Eleventh International Conference on Computational Fluid Dynamics (ICCFD11), Maui, HI, USA, July 11-15, 2022

ICCFD11-3301

Physics-Based Regression vs. CFD for Hagen-Poiseuille and Womersley Flows and Uncertainty Quantification

H, Li*, M. Islam*, H, Yu*, and X. Du*
Corresponding author: duxi@iu.edu

* Indiana University–Purdue University Indianapolis, USA

Abstract: Computational fluid dynamics (CFD) and its uncertainty quantification are computationally expensive. We use Gaussian Process (GP) methods to demonstrate that machine learning can build efficient and accurate surrogate models to replace CFD simulations with significantly reduced computational cost without compromising the physical accuracy. We also demonstrate that both epistemic uncertainty (machine learning model uncertainty) and aleatory uncertainty (randomness in the inputs of CFD) can be accommodated when the machine learning model is used to reveal fluid dynamics. The demonstration is performed by applying simulation of Hagen-Poiseuille and Womersley flows that involve spatial and spatial-tempo responses, respectively. Training points are generated by using the analytical solutions with evenly discretized spatial or spatial-temporal variables. Then GP surrogate models are built using supervised machine learning regression. The error of the GP model is quantified by the estimated epistemic uncertainty. The results are compared with those from GPU-accelerated volumetric lattice Boltzmann simulations. The results indicate that surrogate models can produce accurate fluid dynamics (without CFD simulations) with quantified uncertainty when both epistemic and aleatory uncertainties exist.

Keywords: Supervised Machine Learning, Computational Fluid Dynamics, Volumetric lattice Boltzmann method, Surrogate Model, Uncertainty Quantification.

1 Introduction

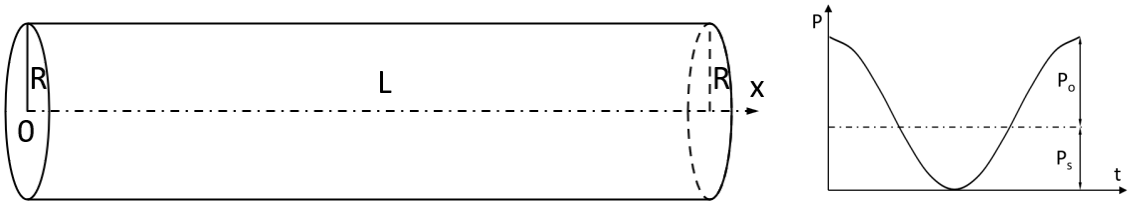
Computational Fluid Dynamics (CFD) plays an important role in solving various real-world flow systems. On the other hand, its heavy computational burden often needs a trade-off between accuracy and efficiency. To this end, surrogate models [1, 2] are commonly used to replace the expensive CFD models in engineering analysis and design. Recently, machine learning regression has been increasingly employed to build surrogate CFD models [3]. Surrogates built from machine learning regression have the potential to achieve both fast fluid simulations and high accuracy. With the high efficiency of the surrogate models, a CFD simulation can be performed in a matter of seconds or minutes, enabling fast what-if analysis, more design option evaluations, wider design space exploration, more reliable decision making, and rapid optimization.

Although surrogate models are inexpensive, they still have some model errors. When they are used for applications, the model input may also have uncertainties [4-7], such as random geometry, loading, material properties, and manufacturing imprecision. There is therefore a need to quantify the effects of uncertainty on the model prediction. In this study, in addition to demonstrating that physics-supervised regression can produce efficient and accurate surrogate models, which can significantly reduce the computational time without compromising the accuracy of CFD, we also demonstrate that the

uncertainties in both surrogate model structure and model input can be quantified so that their effects on predictions (outputs) of a CFD simulation can be estimated.

2 Problem Statement

In this study, we develop two surrogate models for solving steady Hagen-Poiseuille flow and unsteady Womersley flow. Both are Newtonian incompressible pipe flows, shown in Fig. 1(a), with analytical solutions. The existence of analytical solutions allows us to demonstrate the effectiveness of surrogate models and the use of uncertainty quantification. The flow domain is a long perfect pipe with a radius R and length L. The flow is driven by a pressure gradient, $P = P_s + P_o \cos(\omega t)$, where P_s and P_0 represent the steady pressure gradient and the amplitude of the oscillating pressure gradient, respectively, and ω is the angular oscillating frequency, as shown in Fig. 1(b).



(a) (b)

Fig. 1 (a) Isometric view of a pipe, (b) A waveform of spatiotemporal pressure gradient

2.1 Hagen-Poiseuille Flow

The Hagen-Poiseuille flow is a steady flow driven by a steady (constant) pressure gradient (P_s). The analytical solution of the Hagen-Poiseuille flow is given by

$$u(r) = \frac{P_s R^2}{4\mu}\left(1 - \frac{r^2}{R^2}\right) \tag{1}$$

where μ is the dynamic viscosity, and r is the radial distance to the center of the cross-section. If we denote the maximum velocity by $u_{max} = \frac{P_s R^2}{4\mu}$, the normalized solution is found as

$$u'(r') = 1 - r'^2 \tag{2}$$

where the normalized velocity is $\boldsymbol{u'} = \frac{\boldsymbol{u}}{\boldsymbol{u_{max}}}$, and the normalized distance to the cross-section center is $\boldsymbol{r'} = \frac{\boldsymbol{r}}{\boldsymbol{R}}$.

2.2 Womersley Flow

A Womersley flow is driven by an oscillating pressure gradient $P = P_s + P_o\, e^{i\omega t}$. The analytical solution is given by

$$u(r,t) = \frac{P_s R^2}{4\mu}\left(1 - \frac{r^2}{R^2}\right) + Real\left\{\frac{P_o R^2}{i\mu\alpha^2}\left[1 - \frac{J_0\left(\frac{\alpha r}{R} i^{\frac{3}{2}}\right)}{J_0\left(\alpha i^{\frac{3}{2}}\right)}\right] e^{i\omega t}\right\} \tag{3}$$

It is noted that Eq. (3) recovers the solution of Hagen-Poiseuille flow. i.e. Eq. (2), when no oscillating pressure gradient exists, e.g. $P_0 = 0$.

The velocity $u(r,t)$ can be normalized as

$$u'(r',t') = (1-r'^2) + \frac{4A}{\alpha^2} Real\left\{\frac{1}{i}\left[1 - \frac{J_0\left(\alpha r' i^{\frac{3}{2}}\right)}{J_0\left(\alpha i^{\frac{3}{2}}\right)}\right] e^{it'}\right\} \quad (4)$$

where the angular velocity is $t' = \omega t$, and $A = \frac{P_0}{P_s}\mu$. $\alpha = L\sqrt{\frac{\omega\rho}{\mu}}$ and J_0 are the Wormsley number and the Bessel function of the first kind of order zero, respectively.

2.3 Methodology

We generate surrogate models for both Hagen-Poiseuille flow and Womersley flow by regression. We compare the surrogate models with numerical solutions from CFD. CFD is performed using the volumetric lattice Boltzmann method [8] through an in-house GPU accelerated code [9]. We impose pressure gradient as a body force, rigid wall, and periodic conditions at the inlet and outlet. Both regression and CFD share the same parameters. The physical quantities used in the study is given in Table 1.

Table 1 Physical quantities for the Hagen-Poiseuille and Womersley flow

Variables	Values
Angular frequency ω	$7.85\ s^{-1}$
Kinematic viscosity μ	3.415×10^{-6}
Density ρ	$1025\ \mathrm{kg/m^3}$
Steady pressure gradient P_s	$280.00\ \mathrm{Pa/m}$
Amplitude of oscillating pressure gradient P_0	$488.3525\ \mathrm{Pa/m}$

The surrogate models are built via Gaussian Process (GP) regression. For the unsteady Womersley flow, time is also included as a dimension of the input training points. The prediction error can be estimated by the model uncertainty represented by the standard deviation of the prediction. It is also possible to estimate prediction error for a surrogate model built with neural network regression [10].

3 Gaussian Process Regression

There are many machine learning regression methods, such as Neural Network Regression [11] and GP regression [12]. In this study, we use GP since it can easily quantify epistemic (model) uncertainty for the model error.

The aim of GP is to build a surrogate model using training points for a general nonlinear CFD model. The training points include both the model input and output, and GP is therefore a supervised machine learning method. GP has several benefits. In addition to the ability of quantifying epistemic uncertainty, GP works well on small and medium size datasets and can provide high accuracy for nonlinear functions.

Denote the training dataset by $\{(\boldsymbol{x}_i, y_i);\ i = 1, 2, \dots, n\}$, $\boldsymbol{x} \in \mathbb{R}^d$ and $\boldsymbol{y} \in \mathbb{R}$. The dataset is used to train a nonlinear model given by

$$\hat{y}(\boldsymbol{x}) = h(\boldsymbol{x})^T\boldsymbol{\beta} + \varepsilon \quad (5)$$

where $h(\boldsymbol{x})$ is a vector of basic functions, $\boldsymbol{\beta}$ is a vector of to-be-determined coefficients, and $\varepsilon(\boldsymbol{x}) \sim N(0, \sigma^2)$ is a noise term, $\boldsymbol{x}$ is a vector of input variables, and $\hat{y}$ is the model output. The responses (outputs) from the model at different input points are assumed to follow a joint Gaussian distribution; or the response $\hat{y}$ follows a Gaussian Process (GP).

$$\hat{y}(\boldsymbol{x}) \sim N(\mu(\boldsymbol{x}), k(\boldsymbol{x}, \boldsymbol{x}')) \quad (6)$$

where $\mu(\boldsymbol{x})$ and $k(\boldsymbol{x}, \boldsymbol{x}'))$ are the mean and covariance of the GP, respectively. The GP model is therefore a probabilistic model. A linear transformation is performed so that the mean function becomes zero. The prior joint distribution of observed responses $\boldsymbol{y}$ can be modeled as

$$p(\boldsymbol{y}|f(\boldsymbol{x}), \boldsymbol{x}) \sim N(\boldsymbol{y}|\boldsymbol{h}(\boldsymbol{x})^T\boldsymbol{\beta} + f(\boldsymbol{x}), \sigma_n^2 I_n) \quad (7)$$

where I is a $n \times n$ identity matrix. $f(\boldsymbol{x}) \sim N(0, k(\boldsymbol{X}, \boldsymbol{X}))$ follows a multivariate normal distribution.

$k(\boldsymbol{X},\boldsymbol{X})$ is a symmetric covariance matrix, where $\boldsymbol{X}=\left[\boldsymbol{x_1^T},\boldsymbol{x_2^T},\dots,\boldsymbol{x_n^T}\right]$. Thus, the distribution of prior observed values of $\boldsymbol{y}$ and the new responses at test points $\boldsymbol{y}^*$ becomes

$$\begin{pmatrix}\boldsymbol{y}\\ \boldsymbol{y}^*\end{pmatrix}\sim N\left(0,\begin{matrix}\boldsymbol{K}+\sigma_n^2 I & \boldsymbol{K}_*^T\\ \boldsymbol{K}_* & \boldsymbol{K}_{**}\end{matrix}\right) \tag{8}$$

where $K_{ij}=K(\boldsymbol{x}_i,\boldsymbol{x}_j)$ is a symmetric covariance matrix of response at the training points; K_{**} is the covariance matrix of predicted responses to the test points; $\boldsymbol{K}_*=\boldsymbol{K}_*^T$ is the covariance matrix of the responses at test and training points. The covariance kernel function is critical to the accuracy of the model predictions. The most widely used kernel function is the squared exponential kernel. The distribution of the prediction is a conditional distribution and is given by

$$\boldsymbol{y}^*|\boldsymbol{X},\boldsymbol{y},\boldsymbol{X}^*\sim N(\boldsymbol{\mu}^*,\boldsymbol{\Sigma}^*) \tag{9}$$

where $\boldsymbol{\mu}^*$ is the mean of the predicted responses and $\boldsymbol{\Sigma}^*$ is the variance of the prediction, which are given by

$$\begin{aligned}\boldsymbol{\mu}^* &= \boldsymbol{K}_*[\boldsymbol{K}+\sigma_n^2 I_n]^{-1}\boldsymbol{y}\\ \boldsymbol{\Sigma}^* &= \boldsymbol{K}_{**}-\boldsymbol{K}_*[\boldsymbol{K}+\sigma_n^2 I]^{-1}\boldsymbol{K}_*^T\end{aligned} \tag{10}$$

$\boldsymbol{\Sigma}^*$ provides not only the standard deviation of the response at a test point, but also the dependence of responses at all the test points. The former indicates the model or epistemic uncertainty. The higher is the standard deviation of the predicted response, the higher is the epistemic uncertainty.

4 Uncertainty Quantification

When we use the surrogate CDF model for predictions, we will encounter two types of uncertainty: epistemic uncertainty and aleatory uncertainty [13]. Epistemic uncertainty is the model uncertainty in the surrogate model as we discussed in Sec. 3. Aleatory uncertainty is due to the intrinsic randomness in the model input, such as random loading, material properties, and boundary conditions. When we use the surrogate model for real applications, we should consider the effects of both types of uncertainty. They can be modeled by random variables. We now discuss how to predict the probability distribution of CFD output from the surrogate model with the two types of uncertainty.

The response from a GP surrogate model is $y=g^{GP}(\boldsymbol{x})$, where $\boldsymbol{x}=(x_1,x_2,\dots,x_d)^T$ with independent random variables $x_i,\ i=1,2,\dots,d$, which follow certain distributions with probability density function $f_{x_i}(\cdot)$. The input variables have aleatory uncertainty. If the input variables are not independent, they can be converted into independent ones. The conversion may be performed before the regression or after.

The conditional distribution of the predicted y given $\boldsymbol{x}$ is a normal distribution.

$$y|\boldsymbol{x}=g^{GP}(\boldsymbol{x})\sim N(\hat{\mu}_y(\boldsymbol{x}),\hat{\sigma}_y^2(\boldsymbol{x})) \tag{11}$$

The conditional distribution is caused by the epistemic uncertainty in the surrogate model. Let $f(\boldsymbol{x},y)$ be the joint PDF of $\boldsymbol{x}$ and y. Given both types of uncertainty, the mean of y is calculated by

$$\mathbb{E}(y)=\int \hat{\mu}_y(\boldsymbol{x})f(\boldsymbol{x})d\boldsymbol{x}=\int \hat{\mu}_y(\boldsymbol{x})f(\boldsymbol{x})d\boldsymbol{x}=\mathbb{E}_{\boldsymbol{x}}\left[\hat{\mu}_y(\boldsymbol{x})\right] \tag{12}$$

where $\mathbb{E}(\cdot)$ stands for an expectation, and $f(\boldsymbol{x})$ is the joint PDF of $\boldsymbol{x}$.

The variance of the response is

$$\mathbb{V}(y)=\mathbb{E}[y^2]-(\mathbb{E}[y])^2=\mathbb{E}_{\boldsymbol{x}}\left[\hat{\sigma}_y^2(\boldsymbol{x})\right]+\mathbb{V}_{\boldsymbol{x}}\left[\hat{\mu}_y(\boldsymbol{x})\right] \tag{13}$$

The first and second terms on the right-hand side indicate the contributions to the overall uncertainty from the model uncertainty and data uncertainty, respectively.

Next, we discuss how to estimate the cumulative distribution function (CDF) of the response.

Without losing generality, we find the following probability:

$$p_f=\mathrm{P_r}(y<0) \tag{14}$$

This probability is called the probability of failure if a failure occurs when $y<0$ [14]. This probability is commonly used in reliability-based design [15]. It is also the CDF at $y=0$.

$$p_f=\Pr(y<0)=\int\limits_{y(\boldsymbol{x})<0} f_{xy}(\boldsymbol{x},y)d\boldsymbol{x}dy=\int\left(\Phi\left(-\frac{\mu_y(\boldsymbol{x})}{\sigma_y(\boldsymbol{x})}\right)f(\boldsymbol{x})\right)d\boldsymbol{x} \tag{15}$$

where $f_{xy}(\boldsymbol{x},y)$ is the joint PDF of $\boldsymbol{x}$ and y, and $\Phi(\cdot)$ is the CDF of a standard normal variable. Let

$$e(\boldsymbol{x}) = \Phi\left(-\frac{\mu_y(\boldsymbol{x})}{\sigma_y(\boldsymbol{x})}\right) \tag{16}$$

Then

$$p_f = \int e(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x} = \mathbb{E}_{\boldsymbol{x}}[e(\boldsymbol{x})] \tag{17}$$

$\mathbb{E}_{\boldsymbol{x}}(\cdot)$, $\mathbb{V}(\cdot)$, and p_f are all expectations with respect to $\boldsymbol{x}$. There are two methods to estimate the expectation. The first way is Monte Carlo simulation (MCS) [16-18], and the second is Taylor expansion at means $\boldsymbol{\mu_x} = \left(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_d}\right)^{\mathrm{T}}$. The use of MCS is straightforward. We discuss the second method briefly. Denote a nonlinear function by $v(\boldsymbol{x})$, and $v(\boldsymbol{x})$ could be $\mu_y(\boldsymbol{x})$, $\sigma_y(\boldsymbol{x})$, or $e(\boldsymbol{x})$.

$$v(\boldsymbol{x}) \approx v(\boldsymbol{\mu_x}) + \nabla v(\boldsymbol{\mu_x})(\boldsymbol{x} - \boldsymbol{\mu_x}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu_x})^T H(\boldsymbol{\mu_x})(\boldsymbol{x} - \boldsymbol{\mu_x}) \tag{18}$$

where $\nabla h(\cdot)$ and $H(\cdot)$ are the gradient and the Hessian of $v(\cdot)$, respectively. Then

$$E(y) \approx h(\boldsymbol{\mu_x}) + \frac{1}{2} T_r(H(\boldsymbol{\mu_x})\Sigma_{\boldsymbol{x}}) \tag{19}$$

where $T_r(\cdot)$ is the trace of a matrix.

Since the components of $\boldsymbol{x}$ are independent,

$$E(y) \approx g^{GP}(\boldsymbol{\mu_x}) + \sum_{i=1}^{d} \frac{\partial^2 g^{GP}}{\partial x_i^2}(\boldsymbol{\mu_x})\sigma_{x_i}^2 \tag{20}$$

If we use the first-order Taylor expansion,

$$\mathbb{V}(y) \approx \sum_{i=1}^{d} \left(\frac{\partial g^{GP}}{\partial x_i}(\boldsymbol{\mu_x})\sigma_{x_i}\right)^2 \tag{21}$$

The probability of failure is approximated by

$$p_f = e(\boldsymbol{\mu_x}) + \sum_{i=1}^{d} \frac{\partial^2 e}{\partial x_i}(\boldsymbol{\mu_x})\sigma_{x_i}^2 \tag{22}$$

The derivatives are evaluated by the finite difference method.

5 Numerical Results

In this section, we compare the GP surrogate model prediction and CFD simulation for the velocity profile along a radius. In addition, we perform uncertainty quantification for Womersley flow. The analysis includes model uncertainty and data uncertainty.

The detailed information of input variables in training the GP models is shown in Table 2, and the observed responses are obtained by the normalized analytical solution in Eqs. (2) and (4).

Table 2 Distributions of input variables

Variable		Minimum	Maximum	Distribution	Number of points
Hagen-Poiseuille flow	$u_{\max}\ (m/s)$	0.1	5.0	Uniform	450
	$R(m)$	0.0025	0.5	Single	1
Womersley flow	A	1.0	2.0	Uniform	10
	α	2.0	5.0	Uniform	10
	r′	0.0	1.0	Uniform	30
	t′	4.7414	10.2364	Uniform	10

The results are plotted in Fig. 2 for steady Hagen-Poiseuille flow and Fig. 3 for unsteady Womersley flow. For the steady flow, both GP predictions and CFD simulation achieve nearly identical profiles to the analytical solutions. For the unsteady Womersley flow, we use four representative time instants $t' = [5.9625,\ 7.0964, 8.6664, 10.2364]$ in an oscillation, and all the velocity profiles from

GP at the time instants in oscillation are again identical to the analytical solutions, whereas CFD results have noticeable deviations from the analytical solutions.

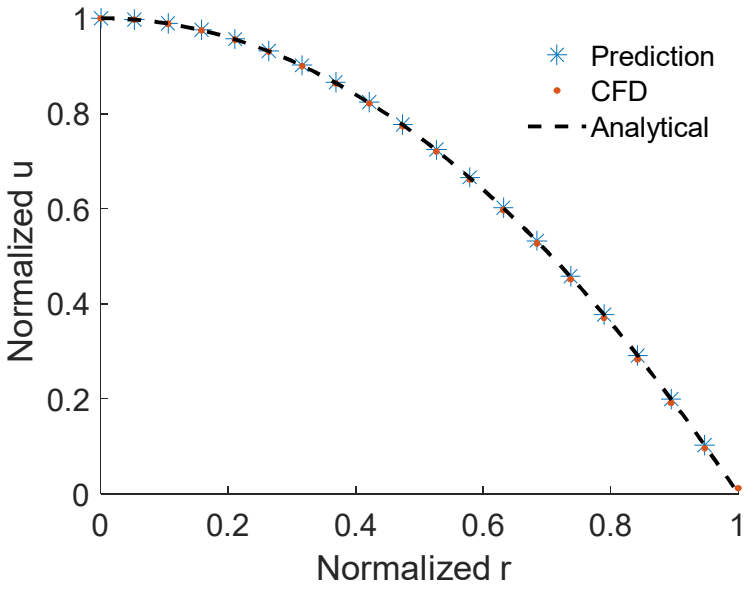


Fig. 2 The comparison of velocity profiles of Hagen-Poiseuille flow along a radial direction.

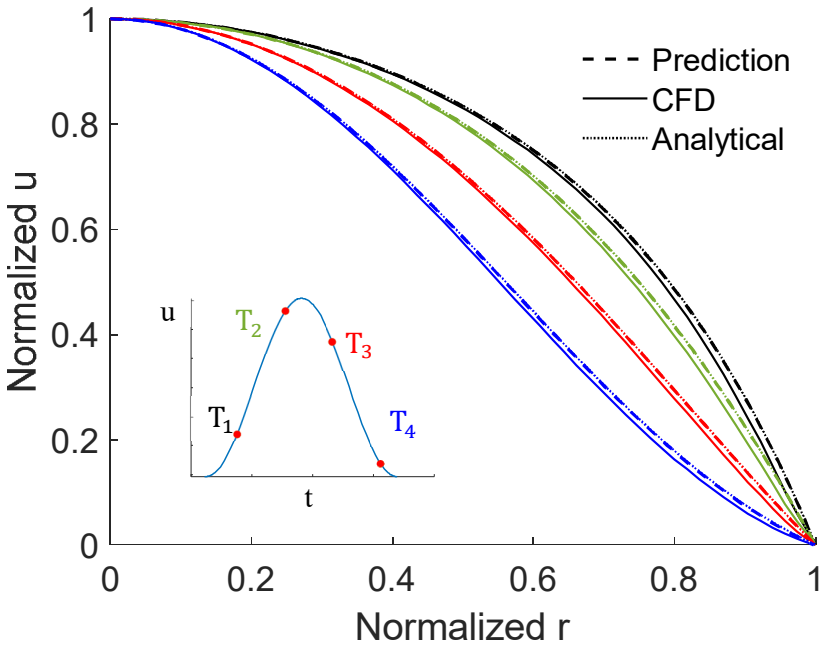


Fig. 3 Comparisons of velocity profiles of Womersley flow along a radial direction at 4 representative time instants in an oscillating cycle.

We use MCS to quantify the model uncertainty, and the sample size is $1e^4$. There are two independent input random variables $\boldsymbol{x} = [A, \alpha]$, and they follow normal distributions. Their distribution parameters are given in Table 3.

Table 3 Distributions of random variables

Variable	Mean	Std	Distribution
A	1.7441	0.0750	Normal
α	3.7523	0.1250	Normal

We compare the predictions from GP with analytical solutions at the time instant 5.9625 s. The mean and standard deviations are given in Tables 4 and 5, respectively. The 95% confidence intervals are also plotted in Fig. 4. The two tables and Fig. 4 indicate good accuracy of the GP model.

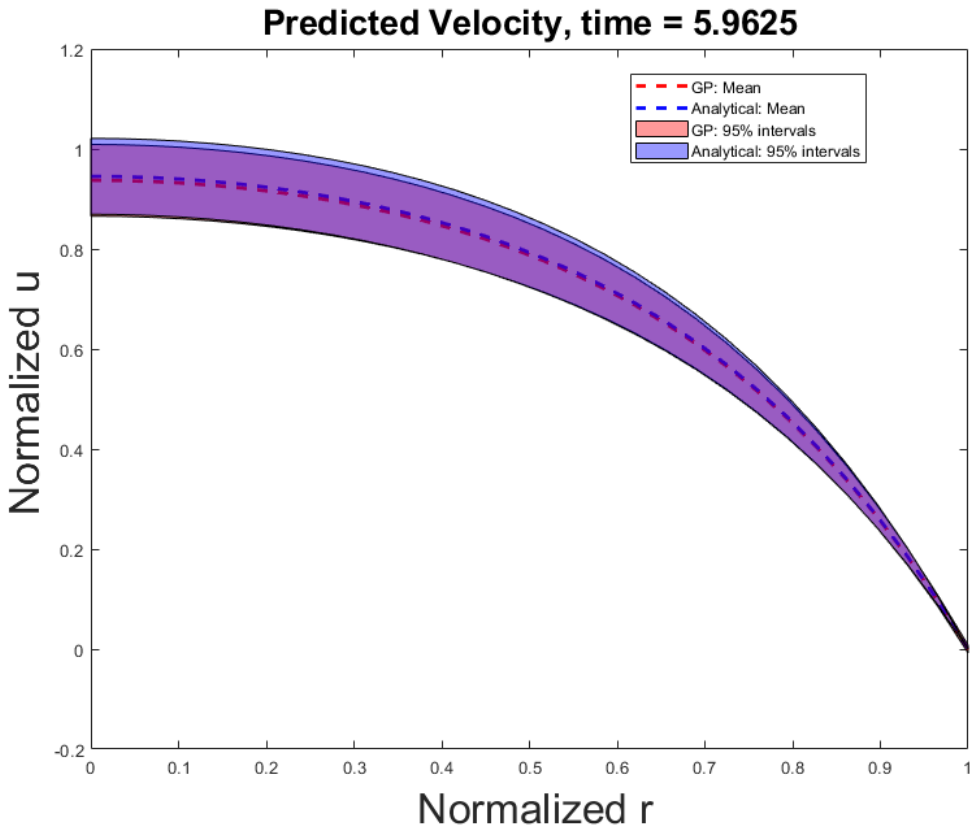


Fig. 4 Predictions from GP and analytical model

The contributions of the model and data uncertainty are provided in Table 6. The second column shows the standard deviations of the contributions from data uncertainty to the overall uncertainty in the last column. It corresponds to the second term on the right-hand side of Eq. (13). The third column shows the contributions from the model uncertainty, corresponding to the first term on the right-hand side of Eq. (13). The uncertainty of the prediction is mainly due to the uncertainty from the data, which is much larger than the model uncertainty.

6 Conclusions and Future Work

Using steady Hagen-Poiseuille and unsteady Womersley flows, we demonstrated that the surrogate predictions are much faster than CFD and equivalent (Hagen-Poiseuille flow) to and more accurate (unsteady Womersley flow) than CFD simulations. We also demonstrated that the model error of the surrogate model can be estimated by the quantification of model uncertainty. Given sufficient training points, the prediction from the surrogate model is accurate with small model uncertainty. Uncertainty

quantification can also accommodate uncertainty in the model input. It helps users understand the effects of uncertainty on the prediction and contributions from different sources of uncertainty. It will assist in more reliable decision-making. Our further work will be uncertainty quantification for more complex CFD simulations without analytical solutions, aiming to significantly reduce the computational time. We will train surrogate models for more realistic steady and unsteady flows and perform a full-scale uncertainty quantification analysis.

Table 4 Mean predictions from GP and analytical solution

r	Mean (GP)	Mean (analytical)	Error	Relative error (%)
0	0.93743	0.94538	0.00794	0.8403
0.0345	0.93681	0.94475	0.00793	0.8394
0.0690	0.93496	0.94285	0.00789	0.8370
0.1034	0.93185	0.93968	0.00782	0.8329
0.1379	0.92747	0.93520	0.00773	0.8273
0.1724	0.92177	0.92939	0.00762	0.8201
0.2069	0.91472	0.92220	0.00748	0.8114
0.2414	0.90625	0.91357	0.00732	0.8012
0.2759	0.89629	0.90342	0.00713	0.7896
0.3103	0.88476	0.89168	0.00692	0.7766
0.3448	0.87156	0.87825	0.00669	0.7624
0.3793	0.85656	0.86301	0.00644	0.7469
0.4138	0.83964	0.84582	0.00617	0.7303
0.4483	0.82063	0.82652	0.00589	0.7127
0.4828	0.79936	0.80495	0.00558	0.6942
0.5172	0.77562	0.78089	0.00527	0.6748
0.5517	0.74918	0.75412	0.00493	0.6548
0.5862	0.71978	0.72438	0.00459	0.6340
0.6207	0.68714	0.69138	0.00423	0.6128
0.6552	0.65095	0.65482	0.00387	0.5912
0.6897	0.61084	0.61434	0.00349	0.5692
0.7241	0.56644	0.56956	0.00311	0.5470
0.7586	0.51734	0.52007	0.00272	0.5246
0.7931	0.46308	0.46541	0.00233	0.5021
0.8276	0.40317	0.40512	0.00194	0.4797
0.8621	0.33712	0.33867	0.00154	0.4573
0.8966	0.26436	0.26551	0.00115	0.4351
0.9310	0.18432	0.18509	0.00076	0.4128
0.9655	0.09641	0.09679	0.00037	0.3890

Table 5 Standard deviations from GP and analytical model

r	Standard deviation (GP)	Standard deviation (analytical)	Error	Relative error (%)
0	0.03585	0.03792	0.00794	5.4482
0.0345	0.03584	0.03792	0.00793	5.4805
0.0690	0.03580	0.03791	0.00789	5.5765
0.1034	0.03573	0.03791	0.00782	5.7326
0.1379	0.03564	0.03789	0.00773	5.9425
0.1724	0.03551	0.03786	0.00762	6.1980
0.2069	0.035351	0.03781	0.00748	6.4894
0.2414	0.03516	0.03772	0.00732	6.8060

0.2759	0.03491	0.03760	0.00713	7.1371
0.3103	0.03462	0.03742	0.00692	7.4719
0.3448	0.03427	0.03717	0.00669	7.8007
0.3793	0.03385	0.03684	0.00644	8.1145
0.4138	0.03335	0.03641	0.00617	8.4058
0.4483	0.03275	0.03586	0.00589	8.6683
0.4828	0.03206	0.03519	0.00558	8.8968
0.5172	0.03124	0.03436	0.00527	9.0868
0.5517	0.03029	0.03337	0.00493	9.2347
0.5862	0.02920	0.03220	0.00459	9.3362
0.6207	0.02795	0.03084	0.00423	9.3864
0.6552	0.02653	0.02928	0.00387	9.3782
0.6897	0.02494	0.02749	0.00349	9.3007
0.7241	0.02316	0.02549	0.00311	9.1362
0.7586	0.02119	0.02325	0.00272	8.8547
0.7931	0.01902	0.02076	0.00233	8.4016
0.8276	0.01664	0.01803	0.00194	7.6686
0.8621	0.01406	0.01503	0.00154	6.4096
0.8966	0.01128	0.01175	0.00115	3.9488
0.9310	0.00835	0.00817	0.00076	2.2266
0.9655	0.00549	0.00426	0.00037	28.735

Table 6 Contributions to overall uncertainty from model and data uncertainty

r	Standard deviation from data uncertainty	Standard deviation from model uncertainty	Standard deviation of the prediction
0	0.03564	0.00388	0.03585
0.0345	0.03563	0.00388	0.03584
0.0690	0.03559	0.00388	0.03580
0.1034	0.03552	0.00388	0.03573
0.1379	0.03543	0.00388	0.03564
0.1724	0.03530	0.00388	0.03551
0.2069	0.03514	0.00388	0.03535
0.2414	0.03494	0.00388	0.03516
0.2759	0.03470	0.00388	0.03491
0.3103	0.03440	0.00388	0.03462
0.3448	0.03405	0.00388	0.03427
0.3793	0.03362	0.00388	0.03385
0.4138	0.03312	0.00388	0.03335
0.4483	0.03252	0.00388	0.03275
0.4828	0.03182	0.00388	0.03206
0.5172	0.03100	0.00388	0.03124
0.5517	0.03004	0.00388	0.03029
0.5862	0.02894	0.00388	0.02920
0.6207	0.02768	0.00388	0.02795
0.6552	0.02624	0.00388	0.02653
0.6897	0.02463	0.00388	0.02494
0.7241	0.02283	0.00388	0.02316
0.7586	0.02083	0.00388	0.02119
0.7931	0.01862	0.00388	0.01902
0.8276	0.01618	0.00388	0.01664
0.8621	0.01351	0.00388	0.01406
0.8966	0.01059	0.00388	0.01128

0.9310	0.00739	0.00388	0.00835
0.9655	0.00388	0.00388	0.00549

Acknowledgements

The research is supported by National Science Foundation through grant CBET 1803845. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation Grant No. ACI-1548562.

References

[1] B. A. Craven, K. I. Aycock, L. H. Herbertson, and R. A. Malinauskas, "A CFD-based Kriging surrogate modeling approach for predicting device-specific hemolysis power law coefficients in blood-contacting medical devices," *Biomechanics and Modeling in Mechanobiology,* vol. 18, no. 4, pp. 1005-1030, 2019/08/01 2019, doi: 10.1007/s10237-019-01126-4.

[2] Z. Qian, C. C. Seepersad, V. R. Joseph, J. K. Allen, and C. F. Jeff Wu, "Building Surrogate Models Based on Detailed and Approximate Simulations," *Journal of Mechanical Design,* vol. 128, no. 4, pp. 668-677, 2005, doi: 10.1115/1.2179459.

[3] Y. Morita, S. Rezaeiravesh, N. Tabatabaei, R. Vinuesa, K. Fukagata, and P. Schlatter, "Applying Bayesian optimization with Gaussian process regression to computational fluid dynamics problems," *Journal of Computational Physics,* vol. 449, p. 110788, 2022.

[4] P. J. Roache, "Quantification of uncertainty in computational fluid dynamics," *Annual review of fluid Mechanics,* vol. 29, no. 1, pp. 123-160, 1997.

[5] W. L. Oberkampf and T. G. Trucano, "Verification and validation in computational fluid dynamics," *Progress in aerospace sciences,* vol. 38, no. 3, pp. 209-272, 2002.

[6] H. Yu *et al.*, "Inlet and Outlet Boundary Conditions and Uncertainty Quantification in Volumetric Lattice Boltzmann Method for Image-Based Computational Hemodynamics," *Fluids,* vol. 7, no. 1, p. 30, 2022.

[7] H. Yu *et al.*, "A new noninvasive and patient-specific hemodynamic index for assessing the severity of renal arterial stenosis," *International Journal for Numerical Methods in Biomedical Engineering,* vol. e3611, no. PMID: 35509229, pp. 1-42, 2022.

[8] H. Yu *et al.*, "Mass-conserved volumetric lattice Boltzmann method for complex flows with willfully moving boundaries," *Physical Review E,* vol. 89, no. 6, p. 063304, 2014.

[9] X. Zhang *et al.*, "Volumetric lattice Boltzmann method for wall stresses of image-based pulsatile flows," *Scientific Reports,* vol. 12, no. 1, pp. 1-15, 2022.

[10] H. Li, J. Yin, and X. Du, "Label Free Uncertainty Quantification," in *AIAA SCITECH 2022 Forum*, (AIAA SciTech Forum: American Institute of Aeronautics and Astronautics, 2021.

[11] G. Calzolari and W. Liu, "Deep learning to replace, improve, or aid CFD analysis in built environment applications: A review," *Building and Environment,* vol. 206, p. 108315, 2021.

[12] H. Wu, Z. Zhu, and X. Du, "System Reliability Analysis With Autocorrelated Kriging Predictions," *Journal of Mechanical Design,* vol. 142, no. 10, 2020, doi: 10.1115/1.4046648.

[13]J. Yin and X. Du, "Uncertainty Quantification by Convolutional Neural Network Gaussian Process Regression with Image and Numerical Data," in *AIAA SCITECH 2022 Forum*, 2022, p. 1100.

[14] J. Yin and X. Du, "High-Dimensional Reliability Method Accounting for Important and Unimportant Input Variables," *Journal of Mechanical Design,* vol. 144, no. 4, 2021, doi: 10.1115/1.4051982.

[15] Z. Hu and X. Du, "Efficient reliability-based design with second order approximations," *Engineering Optimization,* vol. 51, no. 1, pp. 101-119, 2019/01/02 2019, doi: 10.1080/0305215X.2018.1440292.

[16] L. Meng, X. Du, B. McWilliams, and J. Zhang, "Probabilistic Feasibility Design of a Laser Powder Bed Fusion Process Using Integrated First-Order Reliability and Monte Carlo Methods," *Journal of Manufacturing Science and Engineering,* vol. 143, no. 9, 2021, doi: 10.1115/1.4050544.

[17] H. Li and X. Du, "A Bayesian Approach to Recovering Missing Component Dependence for System Reliability Prediction via Synergy Between Physics and Data," in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 2021, vol. 85390: American Society of Mechanical Engineers, p. V03BT03A009.
[18] H. Li and X. Du, "Recovering Missing Component Dependence for System Reliability Prediction via Synergy Between Physics and Data," *Journal of Mechanical Design,* vol. 144, no. 4, 2022.