2) | Y_2, Y_3)$ is of interest. For arbitrary sensor positions $Y_2 \in \mathbb{R}^{M_2}$, this predictive distribution is computed as

$$p(\hat{f}_{2 \rightarrow 3}(Y_2) | Y_2, Y_3) = \int p(Y_3 | Y_2, \beta) p(\beta | Y_2, Y_3)d\beta,$$

which is again a Gaussian with mean and variance

$$\mathbb{E} \left[ \hat{f}_{2 \rightarrow 3}(Y_2) \right] = m_2(Y_2) + K_2(Y_2, Y_2),$$

$$\text{cov}(\hat{f}_{2 \rightarrow 3}(Y_2)) = K_2(Y_2, Y_2) + \Sigma_{Y_3}^{-1}(Y_3 - m_2(Y_2)),$$

and the elements of $K_2(\bar{Y}_1, Y_1) = K_2(Y_1, \bar{Y}_1) \in \mathbb{R}^{N_2 \times M_2}$ and $K_2(Y_1, \bar{Y}_1) \in \mathbb{R}^{M_2 \times M_2}$ are obtained from evaluating a kernel function $k_i(y_A, y_B)$ for the corresponding values of arbitrary positions $Y_1$ and measurements $\bar{Y}_1$. This kernel has the property that $K_1 = \Phi_1^T \Sigma_{i,\beta} \Phi_i$, and thus relates to the chosen model structure or prior. The prior mean is denoted by $m_1 : \mathbb{R}^{M_1} \rightarrow \mathbb{R}^{M_1}$. See Section 3.3 for details on the choice of $k_i$ and $m_i$.

Crucially, (14) provides an analytic expression of the covariance of $\hat{f}_{2 \rightarrow 3}(Y_2)$, i.e., the uncertainty of the model, illustrated in light purple in the left of Fig. 4. This covariance is instrumental to obtaining a more accurate estimate $\hat{f}_{1 \rightarrow 3}$, as explained in the next section.

### 3.2 Calibration of $S_1$

In this section, a model $\hat{f}_{1 \rightarrow 3}$ is made, based on the model of $\hat{f}_{2 \rightarrow 3}$ obtained in the previous section. Since sensor readings $S_1$ can only be experimentally compared with $S_2$, but not to $S_3$, the model of $\hat{f}_{2 \rightarrow 3}$ is applied to $D_1 = \{\hat{y}_{2,k}, \hat{y}_{2,k}^N\}_{k=1}^N$, to obtain

$$D'_1 = \{\hat{y}_{1,k}, \hat{y}_{3,k}^N\}_{k=1}^N,$$

where

$$\hat{y}_{3,n} := \hat{f}_{2 \rightarrow 3}(\hat{y}_{2,n}),$$

$$\hat{y}_{3,n} := \hat{f}_{2 \rightarrow 3}(\hat{y}_{2,n}) = \hat{f}_{3 \rightarrow 3}(\hat{y}_{3,n}) \in \mathbb{R}^{N_2 \times N_2}.$$

The key insight that distinguishes the regression approach in this paper from traditional methods is that the variance of the ‘observations’ $y_3 = \hat{f}_{2 \rightarrow 3}(y_2) \in D_1$, i.e., the prediction of the model created in the previous section evaluated at measurements $y_2$, is affected by the uncertainty of the model. The covariance matrix corresponding to $\hat{Y}$ follows directly from this model uncertainty and is given by

$$\Sigma_{Y_3} := \text{cov}(\hat{f}_{2 \rightarrow 3}(\hat{Y}_2)).$$

which is computed directly through (14). By assuming a Gaussian prior on $f_{1 \rightarrow 3}$ as before and conditioning on $D'_1$, the predictive distribution $p(f_{1 \rightarrow 3})$ is a Gaussian with

$$\mathbb{E} \left[ f_{1 \rightarrow 3}(Y_1) \right] = m_1(Y_1) + K_1(Y_1, \hat{Y}_1)^T,$$

$$\text{cov}(f_{1 \rightarrow 3}(Y_1)) = K_1(Y_1, \hat{Y}_1) + \Sigma_{Y_3}^{-1}(Y_3 - m_1(\hat{Y}_1)),$$

Indeed, the posterior mean $\mathbb{E}[f_{1 \rightarrow 3}]$, shown in dashed blue in Fig. 4, is a function of $\Sigma_{Y_3} = \text{cov}(f_{2 \rightarrow 3})$, shown in light purple. Clearly, by taking into account the uncertainty of the model $f_{2 \rightarrow 3}$, the model $f_{1 \rightarrow 3}$ is affected. It is shown in Section 4 that this choice results in a more accurate model $f_{1 \rightarrow 3}$ than when the uncertainty is ignored. This concept is shown visually in Fig. 4, where the key observation is that $f_{1 \rightarrow 3}$ (dashed blue) does not rely on $y_3$ (purple dots) at locations where $f_{2 \rightarrow 3}$ has high variance (light purple), but instead relies more on its prior $\mathbb{E}[y_3] = y_1$.

Next, the choice of $k_i$ and $m_i$, which determine the prior of $f_{i \rightarrow j}$, is explained.

### 3.3 Selection of the model structure and hyper-parameters

In this section, the chosen model structure, or prior, is elaborated, and it is shown how the hyper-parameters are chosen automatically using data with empirical Bayes.

First, the choice of the prior is explained. It follows from Equations (14), (18) that the model structure of $f_{i \rightarrow j}$ is uniquely determined by choice of kernel function $k_i$, defining the prior variance, and the prior mean $m_i$.

For a function $y_j = f_{i \rightarrow j}(y_i)$, the prior mean is

$$m_i(y_i) := \mathbb{E}[y_i].$$

An intuitive choice is to pick $m_i(y_i) = y_i$, which assumes that in the absence of observations, sensor $S_i$ is expected to yield identical measurements as $S_i$. If prior information on the sensor accuracy is available, it can also be incorporated into $m_i$.

The prior variance, i.e., the range of possible functions that $f_{i \rightarrow j}$ can take, is determined by the kernel function. For an overview of possible kernel functions, including, e.g.,
polynomial or sinusoidal model structures, see Duvenaud (2014). In this paper, the attention is restricted to squared exponential (SE) kernel functions of the form

\[ k_i(y_i, y_n) = \sigma_f^2 \exp\left(-\frac{1}{2\ell_i^2} (y_i - y_n)^2\right), \]

where hyper-parameters \( \ell_i \) and \( \sigma_f^2 \) are the characteristic length scale and the magnitude of the prior variance, respectively. This model structure can be interpreted as imposing smoothness on \( f_1 \).

Hyper-parameters \( \Theta_i = \{ \ell_i, \sigma_f^2, \sigma_n^2 \} \) can be chosen from prior knowledge of the smoothness and magnitude of the sensor inaccuracies and noise. Alternatively, these can be learned from the data \( \mathcal{D}_i \), also known as empirical Bayes, by maximizing the log marginal likelihood, given by

\[
\log p(Y_j | Y_i, \Theta_i) = -\frac{1}{2} Y_j^\top K_i^{-1} Y_j - \frac{1}{2} \log |K_i| - \frac{N_i}{2} \log 2\pi,
\]

with \( K_i = K_i(Y_i, Y_i) + \Sigma_i \). This expression is maximized with respect to \( \Theta \) using an optimization algorithm for non-convex problems, to find optimal hyper-parameters.

3.4 Summary

The complete algorithm to obtain estimates of \( y_1 \) for arbitrary measurements \( y_i \) is summarized in Algorithm 1. Note that after following steps 1-5, step 6 can be repeated cheaply for any \( Y_i \), since the computation of \( \mathbb{E}[Y_i] = \mathbb{E}[\hat{f}_{1 \to 3}(Y_i)] \) through (18) is simply a matrix-vector multiplication once the inverted matrix in (18) is stored for future use.

Remark 1. The posterior means in (14) and (18) are identical to the solutions of the sub-problems in (5), see Rasmussen and Williams (2006, Section 6.2) for details.

4. RESULTS

In this section, the effectiveness of the developed calibration approach is demonstrated through Monte Carlo simulations. The simulation set-up is given first, and subsequently, the results are presented.

Fig. 4. Example of cascaded calibration via Bayesian inference. Sensor readings \( y_2 \) (——) and \( y_1 \) (—) have different position-dependent inaccuracies w.r.t. \( y_3 \approx y^* \) (—-), while the systems are physically aligned to \( y^* \). \( y_3 \) is measured at two locations (●) and a model \( \hat{f}_{2 \to 3}(y_2) \) (—, —) is fitted through these points. Next, this model is used to obtain \( \hat{y}_3 = \hat{f}_{2 \to 3}(\hat{y}_2) \) ( —, right) for many measurements \( \hat{y}_2 \). Finally, a model \( \hat{f}_{1 \to 3} \) ( —, —−) is fitted on \( \hat{y}_3 \). Because this model takes the uncertainty of \( \hat{f}_{2 \to 3} \) into account, the model \( \hat{f}_{1 \to 3} \) partially disregards the data at locations with high uncertainty, and instead relies more on its prior \( \mathbb{E}[y_3] = y_1 \), leading to a more accurate fit to \( f_{1 \to 3} \) (—).

Algorithm 1 Cascaded calibration via Bayesian inference

Require: Data-sets \( \mathcal{D}_1, \mathcal{D}_2 \), test points \( Y_1 \in \mathbb{R}^{M_1} \).

1. Specify kernel functions \( k_1, k_2 \) with initial hyper-parameters \( \Theta_{i,0} \), see Section (3.3).
2. Find optimal hyper-parameters \( \Theta_2 \) by maximization of \( \log p(Y_2 | \Theta_2), \) see (21).
3. Compute \( \mathbb{E}[Y_3] = \mathbb{E}[\hat{f}_{2 \to 3}(Y_2)] \in \mathbb{R}^{N_1} \) and \( \text{cov}(\hat{Y}_3) = \text{cov}(\hat{f}_{2 \to 3}(Y_2)) \in \mathbb{R}^{N_1 \times N_1} \) with (14).
4. Find optimal hyper-parameters \( \Theta_1 \) by maximization of \( \log p(Y_1 | \hat{Y}_3, \Theta_1), \) see (21).
5. Compute \( \mathbb{E}[\hat{Y}_3] = \mathbb{E}[\hat{f}_{1 \to 3}(Y_1)] \in \mathbb{R}^{M_1} \) with (18).
6. return \( \mathbb{E}[\hat{Y}_3] \)

4.1 Monte Carlo simulation set-up

Suppose the sensor measurements \( y_i \) obtained from sensor \( S_i \) during alignment \( (y_1^* = y_2 = y_3 = y^*) \) are described by

\[
y_3 = y^* + \varepsilon, \quad y_2 = y^* + \sum_{k=1}^{N_1} a_k \sin(\omega_{1,k} y^*) + b_k \cos(\omega_{1,k} y^*) + \varepsilon, \quad y_1 = y^* + \sum_{k=1}^{N_1} c_k \sin(\omega_{2,k} y^*) + d_k \cos(\omega_{2,k} y^*) + \varepsilon,
\]

with \( N_1 = 10 \) and \( \varepsilon \sim \mathcal{N}(0, 10^{-8}) \). The range of interest is \( y^* \in [0, 1] \) m, and hence, the values \( y_1^{\text{min}} \) and \( y_1^{\text{max}} \) in (5) follow from (22).

For the Monte Carlo simulations, \( N = 12000 \) different pairs of functions (22) are generated with \( a_k, b_k, c_k, d_k \sim \mathcal{N}(0, 10^{-4}) \) and \( \omega_{i,k} \sim \mathcal{N}(0, 6) \). The data sets are collected as follows. Data-set \( \mathcal{D}_1 \) is obtained by observing \( y_2 \) for an equally spaced grid of \( N_1 = 100 \) values of \( y_1 \). Subsequently, \( y_2 \) is observed for an equally spaced grid of 100 values of \( y_1 \), but then \( 10\% \) of the data on either edge and \( 20\% \) of the data in the center is removed, leading to \( N_2 = 64 \). This represents a scenario where \( S_2 \) cannot measure the test bed everywhere because it is physically obstructed.

Algorithm 1 is followed for all \( N \) functions. The accuracy of the resulting model \( \hat{f}_{1 \to 3} \) is then assessed by its cost \( J \) in (5) and compared with two alternative techniques:
4.2 Simulation results

The computational times of Algorithm 1 and Alternative 1 were less than five seconds each on a personal computer for each of the \( N = 12000 \) simulations. The results of the Monte Carlo simulations are shown in Fig. 5. The empirical probability distribution functions are normalized to have area 1. It is clear from the Fig. 5 that taking the model uncertainty of \( f_{2-3} \) into account through Algorithm 1 leads to a considerably better fit than when the uncertainty is ignored (Alternative 1). Both kernel-based methods result in more accurate models than a lookup table with linear interpolation (Alternative 2).

The results indicate that Algorithm 1 is a suitable solution to Problem 5, leading to better models of sensor inaccuracies than alternative approaches. With better models of sensor inaccuracies, more accurate calibration is achieved.

5. CONCLUSION AND RECOMMENDATIONS

A cascaded calibration method is developed to accurately model position-dependent inaccuracies of position sensors, enabling more accurate calibration of mass-produced systems in less time. By taking into account the uncertainty resulting from the first regression step using Bayesian inference, more accurate calibration is achieved than conventional methods such as lookup tables. The approach is especially advantageous when the number of calibration points is limited. Moreover, since the model hyperparameters are tuned automatically using the data, the procedure is convenient to implement in practice.

Future work is required to deal with cases when sensor readings are not fully repeatable. In the current framework, this is done by storing only average readings in the data sets, but a proper Bayesian treatment of this spread might further increase the achieved accuracy. Moreover, future efforts will be aimed at experimental validation of the method on the motivating example.

REFERENCES


