A Probabilistic Graphical Model Foundation for Enabling Predictive Digital Twins at Scale

Michael G. Kapteyn\textsuperscript{1}, Jacob V.R. Pretorius\textsuperscript{2}, and Karen E. Willcox\textsuperscript{3}\textsuperscript{*}

\textsuperscript{1} Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139, United States
\textsuperscript{2} The Jessara Group, Austin, TX 78704, United States
\textsuperscript{3} Oden Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, TX 78712, United States

Abstract

A unifying mathematical formulation is needed to move from one-off digital twins built through custom implementations to robust digital twin implementations at scale. This work proposes a probabilistic graphical model as a formal mathematical representation of a digital twin and its associated physical asset. We create an abstraction of the asset-twin system as a set of coupled dynamical systems, evolving over time through their respective state-spaces and interacting via observed data and control inputs. The formal definition of this coupled system as a probabilistic graphical model enables us to draw upon well-established theory and methods from Bayesian statistics, dynamical systems, and control theory. The declarative and general nature of the proposed digital twin model make it rigorous yet flexible, enabling its application at scale in a diverse range of application areas. We demonstrate how the model is instantiated as a Bayesian network to create a structural digital twin of an unmanned aerial vehicle. The graphical model foundation ensures that the digital twin creation and updating process is principled, repeatable, and able to scale to the calibration of an entire fleet of digital twins.

Keywords— Digital Twin, Uncertainty Quantification, Probabilistic Graphical Model, Self-Aware Vehicle

Introduction

A digital twin is a set of coupled computational models that evolve over time to persistently represent the structure, behavior, and context of a unique physical asset such as a component, system, or process. The digital twin paradigm has seen significant interest across a range of application areas as a way to construct, manage, and leverage state-of-the-art computational models and data-driven learning. Digital twins underpin intelligent automation by supporting data-driven decision making and enabling asset-specific analysis. Despite this surge in interest, state-of-the-art digital twins are largely the result of custom implementations that require considerable deployment resources and a high level of expertise. To move from the one-off digital twin to accessible robust digital twin implementations at scale requires a unifying mathematical foundation. This paper proposes such a foundation by drawing on the theoretical foundations and computational techniques of dynamical systems theory and probabilistic graphical models. The result is a mathematical model for what comprises a digital twin and for how a digital twin evolves and interacts with its associated physical asset. Specifically, we propose a probabilistic graphical model of a digital twin and its associated physical asset, providing a principled mathematical foundation for creating, leveraging, and studying digital twins.

Digital twins have garnered attention in a wide range of applications\textsuperscript{[1]}. Structural digital twins have shown promise in virtual health monitoring, certification, and predictive maintenance\textsuperscript{[2,3,4,5]}. In healthcare, digital twins of human beings promise to advance medical assessment, diagnosis, personalized treatment, and \textit{in-silico} drug testing\textsuperscript{[6,7,8]}. Similarly, digital twins of individual students offer a path to personalized education\textsuperscript{[9]}. At a larger scale, smart cities enabled by digital twins and Internet of Things (IoT) devices promise to revolutionize urban planning, resource allocation, sustainability and traffic optimization\textsuperscript{[10]}. Although each of these applications has its own unique requirements, challenges, and desired outcomes, the mathematical foundation we develop in this work focuses on a common thread that runs throughout: the infusion of dynamically updated asset-specific computational models into the data-driven analysis and decision-making feedback loop.

We adopt a view of the physical asset and its digital twin as two coupled dynamical systems, evolving over time through their respective state spaces as shown in Figure\textsuperscript{[1]}. The digital twin acquires and assimilates observational data from the asset (e.g., data from sensors or manual inspections) and uses this information to continually update its internal models so that they reflect the evolving physical system\textsuperscript{[11]}. The digital twin can then use these up-to-date internal models for analysis, prediction, optimization, and control of the physical system.

Motivated by this conceptual model, we develop a probabilistic graphical model\textsuperscript{[12]} that defines the elements comprising this coupled dynamical system, and mathematically describes the interactions that need to be modeled in the digital twin. Our model draws inspiration from classical agent-based models such as the partially observable Markov decision process\textsuperscript{[13]}, but includes important features that are unique to the digital twin context. The graphical model formalism
provides a firm foundation from which to draw ideas and techniques from uncertainty quantification, control theory, decision theory, artificial intelligence, and data-driven modeling in order to carry out complex tasks such as data assimilation, state estimation, prediction, planning, and learning, all of which are required to realize the full potential of a digital twin.

Throughout this paper we demonstrate our proposed graphical model using a motivating application: the development of a structural digital twin of an unmanned aerial vehicle (UAV). In particular, we demonstrate how experimental calibration can be used to transform a baseline structural model into a unique digital twin tailored to the characteristics of a particular asset. We show how this process can be formulated via the proposed graphical model in order to provide a rigorous framework for deciding which calibration experiments to perform, leveraging the experimental data for principled model calibration, and evaluating the performance of the calibrated digital twin.

Once calibrated, the structural digital twin can enter operation alongside the physical asset, where it assimilates sensed structural data to update its internal models of the vehicle structure. The dynamically updated models are used for analysis and evaluation of the vehicle’s structural health and for decision-making. This motivating example has specific application in a number of settings. One is urban air mobility and autonomous package delivery, where UAVs operate in urban environments and are subject to damage. In order to ensure a safe, robust, reliable and scalable system, these vehicles must be equipped with advanced sensing, inference, and decision capabilities that continually monitor and react to the vehicle’s changing structural state. The vehicle uses this capability to decide whether to perform more aggressive maneuvers or fall back to more conservative maneuvers in order to minimize further damage or degradation. Another application is hypersonic vehicles, which operate in extreme environments and thus undergo continual degradation of their structural condition. More generally, the structural digital twin illustrative example is highly relevant to other applications where the condition of the system changes over time due to environmental influences and/or operating wear and tear. Examples include wind turbines, nuclear reactors, gas turbine engines, and civil infrastructure such as buildings and bridges.

Results

We present three key results: the abstraction of a digital twin into a state-space formulation that leads to a probabilistic graphical model, the realization of the digital twin mathematical model as a dynamic Bayesian network, and the dynamic evolution of a UAV structural digital twin as experimental data are acquired.

A Probabilistic Graphical Model for Digital Twins

The first main result of this work is the abstraction of a digital twin and its associated physical asset into a representation comprised of six key quantities, defined in Figure 2. We provide specific examples from the self-aware UAV application but emphasize that this abstraction can be applied to digital twins from any discipline and application, thus providing a unifying framework for describing and defining digital twins. Note the key differences between the physical and digital states. The physical state space encapsulates variation in the state of the asset and could thus be a complex high-dimensional space. The physical state is typically not fully observable. Together, these attributes make the physical state generally intractable to model directly. The digital state is
defined as a set of parameters that characterize the models comprising the digital twin. The digital state is updated as the asset evolves over time or as new information about the asset becomes available. In defining the digital state, one must consider what information is sufficient to support the use case at hand. A well-designed digital twin should be comprised of models that provide a sufficiently complex digital state space, capturing variation in the physical asset that is relevant for diagnosis, prediction, and decision-making in the application of interest. On the other hand, the digital state space should be simple enough to enable tractable estimation of the digital state, even when only partially observable. In general, the digital state space will be only a subset of the physical state space.

This abstraction of the digital twin leads to the second main result of this work, a probabilistic graphical model of a digital twin and its associated physical asset. This graphical model represents the structure in an asset-twin system by encoding the interaction and evolution of quantities defined in Figure 2. In particular, the model encodes the end-to-end digital twin data-to-decisions flow, from observations through inference and assimilation to action. We present here a particular type of graphical model called a dynamic Bayesian network: a directed acyclic graph that grows in size with each discretized time step. Figure 3 shows the resulting dynamic Bayesian network unrolled from the initial time step, denoted \( t = 0 \), to the current non-dimensional time step, denoted \( t_c \), and into the future to the prediction horizon, denoted \( t_p \).

This graphical model serves as a mathematically rigorous counterpart to the conceptual model illustrated in Figure 1. Throughout this section we use upper-case letters to denote random variables, with the corresponding lower-case letter denoting a value of the random variable. For example, at each timestep, \( t \), the digital state, \( D_t \), is estimated probabilistically by defining the random variable, \( D_t \), such that \( D_t \sim p(D_t) \). The upper left-to-right path in Figure 3 represents the time evolution of the physical asset state, \( S_t \), while the lower path represents the time evolution of the digital state, \( D_t \). The graphical model encodes the tight two-way coupling between observations, \( O_t \), and actions, \( U_t \), at each time step, \( t \): information flows from the physical asset to the digital twin in the form of observational data, while information flows from the digital twin to the physical twin in the form of control inputs. The models comprising the digital twin can be used to predict quantities of interest, \( Q_t \). These quantities, along with the digital state, observational data, and control inputs, influence the reward accumulated for the time step, \( R_t \). Note that in the presented formulation both observations and control inputs occur once per time step, with observations occurring prior to control inputs. The model and algorithms we discuss can easily be adapted to situations in which this is not the case via a topological reorganization of the graphical model. For example, in the application presented in the following section observational data are observed after a control input is issued.

Formally, the edges in the graphical model define condi-
t = 0       t = t_c = 2       t = t_p = 4

Figure 3: Our approach results in a dynamic Bayesian network that mathematically represents a physical asset and its digital twin. Nodes in the graph shown with bold outlines are observed quantities (i.e., they are assigned a deterministic value), while other quantities are estimated (typically represented by a probability distribution). Directed edges represent conditional dependence. Dashed edges are removed from the graph after their target control node is selected and enacted.

Note that the proposed probabilistic graphical model does not restrict the nature of the models comprising the digital twin. Each of these models could be physics-based (e.g., models based on discretized partial differential equations), data-driven (e.g., neural networks trained on historical or experimental data), or rule-based (e.g., state-transitions in a finite-state automaton). Note also that these models need not be specific to a single physical asset. Indeed, the models may be shared across a large fleet of similar assets, each with a specific digital state governing the parameters used in these models. This enables efficient, centralized data and model management, and also enables each asset in the fleet to contribute to the continual improvement and enrichment of models by providing performance data relevant to a particular part of the state space.

Dynamic data-driven evolution of a UAV structural digital twin We instantiate the approach for the structural digital twin of a UAV, with the target use case of managing a large fleet of UAV assets (e.g., a fleet of autonomous package delivery vehicles). While many UAVs in the fleet share a common design, variation in material properties, manufacturing processes, and system degradation over time ensures that no two assets are truly identical. This variation is captured by creating a structural digital twin for each unique UAV asset. Such a digital twin would enable accurate structural simulation, which could in turn enable simulation-based evaluation, certification, predictive maintenance of the fleet, and optimized operations.

We focus on a key stage in the development of any digital twin: the calibration, tailoring, and updating of computational models within the digital twin, so that they accurately reflect the current characteristics of a unique physical asset. In our example this is achieved by performing a series of experiments on a physical UAV asset, using the resulting data to calibrate and update the structural models within the digital twin. The physical asset, experimental setup, and finite element structural models are described in the Methods section. While this type of experimental model calibration is commonplace throughout engineering, we here demonstrate how formulating the calibration process using our proposed probabilistic graphical model (Fig. 3) ensures that the process is principled, scalable, and repeatable across the entire fleet of UAVs. Moreover, we will discuss how the same probabilistic graphical model can be naturally extended in time beyond calibration and into the service life of each UAV.

In this example application we define the physical state, $S$, of the UAV asset to be its structural state. The physical state-space encompasses any conceivable structural variation between UAV assets, such as differences in geometry and material properties. The physical state of the asset is determined during design and manufacturing. When the asset enters operation, the physical state evolves over time, for example due to maintenance events, damage, or gradual degradation. We consider here the calibration phase, where the asset has not
yet entered operation; thus, the physical state remains constant during the timesteps for which we present results.

The digital state is defined to be:

\[
d := \begin{bmatrix} g \, e \, m \, \alpha \, \beta \end{bmatrix}
\]

This digital state comprises parameters that account for differences between UAV assets due to material or manufacturing variability (geometry, \(g\), and Young’s modulus scale factor, \(e\)), parameters that represent unmodeled details of the UAV (the point masses \(m\) account for physical UAV components that are not represented in the finite element model), and parameters that represent unmodeled physics (the Rayleigh damping coefficients are widely used in engineering to approximately model internal structural damping). By focusing on these parameters (while fixing all other parameters at a nominal value) we obtain a digital twin that is tractable to calibrate, while sufficiently capturing the key differences in consequent structural response between UAV assets.

Each digital twin begins the calibration process with the same prior belief about the digital state parameters, \(p(D_0)\). This distribution over model parameters defines a baseline UAV structural model based on design specifications, with uncertainty based on the degree of variability or confidence in each parameter. At each step, \(t = 1, 2, 3\) of the calibration process we select an action \(U_t = u_t\), which corresponds to conducting a specific experiment on the physical UAV asset. This experiment generates uncertain data, \(o_t \sim p(O_t)\), which are used to compute an updated belief about the digital state, \(p(D_t | O_t = o_t)\). Specifically, each experiment is focused on calibrating a subset of parameters in the digital state, while keeping the other parameters fixed. Using this updated distribution of parameters in the computational model (described in the Methods section), the digital twin is used to compute quantities of interest, \(p(Q_t | D_t, O_t = o_t)\), which describe the structural response of the UAV. We evaluate the success of the calibration procedure by estimating the reward, \(p(R_t)\), at each stage. In what follows, values that are experimentally measured or derived from experimental data are denoted as hat variables, while the corresponding variable without a hat represents a computational estimate of the quantity produced by digital twin. We use braces (\{·\}) to denote that there is an ensemble of data generated by repeated trials of an experiment. Figure 4 presents a summary of the experimental calibration procedure and its formulation in terms of the graphical model.

At time step \(t = 1\) we calibrate the digital twin geometric parameters, \(g = [l, c_{\text{root}}, c_{\text{tip}}]\), where \(l\) is the wing semi-span, \(c_{\text{root}}\) is the chord length at the root, and \(c_{\text{tip}}\) is the chord length at the tip. The control input, \(u_1\), for this timestep is to physically measure each geometric parameter, producing measurements \(\hat{g}\). Since these measurements are able to be taken accurately, the posterior uncertainty in the geometric parameters is negligibly small. Thus, the posterior distribution, \(p(D_1 | O_1, U_1)\), deterministically sets the geometric parameters to their measured values.

At time step \(t = 2\) we calibrate the static load-displacement behavior of the structural model. In the digital state we update \(e\), which is a scale factor applied to the Young’s modulus (both longitudinal and transverse) of the carbon fiber material used in the wing skin. This scale factor allows us to adjust the computational model to account for material or manufacturing variation in the wing skin. The control input, \(u_2\), is a decision to perform a static load-displacement test on the physical UAV asset. The observed data are a set of applied tip-load and measured tip-displacement pairs, \(o_2 = \{f, \hat{x}\}\), as shown.
in Figure 5 (left). Uncertainty in the applied load, \( \hat{f} \), and measured tip-displacement measurement, \( \hat{x} \), are both modeled by Gaussian distributions, with 95% confidence interval of widths of 20g and 1mm respectively. Using this observed data, we perform a Bayesian update on our prior estimate in order to produce the posterior estimate \( p(D_2|D_1, O_2, U_2) \), as shown in Figure 5 (center). We see that the posterior mean value for this scale factor is 1.0073, indicating that the Young’s modulus of carbon fiber in the digital twin model should be increased by 0.73% to better match this particular physical UAV asset. The uncertainty in this parameter is also greatly reduced post-calibration. The quantity of interest for this timestep is a decision to perform an initial condition response experiment. This is done by applying an initial tip displacement and releasing the wing, recording data \( o_3 \) in the form of strain, \( \epsilon \), as a function of time using a dynamic strain sensor. The strain data are post-processed to extract natural frequencies, \( \hat{\omega}_i \) and damping ratios, \( \hat{\zeta}_i \), for the first two bending modes \( i = 1, 2 \). As described in the Methods section, we formulate an optimization problem to fit a distribution of point masses, \( m \), to the wing so that the natural frequencies of the first two bending modes predicted by the model, \( \omega_1 \) and \( \omega_2 \), match the experimental data, \( \hat{\omega}_2 \) and \( \hat{\omega}_2 \), as closely as possible while also matching the total wing mass. Using the calibrated natural frequencies we then compute the coefficients, \( \alpha \), \( \beta \), in the Rayleigh damping model. Together with the computed masses, this gives an estimate for the final calibrated posterior distribution \( p(D_3|D_2, O_3, U_3) \). The quantities of interest for this step are the numerical estimates for modal frequencies and damping ratios. The reward for this step is the difference between the posterior quantity of interest and the experimentally observed values.

Prior and posterior estimates for each component in the digital state are summarized in Table 1. We use \( \mathcal{N}(\mu, \sigma) \) to denote a Normal distribution with mean \( \mu \) and standard deviation \( \sigma \). For sample distributions we report the sample mean, followed by the sample standard deviation in parentheses.

| \( c_{root} \) [mm] | \( c_{tip} \) [mm] | \( l \) [mm] | \( \epsilon \) [-] | \( m_{servo} \) [g] | \( \alpha \) [s^{-1}] | \( \beta \) [s] |
|------------------|------------------|----------|---------|---------|----------|-------|
| Prior information | \( \mathcal{N}(435.6, 1.3) \) | \( \mathcal{N}(261.1, 1.3) \) | \( \mathcal{N}(1828.8, 1.3) \) | \( \mathcal{N}(1.0, 0.026) \) | \( \frac{2m_{servo} + m_{pitot}}{m_{servo}, m_{pitot} \geq 0} = 472 \) | 0 | 0 |
| Posterior estimate | 433 | 260 | 1828 | 1.0073 (0.0103) | 169.1 (3.9) | 1.030 (0.001) | 7.66×10^{-4} (6.18×10^{-7}) |

Table 1: Prior and posterior estimates for each entry in the digital state.
Discussion

The probabilistic graphical model formulation advances the field of digital twins by clearly defining the elements comprising an abstract representation of an asset-twin system; identifying the interactions between elements that need to be modeled, and incorporating end-to-end uncertainty quantification via a Bayesian inference framework. Our model demonstrates how the digital twin paradigm incorporates dynamic updating and evaluation of computational models into a data assimilation and feedback control loop. These computational models offer valuable insights, unattainable through raw observational data alone, which can be leveraged for improved data-driven modeling and decision making [14].

This paper presented an example application of the model, focused on the experimental calibration of a structural digital twin for a UAV. This application was chosen in order to demonstrate the process of defining quantities and applying an instantiation of the graphical model. In order to extend this application into the operational phase of the UAV, future work will focus on developing a model of the state evolution dynamics (e.g., structural damage evolution), as well as the capability to perform optimal control and planning. With these additions, asset monitoring could continue in-flight, where observational data from structural sensors combined with a model of damage evolution would enable the UAV to monitor its structural health (the digital state) and predict flight capabilities (the quantity of interest). This information would enable the UAV to dynamically re-plan upcoming maneuvers (control inputs) based on the digital state. Beyond this UAV application, the declarative and general nature of the graphical model mean that it can be applied to a wide range of applications across engineering and science. Indeed, this application is just one example of a much broader class of next-generation intelligent physical systems that need to be monitored and controlled throughout their lifecycle.

Limitations of the proposed framework include the challenge of defining and parameterizing the models comprising the digital twin. A central aspect of this challenge is a need to quantify and manage model inadequacy (sometimes referred to as model discrepancy [15]), a topic that is beginning to receive more attention throughout computational science. Another open challenge is sensor design for digital twins. In our graphical model this is formalized through the control theoretic notion of observability, i.e., whether the available observational data combined with carefully selected control inputs is adequate to inform the digital state. Similarly, the notion of controllability could be used to study whether available control inputs are sufficient to drive the physical asset to a desired state. Finally, computational resource limitations remain a challenge for the realization of predictive digital twins. The probabilistic inference procedure developed in this paper, even when accelerated via approximate inference techniques, will require many evaluations of the models comprising the digital twin. Such high-fidelity physics-based models require approximations such as reduced-order modeling, surrogate modeling, and other compression techniques [10].

Methods

Physical UAV asset

The physical asset used for this work is a fixed-wing UAV testbed vehicle [17]. A rendering of the vehicle is shown in Figure 2. The fuselage, empennage, and landing gear are from an off-the-shelf Giant Telemaster airplane. The wings have been custom designed with plywood ribs and carbon fiber skin, in order to mimic the structure of a commercial or military-grade vehicle, albeit at a reduced scale. The wings are connected to the fuselage via two carbon rods that extend 25% of the way into each wing. The electric motor, avionics, and sensor suite are a custom installation. Further details on the sensors used for this work are provided in the experimental methodology section.

The experiments in this work were conducted in a laboratory environment. As the calibration is focused only on the aircraft wings, they are detached from the fuselage and mounted to a fixed support. The wings are mounted upside-down so that typical aerodynamic forces can be more easily applied as downward forces.

Computational models comprising the UAV digital twin

The computational model at the heart of the UAV structural digital twin is a physics-based model for the linear elastic structural response of the aircraft [18]. This model relates a time-varying load on the aircraft (e.g., aerodynamic loads for a given aircraft maneuver), to its structural displacement. We adopt a finite element spatial discretization, so that this relationship can be expressed as the second-order ordinary differential equation

\[ M(d)x(t) + V(d)\dot{x}(t) + K(d)x(t) = f(t). \]

Bolded quantities represent vectors of nodal quantities. In this case, \( x, \dot{x}, \text{ and } \ddot{x} \) denote the displacement, velocity, and acceleration, respectively, while \( f \) denotes the applied force at each node in the finite element mesh. The mass, damping, and stiffness matrices, \( M(d), V(d), \text{ and } K(d) \) respectively, are parametrized by the digital state, \( d \). As defined by [1], the digital state is an asset-specific parameter vector that reflects the unique geometric and structural properties of each individual UAV.

The static load-displacement experiment used to generate data at time step \( t = 2 \) in the calibration application is simulated by the digital twin via the static version of [2].

\[ K(d)x = f. \]

Note that for this particular calibration experiment, we set the applied loads vector, \( f \), to represent a point load of magnitude, \( f \), near the tip. We solve the model to compute the displacement vector, \( x \), and post-process the solution to extract the tip displacement, \( x \). Through a computational study using the static elasticity model [8], it is found that the aggregate wing stiffness, \( k \), i.e., the coefficient of proportionality between an applied tip load, \( f \), and the tip displacement, \( x \),
depends linearly on the Young’s modulus scale factor applied to the model, \( e \). In particular, we establish the relationship

\[
k = 0.5752e + 0.1018. \tag{4}
\]

While the calibration process focuses on a tip load, note that once calibrated the digital twin is able to accurately simulate the deformation of the wing to any static load, an example of the power of computational models for prediction and analysis.

The dynamic response of the aircraft structure can also be characterized by the digital twin via an eigenanalysis. The computational model \( \mathcal{E}_t \) can be adapted to perform an eigenanalysis of the form

\[
K(d)\ddot{x}_i = M(d)\omega_i^2\dddot{x}_i \tag{5}
\]

where \( \dddot{x}_i \) is a nodal displacement vector representation of the mode shape, and \( \omega_i \) is the natural frequency for mode \( i \). This model is used at time step \( t = 3 \) of the calibration procedure.

We adopt a Rayleigh damping model \([19]\) which defines the damping matrix, \( V \), as a combination of mass-proportional damping and stiffness proportional damping as follows:

\[
V(d) = \alpha M(d) + \beta K(d). \tag{6}
\]

Under this model the \( i \)th modal damping ratio, \( \zeta_i \), is given by

\[
\zeta_i = \frac{\alpha}{2} \frac{1}{\omega_i} + \frac{\beta}{2} \omega_i, \tag{7}
\]

where \( \omega_i \) is the \( i \)th natural frequency.

During calibration we focus on the first two bending modes, \( i = 1, 2 \), as these are clearly evident in our experimental data. However, the calibrated computational model is not limited to these modes—indeed the digital twin enables us to compute higher bending modes as well as torsional and skin buckling modes, which are difficult to characterize experimentally.

### Methodology for experimental calibration of the UAV Digital Twin

This section provides additional details on the experimental and computational procedures used for each step of the calibration procedure.

#### Prior estimate for the digital state

Our prior estimate of the geometric parameters, \( g = [l, c_{root}, c_{tip}] \), is based on the nominal or as-designed value for each parameter, as stated on technical drawings for the wing design that were provided to the wing manufacturer. The allowable manufacturing tolerance is stated in these drawings as \( \pm 2.5 \)mm. Thus, our prior estimate of each geometric parameter is modeled as a Gaussian distribution, where the mean corresponds to the nominal design value, and the standard deviation is set to give a 95% confidence interval width equal to the manufacturing tolerance.

We set the prior distribution over the carbon fiber Young’s modulus scale factor, \( e \), to be a Gaussian distribution with mean 1 and a 95% confidence interval equal to \( \pm 5\% \) variability. This prior uncertainty is an expert-driven estimate that could be refined over time as more wings are manufactured and the degree of variability in material properties is characterized based on manufacturing data.

We add three point masses to the model: two have mass \( m_{servo} \) and represent the servomotor hardware that actuates each aileron, while the third has mass \( m_{pitot} \) and represents the pitot tube attached to the wing tip. The location of each of these hardware components is measured on the physical asset and fixed in the digital twin model. It is not possible to measure the mass of each component individually as they are fixed to the wing during manufacturing. Instead, we measure the total weight of the wing hardware and compare this with the mass accounted for in the computational model. A discrepancy of 472g is identified. Thus, our prior information about these point masses comes in the form of the constraints:

\[
2m_{servo} + m_{pitot} = 472, \tag{8}
\]

\[
m_{servo}, m_{pitot} \geq 0. \tag{9}
\]

Finally, in this work we consider the prior estimates on the Rayleigh damping coefficients, \( \alpha \) and \( \beta \), to both be zero. This can be interpreted as having no damping in the model before the damping is experimentally calibrated.

#### Step 1: Calibrate geometry, \( g \)

The first calibration step is focused on updating the components of the digital state estimate corresponding to the geometric parameters \( g = [l, c_{root}, c_{tip}] \).

To update the prior estimate over these parameters into the posterior, \( p(D_1|O_1) \), we measure the as-manufactured wing geometry. The measured values constitute the observational data, \( O_1 = [\hat{l}, \hat{c}_{root}, \hat{c}_{tip}] \). The maximum measurement error is estimated to be \( \pm 0.5 \)mm. While each measurement could be represented by a Gaussian distribution and incorporated via a Bayesian update, the resulting posterior uncertainty would be practically negligible. We instead ignore this uncertainty, setting the posteriors to delta distributions centered on the measured values. This provides computational savings throughout the remainder of the calibration and allows us to use a fixed computational mesh in the digital twin.

#### Step 2: Calibrate material properties, \( e \)

We calibrate the Young’s modulus using a static load-displacement experiment. A known mass is placed on the main spar of the wing, 5cm from the wing tip, generating a static tip load with magnitude \( \hat{f} \). To account for error in the applied force, as well as the position of the force, we model the uncertainty using a Gaussian distribution with a 95% confidence interval equivalent to \( \pm 10 \)g. The resulting static tip displacement, \( \hat{x} \), is then measured, with measurement error modeled as an independent Gaussian with 95% confidence interval equal to \( \pm 1 \)mm. This process is repeated for eight total measurements as shown in Figure 5 (left): two times each for applied masses of 250g, 500g, 750g, and 1000g. Thus, the observational data is a set of eight measured load-displacement
pairs, \( O_2 = \{ \hat{f}, \hat{x} \} \). Each measured load-displacement pair is converted into an estimate of the aggregate wing stiffness, \( k = \hat{f} / \hat{x} \), then each of these estimated stiffness values is converted into an equivalent estimate of the Young’s modulus, \( \hat{e} \), via \( \hat{e} \).

Each value, \( \hat{e} \), can be incorporated into the estimate for \( e \) according to the Bayesian update formula

\[
p(\hat{e}|e) \propto p(\hat{e}|e)p(e)
\]

where \( p(e) \) is the Gaussian prior distribution, while \( p(\hat{e}|e) \) is the likelihood function corresponding to the measurement, which in this case is non-Gaussian (due to the \( 1/\hat{x} \) dependence). We estimate this likelihood density by sampling and fitting a kernel density. In particular, we draw \( 10^6 \) sample pairs from the Gaussian distributions centered on \( \hat{f} \) and \( \hat{x} \) respectively. Each sample pair is transformed into a sample of \( \hat{e} \) using the procedure described above. We then fit a kernel density estimate to these samples, which serves as an estimate of the likelihood \( p(\hat{e}|e) \). Each of these likelihood functions are shown in Figure 5 (center).

The Bayesian update is performed for each measurement iteratively, using a standard particle filter approach. We draw \( 10^6 \) samples from the prior distribution, \( p(e) \), and assign each sample a uniform weight. We then scale the weight of each sample by the corresponding likelihood, \( p(\hat{e}|e) \), before re-normalizing the weights, resulting in a weighted-particle approximation of the posterior density, \( p(e) \). This process is repeated for each of the eight measurements taken, incrementally updating the posterior estimate. After all eight measurements are incorporated, we arrive at the posterior estimate, \( p(e|O_2) \). From this weighted-particle approximation of the posterior density we can estimate the mean and standard deviation (reported in Table 1), as well as a posterior 95% confidence interval for \( e \). This confidence interval can be translated into a confidence interval in terms of the quantity of interest, \( k \), which is shown in Figure 5 (right). Finally, we can draw samples from the posterior using inverse transform sampling applied to the empirical cumulative density function. This enables us to propagate forward the uncertainty in \( e \) into the next calibration step.

**Step 3: Calibrate mass and damping, \( m, \alpha, \beta \)**

In the final calibration step, we seek to calibrate the dynamic response of the digital twin model by adding point masses that represent unmodeled hardware, as well as determining the appropriate coefficients for a Rayleigh damping model.

We first describe the experimental data, \( O_3 \), used for this step. We experimentally characterize the dynamic response of the wing using an initial condition response experiment. We apply an initial tip displacement of 10mm and release, resulting in a decaying oscillation of the tip displacement. During this oscillation we collect dynamic strain data, \( \hat{e}(t) \), at time points, \( \hat{t} \), measured from the moment the wing is released. Dynamic strain is measured using thin flexible piezoelectric patches that can be applied to different regions of the wing. These sensors communicate wirelessly, simplifying integration with the wing system. Data are taken at a sampling frequency of 100kHz, downsampled to 2kHz during post-processing. This experiment is repeated five times. From each raw dataset, we generate a power spectrum via fast Fourier transform. The power spectra for all experiments show two clear peaks corresponding to the first and second bending modes of the wing. We extract the locations of these peaks, which correspond to the first and second damped natural frequencies, \( \hat{\omega}^2 \). Using these frequencies, we seek to construct a two-mode reconstruction of the measured signal, \( \hat{e} \), of the form

\[
e^{\text{reconstructed}} = a_1 e^{-b_1 t} \cos(2\pi \hat{\omega}_1 (t - c_1)) + a_2 e^{-b_2 t} \cos(2\pi \hat{\omega}_2 (t - c_2)).
\]

We fit this model to the measured data by solving a non-linear least-squares problem for the coefficients, \( a_1, a_2, b_1, b_2, c_1, c_2 \), using a Levenberg-Marquardt algorithm. From this model we can extract the experimentally derived damping ratios, \( \zeta_i \), as well as the undamped natural frequencies, \( \hat{\omega}_i \), by solving the following system of equations:

\[
\begin{align*}
\hat{\omega}_1 &= \sqrt{1 - \zeta_1^2} \\
\hat{\omega}_2 &= \sqrt{1 - \zeta_2^2} \\
\end{align*}
\]

This procedure is repeated for each dataset, resulting in a total of five experimental estimates for the modal frequencies and damping ratios, \( O_3 = \{ \hat{\omega}_1, \hat{\omega}_2 \} \), for the first two bending modes, \( i = 1, 2 \). We average across all experimental datasets. In the following we refer to these averages by \( \hat{\omega}_i \) and \( \hat{\zeta}_i \) for simplicity.

Recall that our goal is to update the belief at time step \( t = 2 \), namely \( p(D_2|D_1, O_2, U_2) \), into a belief state at \( t = 3 \), namely \( p(D_3|D_2, O_3, U_3) \). We begin by drawing a sample of the calibrated parameters from the previous belief state. Since the calibrated geometric parameters are deterministic, this amounts to sampling a value for the remaining parameter, \( e \), from the posterior distribution \( p(e|\hat{e}) \) computed during the previous calibration step.

Given the observational data \( O_3 \), and a sampled value for \( e \), we seek to add to our sample a computed value for the point masses \( m \). Modifying these point masses in the computational model changes the mass matrix, \( M(d) \). Through the computational model, \( \hat{e} \), this in turn changes the natural frequencies, \( \omega_i \), predicted by the model. Using this model we formulate an optimization problem to fit the point masses, \( m = [m_{\text{serro}}, m_{\text{pital}}] \), so that the natural frequencies of the first two bending modes predicted by the model, \( \omega_1 \) and \( \omega_2 \), match the experimental data, \( \hat{\omega}_1 \) and \( \hat{\omega}_2 \), as closely as possible, while obeying the constraints (3). The objective function used in this optimization is the sum of relative frequency errors, and the optimization problem is solved using a Nelder-Mead gradient-free optimization method.

Finally, using the calibrated natural frequencies we compute the coefficients, \( \alpha, \beta \), in the Rayleigh damping model according to the system of equations

\[
\begin{bmatrix}
1/\omega_1 & \omega_1 \\
1/\omega_2 & \omega_2
\end{bmatrix}
\begin{bmatrix}
\alpha \\
\beta
\end{bmatrix} =
\begin{bmatrix}
\hat{\zeta}_1 \\
\hat{\zeta}_2
\end{bmatrix}.
\]
Note that this system is derived from (7), and ensures by construction that the computational damping ratios, $ζ_i$, exactly match the experimentally estimated damping ratios, $\hat{ζ}_i$, for both modes $i = 1, 2$. The computed values for $m, α$, and $β$, combined with the sampled value for $e$ and the calibrated geometric parameters $g$, constitute a sample from the updated belief state. We repeat this entire optimization procedure for 100 samples in order to build up a sample-based approximation of the final calibrated estimate of the digital state.

The quantities of interest for this time step are the posterior computational estimates of the modal frequencies and damping ratios, $ω_i$ and $ζ_i$. The reward is the average discrepancy between these values and the corresponding experimentally estimated values, $\hat{ω}_i$ and $\hat{ζ}_i$ respectively.

**Data availability**

The experimental data and computational model output that support the findings of this study are available in [REPOSITORY NAME TO APPEAR] with the identifier(s) [DOI TO APPEAR].

**Code availability**

The code used to generate the results in this study are available in the public repository [REPOSITORY INFORMATION TO APPEAR]. The structural analysis software used to generate the results in this paper is Akselos Integra v4.5.9. Since the Akselos Integra software is proprietary and was used under license, we are unable to provide its source code. Instead, the model input and output data generated in this study is provided directly in the repository.

**Acknowledgments**

This work was supported in part by AFOSR grant FA9550-16-1-0108 under the Dynamic Data Driven Application Systems Program, the SUTD-MIT International Design Center, the AFOSR MURI on managing multiple information sources of multi-physics systems awards FA9550-15-1-0038 and FA9550-18-1-0023, and the AEOLUS center under US Department of Energy Applied Mathematics MMICC award DE-SC0019303.

**Competing Interests**

Jessara Group sensors were used in the UAV experimental work described in this paper. Co-author Jacob Pretorius is a co-founder of Jessara. Purchase of the sensors for use in the research was reviewed and approved in compliance with all applicable MIT policies and procedures.

**References**

[1] Adil Rasheed, Omer San, and Trond Kvamsdal. Digital twin: Values, challenges and enablers from a modeling perspective. *IEEE Access*, 8:21980–22012, 2020.

[2] Eric J. Tuegel, Anthony R. Ingraffea, Thomas G. Eason, and S. Michael Spottswood. Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*, 2011.

[3] Edward Glaessgen and David Stargel. The digital twin paradigm for future NASA and US air force vehicles. In 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA, page 1818, 2012.

[4] Chenzhao Li, Sankaran Mahadevan, You Ling, Sergio Choze, and Liping Wang. Dynamic Bayesian network for aircraft wing health monitoring digital twin. *AAIA Journal*, 55(3):930–941, 2017.

[5] Mat Podskarbi and David J. Knezevic. Digital twin for operations — present applications and future digital thread. In *Offshore Technology Conference*. Offshore Technology Conference, 2020.

[6] Koen Bruynseels, Filippo Santoni de Sio, and Jeroen van den Hoven. Digital twins in health care: Ethical implications of an emerging engineering paradigm. *Frontiers in Genetics*, 9:31, 2018.

[7] Luis F Rivera, Miguel Jiménez, Prashanti Angara, Norha M Villegas, Gabriel Tamura, and Hausi A Müller. Towards continuous monitoring in personalized health care through digital twins. In *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering*, pages 329–335, 2019.

[8] Barbara Rita Barricelli, Elena Casiraghi, Jessica Gliozzo, Alessandro Petrini, and Stefano Valtolina. Human digital twin for fitness management. *IEEE Access*, 8:26637–26664, 2020.

[9] Han Yu, Chunyan Miao, Cyril Leung, and Timothy John White. Towards AI-powered personalization in MOOC learning. *npj Science of Learning*, 2(1):1–5, 2017.

[10] Neda Mohammadi and John E Taylor. Smart city digital twins. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–5. IEEE, 2017.

[11] Frederica Darema. Dynamic data driven applications systems: A new paradigm for application simulations and measurements. In *International Conference on Computational Science*, pages 662–669. Springer, 2004.

[12] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques*. MIT press, 2009.

[13] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson Education Limited, 2002.

[14] Steven L. Brunton and J. Nathan Kutz. *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2019.

[15] Jenný Brynjarsdóttir and Anthony O’Hagan. Learning about physical parameters: The importance of model discrepancy. *Inverse Problems*, 30(11):114007, 2014.
[16] Dirk Hartmann, Matthias Herz, and Utz Wever. Model order reduction a key technology for digital twins. In Reduced-Order Modeling (ROM) for Simulation and Optimization, pages 167–179. Springer, 2018.

[17] Stefanie J. Salinger, Michael G. Kapteyn, Cory Kays, Jacob V.R. Pretorius, and Karen E. Willcox. A hardware testbed for dynamic data-driven aerospace digital twins. In International Conference on Dynamic Data Driven Application Systems, pages 37–45. Springer, 2020.

[18] Michael G. Kapteyn, David J. Knezevic, D.B.P. Huynh, Minh Tran, and Karen E. Willcox. Data-driven physics-based digital twins via a library of component-based reduced-order models. International Journal for Numerical Methods in Engineering, 2020.

[19] Anil K. Chopra et al. Dynamics of structures. Pearson Education Upper Saddle River, NJ, 2012.