Parameter estimation for complex thermal-fluid flows using approximate Bayesian computation

Jason D. Christopher,[1] Nicholas T. Wimer,[1] Caelan Lapointe,[1] Torrey R. S. Hayden,[1] Ian Grooms,[2] Gregory B. Rieker,[1] and Peter E. Hamlington[1,*]

[1]*Department of Mechanical Engineering, University of Colorado, Boulder, Colorado 80309, USA*

[2]*Department of Applied Mathematics, University of Colorado, Boulder, Colorado 80309, USA*

(Received 7 May 2018; published 11 October 2018)

Approximate Bayesian computation (ABC) is a data-driven technique that uses many low-cost numerical simulations to estimate unknown physical or model parameters (e.g., boundary conditions and material properties), as well as their uncertainties, given reference data from real-world experiments or higher-fidelity numerical simulations. In this study, ABC is used to estimate unknown parameters in complex thermal-fluid flows, and the technique is demonstrated on a periodically forced high-temperature jet and a steadily forced helium-air plume. In the first case, computational reference data are used to assess the accuracy of ABC when estimating the frequency, amplitude, and mean of the periodic velocity forcing at the jet inlet. These tests show that ABC provides accurate and reasonably certain estimates of inflow parameters even when the model simulations imperfectly represent the physics underlying the reference data. These tests also show that measurements far from the inlet can be used to perform the estimation, and that temperature measurements can be used to infer velocity inflow parameters. For the second case, ABC is used to estimate the inlet Richardson number and its uncertainty given experimental measurements of the Strouhal number within the plume. Once again, the approach is able to accurately estimate unknown parameters with relatively low uncertainty. As a result, ABC is shown to be a versatile technique for estimating unknown physical parameters when knowledge of a real-world system is limited or incomplete.

DOI: 10.1103/PhysRevFluids.3.104602

I. INTRODUCTION

Computational simulations are widely used to design and analyze systems involving complex thermal-fluid flows, from microscale heat transfer in thermal management devices to liquid propulsion in heavy-lift rockets. Nearly all such simulations are intended to provide three-dimensional (3D) spatially and temporally resolved numerical solutions of physics-based governing equations for realistic geometries, boundary conditions, and material properties. Prior to making simulation-based design and operational decisions, however, computational accuracy must be demonstrated by validating against (typically experimental) reference data for realistic or canonical test cases. As simulation fidelity has improved over the past decade with the development of higher-order and geometry-resolving numerical techniques, as well as with the increasing use of petascale computing resources to achieve very fine spatial resolutions and improved physical realism, a long-standing difficulty in such validation efforts has come into sharper focus.

Namely: *Even if a simulation is able to solve physically realistic governing equations with high accuracy, unavoidable uncertainties in real-world boundary conditions, material properties, and*

*peh@colorado.edu

other parameters result in ambiguity as to whether the computational and real-world systems are actually equivalent. That is, there are few assurances that discrepancies between computational and experimental results during a validation test are not simply due to limitations in the characterization of the real-world system. Cases where knowledge of a real-world system is limited include experiments that were not originally intended for validation purposes, systems with limited access, or parameters that are difficult to measure directly [1].

In broad terms, this difficulty can be addressed by proposing distributions of likely parameter values, performing simulations with parameters sampled from those distributions to determine the spread of outcomes in a particular quantity of interest within the flow, and then using statistical inference to determine distributions for unknown parameter values [1–4]. This approach can, in principle, be attempted using full Bayesian analyses, which have recently gained popularity for parameter estimation in engineering applications [5–9]. However, in nearly all such analyses, the requisite components of Bayes' theorem (specifically, the likelihood function) may be unknown or enormously costly to compute, and, consequently, non-physics-based reduced-order surrogate models have often been used to sample the unknown parameter space [6,7,10–12]. Optimization techniques have also been used for parameter estimation (e.g., Refs. [13–17]), but these methods seek to provide single values of unknown parameters, with no intrinsic measure of uncertainty when using potentially imperfect computational models and real-world data.

In the present study, approximate Bayesian computation (ABC), a data-driven Bayesian technique adopted from the biological and geophysical sciences, is used to estimate unknown parameters for complex, turbulent thermal-fluid flows. The power of ABC lies in the fact that far fewer simulations are required than in full Bayesian analyses since ABC does not require a likelihood function, thus permitting the use of physics-based models. Rather than attempting to match the reference data at all locations and times, as in full Bayesian analyses, the ABC method develops an approximate posterior distribution for unknown parameters through comparison of summary statistics from the reference data and model simulations. A wide variety of reference data can be used to drive the estimation, including measurements that are only indirectly related to parameters of interest. Because the technique naturally provides posterior distributions for unknown parameters, Bayesian confidence intervals can be obtained along with parameter estimates. Once parameter estimates are obtained, either they can be used in higher-fidelity numerical simulations of the same system, or they can be considered as part of the description of the real-world system itself (with appropriate uncertainty caveats).

This study is one of a growing number of efforts to use data-driven methods with either experimental or higher-fidelity computational reference data to improve accuracy and quantify uncertainty in turbulent flow simulations. For example, ensemble Kalman filtering (EnKF) approaches have been used to assimilate reference data in simulations [18–21] and to infer turbulence model discrepancies [22]. In such approaches, reference data are used to update the state of a simulation or parameters in the simulation as the simulation progresses. Generally, however, EnKF methods require extensive and high-quality reference data to provide reliable state and parameter estimates. The maximum *a posteriori* (MAP) method has also been used to infer both parameter [23] and model [24] uncertainties. This method maximizes the posterior using either an analytical function or a Monte Carlo approach, but the analytical function used to represent the posterior (or the likelihood function) is often known only approximately, and the number of simulations required in even Markov chain Monte Carlo approaches can be enormous. Recently, the ABC method has been used to infer Arrhenius parameters for chemical kinetics rate coefficients in the context of combustion [25].

The present study is specifically targeted at problems for which the reference data are spatially and temporally sparse, statistical in nature, or otherwise lacking in sufficient detail to update the model simulations at all locations and times (precluding the use of EnKF methods) and for which the true posterior is either intractable or prohibitively expensive to compute (precluding the use of MAP methods). As a demonstration of ABC, this paper outlines the estimation of inflow boundary conditions for two compressible turbulent flows: (1) a periodically forced high-temperature jet and

(2) a weakly but steadily forced helium-air plume. Both cases are two-dimensional (2D) but are physically complex due to compressibility and the corresponding coupling between temperature, density, buoyancy, and forced advection. The first case is intended to demonstrate the utility and breadth of ABC by using computational reference data to estimate the frequency, amplitude, and mean of the forced inflow. Since these parameters and the underlying physical model are precisely known for the computational data, the success of the ABC approach can be quantified. The second case is a demonstration of ABC using experimental measurements in an engineering context.

It should be noted that the present study is focused on demonstrating the use of ABC for estimating *physical* parameters that are difficult to obtain for real-world systems, as opposed to calibrating *model* parameters used in reduced-order or engineering models of more complex phenomena. Such model parameter calibrations are also possible using ABC and are the subject of future research [26].

II. APPROXIMATE BAYESIAN COMPUTATION

The ABC framework is a relatively new approach for linking reference data to physical and model parameters [27–33]. It was first conceived by Rubin [34] and then subsequently applied by Pritchard *et al.* [35] to the study of population dynamics. Although the technique has traditionally been applied in biology [35–41], it has recently gained traction in geophysics [42–48], primarily for calibrating parameters in reduced-order models of complex phenomena. It has also been used recently in the context of chemical kinetics modeling for combustion [25].

From a fundamental perspective, the ABC method produces a posterior distribution of unknown parameters given some reference data. This occurs by comparing observations or summary statistics from the reference data to corresponding data or summary statistics from lower-cost physics-based "model" simulations (e.g., simulations with coarse spatial resolutions or lower-order numerical schemes). The model simulations are repeated many times, with each simulation using parameters drawn from a prior distribution. The prior distribution is the best guess at the span of the unknown parameter space; it must be wide enough to contain the true parameter values, but narrow enough to keep the task computationally manageable. Parameter values are retained if the model data or summary statistics are similar to the reference observations. Once many such candidates are obtained, a posterior distribution is formed, providing estimates for the most likely parameter values as well as their uncertainties, given the reference data.

The specific ABC algorithm used in this work is Method D from Marjoram *et al.* [28]. Given the summary statistic $\boldsymbol{\mathcal{S}}$ obtained from reference data $\boldsymbol{D}$, this algorithm involves the following steps (see also the schematic in Fig. 1):

(1) Generate parameters $\boldsymbol{\theta}$ from the prior distribution $P(\boldsymbol{\theta})$.

(2) Simulate approximate data $\hat{\boldsymbol{D}}$ using parameters $\boldsymbol{\theta}$ and compute the approximate statistic $\hat{\boldsymbol{\mathcal{S}}}$.

(3) Calculate the distance $\delta(\boldsymbol{\mathcal{S}}, \hat{\boldsymbol{\mathcal{S}}})$ between the reference, $\boldsymbol{\mathcal{S}}$, and simulated, $\hat{\boldsymbol{\mathcal{S}}}$, statistics.

(4) Accept $\boldsymbol{\theta}$ if $\delta(\boldsymbol{\mathcal{S}}, \hat{\boldsymbol{\mathcal{S}}}) \leqslant \varepsilon$ (where ε is the "rejection distance") and build the posterior distribution of accepted parameters, denoted $P[\boldsymbol{\theta} \mid \delta(\boldsymbol{\mathcal{S}}, \hat{\boldsymbol{\mathcal{S}}}) \leqslant \varepsilon]$.

(5) Return to step 1 and repeat a total of N times until a reasonable estimate is obtained for the posterior distribution.

In general, the parameters $\boldsymbol{\theta}$, data $\boldsymbol{D}$ and $\hat{\boldsymbol{D}}$, and summary statistics $\boldsymbol{\mathcal{S}}$ and $\hat{\boldsymbol{\mathcal{S}}}$ are multidimensional and are correspondingly indicated by boldface notation.

In broad terms, the prior distribution is simply the initial guess for the values of the true parameters. Greater confidence in the values of unknown parameters permits a more concentrated prior distribution (e.g., a Gaussian instead of a uniform distribution). The only requirement on the prior is that its range spans the true values of the unknown parameters. In many practical cases, a wide prior can be used initially to gain an approximate idea of the true parameter values, and then a narrower prior can be used to determine parameter values with greater precision.

The use of summary statistics is intended to reduce the overall computational cost of the parameter estimation. Although it would be ideal to compare all of the available data $\boldsymbol{D}$ to

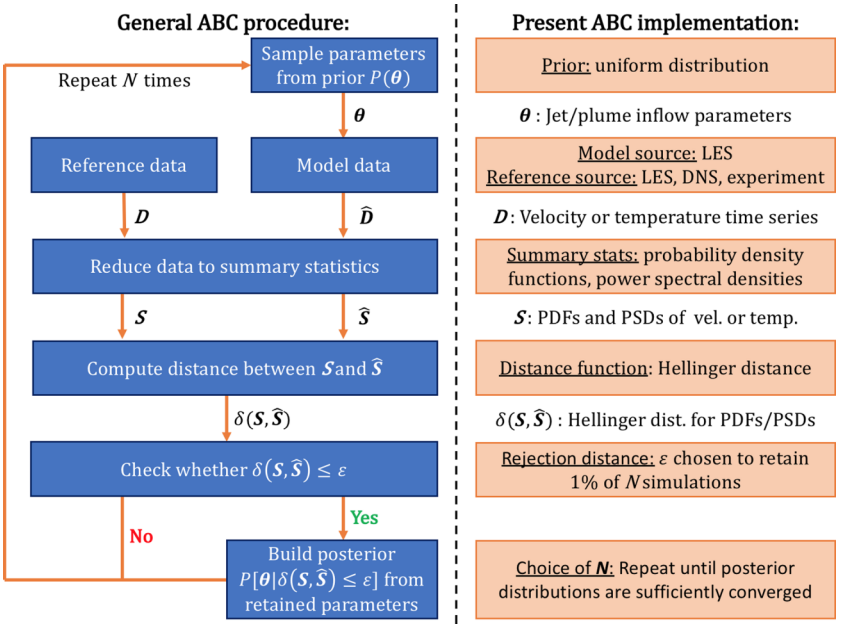

FIG. 1. Schematic of the general ABC approach corresponding to Method D from Marjoram *et al.* [28] and of the specific ABC implementation used in the present study.

all of the simulation data $\hat{\boldsymbol{D}}$, this is typically a very high-dimensional problem and is not computationally feasible in general. Thus, instead, summary statistics are used. A key challenge in successfully implementing ABC is to identify relevant summary statistics that significantly reduce the dimensionality of the data, while still maintaining sufficient identifiability of unknown parameters [49]. Examples of summary statistics include averages, standard deviations, probability distribution functions (PDFs), and power spectral densities (PSDs).

The choice of distance function is typically based on the form of the summary statistic. For example, the root-mean-square error can be used to compare spatial profiles of average quantities, and the Hellinger distance [50] or Kullback-Leibler divergence [51] can be used to compare PDFs. The specific choice of rejection distance, ε, is typically determined by a balance between computational cost and parameter estimate uncertainty. Smaller values of ε lead to higher rejection rates and reduced uncertainty in estimated parameter values [29] but also increase the number of model simulations necessary to generate a sufficiently converged posterior.

With respect to evaluating the success of the ABC parameter estimation, "accuracy" can be assessed either by comparing the approximate posteriors generated using ABC to the true posteriors, or by comparing the estimated parameter values to the true parameter values, along with a measure of certainty in the estimated values. In the present study, the true posteriors are not known analytically, and they are prohibitively expensive to compute numerically. As a result, the accuracy of the ABC approach is quantified here by computing 95% Bayesian confidence intervals (sometimes called credible intervals), denoted C_B, from the approximate parameter posteriors, then determining whether C_B contains the true parameter value. The ultimate measure of accuracy in the present study is indicated by how closely a central tendency of the posterior (such as the mean) matches the true parameter value. The degree of certainty in the estimated parameter value is determined by the width of C_B, where narrower intervals indicate more certain parameter estimates.

Additional details on the specific choices for the prior, summary statistics, distance function, and rejection distances used in the present ABC implementation are outlined in the following sections and are summarized in Fig. 1.

III. DEMONSTRATION OF ABC USING COMPUTATIONAL REFERENCE DATA

As a demonstration of ABC, inflow parameters for a periodically forced high-temperature turbulent buoyant jet [52–56] are estimated using computational reference data. This case is examined due to the simplicity of the geometry combined with the complexity of the high-temperature unsteady compressible flow physics. Large eddy and direct numerical simulations (LES and DNS, respectively) are used as reference data, and model simulations are performed using LES.

The use of computational reference data allows the success of ABC to be assessed when the physics underlying the reference data, as well as all system parameters, are known exactly. Three questions, in particular, are addressed: (1) How accurately can ABC estimate unknown parameters when the reference data physics are exactly reproduced by the model simulations? (2) How accurately can ABC estimate unknown parameters using model simulations that imperfectly represent the physics of the reference case? (3) How accurately can ABC estimate unknown parameters using reference data that are only indirectly connected to the parameters of interest?

The first question represents a "best case scenario" where there is zero model error, and it is examined using both model and reference data from LES. Despite the exact correspondence between the governing equations solved in the model and reference LES, however, the flow fields obtained from each of the simulations differ on a local and instantaneous basis due to the use of different random initial conditions. The second question pertains to a more realistic application of ABC where the physics governing the reference data are not exactly represented by the model simulations. This question is examined here using LES model simulations and DNS reference data and is further addressed in Sec. IV using experimental reference data. The third question addresses the identifiability of unknown parameters using different types of reference data and is examined by performing ABC at different heights above the inlet using either velocity or temperature measurements.

It should be noted that the test where both the model and reference data are obtained from LES most closely resembles an observing system simulation experiment (OSSE), which is a common technique used for the validation of data assimilation methods [57,58]. In the present context, the OSSE represented by this test allows the choices of prior, summary statistic, distance function, and rejection distance to be evaluated in the absence of model error. In this sense, these tests indicate whether the ABC approach can ever be expected to succeed and to what extent. The following analysis illustrates that, even in such a "best case scenario," the ABC parameter estimation does not perfectly recover the true parameter values due to the presence of stochasticity in the model and reference data. That is, the model and reference data are never in perfect local and instantaneous agreement because each of the simulations are initialized with different random temperature fields (even when using identical boundary parameter values). Thus, the true parameter values are only ever recovered approximately, even when the governing equations used to generate the model and reference data are identical.

The test where the model data are obtained from LES and the reference data from DNS is intended to address a common concern when parameter estimation methods are evaluated using OSSEs: namely, that assessments of the method are likely to be overly optimistic due to the exact correspondence between the physics and numerics underlying the model and reference simulations. This issue, sometimes referred to as an "inverse crime" [59,60] or the "identical/fraternal twin problem" [61,62], is typically resolved by using different models, for example with different discretizations, numerics, or physics, for the generation of the reference data and for the parameter estimation. Here the LES-DNS tests are intended to provide a more realistic assessment of the accuracy of the ABC method than the LES-LES tests, since both the grid resolution and underlying governing equations in the LES and DNS are different. The test using experimental reference data and LES model data in Sec. IV is similarly intended to enable a more realistic assessment of the accuracy of the ABC approach.

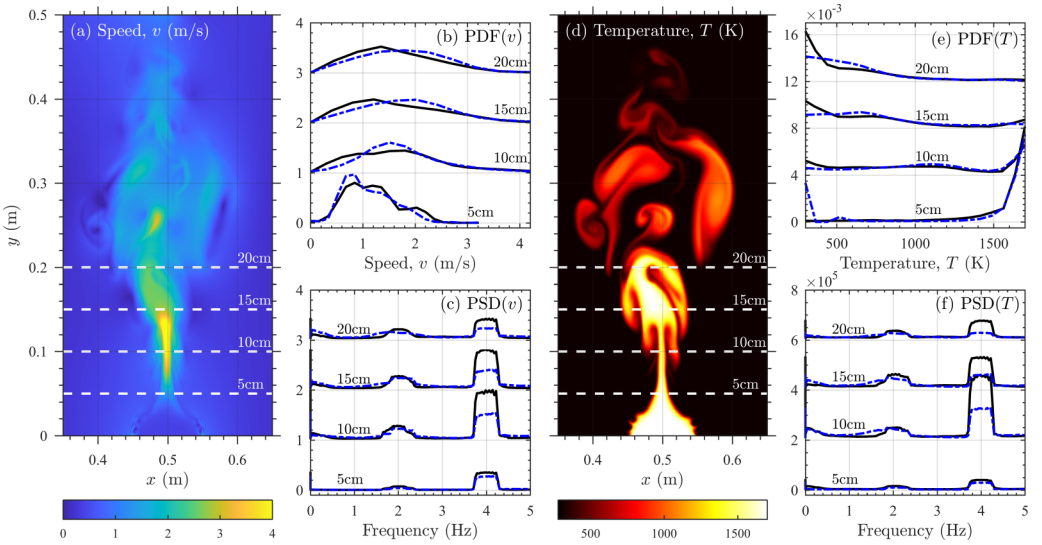

FIG. 2. Representative fields of speed v (a) and temperature T (d) from DNS for the periodically forced turbulent buoyant jet. Panels (b) and (c) show reference probability density functions (PDFs) and power spectral densities (PSDs), respectively, of v at heights of 5, 10, 15, and 20 cm. PDFs and PSDs at different heights are shifted vertically for clarity. Panels (e) and (f) show corresponding PDFs and PSDs of T. Solid black lines in panels (b), (c), (e), and (f) show DNS reference data and dash-dot blue lines show LES reference data. PSDs are computed using Thomson's multitaper estimate [69].

A. Physical and numerical setup

The reference buoyant jet has a sinusoidally varying inlet velocity with a frequency of 4 Hz, an amplitude of 0.2 m/s, and a mean of 0.5 m/s. The inflow velocity is spatially uniform across the inlet. The physical domain is 1×2 m, and the inlet is 0.1 m wide. The ambient and inlet temperatures are 300 K and 1700 K, respectively, resulting in substantial density differences and buoyancy-driven flow. Instantaneous fields of velocity magnitude (i.e., speed) and temperature are shown in Fig. 2.

All of the computations were performed using the FireFoam solver [63] within OpenFOAM [64,65]. The simulations were restricted to two dimensions to minimize computational cost, although the ABC method is equally applicable in three dimensions. For the LES, the compressible filtered Navier-Stokes equations were solved with second-order accuracy in space and time using the one-equation eddy viscosity model [66] for closure of subgrid-scale stresses. The LES domain was discretized using 61 088 cells, with grid stretching in the vertical direction and two levels of refinement near the jet inlet. For the DNS computations, the compressible Navier-Stokes equations were solved in conjunction with the total energy equation at a high spatial resolution, again using a second-order accurate numerical scheme in space and time. The DNS domain was discretized using 745 472 cells with two levels of grid refinement, giving a grid that was more than an order of magnitude larger than the LES grid.

For both the LES and DNS, the ideal gas equation was used to relate state variables, and fluid viscosity and specific heat varied with temperature according to the Sutherland model [67] and JANAF tables [68], respectively. The LES and DNS numerical schemes were also the same, with Crank-Nicolson time stepping and Gauss integration with linear interpolation for spatial derivatives. Pressure-velocity coupling was accomplished using the PIMPLE algorithm, which combines the pressure-implicit split-operator and the semi-implicit method for pressure-linked equations. An adaptive time step was used to limit the maximum CFL condition to 0.5 for the LES and 0.2 for the DNS.

B. ABC implementation

In this demonstration, ABC is used to estimate the frequency, amplitude, and mean of the periodic forcing at the inlet of the buoyant jet given either LES or DNS reference data, with model simulations from LES. The reference data consist of time series of speed or temperature measured at nine heights above the inlet, ranging from roughly 0.5 mm up to 0.2 m. Time series at each height were recorded over approximately 14 s after a 1 s initialization period, and PDFs and PSDs were computed for each time series; these are the ABC summary statistics. The PSDs were calculated using Thomson's multitaper estimate [69,70], which provides a robust estimate of the PSD by reducing energy leakage across frequencies and reducing variance [71].

To ensure statistical convergence of the reference data, ensembles of 100 statistically independent simulations were created using both LES and DNS. Each simulation had the same periodically forced inflow at a frequency of 4 Hz, an amplitude of 0.2 m/s, and a mean of 0.5 m/s, but was initialized using different stochastically generated temperature fields. The final reference data for both the LES and DNS cases were then created by averaging the PDFs and PSDs for each of the 100 simulations in each ensemble, at each of the nine heights, giving the reference data sets shown, for example, in Figs. 2(b) and 2(c) for speed and in Figs. 2(e) and 2(f) for temperature. Moreover, Fig. 2 shows that the LES and DNS reference data are in reasonably good agreement.

Approximately 10^4 additional LES model cases were generated for the parameter estimation. Each simulation was executed with a unique set of inlet parameters randomly sampled from prior distributions for the frequency, amplitude, and mean of the sinusoidal velocity oscillations at the inlet. The priors were uniform with bounds of 3.6–4.4 Hz for the frequency, 0.0–0.4 m/s for the amplitude, and 0.0–1.0 m/s for the mean. Uniform priors were chosen to avoid the creation of anomalous biases in the posterior distributions. The width of the priors for the amplitude and mean were selected to be $\pm 100\%$ the true parameter values, while the width of the frequency prior is $\pm 10\%$ the true value; it will be seen in the following that even with such a narrow frequency prior, the posteriors go to zero at the edges of the prior for all heights. Although wider priors would result in greater recovery of the edges of the posterior distributions, nearly all of the posteriors obtained in this study display a pronounced maximum within the range of values present in the priors. It should also be noted that the current priors are centered on the true values and that off-centered priors could be formed, for example, by using widths that are $[-100\%, +200\%]$ the true values for the amplitude and mean and $[-10\%, +20\%]$ for the frequency. However, once more, the success of the ABC method will not be affected by the use of such off-centered priors provided that the posteriors already display a distinct maximum within the bounds of the current centered priors, as is the case for nearly all of the tests performed here.

The Hellinger distance [50] was used to quantify the agreement between summary statistics from the reference and model data. Values for ε were chosen such that a fixed percentage of the tested parameters were retained in the final posterior [37,40]. Parameter rejection was first performed using the Hellinger distance for the PSD, where ε was selected to reject 80% of all parameters tested. Subsequently, rejection was performed using the Hellinger distance for the PDF, with ε chosen to reject 95% of the remaining parameters. With this sequential approach, only 1% of the parameters tested were included in the final posteriors.

C. Inflow parameter estimation using LES reference data

Figure 3 shows marginal posterior distributions for each inflow parameter using both speed and temperature reference data from LES at different heights. The posteriors are represented by Gaussian kernel estimations with bandwidths from Scott's normal reference rule [72]. As noted previously, these tests represent a best-case scenario for the ABC method where both the model and reference data physics are identical. Correspondingly, Figs. 3(a), 3(c), and 3(e) show that the true parameter values are captured by the ABC posteriors when using speed reference data from LES. In general, the uncertainty is smallest at locations close to the inlet, although reasonable estimates of the true parameters are still obtained at higher locations.

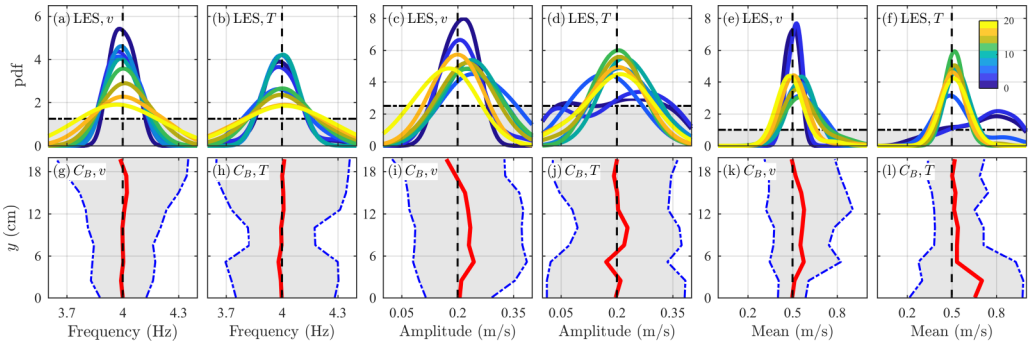

FIG. 3. Top row: Marginal posterior distributions from ABC (visualized using Gaussian kernel estimation) for the frequency (a, b), amplitude (c, d), and mean (e, f) of the periodically forced turbulent buoyant jet using speed (a, c, e) and temperature (b, d, f) reference data from LES. Posteriors are calculated using measurements at heights from 0 to 20 cm (indicated by colors from blue to yellow). True parameter values are shown by vertical black dashed lines. Horizontal dash-dot lines and gray regions show the priors. Bottom row: Vertical profiles of 95% Bayesian confidence intervals, C_B, for the marginal posterior distributions in all panels. Blue dash-dot lines and gray shading indicate the 95% Bayesian confidence regions, and the solid red lines show the mean values of the posteriors. True parameter values are shown by vertical black dashed lines.

For the speed reference data, Figs. 3(g), 3(i), and 3(k) show that C_B is narrowest near the inlet when using LES reference data. This indicates that identifiability of the unknown parameters is greatest near the inlet when using the speed as reference data, although even higher in the domain it is still possible to estimate inflow parameters, albeit with less confidence. Figures 3(g), 3(i), and 3(k) also show that mean values from the posteriors are in good agreement with the true values, particularly near the inlet for the amplitude and mean, and at all heights for the frequency.

The identifiability of unknown parameters using different measured quantities can be assessed by repeating the ABC procedure using PDFs and PSDs computed from temperature time series. Figure 3(b) shows that frequency predictions based on temperature reference data are all roughly centered on the true parameter value, and Fig. 3(h) shows that the posterior means are in good agreement with the true value. The amplitude and mean posteriors in Figs. 3(d) and 3(f) are centered around the true parameter values in the region from approximately 0.5 to 1.5 jet widths above the inlet. Below these heights, however, there is not enough variation in the temperature field to aid in parameter estimation, given the uniform inflow temperature. Because the PDFs were nearly identical at such low heights, the ABC procedure was modified such that 15% of the simulated parameters were accepted based solely on the PSD comparison. Higher than 1.5 jet widths above the inlet, the connection between the temperature field and the inlet velocity becomes weaker, resulting in reduced identifiability of velocity boundary conditions using temperature measurements. Despite these limitations, as shown in Figs. 3(h), 3(j), and 3(l), all parameters are predicted for all heights within the C_B intervals for each posterior.

The amount of information gained about the unknown parameters during the ABC procedure can be quantified using the Kullback-Leibler (KL) divergence [51] between the prior and posterior distributions. Higher values of the KL divergence indicate that the posterior is significantly different than the prior, while values close to zero indicate that the posterior is similar to the prior and that little information about unknown parameters has been gained during the ABC procedure.

Figure 4 shows vertical profiles of the KL divergence for speed and temperature reference data from LES, revealing that, in all cases, at least some information about the unknown parameters is gained at nearly all locations within the domain. For the speed reference data in Fig. 4(a), the greatest information about the unknown parameters is obtained near the jet inlet, and the posteriors for the frequency generally provide more information higher in the domain than the posteriors for the amplitude and mean. By contrast, for the temperature reference data in Fig. 4(b), the greatest

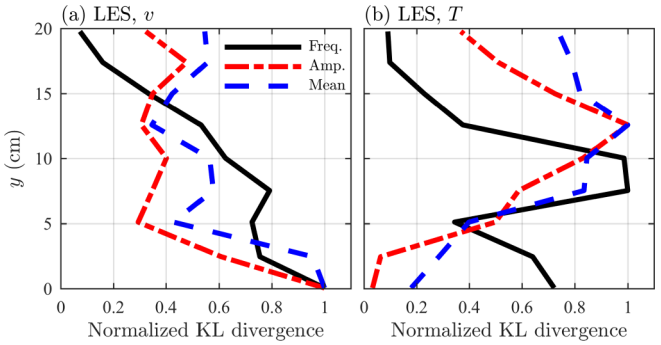

FIG. 4. Vertical profiles of Kullback-Leibler (KL) divergences for the marginal posterior distributions in Fig. 3 obtained using LES reference data. Divergences for the frequency (black solid lines), amplitude (red dash-dot lines), and mean (blue dashed lines) posteriors are shown for speed (a) and temperature (b) reference data. Each KL divergence vertical profile is normalized by its respective maximum value.

information is obtained higher in the domain, close to 10 cm (corresponding to one jet width above the inlet). In general, only the temperature reference data near the jet inlet fail to provide significant information regarding unknown parameter values.

D. Inflow parameter estimation using DNS reference data

As noted previously, evaluations of the ABC method based on tests where the same model is used for the reference data and the parameter estimation (as in the LES-LES tests described in the previous section) are likely to be overly optimistic. To address this issue, tests where the reference data are generated by DNS and the parameter estimation is performed using LES have also been carried out.

Figure 5 shows marginal posterior distributions for each inflow parameter using both speed and temperature reference data from DNS at different heights. As with the LES-LES results shown

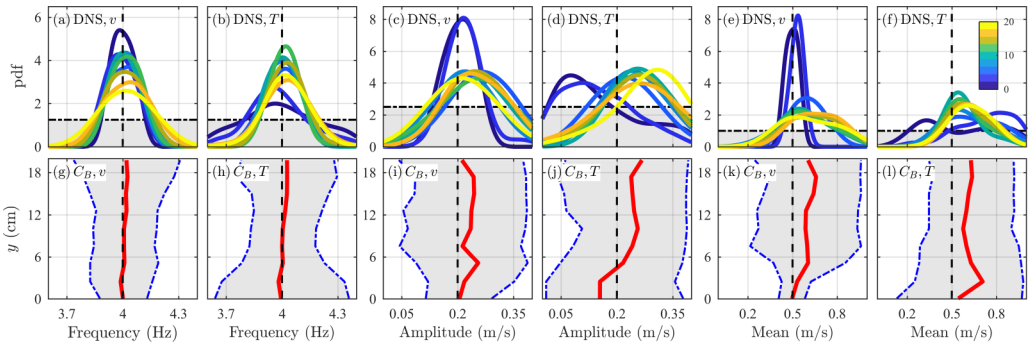

FIG. 5. Top row: Marginal posterior distributions from ABC (visualized using Gaussian kernel estimation) for the frequency (a, b), amplitude (c, d), and mean (e, f) of the periodically forced turbulent buoyant jet using speed (a, c, e) and temperature (b, d, f) reference data from DNS. Posteriors are calculated using measurements at heights from 0 to 20 cm (indicated by colors from blue to yellow). True parameter values are shown by vertical black dashed lines. Horizontal dash-dot lines and gray regions show the priors. Bottom row: Vertical profiles of 95% Bayesian confidence intervals, C_B, for the marginal posterior distributions in all panels. Blue dash-dot lines and gray shading indicate the 95% Bayesian confidence regions, and the solid red lines show the mean values of the posteriors. True parameter values are shown by vertical black dashed lines.

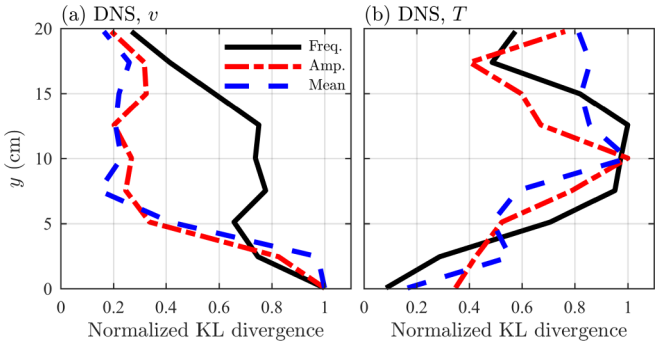

FIG. 6. Vertical profiles of Kullback-Leibler (KL) divergences for the marginal posterior distributions in Fig. 3 obtained using DNS reference data. Divergences for the frequency (black solid lines), amplitude (red dash-dot lines), and mean (blue dashed lines) posteriors are shown for speed (a) and temperature (b) reference data. Each KL divergence vertical profile is normalized by its respective maximum value.

in Fig. 3, Figs. 5(a), 5(c), and 5(e) show that the true parameter values are captured by the ABC posteriors when using speed reference data, with generally narrower posteriors close to the inlet. This is also indicated by the variations of C_B in Figs. 5(g), 5(i), and 5(k), where it can be seen that C_B is narrowest near the inlet when using speed reference data. The width of C_B consistently increases with height as the distance from the inlet increases. For each of the unknown parameters, posterior means are generally close to the true parameter values.

For the temperature reference data from DNS, Figs. 5(b), 5(d), and 5(f) show that the marginal posterior distributions once again capture the true parameter values with reasonable success. There is some bias in the posteriors for the DNS reference data (particularly for the estimates of the amplitude and mean), but the true parameter values are nevertheless contained within reasonably narrow 95% Bayesian confidence intervals. This is also shown in Figs. 5(h), 5(j), and 5(l), where C_B is widest for the temperature-based DNS reference data near the inlet and narrowest roughly one jet width above the inlet.

Vertical profiles of the KL divergence for the DNS reference data are shown in Fig. 6. As with the LES-LES results shown in Fig. 4(a), Fig. 6(a) shows that speed reference data from DNS provides the greatest information about unknown parameters near the jet inlet, with the frequency posteriors providing more information higher in the domain than the posteriors for the amplitude and mean. For the temperature reference data in Fig. 6(b), the greatest information is once again obtained roughly one jet width above the inlet. In general, the trends shown in Figs. 4 and 6 are similar for the LES and DNS reference data, respectively.

Taken together, the results for the DNS reference data in Figs. 5 and 6 indicate that the ABC approach is successful even in non-overly optimistic tests where the reference and model data are significantly different. Overall, the parameter estimation is generally more accurate using LES, as opposed to DNS, reference data, but this is unsurprising given the additional model error introduced when using the DNS reference data. As a result, ABC is able to provide predictions for unknown parameters using either exact or imperfect model simulations (as compared to the reference data), and for locations that are indirectly related to the parameters of interest (i.e., using measurements at locations far from the inlet).

IV. DEMONSTRATION OF ABC USING EXPERIMENTAL REFERENCE DATA

To demonstrate the utility of the ABC approach when the reference data are obtained from an experiment, the ABC technique is next applied to a steadily forced planar plume for which experimental data are available from Cetegen *et al.* [73]. In this experiment, a helium-air mixture is weakly forced into 300 K ambient air, resulting in a natural oscillatory behavior of the plume.

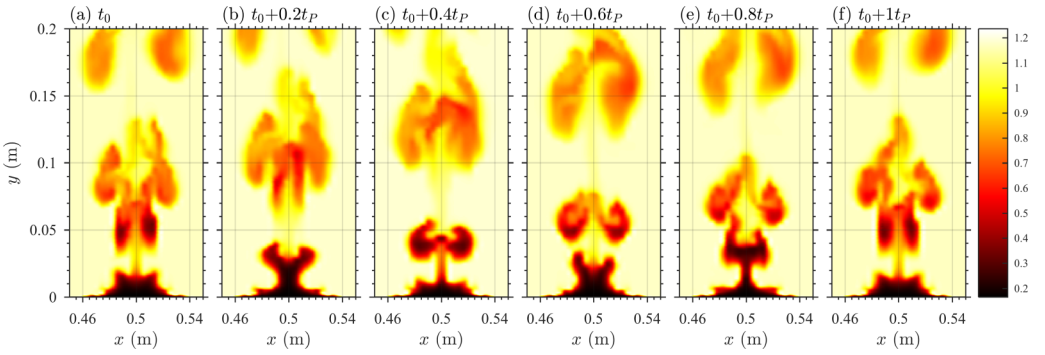

FIG. 7. Representative 2D fields of density (kg/m^3) from LES at six different times (a–f) for the steadily forced plume described in Sec. IV. The fields shown begin at an arbitrary time t_0 late in the LES and are separated by a time interval spanning one complete period, t_P ($t_P = 1/f = 0.22$ s).

By spanning a range of inflow conditions, Cetegen *et al.* [73] determined the empirical relation $\mathrm{St} = 0.55\mathrm{Ri}^{0.45}$ between the Strouhal, $\mathrm{St} = fw/v_i$, and Richardson, $\mathrm{Ri} = (1 - \rho_i/\rho_\infty)gw/v_i^2$, numbers, where ρ_i is inlet density, ρ_∞ is ambient density, g is gravity, w is inlet width, v_i is inlet velocity, and f is the frequency of the natural oscillation.

In this demonstration, ABC is performed assuming that the oscillation frequency above the jet inlet, f, has been measured experimentally and that the density of the inflow mixture, ρ_i/ρ_∞, is unknown (where $\rho_i/\rho_\infty = 0.14$ for pure helium and 1 for pure air). All other properties of the inflow, including v_i and w, are assumed to be known. This case thus serves as the first experimental

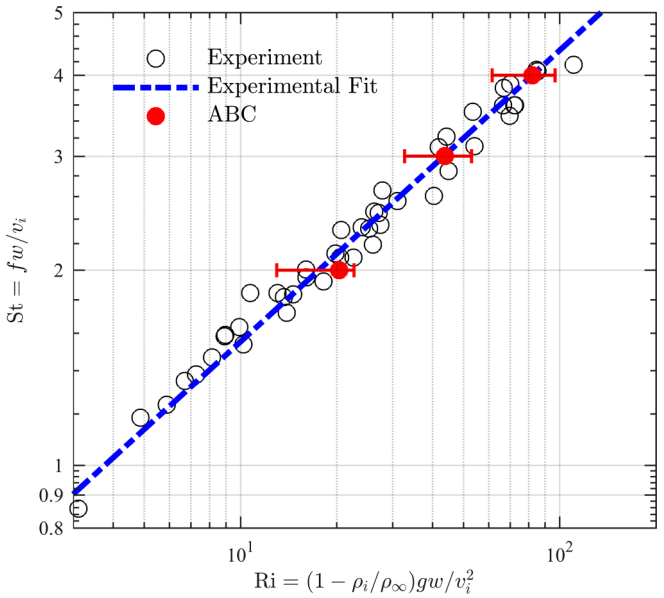

FIG. 8. Relationship between Strouhal number $\mathrm{St} = fw/v_i$ and inlet Richardson number $\mathrm{Ri} = (1 - \rho_i/\rho_\infty)gw/v_i^2$ for the steadily forced helium-air plume. Experimental results from Cetegen *et al.* [73] are indicated by open circles, and the empirical fit $\mathrm{St} = 0.55\mathrm{Ri}^{0.45}$ is shown as a blue dash-dot line. ABC results are indicated by filled red circles located at the mean Ri of the posterior distribution. Uncertainty bars show the minimum and maximum values of Ri in the posteriors.

demonstration that ABC is able to accurately estimate unknown parameters in complex thermal-fluid flows, while also providing measures of uncertainty in the parameter estimates.

The model simulations were once again performed using LES in FireFoam [63]. The plume had a width of $w = 0.07$ m and a steady uniform inflow velocity of $v_i = 0.067$ m/s. The reference Strouhal number was assumed to be $\mathrm{St} = 4$, which is close to several values measured experimentally (see Fig. 8). The inlet width and velocity were held fixed for 100 LES runs while the inlet density ratio varied from 0.14 to 0.6, corresponding to the published range of density ratios from Ref. [73]. The lowest density ratio corresponds to pure helium while the upper value was found in Ref. [73] to be the bounding case beyond which the characteristic natural oscillation was not observed. The Strouhal number was obtained from the peak frequency in a fast Fourier transform of vertical velocity a distance w above the inlet; this was the summary statistic. The distance function consisted of an L1 error norm between reference and model peak frequencies, and ε was chosen to retain 20% of the parameter values from the model simulations. The characteristic oscillation of the plume is captured by the LES, as shown in Fig. 7.

Figure 8 shows that the estimated value of Ri from ABC for $\mathrm{St} = 4$ agrees closely with the experimental data. Additional tests for $\mathrm{St} = 2$ and 3 were also performed using $v_i = 0.1$ m/s and 0.15 m/s, respectively. These different inlet velocities were necessary to span a range of Ri while constraining the physically allowable values of ρ_i/ρ_∞ to between 0.14 and 0.6. Once again, Fig. 8 shows that ABC provides estimates of Ri that are in close agreement with experiments. As a result, ABC correctly identifies the unknown Ri that would provide the desired natural oscillation frequencies for known inlet velocity and width.

V. CONCLUSIONS

Approximate Bayesian computation (ABC) has been used to estimate unknown physical parameters in complex thermal-fluid flows. As a demonstration of the approach, ABC was used to estimate the frequency, amplitude, and mean of the velocity inflow in a periodically forced turbulent buoyant jet, as well as the inlet Richardson number for a steadily forced helium-air plume. In the former case, computational reference data were used to drive the parameter estimation, while in the latter case experimental reference data were used. For both cases, ABC provided accurate estimates of unknown physical parameters even when the model simulations did not exactly match the physics underlying the reference data (e.g., when using LES for the model simulations and reference data from either DNS or experiments), and when the reference data were only indirectly connected to the unknown parameters of interest (e.g., when using measurements distant from the jet and plume inlets, or when using temperature measurements to infer velocity inlet parameters). The primary impacts of using imperfect model simulations or only indirectly connected measurements were observed in the width of the Bayesian confidence intervals. As the physics of the model simulations more closely matched that of the reference data, and as the connection between the measurements and unknown parameters improved, the confidence intervals were found to become narrower.

Although the present demonstration indicates that ABC provides accurate estimates of parameter values with relatively little uncertainty, there are many directions for future research. Different choices for the prior, summary statistic, distance function, and rejection distance can all potentially lead to slightly different parameter estimates. Further work is required to determine the specific effects of each of these choices on the accuracy and uncertainty of ABC. Markov chains can also be used to significantly reduce the cost of ABC by limiting the number of parameter values that are rejected, even for relatively small rejection distances [28]. Such approaches have become more popular in recent years but have yet to be applied in the present context. Finally, the true power of ABC lies in its ability to predict unknown parameters in real-world systems, and future work will more deeply explore the use of experimental reference data to drive parameter estimates, particularly taking into account experimental uncertainty in the context of validation efforts for simulations of complex thermal-fluid problems. An important direction for future research is therefore to apply the ABC method to problems where the true parameter values are unknown.

ACKNOWLEDGMENTS

Helpful discussions with Mark Strobel, Aniruddha Upadhye, Colin Towery, Melvyn Branch, Jeff Anderson, and Daven Henze are gratefully acknowledged. J.D.C., N.T.W., C.L., and T.R.S.H. acknowledge gift support from the 3M Company. J.D.C. and G.B.R. were supported, in part, by NSF Award No.CBET 1454496. P.E.H. was supported, in part, by AFOSR Award Nos.FA9550-14-1-0273 and FA9550-17-1-0144. Computing resources were provided by DoD HPCMP under a Frontier project award and by the Air Force Research Laboratory.

[1] W. L. Oberkampf and C. J. Roy, *Verification and Validation in Scientific Computing* (Cambridge University Press, Cambridge, UK, 2010).

[2] W. L. Oberkampf and T. Trucano, Validation methodology in computational fluid dynamics, AIAA paper AIAA-2000-2549 (2000).

[3] W. L. Oberkampf and T. G. Trucano, Verification and validation in computational fluid dynamics, Progr. Aerospace Sci. **38**, 209 (2002).

[4] W. L. Oberkampf, T. G. Trucano, and C. Hirsch, Verification, validation, and predictive capability in computational engineering and physics, Appl. Mech. Rev. **57**, 345 (2004).

[5] D. C. Estumano, F. C. Hamilton, M. J. Colaço, A. J. K. Leiroz, H. R. B. Orlande, R. N. Carvalho, and G. S. Dulikravich, Bayesian estimate of mass fraction of burned fuel in internal combustion engines using pressure measurements, in *Engineering Optimization IV*, edited by H. Rodrigues, J. Herskovits, C. M. Soares, J. M. Guedes, A. Araujo, J. Folgado, F. Moleiro, and J. A. Madeira (Taylor & Francis Group, London, UK, 2015), pp. 997–1003.

[6] S. Mosbach, A. Braumann, P. L. W. Man, C. A. Kastner, G. P. E. Brownbridge, and M. Kraft, Iterative improvement of Bayesian parameter estimates for an engine model by means of experimental design, Combust. Flame **159**, 1303 (2012).

[7] P. G. Constantine, Q. Wang, A. Doostan, and G. Iaccarino, A surrogate accelerated Bayesian inverse analysis of the HyShot II flight data, AIAA paper AIAA-2011-2037 (2011).

[8] I. Vrbik, R. Deardon, Z. Feng, A. Gardner, and J. Braun, Using individual-level models for infectious disease spread to model spatio-temporal combustion dynamics, Bayesian Anal. **7**, 615 (2012).

[9] S. H. Cheung, T. A. Oliver, E. E. Prudencio, S. Prudhomme, and R. D. Moser, Bayesian uncertainty analysis with applications to turbulence modeling, Reliabil. Eng. Syst. Safety **96**, 1137 (2011).

[10] A. H. Elsheikh, I. Hoteit, and M. F. Wheeler, Efficient Bayesian inference of subsurface flow models using nested sampling and sparse polynomial chaos surrogates, Comput. Methods Appl. Mech. Eng. **269**, 515 (2014).

[11] S. Yeşilyurt and A. Patera, Surrogates for numerical simulations: Optimization of eddy-promoter heat exchangers, Comput. Methods Appl. Mech. Eng. **121**, 231 (1995).

[12] L. Zeng, L. Shi, D. Zhang, and L. Wu, A sparse grid based Bayesian method for contaminant source identification, Adv. Water Resour. **37**, 1 (2012).

[13] A. G. Salinger, R. P. Pawlowski, J. N. Shadid, and B. G. van Bloemen Waanders, Computational analysis and optimization of a chemical vapor deposition reactor with large-scale computing, Ind. Eng. Chem. Res. **43**, 4612 (2004).

[14] W. Jahn, G. Rein, and J. L. Torero, Forecasting fire dynamics using inverse computational fluid dynamics and tangent linearisation, Adv. Eng. Softw. **47**, 114 (2012).

[15] H. Madsen, Parameter estimation in distributed hydrological catchment modeling using automatic calibration with multiple objectives, Adv. Water Resour. **26**, 205 (2003).

[16] S. Wang and X. Xu, Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm, Energy Convers. Manage. **47**, 1927 (2006).

[17] E. Pemha and E. Ngo Nyobe, Genetic algorithm approach and experimental confirmation of a laser-based diagnostic technique for the local thermal turbulence in a hot wind tunnel jet, Prog. Electromagn. Res. **28**, 325 (2011).

[18] H. Kato, A. Yoshizawa, G. Ueno, and S. Obayashi, A data assimilation methodology for reconstructing turbulent flows around aircraft, J. Comput. Phys. **283**, 559 (2015).
[19] X. Gao, Y. Wang, N. Overton, M. Zupanski, and X. Tu, Data-assimilated computational fluid dynamics modeling of convection-diffusion-reaction problems, J. Comput. Sci. **21**, 38 (2018).
[20] J. Sousa, C. García-Sánchez, and C. Gorlé, Improving urban flow predictions through data assimilation, Building Environ. **132**, 282 (2018).
[21] J. R. Stroud, M. Katzfuss, and C. K. Wikle, A Bayesian adaptive ensemble Kalman filter for sequential state and parameter estimation, Mon. Weather Rev. **146**, 373 (2018).
[22] H. Xiao, J.-L. Wu, J.-X. Wang, R. Sun, and C. J. Roy, Quantifying and reducing model-form uncertainties in Reynolds-averaged Navier-Stokes simulations: A data-driven, physics-informed Bayesian approach, J. Comput. Phys. **324**, 115 (2016).
[23] W. N. Edeling, M. Schmelzer, R. P. Dwight, and P. Cinnella, Bayesian predictions of Reynolds-averaged Navier–Stokes uncertainties using maximum a posteriori estimates, AIAA J. **56**, 2018 (2018).
[24] A. P. Singh and K. Duraisamy, Using field inversion to quantify functional errors in turbulence closures, Phys. Fluids **28**, 045110 (2016).
[25] M. Khalil and H. N. Najm, Probabilistic inference of reaction rate parameters from summary statistics, Combust. Theory Model. **22**, 635 (2018).
[26] O. A. Doronina, J. D. Christopher, C. A. Z. Towery, P. E. Hamlington, and W. J. A. Dahm, Autonomic closure for turbulent flows using approximate Bayesian computation, AIAA paper AIAA-2018-0594 (2018).
[27] M. Sunnåker, A. G. Busetto, E. Numminen, J. Corander, M. Foll, and C. Dessimoz, Approximate Bayesian computation, PLoS Comput. Biol. **9**, e1002803 (2013).
[28] P. Marjoram, J. Molitor, V. Plagnol, and S. Tavaré, Markov chain Monte Carlo without likelihoods, Proc. Natl. Acad. Sci. USA **100**, 15324 (2003).
[29] J. M. Marin, P. Pudlo, C. P. Robert, and R. J. Ryder, Approximate Bayesian computational methods, Stat. Comput. **22**, 1167 (2012).
[30] F. Hartig, J. M. Calabrese, B. Reineking, T. Wiegand, and A. Huth, Statistical inference for stochastic simulation models—Theory and application, Ecol. Lett. **14**, 816 (2011).
[31] A. B. Abdessalem, N. Dervilis, D. Wagg, and K. Worden, Model selection and parameter estimation in structural dynamics using approximate Bayesian computation, Mech. Syst. Signal Proc. **99**, 306 (2018).
[32] J. D. Christopher, N. T. Wimer, T. R. S. Hayden, C. Lapointe, I. Grooms, G. B. Rieker, and P. E. Hamlington, Parameter estimation for a turbulent buoyant jet using approximate Bayesian computation, AIAA paper AIAA-2017-0531 (2017).
[33] J. D. Christopher, C. Lapointe, N. T. Wimer, T. R. S. Hayden, I. Grooms, G. B. Rieker, and P. E. Hamlington, Parameter estimation for a turbulent buoyant jet with rotating cylinder using approximate Bayesian computation, AIAA paper AIAA-2017-3629 (2017).
[34] D. B. Rubin, Bayesianly justifiable and relevant frequency calculations for the applied statistician, Ann. Stat. **12**, 1151 (1984).
[35] J. K. Pritchard, M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman, Population growth of human Y chromosomes: A study of Y chromosome microsatellites, Mol. Biol. Evol. **16**, 1791 (1999).
[36] S. Tavaré, D. J. Balding, R. C. Griffiths, and P. Donnelly, Inferring coalescence times from DNA sequence data, Genetics **145**, 505 (1997).
[37] M. A. Beaumont, W. Zhang, and D. J. Balding, Approximate Bayesian computation in population genetics, Genetics **162**, 2025 (2002).
[38] T. Toni, D. Welch, N. Strelkowa, A. Ipsen, and M. P. H. Stumpf, Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems, J. R. Soc., Interface **6**, 187 (2009).
[39] K. Csilléry, M. G. B. Blum, O. E. Gaggiotti, and O. François, Approximate Bayesian computation (ABC) in practice, Trends Ecol. Evol. **25**, 410 (2010).

[40] M. A. Beaumont, Approximate Bayesian computation in evolution and ecology, Annu. Rev. Ecol. Evol. System. **41**, 379 (2010).
[41] N. J. R. Fagundes, N. Ray, M. A. Beaumont, S. Neuenschwander, F. M. Salzano, S. L. Bonatto, and L. Excoffier, Statistical evaluation of alternative models of human evolution, Proc. Natl. Acad. Sci. USA **104**, 17614 (2007).
[42] J. A. Vrugt and M. Sadegh, Toward diagnostic model calibration and evaluation: Approximate Bayesian computation, Water Resour. Res. **49**, 4335 (2013).
[43] D. J. Nott, L. Marshall, and J. Brown, Generalized likelihood uncertainty estimation (GLUE) and approximate Bayesian computation: What's the connection?, Water Resour. Res. **48**, W12602 (2012).
[44] M. Sadegh and J. A. Vrugt, Bridging the gap between GLUE and formal statistical approaches: Approximate Bayesian computation, Hydrol. Earth Syst. Sci. **17**, 4831 (2013).
[45] B. Olson, Stochastic weather generation with approximate Bayesian computation, Master's thesis, University of Colorado at Boulder, 2016.
[46] B. Olson and W. Kleiber, Approximate Bayesian computation methods for daily spatiotemporal precipitation occurrence simulation, Water Resour. Res. **53**, 3352 (2017).
[47] D. D. Lucas, A. Gowardhan, P. Cameron-Smith, and R. L. Baskett, Impact of meteorological inflow uncertainty on tracer transport and source estimation in urban atmospheres, Atmos. Environ. **143**, 120 (2016).
[48] J. Rohmer, M. Rousseau, A. Lemoine, R. Pedreros, J. Lambert, and A. Benki, Source characterisation by mixing long-running tsunami wave numerical simulations and historical observations within a metamodel-aided ABC setting, Stochastic Environmental Research and Risk Assessment **32**, 967 (2018).
[49] R. C. Smith, *Uncertainty Quantification: Theory, Implementation, and Applications* (SIAM, Philadelphia, PA, 2014).
[50] L. Le Cam and G. L. Yang, *Asymptotics in Statistics: Some Basic Concepts* (Springer Science & Business Media, CITY, 2012).
[51] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, Cambridge, UK, 2003).
[52] S. C. Crow and F. H. Champagne, Orderly structure in jet turbulence, J. Fluid Mech. **48**, 547 (1971).
[53] J. A. Lovett and S. R. Turns, Experiments on axisymmetrically pulsed turbulent jet flames, AIAA J. **28**, 38 (1990).
[54] R. B. Farrington and S. D. Claunch, Infrared imaging of large-amplitude, low-frequency disturbances on a planar jet, AIAA J. **32**, 317 (1994).
[55] S. Marzouk, H. Mhiri, S. El Golli, G. Le Palec, and P. Bournot, Numerical study of momentum and heat transfer in a pulsed plane laminar jet, Int. J. Heat Mass Transf. **46**, 4319 (2003).
[56] W. Kriaa, H. B. Cheikh, H. Mhiri, G. Le Palec, and P. Bournot, Numerical study of free pulsed jet flow with variable density, Energy Convers. Manage. **49**, 1141 (2008).
[57] C. P. Arnold Jr. and C. H. Dey, Observing-systems simulation experiments: Past, present, and future, Bull. Am. Meteorol. Soc. **67**, 687 (1986).
[58] M. Xue, M. Tong, and K. K. Droegenmeier, An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting, J. Atmos. Ocean. Technol. **23**, 46 (2006).
[59] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, *Applied Mathematical Series*, Vol. 160 (Springer, Berlin, 2004).
[60] J. Kaipio and E. Somersalo, Statistical inverse problems: Discretization, model reduction and inverse crimes, J. Comput. Appl. Math. **198**, 493 (2007).
[61] A. Stoffelen, G. J. Marseille, F. Bouttier, D. Vasiljevic, S. De Haan, and C. Cardinali, ADM-Aeolus Doppler wind Lidar observing system simulation experiment, Q. J. R. Meteorol. Soc. **132**, 1927 (2006).
[62] W. Lahoz, B. Khattatov, and R. Menard, editors, *Data Assimilation: Making Sense of Observations* (Springer, Berlin, 2010).
[63] Y. Wang, P. Chatterjee, and J. L. de Ris, Large eddy simulation of fire plumes, Proc. Combust. Inst. **33**, 2473 (2011).
[64] C. Greenshields, OpenFOAM, OpenFOAM Foundation, March (2016).

[65] H. G. Weller, G. Tabor, H. Jasak, and C. Fureby, A tensorial approach to computational continuum mechanics using object-oriented techniques, Comput. Phys. **12**, 620 (1998).
[66] A. Yoshizawa, Statistical theory for compressible turbulent shear flows, with the application to subgrid modeling, Phys. Fluids **29**, 2152 (1986).
[67] W. Sutherland, LXXV. A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin, London Edinburgh Dublin Philos. Mag. J. Sci. **9**, 781 (1905).
[68] D. R. Stull and H. Prophet, JANAF thermochemical tables, Technical report, DTIC document (1971).
[69] D. J. Thomson, Spectrum estimation and harmonic analysis, Proc. IEEE **70**, 1055 (1982).
[70] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications* (Cambridge University Press, Cambridge, UK, 1993).
[71] W. Van Drongelen, *Signal Processing for Neuroscientists: An Introduction to the Analysis of Physiological Signals* (Academic Press, London, UK, 2006).
[72] D. W. Scott, On optimal and data-based histograms, Biometrika **66**, 605 (1979).
[73] B. M. Cetegen, Y. Dong, and M. C. Soteriou, Experiments on stability and oscillatory behavior of planar buoyant plumes, Phys. Fluids **10**, 1658 (1998).