WearGP: A computationally efficient machine learning framework for local erosive wear predictions via nodal Gaussian processes

Anh Tran[a,b], John M. Furlan[b], Krishnan V. Pagalthivarthi[b], Robert J. Visintainer[b], Tim Wildey[c], Yan Wang[a,*]

[a]*Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332*
[b]*GIW Industries Inc., Grovetown, GA 30813*
[c]*Optimization and Uncertainty Quantification Department, Sandia National Labs, Albuquerque, NM*

Abstract

Computational fluid dynamics (CFD)-based wear predictions are computationally expensive to evaluate, even with a high-performance computing infrastructure. Thus, it is difficult to provide accurate local wear predictions in a timely manner. Data-driven approaches provide a more computationally efficient way to approximate the CFD wear predictions without running the actual CFD wear models. In this paper, a machine learning (ML) approach, termed WearGP, is presented to approximate the 3D local wear predictions, using numerical wear predictions from steady-state CFD simulations as training and testing datasets. The proposed framework is built on Gaussian process (GP) and utilized to predict wear in a much shorter time. The WearGP framework can be segmented into three stages. At the first stage, the training dataset is built by using a number of CFD simulations in the order of $\mathcal{O}(10^2)$. At the second stage, the data cleansing and data mining processes are performed, where the nodal wear solutions are extracted from the solution database to build a training dataset. At the third stage, the wear predictions are made, using trained GP models. Two CFD case studies including 3D slurry pump impeller and casing are used to demonstrate the WearGP framework, in which 144 training and 40 testing data points are used to train and test the proposed method, respectively. The numerical accuracy, computational effiency and effectiveness between the WearGP framework and CFD wear model for both slurry pump impellers and casings are compared. It is shown that the WearGP framework can achieve highly accurate results that are comparable with the CFD results, with a relatively small size training dataset, with a computational time reduction on the order of 10^5 to 10^6.

Keywords: wear, machine learning, Gaussian process, multiphase computational fluid dynamics, slurry pump

---

*Corresponding author. Email: yan.wang@me.gatech.edu

1. Introduction

Wear predictions using computational fluid dynamics (CFD) simulation are accurate and reliable, but computationally expensive. Even if parallelism is enabled on multi-core and multi-processor high-performance computing (HPC) systems, it is difficult to obtain accurate 3D wear predictions within a short period of time. The limitation in parallelism is rooted in the bounded upper limit of computational speedup from Amdahl's law, which is approached asymptotically as the number of processors increases [1]. While the local wear predictions with details are very useful, its expensive computational cost prohibits the usage among related personnel who do not have assess to HPC systems. Thus, parallelism does not provide a satisfactory solution in bringing down the computational cost for intensive computing applications, such as wear predictions using CFD. Reducing the computational cost of CFD wear predictions, from hours to seconds, would enable the results to be used in a much more expansive manner, from research and design for research engineers and design engineers, to troubleshooting and pump selections for sales engineers. Furthermore, in mining applications, allowing wear predictions within seconds enables field engineers to change operating conditions appropriately according to the mining operations, such that the life cycle of different slurry pump components can match, and maintenance can be scheduled all at once to replace most of the old pump components. Predicting life time and avoiding sudden shutdowns in mining applications are very important, as a few days of pipeline shutdown would result in millions dollar of loss in revenue.

Data-driven approaches, also known as machine learning (ML), have been shown to be a powerful paradigm with predictive capabilities, and does not require direct access to HPC systems. The ML approaches have been shown to be a promising candidate to replace the actual CFD predictions. They are not only fast and robust, but also accurate if trained properly. Simply speaking, data-driven approaches, or ML algorithms, allow one to build a surrogate model, or metamodel, to predict the quantity of interest (QoI) within a much shorter time.

Computational efficiency, accuracy, and effectiveness are the three major factors to determine the quality of ML algorithms. The computational efficiency measures how much computational time has been reduced. The accuracy describes how well the QoI has been approximated, for example, by measuring the difference between the predictions from CFD simulations and from the ML model. In this paper, the QoI is the predicted wear rate of different slurry pump components under specific operating conditions. The effectiveness describes how fast the ML algorithm can approximate the QoI, with respect to the size of the training dataset. Increasing the size of the training dataset typically yields a smaller approximation error. Highly effective algorithms are

2

associated with high convergence rate, where the approximation error decays with respect to the number of training data points [2].

Neural networks, including traditional artificial neural networks (ANN) and more recent convolutional neural networks, have been used to approximate CFD solutions. While the CFD community has been integrating ML to accelerate its solvers, multiphase flow, such as slurry particulate flow at high concentrations, is more complex and poses a challenging problem. The complexity is rooted in the wear mechanism, and the interaction between the carrier fluid and the solid particles. Furthermore, randomness also plays an important role in capturing erosive wear. For example, a particle population can only be captured by a statistical distribution, which in turns introduces stochasticity into wear predictions.

In this paper, a Gaussian process (GP) based framework, called WearGP, is proposed to approximate wear predictions from CFD to improve the computational efficiency of traditional simulation approaches. The wear prediction dataset is decomposed according to the wall nodes of the mesh. At each wall node, a GP is constructed, where the GP inputs describe the operating conditions, whereas the GP output is the nodal wear solution at that particular node.

The advantage of the WearGP framework is three-fold. First, we show that the the WearGP framework is highly effective in the sense that fewer CFD simulations in the order of $\mathcal{O}(10^2)$ are needed to train, compared to $\mathcal{O}(10^3) - \mathcal{O}(10^6)$ training points required in deep learning algorithms. Second, the computational time is reduced five to six orders of magnitude, enabling the wear predictions to be made within seconds, compared to hours on a HPC platform. Third, the wear predictions by WearGP framework and the CFD simulations are comparable, both quantitatively and qualitatively. Our comparison demonstrates the high accuracy of the proposed WearGP framework.

The state-of-the-art ML methods predict wear mainly based on temperature, particle velocity, particle impingement angle, wall shear stress, turbulent kinetic energy, concentration, and experimental erosion rate. However, these methods do not predict wear in local regions with respect to geometric models of components. In other words, the state-of-the-art ML methods mainly consider wear predictions as a global QoI, as opposed to a local QoI. Compared to the global QoI, the local QoI level contains much more information. Our work differs from other related work in literature by predicting wear at the local QoI level. In this paper, the geometric models of the object of interest are slurry pump impellers and casings. The goal of the WearGP framework is to accurately approximate the wear predictions from CFD simulations, at a significantly lower cost in terms of computational time. To the best of the authors' knowledge, this is the first work attempting

3

to predict local wear with respect to a geometric model, such that the wear predictions can be visualized in 3D, based on an ML algorithm. Compared to other methods, our proposed WearGP framework provides an accurate and computationally efficient approach to predict wear solutions of the same quality as CFD, including 3D visualization.

In the rest of this paper, Section 2 provides the literature review on ML approaches for CFD and wear problems. Section 3 describes the WearGP formulation and its potential applications. Section 4.1 describes the impeller CFD wear model and compares the CFD model with the WearGP framework. Section 5.1 describes the casing CFD wear model and compares the CFD model with the WearGP framework. Section 6 discusses the WearGP framework advantages and disadvantages. Section 7 concludes and summarizes the paper.

2. Literature review

In this section, we review existing work that combines CFD and ML methods. Applications of ML and CFD wear predictions are discussed, including the use of deep learning methods to approximate unsteady CFD solutions.

2.1. Data-driven CFD solutions and wear predictions

Deep learning approaches, such as convolutional neural networks, have been one of the most common choices in using ML to solve CFD problems. Guo et al. [3] successfully deployed convolutional neural networks to approximate steady flow solution. Tompson et al. [4] developed a deep neural network architecture to approximate the CFD solution for Eulerian fluid simulation. Miyanawala and Jaiman [5] also deployed deep learning to approximate the unsteady fluid forces for different bluff body shapes at low Reynolds number. Chu and Thuerey [6] proposed a deep learning architecture that yields high-resolution space-time solutions for smoke simulations.

Likewise, ANN is one of the most common choice in data-driven wear predictions. Friedrich et al. [7] [8] first applied ANN to predict wear test measurements for polymers. Danaher et al. [9] included temperature to model wear in high temperature erosion Ni-base alloys, also using the ANN. Suresh et al. [10] adopted ANN approach to model wear on polymeric materials. Shamshirband et al. [11] also employed ANN approach to predict wear in a $90°$ elbow, trained by CFD simulations, where the QoIs (total erosion and maximum erosion rate) are global variables. Qu and Zuo [12] applied support vector machine to classify the wear degree of slurry pump impellers. Pandya et al. [13] employed ANN to predict solid particle erosion based on ANSYS Fluent.

4

Dai et al. [14] used GP for uncertainty quantification in erosive wear predictions following data mining process. However, the work described above limits its scope in wear predictions as a global QoI, rather than a local QoI.

2.2. Gaussian process

Brochu et al. [15] and Shahriari et al. [16] provide a comprehensive review about GP and Bayesian optimization based on the GP surrogate model. Here we adopt the notation from Shahriari et al. [16], and summarize the GP formulation that is used in this paper as the ML solver.

Assume that f is a function of $\boldsymbol{x}$, where $\boldsymbol{x} \in \mathcal{X}$ is the d-dimensional input. A $GP(\mu_0, k)$ is a nonparametric model over functions f, which is fully characterized by the prior mean functions $\mu_0(x) : \mathcal{X} \mapsto \mathbb{R}$ and the positive-definite kernel, or covariance function $k : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$. In GP regression, it is assumed that $\boldsymbol{f} = f_{1:n}$ is jointly Gaussian, and the observation y is normally distributed given f, leading to

$$\boldsymbol{f}|\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{m}, \boldsymbol{K}), \tag{1}$$

$$\boldsymbol{y}|\boldsymbol{f}, \sigma^2 \sim \mathcal{N}(\boldsymbol{f}, \sigma^2 \boldsymbol{I}), \tag{2}$$

where $m_i := \mu(\boldsymbol{x}_i)$, and $K_{i,j} := k(\boldsymbol{x}_i, \boldsymbol{x}_j)$. Equation 1 describes the prior distribution induced by the GP.

The covariance kernel k is a choice of modeling covariance between inputs. One of the most widely used kernel is the squared exponential kernel, where f is implicitly assumed to be smooth. The covariance kernel is

$$\boldsymbol{K}_{i,j} = k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \theta_0^2 \exp\left(-\frac{r^2}{2}\right), \tag{3}$$

where $r^2 = (\boldsymbol{x} - \boldsymbol{x}')\boldsymbol{\Gamma}(\boldsymbol{x} - \boldsymbol{x}')$, and $\boldsymbol{\Gamma}$ is a diagonal matrix of d squared length scale θ_i.

The hyper-parameters θ is determined by maximum likelihood estimation, where the log marginal likelihood is described as

$$\log p(\boldsymbol{y}|\boldsymbol{x}_{1:n}, \theta) = -\frac{1}{2}(\boldsymbol{y} - \boldsymbol{m}_\theta^T)(\boldsymbol{K}^\theta + \sigma^2 \boldsymbol{I})^{-1}(\boldsymbol{y} - \boldsymbol{m}_\theta) - \frac{1}{2}\log|\boldsymbol{K}^\theta + \sigma^2 \boldsymbol{I}| - \frac{n}{2}\log(2\pi). \tag{4}$$

Optimizing the likelihood functions yields the hyper-parameters θ for the GP. Because the optimization process involves computing the inverse of the covariance matrix, the algorithmic complexity is at $\mathcal{O}(n^3)$.

Let the dataset $\mathcal{D} = (\boldsymbol{x}_i, y_i)_{i=1}^n$ denote a collection of n noisy observations and $\boldsymbol{x}$ denote an arbitrary input of dimension d. Under the formulation of GP, given the dataset $\mathcal{D}_n$, the prediction for an unknown arbitrary

test point is characterized by the posterior Gaussian distribution, which can be described by the posterior mean and posterior variance functions, respectively as

$$\mu_n(\boldsymbol{x}) = \mu_0(\boldsymbol{x}) + \boldsymbol{k}(\boldsymbol{x})^T(\boldsymbol{K} + \sigma^2\boldsymbol{I})^{-1}(\boldsymbol{y} - \boldsymbol{m}), \tag{5}$$

and

$$\sigma_n^2 = k(\boldsymbol{x}, x) - \boldsymbol{k}(\boldsymbol{x})^T(\boldsymbol{K} + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{k}(\boldsymbol{x}), \tag{6}$$

where $\boldsymbol{k}(\boldsymbol{x})$ is the covariance vector between the test point $\boldsymbol{x}$ and $\boldsymbol{x}_{1:n}$. Compared to other types of surrogate models and machine learning algorithms, the GP framework has several advantages. First, the framework is adaptive, which does not restrict the locations of the sampling data points. Second, it provides a way to quantify uncertainty in terms of posterior mean and posterior variance.

3. WearGP framework

In this section, we describe the main idea behind the WearGP framework, which is a data-driven method that approximates the steady state CFD wear solution, and its implementation, flexibility, and scalability in industrial environments.

3.1. Methodology

Figure 1 illustrates the process of wear predictions. On one hand, the slurry behaviors are determined by the carrier fluid, which can be modeled by density, temperature, viscosity, and the solid particles, which can be characterized by average solid concentration, particle size distribution (PSD), hardness, sphericity, shape factor, etc. With appropriate boundary conditions, the slurry particulate flow fields can be solved by multiphase CFD simulations to obtain the velocity field, pressure field, wall shear stress, particle impingement angle, particle impingement velocity, and local solid concentration. The collection of these slurry parameters, combining both carrier fluid and solid particles, form an input $(\boldsymbol{x}_{\text{solids}}, \boldsymbol{x}_{\text{liquids}}, \boldsymbol{x}_{\text{BCs}})$. The materials behaviors, on the other hand, describe the wear response of materials, under the imposed shear stress by the slurry. From the process-structure-property perspective, the material behaviors can be captured by another set of parameters, including a number of processing parameters, microstructure descriptors, and constitutive material models. The collection of the parameters, which describe the constitutive model of materials, form an input $\boldsymbol{x}_{\text{materials}}$.

The material characterization is performed experimentally, where the wear response of materials is measured. A constitutive material wear model is then constructed. Using the predicted flow fields for both solid and liquid phases, the wear response of the materials under some specific boundary conditions can be predicted.
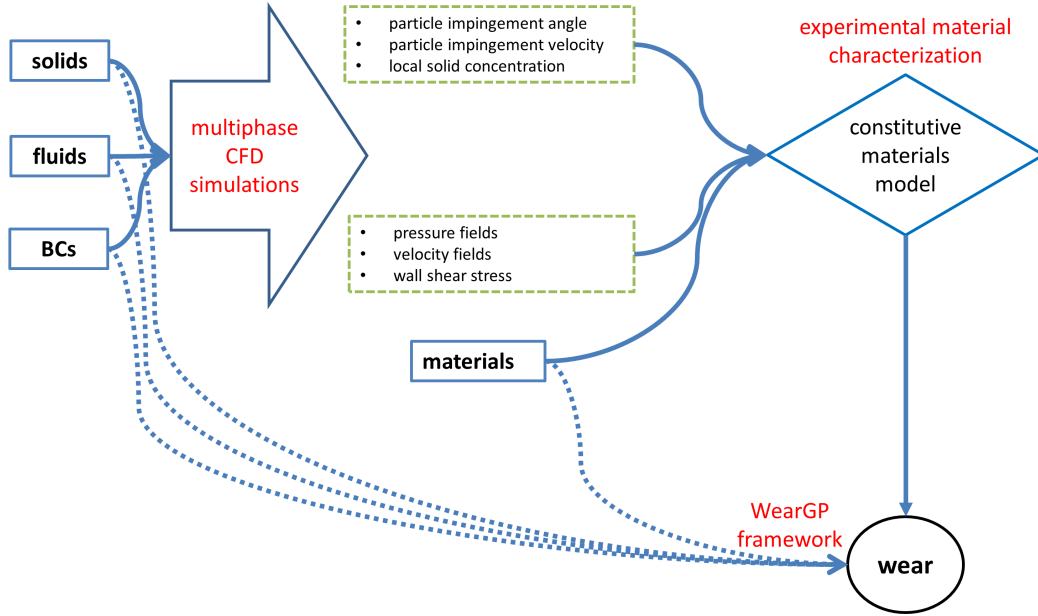
Figure 1: Schematic of wear prediction process involving multiphase CFD simulations.

Wear response of the materials, denoted as $\dot{W}$, can be considered as a function of all of the above-mentioned parameters that model both slurry and materials. That is, the Equation describing wear response

$$\dot{W} = \dot{W}(\boldsymbol{x}_{\text{solids}}, \ \boldsymbol{x}_{\text{fluid}}, \ \boldsymbol{x}_{\text{BCs}}, \ \boldsymbol{x}_{\text{materials}}) \tag{7}$$

holds for every location on the walls of the geometric models within the CFD simulations. Given that the CFD simulations are non-stochastic, i.e. a set of inputs corresponding to a deterministic wear prediction, the WearGP framework can be constructed as a shortcut to predict the wear directly from the set of input parameters $(\boldsymbol{x}_{\text{solids}}, \ \boldsymbol{x}_{\text{fluid}}, \ \boldsymbol{x}_{\text{BCs}}, \ \boldsymbol{x}_{\text{materials}})$ without running the CFD wear models. Consequently, the main idea of the WearGP framework can be equivalently interpreted as approximating the CFD wear predictions with varying input parameters.

Since equation 7 holds locally, we seek to approximate the nodal wear prediction $\dot{W}$ at every node of the mesh using GP. The GP inputs describing solid properties, fluid properties, boundary conditions, and materials properties are in a vector form $(\boldsymbol{x}_{\text{solids}}, \ \boldsymbol{x}_{\text{fluid}}, \ \boldsymbol{x}_{\text{BCs}}, \ \boldsymbol{x}_{\text{materials}})$. The GP output is the nodal wear prediction $\dot{W}$.

The overall workflow of the WearGP framework is illustrated in Figure 2. In order to perform data-mining and training procedure for the proposed method, an initial database is built to train the ML framework, using the CFD wear model as the training model. The mesh settings remain unchanged for a particular object.
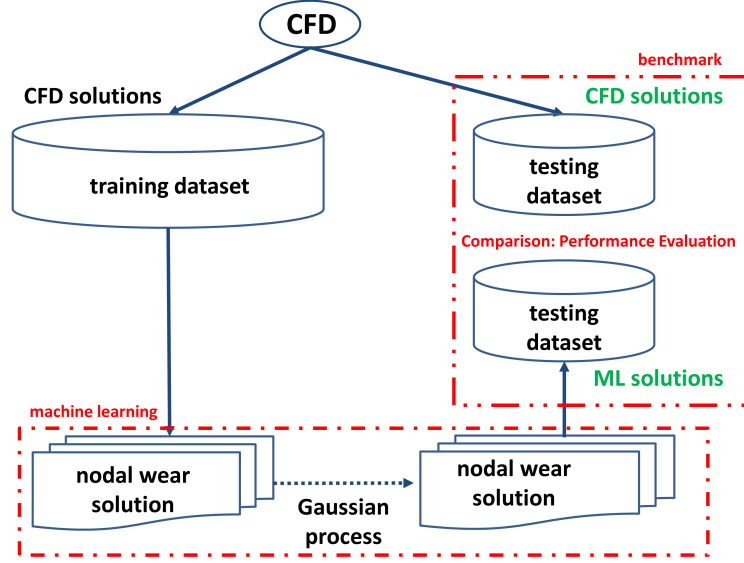
7

Figure 2: Workflow of WearGP framework for a specific pump component.

The CFD simulations are evaluated repeatedly to construct the training and testing dataset for different input parameters. After the training dataset is built, the nodal wear solution is extracted from the training dataset using a data mining approach.

To predict the wear solutions trained using the WearGP framework, at every node of the mesh, the trained GP is used to directly predict the wear rate $\dot{W}$ based on the input parameters ($\boldsymbol{x}_{\text{solids}}$, $\boldsymbol{x}_{\text{fluid}}$, $\boldsymbol{x}_{\text{BCs}}$, $\boldsymbol{x}_{\text{materials}}$). The nodal wear predictions are then assembled together to mimic a 3D CFD wear predictions, which can be visualized by other post-processing packages. To evaluate the performance of the WearGP framework, the WearGP wear predictions are compared with the CFD wear predictions.

3.2. Slurry pump applications

In this study, the geometric models of the slurry pump components, such as the pump impeller and pump casing, are fixed, and the pump operating conditions vary. The slurry pump operating conditions are mainly specified by the head H, flow rate Q, PSD, and concentration by volume C_v. The $\%BEPQ$ is defined as the ratio between the operating flow rate Q and the flow rate at the best efficiency point, Q_{BEP}, at a given head H and pump speed N. Mathematically, $\%BEPQ = \dfrac{Q}{Q_{\text{BEP}}}$. Q_{BEP} is a unique flow rate for a given head H and pump speed N, which can be obtained experimentally by maximizing the pump efficiency. In this case, the PSD is represented by the effective particle size d_{eff}, which is proportional to the mean particle

8

size d_{50}. Varying the head H, flow rate Q, would subsequently change the inlet and outlet velocities of the boundary conditions. Varying the mean PSD d_{50} affects wear predictions in the CFD slurry steady solutions, as well as the material constitutive models. Larger particles typically also result in higher concentration along the wall surfaces, affecting the wear predictions. Higher wear predictions are observed with increasing the particle size. Varying the concentration in volume C_v would completely change the CFD solution, where the Eulerian-Eulerian framework is utilized to solve for a steady flow field solution.

Since many input parameters are repetitive and redundant, a set of parameters is chosen to represent a pump duty condition, such that the dimensionality of the problem is minimized. Reducing the dimensionality based on physics-based knowledge is an important step to avoid the curse of dimensionality, and thus improve the performance of the WearGP framework. Two options for choosing the basis parameters are investigated in this study. The first parameter set is (H, Q, d_{50}, C_v), whereas the second set is $(H, \%BEPQ, d_{50}, C_v)$. The pump speed N, as well as the efficiency η are solved as inputs for CFD simulations based on a particular operating condition, using an in-house program, before the CFD simulations are performed. For each part of the pump, i.e. the impeller and the casing, the CFD simulation is called to evaluate the wear solution. For each wall node of the mesh, a GP is constructed, where the inputs of the GP are the pump operating conditions, and the output of the GP is the nodal wear solution corresponding to a particular set of input operating conditions.

3.3. Implementation

Essentially, the WearGP targets the node of the mesh and constructs a local GP at each node, with respect to different pump operating conditions, or more specifically boundary conditions. In practice, the WearGP is implemented and carried out in three stages. The first stage includes the data cleansing and data mining processes, which seeks to purify the output of CFD simulations. The second stage is the training, which can be massively parallelized, because the nodes of the mesh are treated independently. The advantage of nodal decomposition is two-fold. First, the computational efficiency is guaranteed, based on the nodal decomposition such that each GP is associated with a particular node, and the nodes are treated independently of each other. Second, the discontinuity and the singularity of the CFD simulations is preserved. The drawback of the nodal decomposition approach is that the number of GPs must be trained in on the same scale with the number of nodes. However, in theory, the problem can be mitigated by massively parallelizing and distributing the computation workload across a multi-core computing architecture, since the nodes are treated independently. The third stage involves the forward evaluation to predict the wear solution, including the postprocess for

visualization. The postprocess in the third stage can be considered as the reversal of data cleansing in the first stage.

In this paper, we describe two variants of WearGP framework for two main components of slurry pump: impeller and casing. They are both formulated on the nodal decomposition approach, where GP is employed to predict the wear at each node in the mesh.

4. Impeller wear model

In this section, we apply the WearGP framework on the 3D CFD impeller wear model [17] developed as an in-house wear code to assess the erosive wear rate of the slurry pump. Section 4.1 summarizes the formulation of the CFD impeller wear model. Section 4.2 compares the numerical performance, in terms of both accuracy and efficiency of the proposed WearGP framework and the CFD impeller wear model.

Table 1 shows the variables used to generate the training dataset. The flowrate Q is numerically solved using another in-house pump solver, given the other operating conditions. The training dataset is generated at the following locations according to last column of Table 1. It is noted that the $\%BEPQ$ is not always at 130% for the upper bound, because the in-house pump solver does not always converge for certain operating conditions. To circumvent the problem, the upper bound of $\%BEPQ$ is gradually decreased down to 110% until a convergence solution for pump input is obtained. Consequently, 144 locations are chosen to build the training datasets, in which one location corresponds to one CFD simulations.

Table 1: Parameters used for building training datasets.

Variable	Physical description	Unit	Lower bound	Upper bound	Used Values
H	pump head	m	35	65	35,50,65
$\%BEPQ$	percentage of BEP flow rate	%	70	130	70,100,130
d_{50}	mean particle size	μm	150	600	150,300,450,600
C_v	concentration by volume	%	10	40	10,20,30,40

Table 2 describes the operating conditions of testing datasets. In this table, four operating variables, H, $\%BEPQ$, d_{50}, C_v, are varied independently to generate the testing dataset. The upper bound of $\%BEPQ$ is set at 110%, to avoid any divergence from the in-house pump solver. Both Tables 1 and 2 are used to train and test the CFD impeller wear model and CFD casing wear model, respectively.

10

Table 2: 40 testing cases and the corresponding operating conditions.

Testing case	$H(m)$	$BEPQ(\%)$	$d_{50}(\mu m)$	$C_v(\%)$
1	61.0238	100.7292	166.3271	17.1453
2	59.1267	98.1905	315.0980	32.1904
3	44.7113	108.3772	240.2366	20.7519
4	48.4715	91.1032	559.3353	30.8105
5	61.5799	88.8346	164.6182	29.2931
6	60.6776	93.6561	409.1707	14.3419
7	55.3361	90.3236	543.6216	30.1760
8	49.2122	79.0786	244.4687	33.3295
9	40.8413	96.6661	176.7400	18.0289
10	37.3730	84.1489	194.7891	23.2157
11	52.6503	89.6869	469.7961	35.7262
12	44.7080	95.1341	321.9043	38.8840
13	40.7969	79.9512	234.7505	12.2550
14	40.7849	84.3104	274.8115	34.8312
15	52.0627	86.1245	234.0740	38.4080
16	46.5643	106.2776	291.9244	36.7283
17	42.4220	91.7494	391.4051	36.8881
18	57.2903	89.4059	516.8324	36.8432
19	57.7757	81.3670	404.4440	14.4302
20	63.5732	77.5094	293.5427	18.2543
21	46.8153	105.3642	300.4395	21.8259
22	38.6464	76.4270	252.9711	29.3955
23	51.5555	94.2633	412.6286	19.7275
24	41.3691	81.1006	507.0509	30.9224
25	60.6506	105.4266	287.0631	31.1752
26	44.4207	99.7222	151.2729	29.2686
27	55.1794	76.4686	489.6172	21.6258
28	45.4443	80.8128	596.0691	10.3929
29	41.9834	81.6694	394.4312	21.1014
30	39.9178	105.8456	407.9843	36.4604
31	49.8985	86.2312	171.2106	28.1420
32	51.7241	93.4960	235.8800	35.4772
33	55.7365	79.2394	469.7201	39.0805
34	63.8825	77.7157	418.2133	39.7347
35	54.2585	106.0398	417.2460	29.2489
36	50.3763	85.9893	488.4799	32.5342
37	46.0812	90.4854	412.1844	24.4333
38	37.6830	87.7900	436.1273	22.7111
39	42.0202	96.4432	393.2446	16.7365
40	35.6735	94.1172	367.0682	11.8249

4.1. CFD impeller wear model

The CFD impeller wear used to predict wear in centrifugal slurry pump impellers is described in detail by Pagalthirvarthi et al [17]. The CFD impeller wear model is summarized here for the sake of completeness and convenience of the readers. To capture the multi-size solid-liquid particulate flow, the Eulerian-Eulerian framework is utilized, where the PSD is discretized into finitely many species, each with a different particle size and individual species concentration and shape factor. The continuity and momentum equations of the mixture and different species are formulated using the volume and time-averaged governing equations. In this study, a mono-size solid PSD is used to approximate the multi-size PSD in the CFD wear predictions, where the mono-size effective particle size is taken as $d_{\text{eff}} = 1.65 d_{50}$ [18]. Figure 3 describes the computational domain of the CFD impeller wear model in the Cartesian coordinate system. Due to the angular symmetry, the governing equations are rewritten in the polar coordinate system. The Spalart-Allmaras model [19] is used to describe the turbulence.
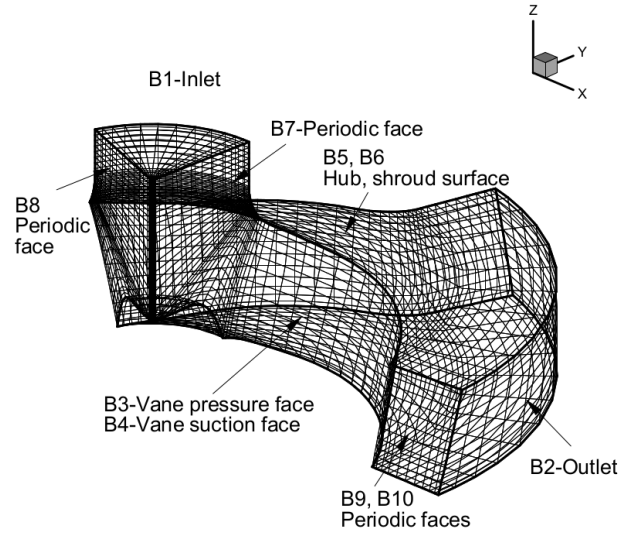
Figure 3: Three-dimensional pump impeller with mesh and its boundary conditions.

Figure 3 also shows the boundary conditions of the CFD impeller wear model. The inlet velocity boundary condition is applied at the inlet surface B1. The stress free boundary condition is applied at the outlet surface B2. The blade surfaces B3 and B4, the hub surface B5, and the shroud surface B6 are treated as wall, where Spalding wall functions [20] are utilized. On the surfaces B7, B8, B9, and B10, periodic boundary conditions are applied.

The solver for CFD impeller wear model is constructed based on the finite element problem formulation,

where the streamwise upwind Petrov Galerkin [21, 22] is utilized. The nonlinear governing equations are then solved iteratively using Newton's method with under-relaxation factors. The CFD solver is implemented and parallelized using the Intel PARADISO solver [23].

A constitutive model for wear prediction is formed based on experimental measurements [24]. The total wear is decomposed as the sum of the sliding wear and impact wear of all the species. For each species, the impact wear and sliding wear are computed using empirically determined wear coefficients in concert with the CFD-predicted flow field; particle concentration, mixture density, velocity magnitude, tangential velocity, shear stress, impingement angle, and particle size [25] are predicted by the CFD, and, together with the wear coefficients, are used to predict the wear rate locally over the wetted surfaces.

4.2. Comparison between WearGP and CFD impeller wear model

To evaluate the performance of the WearGP framework, we compare its prediction with the 3D CFD wear simulation for slurry pump impeller. 144 CFD simulations are performed to build the training dataset, and 40 CFD simulations are used to build a testing dataset. The testing operating conditions are all sampled using Monte Carlo algorithm with uniform distributions.

GIW pump impeller 0509X LSA-36 5525D 7671D-A1 is chosen as a case study in this example. The design impeller diameter is 0.9144 m. The impeller has 3 vanes, with impeller eye diameter of 0.3112 m. The hub entrance radius is 0.2286 m. Figure 4 shows the actual slurry pump impeller in the field.
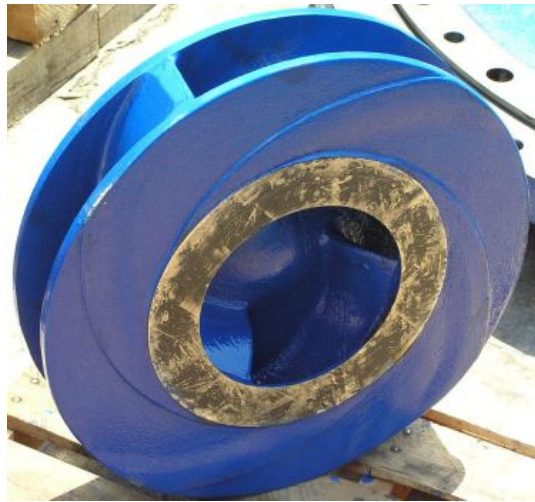
Figure 4: GIW impeller 0509X LSA-36 5525D 7671D-A1 in the field.

For CFD simulation, a structured mesh of 10478 nodes and 8784 elements is used to discretize the com-

putational domain. Among those, the 1664 wall nodes are used to evaluate the wear performance of the
impeller.

In Figures 5 to 8, the wear prediction results are compared for the back shroud of the impeller, front
shroud of the impeller, pressure side of the impeller vane, and suction side of the impeller vane, respectively,
for testing case 16, where $H, \%BEPQ, d_{50}, C_v$ are $46.5643, 106.2776, 291.9244, 36.7283$, respectively. The left
subfigure is the wear predictions result based on the CFD simulation, the middle one is based on the WearGP
implementation using (H, Q, d_{50}, C_v) basis, and the right one is based on the WearGP implementation using
$(H, \%BEPQ, d_{50}, C_v)$ basis. They all show an excellent agreement between each other, not only in predicting
the wear magnitude, but also the location of wear hotspots, where high wear is concentrated in a small area.

(a) CFD-3DImpeller.   (b) WearGP (H, Q, d_{50}, C_v).   (c) WearGP $(H, \%BEPQ, d_{50}, C_v)$.
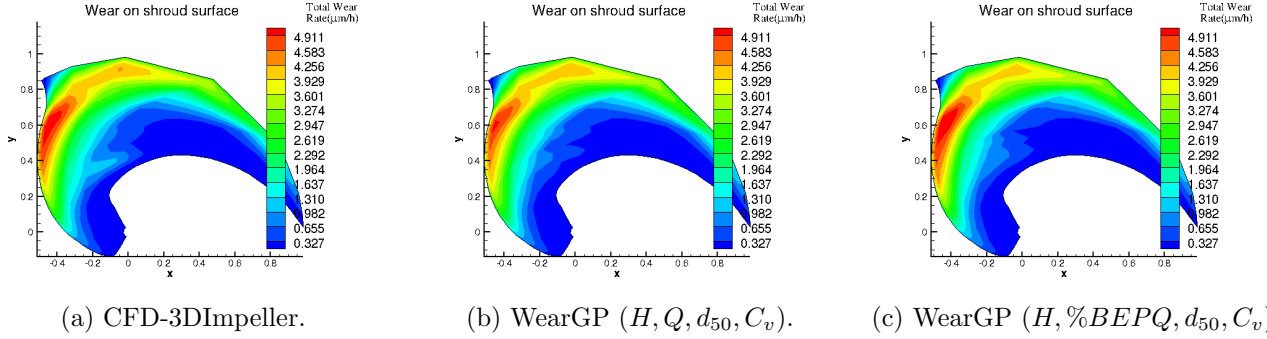
Figure 5: Impeller Case 16: $(H, \%BEPQ, d_{50}, C_v) = (46.5643, 106.2776, 291.9244, 36.7283)$: Comparison of wear predictions on
back shroud surface.

(a) CFD-3DImpeller.   (b) WearGP (H, Q, d_{50}, C_v).   (c) WearGP $(H, \%BEPQ, d_{50}, C_v)$.
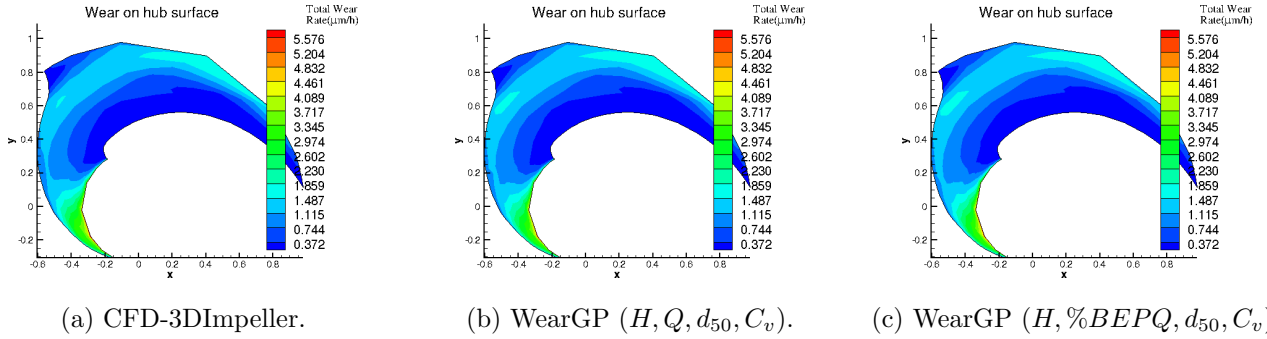
Figure 6: Impeller Case 16: $(H, \%BEPQ, d_{50}, C_v) = (46.5643, 106.2776, 291.9244, 36.7283)$: Comparison of wear predictions on
front shroud surface.

Figures 9 to 12 show the wear predictions for the back shroud, front shroud, pressure side, and suction
side of the impeller, respectively, for the testing case 7/40 of the CFD impeller wear model, where the testing
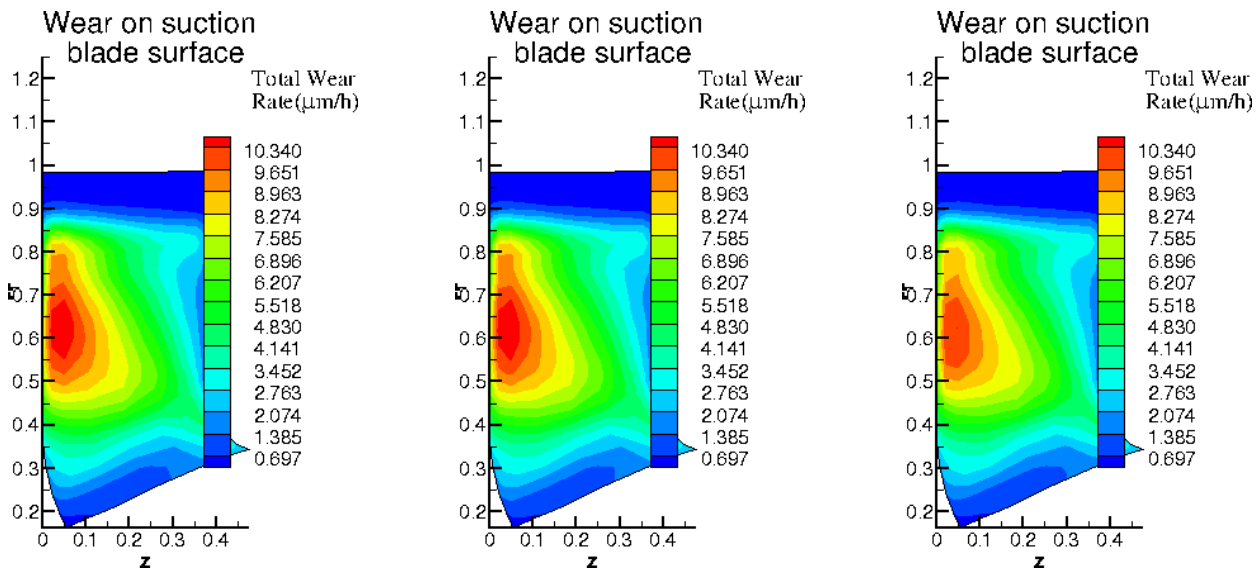
14
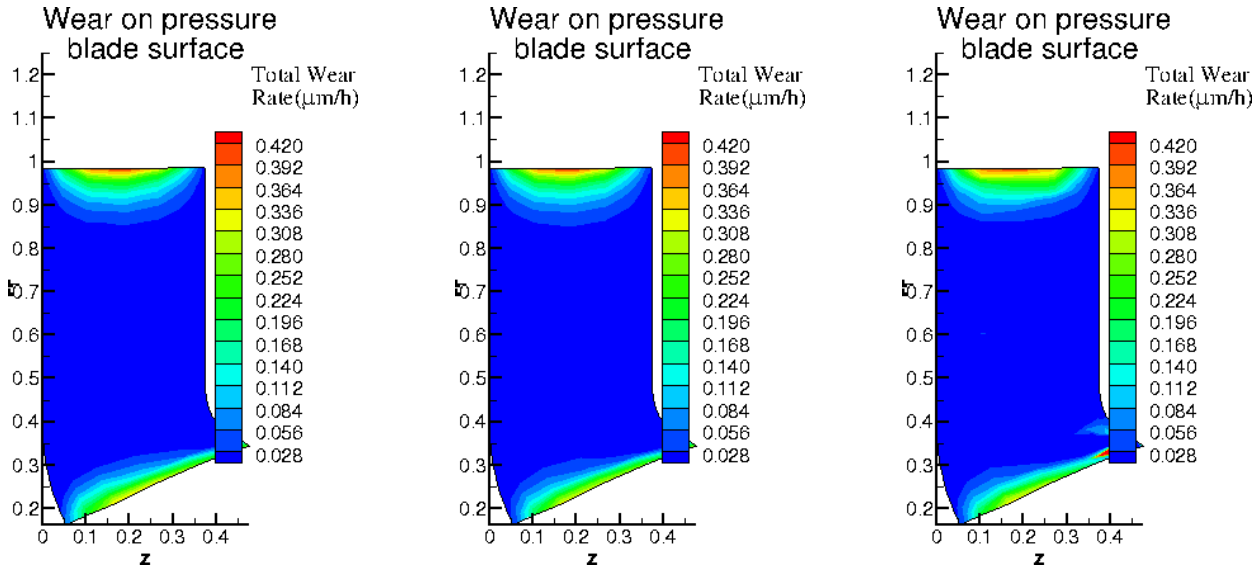
(a) CFD-3DImpeller.

(b) WearGP (H, Q, d_{50}, C_v).

(c) WearGP $(H, \%BEPQ, d_{50}, C_v)$.

Figure 7: Impeller Case 16: $(H, \%BEPQ, d_{50}, C_v) = (46.5643, 106.2776, 291.9244, 36.7283)$: Comparison of wear predictions on pressure surface.
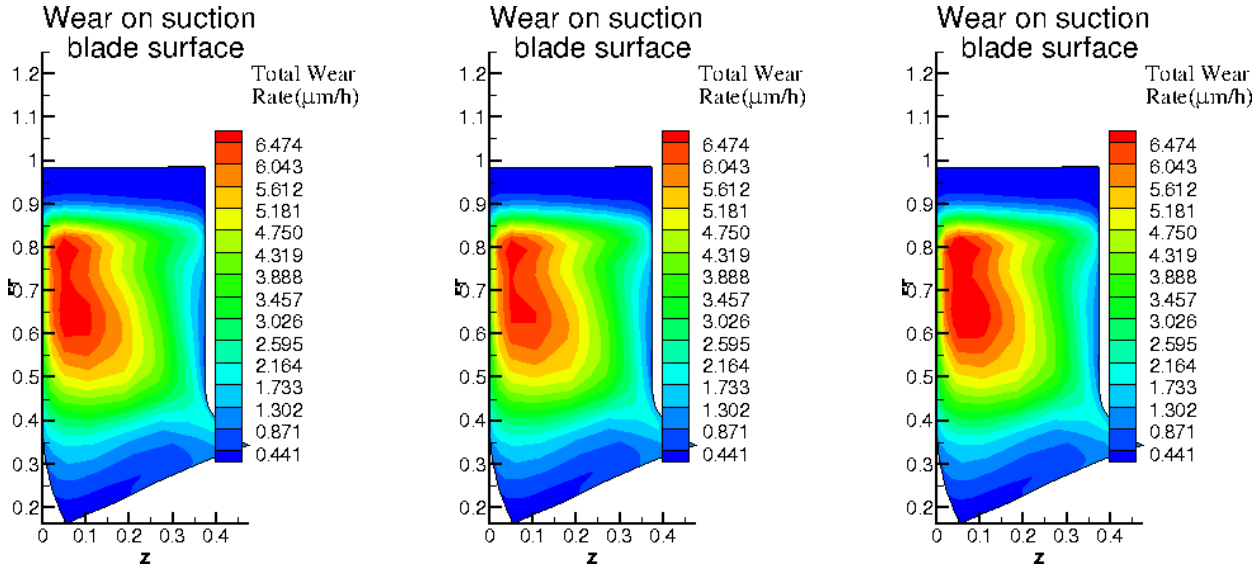
(a) CFD-3DImpeller.

(b) WearGP (H, Q, d_{50}, C_v).

(c) WearGP $(H, \%BEPQ, d_{50}, C_v)$.

Figure 8: Impeller Case 16: $(H, \%BEPQ, d_{50}, C_v) = (46.5643, 106.2776, 291.9244, 36.7283)$: Comparison of wear predictions on suction surface.

15

operating conditions are $H, \%BEPQ, d_{50}, C_v$ are $55.3361, 90.3236, 543.6216, 30.1760$, respectively. Compared to the testing case 16/40, as in Figure 5 to Figure 8, the predicted wear magnitudes are significantly higher. Yet, similar to previous case, case 7 shows a very good agreement between the WearGP framework and the CFD simulation.

(a) CFD-3DImpeller. (b) WearGP (H, Q, d_{50}, C_v). (c) WearGP $(H, \%BEPQ, d_{50}, C_v)$.
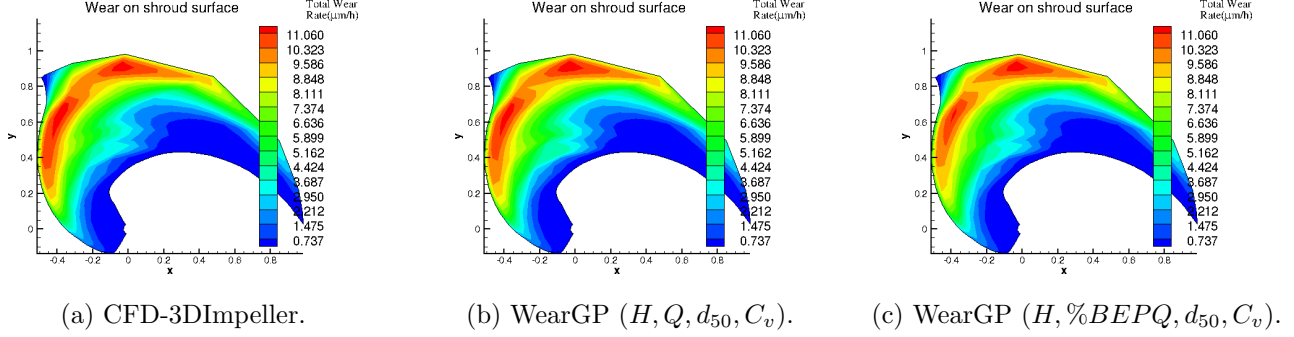
Figure 9: Impeller Case 7: $(H, \%BEPQ, d_{50}, C_v) = (55.3361, 90.3236, 543.6216, 30.1760)$: Back shroud surface comparison.

(a) CFD-3DImpeller. (b) WearGP (H, Q, d_{50}, C_v). (c) WearGP $(H, \%BEPQ, d_{50}, C_v)$.
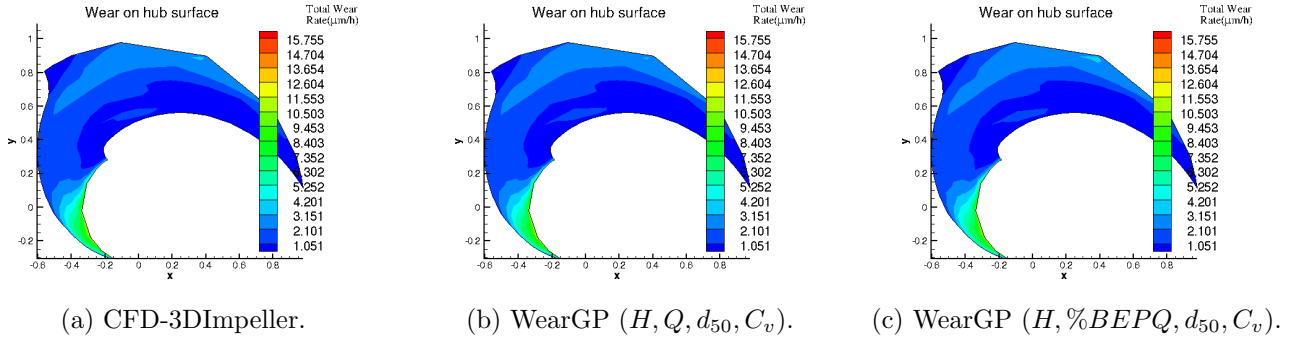
Figure 10: Impeller Case 7: $(H, \%BEPQ, d_{50}, C_v) = (55.3361, 90.3236, 543.6216, 30.1760)$: Hub surface comparison.

To rigorously quantify and compare the error between two methods, the mean square error (MSE) is used to measure and evaluate the numerical prediction between two variants of WearGP framework. The MSE for each node is calculated as

$$MSE = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \left(\dot{W}_T^{(\text{WearGP})} - \dot{W}_T^{(\text{CFD})} \right)^2, \tag{8}$$

where N_{test} is the number of testing cases, which is 40 in this study. $\dot{W}_T^{(\text{WearGP})}$ is the total wear predicted using the WearGP framework, and $\dot{W}_T^{(\text{CFD})}$ is the wear predicted using the CFD impeller wear model.

Figures 13 to 16 plot and compare the MSE between the two WearGP implementation variants on the back shroud, front shroud, pressure vane side, suction vane side of the slurry pump impeller, respectively, where

16
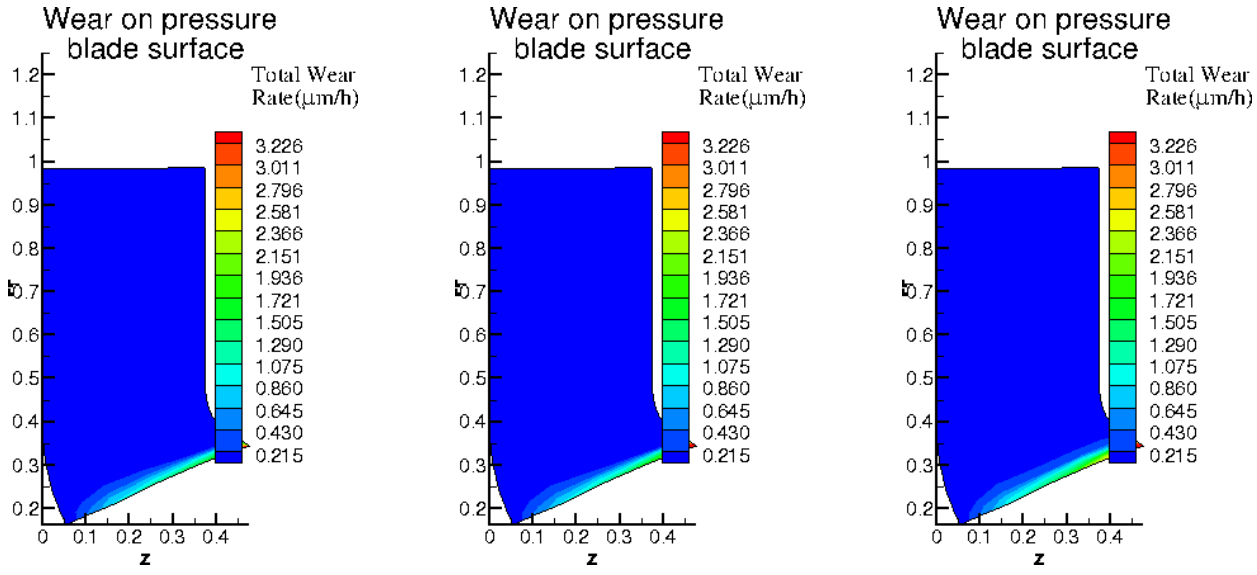
(a) CFD-3DImpeller. (b) WearGP (H, Q, d_{50}, C_v). (c) WearGP $(H, \%BEPQ, d_{50}, C_v)$.

Figure 11: Impeller Case 7: $(H, \%BEPQ, d_{50}, C_v) = (55.3361, 90.3236, 543.6216, 30.1760)$: Pressure surface comparison.

(a) CFD-3DImpeller. (b) WearGP (H, Q, d_{50}, C_v). (c) WearGP $(H, \%BEPQ, d_{50}, C_v)$.

Figure 12: Impeller Case 7: $(H, \%BEPQ, d_{50}, C_v) = (55.3361, 90.3236, 543.6216, 30.1760)$: Suction surface comparison.

17

the color bars are adjusted to the same scale. The comparison figures show that for WearGP framework, the MSE using (H, Q, d_{50}, C_v) as a basis is smaller than the MSE using the $(H, \%BEPQ, d_{50}, C_v)$.
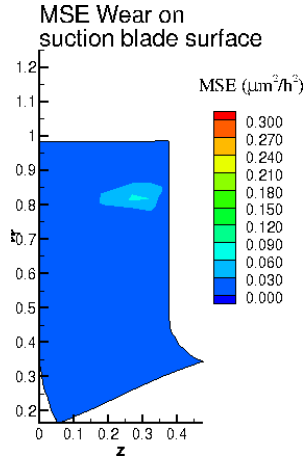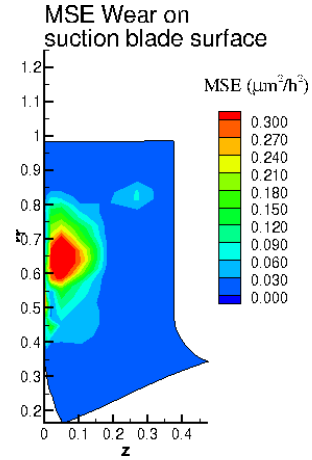

(a) WearGP (H, Q, d_{50}, C_v).

(b) WearGP $(H, \%BEPQ, d_{50}, C_v)$.
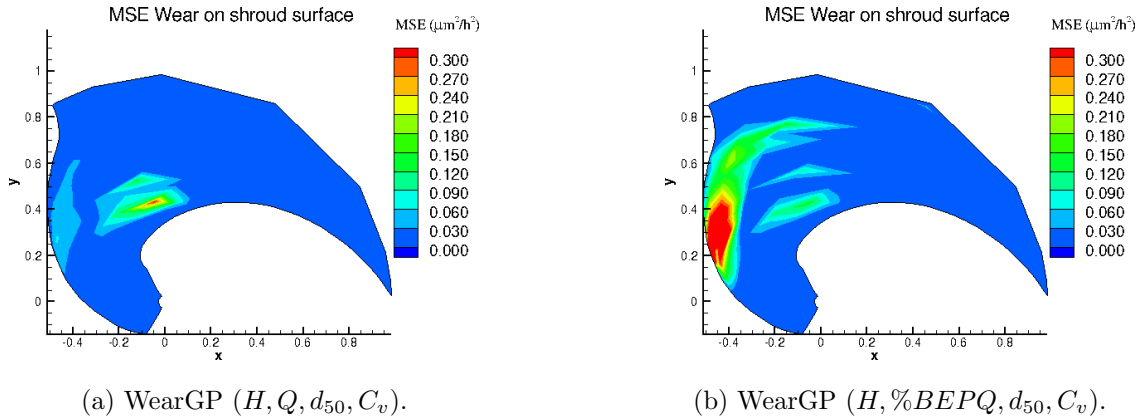
Figure 13: Comparison of MSE (averaged over 40 testing cases) between two variants of WearGP framework and the actual CFD wear predictions on the back shroud surface.

(a) WearGP (H, Q, d_{50}, C_v).

(b) WearGP $(H, \%BEPQ, d_{50}, C_v)$.
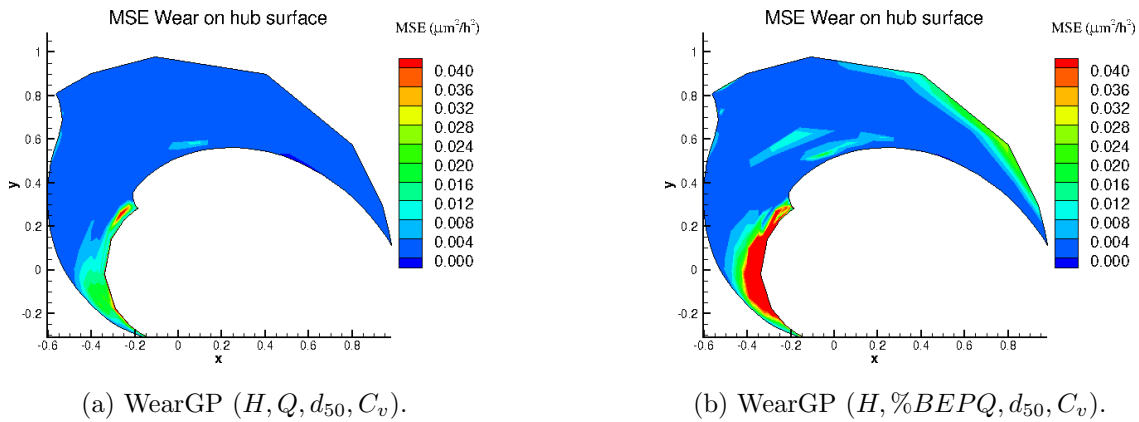
Figure 14: Comparison of MSE (averaged over 40 testing cases) between two variants of WearGP framework and the actual CFD wear predictions on the front shroud surface.

Table 3 shows the descriptive statistics of the MSE between two WearGP variants and the CFD impeller wear model. It is clearly proved that the basis choice of (H, Q, d_{50}, C_v) supersedes the choice of $(H, \%BEPQ, d_{50}, C_v)$ in terms of predicting error, for all surfaces of the slurry pump impeller.

To provide an insight regarding the robustness of the proposed WearGP framework, the mean absolute relative error (MARE) is quantified as a measure for uncertainty in wear predictions. The MARE is calculated

18
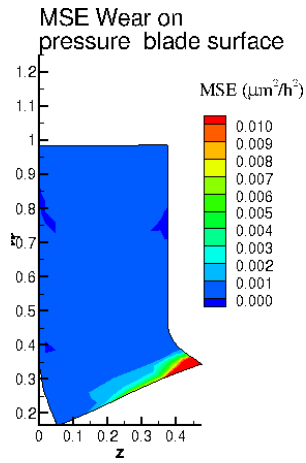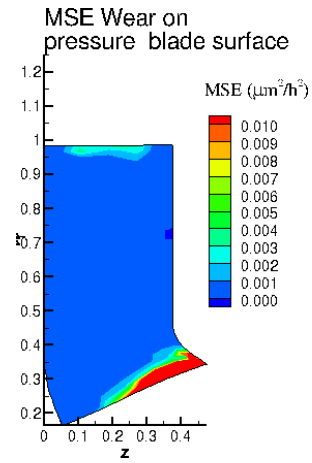
(a) WearGP (H, Q, d_{50}, C_v).

(b) WearGP $(H, \%BEPQ, d_{50}, C_v)$.

Figure 15: Comparison of MSE (averaged over 40 testing cases) between two variants of WearGP framework and the actual CFD wear predictions on the pressure side of the impeller vane.

(a) WearGP (H, Q, d_{50}, C_v).

(b) WearGP $(H, \%BEPQ, d_{50}, C_v)$.

Figure 16: Comparison of MSE (averaged over 40 testing cases) between two variants of WearGP framework and the actual CFD wear predictions on the suction side of the impeller vane.

19

Table 3: Descriptive comparison of the MSE ($\mu m^2/hr^2$) of the total wear $\dot{W}_T$ between two WearGP variants and the CFD impeller wear model prediction, averaged over all 40 testing cases, and over all nodes in each respective region.

		(H, Q, d_{50}, C_v)	$(H, \%BEPQ, d_{50}, C_v)$
	min	0.0000	0.0000
Back shroud	mean	**0.0081**	**0.0151**
	max	0.1441	0.2265
	min	0.0000	0.0000
Front shroud	mean	**0.0062**	**0.0284**
	max	0.3126	0.7994
	min	0.0000	0.0000
Pressure	mean	**0.0020**	**0.0068**
	max	0.1924	0.5195
	min	0.0000	0.0000
Suction	mean	**0.0050**	**0.0262**
	max	0.0647	0.4528

as

$$MARE = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \left| \frac{\dot{W}_T^{(\text{WearGP})} - \dot{W}_T^{(\text{CFD})}}{\dot{W}_T^{(\text{CFD})}} \right| \times 100\%, \tag{9}$$

Figures 17 to 20 plot and compare the MARE between the two WearGP implementation variants on the back shroud, front shroud, pressure vane side, suction vane side of the slurry pump impeller, respectively, where the color bars are again adjusted to the same scale. Once again, the quantitative comparison figures show that for WearGP framework, the MARE using (H, Q, d_{50}, C_v) as a basis is smaller than the MSE using the $(H, \%BEPQ, d_{50}, C_v)$. The relative error on the region with local high wear rate are approximately 8%, averaged across all the cases. The numerical error is likely caused by implementation, retaining only a certain number after decimal points. To circumvent the numerical artifacts which undermines the performance of the proposed WearGP framework, a threshold of 1μm/hr is imposed. The nodes where the CFD wear predictions are smaller than the threshold are excluded from the relative error quantification process. In the regions with locally high wear rate, which are more important, the wear predictions accurately approximate the CFD wear model.

Table 4 shows the MARE, as a measure of relative error. It is noted that as indicated earlier, the high relative error is due to the small wear magnitude in some regions, and the numerical error which was implemented into the WearGP framework. The wall nodes where CFD wear predictions are less than the imposed threshold are excluded from the analysis.
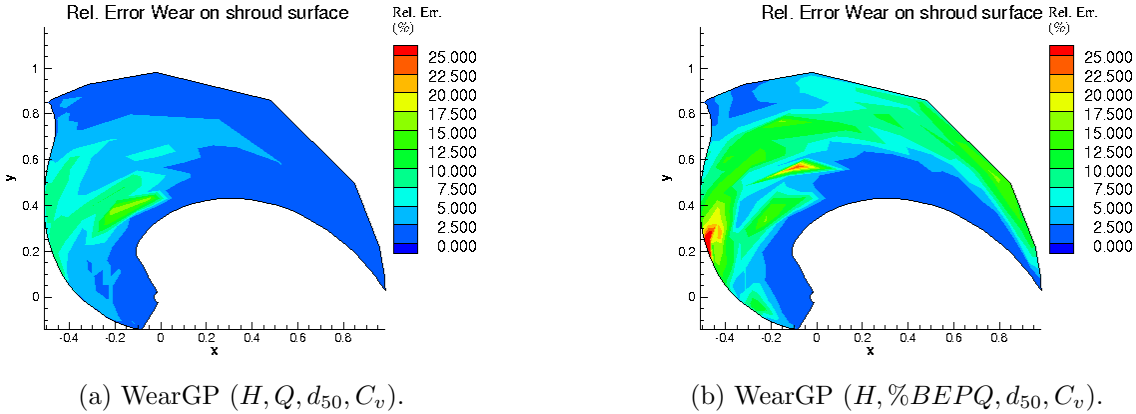

(a) WearGP (H, Q, d_{50}, C_v).   (b) WearGP $(H, \%BEPQ, d_{50}, C_v)$.
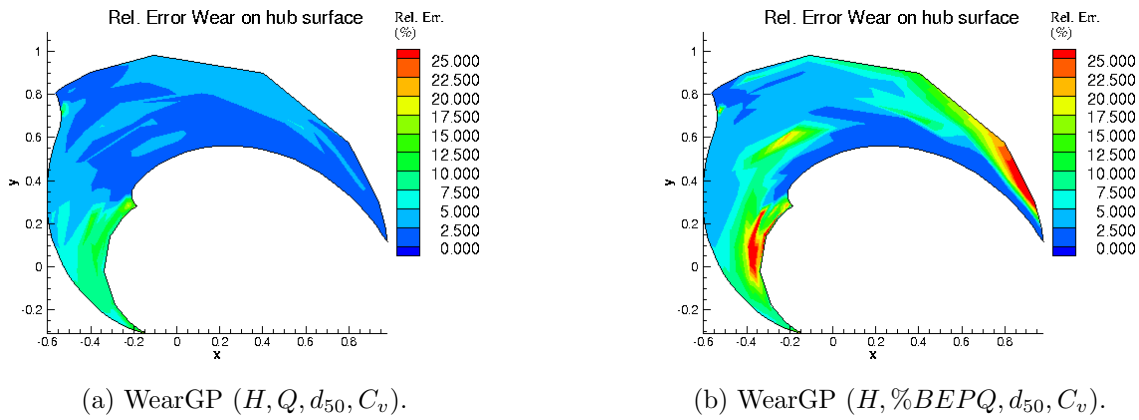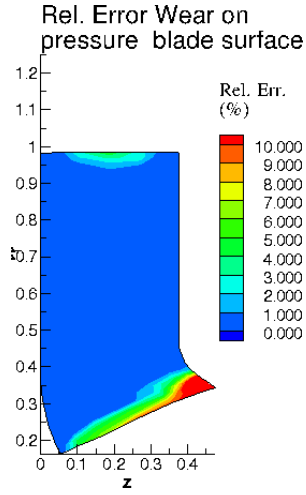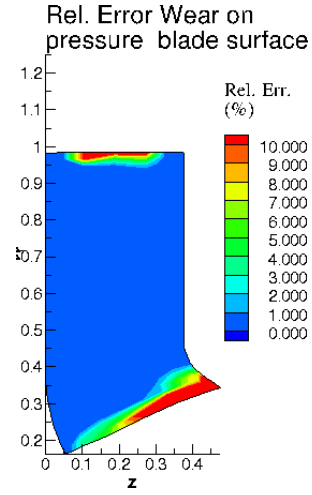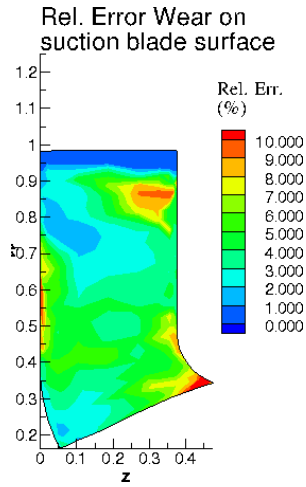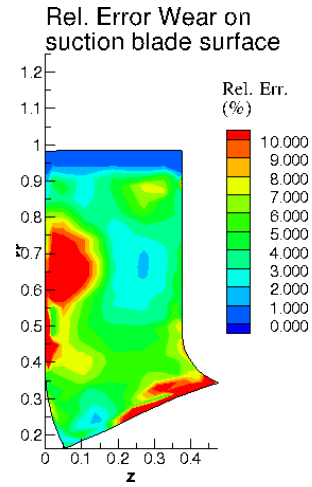
Figure 17: Comparison of the MARE % with threshold 0.1μm/hr enabled between two variants of WearGP framework and the actual CFD wear predictions on the back shroud surface.

Table 5 presents the computational time for training and testing of two WearGP variants, compared with the CFD impeller wear model, for all 40 testing cases. The testing time is computed as the elapsed time to

21

Table 4: Descriptive comparison of the MARE (%) with threshold 0.1μm/hr enabled of the total wear $\dot{W}_T$ between two WearGP variants and the CFD impeller wear model prediction, averaged over all 40 testing cases, and over all nodes in each respective region.

		(H, Q, d_{50}, C_v)	$(H, \%BEPQ, d_{50}, C_v)$
Back shroud	min	0.1273%	0.1503%
	mean	**6.7169%**	**9.0529%**
	max	19.7952%	56.1574%
Front shroud	min	0.0070%	0.0667%
	mean	**3.5324%**	**6.7354%**
	max	15.7389%	30.3777%
Pressure	min	0.0653%	0.0237%
	mean	**3.9786%**	**7.8427%**
	max	16.9161%	78.4961%
Suction	min	0.1039%	0.3798%
	mean	**4.7568%**	**6.7658%**
	max	14.1376%	27.4350%

(a) WearGP (H, Q, d_{50}, C_v).

(b) WearGP $(H, \%BEPQ, d_{50}, C_v)$.

Figure 18: Comparison of the MARE % with threshold 0.1μm/hr enabled between two variants of WearGP framework and the actual CFD wear predictions on the front shroud surface.

(a) WearGP (H, Q, d_{50}, C_v).

(b) WearGP $(H, \%BEPQ, d_{50}, C_v)$.

Figure 19: Comparison of the MARE % with threshold 0.1μm/hr enabled between two variants of WearGP framework and the actual CFD wear predictions on the pressure side of the impeller vane.

(a) WearGP (H, Q, d_{50}, C_v).

(b) WearGP $(H, \%BEPQ, d_{50}, C_v)$.

Figure 20: Comparison of the MARE % with threshold 0.1μm/hr enabled between two variants of WearGP framework and the actual CFD wear predictions on the suction side of the impeller vane.

predict the wear rate for 40 cases. The total amount of time for ML approaches are both approximately at 10^0 order in terms of seconds, whereas the CFD approach is at 10^5 order in terms of seconds. We show that the same type of wear predictions can be achieved in a significantly shorter amount of time. The computational time difference is in five orders of magnitude.

Table 5: Comparison of computational time in predicting wear using WearGP framework and CFD impeller wear model for 40 testing cases. The training time for CFD is computed by totaling the computational time for 144 training cases.

	WearGP		**CFD**
	(H, Q, d_{50}, C_v)	$(H, \%BEPQ, d_{50}, C_v)$	
Training time (s)	876388.2345	876369.0820	876,213
Testing time (s)	**1.2961**	**1.0764**	252,022

5. Casing wear model

In this section, we demonstrate the effectiveness of the proposed WearGP framework to predict the local erosive wear on slurry pump casing, using in-house CFD wear code. The description of the CFD casing wear model is summarized in Section 5.1. The WearGP framework is then deployed to predict the local wear rate, following the data mining. The numerical performance of WearGP is then compared to the CFD casing wear model in Section 5.2, in terms of both accuracy and efficiency.

5.1. CFD casing wear model

The CFD casing wear model used to predict erosive wear in the centrifugal slurry pump casing is also the co-authors' previous work [26, 27]. Similar to the CFD impeller wear model, the CFD casing wear model is summarized as follows. Figure 21 shows the computational domain of the CFD casing wear model. An Eulerian-Eulerian framework is utilized, and the continuity and momentum equations of the mixture and different solids species are derived using volume and time averaged governing equations. The casing inlet is divided into three sections, as shown in Figure 22. The inlet velocity boundary conditions are applied on the region AA', BC, and B'C', where radial and tangential velocities are imposed separately based on the head H and flow rate, Q, of the slurry pump. The regions AB and A'B' are treated as walls, where Spalding wall functions [20] are applied. Similar to the CFD impeller wear model, in CFD casing wear model, the Spalart-Allmaras turbulence model [19] is used, and the set of nonlinear governing equations in the finite

24

element problem are solved iteratively until a given residual level is achieved. Based on the CFD solutions, the constitutive model for wear prediction is then utilized to predict the total wear as a function of concentration, density, velocity magnitude, tangential velocity, shear stress, impingement angle, and the particle size [26, 27].
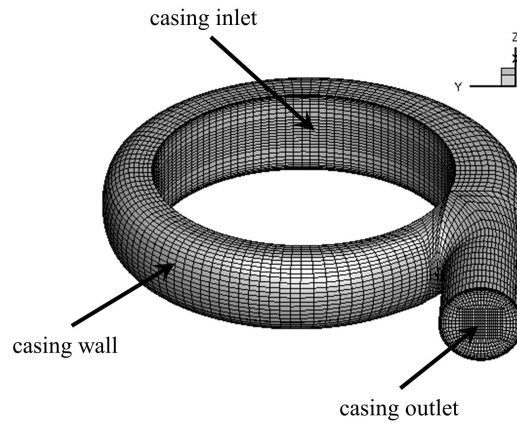

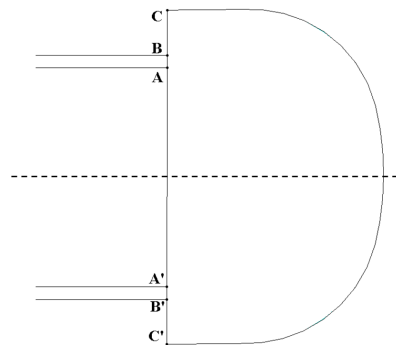Figure 21: Three-dimensional pump casing with mesh and its boundary conditions.

Figure 22: Sections of casing inlet.

335 *5.2. Comparison between WearGP and CFD casing wear model*

GIW pump casing 0509X LSA-36 5525D 7723D-A0 is chosen as a case study in this example. The radius to inside tongue is 0.5583 m. The casing inner radius is 0.9906 m, whereas the casing outer radius is approximately 1.397 m. Figure 23 shows the actual slurry pump casing in the field.

To evaluate the performance of the WearGP framework, we compare its prediction with the 3D CFD wear
340 simulation for slurry pump impeller. 144 CFD simulations are performed to build the training dataset, and 40 CFD simulations are used to build a testing dataset. The testing operating conditions are all sampled using Monte Carlo algorithm. A structured mesh of 59844 nodes and 55600 elements is used to discretize the

25

Figure 23: GIW casing 0509X LSA-36 5525D 7723D-A0 in the field.

computational domain. Among those, the 4719 wall nodes are used to evaluate the wear performance of the impeller.

Figure 24 and Figure 25 show the comparison between the CFD simulation wear predictions and WearGP implementation based on the CFD results for case 7 and case 16 in the 40 testing cases. Both figures show an excellent agreement between the CFD simulation results and the ML prediction, verifying the implementation and validating the proposed WearGP approach. The ML results are obtained using 144 CFD runs with different operating conditions for the pump. In other words, the boundary conditions of the computational domain are changed, while the CFD formulation remains the same.

To evaluate the accuracy of the WearGP framework, the MSE, as described in Equation 8, is used again to quantify the error between the CFD casing wear model and its two WearGP variants. Essentially, it measures the average discrepancy of the predicted wear using WearGP framework and CFD casing wear model, among 40 testing cases. Figure 26 show the comparison of MSE between two WearGP implementation variants, where the MSE is significantly smaller for the (H, Q, d_{50}, C_v) implementation, compared to the $(H, \%BEPQ, d_{50}, C_v)$. The error magnifies the most near the tongue region of the casing.

Figure 27 plots the mean absolute relative error between the WearGP prediction and the CFD prediction, again, showing the high relative error regions typically associate with low local wear regions, which could pos-
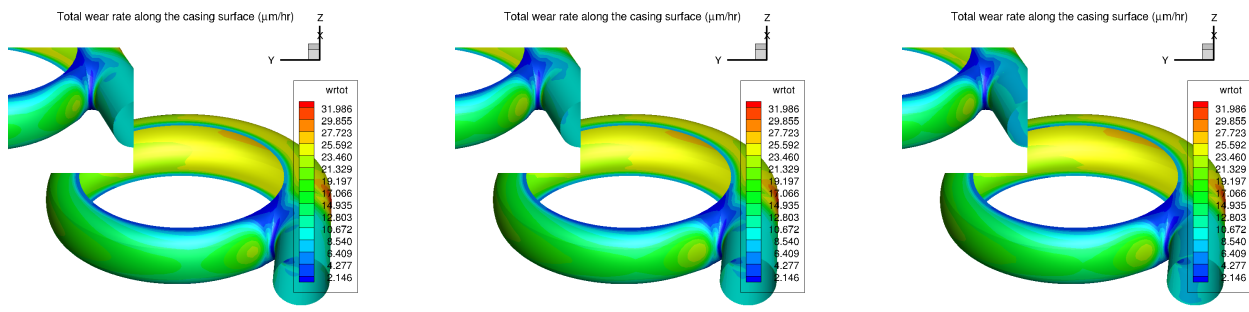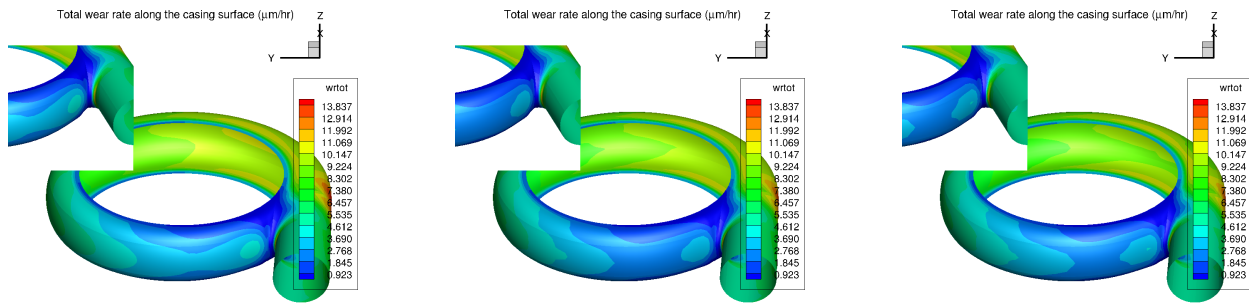
(a) 3D slurry pump casing CFD wear simulation result.

(b) WearGP: (H, Q, d_{50}, C_v) implementation result.

(c) WearGP: $(H, \%BEPQ, d_{50}, C_v)$ implementation result.

Figure 24: Comparison for wear prediction results between CFD simulation and two variants of WearGP implementation for case 7/40 testing cases.
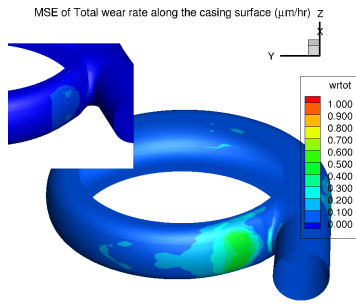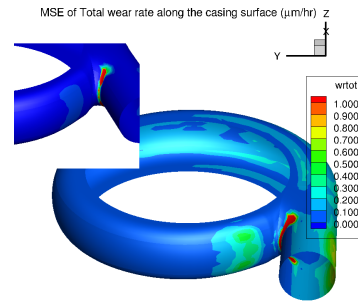

(a) 3D slurry pump casing CFD wear simulation result.

(b) WearGP: (H, Q, d_{50}, C_v) implementation result.

(c) WearGP: $(H, \%BEPQ, d_{50}, C_v)$ implementation result.

Figure 25: Comparison for wear prediction results between CFD simulation and two variants of WearGP implementation for case 16/40 testing cases.

(a) WearGP (H, Q, d_{50}, C_v).   (b) WearGP $(H, \%BEPQ, d_{50}, C_v)$.
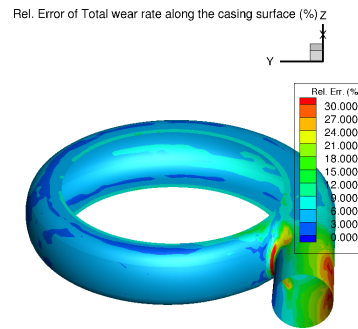
Figure 26: Comparison of MSE (averaged over 40 testing cases) between two variants of WearGP framework and the actual CFD wear predictions in the slurry pump casing.

sibly be due to numerical implementation. However, the regions with low wear rate are relatively unimportant, compared to regions with high wear rate.
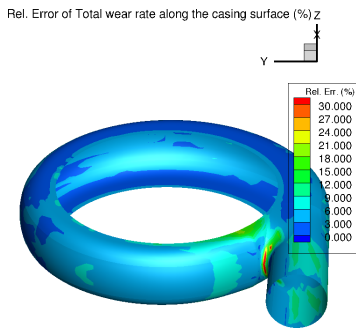
(a) WearGP (H, Q, d_{50}, C_v).   (b) WearGP $(H, \%BEPQ, d_{50}, C_v)$.

Figure 27: Comparison of the MARE (averaged over 40 testing cases) between two variants of WearGP framework and the actual CFD wear predictions in the slurry pump casing.

Table 6 compares the accuracy between two WearGP variants implementation. It is clearly shown that the (H, Q, d_{50}, C_v) basis supersedes the $(H, \%BEPQ, d_{50}, C_v)$ approach. The mean value of the MSE for the former approach, which uses (H, Q, d_{50}, C_v) as the basis, is approximately 3 times lower than of the later approach.

Table 7 compares the MARE, as a measure of accuracy between two WearGP variants implementation. In general, the (H, Q, d_{50}, C_v) basis performs better in most of the regions on the slurry pump casing. Overall, the $(H, \%BEPQ, d_{50}, C_v)$ achieves around 2% of relative error, compared to 3% of relative error using the

28

Table 6: Descriptive comparison of the MSE ($\mu m^2/hr^2$) of the total wear $\dot{W}_T$ between two WearGP variants and the CFD casing wear model prediction.

	(H, Q, d_{50}, C_v)	$(H, \%BEPQ, d_{50}, C_v)$
min	0.0007	0.0024
mean	**0.0531**	**0.1472**
max	0.5293	41.3362

$(H, \%BEPQ, d_{50}, C_v)$ approach. The wall nodes where CFD wear predictions are less than the imposed threshold are excluded from the analysis. The high MARE are associated with only a few nodes, where both the ML and CFD wear predictions.

Table 7: Descriptive comparison of the MARE (%) with threshold 0.1μm/hr enabled of the total wear $\dot{W}_T$ between two WearGP variants and the CFD casing wear model prediction.

	(H, Q, d_{50}, C_v)	$(H, \%BEPQ, d_{50}, C_v)$
min	1.0187%	1.9588%
mean	**4.6578%**	**6.9938%**
max	64.9434%	84.7626%

Table 8 compares the computational time to obtain the wear predictions for all 40 cases, including training and testing time. The training time for CFD simulation is summed over 144 training cases. The testing time for CFD simulation is summed over 40 testing cases.

Table 8: Comparison of computational time in predicting wear using WearGP framework and CFD casing wear model for 40 testing cases. The training time for CFD is computed by summing the computational time for 144 training cases.

	WearGP		CFD
	(H, Q, d_{50}, C_v)	$(H, \%BEPQ, d_{50}, C_v)$	
Training time (s)	4043908.0146	4043879.1943	4,043,374
Testing time (s)	4.3613	4.2793	1,061,203

6. Discussion

Figure 28 and Figure 29 present the computational time to obtain the wear predictions for 40 testing cases in log scale for slurry pump impeller and slurry pump casing, respectively. Five to six orders of magnitude for computational time reduction is observed. It is indicated that the ML approach has the clear advantage of computational efficiency, compared to the CFD approach.
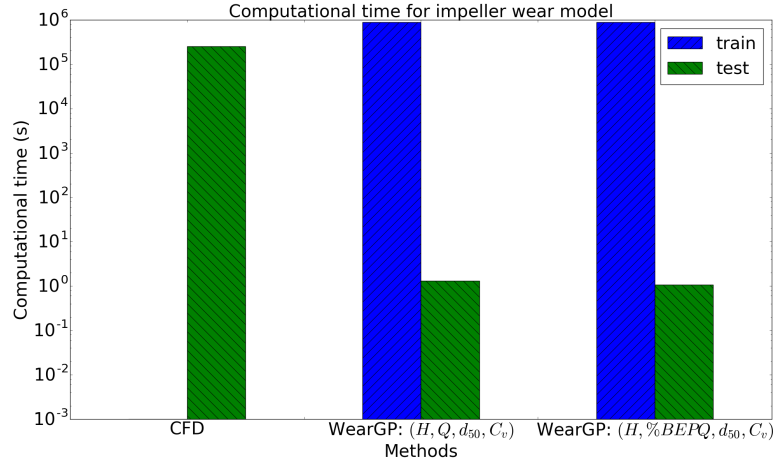

Figure 28: Comparison of computational efficiency for impeller wear predictions between CFD and ML methods.
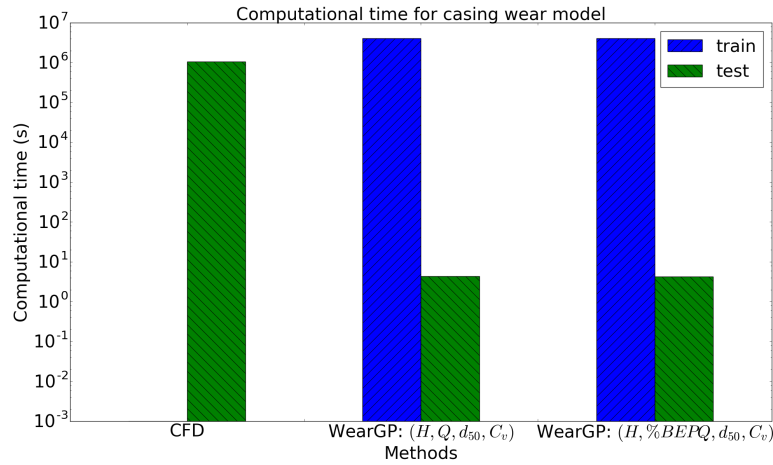

Figure 29: Comparison of computational efficiency for casing wear predictions between CFD and ML methods.

Figure 30a and Figure 30b show the coefficient of determination R^2 between WearGP and CFD wear predictions for slurry pump impeller, using (H, Q, d_{50}, C_v) and $(H, \%BEPQ, d_{50}, C_v)$ basis, respectively. Figure
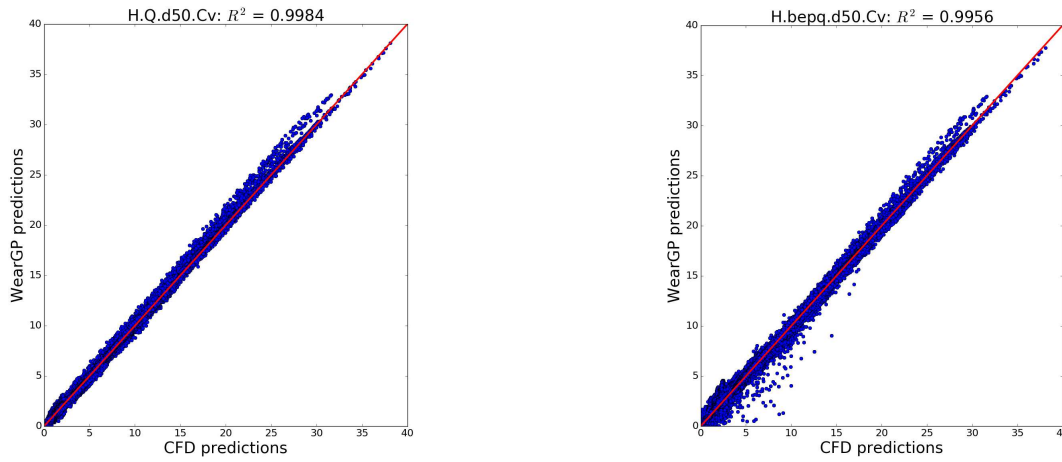
31a and Figure 31b show the coefficient of determination R^2 between WearGP and CFD wear predictions for slurry pump casing, using (H, Q, d_{50}, C_v) and $(H, \%BEPQ, d_{50}, C_v)$ basis, respectively. The CFD wear predictions are plotted as x-axis, whereas the WearGP wear predictions are plotted as y-axis. A solid diagonal line showing ideal case where two predictions are identical, $y = x$, is also plotted. Notably, the WearGP framework shows a strong predictive capability, with $R^2 > 0.99$ in all cases, demonstrating a highly accurate wear predictions.
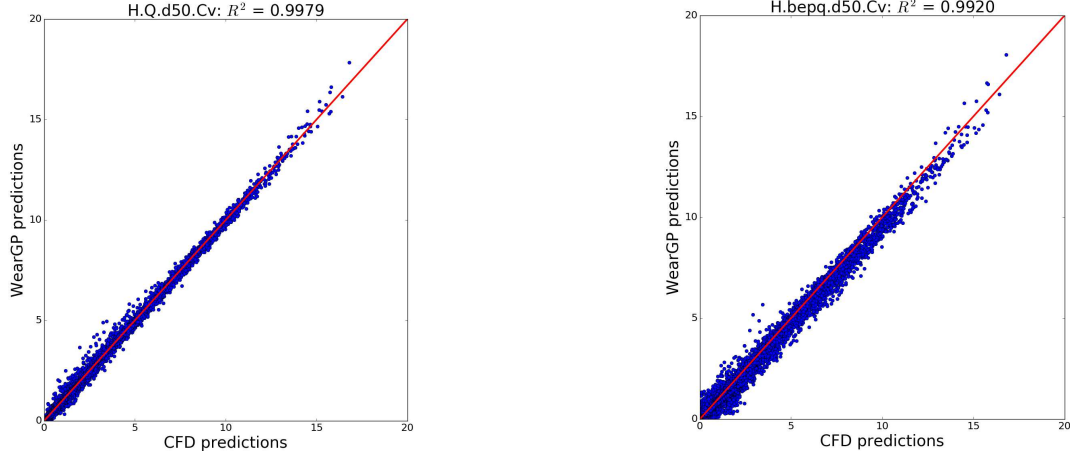
(a) Coefficient of determination R^2 between WearGP and CFD wear predictions for slurry pump impeller, using (H, Q, d_{50}, C_v) basis.

(b) Coefficient of determination R^2 between WearGP and CFD wear predictions for slurry pump impeller, using $(H, \%BEPQ, d_{50}, C_v)$ basis.

Figure 30: Comparison of the coefficient of determination R^2 for WearGP and CFD wear predictions in slurry pump impeller.

Numerical verification and experimental validation studies have been conducted in parallel with the development and implementation of the CFD wear models for slurry pump components. Most of the results have not been published, but some results can be found in the literature. For example, ultrasonic measurement techniques have been employed by Furlan et al. [28, 29] to investigate particle velocity and concentration and compare against the CFD wear models described in this paper. Wear coefficients in the CFD wear models used in this study are obtained from Coriolis and impact wear testers [30]. Experimental results in laboratories have been shown to agree fairly well with numerical predictions of pump wear within casings and impellers. Experimental field data are much harder to validate because the operating conditions are not perfectly controlled, and many assumptions no longer hold. Furthermore, as components in the pump wear out, the geometry of components evolve gradually, and thereby altering the flow. Thus, the numerical wear predictions are only technically valid for predicting wear rates on unworn pumps. However, wear rate ratios predicted from one

31

(a) Coefficient of determination R^2 between WearGP and CFD wear predictions for slurry pump casing, using (H, Q, d_{50}, C_v) basis.

(b) Coefficient of determination R^2 between WearGP and CFD wear predictions for slurry pump casing, using $(H, \%BEPQ, d_{50}, C_v)$ basis.

Figure 31: Comparison of the coefficient of determination R^2 for WearGP and CFD wear predictions in slurry pump casing.

pump to another or from one set of operating conditions to another set have been found to be reliable in making wear life predictions in cases where a baseline, field observed wear life is known. In this paper, none of the training or testing conditions has been thoroughly validated.

For the pump impeller code, a representative case for experimental validation of the CFD impeller wear model for a cyclone feed pump (mill discharge) in a copper mine is shown in Figure 32. Figures 32a and 32b show the wear prediction on the pressure and suction sides of the impeller vane, respectively. The localized wear regions are located near the trailing edge of the impeller on the suction side, as well as midway of the impeller on the pressure side.

For the pump casing code, Figure 33 shows the inside of a pump casing with significant wear at the tongue and on the casing side wall, which is typically referred to as side wall gouging, for a pump that was operated at low flow rates relative to the pump best efficiency point flow rate. Figure 33a presents the numerical predictions using the CFD casing wear model at the same operating conditions, showing a substantial wear rate on the side wall. Figures 32 and 33 demonstrate that qualitatively, numerical predictions and experimental observations are very close. Quantitative wear analyses with high accuracy in the field are harder to obtain because the operating conditions in practice often are transient. However, as mentioned previously, quantitative validation between the CFD wear codes and experimental data have been conducted on liquid and solids velocity, as well as local particle concentration.
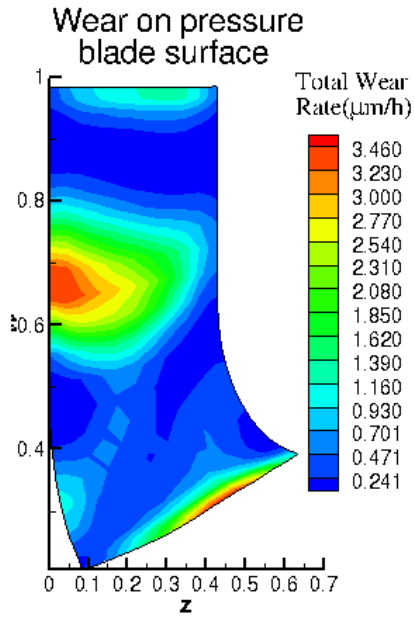
(a) CFD wear prediction on pressure side of impeller vane.
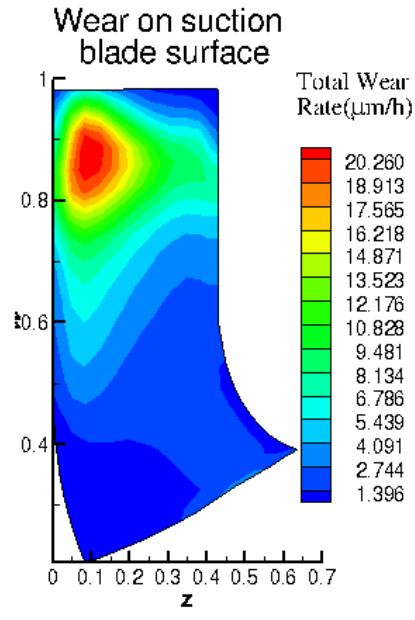
(b) CFD wear prediction on suction side of impeller vane.

(c) Field observation of impeller wear.

Figure 32: A qualitative experimental validation of CFD impeller wear model.
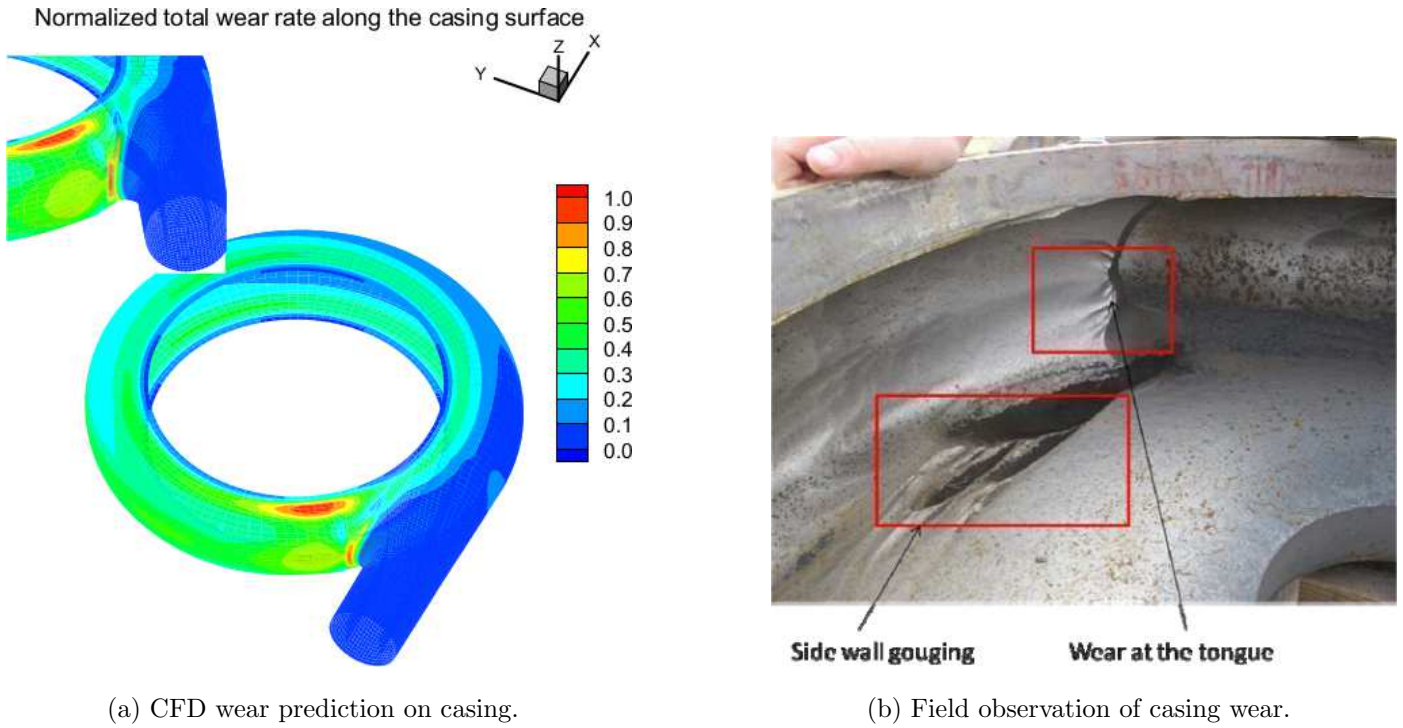
33

(a) CFD wear prediction on casing.

(b) Field observation of casing wear.

Figure 33: A qualitative experimental validation of CFD casing wear model [27].

The GP formulation is well-known to have scalability problems. If the number of observations n is larger than 10^4, it is computationally very expensive to obtain the inverse of the covariance matrix, with a size of $n \times n$. In WearGP framework, this scalability issue is avoided by decomposing the large dataset based on the nodes of the mesh, and treating the nodal wear solution as the GP output, where the pump operating conditions or boundary conditions are considered as GP inputs. This significantly reduces the dimensionality of the WearGP approach, allowing the nodal wear solution to be predicted with an acceptable accuracy. The relative errors of wear predictions using the proposed WearGP framework fall within another validation studies [29], where concentration and particle velocities are compared between experimental and computational model. Compared to deep learning approaches, such as Guo et al. [3], the size of the training dataset is significantly smaller. In this work, the WearGP framework is trained using $\mathcal{O}(10^2)$ samples, whereas for example, Guo et al. [3] used a large number of samples in the order of $\mathcal{O}(10^4)$ to $\mathcal{O}(10^5)$. The main difference of computational effectiveness is originated from the solid theoretical foundation of GP, which yields a smaller approximation error, compared to the deep learning and convolutional neural network approaches.

Another advantage of the nodal decomposition approach is that the local discontinuity in wear predictions

is preserved, because the nodes are treated independently from each other. For example, near the tongue region of the slurry pump casing, there are small local regions that typically correspond to high wear rates. If the nodes are not considered as independent from each other, the GP kernel would induce wear predictions to be overly smoothed, compared to the CFD wear model. Thus, nodal decomposition is appropriate for CFD wear predictions with highly varied local wear, where local wear are concentrated at the locations with high solid particle velocities.

There are two shortfalls of the nodal decomposition approach. First, the number of GPs that are required to be trained is in the same order of the number of wall nodes in the mesh. Second, the current WearGP algorithm requires the same mesh settings, only allowing boundary conditions to be varied simultaneously. Relaxing this constraint would leverage the flexibility of the current framework, which consequently requires further studies. There are two possible extensions for the current work. The first extension is the adaptation to multi-fidelity meshes, where constraints on the mesh are completely removed. Such relaxation allows the results from coarse mesh and fine mesh to be combined. However, a significant modification must be applied in order to account for extra nodes that do not exist in the coarse mesh, but do exist in the fine mesh. The second extension is to use the WearGP framework to propose a good initial guess for faster CFD simulations with fewer iterations.

The current WearGP framework can also be easily extended for wear prediction with different meshes, for example, by constructing a global GP or equivalent ML framework that includes the 3D Cartesian coordinates of the nodes as inputs, as well as other inputs in Equation 7. Such an extension is more practical, but requires a global ML framework, instead of a local and nodal GP approach proposed in this paper. Because the dimensionality of the problem increases, a larger dataset might be needed to obtain the same accuracy achieved using the local WearGP approach. A scalable GP framework can also be adopted to cope with scalability issues and extend the current WearGP framework, such as the subset of data, the subset of regressor, the deterministic training conditional, and the partially and fully independent training conditions approximations [31] [32]. Mesh independence studies were performed during the development and implementation of the CFD wear codes. The final mesh used in this study has been selected based on the mesh independence results.

7. Conclusion

In this paper, the WearGP framework that predicts 3D wear for slurry pump components, impellers and casings, in times on the order of seconds is introduced. The proposed WearGP framework is investigated

by training using 144 CFD simulations, and testing using 40 CFD simulations. A data mining process is applied to extract the nodal wear solution, and a GP is constructed at each node to predict the wear rate at the testing operating conditions. Four parameters are used to vary the pump operating conditions, which in turns is interpreted as boundary conditions for CFD simulations. Two variants of the WearGP are proposed: (H, Q, d_{50}, C_v) and $(H, \%BEPQ, d_{50}, C_v)$. The accuracy performance between two variants are tested and analyzed in terms of MSE and MARE. The MARE is shown to be 1.4-8.7% for slurry pump impellers, and 1.9-3.3% for slurry pump casings. It is concluded that the (H, Q, d_{50}, C_v) approach has smaller MSE and MARE in general, and thus is more accurate, compared to the $(H, \%BEPQ, d_{50}, C_v)$ approach.

A significant computational efficiency of five or six orders of magnitude is observed, where the prediction time is reduced from $\mathcal{O}(10^5) - \mathcal{O}(10^6)$ seconds to $\mathcal{O}(10^0)$ seconds. Additionally, it is shown that the wear predictions using the ML approach approximates well the CFD wear solutions. Visually, the wear pattern is essentially similar, comparing between the ML-based and CFD approaches.

Compared to other deep learning approaches where $\mathcal{O}(10^4)$ to (10^6) data points are required, our proposed WearGP framework can be trained more efficiently using only $\mathcal{O}(10^2)$ data points. Yet, the relative errors of the WearGP predictions are less than 9%, compared to the CFD predictions, because of the rigorous formulations and properties of GP. The proposed framework is proven to be feasible for industrial applications.

8. Acknowledgment

References

[1] M. D. Hill, M. R. Marty, Amdahl's law in the multicore era, Computer 41 (7) (2008) 33–38.

[2] C. E. Rasmussen, Gaussian processes in machine learning, in: Advanced lectures on machine learning, Springer, 2004, pp. 63–71.

[3] X. Guo, W. Li, F. Iorio, Convolutional neural networks for steady flow approximation, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2016, pp. 481–490.

[4] J. Tompson, K. Schlachter, P. Sprechmann, K. Perlin, Accelerating Eulerian fluid simulation with convolutional networks, arXiv preprint arXiv:1607.03597.

[5] T. P. Miyanawala, R. K. Jaiman, An efficient deep learning technique for the Navier-Stokes equations: application to unsteady wake flow dynamics, arXiv preprint arXiv:1710.09099.

[6] M. Chu, N. Thuerey, Data-driven synthesis of smoke flows with CNN-based feature descriptors, ACM Transactions on Graphics (TOG) 36 (4) (2017) 69.

[7] K. Velten, R. Reinicke, K. Friedrich, Wear volume prediction with artificial neural networks, Tribology International 33 (10) (2000) 731–736.

[8] Z. Zhang, N.-M. Barkoula, J. Karger-Kocsis, K. Friedrich, Artificial neural network predictions on erosive wear of polymers, Wear 255 (1-6) (2003) 708–713.

[9] S. Danaher, S. Datta, I. Waddle, P. Hackney, Erosion modelling using bayesian regulated artificial neural networks, Wear 256 (9-10) (2004) 879–888.

[10] A. Suresh, A. Harsha, M. Ghosh, Solid particle erosion studies on polyphenylene sulfide composites and prediction on erosion data using artificial neural networks, Wear 266 (1-2) (2009) 184–193.

[11] S. Shamshirband, A. Malvandi, A. Karimipour, M. Goodarzi, M. Afrand, D. Petković, M. Dahari, N. Mahmoodian, Performance investigation of micro-and nano-sized particle erosion in a 90 elbow using an ANFIS model, Powder Technology 284 (2015) 336–343.

[12] J. Qu, M. J. Zuo, Support vector machine based data processing algorithm for wear degree classification of slurry pump systems, Measurement 43 (6) (2010) 781–791.

[13] D. Pandya, B. Dennis, R. Russell, A computational fluid dynamics based artificial neural network model to predict solid particle erosion, Wear 378 (2017) 198–210.

[14] W. Dai, S. Cremaschi, H. J. Subramani, H. Gao, Uncertainty quantification in erosion predictions using data mining methods, Wear 408 (2018) 108–119.

37

[15] E. Brochu, V. M. Cora, N. De Freitas, A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning, arXiv preprint arXiv:1012.2599.

[16] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, N. de Freitas, Taking the human out of the loop: A review of Bayesian optimization, Proceedings of the IEEE 104 (1) (2016) 148–175.

[17] K. V. Pagalthivarthi, J. M. Furlan, R. J. Visintainer, Wear rate prediction in multi-size particulate flow through impellers, in: ASME 2013 Fluids Engineering Division Summer Meeting, American Society of Mechanical Engineers, 2013.

[18] K. V. Pagalthivarthi, J. M. Furlan, R. J. Visintainer, Effective particle size representation for erosion wear in centrifugal pump casings, in: ASME 2017 Fluids Engineering Division Summer Meeting, American Society of Mechanical Engineers, 2017, pp. V01CT15A004–V01CT15A004.

[19] P. R. Spalart, S. R. Allmaras, et al., A one equation turbulence model for aerodynamic flows, Recherche Aerospatiale-French Edition (1994) 5–5.

[20] F. White, Viscous Fluid Flow 2nd Edition, McGraw-Hill New York, 1991.

[21] T. J. Hughes, A. N. Brooks, A theoretical framework for petrov-galerkin methods with discontinuous weighting functions: Application to the streamline-upwind procedure, Finite elements in fluids (1982) 47–65.

[22] A. N. Brooks, T. J. Hughes, Streamline upwind/petrov-galerkin formulations for convection dominated flows with particular emphasis on the incompressible navier-stokes equations, Computer methods in applied mechanics and engineering 32 (1-3) (1982) 199–259.

[23] O. Schenk, K. Gärtner, Solving unsymmetric sparse systems of linear equations with PARDISO, Future Generation Computer Systems 20 (3) (2004) 475–487.

[24] H. H. Tian, G. R. Addie, K. V. Pagalthivarthi, Determination of wear coefficients for erosive wear prediction through coriolis wear testing, Wear 259 (1-6) (2005) 160–170.

[25] K. Pagalthivarthi, R. Visintainer, Finite element prediction of multi-size particulate flow through three-dimensional channel: Code validation, The Journal of Computational Multiphase Flows 5 (1) (2013) 57–72.

[26] K. V. Pagalthivarthi, R. J. Visintainer, Solid-liquid flow-induced erosion prediction in three-dimensional pump casing, in: ASME 2009 Fluids Engineering Division Summer Meeting, American Society of Mechanical Engineers, 2009, pp. 611–617.

[27] K. V. Pagalthivarthi, J. M. Furlan, R. J. Visintainer, Finite element prediction of multi-size particulate flow through three-dimensional pump casing, in: ASME/JSME/KSME 2015 Joint Fluids Engineering Conference, American Society of Mechanical Engineers, 2015, pp. V001T31A002–V001T31A002.

[28] J. M. Furlan, V. Mundla, J. Kadambi, N. Hoyt, R. Visintainer, G. Addie, Development of A-scan ultrasound technique for measuring local particle concentration in slurry flows, Powder technology 215 (2012) 174–184.

[29] J. M. Furlan, M. Garman, J. Kadambi, R. J. Visintainer, K. V. Pagalthivarthi, Ultrasonic measurements of local particle velocity and concentration within the casing of a centrifugal pump, in: ASME/JSME/KSME 2015 Joint Fluids Engineering Conference, American Society of Mechanical Engineers, 2015, pp. V001T31A003–V001T31A003.

[30] K. V. Pagalthivarthi, J. M. Furlan, R. J. Visintainer, Consistent evaluation of wear coefficients from the experiments for use in CFD simulations, in: ASME 2017 Fluids Engineering Division Summer Meeting, American Society of Mechanical Engineers, 2017, pp. V01CT15A005–V01CT15A005.

[31] J. Quinonero-Candela, C. E. Rasmussen, C. K. Williams, Approximation methods for gaussian process regression, Large-scale kernel machines (2007) 203–224.

[32] K. Chalupka, C. K. Williams, I. Murray, A framework for evaluating approximation methods for Gaussian process regression, Journal of Machine Learning Research 14 (Feb) (2013) 333–350.