UNCECOMP 2023
5th ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Sciences and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Athens, Greece, 12—14 June 2023

EVALUATION OF MODEL BIAS IDENTIFICATION APPROACHES BASED ON BAYESIAN INFERENCE AND APPLICATIONS TO DIGITAL TWINS

Daniel Andrés Arcones[1], Martin Weiser[2], Faidon-Stelios Koutsourelakis[3] and Jörg F. Unger[1]

[1]Bundesanstalt für Materialforschung und -prüfung (BAM)
Unter den Eichen 87, 12205 Berlin, Germany
e-mail: {daniel.andres-arcones, joerg.unger}@bam.de

[2] Zuse Institute Berlin (ZIB)
Takustraße 7, 14195 Berlin, Germany
e-mail: weiser@zib.de

[3] Technical University of Munich (TUM)
Boltzmannstraße 15, 85748 Garching bei München, Germany
e-mail: p.s.koutsourelakis@tum.de

Keywords: Model bias, Bayesian Uncertainty Quantification, Digital Twins, Gaussian Processes, Statistical Finite Element Method

Abstract. *In recent years, the use of simulation-based digital twins for monitoring and assessment of complex mechanical systems has greatly expanded. Their potential to increase the information obtained from limited data makes them an invaluable tool for a broad range of real-world applications. Nonetheless, there usually exists a discrepancy between the predicted response and the measurements of the system once built. One of the main contributors to this difference in addition to miscalibrated model parameters is the model error. Quantifying this so-called model bias (as well as proper values for the model parameters) is critical for the reliable performance of digital twins. Model bias identification is ultimately an inverse problem where information from measurements is used to update the original model. Bayesian formulations can tackle this task. Including the model bias as a parameter to be inferred enables the use of a Bayesian framework to obtain a probability distribution that represents the uncertainty between the measurements and the model. Simultaneously, this procedure can be combined with a classic parameter updating scheme to account for the trainable parameters in the original model. This study evaluates the effectiveness of different model bias identification approaches based on Bayesian inference methods. This includes more classical approaches such as direct parameter estimation using MCMC in a Bayesian setup, as well as more recent proposals such as statFEM or orthogonal Gaussian Processes. Their potential use in digital twins, generalization capabilities, and computational cost is extensively analyzed.*

1 INTRODUCTION

Since the first introduction of the term in 2002 [1], the use of digital twins has gained considerable traction across multiple industries in recent years. Initially conceived as a digital representation of a physical object or system that is continuously updated with real-time data in the context of Product Lifecycle Management (PLM), nowadays digital twins are applied to many different fields, ranging from manufacturing process control to augmented reality (AR) or Internet of Things (IoT) paradigms. In the field of engineering, digital twins have become a valuable tool for the assessment and management of physical systems. Despite their rapid adoption, there is still much to be accomplished regarding homogenization and standardization [2]. Under the definition of digital twin, very distinct types of digital representations of physical objects are included. In this study, we will focus on digital twins that rely on a computational model to simulate the behaviour of the physical system under different conditions.

This type of digital twin offers an invaluable tool for the control of complex mechanical systems. A remarkably relevant application of such a model is the monitoring and assessment of large structures such as bridges [3, 4]. These systems typically provide few opportunities for the placement of sensors due to their cost and nature, especially in the case of already-standing structures. The potential of digital twins to increase the information obtained from limited data makes them an invaluable tool to provide predictions and evaluations for these critical systems in a non-intrusive manner. Additionally, the use of a simulation model for the virtual system allows for combining the previous knowledge of the physics that governs the real system with the measurement data obtained from the sensors to generate accurate predictions even where observations are not available.

A common first step in the implementation of simulation-based digital twins is fitting the virtual model parameters to a set of measurements. This is a classical problem where Bayesian inference excels [5]. Despite this fitting, there usually exists a discrepancy between the predicted response and the measurements of the system once built. One of the main contributors to this difference, in addition to potentially unfit model parameters, is the model error or model bias. It appears as a consequence of the inherent inability of the computer model to reflect the system's response in full. Among other causes, this can result from a limited set of equations to represent the physics of the system, or due to modelling assumptions necessary for the virtual representation. Every model that aims to replicate a real system is susceptible to bias independently of its complexity due to the choices and assumptions necessary for its definition. Therefore, quantifying the model bias is critical for the reliable performance of digital twins.

Model bias identification is ultimately an inverse problem where information from measurements is used to update the original model and to evaluate the fitness of the predictions simultaneously. Bayesian formulations can tackle this task. Including the model bias as a parameter to be inferred enables the use of a Bayesian framework to obtain a probability distribution that represents the uncertainty between the measurements and the model. Additionally, this procedure can be combined with a classic parameter updating scheme to account for the trainable parameters in the original model. Several approaches have arisen to tackle this problem, however, not all of them are equally suitable to be used in a digital twin framework.

Precisely, this study aims to evaluate the effectiveness of different model bias identification approaches based on Bayesian inference methods for simulation-based digital twins. The classical Bayesian parameter estimation using Markov chain Monte Carlo (MCMC) is taken as a basis. It will be compared with model bias identification approaches, such as Kennedy and O'Hagan's proposal in its original [6] and modularized formulation [7], orthogonal Gaussian

Processes (GP) [8] and the novel statFEM [9]. The potential use of these methods in digital twins, their generalization capabilities, and computational cost will also be extensively evaluated. The results of this study will help identify the most efficient and reliable approach for quantifying model bias in simulation-based digital twin applications, contributing to the homogenization and standardization of this valuable technology.

This paper first introduces the different approaches and their associated algorithms (Section 2), then applies them to a simple 1D benchmark case for its comparison (Section 3), and finally draws conclusions on the results (Section 4).

2 MODEL BIAS IDENTIFICATION APPROACHES

This section presents the different approaches to be analyzed in the context of digital twins.

2.1 Classical Bayesian inference without bias

The classical Bayesian inference framework is the basis for all the other approaches analyzed in this study. There exist numerous texts that describe it extensively (e.g. [5]), but here the focus is on a representative use-case as an implementation of a Monte Carlo-Markov Chain (MCMC) algorithm with Metropolis-Hastings selection criteria. The general workflow of such an implementation can be observed at Algorithm 1. The main objective is to apply Bayes' theorem to infer a set of parameters θ of a computational model f based on a set of real measurements y obtained at the points x. To do so, prior distributions are assigned to the parameters θ, together with a likelihood function, noise and correlation structures. This allows applying Bayes' theorem to obtain posterior distributions of the parameters. Through an MCMC approach, the stochastic process eventually converges to an estimation of the posterior distribution given the observed measurements.

Algorithm 1 Bayesian Inference for Parameter Estimation using Metropolis-Hastings MCMC

1: Define the forward model $f(x, \theta)$ that maps parameters θ at points x to measurements y
2: Define the prior distribution $p(\theta)$ that describes our beliefs about the values of θ before seeing the data
3: Define the likelihood function $p(y|\theta)$ that describes the probability of observing the data y given the parameters θ. The map is done through the forward model $f(x, \theta)$
4: Define the noise structure that characterizes the errors in the measurements y
5: Define the correlation structure that characterizes any dependencies or relationships between the measurements y
6: Initialize the MCMC sampler with an initial parameter value θ_0 and a desired number of posterior samples N
7: **for** $i \leftarrow 1$ to N **do**
8: Propose a new parameter value θ^* by sampling from a proposal distribution $q(\theta^*|\theta_{i-1})$
9: Compute the acceptance probability $A = \min\left(1, \frac{p(y|\theta^*)p(\theta^*)q(\theta_{i-1}|\theta^*)}{p(y|\theta_{i-1})p(\theta_{i-1})q(\theta|\theta_{i-1})}\right)$
10: Sample a random number u from a uniform distribution on $[0, 1]$
11: Apply selection criteria (e.g. Metropolis-Hastings)
12: **end for**
13: Compute the posterior distribution $p(\theta|y)$ using the set of posterior samples $\theta_{i\,i=1}^{N}$
14: Compute summary statistics and/or generate plots to analyze the posterior distribution $p(\theta|y)$

The following requirements must be fulfilled to implement a classical Bayesian inference approach without bias on the set of latent parameters θ using a sampling procedure such as Metropolis-Hastings MCMC:

- the full description of the forward model must be specified beforehand.
- The forward model must be evaluated at every sampling point. If it is computationally expensive, a surrogate model can be generated. Alternatively, a different inference procedure can be implemented.
- Information of the prior distributions of θ must be available.
- Knowledge about the correlation and noise structures for the measurements must be available.

These requirements present a significant set of challenges to the simplest version of this framework in the context of digital twins. First, real-world systems are extremely complex, therefore modelling assumptions are necessary to generate computer models that can provide useful information in a reasonable setting. This approach acts on the parameters of such models, which will lead inevitably to errors derived from the modelling assumptions. This happens both if the description of the forward model is incomplete or if we consider it fully detailed. If a surrogate is introduced as a solution for expensive models, further deviations from the true system must be expected. Furthermore, digital twins often require almost-real time response in their predictions as a source of information, which can only be achieved if the evaluation of the forward model with the estimated parameters is performed in a reasonable time, giving more arguments for a surrogate approach.

Additionally, reliable information on the prior distributions of the latent parameters may not be available for the real model. Analogously, noise and correlation structures are often inferred from the model and the measurements themselves, as they are usually not known beforehand, increasing the number of potential sources of bias that the approach does not account for. These assumptions on the prior distributions have a notorious impact on the performance of the model, potentially increasing or reducing the uncertainty and accuracy of the predictions. Inverse problems like this, especially when the number of latent parameters grows, tend to be ill-posed, which is at its time affected by the choice of priors.

Nevertheless, such an approach can be useful for digital twins where enough reliable information about the real system is available. It allows estimating the posterior distribution of the latent parameters to be inferred, which at its time enables generating easily predictions on the posterior distribution of the system's response. The prediction time and computational costs depend exclusively on the definition of the forward model, which can be controlled through the use of surrogates. When the latent parameters have physical meaning, e.g. they represent the material parameters of a structure, knowing their posterior distribution may be critical for identifying potential failures or damages in the real system.

2.2 Kennedy and O'Hagan (KOH)

Kennedy and O'Hagan's approach [6] has risen as the canonical method for introducing bias quantification in a Bayesian inference procedure. The main difference is the inclusion of a bias term b in the comparison with the measurements to account for the part of the real system that

the computer model cannot reliably represent. Therefore, for an set of parameters θ and a noise term e:

$$y(x) = f(x, \theta) + b(x) + e \tag{1}$$

Notice how the bias b does not depend on the parameters θ. In effect, the definition of the bias as a measurement of the model's lack of fitness is only possible when considering the optimal set of θ parameters. Kennedy and O'Hagan suggest representing this bias term as a Gaussian Process (GP) defined by a mean μ_b and a covariance function $k(x, x')$, which using a stationary kernel function can be defined through a standard deviation σ_b and a length-scale ℓ_b. The choice of kernel depends on the system to study. Coincidentally, they recommend surrogating f by a GP in an analogous way. However, f depends on θ, which are to be inferred and vary during the sampling procedure. Therefore, a training dataset Y_M must be generated with combinations of (x, θ) parameters to train it. To generate such a dataset efficiently, a Design of Experiments (DoE) approach, such as Latin Hypercube sampling, can be followed. Algorithm 2 adapts the workflow proposed for this approach. During the inference procedure, following for example the previous MCMC scheme, both the GPs for b and f are fitted to the set of measurements $Y = Y_M + Y_F$, where Y_F are the real measurements. This is done in the same process as the latent parameters θ (and eventually the noise term σ) are estimated. Substituting these parameters in Equation 1, it is possible to regenerate the full biased response.

Algorithm 2 Kennedy and O'Hagan's (2001) approach for Bayesian inference with model bias

1: Approximate the computer model $f(x, \theta)$ as a GP with mean function $\mu_f(x, \theta)$ and covariance function $k((x, \theta), (x', \theta'))$ with standard deviation σ_f and kernel function with lengthscale ℓ_f
2: Generate model samples Y_M from the original computer model f evaluations design of experiments (DOE) methods that cover evenly the space (x, θ)
3: Define the bias model $b(x)$ as a GP with mean function $\mu_b(x)$ and covariance function $k_b(x, x')$ with standard deviation σ_b and kernel function with length-scale ℓ_b
4: Define the joint model as $y(x) = f(x, \theta) + b(x) + e$ with the best fit at θ^* giving $y(x) = f(x, \theta^*) + b_\theta(x) + \sigma^2 I$, where σ is a noise term. Define its likelihood, noise and correlation models
5: Use Y_M and original samples Y_F to fit the parameters θ, and e using MCMC at the same time as the GPs for f and b
6: Being $Y = (Y_M, Y_F)^T$, compute the posterior distribution $p(\theta|Y)$ using the set of posterior samples obtained from the MCMC sampler
7: Compute summary statistics and/or generate plots to analyze the posterior distribution $p(\theta|Y)$
8: Regenerate the full biased response $f(x, \theta^*) + b_\theta(x)$

The only additional requirement when compared with a classical Bayesian inference approach is the need of defining the GP for the bias term. This method incurs the same pitfalls of ill-posedness and extreme dependency on the prior's choice. In the case of KOH, the identifiability issues are even more critical. Inaccurately defined system priors and inference procedures may tend to compensate for the model error by increasing the bias instead of modifying θ, which leads to posterior distributions with increased variance. Alternatively, the recommendation of substituting the computational model with a surrogate is often already implemented in the context of simulation-based digital twins in order to speed up the response times of the

virtual representation. Additionally, information on the distribution of the bias can be used by the user to identify deficiencies in the virtual model and its assumptions.

2.3 Modularized KOH

Despite its status as a seminar paper [10], criticism of Kennedy and O'Hagan's approach has arisen over the years. One of the main points of contention is the choice of inferring concurrently the GPs for the computational model f with the bias b and latent parameters θ. As exposed by Bayarri in [7], this approach potentially leads to the contamination of the whole inference process due to badly defined parts of it. The proposed alternative is a modularization of Kennedy and O'Hagan's approach by separating the training of the GP for the computational model and the inference on the bias and the latent parameters. An algorithm for such an approach can be observed in Algorithm 3.

Algorithm 3 Modularized version of Kennedy and O'Hagan by Bayarri (2009)

1: Approximate the computer model $f(x, \theta)$ as a GP with mean function $\mu_f(x, \theta)$ and kernel function $k((x, \theta), (x', \theta'))$
2: Generate model samples Y_M from the original computer model f evaluations design of experiments (DOE) methods that cover evenly the space (x, θ)
3: Fit the GP for $f(x, \theta)$ using Y_M, for example using Maximum a Posteriori (MAP) estimation
4: Define the bias model $b(x)$ as a GP on the residuals $R = Y_F - f(X, \theta)$ for a given evaluation of θ at measurement positions X
5: Use only original samples Y_F to fit the parameters of b, θ, and σ using MCMC
6: Compute the posterior distribution $p(\theta|Y_F)$ using the set of posterior samples obtained from the MCMC sampler
7: Compute summary statistics and/or generate plots to analyze the posterior distribution $p(\theta|Y_F)$
8: Regenerate the full biased response $f(x, \theta^*) + b_\theta(x)$

The modularized version of Kennedy and O'Hagan's approach does not require additional information on the real system, while exhibiting a more robust behaviour. According to Bayarri [7], modularization is a good strategy to isolate good modules from suspect ones to avoid contamination. This is especially relevant when applying the approach to an already existing virtual representation of a digital twin. For example, with a modularized approach, the surrogate of the computational model is trained independently from the inference procedure, which allows using seamlessly an already existing surrogate. Additionally, this compartmentalization can be advantageous in the context of digital twins, as any of the modules can be improved or modified as needed without having to modify the others. Understanding digital twins as tools under constant revision throughout the life of the real system, they are expected to require updates and changes in the virtual modules. Purely from the point of view of estimating the model bias, the modularized approach results advantageous when dealing with identifiability and confounding issues present in the original one. Due to these advantages, the modularized version has been adopted as the preferred implementation of Kennedy and O'Hagan's approach to model bias in Bayesian inference problems.

2.4 Orthogonal Gaussian Processes

Orthogonal Gaussian Processes are formulated by Plumlee [8] as a response to the identifiability problems exhibited by Kennedy and O'Hagan's approach with respect to the model bias. To avoid that during the inference procedure the bias is overestimated, Plumlee established an orthogonality restriction such that the GP for the bias b represents exclusively the part of the measurements that cannot be explained by the best-fitted computer model f. The orthogonality condition is defined with respect to a given loss function on the residuals between measurements and model observations. This condition is implemented as a modification on the prior of the GP for the bias. The new parameters for the GP prior defined as $b(x) \sim GP(\mu_b, C_b)$ are:

$$\mu_{b_\theta}(x) = w(x)^T M_\theta (M_\theta^T W M_\theta)^{-1} Q_\theta \tag{2}$$

and

$$C_{b_\theta}(x, x') = k_b(x, x') - w(x)^T M_\theta (M_\theta^T W M_\theta)^{-1} M_\theta^T w(x') \tag{3}$$

where k_b is the original correlation function prescribed by a kernel and a standard deviation, W is the realization of k_b over the set of measurements ξ where f is evaluated, w is the correlation vector obtained from evaluating k_b between x and ξ, M is a matrix of the derivatives of the mean μ_f of the GP of f with respect of the parameters θ and Q a vector of the sum of the derivatives of the variances $v_f(\xi, \theta) = k((\xi, \theta), (\xi, \theta))$ of the GP of f with respect to each latent parameters. Details on the formulation can be observed in [8]. With these modifications, Algorithm 4 represents the workflow for the implementation of the inference procedure using orthogonal GPs.

Algorithm 4 Orthogonal Gaussian Processes approach by Plumlee (2017)

1: Approximate the computer model $f(x, \theta)$ as a GP with mean function $\mu_f(x, \theta)$ and kernel function $k((x, \theta), (x', \theta'))$
2: Generate model samples Y_M from the original computer model f evaluations design of experiments (DOE) methods that cover evenly the space (x, θ)
3: Fit the GP for $f(x, \theta)$ using Y_M, for example using Maximum a Posteriori (MAP) estimation
4: Define the bias model $b(x)$ as a GP on the residuals $R = Y_F - f(X, \theta)$ for a given evaluation of θ
5: Impose an orthogonal prior to the GP of b by defining $b(x) \sim GP(\mu_b, C_b)$ where $\mu_{b_\theta}(x) = w(x)^T M_\theta (M_\theta^T W M_\theta)^{-1} Q_\theta$ and $C_{b_\theta}(x, x') = k_b(x, x') - w(x)^T M_\theta (M_\theta^T W M_\theta)^{-1} M_\theta^T w(x')$ (see Equations 2 and 3)
6: Use only original samples Y_F to fit the parameters of b, θ, and σ using MCMC
7: Compute the posterior distribution $p(\theta|Y_F)$ using the set of posterior samples obtained from the MCMC sampler
8: Compute summary statistics and/or generate plots to analyze the posterior distribution $p(\theta|Y_F)$
9: Regenerate the full biased response $f(x, \theta^*) + b_\theta(x)$

The main difference regarding the implementation of the inference procedure between Plumlee's and Kennedy and O'Hagan's approach is the requirement of the derivatives of the distribution of f with respect to the latent parameters. If f is expressed as a GP surrogate, that only involves a slight extra computational effort, as they are usually already known and calculated

in the inference step. Other computational models may present themselves as more challenging for this approach, but it should always be possible to apply finite differences differentiation during the sampling procedure. In exchange for the extra computational cost, the inferred latent parameters present reduced variability than with the original approach, as shown in [8]. This allows for more precise predictions and reduces the identifiability issue present in the inverse problem. Regarding its integration in a digital twin, the resources are only additionally needed in the inference step, therefore any other part of the implementation is analogous to the original approach.

2.5 StatFEM

One of the most recent approaches for adding model bias to a more complex computational model is the so-called statistical Finite Element Method (statFEM) developed by Girolami et al. [9]. Contrary to the classical Bayesian approaches, the objective of statFEM is to estimate the system's response u from a Finite Element (FE) approximation instead of inferring the parameters θ that are present in such a model. This allows the use of inaccurate models and assumptions, as the modelling choices have a smaller impact on the inference results. However, this is possible only under strong impositions and assumptions on the nature of the statistical model. Respecting the notation from [9] such statistical model is defined as:

$$y = z + e = \rho P u + d + e \tag{4}$$

where y is the set of real measurements, z is the real underlying system, u is the system's FE response, P is the projection matrix from u to the measurements space, ρ is a scalar regression term, d is a bias term represented as a zero-mean GP with covariance defined by a stationary kernel, standard deviation σ_d and length-scale ℓ_d, and e a noise term. Supposing normality in all the terms in the probabilistic sense, it is possible to apply Bayes' theorem to the statistical model and obtain directly the posterior distributions of $u|y$ and $z|y$ from the prior distribution of u and the statistical hyperparameters $(\rho, \sigma_d, \ell_d, e)$.

Let u be a stochastic FEM model $u(\kappa, f)$ dependant on the parameters κ and the source terms f such that it is governed by a linear PDE. Then, the solution of the PDE can be expressed as $A(\kappa)u = f$, where A is the linear transformation matrix from the discretization of u. As either κ or f are defined as stochastic variables and A is a linear transformation, it is possible to obtain a prior distribution of u from the imposed prior distributions of κ or f. Given that all components of Equation 4 are modelled as normal (either multivariate normal distributions or GPs), it is possible to calculate analytically $\mu_{u|y}$ and $C_{u|y}$. Therefore, using this analytical expression of the posterior of u for a given set of statistical hyperparameters, it is possible to apply an inference procedure to those hyperparameters as a first step. Once their posterior distributions are obtained, a set of estimators (e.g. maximum likelihood estimators) can be input in the expressions for $\mu_{u|y}$ and $C_{u|y}$ to recover a posterior distribution of the response field. This procedure is shown in Algorithm 5.

Algorithm 5 StatFEM by Girolami et al. (2019)

Require: Assume all variables can be modelled as Gaussians

1: Define the statistical model $y = z + e = \rho Pu + d + e$
2: Define the stochastic FEM model $u(\kappa, f)$ by defining a probability distribution on the source term f or on the parameters κ.

Require: Linearity of the PDE that defines u, giving $A(\kappa)u = f$

3: Apply the perturbation method (1st order Taylor series expansion of u) to obtain the mean μ_u and covariance matrix C_u for the prior distribution of u modelled as a Gaussian
4: Calculate analytically $\mu_{u|y}$ and $C_{u|y}$ as a function of ρ, σ_d and l_d. This is possible because of the assumption of normality
5: Infer ρ, σ_d and ℓ_d from y using the analytical posterior distribution of u for a given set of statistical hyperparameters
6: Use (point) estimators of ρ, σ_d and ℓ_d to calculate $p(u|y)$ and $p(z|y)$ at the data points.
7: Predict values of the fitted distributions at points without data.

One of the main advantages of statFEM is the possibility of using deficient or uncertain models, as is typically the case with representations of real systems. The variability is introduced by the probability model of the system's components, which allows the inclusion of previous information into the model. Nevertheless, a defective estimation of this variability may lead to an overestimation of the bias term, as with traditional Bayesian approaches. Additionally, the FE model is only evaluated when calculating the prior distribution of the response field, making this a very well-suited approach for expensive models that would require complex surrogates. However, this evaluation of the prior scales very poorly with the number of variables and data positions, which can hinder its performance in some applications. Each new data position where predictions are requested requires additional evaluations of the FE model, therefore it would not be recommended for digital twins where the location of the predictions is not known beforehand or a large set of predictive data positions is required. Nevertheless, new data points at known positions can be added to the dataset easily, enabling its training to represent the new conditions of the system. It must be noted that no information on the computational model parameters is obtained, which can be undesirable for some digital twin applications that rely on the analysis of such parameters.

Arguably, excluding the number of data positions and variables, the main limitations come from the type of systems that can be treated with the current formulation. The original implementation of statFEM requires a linear static PDE and assuming normality in the response field. Duffin [11] modified statFEM to deal with non-linear time-dependent problems at the expense of computational speed, which may limit its applicability to digital twins that require a fast response. Narouie et al. [12] avoided the limitation of linearity in the PDE by applying a Polynomial Chaos Expansion of u to obtain its prior distributions instead of the original perturbation method with a 1st-order Taylor expansion. This approach introduces a further approximation on u, increasing the uncertainty or the computational cost depending on the number of modes to be evaluated from the PCE. In any case, the ability to infer the response of a model independently of its assumptions and additionally include information on the bias makes statFEM a promising approach for specific digital twins that can be accommodated to its requirements.

3 APPLICATION AND RESULTS

The chosen benchmark is the "Pedagogical example" presented in [8]. It is a simple 1D case that should be suitable for the different approaches. Taking $X = [0, 1]$, the original measurements are taken from a "true" model governed by $y(x) = 4x + x \sin 5x$, while the computer model is $f(x, t) = tx$. Measurements are collected for $\hat{X} = [0, 0.8]$ and are perturbated with Gaussian noise of variance $v^2 = 0.02$. The objective is to fit the variable t such that the computer model reflects the measurements. There is a clear bias between y and f, therefore it can potentially be quantified. In the original paper, the authors limit themselves to a reduced comparison between the original KOH and orthogonal GPs with norms L_{L^2} and $L_{W_1^2}$. In this study, we extend the analysis to a modularized version of KOH and statFEM, focussing on the applicability of each method to the implementation of a digital twin.

3.1 Bias-free Bayesian inference

For a bias-free approach, a normal prior with mean 3.0 and standard deviation 1.0 is prescribed for t. Then, the MCMC algorithm from Section 2.1 is applied for 2000 steps after 200 more of burn-in. The software package `probeye` is used to perform the inference procedure. The results can be observed in Figure 1. As expected, the best-fitted model, though certain,

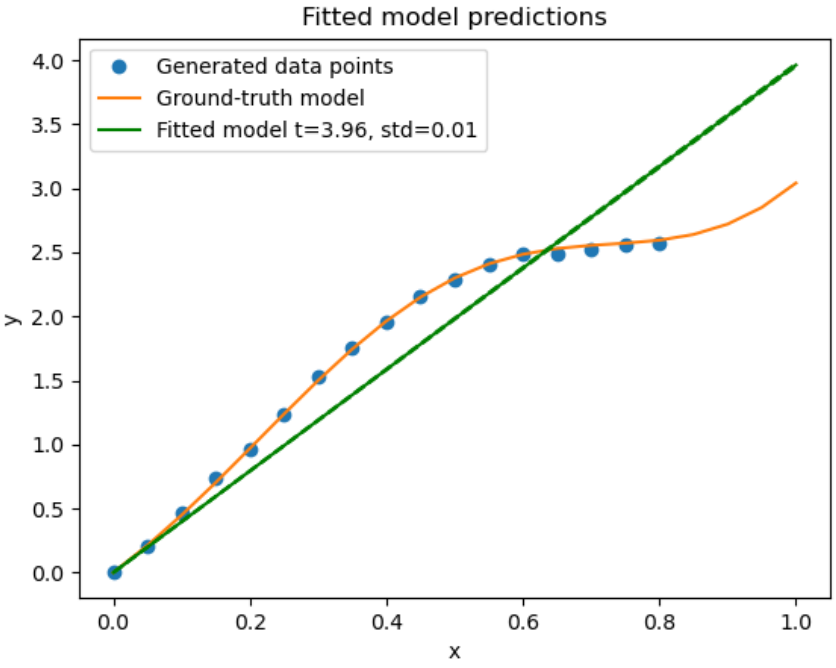
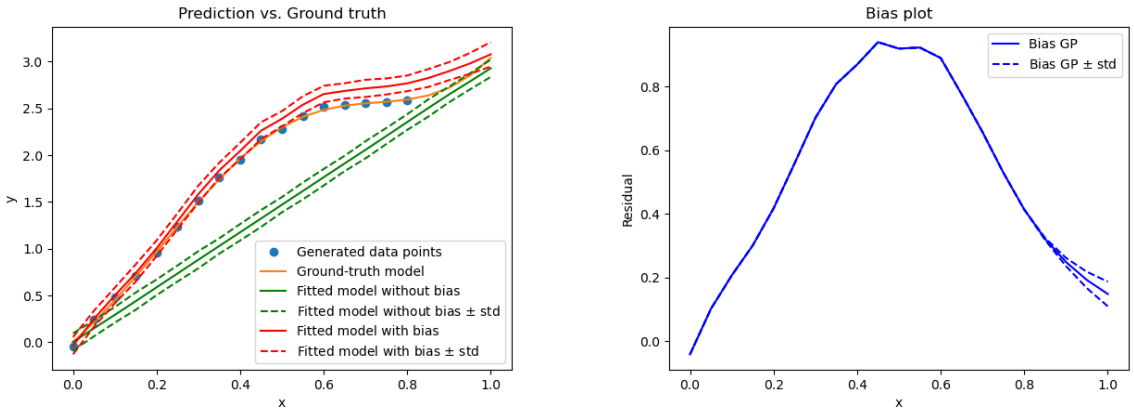


Figure 1: Fitted system for bias-free Bayesian inference.

includes the bias inherent to its definition. Additionally, there is no influence on the computer model by the lack of data with $x > 0.8$. These results prove the necessity of accounting for model bias, as the computer model could only be used for predictions when the certainty of its validity is very high.

3.2 Modularized KOH approach

Next, a modularized version of KOH's approach is implemented. In a first step, a GP is trained for the computer model f. In such a simple case, training a GP should not be required, but it is included here to account for the potential variability that appears when surrogating a more complex system. Following the methodology indicated in [8], in this benchmark the kernels for every GP would be a Matérn kernel with $\nu = 3/2$ and correlation length $\ell = 0.5/\sqrt{3}$. A grid of computational observations of 10×10 regularly spaced samples in $(x, t) = [-0.2, -0.2] \times [1.2, 5.0]$ is evaluated as the training dataset for the GP of f. Once trained, it is used as a predictor for the inference procedure. For t, an analogous normal distribution is provided as prior. This process is implemented using the software package `PyMC` [13]. The results obtained are shown in Figures 2 and 3.



(a) Fitted results

(b) Bias

Figure 2: Fitted system and bias for Kennedy and O'Hagan's approach.

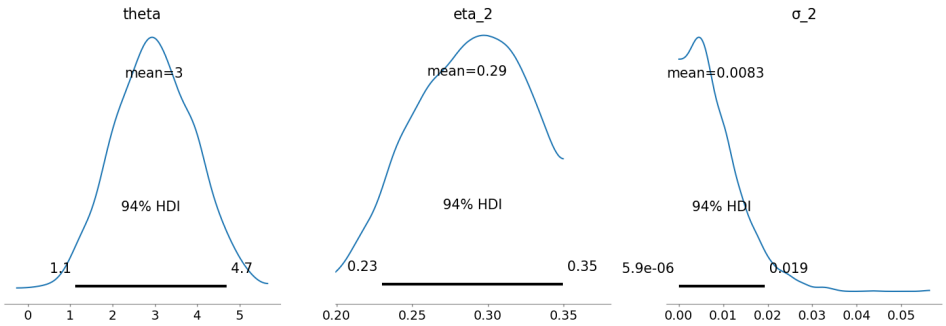


Figure 3: Posterior distribution for latent parameters in KOH. The plot for "theta" represents the posterior distribution of the best fitted t, "eta_2" represents the scale factor of the GP for the bias term and "σ_2" represents a Gaussian noise scale.

The corrected results clearly align better with the measurements than using the bias-free approach. There is however more variability in the predictions. This comes from the introduction of a GP as a surrogate model for f and the model bias inference itself. Modifications in the

prior distributions had a significant impact on the fitted posterior distributions. Additionally, it can be observed that the further from the training data points, the larger the variance in the bias.

3.3 Orthogonal GPs approach

The procedure followed to implement the is analogous to the previous one, with a difference in the definition of the correlation function for the GP of the bias. As the computer model is simple, it is possible to express analytically the new orthogonal priors for the correlation functions [8]. Being $r(x, x')$ the original correlation function with a Matérn kernel 3/2 and a loss L_{L^2}, the new correlation for the bias is:

$$r_\theta(x, x') = 5.19r(x, x') - 23.52h(x)h(x') \tag{5}$$

where

$$h(x) = 2x - \frac{2x+3}{4}\exp(-2x) + \frac{6x-13}{4}\exp(2x-2) \tag{6}$$

Meanwhile, for the loss $L_{W_1^2}$:

$$r_\theta(x, x') = 1.27r(x, x') - 0.78h(x)h(x') \tag{7}$$

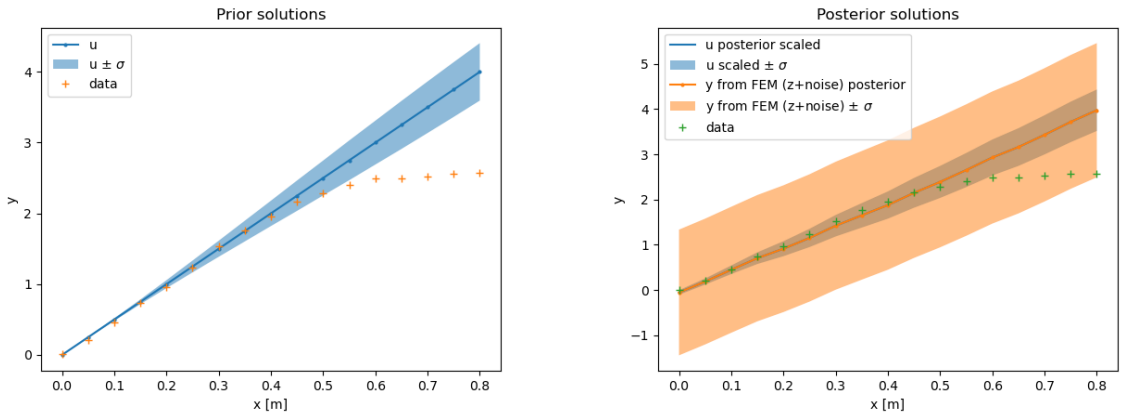
where

$$h(x) = 2x - \frac{6x+1}{4}\exp(-2x) + \frac{2x+1}{4}\exp(2x-2) \tag{8}$$

Applying these new priors for the correlation of the bias, it is possible to obtain the results indicated in [8]. These results show lower variance than the KOH approach, where θ (the best estimator for t) is centered around the minimizer of the loss. Therefore, the model fitted using orthogonal GPs provides predictions closer to the measurements, while the bias corrects exclusively the part of the real system that cannot be explained by the computational model. The choice of loss weights which characteristics of the measurements must be considered stronger. The loss L_{L^2} will weigh more fitting to the actual points, which provides higher θ due to the initial slope of the real measurements. Alternatively, the loss $L_{W_1^2}$ intuitively weighs in additionally the first derivative of the measurements underlying process, providing lower values for θ. The choice of this loss is an additional consideration for practical implementations, as it will vary depending on the application of the digital twin.

3.4 StatFEM approach

To implement statFEM it is required a FEM model of the system that can be represented as a linearized version of a PDE, therefore the computer model from the benchmark must be modified accordingly. In this case, the example can be approximated by a truss under axial load. The solution of such a system is a linear function $ku = f$, where u is the deformation and f is the force. Adjusting k and f, we obtain the same linear system as in the other examples. The variability is introduced as a distribution on the source term, recreating the case when the load is unknown beforehand and potentially not uniform across its length. The prior results can be observed in Figure 4a. The posterior after the inference procedure is presented in Figure 4b. As it can be appreciated, the influence of the initial model is still dominant over the measurements. This is a typical case for statFEM implementations, as the differences between the output u and the measurements tend to be seen by the inference model as a product of the variability of u. This effect is reduced when more than one sample is present for every value of x (see [9]), as the certainty of the model at the data points is increased. It can be observed that the bias is added



(a) Prior

(b) Posterior

Figure 4: Solutions for statFEM.

as if the fitted model were to remain close to the original region of the prior, while the corrected system includes the data points. Despite depending on the variability introduced by the source term, the output is less sensitive to the defects in the model, and the predictions generated by the fitted model will reflect this stability. This makes statFEM the appropriate method when the reliability of the measurements is decisive, as the real values for the prediction will likely be found inside the region predicted by the model. Additionally, the uncertainty in the predictions comes directly from that introduced to the initial model, which reflects faithfully the previous knowledge in the system.

3.5 Comparison

First, it must be noted that all the approaches that include model bias allow for quantification of the uncertainty introduced by the choice of an incomplete computer model. Nevertheless, this is done by adding further variance in the results. This may seem an undesirable effect, but being aware of this variance while obtaining results closer to reality allows for better-informed decision-making. The focus now resides on 1) reducing the variability of the results as much as possible, and 2) providing reliable, repeatable and stable predictions with the fitted model.

Regarding the smallest variability, orthogonal GPs triumph over the other methods analyzed with the simple example. However, this variance is intrinsically related to the choice in the priors, therefore a real application of a digital twin will depend almost exclusively on the information available and the measurements observed. Kennedy and O'Hagan's approach is the one that suffers the most due to this dependency, in exchange for being the simplest to integrate and that requires the least information about the system. Alternatively, statFEM introduces the variance through prior knowledge of measurable parameters of the system instead of introducing a reasonable expectation for the prior distribution, which arguably improves the insight provided by it. The spread of statFEM's results is generally larger than in the other approaches, but it would greatly benefit from measurements repetition.

Concerning the reliability of results, statFEM offers the most stable predictions due to its closeness to the prescribed system. The other approaches tend to overfit more to the model, relying on the available measurements. KOH seems to be the one with the least robust results, as the non-directed stochastic nature of the inference procedure does not put any restriction on the fitness of the bias, suffering from the ill-posedness of the problem.

Finally, an analysis of how each of the models would be implemented in a digital twin is required. The KOH approach is the easiest to implement and can be adapted to virtually any model and measurements without many considerations. Orthogonal GPs require access to derivatives, therefore surrogating the computer model is usually a hard requirement that may reduce its range of application. Furthermore, manipulating the covariance matrix can result in larger computational times, which should be considered carefully for problems highly dimensional in the latent parameters. Alternatively, statFEM requires a FE model available for the system, and its original formulation imposes strong assumptions of linearity and time independency. This makes it unsuitable for many applications in its current form. Nonetheless, statFEM can be a good alternative for those cases that fulfil these requirements, as it is independent of the choice of latent parameters and relatively robust against deficient modelling assumptions. It is in fact invariant to the choice of latent parameters, but it escalates very poorly with the increase in dimensions and number of measurement positions, which makes it paradoxically a better alternative the fewer sensor points are available.

4 CONCLUSIONS

In this study, we compared the effectiveness of different model bias identification approaches regarding their application to digital twins. Namely, using the classical Bayesian inference implementation as the basis, we analyzed Kennedy and O'Hagan's approach in its original and modularized version, orthogonal GPs and statFEM. They were applied to a simple 1D example as a benchmark.

Through extensive evaluations, we found that every implementation outperforms bias-free methods, having different potential applications. Orthogonal GPs provide the least variance in the estimation of the updated model parameters with bias. StatFEM is proposed as the choice when reliability is to be prioritize and its requirements can be satisfied. Nevertheless, each model bias identification approach requires knowledge about the behaviour of both the real and computational model, and this information will greatly influence the performance once implemented in the digital twin.

Nevertheless, the findings of this study have significant implications for the development of digital twins as a valuable technology, as they contribute to a safe, efficient and reliable implementation of digital twins by combining simulations and measurement data. A model whose bias is adequately quantified provides invaluable information to the end users, who would be informed not only on the measurements and predictions but on the fitness of the model itself.

In conclusion, the results of this study have identified the most efficient and reliable approach for quantifying model bias in simulation-based digital twin applications. In future work, two main research paths can be discerned. First, the different models should be applied to a larger variety of cases and real measurement data. Implementation concerns such as the data structure or behaviour when providing predictions cannot be studied in a simple example as the one proposed. Alternatively, there is still room for improvement in the implementation of statFEM and similar models, as its applicability is reduced by its requirements.

REFERENCES

[1] M. W. Grieves, Virtually Intelligent Product Systems: Digital and Physical Twins, *Complex Systems Engineering: Theory and Practice*, American Institute of Aeronautics and Astronautics, Inc., 2019, pp. 175–200.

[2] F. Tao and Q. Qi, Make more digital twins. *Nature*, **573**, Art. no. 7775, Sep. 2019.

[3] Y. Cong et al., A digital twin of bridges for structural health monitoring, *Structural Health Monitoring 2019*, Stanford, USA, Nov. 2019.

[4] F. Sanfilippo, R. T. Thorstensen, A. Jha, Z. Jiang, and K. G. Robbersmyr, A Perspective Review on Digital Twins for Roads, Bridges, and Civil Infrastructures. *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, Maldives, November 16-18, 2022.

[5] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*. Springer London, Limited, 2006.

[6] M. C. Kennedy and A. O'Hagan, Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*,**63**, Art. no. 3, 2001.

[7] M. J. Bayarri, J. O. Berger, and F. Liu, Modularization in Bayesian analysis, with emphasis on analysis of computer models. *Bayesian Analysis*, **4**, Art. no. 1, Mar. 2009.

[8] M. Plumlee, Bayesian Calibration of Inexact Computer Models. *Journal of the American Statistical Association*, **112**, Art. no. 519, Jun. 2017.

[9] M. Girolami, E. Febrianto, G. Yin, and F. Cirak, The statistical finite element method (statFEM) for coherent synthesis of observation data and model predictions. *Computer Methods in Applied Mechanics and Engineering*, **375**, p. 113533, May 2019.

[10] R. B. Gramacy et al., Calibrating a large computer experiment simulating radiative shock hydrodynamics. *Annals of Applied Statistics 2015*, **9**, No. 3, 1141-1168, Oct. 2014.

[11] C. Duffin, *Statistical finite element methods for nonlinear PDEs.* PhD. Thesis, University of Western Australia, 2022.

[12] V. B. Narouie, H. Wessels, and U. Römer, Inferring Displacement Fields from Sparse Measurements Using the Statistical Finite Element Method. *ArXiv*, Dec. 2022.

[13] J. Salvatier, T.V. Wiecki, C. Fonnesbeck, Probabilistic programming in Python using PyMC3. *Peerj Computer Science*, **2**:e55, 2016.