

Article

A Mechanism-Informed Gaussian Process Surrogate Model for Solid-Particle Erosion Prediction in Gas–Solid Bent Pipe Flows

Junyan Ma ^{1,†} , Jiafu Yang ^{2,†}, Wenwen Yang ³, Yonggang Song ³, Adilanmu Sitahong ^{2,*}, Duoming Pan ³ and Yong Huang ¹

¹ School of Energy Engineering, Xinjiang Institute of Engineering, Urumqi 830023, China; majunyanbyr@163.com (J.M.); 13321366105@163.com (Y.H.)

² School of Mechanical Engineering, Xinjiang University, Urumqi 830017, China; ridiculous151@163.com

³ Xinte Silicon Materials Co., Ltd., No. 19 Huayi 1st Road, Xiheishan Industrial Park, Zhundong Economic and Technological Development Zone, Changji Hui Autonomous Prefecture 831100, Xinjiang, China; 19171160669@163.com (W.Y.); 18337678570@163.com (Y.S.); 18625427728@163.com (D.P.)

* Correspondence: adilanmujx@xju.edu.cn

† These authors contributed equally to this work.

Abstract

In cold hydrogenation processes, bent pipes are highly susceptible to severe localized erosion under hydrogen–silica powder gas–solid two-phase flow. However, high-fidelity numerical simulations are computationally expensive and thus inadequate for rapid assessment under multiple operating conditions. To overcome this limitation, an MI-UK-GPR-based method is proposed for predicting the erosion rate of cold hydrogenation bent pipes. Based on a validated CFD model, six input variables, namely pipe inner diameter, curvature ratio, bend angle, particle mass flow rate, particle size, and particle velocity, were selected. Latin hypercube sampling was employed to generate parameter combinations, and the corresponding maximum erosion rates were obtained through high-fidelity CFD simulations to construct an LHS-CFD sample database. The input variables were then normalized, and the maximum erosion rates were log-transformed. On this basis, an MI-UK-GPR model integrating a mechanistic trend term with a Gaussian process residual term was developed to capture both the global trend of erosion peaks and local nonlinear deviations. Model performance was assessed using leave-one-out cross-validation with MAE, RMSE, MAPE, R^2 , and PICP as evaluation metrics. The results show that, under leave-one-out cross-validation, the proposed MI-UK-GPR model achieved an MAE of 7.10×10^{-5} , an RMSE of 1.29×10^{-4} , a MAPE of 14.53%, an R^2 of 0.9573, and a PICP of 88.33%, outperforming RSM, SVR, and ordinary GPR in terms of overall prediction performance. In addition, for 50 independent operating conditions, the total computational time of parameterized CFD batch simulations was 5083.51 s, whereas the trained MI-UK-GPR model required only 0.004860 s, corresponding to a speedup of approximately 1.05×10^6 . Overall, the proposed method provides a physically consistent, uncertainty-aware, and computationally efficient framework for rapid erosion assessment of cold hydrogenation elbows under multiple operating conditions.

Keywords: cold hydrogenation; gas-solid two-phase flow; elbow erosion; high-fidelity CFD; MI-UK-GPR; rapid prediction



Received: 6 May 2026

Revised: 22 June 2026

Accepted: 24 June 2026

Published: 27 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Polysilicon is a critical foundational material for the photovoltaic and semiconductor industries, and its high-purity production process directly affects the performance and manufacturing cost of downstream devices [1]. At present, industrial polysilicon production still primarily relies on the Modified Siemens Process. This process realizes a closed-loop raw material cycle through the recovery and conversion of the by-product silicon tetrachloride. As a key reaction step in this process, cold hydrogenation plays an essential role in converting silicon tetrachloride into trichlorosilane, thereby significantly improving raw material utilization efficiency and reducing overall process costs [1–4]. Under cold hydrogenation conditions, the reaction medium typically operates at high temperature and high pressure in a gas–solid two-phase environment. Continuous particle impingement on the walls of elbow sections can readily cause localized erosion and wall thinning, thereby compromising the operational reliability of the equipment [2–4]. Owing to the substantial geometric curvature variation of elbow sections, the strong particle deflection and migration, and the complexity of local flow structures, erosion behavior in these regions often exhibits pronounced spatial non-uniformity. Compared with direct in situ measurements, erosion assessment based on high-fidelity numerical simulations can provide more detailed information on wall response; however, the computational cost becomes prohibitive under multiple operating conditions. Therefore, it is necessary to develop a rapid prediction method that combines physical consistency with computational efficiency.

For engineering erosion assessment, empirical and semi-empirical models have long been used because of their simplicity and low computational cost. Representative examples include DNV-RP-O501 and the Tulsa model. DNV-RP-O501 provides recommended-practice guidance for managing sand production and erosion in oil and gas production facilities and has been widely used as an engineering reference for erosion-risk assessment in piping systems [5]. The Tulsa model, represented by the work of Shirazi, Shadley, McLaury, and Rybicki, provides a semi-empirical procedure for predicting solid-particle erosion in typical pipe fittings such as elbows and tees [6]. These models are valuable for preliminary erosion screening, safe-operation assessment, and rapid engineering estimation. Although these empirical and semi-empirical models are useful for rapid engineering estimation, their simplified assumptions make it difficult to fully describe localized erosion distributions and complex parameter coupling in cold hydrogenation elbows. Therefore, experimental investigation and high-fidelity numerical simulation remain important approaches for revealing detailed erosion mechanisms and wall-response characteristics.

At present, extensive research on pipeline erosion has been carried out both in China and internationally, mainly along two lines: experimental investigation and numerical simulation. Experimental studies have generally relied on pipeline erosion test systems to examine erosion mechanisms in gas–solid two-phase flow and related elbow erosion scenarios, whereas numerical studies have primarily focused on evaluating the effects of various factors on erosion rates in elbow sections under gas–solid and liquid–solid two-phase flow conditions. In experimental research, Othayq et al. [7] constructed a series of standard elbow loops, combined with paint film removal and ultrasonic thickness measurements, to compare erosion patterns under gas–sand and multiphase flow conditions, and found that the elbow spacing ratio (L/D) influenced both the erosion magnitude and the erosion location in the second elbow; under gas–sand conditions, the overall erosion of the second elbow was lower than that of the first elbow, consistent with CFD results. Wang et al. [8] experimentally and numerically investigated slurry erosion behavior in a horizontal elbow and elbows in series, providing useful evidence for understanding erosion distribution and particle-impact behavior in elbow structures. Zhou F. et al. [9] carried out orthogonal experiments using 7–15 mm coarse coal particles in a 90° elbow loop

of a pneumatic conveying system and reported that the erosion center was located near the inlet tangent region at approximately $25\text{--}30^\circ$, with elbow type and gas velocity identified as the dominant influencing factors, while typical micro-cutting morphologies were also observed. Liu E.-B. et al. [10] combined experiments with CFD simulations based on the Oka model using elbows in shale gas gathering and transportation systems as the test object, confirming that the Oka model provided the best fit for the interaction between quartz sand and carbon steel and that a spiral structure could reduce erosion by approximately 34% under identical operating conditions. Hong et al. [11] established a gas–solid two-phase erosion test bench and found that erosion was most severe at an impingement angle of approximately 30° ; they further reported that L245 carbon steel exhibited a higher erosion rate than 304 stainless steel because of its lower hardness, and on this basis proposed a modified erosion model to better fit the measured data.

In numerical simulation research, Yu Y. et al. [12] employed dynamic meshing combined with two-way fluid–structure interaction to simulate erosion in high-pressure gas–solid two-phase flow within elbows while simultaneously accounting for wall deformation and stress, and quantified the effects of pressure, velocity, bending radius, bend angle, and wall thickness on erosion and residual strength. Peng et al. [13] experimentally and numerically studied sand-particle erosion under slug-flow conditions in a horizontal pipe bend, showing that multiphase flow patterns can significantly affect particle-impact behavior and local erosion distribution. Mu L. et al. [14] combined CFD with reduced-order modeling to rapidly predict the erosion field in gas–solid two-phase pipe bend systems, and validation based on a 135° industrial pipeline demonstrated that the model could predict both erosion location and erosion rate in a timely and accurate manner while significantly reducing computational cost. Zhao et al. [15] investigated the influence and optimization of swirler parameters on elbow wear in a deep coal fluidization pipeline transportation system, indicating that flow-guiding structures can change particle trajectories and reduce local wear in elbow sections. Li Y. et al. [16] developed a large eddy simulation–discrete phase model for predicting gas–solid two-phase flow in elbows, and comparisons with RNG $k\text{--}\epsilon$ simulations and experimental data showed that increasing inlet velocity led to a monotonic rise in total pressure drop and a reduction in particle motion stability. Liu et al. [17] numerically analyzed solid-particle erosion in elbows under annular-flow conditions and discussed the influence of flow and particle behavior on the erosion characteristics of elbow sections.

In recent years, data-driven machine learning (ML) [18] has provided a wide range of methods for knowledge extraction from data and offers substantially higher computational efficiency than traditional experiments or numerical simulations, thereby attracting considerable attention in intelligent modeling and rapid prediction in fluid mechanics [19]. Wang et al. [20] combined machine learning with swarm intelligence optimization to model the erosion rate of solid particles in elbows, and found that hybrid optimization significantly reduced prediction errors while identifying flow velocity and sand content as the dominant influencing factors. Chen et al. [21] developed an AI-driven prediction model for gas–solid two-phase erosion in elbows and performed hyperparameter optimization, which significantly reduced prediction errors and mitigated overfitting compared with empirical correlations and standalone ML models. Chi et al. [22] applied response surface methodology (RSM) and a genetic algorithm-optimized backpropagation neural network (GA-BP) to predict the maximum erosion rate of elbows in shale gas gathering and transportation systems, and the results showed that GA-BP was more capable of capturing variable interactions and significantly outperformed RSM. Li et al. [23] employed explainable machine learning based on tree models and SHAP to predict gas–solid erosion in series-connected elbows, revealing that inlet velocity and bend radius are the key control-

ling variables and that the model exhibits good generalization ability between upstream and downstream elbows.

Although the above studies have significantly advanced the understanding and prediction of elbow erosion, several challenges remain for cold hydrogenation applications. First, empirical and semi-empirical engineering models are computationally efficient but may not fully capture the coupled effects of cold hydrogenation geometry, gas–solid flow conditions, and localized particle impingement. Second, high-fidelity CFD simulations can provide detailed erosion distributions and physically consistent erosion-response data, but their computational cost becomes prohibitive when a large number of geometric and operating conditions must be evaluated. Third, purely data-driven machine learning models can improve prediction efficiency, but they generally require sufficient calibration data and may lack physical interpretability under small-sample conditions. Finally, many existing prediction methods mainly provide deterministic point estimates, whereas uncertainty information is also important for engineering risk assessment and high-risk operating-condition identification. Therefore, there remains a need for a rapid erosion prediction method that can combine CFD-based physical consistency, small-sample learning capability, computational efficiency, and uncertainty quantification.

To address these issues, this study focuses on the feed elbow of a cold hydrogenation reactor in a polysilicon chemical plant in Xinjiang. Under long-term high-pressure operating conditions involving hydrogen–particle interactions, this pipeline is highly susceptible to repeated impingement by high-velocity particles, which gradually removes surface material and thereby compromises equipment safety and service life. Moreover, accurate wall-thickness measurement on curved surfaces such as elbows remains difficult, and manual inspection under high-pressure, hydrogen-rich conditions poses considerable safety risks, resulting in limited accumulation of wall-thinning data. From the perspective of plant operation, erosion-induced elbow failure can trigger unplanned shutdowns, which not only cause substantial direct economic losses but also disrupt the overall production schedule through capacity interruption. Therefore, this study first establishes a CFD-based gas–solid two-phase erosion model to obtain the maximum erosion rate, E_{max} , on the inner wall of the elbow, and subsequently constructs a training dataset under multiple operating conditions to develop an MI-UK-GPR surrogate model for predicting E_{max} . By reducing the prediction cost under multiple operating conditions while preserving physical consistency, this study provides a reproducible methodological framework for rapid erosion assessment of cold hydrogenation elbows, parameter screening, and the identification of operating conditions associated with relatively high erosion rates. The technical roadmap of this study is shown in Figure 1.

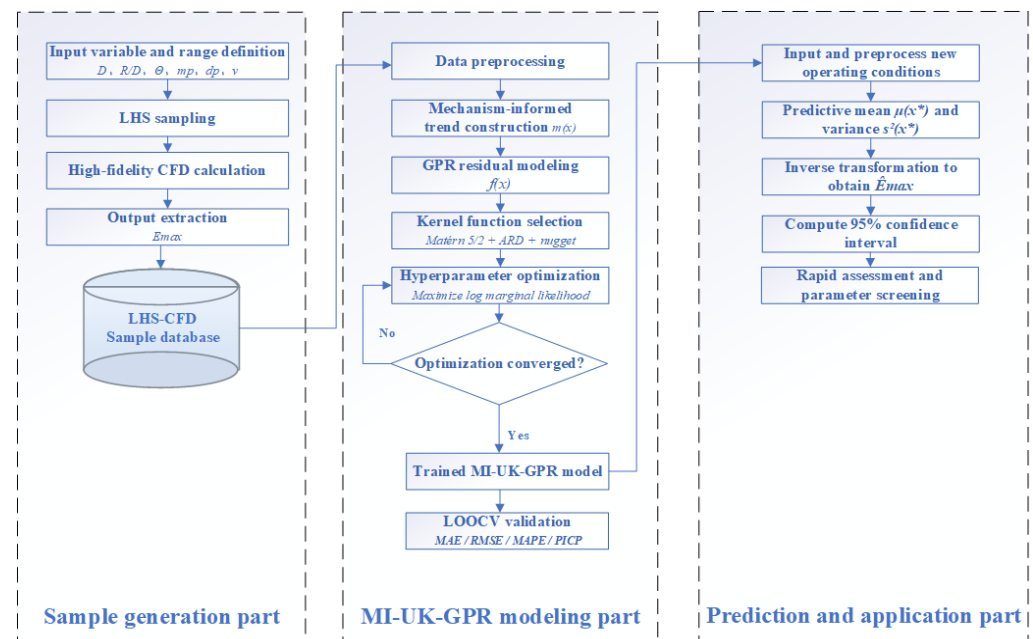


Figure 1. Technical roadmap for erosion rate prediction in cold hydrogenation elbows based on the MI-UK-GPR model. Note: x^* denotes a new input vector corresponding to an operating condition to be predicted.

2. Methodology

2.1. Overall Workflow

Following the technical roadmap presented in Figure 1, the methodology of this study comprises three major stages: generation of erosion-response data using a validated CFD model, construction and preprocessing of the LHS-CFD sample database, and development of the MI-UK-GPR surrogate model for rapid prediction. The detailed procedures for geometric parameterization, numerical simulation, database construction, model formulation, and performance evaluation are described in the following subsections.

2.2. Numerical Simulation of Erosion Rates

To improve modeling efficiency and focus on the dominant mechanism of particle inertia-driven erosion, this study neglects chemical reactions and compositional variations by treating the continuous phase as an isothermal equivalent gas. This treatment preserves the first-order effects of fluid density and viscosity on particle drag and on the flow structure within the elbow, while excluding the influences of exothermic reactions, species transport, and phase change. Based on production data from the cold hydrogenation unit of a representative polysilicon plant, the ranges of six factors were defined, including pipe inner diameter D , curvature ratio R/D , bend angle θ , particle mass flow rate $\dot{m}_p$, particle diameter d_p , and particle velocity v . The detailed parameter settings are listed in Table 1.

Table 1. Parameters of Factors Considered in the Simulation.

Factor	Symbol	Range	Unit
Pipe inner diameter	D	50–200	mm
Curvature ratio	R/D	1.5–5.0	—
Bend angle	θ	45–90	°
Particle mass flow rate	$\dot{m}_p$	0.01–0.30	kg/s
Particle diameter	d_p	80–250	μm
Particle velocity	v	12–30	m/s

2.2.1. Mathematical Model

Turbulence Model

This study adopts the SST k - ω turbulence model [24] for the numerical simulations. The corresponding governing equations are presented in Equations (1)–(3).

$$\frac{\partial(\rho k)}{\partial t} + \nabla \cdot (\rho \mathbf{u} k) = \nabla \cdot \left[\left(\mu + \frac{\mu_t}{\sigma_k} \right) \nabla k \right] + P_k - \beta^* \rho k \omega \quad (1)$$

$$\frac{\partial(\rho \omega)}{\partial t} + \nabla \cdot (\rho \mathbf{u} \omega) = \nabla \cdot \left[\left(\mu + \frac{\mu_t}{\sigma_\omega} \right) \nabla \omega \right] + \alpha \frac{\omega}{k} P_k - \beta \rho \omega^2 + 2(1 - F_1) \frac{\rho \sigma_{\omega 2} \nabla k \cdot \nabla \omega}{\omega} \quad (2)$$

$$\mu_t = \rho \frac{a_1 k}{\max(a_1 \omega, S F_2)} \quad (3)$$

where ρ is the fluid density, in kg/m^3 ; $\mathbf{u}$ denotes the velocity vector, in m/s ; t is time, in s ; μ is the dynamic viscosity, in $\text{Pa} \cdot \text{s}$; μ_t is the turbulent viscosity, in $\text{Pa} \cdot \text{s}$; k is the turbulent kinetic energy, in m^2/s^2 ; ω is the specific dissipation rate, in s^{-1} ; σ_k and σ_ω are the turbulent Prandtl numbers associated with the k and ω equations, respectively; P_k is the production term of turbulent kinetic energy induced by the mean velocity gradient; β^* and β are empirical model constants controlling the dissipation terms in the k and ω equations, respectively; α is the coefficient of the production term in the ω equation; F_1 and F_2 are blending functions used to ensure a smooth transition between the near-wall region and the free-stream region; a_1 is the model constant used to limit the turbulent viscosity; S is the magnitude of the strain rate; ∇ denotes the gradient operator; and $\nabla \cdot$ denotes the divergence operator.

Force Equations for the Discrete Phase

Discrete-phase particles in the flow field are subjected to multiple forces, including drag, lift, gravity, and virtual mass force, and their motion is governed by Equations (4)–(10).

$$\frac{d\mathbf{u}_p}{dt} = \mathbf{F}_{\text{drag}} + \mathbf{F}_{\text{lift}} + \mathbf{F}_{\text{gravity}} + \mathbf{F}_{\text{vm}} \quad (4)$$

$$\mathbf{F}_{\text{drag}} = \frac{18\mu}{\rho_p d_p^2} \cdot \frac{C_D \text{Re}_p}{24} \cdot (\mathbf{u} - \mathbf{u}_p) \quad (5)$$

$$C_D = \frac{24}{\text{Re}_p} \left(1 + 0.15 \text{Re}_p^{0.687} \right) \quad (6)$$

$$\text{Re}_p = \frac{\rho_p d_p |\mathbf{u} - \mathbf{u}_p|}{\mu} \quad (7)$$

$$\mathbf{F}_{\text{lift}} = 1.615 \cdot d_p^2 \cdot \sqrt{\mu \rho} \cdot \nabla \times \mathbf{u} \quad (8)$$

$$\mathbf{F}_{\text{gravity}} = \rho_p \cdot V_p \cdot \mathbf{g} \quad (9)$$

$$V_p = \frac{\pi d_p^3}{6} \quad (10)$$

In these equations, $\mathbf{u}_p$ denotes the particle velocity, in m/s ; $\mathbf{F}_{\text{drag}}$ is the drag force exerted by the gas flow on the particle; μ is the gas viscosity, in $\text{Pa} \cdot \text{s}$; ρ_p is the particle density, in kg/m^3 ; d_p is the particle diameter, in m ; C_D is the drag coefficient; Re_p is the particle Reynolds number; $\mathbf{F}_{\text{lift}}$ is the lift force acting on the particle in the flow field; $\nabla \times \mathbf{u}$ represents the vorticity of the gas flow; $\mathbf{F}_{\text{gravity}}$ is the gravitational force acting on the particle; V_p is the particle volume; $\mathbf{g}$ is the gravitational acceleration; and $\mathbf{F}_{\text{vm}}$ represents the effect associated with fluid acceleration around the particle. For particles with relatively high density, the virtual mass force is generally negligible.

Oka Erosion Model

For erosion prediction in cold hydrogenation feed elbows, the Oka erosion model [25,26] was adopted. Originally proposed by Oka et al. [25,26] at Hiroshima University in 2005, the model is described by Equations (11) and (12).

$$R_E(\chi) = J_p(\chi) C \left(\frac{V}{V_{ref}} \right)^n f(\theta) \left(\frac{d}{d_{ref}} \right)^m \left(\frac{H_w}{H_{ref}} \right)^{-k} \quad (11)$$

$$f(\theta) = \frac{\sin^a \theta \cdot \cos^b \theta}{\left(\frac{a}{a+b} \right)^{\frac{a}{2}} \cdot \left(\frac{b}{a+b} \right)^{\frac{b}{2}}} \quad (12)$$

where $R_E(\chi)$ is the erosion mass flux, in $\text{kg}/\text{m}^2\cdot\text{s}$; J_p is the particle impact mass flux, in $\text{kg}/(\text{s}\cdot\text{m}^2)$; V is the particle impact velocity, in m/s ; $f(\theta)$ is the angle function; d is the particle diameter, in m ; H_w is the hardness of the target material, in GPa ; and C , n , m , k , a , b , V_{ref} , H_{ref} are model constants.

2.2.2. Geometric Model and Mesh Generation

This study takes the feed elbow of a cold hydrogenation reactor in a polysilicon chemical plant in Xinjiang as the prototype for geometric modeling. The elbow specifications are $\text{DN}100 \times \text{R}400 \times 90^\circ$. Straight pipe sections with a length of five diameters were arranged upstream and downstream of the elbow to minimize the direct influence of boundary conditions on the local secondary flow and particle impingement behavior within the elbow, while maintaining a consistent relative development length across different geometric configurations. The geometry was created using ANSYS SpaceClaim 2022R1. To ensure the accuracy of flow-field and particle erosion calculations in the elbow region, a structured sweep meshing method was adopted. Local mesh refinement was applied in the elbow region to improve the resolution of secondary flow structures and erosion hotspots. Mesh quality was evaluated using skewness and orthogonal quality. To provide a more complete description of the mesh quality, the minimum, maximum, average, and standard deviation values of these two criteria were extracted from the selected mesh used for the subsequent CFD-DPM simulations. The detailed mesh-quality statistics are summarized in Table 2. The skewness ranged from 5.522×10^{-3} to 0.4088, with an average value of 7.961×10^{-2} and a standard deviation of 6.8384×10^{-2} . The orthogonal quality ranged from 0.59341 to 0.99975, with an average value of 0.99147 and a standard deviation of 2.3942×10^{-2} . These results indicate that the generated mesh exhibited low element distortion and good orthogonality, satisfying the quality requirements for resolving the gas–solid two-phase flow and particle-induced erosion behavior in the elbow region. The geometric model and the corresponding mesh are shown in Figure 2.

Table 2. Mesh-quality statistics of the selected mesh used in the CFD-DPM erosion simulation.

Mesh-Quality Criterion	Minimum	Maximum	Average	Standard Deviation
Skewness	0.005522	0.4088	0.07961	0.068384
Orthogonal quality	0.59341	0.99975	0.99147	0.023942

2.2.3. Boundary Condition Settings and Material Parameters

The continuous phase was treated as hydrogen with constant thermophysical properties under operating conditions of 3.2 MPa and 80 °C to represent the primary flow characteristics of the cold hydrogenation feed gas. No-slip boundary conditions were imposed on the flow field, and the near-wall region was treated using the automatic wall-function approach of the turbulence model. A velocity inlet boundary condition was specified at the inlet, a pressure outlet boundary condition was applied at the outlet, and a

wall boundary condition was imposed on the pipe wall. The discrete phase consisted of silicon powder particles with a density of 2330 kg/m^3 . Since the particle volume fraction was less than 1%, inter-particle collisions were neglected. In the DPM model, escape conditions were specified at both the inlet and outlet, whereas reflect conditions were imposed at the wall. The particle injection was modeled using a surface injection source, with particles randomly released from the inlet boundary, and the initial particle velocity was set equal to the fluid inlet velocity. Particle–wall collisions were described using the Grant and Tabakoff reflection model. During the FLUENT simulations, the SIMPLER algorithm was adopted for pressure–velocity coupling, and second-order upwind schemes were used for the discretization of momentum, turbulent kinetic energy, and specific dissipation rate.

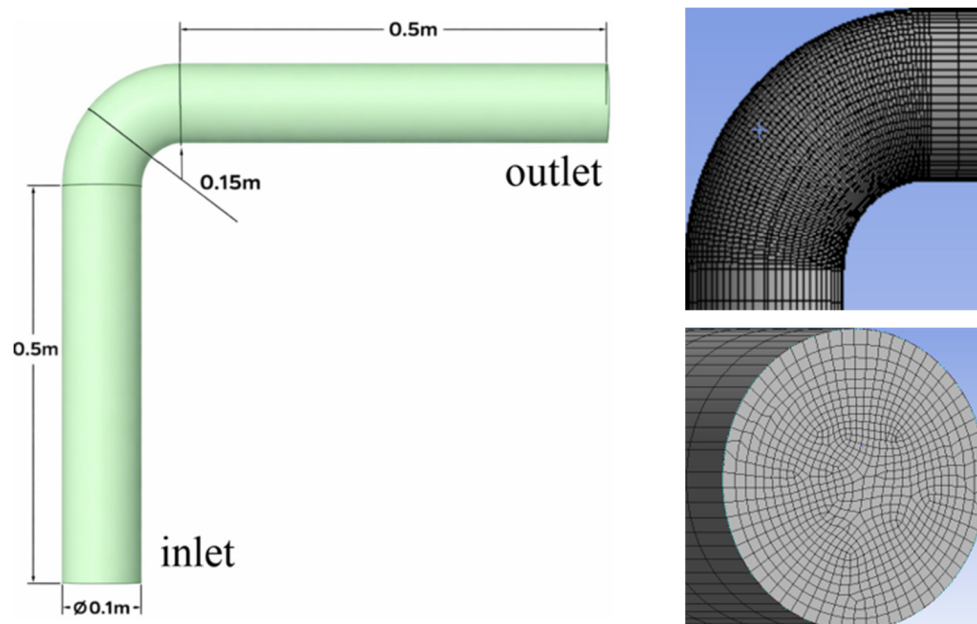


Figure 2. Geometric Modeling and Mesh Generation of the Bent Pipe.

In addition, the material properties used for the wall and discrete phase were specified to ensure consistency in the CFD-DPM erosion calculation. The actual elbow material used in the cold hydrogenation conveying system was 316 stainless steel, while the discrete phase was silicon powder. In the Oka erosion model implemented in Fluent, the wall Vickers hardness, (H_w), was set to 1.8 GPa, which was consistently used for all CFD cases in the construction of the LHS-CFD sample database. The relevant wall and particle material parameters are summarized in Table 3. It should be noted that the present simulation mainly focuses on mechanical erosion caused by particle impingement, while chemical corrosion, material degradation during long-term service, and temperature-dependent variation of wall material properties were not considered.

Table 3. Material parameters used in the CFD-DPM erosion simulation.

Category	Parameter	Symbol	Value	Unit	Description
Wall material	Elbow material	—	316 Stainless steel	—	Actual material of the cold hydrogenation conveying elbow
Wall material	Density	ρ_w	8000	kg/m^3	Engineering value for 316 stainless steel
Wall material	Wall Vickers hardness	H_w	1.8	GPa	Hardness parameter used in the Fluent Oka erosion model
Wall material	Young's modulus	E_w	193	GPa	Engineering value for 316 stainless steel

Table 3. Cont.

Category	Parameter	Symbol	Value	Unit	Description
Wall material	Poisson's ratio	ν_w	0.3	—	Engineering value for 316 stainless steel
Discrete phase	Particle material	—	Silicon powder	—	Solid particles in the cold hydrogenation feed gas
Discrete phase	Particle density	ρ_p	2330	kg/m ³	Same value as that used in the DPM simulation
Discrete phase	Particle diameter	d_p	80–250	μm	Same range as that listed in Table 1

2.2.4. Verification of Mesh Independence

Because the objective of this study is to predict and compare the maximum erosion rate, E_{max} , on the inner wall of the elbow—a quantity highly sensitive to local mesh resolution— E_{max} was adopted as the primary criterion, while the inlet–outlet pressure drop, ΔP , was used as the secondary criterion for mesh independence verification. To ensure comparability among different mesh systems, multiple sets of structured meshes with progressively increasing density were generated by adjusting the number of layers in the sweep direction, while keeping the geometric model, mesh topology, and boundary conditions unchanged. To quantify the influence of mesh refinement on the numerical results, the variations in the maximum erosion rate, E_{max} , on the inner wall of the elbow and in the inlet–outlet pressure drop, ΔP , were compared across different meshes. Here, E_{max} is defined as the maximum value of the wall erosion-rate field, as given in Equation (13).

$$E_{max} = \max_{s \in \Omega_w} E(s) \quad (13)$$

where Ω_w denotes the inner-wall region of the elbow, and $E(s)$ is the erosion rate at wall position s . The mesh independence criterion is evaluated in terms of relative change rates. For any two mesh levels, the relative differences in the maximum erosion rate and pressure drop are defined by Equation (14).

$$\varepsilon_E = \frac{|E_{max}^{(i)} - E_{max}^{(i-1)}|}{E_{max}^{(i)}} \times 100\% \quad \varepsilon_P = \frac{|\Delta P^{(i)} - \Delta P^{(i-1)}|}{\Delta P^{(i)}} \times 100\% \quad (14)$$

where the superscript i denotes the finer mesh level, ε_E denotes the relative change rate of E_{max} , and ε_P denotes the relative change rate of ΔP . The specific results are presented in Table 4, which lists the number of elements, pressure drop, maximum erosion rate, and the corresponding relative change rates for each mesh level, thereby allowing the variation of E_{max} with mesh refinement to be examined.

The calculation results indicate that, with progressive mesh refinement, the pressure drop, ΔP , remains approximately constant at around 176 Pa, and the variation rates between adjacent mesh levels are relatively small. Specifically, the values of ε_P for G2–G6 are 0.6186%, 0.0379%, 0.1211%, 0.1529%, and 0.0067%, respectively, indicating that the overall flow resistance is no longer sensitive to mesh refinement. Meanwhile, as shown in Figure 3, the location of the maximum erosion rate remains essentially unchanged under different mesh configurations, demonstrating good consistency in the spatial distribution of the primary erosion zone. Figure 4 presents the maximum erosion rate obtained under different mesh conditions. As the number of mesh elements increases from 98,397 to 196,267, the maximum erosion rate rises markedly from $7.81 \times 10^{-5} \text{ kg} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ to $9.90 \times 10^{-5} \text{ kg} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$, with the corresponding relative change rates between adjacent mesh levels, ε_E , being 14.6244%, 4.4071%, and 3.4280%, respectively. This indicates that, within the relatively coarse mesh range, the prediction of the erosion peak is still affected by mesh resolution.

However, when the mesh density is further increased to 240,487 and 273,277 cells, E_{max} changes only slightly, from $1.019 \times 10^{-4} \text{ kg}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ to $1.022 \times 10^{-4} \text{ kg}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, while the corresponding relative change rates decrease to 2.7905% and 0.2573%, indicating that the numerical error has largely converged. A comprehensive examination of the variation characteristics of ΔP and E_{max} shows that, once the mesh reaches level G5, further refinement yields only marginal improvements in the overall results while significantly increasing the computational cost. Therefore, the G5 mesh satisfies the accuracy requirements for elbow erosion prediction in this study, and all subsequent model validation and result analysis are based on this mesh.

Table 4. Results of the mesh independence analysis for the reference elbow model.

Mesh Level	Cells	$\Delta P/\text{Pa}$	$E_{max}/(\text{kg}\cdot\text{m}^{-2}\cdot\text{s}^{-1})$	$\varepsilon_P/\%$	$\varepsilon_E/\%$
G1	98,397	176.9719	7.809995×10^{-5}	—	—
G2	142,117	175.8839	9.14781×10^{-5}	0.6186	14.6244
G3	172,721	175.8173	$9.56955384 \times 10^{-5}$	0.0379	4.4071
G4	196,267	176.0305	9.900924×10^{-5}	0.1211	3.4280
G5	240,487	175.7617	1.01937×10^{-4}	0.1529	2.7905
G6	273,277	175.7500	1.022×10^{-4}	0.0067	0.2573

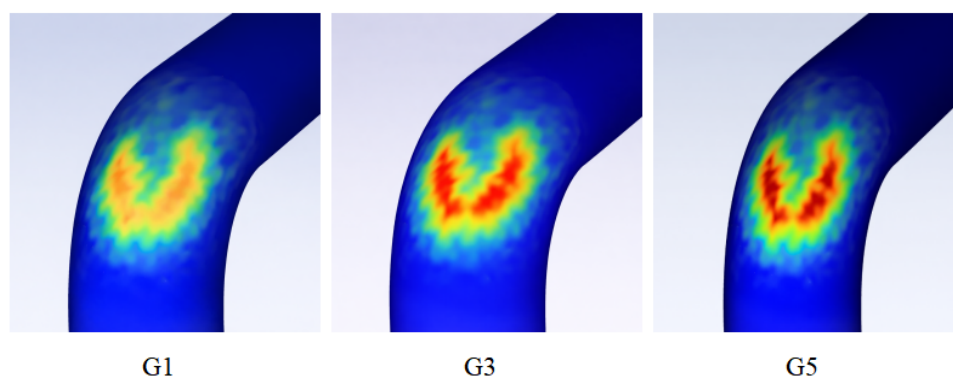


Figure 3. Locations of erosion hotspots under different mesh levels (G1, G3, and G5 as examples). The color scale represents erosion intensity, where blue indicates low erosion and red indicates high erosion.

2.2.5. Model Validation

After completing the mesh independence analysis, the reliability of the numerical model developed in this study was further validated using the experimental and numerical results reported by Liu et al. [10] on the erosion behavior of elbows in shale gas gathering and transportation systems as an external reference case. In that study, the 50–70° region on the outer-curvature side of the elbow was first identified as the primary erosion zone through pre-erosion experiments. Subsequently, nine specimens were extracted from this region, and the local erosion rates were measured using the gravimetric method. In addition, four erosion models, namely the Generic, Finnie, McLaury, and Oka models, were compared, and the results showed that the Oka model was in the closest agreement with the experimental data. Since the research object of the present study does not fully coincide with that reported in the literature in terms of medium composition and operating conditions, the validation in this section does not aim to reproduce the absolute operating conditions exactly, but rather to examine whether the proposed model can adequately characterize the distribution pattern of the primary erosion zone in the elbow. To this end, following the specimen partitioning strategy adopted by Liu et al., the primary erosion zone on the outer side of the elbow was divided into nine corresponding regions, and

the average erosion loss of each specimen region was extracted and compared with the experimental values reported in the literature.

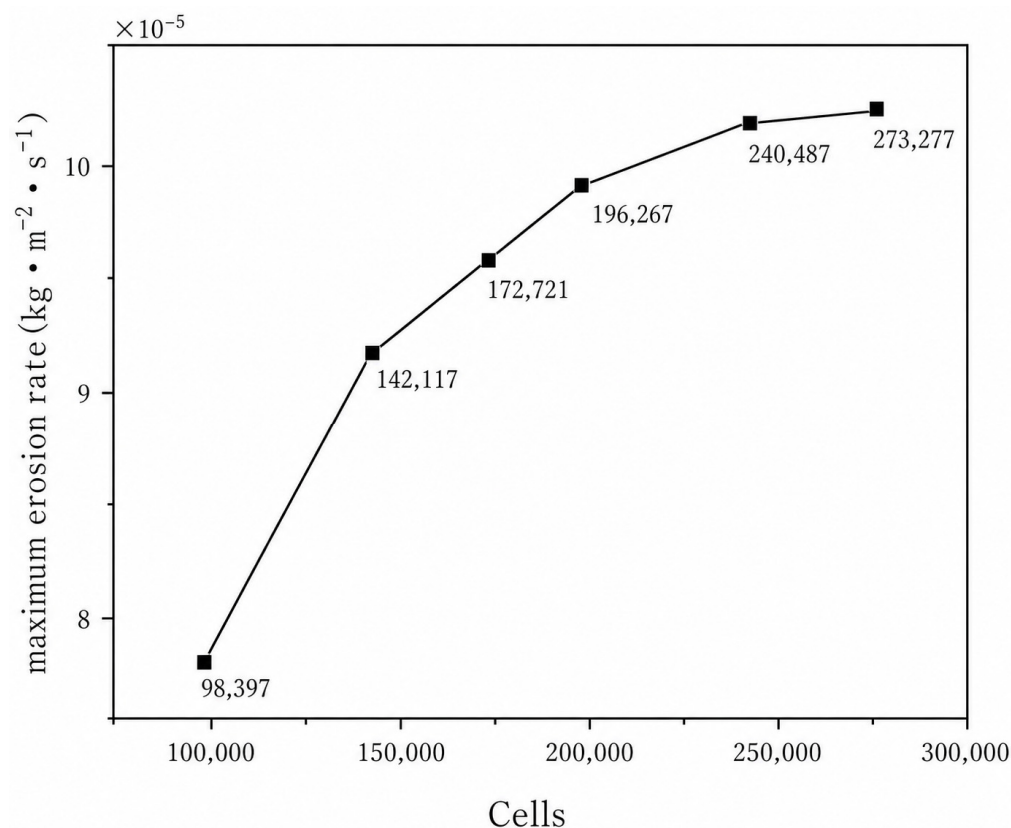


Figure 4. Maximum erosion rates under different mesh sizes.

Figure 5 and Table 5 present a comparison between the experimental data obtained at different specimen locations and the simulation results of the present study. It can be seen that both the experimental and numerical results exhibit a trend of first increasing and then decreasing, with the peak values located in the central primary erosion zone, indicating that the proposed model can effectively reproduce the spatial distribution characteristics of the main erosion zone on the outer side of the elbow. The average relative error across the nine specimen locations was 12.48%. Moreover, the errors were generally smaller for the specimens located within the primary erosion zone, suggesting that the model provides relatively stable predictions in the peak erosion region. Previous studies have shown that, within the 50–90° range of the primary erosion zone, the relative error of the Oka model with respect to experimental results is approximately between −11% and +18%, indicating good predictive performance [10]. The results obtained in this study are generally consistent with this error range, and the locations of the peak-value specimens agree well with the experimental observations. It should be emphasized that the purpose of the validation in this section is to verify the capability of the model to characterize the distribution pattern of the primary erosion zone in gas–solid elbows, rather than to directly equate the absolute erosion values reported in the literature with the actual operating conditions of cold hydrogenation. Considering the consistency in peak location, distribution trend, and overall error level, the CFD model established in this study is considered to be reliable and capable of meeting the requirements for subsequent parametric sample generation and surrogate model development.

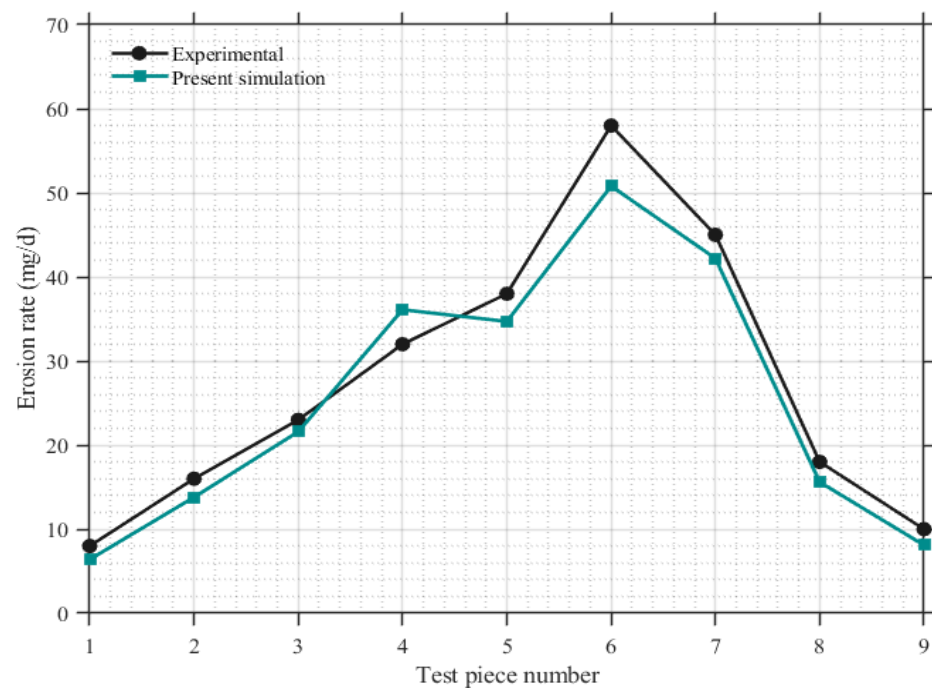


Figure 5. Comparison between experimental erosion rates and numerical predictions at different specimen locations.

Table 5. Comparison of experimental and simulated erosion values at different specimen locations.

Test Piece Number	Experimental Value/(mg·d ^{−1})	Present Simulation/(mg·d ^{−1})	Relative Error/%
1	8	6.4	20.00
2	16	13.8	13.75
3	23	21.6	6.09
4	32	36.1	12.81
5	38	34.7	8.68
6	58	50.8	12.41
7	45	42.2	6.22
8	18	15.6	13.33
9	10	8.1	19.00
MRE		12.48	

2.3. Sample Construction and Data Preprocessing

The predictive performance of a surrogate model depends strongly on the coverage of the training samples and the quality of data representation. Considering that the maximum erosion rate, E_{max} , of elbows is jointly governed by geometric parameters and particle-related conditions, while high-fidelity CFD simulations are computationally expensive, this study employs Latin hypercube sampling to construct a representative LHS-CFD sample library. On this basis, the input and output variables are further preprocessed to improve the numerical stability of subsequent model training.

2.3.1. Construction of the LHS-CFD Sample Library

The maximum erosion rate, E_{max} , of an elbow is jointly governed by geometric parameters and particle-flow conditions. To obtain training samples with uniform coverage over the engineering parameter space, this study employs Latin hypercube sampling (LHS) to construct input–output sample pairs. Compared with simple random sampling, LHS stratifies the marginal distribution of each input variable with equal probability and then

randomly combines the samples, thereby improving the coverage of the parameter space under a limited sample size [27]. Based on the variable ranges specified in Table 1, six input factors are considered in this study: pipe inner diameter D (50–200 mm), curvature ratio R/D (1.5–5.0), bend angle θ (45–90°), particle mass flow rate $\dot{m}_p$ (0.01–0.30 kg/s), particle diameter d_p (80–250 μm), and particle velocity v (12–30 m/s). For each LHS sample point, the previously validated CFD–DPM erosion model is employed to calculate the wall erosion-rate field $E(s)$, and the target label is defined as the maximum erosion rate over the inner-wall region Ω of the elbow, thereby yielding the training dataset $\mathcal{D} = \{(x_i, E_{\max,i})\}_{i=1}^{60}$. The LHS design and the corresponding CFD simulation results are summarized in Table 6. Because each sample requires a high-fidelity CFD–DPM simulation to obtain its label, sample construction is computationally expensive. Therefore, to balance the coverage of the six-dimensional parameter space against computational cost, this study constructs a small-sample surrogate modeling dataset consisting of 60 LHS samples and further improves sample efficiency through leave-one-out cross-validation (LOOCV). This dataset preserves the physical consistency of high-fidelity numerical simulations while providing a controllable and reproducible data basis for subsequent surrogate model training.

Table 6. LHS sample set and corresponding CFD simulation results.

Sample_ID	D (mm)	R/D (-)	θ (deg)	$\dot{m}_p$ (kg/s)	d_p (μm)	v (m/s)	$E_{\max}$ ($\text{kg}/(\text{m}^2 \cdot \text{s})$)
1	148	2.52	62.4	0.103	88.23	24.6	0.000149805
2	78	3.55	63.7	0.129	84.62	16.2	0.000308447
3	101	4.84	67.2	0.110	214.43	21.0	0.000232976
4	108	2.80	56.4	0.293	142.42	29.1	0.001204266
5	81	1.72	66.4	0.242	124.90	28.6	0.001918827
6	131	4.71	50.8	0.298	152.34	13.7	0.000343280
7	125	3.92	68.8	0.065	146.76	20.2	0.000102784
8	68	2.65	71.4	0.151	98.74	20.0	0.000661970
9	63	4.06	75.6	0.058	114.60	23.8	0.000402326
10	111	3.44	88.9	0.017	190.10	13.2	0.000034354
11	96	1.94	76.8	0.190	174.54	23.2	0.000707484
12	99	1.66	52.5	0.201	246.56	28.1	0.001023334
13	50	2.36	54.0	0.270	249.17	16.9	0.003942515
14	191	2.15	65.5	0.095	176.86	23.6	0.000139717
15	170	4.94	61.0	0.114	243.88	17.8	0.000122076
16	177	3.68	60.4	0.226	141.48	14.5	0.000102013
17	164	2.02	73.4	0.213	148.26	24.4	0.000353906
18	62	1.50	67.7	0.137	231.73	12.2	0.000822569
19	127	3.61	70.9	0.082	223.63	21.4	0.000141180
20	77	4.31	48.9	0.277	226.11	19.6	0.001376579
21	86	3.33	59.2	0.163	196.05	13.4	0.000543182
22	154	1.90	58.4	0.205	184.35	21.9	0.000339338
23	114	4.40	74.6	0.126	88.62	21.2	0.000181821
24	118	3.83	69.4	0.143	102.84	27.4	0.000374215
25	59	4.17	80.8	0.036	158.09	26.1	0.000341969
26	158	1.80	45.6	0.220	187.36	29.8	0.000710249
27	117	4.44	87.2	0.059	235.53	26.4	0.000151394
28	54	1.79	88.1	0.090	154.15	25.6	0.001483421
29	189	2.73	51.5	0.157	113.28	27.0	0.000237895
30	172	3.01	57.4	0.257	91.64	21.7	0.000267457
31	198	3.02	59.6	0.023	137.31	20.5	0.000016030
32	174	2.31	48.2	0.212	117.01	27.8	0.000334958

Table 6. Cont.

Sample_ID	D (mm)	R/D (-)	θ (deg)	$\dot{m}_p$ (kg/s)	d_p (μm)	v (m/s)	E_{max} ($\text{kg}/(\text{m}^2 \cdot \text{s})$)
33	142	2.91	65.1	0.173	228.47	18.9	0.000337073
34	197	2.88	47.4	0.194	126.75	26.4	0.000194804
35	185	4.11	85.4	0.099	102.32	14.8	0.000030042
36	83	3.16	78.6	0.045	199.81	28.3	0.000237101
37	57	4.99	46.6	0.076	181.15	17.3	0.000650203
38	143	2.67	80.0	0.180	213.10	25.3	0.000365452
39	138	3.13	76.4	0.177	128.26	16.6	0.000268749
40	90	4.49	84.4	0.166	96.95	15.9	0.000327629
41	178	4.25	53.1	0.288	120.78	13.8	0.000143159
42	156	2.41	82.2	0.134	191.49	26.7	0.000246810
43	121	4.81	86.8	0.282	106.02	14.4	0.000366127
44	105	2.57	83.1	0.011	81.95	18.6	0.000016129
45	180	3.24	72.5	0.232	133.51	18.3	0.000216204
46	152	2.25	86.2	0.254	202.74	25.1	0.000487294
47	187	4.56	54.1	0.040	170.54	19.4	0.000032768
48	147	3.29	64.1	0.148	218.69	23.0	0.000241703
49	134	3.89	83.4	0.053	162.38	15.6	0.000071533
50	135	2.03	78.8	0.025	196.79	12.4	0.000033839
51	166	2.13	73.8	0.070	205.07	12.8	0.000051204
52	162	1.58	89.8	0.033	166.48	22.4	0.000052848
53	74	4.63	55.2	0.274	161.09	22.6	0.001425215
54	72	2.44	77.3	0.264	135.74	15.1	0.001196203
55	93	3.74	55.6	0.186	208.54	24.2	0.000621809
56	106	4.70	61.7	0.117	110.79	18.7	0.000215292
57	67	3.38	50.2	0.233	221.28	17.4	0.001444117
58	88	3.97	45.9	0.243	170.82	29.6	0.001108505
59	194	3.54	70.0	0.087	239.90	16.0	0.000070277
60	129	4.21	81.2	0.248	237.27	28.9	0.000502474

2.3.2. Parameter-Space Coverage and Sample-Size Rationale

To further verify the representativeness of the 60 LHS-CFD samples, the coverage of the sampled parameter space was analyzed from two aspects: the distribution range of each input variable and the statistical contribution of each factor to the erosion response. As shown in Table 7, the sampled values of the six input variables were distributed over nearly the entire predefined engineering range. Specifically, the sampled pipe inner diameter, (D), ranged from 50 to 198 mm, the curvature ratio, (R/D), ranged from 1.50 to 4.99, the bend angle, (θ), ranged from 45.6° to 89.8° , the particle mass flow rate, ($\dot{m}_p$), ranged from 0.011 to 0.298 kg/s, the particle diameter, (d_p), ranged from 81.95 to 249.17 μm , and the particle velocity, (v), ranged from 12.2 to 29.8 m/s. These ranges are close to the lower and upper bounds listed in Table 1, indicating that the LHS design provided broad coverage of the six-dimensional parameter space under a limited number of CFD simulations.

Table 7. Coverage of the LHS-CFD samples within the predefined parameter space.

Input Variable	Symbol	Predefined Range	Sampled Range	Unit
Pipe inner diameter	D	50–200	50–198	mm
Curvature ratio	R/D	1.5–5.0	1.50–4.99	—
Bend angle	θ	45–90	45.6–89.8	$^\circ$
Particle mass flow rate	$\dot{m}_p$	0.01–0.30	0.011–0.298	kg/s
Particle diameter	d_p	80–250	81.95–249.17	μm
Particle velocity	v	12–30	12.2–29.8	m/s

In addition, an ANOVA-based sensitivity analysis was conducted using the logarithmically transformed maximum erosion rate as the response variable, which is consistent with the output preprocessing strategy used in this study. A first-order regression model containing the six input variables was used for the ANOVA screening. The model achieved an R^2 value of 0.9063, indicating that the six selected variables could explain most of the variation in the logarithmic erosion response. The corresponding ANOVA results are listed in Table 8.

Table 8. ANOVA results for the logarithmically transformed maximum erosion rate.

Factor	Sum of Squares	F-Value	p-Value	Contribution
D	5.5422	194.57	2.23×10^{-19}	43.88%
$\dot{m}_p$	4.2964	150.84	3.96×10^{-17}	34.02%
v	0.7721	27.11	3.20×10^{-6}	6.11%
d_p	0.2028	7.12	1.01×10^{-2}	1.61%
R/D	0.1921	6.75	1.21×10^{-2}	1.52%
θ	0.1144	4.02	5.02×10^{-2}	0.91%
Residual	1.5096	—	—	11.95%

The ANOVA results show that (D) and ($\dot{m}_p$) are the two dominant factors affecting the maximum erosion rate, with contribution ratios of 43.88% and 34.02%, respectively. Particle velocity, (v), also has a significant influence, with a contribution ratio of 6.11%. By comparison, (d_p), (R/D), and (θ) show smaller but still observable effects on the erosion response. These results are physically reasonable: under a given particle mass flow rate, a smaller pipe diameter tends to increase the particle concentration and local impact intensity, while a higher particle mass flow rate and particle velocity directly increase the momentum input and impact energy of particles. The residual contribution of 11.95% indicates that nonlinear effects and interactions among the variables still exist, which further supports the use of the MI-UK-GPR surrogate model rather than a purely linear regression model.

The choice of 60 CFD samples was therefore based on a balance among the dimensionality of the input space, the computational cost of high-fidelity CFD-DPM simulations, and the need for adequate coverage of the engineering parameter ranges. For a six-dimensional parameter space, a full-factorial design would require a much larger number of simulations and would be computationally prohibitive. In contrast, the LHS strategy allows each variable range to be stratified and sampled efficiently, thereby improving the space-filling property under a limited sample size. Moreover, leave-one-out cross-validation was adopted in the subsequent model evaluation to make full use of the small-sample dataset. Therefore, the 60-sample LHS-CFD database provides a reasonable compromise between computational feasibility and parameter-space representativeness for the present surrogate-modeling task.

2.3.3. Input and Output Preprocessing

Since the six input variables differ substantially in both physical units and numerical magnitude, min–max normalization is applied to the input variables in this study to avoid the adverse effects of inconsistent scales on model training. For any input variable x_j , its normalized form is given by Equation (15).

$$x_j^* = \frac{x_j - x_{j,min}}{x_{j,max} - x_{j,min}}, j = 1, 2, \dots, 6 \quad (15)$$

where $x_{j,min}$ and $x_{j,max}$ denote the minimum and maximum values of the j -th input variable in the training samples, respectively. After this transformation, each input variable is mapped to the interval $[0, 1]$.

For the output variable E_{max} , considering that the erosion rate is strictly positive and typically exhibits pronounced nonlinear scaling behavior with respect to velocity, particle size, and mass flow rate, a logarithmic transformation is adopted in the modeling process, as given in Equation (16).

$$y = \ln(E_{max}) \quad (16)$$

This transformation, on the one hand, compresses the range of output values and thereby reduces the influence of high-value samples on training stability; on the other hand, it helps prevent the model from producing nonphysical negative predictions in the original space. All subsequent surrogate models are trained and evaluated in the log-transformed output space, and the final results are mapped back to the original physical space through an exponential transformation.

After sample construction and preprocessing, the next step is to develop a surrogate model capable of simultaneously capturing the global pattern of erosion response and local nonlinear deviations. To this end, this study incorporates a mechanistic trend term into the Gaussian process regression (GPR) framework to construct an MI-UK-GPR model for predicting the maximum erosion rate of cold hydrogenation elbows.

2.4. Development of the MI-UK-GPR

Unlike purely black-box surrogate models, the MI-UK-GPR achieves a balance between physical interpretability and predictive accuracy under small-sample conditions by incorporating a mechanism-inspired trend term into the mean structure and using a Gaussian process to model the remaining complex nonlinear deviations. This section describes the model from five perspectives: overall structure, mechanism-inspired trend term, Gaussian process residual term, parameter estimation, and prediction output. All MI-UK-GPR modeling, parameter estimation, leave-one-out cross-validation, and prediction procedures were implemented in MATLAB R2024a.

2.4.1. Overall Model Structure

In surrogate modeling for computer simulations, Kriging/Gaussian process regression (GPR) exhibits strong capabilities in small-sample fitting and uncertainty quantification [28,29]. To address the trend drift and unstable extrapolation commonly encountered in purely data-driven models under small-sample conditions, this study develops a mechanism-informed universal Kriging–Gaussian process regression (MI-UK-GPR) model, in which the output response is decomposed into the sum of a mechanistic trend term and a Gaussian process residual term, as expressed in Equation (17).

$$y(x) = m(x) + f(x) \quad (17)$$

where $m(x)$ represents the interpretable trend structure, which is used to encode the overall direction of variation in the erosion peak with respect to different input factors, whereas $f(x)$ denotes a zero-mean Gaussian process used to characterize the complex nonlinear residual component not captured by the trend term. To preserve the physical interpretability of the mechanistic trend term, $m(x)$ is constructed on the basis of the original physical variables, whereas min–max normalization is applied only to the inputs of the kernel function and to the numerical optimization procedure. This structure balances physical insight with nonparametric flexibility and is therefore well suited to the present small-sample surrogate modeling task. The architecture of the MI-UK-GPR model is shown in Figure 6.

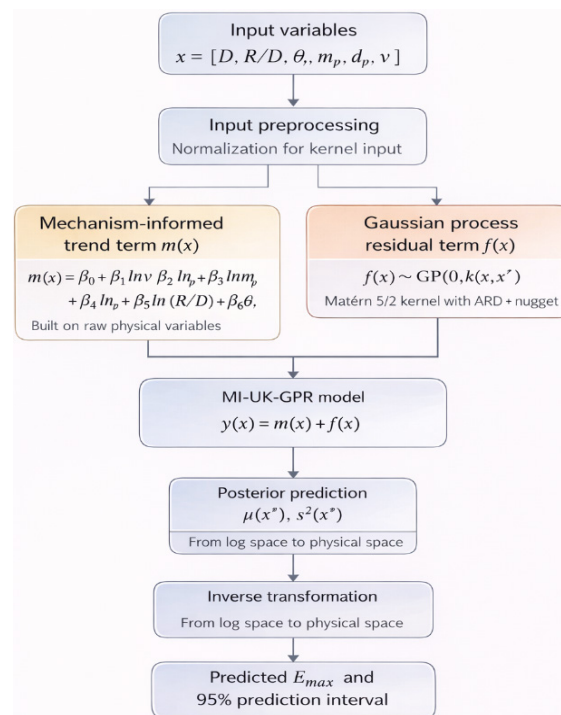


Figure 6. Schematic diagram of the MI-UK-GPR model structure.

2.4.2. Mechanistic Trend Term $m(x)$

The mechanistic trend term is not intended to provide a highly accurate approximation of local details; rather, it is designed to stably capture the direction of macroscopic variation. From the perspective of the erosion mechanism, velocity v , particle diameter d_p , and particle mass flux $\dot{m}_p$ are the primary factors governing particle momentum input and impact energy, whereas the geometric parameters D , R/D , and θ affect the erosion peak by modifying the degree of particle deflection, the distribution of impingement angles, and the focusing effect on the outer-curvature side. Therefore, this study adopts the following low-order trend form in logarithmic space, as given in Equation (18).

$$m(x) = \beta_0 + \beta_1 \ln v + \beta_2 \ln d_p + \beta_3 \ln \dot{m}_p + \beta_4 \ln D + \beta_5 \ln(R/D) + \beta_6 \theta_r \quad (18)$$

Here, $\theta_r = \theta\pi/180$ denotes the bend angle in radians, and $\beta_0 \sim \beta_6$ are the parameters to be estimated.

2.4.3. Gaussian Process Residual Term $f(x)$

The residual term is modeled as a zero-mean Gaussian process, as expressed in Equation (19).

$$f(x) \sim \mathcal{GP}(0, k(x, x')) \quad (19)$$

The Matérn 5/2 kernel is adopted, and its automatic relevance determination (ARD) form is given in Equation (20).

$$k(x, x') = \sigma_f^2 \left(1 + \sqrt{5}r + \frac{5}{3}r^2 \right) \exp(-\sqrt{5}r), \quad r = \sqrt{\sum_{j=1}^6 \frac{(x_j - x'_j)^2}{\ell_j^2}} \quad (20)$$

where σ_f^2 is the signal variance, and ℓ_j^2 is the characteristic correlation length scale associated with the j -th input dimension. In addition, to improve numerical stability and account for

CFD discretization errors or iterative noise, a nugget term, σ_n^2 , is added to the diagonal of the covariance matrix, as given in Equation (21).

$$K = K_f + \sigma_n^2 I \quad (21)$$

Here, K_f denotes the covariance matrix constructed from the kernel function, and I is the identity matrix. This treatment prevents the kernel matrix from becoming singular and improves the robustness of model training under small-sample conditions.

2.4.4. Parameter Estimation

The parameters to be estimated in the MI-UK-GPR model include the trend coefficients, $\sigma_f^2, \ell_1, \dots, \ell_6, \sigma_n^2$. In this study, parameter training is performed by maximizing the log marginal likelihood [29]. Let the output vector of the training set be y , and let the trend design matrix be H . Then, the model satisfies Equation (22):

$$y \sim \mathcal{N}(H\beta, K) \quad (22)$$

The corresponding negative log marginal likelihood can be written as Equation (23):

$$\mathcal{L} = \frac{1}{2}(y - H\beta)^\top K^{-1}(y - H\beta) + \frac{1}{2} \ln |K| + \frac{N_s}{2} \ln(2\pi) \quad (23)$$

where N_s is the number of samples. By minimizing $\mathcal{L}$, the trend coefficients and kernel hyperparameters are estimated simultaneously, thereby yielding the optimal mean and covariance structures. Since the input variables have been normalized, the ARD length scales can more directly reflect the sensitivity of the response to variations in different input factors.

2.4.5. Prediction Output

For any new operating condition x_* , the MI-UK-GPR model can provide both point predictions and uncertainty estimates simultaneously. Its posterior mean and posterior variance are given in Equations (24) and (25), respectively.

$$\mu(x_*) = m(x_*) + k_*^\top K^{-1}(y - H\beta) \quad (24)$$

$$s^2(x_*) = k(x_*, x_*) - k_*^\top K^{-1}k_* + u^\top (H^\top K^{-1}H)^{-1}u \quad (25)$$

where k_* denotes the covariance vector between the test point and the training samples, and $u = h(x_*) - H^\top K^{-1}k_*$. In the original physical space, the predicted maximum erosion rate can be written as Equation (26):

$$\hat{E}_{max} = \exp(\mu(x_*)) \quad (26)$$

The corresponding 95% prediction interval is given by Equation (27)

$$[\exp(\mu(x_*) - 1.96s(x_*)), \exp(\mu(x_*) + 1.96s(x_*))] \quad (27)$$

Therefore, the MI-UK-GPR model provides not only point estimates but also prediction uncertainty, thereby offering additional support for subsequent rapid multi-scenario assessment, parameter screening, and reliability evaluation of the prediction results.

It should be emphasized that the MI-UK-GPR model developed in this study was used for interpolation within the predefined parameter space rather than for extrapolation beyond the sampled domain. Specifically, the applicable ranges of the input variables were limited to those listed in Table 1. Therefore, all new operating conditions predicted in

this work were constrained within the above parameter ranges. Predictions outside these ranges were not considered in the present study, because extrapolation may reduce the reliability of the surrogate model. If the model is to be applied to operating conditions beyond the current parameter space, additional CFD samples should be generated and incorporated into the training database to update the surrogate model.

3. Results and Discussion

3.1. Erosion Characteristics of the Reference Elbow

3.1.1. Spatial Distribution Characteristics of Erosion in the Reference Elbow

Based on the erosion-rate contour of the inner wall predicted by the Oka model (Figure 7), the erosion pattern exhibits pronounced spatial non-uniformity, with the primary erosion zone concentrated mainly in the middle and downstream portions of the outer-curvature side of the elbow. The peak erosion is concentrated in the mid-to-downstream outer-curvature region of the elbow, indicating that particle impingement is strongly focused on the outer-curvature wall. The high-erosion region extends circumferentially in a band-like form and exhibits a certain V-shaped spreading characteristic, whereas the erosion level on the inner-curvature side is markedly lower. Under the reference operating conditions considered in this study, the erosion intensity within the primary erosion zone is generally on the order of 10^{-4} , suggesting that the middle and downstream regions of the outer-curvature side of the elbow are susceptible to localized wall thinning and should therefore be prioritized in condition monitoring and integrity assessment.

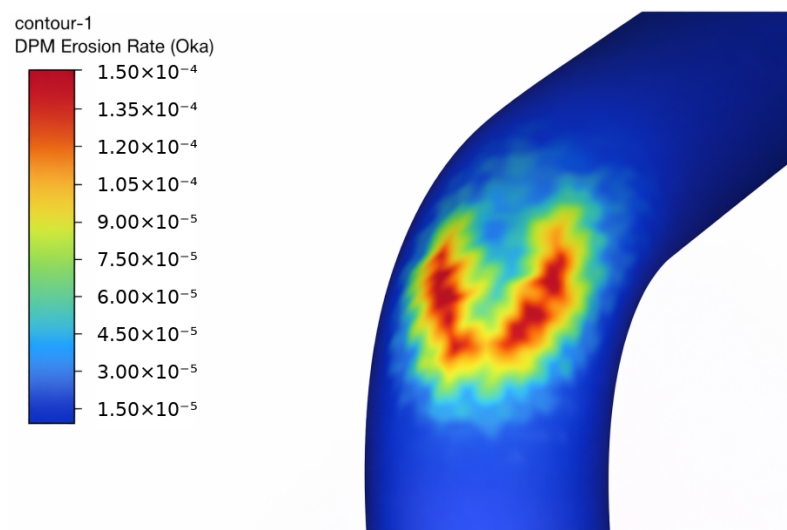


Figure 7. Erosion-rate contour on the inner wall of the elbow.

3.1.2. Erosion Response Characteristics Along the Axial Angle of the Elbow

To quantitatively characterize the angular distribution of erosion in the elbow and to avoid the possible numerical sensitivity associated with a single-cell maximum value, the erosion response along the axial angle was evaluated using the local area-averaged erosion rate within each angular interval. Specifically, the elbow wall was divided into several axial-angle intervals with an interval width of 10° , and the erosion rate was averaged over the wall surface area belonging to each interval. The local area-averaged erosion rate can be calculated using Equation (28).

$$\bar{E}_{\phi_i} = \frac{1}{A_{\phi_i}} \int_{A_{\phi_i}} E(s) dA \quad (28)$$

where $\bar{E}_{\phi_i}$ is the local area-averaged erosion rate in the (i)-th axial-angle interval, A_{ϕ_i} is the corresponding wall surface area, and $E(s)$ is the local erosion rate at wall position (s). Therefore, the erosion peak in the 40–50° interval discussed in this section was determined based on the local area-averaged erosion rate rather than a single-point maximum erosion rate. This treatment provides a more stable representation of the erosion distribution along the elbow angle and reduces the influence of local mesh-scale fluctuations on the identification of the primary erosion region.

The results are presented in Figure 8. It can be seen that the local area-averaged erosion rate generally increases first and then decreases with increasing elbow angle. The highest local area-averaged erosion rate occurs in the 40–50° interval, with a value of $1.002083 \times 10^{-4} \text{ kg} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$. By contrast, the corresponding value in the 0–10° interval near the inlet is only $1.353050 \times 10^{-5} \text{ kg} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$, whereas that in the 80–90° interval near the outlet further decreases to $3.725909 \times 10^{-6} \text{ kg} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$. These results indicate that the most severe erosion does not occur immediately after particles enter the elbow, but instead develops into a localized intensified region in the middle and downstream portions of the bend under the combined effects of flow deflection and particle migration.

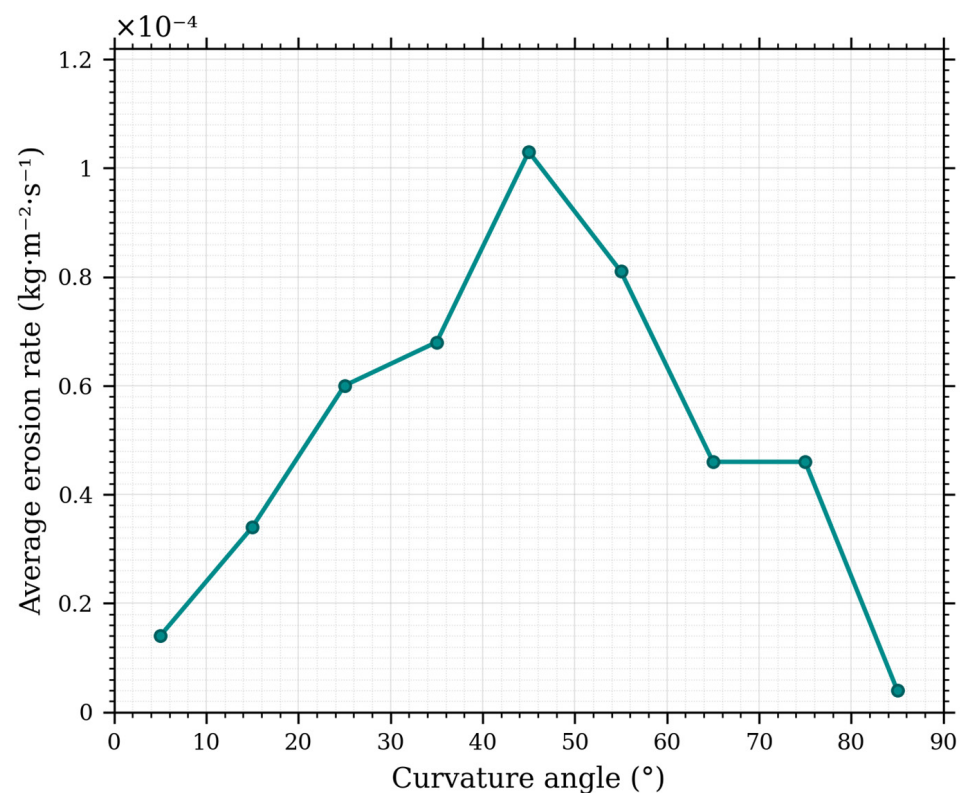


Figure 8. Distribution of local area-averaged erosion rate along the axial angle of the elbow.

It should be noted that E_{max} was used as the target response variable for mesh-independence verification and surrogate-model training, because the maximum wall erosion rate is directly relevant to conservative safety assessment. However, the angular location of the primary erosion region in Figure 8 was identified using the local area-averaged erosion rate over each axial-angle interval. Thus, E_{max} and $\bar{E}_{\phi_i}$ serve different purposes: the former represents the most severe local erosion intensity for prediction and safety evaluation, whereas the latter is used to describe the spatial distribution and angular evolution of erosion along the elbow. The location of the 40–50° high-erosion interval is also consistent with the continuous high-erosion band observed in the erosion-rate con-

tour, further supporting the rationality of using the local area-averaged erosion rate for identifying the primary erosion zone.

3.1.3. Particle Motion Behavior and Hotspot Formation Mechanism

Figure 9 illustrates the particle trajectories and velocity distribution within the elbow. It can be seen that, in the inlet straight section, particles generally migrate along the main flow direction and exhibit a relatively concentrated trajectory distribution. After entering the elbow, however, the particle trajectories begin to deviate markedly, indicating that, because of inertial effects, the particles cannot fully follow the continuous turning motion of the carrier phase. In the transition region of the elbow and the initial section of the outlet, the particle trajectory bundle undergoes significant redistribution, while a large number of high-velocity particles are still present, suggesting that the particles retain strong momentum transport capacity within the elbow. Combined with the above erosion contour and the distribution of local average erosion rate along the elbow angle, these results indicate that the deflected migration and high-momentum transport of particles in the middle and downstream sections of the elbow constitute the key dynamic mechanisms responsible for the intensified local erosion on the outer-curvature side.

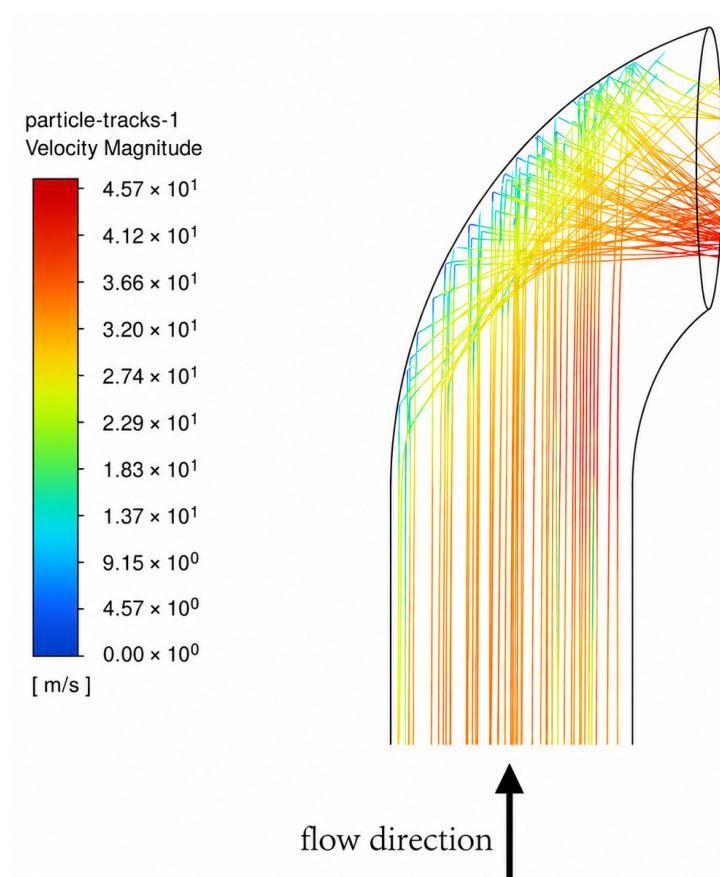


Figure 9. Particle trajectories and velocity distribution within the elbow.

3.2. Predictive Performance and Comparative Analysis of MI-UK-GPR Model

To comprehensively evaluate the overall performance of the proposed model, this study conducts a unified comparison of the MI-UK-GPR model with RSM, SVR, and ordinary GPR under the same dataset, input–output definitions, and validation protocol. Since the proposed method is designed to achieve both high predictive accuracy under small-sample conditions and efficient post-training evaluation, this section analyzes the

results from two perspectives: LOOCV-based performance comparison and computational efficiency under multiple scenarios [30].

3.2.1. Comparison of Models and Validation Protocols

To systematically evaluate the effectiveness of the proposed MI-UK-GPR model in predicting the maximum erosion rate of elbows, a comparative study was conducted against the quadratic response surface model (RSM), support vector regression (SVR), and ordinary Gaussian process regression (ordinary GPR) using the same 60 LHS-CFD samples, identical input variables (D , R/D , θ , $\dot{m}_p$, d_p , and v), the same output target (E_{max}), and a unified leave-one-out cross-validation (LOOCV) protocol [30]. All models were trained in the $\ln(E_{max})$ space, and the prediction results were subsequently mapped back to the original physical space to calculate the mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). For models capable of generating interval predictions, the 95% prediction interval coverage probability (PICP) was also reported as a supplementary metric. In each LOOCV fold, the normalization parameters of the input variables were computed exclusively from the training subset of that fold and then applied to the corresponding training and test samples. Likewise, the hyperparameters of each surrogate model were determined only within the training subset, so as to avoid information leakage and ensure the fairness of model comparison. The LOOCV results of the different surrogate models are presented in Table 9.

Table 9. Comparison of predictive performance of different surrogate models under LOOCV.

Model	MAE	RMSE	MAPE (%)	R^2	PICP (%)
RSM	1.57×10^{-4}	3.23×10^{-4}	28.32	0.7334	85.00
SVR	1.00×10^{-4}	2.37×10^{-4}	27.29	0.8564	88.33
Ordinary GPR	9.70×10^{-5}	1.90×10^{-4}	22.00	0.9081	90.00
MI-UK-GPR	7.10×10^{-5}	1.29×10^{-4}	14.53	0.9573	88.33

Note: All models predict E_{max} based on the same six input variables and are evaluated under the same LOOCV protocol. All models are trained in the $\ln(E_{max})$ space, and MAE, RMSE, MAPE, and R^2 are calculated after the final predictions are mapped back to the original physical space. The prediction interval of the SVR model is approximated from the training residuals; therefore, its PICP is provided only as a supplementary reference and should not be interpreted as fully equivalent to the probabilistic prediction intervals produced by GPR-based models.

3.2.2. Comparison and Analysis of LOOCV Results

As shown in Table 9, the RSM exhibits the weakest predictive performance, with an R^2 of only 0.7334 and a MAPE of 28.32%, indicating that the relationship between the maximum erosion rate and the multi-factor input variables cannot be adequately captured by a global low-order polynomial. After cross-validation-based hyperparameter optimization, the SVR model improves the R^2 value to 0.8564 and reduces the MAPE to 27.29%, indicating that kernel-based methods possess stronger nonlinear fitting capability than conventional response surface models. Ordinary GPR further achieves $R^2 = 0.9081$ and a MAPE of 22.00%, suggesting that the probabilistic kernel regression framework is more suitable for the present problem under small-sample conditions. In contrast, the MI-UK-GPR model delivers the best performance across all four core predictive metrics, namely MAE, RMSE, MAPE, and R^2 , with corresponding values of 7.10×10^{-5} , 1.29×10^{-4} , 14.53%, and 0.9573, respectively. Compared with ordinary GPR, MI-UK-GPR preserves the flexible fitting capability of the Gaussian process while explicitly constraining the dominant directions of variation in erosion peaks induced by velocity, particle size, particle mass flux, and geometric factors through a mechanistic trend term. As a result, the Gaussian process residual component can focus more effectively on learning the remaining complex nonlinear patterns, thereby further improving the overall predictive accuracy.

In terms of interval coverage, the PICP values of SVR, ordinary GPR, and MI-UK-GPR are 88.33%, 90.00%, and 88.33%, respectively. Although MI-UK-GPR does not achieve the highest PICP, it should be noted that the prediction interval of SVR is an empirical interval constructed from training residuals and therefore does not have strict probabilistic meaning. Consequently, its coverage should not be regarded as fully equivalent to that of GPR-based models. Considering point prediction accuracy, fitting consistency, and structural interpretability together, MI-UK-GPR shows the most favorable overall performance among the surrogate models examined in this study.

To further compare the models, Figures 10–14 present the parity plots, relative error distributions, prediction intervals, sample-level prediction results, and metric bar charts, respectively.

Figure 10 presents the parity plots of RSM, SVR, ordinary GPR, and MI-UK-GPR, in which the dashed line represents the ideal prediction condition, $y = x$. It can be seen that the scatter points of the RSM deviate most markedly from the reference line, indicating its limited ability to capture complex nonlinear relationships. Although the scatter distributions of SVR and ordinary GPR show a certain degree of convergence toward the reference line, noticeable deviations still remain in the high-value tail region. By contrast, the data points of MI-UK-GPR are generally more concentrated around the reference line, indicating that this model achieves higher fitting consistency under the current small-sample conditions.

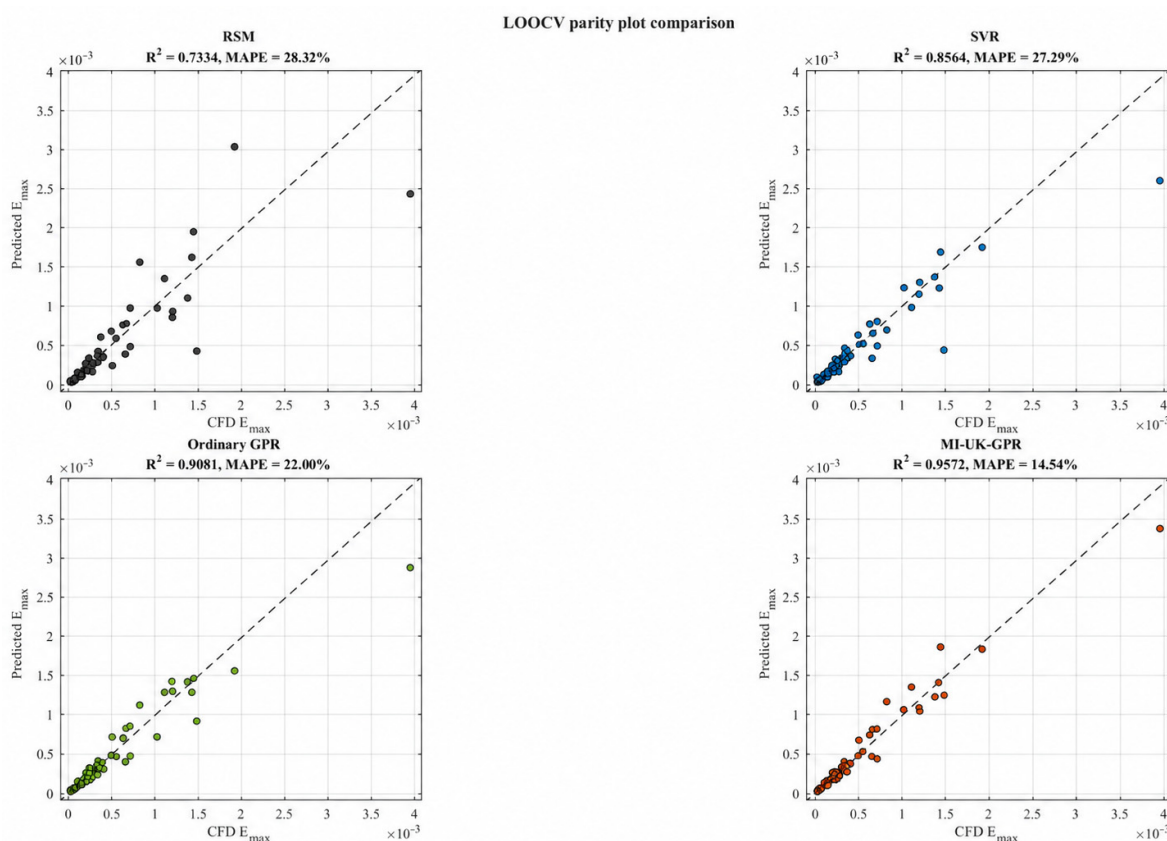


Figure 10. Comparison of parity plots for different surrogate models under LOOCV.

Figure 11 shows the distribution of relative errors across samples for different surrogate models. The RSM exhibits the widest error fluctuation range, indicating its limited adaptability to complex local responses and high-value samples. SVR and ordinary GPR reduce the error fluctuation to some extent, but pronounced error peaks still appear in a few high-erosion tail samples. By contrast, the relative errors of MI-UK-GPR are generally lower, and the error amplitude is effectively suppressed in both the low-to-medium

and medium-to-high value ranges, indicating that the mechanistic trend term contributes positively to the overall stability of prediction.

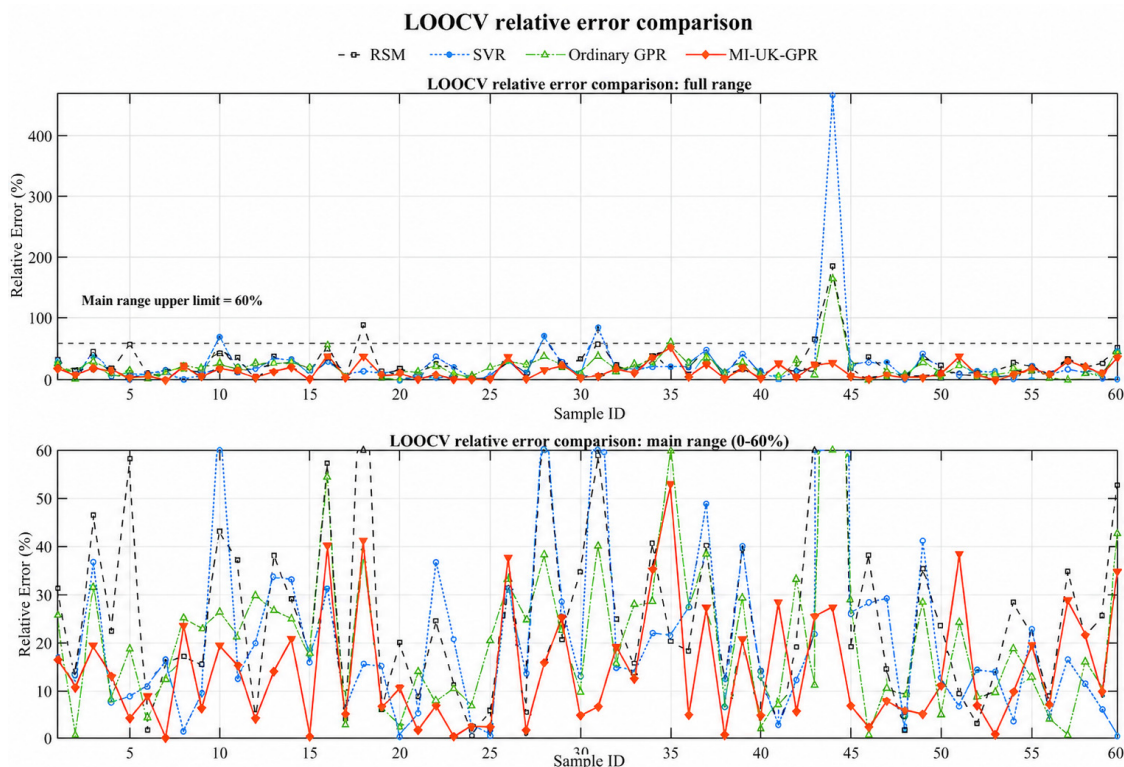


Figure 11. Comparison of sample-level relative errors for different surrogate models under LOOCV conditions.

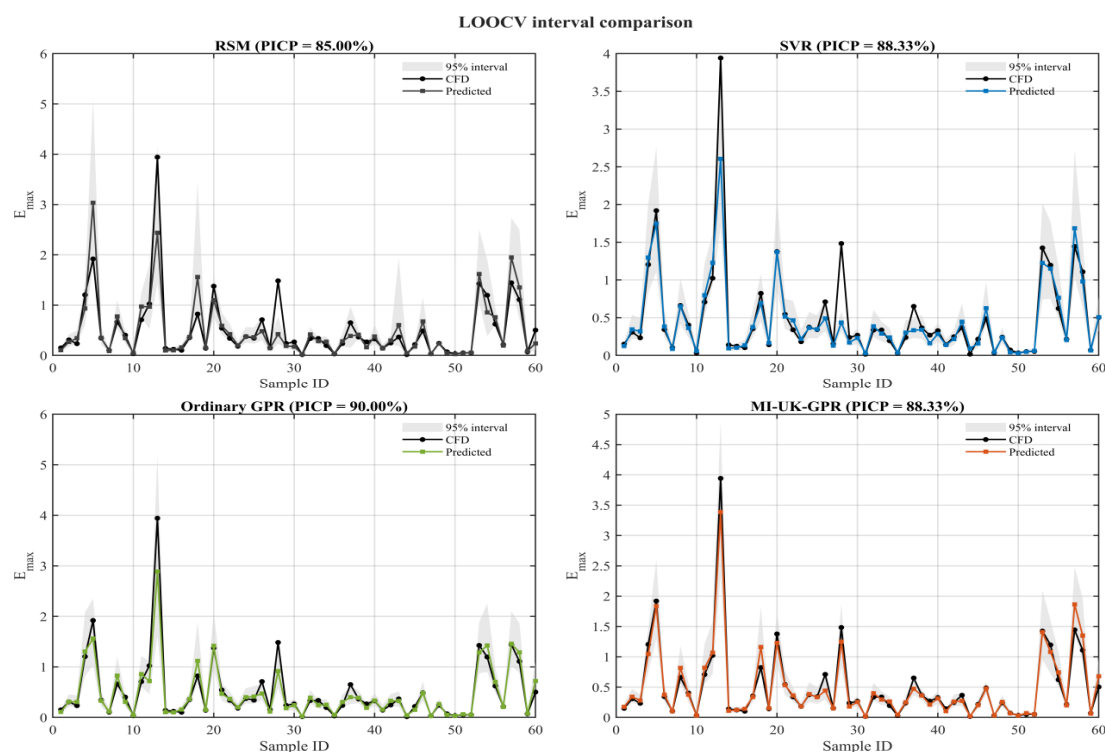


Figure 12. Comparison of prediction intervals for different surrogate models under LOOCV conditions.

Figure 12 presents the sample-level prediction intervals and the corresponding distributions of true values for different surrogate models. For GPR-based models, the gray band represents the 95% prediction interval. For SVR, the interval is approximated from the training residuals and is provided for comparative purposes only. Overall, the RSM interval exhibits weaker coverage of the true values. The PICP values of SVR and MI-UK-GPR are both 88.33%, indicating similar interval coverage performance, whereas the PICP of ordinary GPR reaches 90.00%, which is the highest among the four models. Meanwhile, MI-UK-GPR performs best in terms of point prediction accuracy, although its 95% prediction interval still exhibits some undercoverage, indicating that uncertainty calibration remains to be improved under the current condition of 60 small-sample groups. It should be noted that the interval construction methods differ among models; therefore, the interval coverage shown in the figure is intended primarily for auxiliary analysis and should not be used for strict direct comparison.

Figure 13 presents the sample-wise prediction results for the 60-sample dataset under LOOCV, where the black line represents the true E_{max} obtained from the CFD simulations, and the other curves correspond to the predictions of RSM, SVR, ordinary GPR, and MI-UK-GPR, respectively. The sample numbers on the x-axis indicate only sample indices and do not represent a continuous evolution sequence of the physical process. Overall, the RSM shows the largest deviation in the medium-to-high-value range. SVR and ordinary GPR track the overall trend relatively well, but still exhibit noticeable dispersion around some high-value samples. By contrast, the MI-UK-GPR prediction curve is closer to the true values, indicating that this model achieves a better balance between preserving the global trend and capturing local peaks.

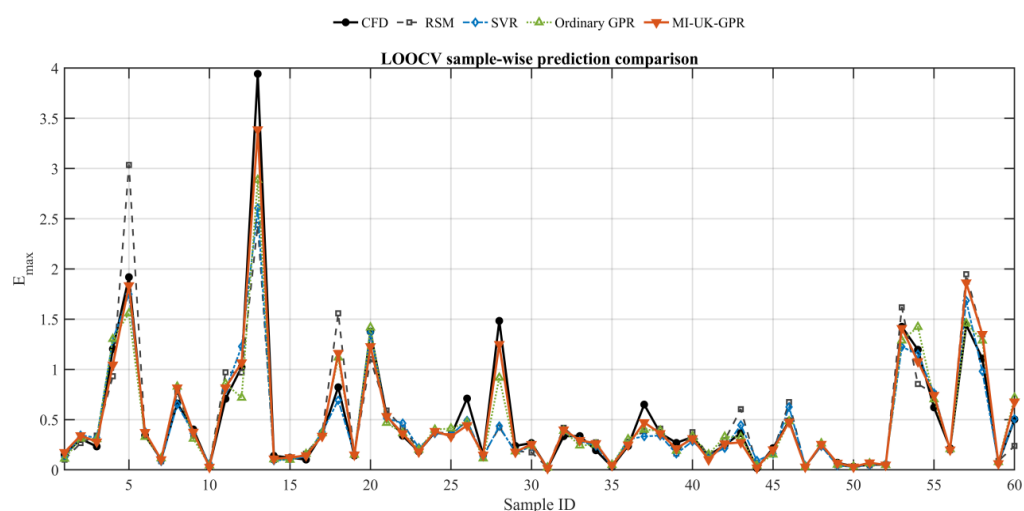


Figure 13. Comparison of sample-level prediction results of different surrogate models under LOOCV.

Figure 14 compares the performance of different surrogate models in terms of four core metrics, namely MAE, RMSE, MAPE, and R^2 . The RSM performs worst across all metrics, indicating that the traditional quadratic response surface is unable to satisfy the accuracy requirements of the present study. The results of SVR and ordinary GPR are markedly better than those of the RSM, suggesting that kernel-based methods are more suitable for the current small-sample nonlinear regression task. Among all the models, MI-UK-GPR achieves the best results in terms of MAE, RMSE, MAPE, and R^2 , demonstrating its superior overall performance in point prediction accuracy.

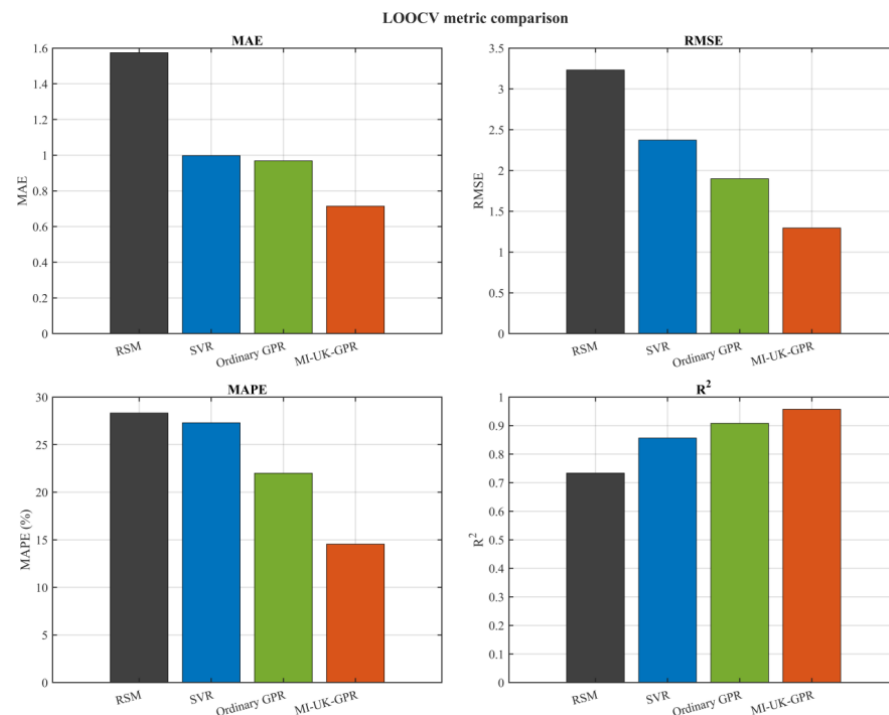


Figure 14. Comparison of performance metrics of different surrogate models under LOOCV.

A combined examination of Figures 10–14 shows that MI-UK-GPR achieves the best balance between global trend preservation and local peak capture. However, a few high-erosion tail samples are still underestimated, mainly because the current sample size is limited to 60 and strict LOOCV leaves too few similar high-value neighbors in the training set. This suggests that future improvements should focus on enhancing the coverage of extreme operating conditions and refining interval calibration.

3.3. Comparison with Existing Engineering Erosion Models

To further clarify the applicability of the proposed method relative to existing engineering erosion models, DNV-RP-O501, the Tulsa model, RSM, SVR, ordinary GPR, and the proposed MI-UK-GPR model were compared from the perspectives of model type, prediction output, and main features or limitations. The comparison is summarized in Table 10.

As shown in Table 10, DNV-RP-O501 and the Tulsa model provide useful engineering references for rapid erosion assessment in pipelines and fittings. However, they are mainly empirical or semi-empirical and usually require material- and condition-dependent parameters. In contrast, the proposed MI-UK-GPR model was trained using CFD-DPM erosion responses generated within the cold hydrogenation parameter space considered in this study. Therefore, it can provide rapid point predictions and prediction intervals for the localized maximum erosion rate within the predefined parameter domain.

It should be noted that the comparison in Table 10 focuses on model characteristics and applicability rather than a direct numerical benchmark of DNV-RP-O501 or the Tulsa model. This is because the present study aims to construct a CFD-informed surrogate model for cold hydrogenation elbows, and the calibrated empirical parameters required for directly applying these engineering models to the hydrogen–silicon-powder material system are not fully available.

Table 10. Comparison between existing erosion-prediction approaches and the proposed MI-UK-GPR model.

Model	Model Type	Output	Main Feature and Limitation
DNV-RP-O501	Empirical engineering method	Erosion-risk or erosion-rate estimate	Fast screening method, but limited for localized erosion in cold hydrogenation elbows
Tulsa model	Semi-empirical erosion model for elbows and tees	Erosion-rate estimate	Classical reference model, but requires condition-dependent calibration
RSM	Polynomial surrogate model	Point prediction of E_{max}	Interpretable, but weak for strong nonlinear responses
SVR	Data-driven machine-learning surrogate model	Point prediction of E_{max}	Nonlinear fitting capability, but no direct uncertainty output
Ordinary GPR	Probabilistic surrogate model	Point prediction and interval	Provides uncertainty, but lacks an explicit mechanism trend
MI-UK-GPR	Mechanism-informed probabilistic surrogate model	Point prediction and interval of E_{max}	Combines mechanism trend and residual modeling for small-sample rapid prediction

3.4. Comparative Analysis of Computational Efficiency

In addition to predictive accuracy, the computational efficiency of the proposed model under multiple operating conditions is also a key indicator of its engineering value. To further demonstrate this advantage, this study compares the computational efficiency of automated batch CFD simulations with that of batch prediction using the trained MI-UK-GPR model within a parameterized framework. In this study, “rapid prediction” refers to the post-training evaluation of new operating conditions using the trained surrogate model rather than to the initial CFD sample generation and model training stages. To ensure a fair comparison, 50 independent operating conditions not included in the training set were selected as a unified test set. For CFD, computational time is defined as the total time required to complete all simulations and extract the corresponding E_{max} values after batch execution is initiated on the parameterization platform. For MI-UK-GPR, computational time is defined as the total time required for the trained model to output the corresponding E_{max} predictions for the same 50 operating conditions. On this basis, a computational efficiency comparison was conducted between parameterized batch CFD simulations and batch prediction using the trained MI-UK-GPR model. The results are presented in Table 11.

Table 11. Comparison of computational efficiency between parameterized CFD and MI-UK-GPR for 50 independent operating conditions.

Number of Cases	CFD Total Time/s	MI-UK-GPR Total Time/s	Average CFD Time/s	Average MI-UK-GPR Time/s	Speedup
50	5083.51	0.004860	101.67	9.72×10^{-5}	1.05×10^6

The results show that the total runtime of the parameterized CFD analysis for the 50 operating conditions was 5083.51 s, with an average of 101.67 s per condition. By contrast, the total runtime for batch prediction of the same 50 operating conditions using the trained MI-UK-GPR model was only 0.004860 s, corresponding to an average prediction time of 9.72×10^{-5} s and a speed-up ratio of approximately 1.05×10^6 . These results indicate that, even when geometric modeling, meshing, and the solution procedure have all been parameterized and can be executed automatically in batch mode, the surrogate model still exhibits substantially higher computational efficiency in evaluating new operating

conditions. This advantage arises because high-fidelity CFD methods still require flow-field solution, particle trajectory computation, and post-processing of wall erosion for each operating condition, whereas the MI-UK-GPR model, once trained, only requires the operating-condition parameters as input to directly provide the predicted maximum erosion rate. Therefore, compared with repeatedly performing high-fidelity numerical simulations, the proposed method is better suited for rapid evaluation under multiple operating conditions, parameter screening, and the identification of relatively high-erosion conditions in cold hydrogenation elbows.

Overall, the trained MI-UK-GPR model retains the accuracy advantage required for small-sample prediction while markedly reducing the marginal cost of evaluating new operating conditions. Therefore, compared with repeated high-fidelity CFD simulations, it is more suitable for rapid multi-condition evaluation and parameter screening of cold hydrogenation elbows.

4. Conclusions and Future Work

4.1. Conclusions

This study proposes an MI-UK-GPR-based model for rapid prediction of the maximum erosion rate of cold hydrogenation elbows under multiple operating conditions. The main conclusions are summarized as follows.

- (1) A CFD model for erosion in elbows under gas–solid two-phase flow was established. Through mesh independence analysis and comparison with published literature, the model was verified to effectively characterize the spatial distribution pattern and peak-response characteristics of the primary erosion zone on the outer-curvature side of the elbow, thereby providing a reliable numerical basis for subsequent sample generation.
- (2) A set of multi-parameter operating conditions was generated using Latin hypercube sampling, and the maximum erosion rate, E_{max} , was obtained through CFD simulations to establish the LHS-CFD sample database. During sample construction, parameter-space coverage and computational cost were balanced, thereby providing effective data support for surrogate modeling under small-sample conditions.
- (3) An MI-UK-GPR model consisting of a mechanistic trend term and a Gaussian process residual term was developed. The trend term, constructed from physical variables, describes the global variation pattern of the erosion response, whereas the residual term captures local nonlinear deviations. The results show that this model effectively balances predictive accuracy and physical interpretability under small-sample conditions.
- (4) Under the unified LOOCV framework, the proposed MI-UK-GPR model achieved the best overall prediction performance among the compared surrogate models. Its MAE, RMSE, MAPE, and R^2 were 7.10×10^{-5} , 1.29×10^{-4} , 14.53%, and 0.9573, respectively, and the corresponding PICP was 88.33%. Compared with RSM, SVR, and ordinary GPR, the proposed model showed lower prediction errors and better overall fitting accuracy, indicating that the mechanism-inspired trend term and Gaussian-process residual modeling can improve prediction performance under small-sample conditions.
- (5) A computational efficiency comparison based on 50 independent operating conditions within a parameterized framework showed that the total runtime of batch parameterized CFD simulations was 5083.51 s, whereas the total runtime of batch prediction for the same operating conditions using the trained MI-UK-GPR model was only 0.004860 s, corresponding to a speed-up ratio of approximately 1.05×10^6 . This confirms that the computational advantage of the proposed method is realized primarily in the post-training evaluation of new operating conditions. Compared with

repeatedly performing high-fidelity numerical simulations, the proposed method can substantially reduce the marginal computational cost of multi-condition analysis and is therefore better suited for rapid evaluation, parameter screening, and the identification of relatively high-erosion conditions in cold hydrogenation elbows.

It should be noted that the surrogate model proposed in this study is trained on responses generated by a validated CFD model. Therefore, its predictive capability should be understood as an efficient approximation of high-fidelity numerical results rather than direct experimental calibration of absolute erosion rates under field conditions. Accordingly, the model is more suitable for rapid multi-condition evaluation and parameter screening in engineering applications.

4.2. Future Work

- (1) Incorporation of experimental or industrial monitoring data for model calibration and cross-domain validation

The training labels of the surrogate model developed in this study are derived from a validated CFD model, and its predictive capability is mainly reflected in the efficient approximation of high-fidelity numerical responses. In future work, online monitoring data or experimental measurements from cold hydrogenation units could be incorporated to perform bias correction and output recalibration, thereby improving the quantitative predictive capability and generalization performance of the model in practical engineering applications.

- (2) Consideration of unsteady operating conditions and temporal evolution characteristics

This study establishes an erosion prediction model under steady-state operating conditions, whereas actual cold hydrogenation processes may involve operating fluctuations and transient variations. In future work, time-series-based surrogate models could be developed on the basis of unsteady CFD simulations to investigate the dynamic evolution of erosion rates, thereby extending the scope of the present framework from static prediction to dynamic assessment.

Author Contributions: Conceptualization, J.M., J.Y. and A.S.; methodology, J.M. and J.Y.; software, J.M. and J.Y.; validation, J.M. and J.Y.; formal analysis, J.M. and J.Y.; investigation, J.M., W.Y., Y.S. and D.P.; resources, A.S. and Y.H.; data curation, J.M. and J.Y.; writing—original draft preparation, J.Y.; writing—review and editing, J.M., J.Y., A.S. and Y.H.; visualization, J.M. and J.Y.; supervision, J.M. and A.S.; project administration, J.M., W.Y., Y.S. and D.P.; funding acquisition, J.M., W.Y. and Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Xinjiang Uygur Autonomous Region Department of Science and Technology through the project “Research on Key Technologies for Digital Twin Modeling and Intelligent Operation and Maintenance of Silicon-Based New Material Production Equipment” (grant number 2024B04003-1), the project “Research on Key Technologies for Intelligent Operation and Maintenance of Wind Turbine Units Based on Data–Model Joint Driving” (grant number 2025XGYT-CYC10), and the project “Development of Key Corrosion Protection Technologies for Hydrogen Production Equipment from Semi-Coke Crude Gas” (grant number 2024B04021-2).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated and/or analysed during the current study, including the LHS-CFD sample database and the independent test cases, are available from the corresponding author on reasonable request.

Conflicts of Interest: Authors W.Y., Y.S. and D.P. were employed by Xinte Silicon Materials Co., Ltd. The remaining authors declare no conflicts of interest. The authors declare that this employment did not influence the design, analysis, interpretation, or reporting of the research.

References

1. Lv, Q.; Yuan, X.; Guo, L.; Zhao, D.; Ma, W.; Xie, G.; Hou, Y.; Shen, J.; Yang, N. Advances in production and optimization of electronic-grade polysilicon: A review of modified Siemens and silane methods. *Sol. Energy Mater. Sol. Cells* **2025**, *283*, 113446. [\[CrossRef\]](#)
2. Yadav, S.; Chattopadhyay, K.; Singh, C.V. Solar grade silicon production: A review of kinetic, thermodynamic and fluid dynamics based continuum scale modeling. *Renew. Sustain. Energy Rev.* **2017**, *78*, 1288–1314. [\[CrossRef\]](#)
3. Yan, D. Siemens process. In *Handbook of Photovoltaic Silicon*; Yang, D., Ed.; Springer: Berlin/Heidelberg, Germany, 2019; pp. 37–68.
4. Ding, W.J.; Yan, J.M.; Xiao, W.D. Hydrogenation of silicon tetrachloride in the presence of silicon: Thermodynamic and experimental investigation. *Ind. Eng. Chem. Res.* **2014**, *53*, 10943–10953. [\[CrossRef\]](#)
5. Det Norske Veritas. *DNV-RP-O501: Managing Sand Production and Erosion*; amended 2025; DNV: Høvik, Norway, 2015.
6. Shirazi, S.A.; Shadley, J.R.; McLaury, B.S.; Rybicki, E.F. A procedure to predict solid particle erosion in elbows and tees. *J. Press. Vessel. Technol.* **1995**, *117*, 45–52. [\[CrossRef\]](#)
7. Othayq, M.M.; Bilal, F.S.; Sedrez, T.A.; Shirazi, S.A. Experimental and numerical assessments on solid particle erosion in upward vertical–horizontal and horizontal–vertical downward elbows for multiphase and gas–sand flows. *Wear* **2023**, *524–525*, 204812. [\[CrossRef\]](#)
8. Wang, Q.; Huang, Q.; Wang, N.; Wen, Y.; Ba, X.; Sun, X.; Zhang, J.; Karimi, S.; Shirazi, S.A. An experimental and numerical study of slurry erosion behavior in a horizontal elbow and elbows in series. *Eng. Fail. Anal.* **2021**, *130*, 105779. [\[CrossRef\]](#)
9. Zhou, F.; Li, J.; Yang, D.; Li, Y.; Luo, J. Experimental study on collision characteristics of large coal particles (7–15 mm) in 90° elbows of pneumatic conveying systems. *Powder Technol.* **2022**, *396*, 305–315. [\[CrossRef\]](#)
10. Liu, E.B.; Huang, S.; Tian, D.C.; Shi, L.M.; Peng, S.B.; Zheng, H. Experimental and numerical simulation study on the erosion behavior of the elbow of gathering pipeline in shale gas field. *Pet. Sci.* **2024**, *21*, 1257–1274. [\[CrossRef\]](#)
11. Hong, B.; Li, Y.; Li, X.; Li, G.; Huang, A.; Ji, S.; Li, W.; Gong, J.; Gu, J. Experimental investigation of erosion rate for gas-solid two-phase flow in 304 stainless/L245 carbon steel. *Pet. Sci.* **2022**, *19*, 1347–1360.
12. Yu, Y.; Ma, W.; Li, Z.; Yang, Z.; Jin, Z.; Liu, X.; Xu, Q. Research of erosion and buckling of pipe bends under gas-solid two-phase flow. *Ocean Eng.* **2025**, *320*, 120336.
13. Peng, W.; Cao, X.; Hou, J.; Xu, K.; Fan, Y.; Xing, S. Experiment and numerical simulation of sand particle erosion under slug flow condition in a horizontal pipe bend. *J. Nat. Gas Sci. Eng.* **2020**, *76*, 103175. [\[CrossRef\]](#)
14. Mu, L.; He, S.; Chen, S.; Chen, X.; Shi, J.; Zhang, Y.; Liu, C. Prediction of bend-pipeline erosion field based on CFD and reduced order method. *Chem. Eng. Process.-Process Intensif.* **2025**, *213*, 110320. [\[CrossRef\]](#)
15. Zhao, S.; Bao, J.; Ge, S.; Wang, Z.; Yin, Y.; Huang, C. Influence and optimization of swirler parameters on elbow wear of deep coal fluidization pipeline transportation system. *Friction* **2024**, *12*, 1408–1433. [\[CrossRef\]](#)
16. Li, Y.; Zhou, Y.; Xiao, Y.; Yang, W. Study of gas-solid two-phase flow in pipeline elbows using an LES-DPM coupling method. *Powder Technol.* **2023**, *413*, 118012.
17. Liu, M.; Liu, H.; Zhang, R. Numerical analyses of the solid particle erosion in elbows for annular flow. *Ocean Eng.* **2015**, *105*, 186–195. [\[CrossRef\]](#)
18. Mitchell, T.M. *Machine Learning*; McGraw-Hill: New York, NY, USA, 1997.
19. Brunton, S.L.; Noack, B.R.; Koumoutsakos, P. Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* **2020**, *52*, 477–508. [\[CrossRef\]](#)
20. Wang, Z.; Chen, H.; Meng, W.; Zhang, X.; Dou, Y. Solid particle erosion prediction in elbows based on machine learning and swarm intelligence algorithm. *J. Pet. Sci. Eng.* **2022**, *218*, 111042. [\[CrossRef\]](#)
21. Chen, H.; Huang, H.; Wei, R.; Wang, Z. A novel AI-driven model for erosion prediction for elbow in gas-solid two-phase flows. *Wear* **2024**, *540–541*, 205241.
22. Chi, M.; Zeng, X.; Gao, Y.; Li, W.; Jiang, H.; Sun, R. The erosion rate prediction for the elbow in shale gas gathering and transportation system: RSM and GA-BP-ANN modeling. *Powder Technol.* **2024**, *428*, 119429. [\[CrossRef\]](#)
23. Li, W.; Feng, J.; Deng, J.; Jiao, Q.; Fang, Z. Research on predicting gas–solid erosion in series elbows using interpretable machine learning. *ASME J. Fluids Eng.* **2025**, *147*, 111402. [\[CrossRef\]](#)
24. Menter, F.R. Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA J.* **1994**, *32*, 1598–1605. [\[CrossRef\]](#)
25. Oka, Y.; Okamura, K.; Yoshida, T. Practical estimation of erosion damage caused by solid particle impact: Part 1: Effects of impact parameters on a predictive equation. *Wear* **2005**, *259*, 95–101.
26. Oka, Y.; Okamura, K.; Yoshida, T. Practical estimation of erosion damage caused by solid particle impact: Part 2: Mechanical properties of materials directly associated with erosion damage. *Wear* **2005**, *259*, 102–109.

27. McKay, M.D.; Beckman, R.J.; Conover, W.J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **2000**, *42*, 55–61. [[CrossRef](#)]
28. Sacks, J.; Welch, W.J.; Mitchell, T.J.; Wynn, H.P. Design and analysis of computer experiments. *Stat. Sci.* **1989**, *4*, 409–423. [[CrossRef](#)]
29. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.
30. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**, *4*, 40–79. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.