AIAA SciTech Forum
11–15 & 19–21 January 2021, VIRTUAL EVENT
AIAA Scitech 2021 Forum
10.2514/6.2021-1477

A Comparison of Reification and Cokriging for Sequential Multi-Information Source Fusion

Danial Khatamsaz * and Douglas L. Allaire.†
J. Mike Walker ’66 Department of Mechanical Engineering, Texas A&M University, College Station, TX 77843

Many engineering tasks, such as optimization, analysis, model development, model calibration, and others, can potentially exploit information from many sources. These sources include numerical models, expert opinion, and experimental data. Information fusion over these sources of information has the potential to provide a more complete quantitative picture of the current state of knowledge of a given ground truth quantity of interest. This state of knowledge can be updated as new information from any given source is acquired. In this work, we compare two information fusion approaches that both seek to combine all available information to form a surrogate model of the ground truth. These are model reification and cokriging. The comparison considers several test functions as well as a real world NACA 0012 analysis. A quantity of interest is considered for each test case, as well as the derivative of the quantity of interest in some cases. Each fusion approach performs well generally, with each being superior to the other under certain conditions.

I. Introduction

In most engineering applications there are available many different models for the task at hand. These can include computational models of varying fidelity, historical experimental data, new experimental data to be acquired, and expert opinion. These different models, comprising different physical or mathematical assumptions and having potentially widely varying execution costs in terms of certain essential resources, are referred to collectively as information sources. Since every information source may provide valuable information to the task at hand, it is useful and at times critical, that all information be gathered from these sources efficiently and fused dynamically to ensure that all information is used throughout the given task. Many information fusion approaches have been proposed, such as Bayesian modeling averaging [1–6], the use of adjustment factors [7–10], covariance intersection methods [11, 12], and fusion under known correlation [13–15] and cokriging [16–18]. Each of these have their pros and cons depending on whether the task requires a physics-based fused model, a conservative approach to fusion, or knowledge about information source correlation. In this work, we approach the information fusion process from the perspective of dynamically constructing a surrogate model that encompasses all available information at a given point in the engineering task.

Surrogate modeling is often a critical component in the execution of many engineering tasks. These tasks include analysis and optimization of systems, model development, parameter estimation, and many others. The appeal of surrogate models lies in the ability to tradeoff potentially prohibitive computational expensive for diminished accuracy. Understanding this tradeoff as the task unfolds allows for the dynamic selection of which source to query and where to query to ensure that the fused surrogate model comprising all available information is as effective as possible. In this work, we consider a Bayesian approach to surrogate modeling using information gathering policies based on information gain. For this, Gaussian processes are used as surrogate models for each information source. Discrepancy is quantified between each information source and available ground truth information, which often is only given by expert opinion or an assumed highest fidelity model (that is often computational). The fusion takes place over these Gaussian process surrogates and leads to another Gaussian process model, which we refer to as the fused model. This fused model comprises our best quantification of the current available information. By generating synthetic data from prior predictive Gaussian processes of the information sources, we temporarily update the fused model. This begets a new *potential* posterior distribution of our data that could follow an actual query to a true underlying information source (and not the source’s Gaussian process). These potential posterior distributions are compared to each other to measure information gain using the Kullback–Leibler divergence [19, 20]. The information source and query location that provide the largest expected information gain are selected to perform the next experiment (physical or computer). The fused model is then updated. This process can be carried out by the well-known cokriging method or by our newly

*Ph.D. Candidate, danialkh26@tamu.edu, AIAA Student Member.
†Associate Professor, dallaire@tamu.edu, AIAA Senior Member.

developed sequential model reification procedure presented in Ref. [21], which was originally developed to handle a lack of ground truth information.

The goal of this work is to compare the sequential fusion techniques of model reification [21–23] and cokriging [16–18]. The comparison will study the advantages and disadvantages of each technique under different conditions, such as noisy sampling, availability of online sampling from the highest fidelity model, or presence of prior information of highest fidelity model. First, we apply these fusion techniques on several test conditions using analytical test functions and then consider each approach on an aerodynamic design problem.

The rest of the paper is organized as follows. Sec. II presents the different ingredients of each approach. In Sec. III, each approach is demonstrated on test functions in different design conditions and then on an aerodynamic design problem. Finally, conclusions are drawn in Sec. IV.

II. Approach

A. Gaussian Process Regression Surrogates

To estimate a quantity of interest, there may be several information sources with different mathematical formulations, costs, and varying fidelity. Making decision in different steps of the process needs predictions of the response surface of these information sources. Following Refs. [24, 25], we assume there are available some information sources, $f_i(\mathbf{x})$, where $i \in \{1, 2, \ldots, S\}$, used to estimate a quantity of interest, $f(\mathbf{x})$, at a design point $\mathbf{x}$. By constructing a surrogate model, we are able to predict the output of information sources at design points that have not been executed yet. The surrogate models are built for each information source using Gaussian process regression [26]. A Gaussian process is a powerful mathematical modeling tool due to its easy manipulation and implementation. This is important, particularly in Bayesian analysis and optimization frameworks, since the models need to be updated as new information is added to the system. A Gaussian process is a probabilistic model that provides a normal distribution as the prior and posterior prediction of the function value corresponding to a query location. Gaussian processes are constructed with kernels that make certain assumptions regarding the correlation between data points. Depending on the nature of the function being modeled by a Gaussian process, different kernel functions can be utilized.

Using N_i previously evaluations of information source i described by $\{\mathbf{X}_{N_i}, \mathbf{y}_{N_i}\}$, where $\mathbf{X}_{N_i} = (\mathbf{x}_{1,i}, \ldots, \mathbf{x}_{N_i,i})$ represents the N_i input samples to information source i and $\mathbf{y}_{N_i} = \left(f_i(\mathbf{x}_{1,i}), \ldots, f_i(\mathbf{x}_{N_i,i})\right)$ are the corresponding outputs from information source i, the posterior distribution of information source i at a design point $\mathbf{x}$ is represented as

$$f_{\mathrm{GP},i}(\mathbf{x}) \mid \mathbf{X}_{N_i}, \mathbf{y}_{N_i} \sim \mathcal{N}\left(\mu_i(\mathbf{x}), \sigma^2_{\mathrm{GP},i}(\mathbf{x})\right) \tag{1}$$

where

$$\mu_i(\mathbf{x}) = K_i(\mathbf{X}_{N_i}, \mathbf{x})^T\,[K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i}) + \sigma^2_{n,i} I]^{-1}\mathbf{y}_{N_i} \tag{2}$$

$$\sigma^2_{\mathrm{GP},i}(\mathbf{x}) = k_i(\mathbf{x}, \mathbf{x}) - K_i(\mathbf{X}_{N_i}, \mathbf{x})^T\,[K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i}) + \sigma^2_{n,i} I]^{-1} K_i(\mathbf{X}_{N_i}, \mathbf{x}) \tag{3}$$

and $K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i})$ is the $N_i \times N_i$ matrix whose (m, n) entry is $k_i(\mathbf{x}_{m,i}, \mathbf{x}_{n,i})$, and $K_i(\mathbf{X}_{N_i}, \mathbf{x})$ is the $N_i \times 1$ vector whose m^{th} entry is $k_i(\mathbf{x}_{m,i}, \mathbf{x})$ for information source i, where $k_i(\mathbf{x}, \mathbf{x}')$ is a real-valued kernel function over the input space. For the kernel function, a common choice is the squared exponential covariance function

$$k_i(\mathbf{x}, \mathbf{x}') = \sigma_s^2 \exp\left(-\sum_{h=1}^{d} \frac{(x_h - x'_h)^2}{2 l_h^2}\right) \tag{4}$$

where d is the dimensionality of the input space, σ_s^2 indicates signal variance, and l_h, where $h = 1, 2, \ldots, d$, is the characteristic length-scale that is used to define the correlation between points within dimension h. By putting the term $\sigma^2_{n,i}$, we can model observation error for information sources based on experiments as well.

To determine the uncertainty for each information source with respect to the ground truth quantity of interest, we add a term associated with the fidelity of the information source at a given location in the input space, $\sigma^2_{f,i}(\mathbf{x})$. Then, the total variance is calculated including the variance associated with Gaussian process and the variance associated with the information source fidelity over the input space.

$$\sigma_i^2(\mathbf{x}) = \sigma^2_{\mathrm{GP},i}(\mathbf{x}) + \sigma^2_{f,i}(\mathbf{x}), \tag{5}$$

and the term $\sigma^2_{f,i}(\mathbf{x})$ can be estimated from, for example, expert opinion or available real-world data. Here, our method to find this variance is to calculate the square of the difference between available data from the ground truth quantity of interest and an information source. A Gaussian process is then performed using the values as training points to estimate the fidelity variance over the input space as mean of the Gaussian process. More details are available in Ref. [27].

B. Reification-based Fusion of Multiple Information Sources

Every employed information source contains potentially useful information about the quantity of interest that may not be accessible from other information sources [22, 23, 25, 27, 28]. The goal is to fuse available new information from every information source query to make sure we are using the resources wisely. As noted previously, several approaches exist for fusing multiple sources of information such as Bayesian modeling averaging [1–6], the use of adjustment factors [7–10], covariance intersection methods [11, 12], and fusion under known correlation [13–15]. By assuming that every information source provides useful information regarding the ground truth quantity of interest, when incorporating more information sources into a fusion process , the expectation is to have the variance of the quantity of interest estimates decreased. This is not guaranteed for the mentioned fusion techniques with the exception of fusion under known correlation. Thus, it is important to determine the correlations prior to fusion.

Following the Refs. [22, 23, 25, 27, 28], by representing information sources using intermediate Gaussian processes, their fusion suggests the normally distributed information. If correlations between the discrepancies of information sources are known, the fused mean and variance are found using the following equations [15]

$$\mathbb{E}[\hat{f}(\mathbf{x})] = \frac{\mathbf{e}^{\mathrm{T}}\tilde{\Sigma}(\mathbf{x})^{-1}\boldsymbol{\mu}(\mathbf{x})}{\mathbf{e}^{\mathrm{T}}\tilde{\Sigma}(\mathbf{x})^{-1}\mathbf{e}} \tag{6}$$

$$\mathrm{Var}\left(\hat{f}(\mathbf{x})\right) = \frac{1}{\mathbf{e}^{\mathrm{T}}\tilde{\Sigma}(\mathbf{x})^{-1}\mathbf{e}} \tag{7}$$

where $\mathbf{e} = [1,\dots,1]^T$, $\boldsymbol{\mu}(\mathbf{x}) = [\mu_1(\mathbf{x}),\dots,\mu_S(\mathbf{x})]^{\mathrm{T}}$ given S models, and $\tilde{\Sigma}(\mathbf{x})^{-1}$ is the inverse of the covariance matrix between the information sources. Note that there is no assumption of hierarchy of the information sources in contrast with some traditional approaches [29–36].

Before using equations (6) and (7), we need to estimate the correlation coefficients between information sources over the input domain. This task is described in [22, 23] as the reification process. To estimate the correlation coefficients between the deviations of information sources i and j, each of the information sources i and j, one at a time, is reified meaning that it is taken as the ground truth. If information source i is reified, the correlation coefficients between the information sources i and j, for $j = 1,\dots,i-1,i+1,\dots,S$, are given as

$$\rho_{ij}(\mathbf{x}) = \frac{\sigma_i^2(\mathbf{x})}{\sigma_i(\mathbf{x})\sigma_j(\mathbf{x})} = \frac{\sigma_i(\mathbf{x})}{\sqrt{\left(\mu_i(\mathbf{x}) - \mu_j(\mathbf{x})\right)^2 + \sigma_i^2(\mathbf{x})}} \tag{8}$$

where $\mu_i(\mathbf{x})$ and $\mu_j(\mathbf{x})$ are the mean values of the Gaussian processes of information sources i and j accordingly, at $\mathbf{x}$, and $\sigma_i^2(\mathbf{x})$ and $\sigma_j^2(\mathbf{x})$ are the total variances at $\mathbf{x}$. Then, information source j is reified to estimate $\rho_{ji}(\mathbf{x})$. To estimate the correlation between the errors, the variance weighted average of the two estimated correlation coefficients is used

$$\bar{\rho}_{ij}(\mathbf{x}) = \frac{\sigma_j^2(\mathbf{x})}{\sigma_i^2(\mathbf{x}) + \sigma_j^2(\mathbf{x})}\rho_{ij}(\mathbf{x}) + \frac{\sigma_i^2(\mathbf{x})}{\sigma_i^2(\mathbf{x}) + \sigma_j^2(\mathbf{x})}\rho_{ji}(\mathbf{x}) \tag{9}$$

Finally, the covariance matrix is computed using the average correlations, Then the fused mean and variance in equations (6) and (7) are estimated. Details on how to build the covariance matrix is discussed in Ref. [37].

C. Cokriging

In design applications with different computational models available to estimate the same quantity of interest, cokriging is considered as an approach to employ all models predicting the response of the highest fidelity model at a new design point. Most often, when a function is more expensive to evaluate, there are fewer samples available to provide information about the function. By assuming this, a cokriging system is designed to estimate the output of the highest fidelity model by combination of cheaper functions with more samples available. Although the traditional cokriging method uses one primary and one secondary function to estimate the output of the primary function, there have been several works to extend the idea to use any number of secondary functions or information sources available to do the estimation. We show how to build the cokriging system using multiple information sources based on the proposed method in Ref. [38].

To construct the cokriging system equations, we need a number of samples from each model. In general, the number of available samples for a model decrease as the fidelity of the model increases due to the cost of evaluations. We take samples as $(\mathbf{X}_{\mathrm{e}}^{\mathrm{T}}, \mathbf{Y}_{\mathrm{e}}^{\mathrm{T}}) = [(x_{\mathrm{e}}^1, y_{\mathrm{e}}^1), ..., (x_{\mathrm{e}}^{\mathrm{n}}, y_{\mathrm{e}}^{\mathrm{n}})]^{\mathrm{T}}$ for highest fidelity model and $(\mathbf{X}_k^{\mathrm{T}}, \mathbf{Y}_k^{\mathrm{T}}) = [(x_k^1, y_k^1), ..., (x_k^{n_k}, y_k^{n_k})]^{\mathrm{T}}$ with $k = 1, ..., N$ for other lower fidelity models where we assume there are n samples for highest fidelity model and n_k samples for lower fidelity model k. Following Refs. [16, 18], the multifidelity cokriging system is constructed as

$$\mathbf{y}_e(x) = \sum_{k=1}^{N} \rho_k \mathbf{y}_k(x) + \mathbf{y}_d(x), \tag{10}$$

where $\mathbf{y}_e(x)$ is the estimate of the highest fidelity model, ρ_k is the scaling factor for model k, and $\mathbf{y}_d(x)$ indicating a stationary random process independent of $\mathbf{y}_k(x)$ with variance σ_d^2 at location x. Another assumption in terms of stochastic processes is independence of different lower fidelity models in estimating $y_{\mathrm{e}}(x)$, leading to

$$\mathrm{cov}(\mathbf{y}_i(x), \mathbf{y}_j(x)) = 0,\ i \neq j,\ i, j = 1, ..., N. \tag{11}$$

The covariance matrix in cokriging is defined as

$$\mathrm{C} = \begin{pmatrix} \sigma_1^2 \mathrm{K}_1(\mathbf{X}_1, \mathbf{X}_1) & 0 & \dots & \rho_1 \sigma_1^2 \mathrm{K}_1(\mathbf{X}_1, \mathbf{X}_e) \\ 0 & \sigma_2^2 \mathrm{K}_2(\mathbf{X}_2, \mathbf{X}_2) & \dots & \rho_2 \sigma_2^2 \mathrm{K}_2(\mathbf{X}_2, \mathbf{X}_e) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_1 \sigma_1^2 \mathrm{K}_1(\mathbf{X}_e, \mathbf{X}_1) & \rho_2 \sigma_2^2 \mathrm{K}_2(\mathbf{X}_e, \mathbf{X}_2) & \dots & \sum_{k=1}^{N} \rho_k^2 \sigma_k^2 \mathrm{K}_k(\mathbf{X}_e, \mathbf{X}_e) + \sigma_d^2 \mathrm{K}_{\mathrm{d}}(\mathbf{X}_e, \mathbf{X}_e) \end{pmatrix}, \tag{12}$$

where $\mathrm{K}_1, \dots, \mathrm{K}_k, \mathrm{K}_{\mathrm{d}}$ are the covariance matrices indicating correlations between set of data points using hyper parameters according to the model. We use Eq. (4) to define the covariance matrices, K_i. Note that in Eq. (4), σ_s^2 is included in the function while here we are showing the term as a coefficient multiplied to a covariance matrix. In the multifidelity cokriging, we have to estimate $\Theta_k, \rho_k, k = 1, \dots, N$ and Θ_{d} and ρ_{d} as the hyper parameters. To define hyper parameters, we need to maximize the log likelihood of $\mathbf{y}_k$ given as

$$-\frac{n_k}{2} \ln(\hat{\sigma}_k^2) - \frac{1}{2} \ln|\det(\mathrm{K}_k(\mathbf{X}_k, \mathbf{X}_k)|, \tag{13}$$

where

$$\hat{\sigma}_k^2 = \frac{(\mathbf{y}_k - \mathbf{1}\hat{\mu}_k)^{\mathrm{T}} \mathrm{K}_k(\mathbf{X}_k, \mathbf{X}_k)^{-1} (\mathbf{y}_k - \mathbf{1}\hat{\mu}_k)}{n_k}, \tag{14}$$

and

$$\hat{\mu}_k = \frac{\mathbf{1}^{\mathrm{T}} \mathrm{K}_k(\mathbf{X}_k, \mathbf{X}_k)^{-1} \mathbf{y}_k}{\mathbf{1}^{\mathrm{T}} \mathrm{K}_k(\mathbf{X}_k, \mathbf{X}_k)^{-1} \mathbf{1}}, \tag{15}$$

with $\mathbf{1}$ as a column vector of ones. The next step is to find scaling factors and Θ_{d}. To do so, first we build the following relation

$$\mathbf{d} = \mathbf{y}_e - \sum_{k=1}^{N} \rho_k \mathbf{y}_k(\mathbf{X}_k). \tag{16}$$

The same process is followed by maximizing the Eq. (13) for $\hat{\sigma}_k^2$, $\mathbf{X}_e$, K_{d}, and replacing $\mathbf{y}_k$ by $\mathbf{d}$ in Eqs. (14) and (15). Finally

$$\hat{y}_e(x) = \hat{\mu} + \mathrm{c}^{\mathrm{T}} \mathrm{C}^{-1} (\mathbf{y} - \mathbf{1}\hat{\mu}), \tag{17}$$

where $\mathbf{y} = \{\mathbf{y}_1, \dots, \mathbf{y}_k, \mathbf{y}_e\}$ and

$$\hat{\mu} = \frac{\mathbf{1}^{\mathrm{T}} \mathrm{C}^{-1} \mathbf{y}}{\mathbf{1}^{\mathrm{T}} \mathrm{C}^{-1} \mathbf{1}}, \tag{18}$$

$$c = \begin{pmatrix} \hat{\rho}_1 \hat{\sigma}_1^2 \mathrm{K}_1(\mathbf{X}_1, x) \\ \hat{\rho}_2 \hat{\sigma}_2^2 \mathrm{K}_2(\mathbf{X}_2, x) \\ \vdots \\ \sum_{k=1}^{N} \rho_k^2 \sigma_k^2 \mathrm{K}_k(\mathbf{X}_e, x) + \sigma_d^2 \mathrm{K}_\mathrm{d}(\mathbf{X}_e, x) \end{pmatrix}. \tag{19}$$

Note the similarity of Eq. (17) and Eq. (2). Using Eq. (17), we estimate the response of the highest fidelity model for locations in the input space not yet executed based on the samples available from other lower fidelity models. The estimated mean squared error of the predictor using standard stochastic process then is calculated as [16, 18]

$$\hat{S}^2(x) = \sum_{k=1}^{N} \rho_k^2 \sigma_k^2 + \sigma_d^2 - \mathrm{c}^\mathrm{T} \mathrm{C}^{-1} \mathrm{c} + 1 - \frac{\mathbf{1}^\mathrm{T} \mathrm{C}^{-1} \mathbf{c}}{\mathbf{1}^\mathrm{T} \mathrm{C}^{-1} \mathbf{1}}. \tag{20}$$

D. Kullback–Leibler divergence

The next step is to use a systematic approach to determine the next point and information source to query. The process begins with generating Latin hypercube samples in the design space as alternative points, $\mathbf{X}_\mathrm{alt}$. We need to find an alternative point and an information source to query that provides us with the most information about the ground truth considering the cost of query. Among different policies developed to measure the information gain term, we use the Kullback-Leibler divergence method [19–21]. The term $D_\mathrm{KL}(P||Q)$, which denotes the Kullback-Leibler (KL) divergence, is a criterion to evaluate the difference between two probability distributions P and Q with densities $p(x)$ and $q(x)$ given as

$$D_\mathrm{KL}(P||Q) = \int_{-\infty}^{+\infty} p(x) \log \frac{p(x)}{q(x)}\, dx. \tag{21}$$

Now, to find the an alternative point and an information source to query, we generate N_alt alternative Latin hypercube samples in the input space and calculate the mean and variance using Eqs. (2) and (3). Note that we are constructing temporary Gaussian processes for the information sources in both the reification fusion and cokriging approaches to be able to predict the response surface of the information sources. Then, for each alternative point and each information source, we draw N_q samples by leveraging the fact that the function value of information source i at design point $\mathbf{x}$ is normally distributed with mean $\mu_i(\mathbf{x})$ and variance $\sigma^2_{\mathrm{GP},i}(\mathbf{x})$, according to the Eq. (1). For example, at alternative point $\mathbf{x}$ and information source i we sample such that

$$f_i^q(\mathbf{x}) \sim \mathcal{N}\left(\mu_i(\mathbf{x}), \sigma^2_{\mathrm{GP},i}(\mathbf{x})\right), \quad \text{for } q = 1, \ \ldots, N_q, \tag{22}$$

and the alternative point $\mathbf{x}$ and the objective value $f_i^q(\mathbf{x})$ are temporarily augmented to the available samples of the corresponding information source one at a time. The current Gaussian process of the given source is then updated temporarily with this new data point. Note that our predictive model can be constructed using reification fusion or the cokriging approach. Denoting the mean and variance of this model at N_alt previously generated samples before update as $\mu_{j,\mathrm{curr}}$ and $\sigma^2_{j,\mathrm{curr}}$ and the mean and variance after temporarily updating the information sources as $\mu_{j,\mathrm{temp}}$ and $\sigma^2_{j,\mathrm{temp}}$, the Kullback-Leibler divergence between current and temporary estimates is computed as

$$D_\mathrm{KL}^{i,q}(\mathrm{curr}||\mathrm{temp}) = \frac{1}{2N_p} \sum_{j=1}^{N_p} \left[\frac{\sigma^2_{j,\mathrm{curr}}}{\sigma^2_{j,\mathrm{temp}}} + \frac{(\mu_{j,\mathrm{temp}} - \mu_{j,\mathrm{curr}})^2}{\sigma^2_{j,\mathrm{temp}}} - d + \ln \frac{\sigma^2_{j,\mathrm{temp}}}{\sigma^2_{j,\mathrm{curr}}} \right], \tag{23}$$

where superscript i, q indicates augmentation sample q to the available data of information source i and d is dimension of the design space. Repeating the process for all N_q samples drawn for the alternative point $\mathbf{x}$ and information source i, the expected Kullback-Leibler divergence at alternative $\mathbf{x}$ and information source i is calculated as

$$\mathbb{E}(D_\mathrm{KL}^{i,\mathbf{x}}) = \frac{1}{N_q} \sum_{q=1}^{N_q} D_\mathrm{KL}^{i,q}(\mathrm{curr}||\mathrm{temp}). \tag{24}$$

After computing the expected Kullback-Leibler divergence for all alternatives and information sources, denoting that the cost of querying information source i at alternative point $\mathbf{x}$ is defined as $C_{i,\mathbf{x}}$, the next design point and information source to query is selected as

$$(i_{N+1}, \mathbf{x}_{N+1}) = \arg\max_{\mathbf{x}\in\mathcal{X}} \frac{\mathbb{E}(D_{\mathrm{KL}}^{i,\mathbf{x}})}{C_{i,\mathbf{x}}}. \tag{25}$$

After querying the selected alternative and information source, it will be added to the available data of that information source and the corresponding Gaussian process is updated. This process continues until a termination criterion is met.

III. Demonstration

In this section, we use and compare the reification and cokriging techniques to perform fusion of multiple information sources on several test conditions and a real-world application. The aim of each technique is to build accurate fused models for predicting the response of the highest fidelity model (ground truth or truth model). Since the fusion process can be influenced by samples from the truth model, a variety of conditions are studied to investigate the characteristics of each approach in different settings. Accessibility to the truth model for online sampling, availability of prior truth samples, and a noisy truth model are settings considered. Samples from the truth models help to quantify model discrepancy for each source and to make corrections/updates as new data arrive. To compare the performance of each fusion approach in different conditions, the mean squared error is calculated and averaged over 200 replications of simulations for each fusion approach. In what follows, we assume the cost of executing a medium fidelity model is 100 units and the low fidelity model is 10 units. We limit the budget to 4000 units. These values are chosen to ensure we can visualize trends prior to exhausting the total budget.

A. Test problems

In Fig. 1, the response of two information sources varying in fidelity to estimate a truth model are illustrated. The shaded area around the truth function shows the lower and upper bounds of the truth model's response when the noise is added to the system to study how it influences the fused model estimation accuracy and capability of a fusion technique to handle the noise. In presence of noise, the truth response is a uniform distribution in the bounds for a particular design point.

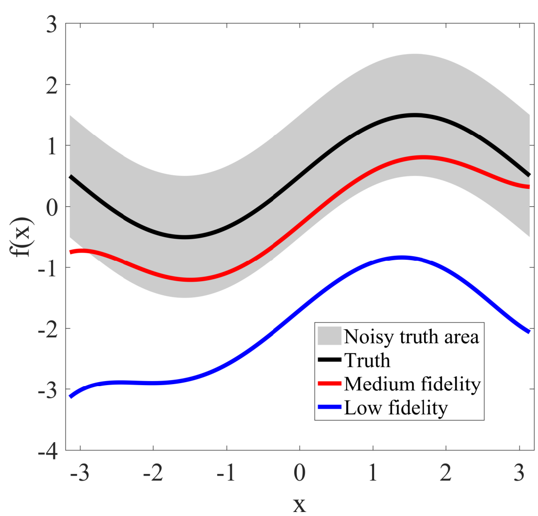

Fig. 1 Two information sources and the truth model. When sampling the truth model in noisy condition, the truth objective value can be any value inside the shaded area for a particular design point.

Absence of any information about the truth model, for example, unavailability of prior evaluations or online sampling due to the cost of evaluations or lack of a function or experiment to evaluate the ground truth objective value for a specific design, is a common condition in engineering design. For example, to design a new airplane, initially there is no physical model built yet to study the drag and lift coefficients during different flight conditions to provide real data to help the design process. In this case, the designer has to use lower fidelity information sources to approximate the real data values based on his opinion about the performance of cheaper information sources to make inferences about the truth model.

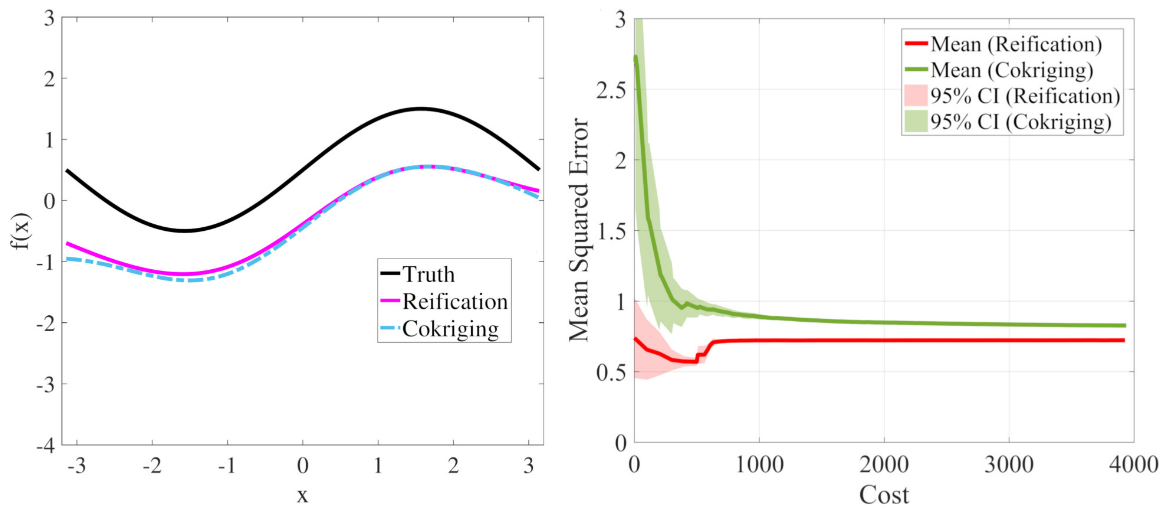

Fig. 2 Fused models obtained by reification and cokriging techniques when there is no information available about the truth model. Reification is showing quietly better approximations of the truth model as it suggests lower mean squared error.

In Fig. 2, the final estimates of the truth model's response using reification and cokriging fusion techniques are shown in the case where there is not any information about the truth model available and the model discrepancies are set by expert opinion. The point here is with reification, the fused model predicts more accurately even in the initial iterations when there is only a few samples available from the information sources while cokriging needs more samples to increase the estimation accuracy.

Another condition to consider is when we have access to some prior evaluations of the truth model, however, online sampling for new design points is not possible. Using the prior samples, we can calculate each model discrepancy at some locations in the objective space with respect to the truth model to make the approximation more accurate around those regions but we should also consider the situation in which the available data of the truth model might be noisy. In fact, for a real-world design problem, the presence of noise in the data is inevitable and should be taken into account. While reification based method performs quite consistently in fusion of information even in noisy conditions, the fused model constructed by cokriging shows high dependency on samples from the truth model. When the data is corrupted by noise, cokriging is unable to correct the estimations using the information from lower fidelity models.

In some simulations and experiments, the truth model is defined and is accessible or the conditions to run an expensive experiment is feasible but we are not given enough resources to evaluate many design points. Therefore, we run the expensive evaluations only to update the discrepancy associated with each information source in highly uncertain regions in the objective space to decrease the fused model's estimation error as much as possible. However, there are cases in which the sampling is done in presence of some level of noise and the fusion methods have to tackle this fact too. Figure 4 shows the superiority of cokriging when it has access to online sampling of the truth model. However, when the truth model is noisy, it adversely affects the cokriging performance and the fused model is misleading, not representing the truth model correctly, although it is estimating the ground truth value better than reification based fusion. On the other side, reification based fusion is not influenced remarkably by noise and it has the capability to represent the truth model nicely but it is not estimating very close to the ground truth values.

In Fig. 5, the mean squared error is shown as a function of cost for all situations. The left plot presents the mean squared error when using some prior information about the truth model. The right plot considers when online sampling is possible. Again, the results show that cokriging is suggesting a better estimate considering the mean squared error, but its performance can be affected significantly in presence of noise. Also, note that the 95% confidence intervals have been grown in noisy conditions.

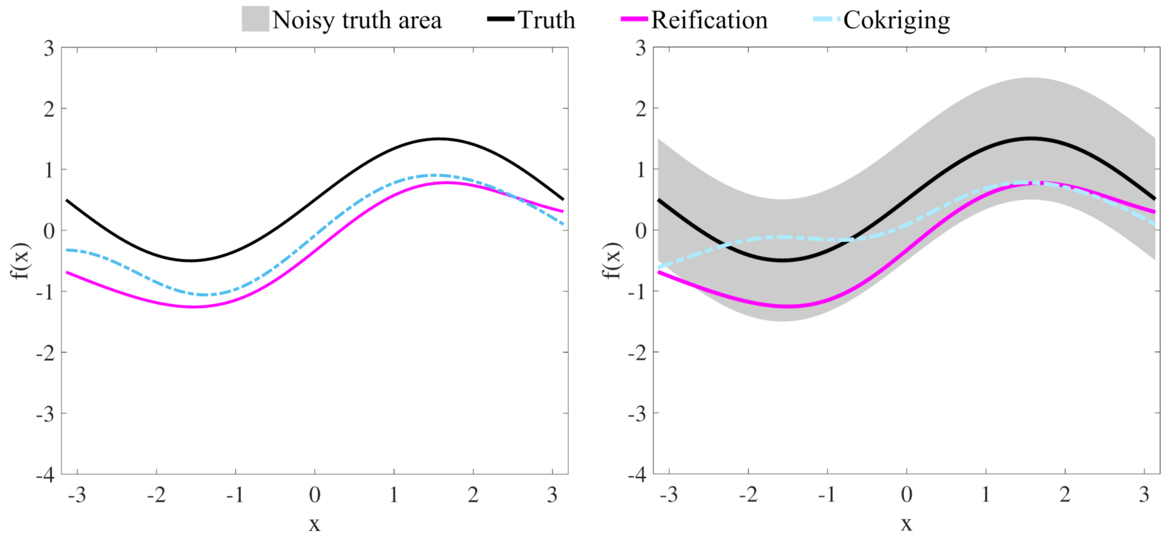

Fig. 3 Fused models built with reification and cokriging using prior information of the truth model with and without noise in the truth samples.

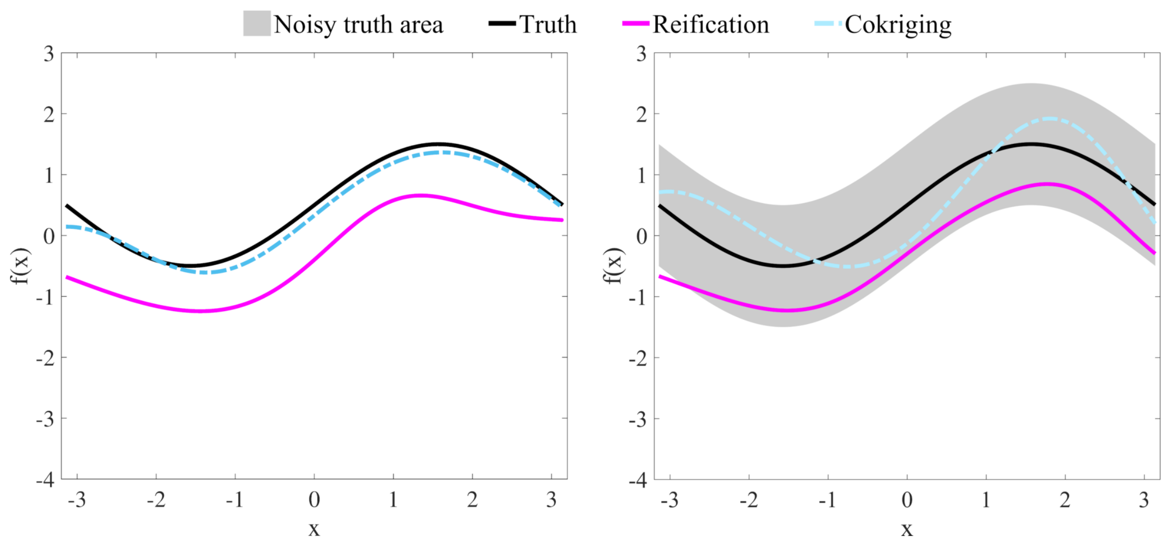

Fig. 4 Fused models built with reification and cokriging when online sampling from the truth model is possible considering noiseless and noisy sampling situations.

As seen in Figs. 3, 4, and 5, noise is an important factor influencing the fused model estimation of the truth model. In Fig. 6, a set of 5 experiments for building the fused model using the reification and cokriging approaches is designed to show how the truth model estimate is affected in the presence of noise. For both approaches, two conditions of using prior samples or online sampling is considered. The results show that with reification, we are able to handle the noisy information about the ground truth and obtain more consistent estimates of the truth model. However, the cokriging is highly dependent on the truth data. In the presence of noise, while the estimates are closer to the truth model, the overall shape of the fused model is not representing the correct variation of the objective value with respect to the design space. This could be very problematic for certain tasks requiring derivative estimates.

Sometimes it is not only the estimated value that is important, but also we need to gain information about the behavior of a black-box function when moving through different regions in the input design space. In other words, the gradient might give important information about the function in some applications as well. It has been shown

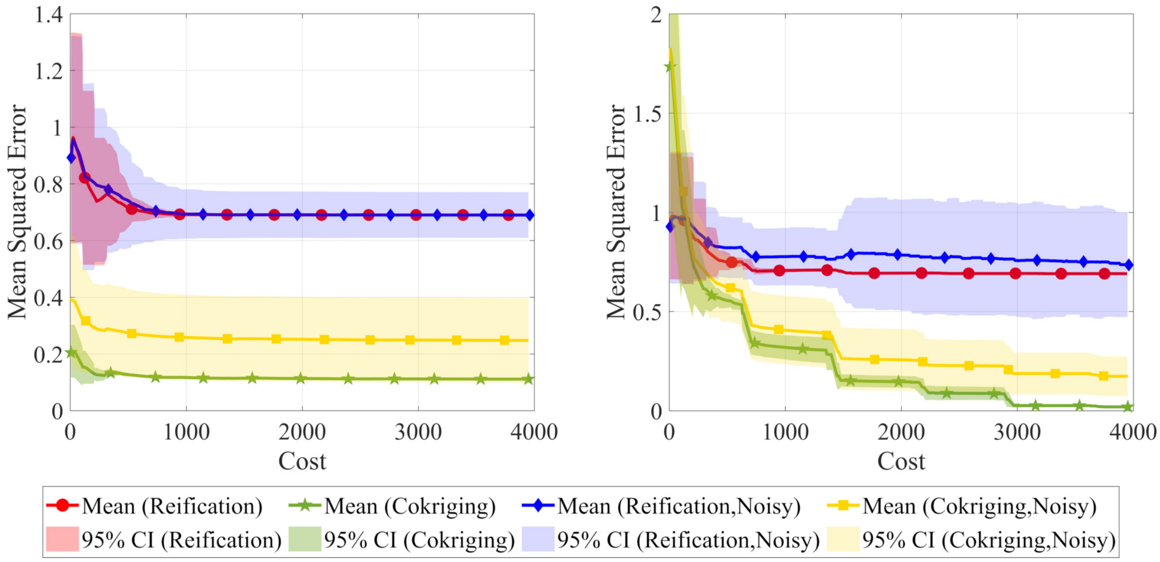

Fig. 5 The mean squared error as a function of cost. Left: using prior information. Right: online sampling of the truth model.

that in presence of noise, the cokriging method suggests a better estimate of the ground truth value but not a correct representation of the response surface shape. We define another criterion to address this aspect of a fused model performance. As the shape of the objective function is defined via the gradient of the objective function, we compare the error in derivatives to check how close the shape is to the correct representation of the truth model. The derivative mean squared error is shown versus cost in Fig. 7. As expected, the reification based fusion is performing better in the matter of how close the fused model approximates the derivative of the truth model.

Overall, the reification based approach performs more consistently over the different conditions tested. However, when there is information about the truth model, cokriging appears to provide better estimates. When considering the derivative of the truth model, which at times can be essential, reification appears to provide better estimates.

B. Real world problem application

In this section, both reification and cokriging approaches are applied on a computational fluid dynamics problem to build fused models to estimate the drag and lift coefficients of the NACA 0012 airfoil [39, 40] in different flight conditions. We use computational fluid dynamic simulator programs XFOIL [41] and SU2 [42] as the lower fidelity information sources and the real-world tunnel data of the NACA 0012 airfoil as the truth model to validate the methods. The problem has a two-dimensional input space with Mach number and angle of attack as the input variables. For this problem, the only information about the truth model are 68 data points previously evaluated. Thus, we are unable to do online sampling. Additionally, there is no evidence of how the data is collected and if it is corrupted by noise.

In Fig. 8, the drag and lift coefficients are plotted by fixing the Mach number to 0.3 and varying the angle of attack. As seen here, the simulations done by SU2 estimate the objective values closer to the truth value but the cost of a simulation in SU2 is more than XFOIL program. Based on the simulation time, the evaluation cost assigned to the cheap and expensive information sources are set to 30 and 50 units respectively.

Initially, we assume that there is no data point providing information about the truth model and there are only two information sources to estimate the quantities of interest. Simulated expert opinion is used to set each model discrepancy with respect to the truth model, where we assume SU2 is more accurate than XFOIL. By calculating the mean squared error associated to the fused models built with different fusion techniques, the results in Fig. 9 show the reification based method is estimating the coefficient of drag (left) better than cokriging but for coefficient of lift (right), both approaches suggest a very small error with cokriging slightly better. Looking at the errors in estimating the drag coefficient, we see the same behavior as in Fig. 2 that the fused model built with reification based approach is more accurate than cokriging. However, both reification based fusion and cokriging are approximating the truth model well. Note that a potential factor to influence the results is how the truth values are distributed in relative to other information sources.

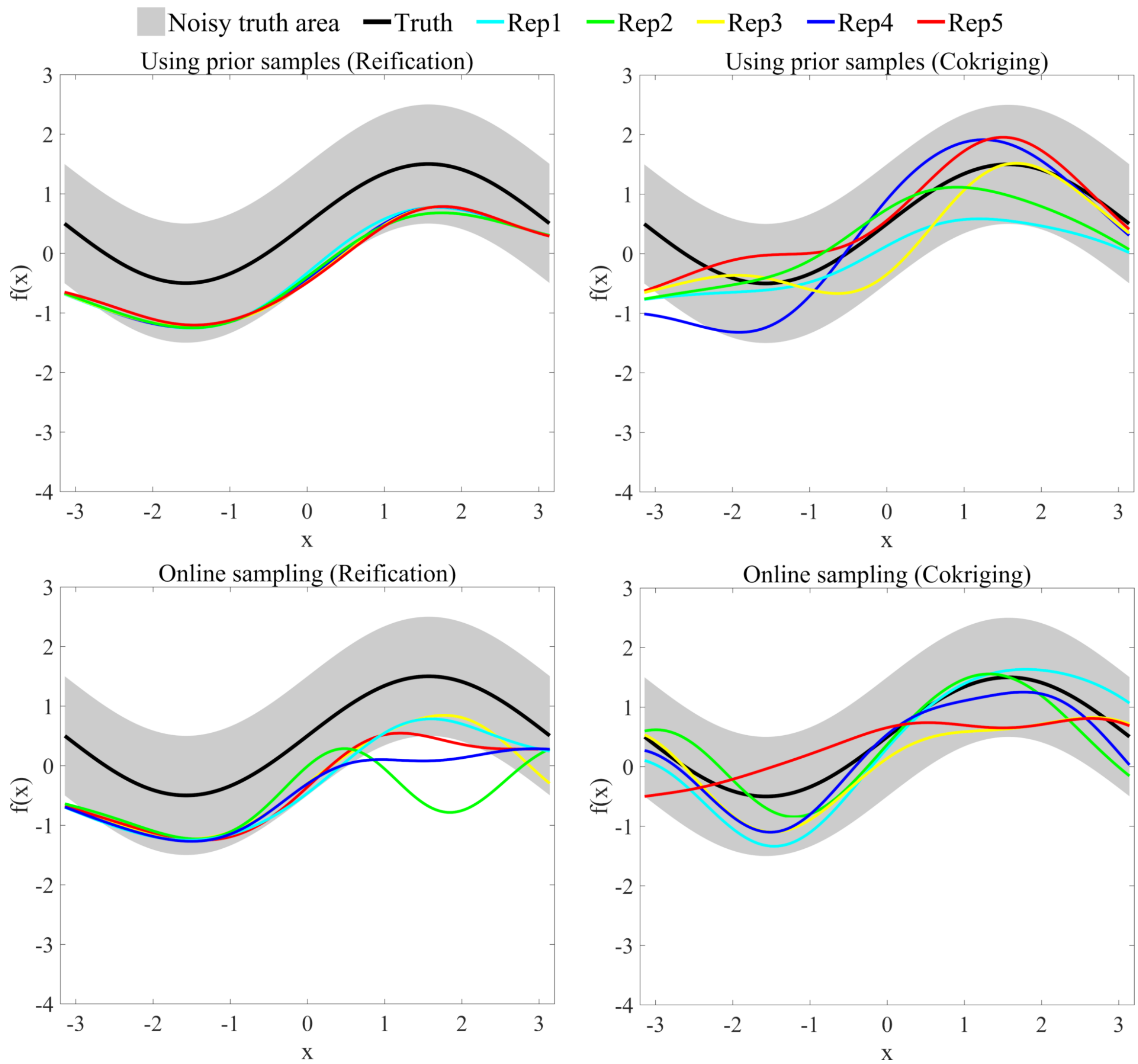

Fig. 6 Five different realizations of the fusion process using prior samples and online sampling for both methods.

For example, in Fig. 8, the drag coefficient is suggesting the truth values in between the values suggested by information sources while for lift coefficient, both information sources are estimating values larger than the truth values for most of the design region.

Finally, we use 10 data points as the prior information about the truth model and do the same process. However, this time the discrepancy of each information source is defined by the prior information we have added to the system. The results are shown in Fig. 10. Similar to the results seen in the test problems, the performance of cokriging performance is improved significantly and is slightly outperforming the reification based approach. Again, in approximating the lift coefficient, both approaches are acting similarly, suggesting close mean squared error estimating the lift coefficient. A point of interest here is that the addition of a few truth data points has resulted in a notable improvement in the cokriging initial esimation error. This does not occur in reification.

When using prior samples to add information about the truth model to the system, it is not surprising to see the error is increased in some cases, particularly if the design space is high dimensional. In general, we do not have many evaluations from the truth model to cover the design space sufficiently to calculate the discrepancy associated to every information source accurately everywhere. Consequently, the estimation error around areas where there is not enough training points grows due to miscalculation of models discrepancy. A subject of future study could be investigating how

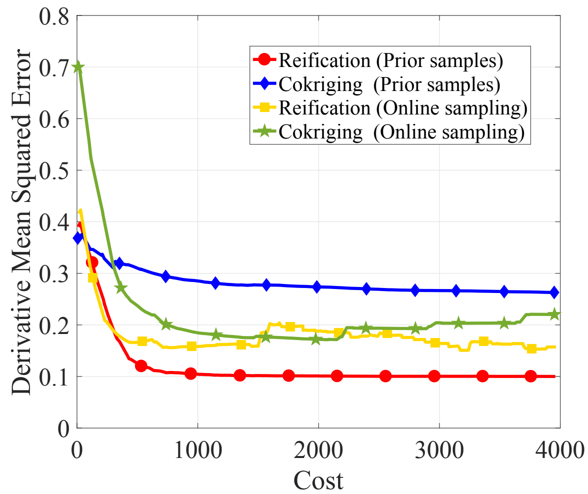

Fig. 7 Comparing the mean squared error in the derivative of the function suggested by each fused model.

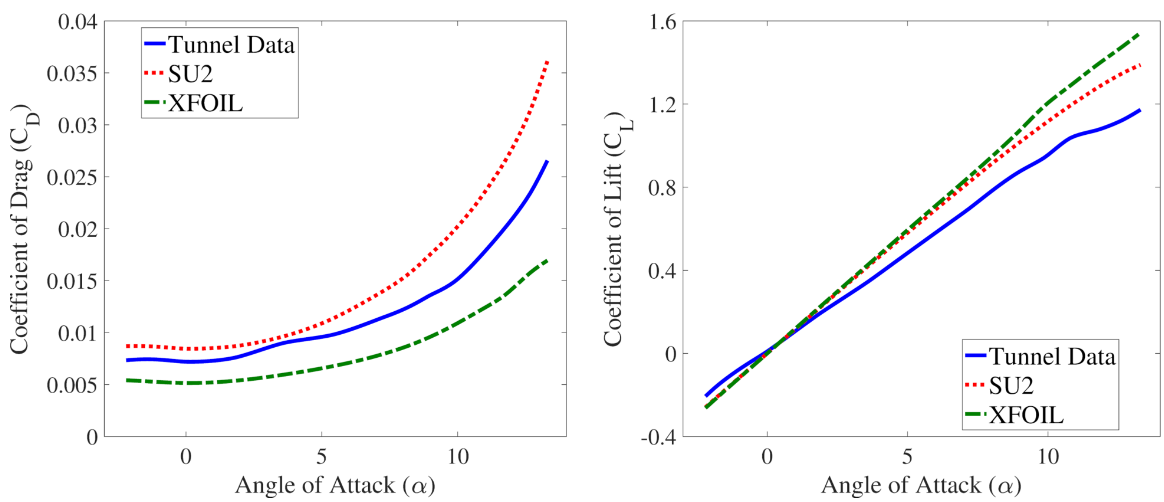

Fig. 8 Coefficients of drag and lift estimates from SU2 and XFOIL in comparison to the truth model determined by wind tunnel data for fixed Mach number of 0.3.

to import expert opinion in areas where information to calculate the model discrepancy is not sufficient, so that we can take advantage of both expert opinion and prior information in the process.

IV. Conclusions

In this study, the goal was to compare two known fusion approaches in a variety of conditions to investigate their advantages and disadvantages in different situations. First, by creating a series of test problems, it has been shown that in absence of any information about the ground truth quantity of interest and the truth model, the reification based approach is a more reliable choice for building a fused model to represent ground truth. When prior evaluations of a truth model are available we conclude that cokriging is a better option to estimate the truth values more accurately. However, in the presence of noise, the cokriging approach might be unable to represent the correct shape of the truth model, leading to potential degradation when gradient information is required. When online sampling is possible for sequential fusion, we have found that cokriging provides better estimates early in the process than reification. However, under noisy conditions, cokriging's reliance on the truth model again leads to inappropriate derivative estimates. Similar

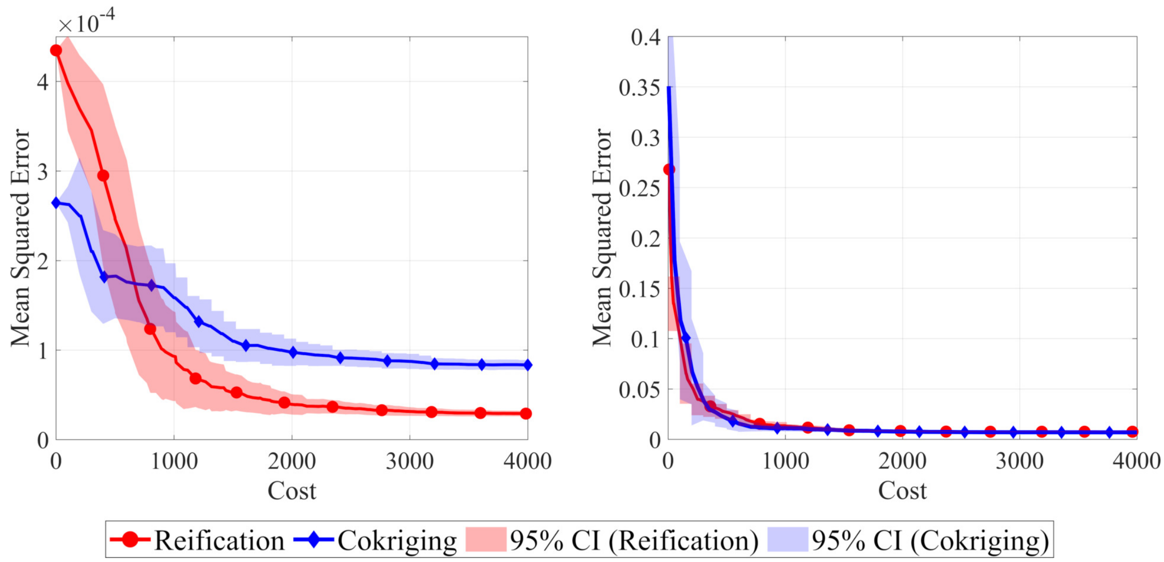

Fig. 9 Fused models estimation error with respect to the truth data for drag (left) and lift (right) coefficients when system has no information about the truth model.

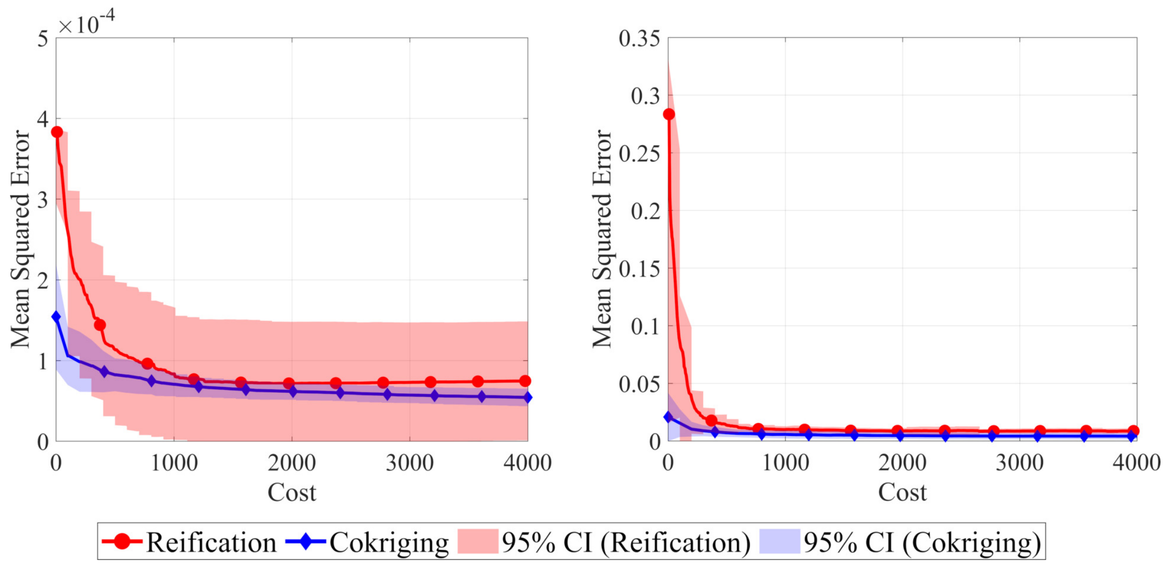

Fig. 10 Fused models estimation error with respect to the truth data for drag (left) and lift (right) coefficients when system has prior information as data points from the truth model.

results were found when each method was used for the NACA 0012 airfoil analysis. We note here that the conclusions made here have been made only for the test problems considered in this work. In future work we will consider robustness against faulty ground truth data and consider a more theoretic comparison. As it stands now, either fusion approach can be used for estimating quantities of interest under a variety of conditions. We note that reification should be preferred when only expert opinion exists regarding ground truth or when the shape, in terms of derivatives, of the quantity of interest are important. Cokriging is preferred when ground truth data is available and is accurate.

Acknowledgments

This work was supported by the AFOSR MURI on multi-information sources of multi-physics systems under Award Number FA9550-15-1-0038, program manager, Dr. Fariba Fahroo and by the National Science Foundation under grant no. CMMI-1663130. Opinions expressed in this paper are of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] Clyde, M., "Model Averaging," In Subjective and Objective Bayesian Statistics, 2nd ed., Chapter 13, Wiley-Interscience, 2003.

[2] Clyde, M., and George, E., "Model Uncertainty," *Statistical Science*, Vol. 19, 2004, pp. 81–94.

[3] Draper, D., "Assessment and Propagation of Model Uncertainty," *Journal of the Royal Statistical Society Series B*, Vol. 57, No. 1, 1995, pp. 45–97.

[4] Hoeting, J., Madigan, D., Raftery, A., and Volinsky, C., "Bayesian Model Averaging: A Tutorial," *Statistical Science*, Vol. 14, No. 4, 1999, pp. 382–417.

[5] Leamer, E., *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, John Wiley & Sons, New York, NY, 1978.

[6] Madigan, D., and Raftery, A., "Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam's Window," *American Statistical Association*, Vol. 89, No. 428, 1994, pp. 1535–1546.

[7] Mosleh, A., and Apostolakis, G., "The Assessment of Probability Distributions from Expert Opinions with an Application to Seismic Fragility Curves," *Risk Analysis*, Vol. 6, No. 4, 1986, pp. 447–461.

[8] Reinert, J., and Apostolakis, G., "Including Model Uncertainty in Risk-informed Decision Making," *Annals of Nuclear Energy*, Vol. 33, No. 4, 2006, pp. 354–369.

[9] Riley, M., and Grandhi, R., "Quantification of Modeling Uncertainty in Aeroelastic Analyses," *Journal of Aircraft*, Vol. 48, No. 3, 2011, pp. 866–873.

[10] Zio, E., and Apostolakis, G., "Two Methods for the Structured Assessment of Model Uncertainty by Experts in Performance Assessments of Radioactive Waste Repositories," *Reliability Engineering & System Safety*, Vol. 54, No. 2-3, 1996, pp. 225–241.

[11] Julier, S., and Uhlmann, J., "General Decentralized Data Fusion with Covariance Intersection," In D. Hall and J. Llinas, editor, Handbook of Data Fusion. CRC Press, Boca Raton FL, USA, 2001, ????

[12] Julier, S., and Uhlmann, J., "A Non-Divergent Estimation Algorithm in the Presence of Unknown Correlations," In proceedings of the American Control Conference, pp. 2369-2373, 1997.

[13] Geisser, S., "A Bayes Approach for Combining Correlated Estimates," *Journal of the American Statistical Association*, Vol. 60, 1965, pp. 602–607.

[14] Morris, P., "Combining Expert Judgments: A Bayesian Approach," *Management Science*, Vol. 23, 1977, pp. 679–693.

[15] Winkler, R., "Combining Probability Distributions from Dependent Information Sources," *Management Science*, Vol. 27, No. 4, 1981, pp. 479–488.

[16] Forrester, A. I., Sóbester, A., and Keane, A. J., "Multi-fidelity optimization via surrogate modelling," *Proceedings of the royal society a: mathematical, physical and engineering sciences*, Vol. 463, No. 2088, 2007, pp. 3251–3269.

[17] Tran, A., Wildey, T., and McCann, S., "sMF-BO-2CoGP: A sequential multi-fidelity constrained Bayesian optimization framework for design applications," *Journal of Computing and Information Science in Engineering*, Vol. 20, No. 3, 2020.

[18] Xiao, M., Zhang, G., Breitkopf, P., Villon, P., and Zhang, W., "Extended Co-Kriging interpolation method based on multi-fidelity data," *Applied Mathematics and Computation*, Vol. 323, 2018, pp. 120–131.

[19] Kullback, S., and Leibler, R. A., "On information and sufficiency," *The annals of mathematical statistics*, Vol. 22, No. 1, 1951, pp. 79–86.

[20] Ghoreishi, S., and Allaire, D., "Adaptive uncertainty propagation for coupled multidisciplinary systems," *AIAA journal*, 2017, pp. 3940–3950.

[21] Ghoreishi, S., Thomison, W., and Allaire, D., "Sequential Information-Theoretic and Reification-Based Approach for Querying Multi-Information Sources," *Journal of Aerospace Information Systems*, Vol. 16, No. 12, 2019, pp. 575–587.

[22] Allaire, D., and Willcox, K., "Fusing information from multifidelity computer models of physical systems," *2012 15th International Conference on Information Fusion*, IEEE, 2012, pp. 2458–2465.

[23] Thomison, W. D., and Allaire, D. L., "A model reification approach to fusing information from multifidelity information sources," *19th AIAA non-deterministic approaches conference*, 2017, p. 1949.

[24] Ghoreishi, S. F., and Allaire, D., "Multi-information source constrained Bayesian optimization," *Structural and Multidisciplinary Optimization*, Vol. 59, No. 3, 2019, pp. 977–991.

[25] Ghoreishi, S. F., Molkeri, A., Srivastava, A., Arroyave, R., and Allaire, D., "Multi-information source fusion and optimization to realize icme: Application to dual-phase materials," *Journal of Mechanical Design*, Vol. 140, No. 11, 2018, p. 111409.

[26] Rasmussen, C. E., and Williams, C. K. I., *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*, The MIT Press, 2005.

[27] Ghoreishi, S. F., Thomison, W. D., and Allaire, D., "Sequential Information-Theoretic and Reification-Based Approach for Querying Multi-Information Sources," *Journal of Aerospace Information Systems*, Vol. 16, No. 12, 2019, pp. 575–587.

[28] Khatamsaz, D., Peddareddygari, L., Friedman, S., and Allaire, D. L., "Efficient Multi-Information Source Multiobjective Bayesian Optimization," *AIAA Scitech 2020 Forum*, 2020, p. 2127.

[29] Alexandrov, N., Lewis, R., Gumbert, C., Green, L., and Newman, P., "Approximation and Model Management in Aerodynamic Optimization with Variable-Fidelity Models," *AIAA Journal*, Vol. 38, No. 6, 2001, pp. 1093–1101.

[30] Allaire, D., and Willcox, K., "Surrogate modeling for uncertainty assessment with application to aviation environmental system models," *AIAA journal*, Vol. 48, No. 8, 2010, pp. 1791–1803.

[31] Balabanov, V., Haftka, R., Grossman, B., Mason, W., and Watson, L., "Multifidelity Response Surface Model for HSCT Wing Bending Material Weight," *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, St. Louis, MO, 1998. AIAA 1998-4804.

[32] Balabanov, V., and Venter, G., "Multi-Fidelity Optimization with High-Fidelity Analysis and Low-Fidelity Gradients," *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Albany, New York, 2004. AIAA 2004-4459.

[33] Choi, S., Alonso, J. J., and Kroo, I. M., "Two-level multifidelity design optimization studies for supersonic jets," *Journal of Aircraft*, Vol. 46, No. 3, 2009, pp. 776–790.

[34] Eldred, M., Giunta, A., and Collis, S., "Second-order corrections for surrogate-based optimization with model hierarchies," 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 2004. AIAA 2004-4457.

[35] March, A., and Willcox, K., "Convergent multifidelity optimization using Bayesian model calibration," *Structural and Multidisciplinary Optimization*, Vol. 46, No. 1, 2012, pp. 93–109.

[36] March, A., and Willcox, K., "Provably convergent multifidelity optimization algorithm not requiring high-fidelity derivatives," *AIAA Journal*, Vol. 50, No. 5, 2012, pp. 1079–1089.

[37] Ghoreishi, S. F., Molkeri, A., Arróyave, R., Allaire, D., and Srivastava, A., "Efficient use of multiple information sources in material design," *Acta Materialia*, Vol. 180, 2019, pp. 260–271.

[38] Kennedy, M. C., and O'Hagan, A., "Predicting the output from a complex computer code when fast approximations are available," *Biometrika*, Vol. 87, No. 1, 2000, pp. 1–13.

[39] Rumsey, C., "Turbulence modeling resource," *2D NACA 0012 Airfoil Validation Case*, 2014.

[40] Thibert, J., Granjacques, M., and Ohman, L., "NACA0012 Airfoil AGARD Advisory Report No. 138," *Experimental Data Base for Computer Program Assessment*, 1979, pp. A1–9.

[41] Drela, M., "XFOIL: An analysis and design system for low Reynolds number airfoils," *Low Reynolds number aerodynamics*, Springer, 1989, pp. 1–12.

[42] Palacios, F., Alonso, J., Duraisamy, K., Colonno, M., Hicken, J., Aranake, A., Campos, A., Copeland, S., Economon, T., Lonkar, A., et al., "Stanford University unstructured (SU 2): An open-source integrated computational environment for multi-physics simulation and design," *51st AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition*, 2013, p. 287.