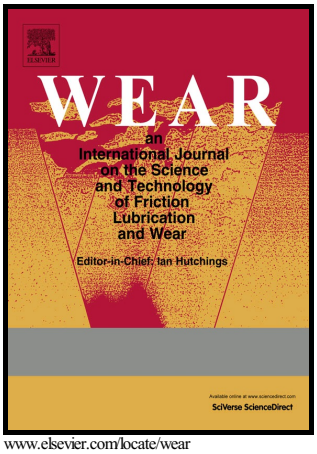
# Uncertainty Quantification in Erosion Predictions using Data Mining Methods

Wei Dai, Selen Cremaschi, Hariprasad J. Subramani, Haijing Gao




www.elsevier.com/locate/wear

PII: S0043-1648(17)31757-X
DOI: https://doi.org/10.1016/j.wear.2018.05.009
Reference: WEA102420

To appear in: *Wear*

Received date: 8 December 2017
Revised date: 4 May 2018
Accepted date: 9 May 2018

Cite this article as: Wei Dai, Selen Cremaschi, Hariprasad J. Subramani and Haijing Gao, Uncertainty Quantification in Erosion Predictions using Data Mining Methods, *Wear,* https://doi.org/10.1016/j.wear.2018.05.009

# Uncertainty Quantification in Erosion Predictions using Data Mining Methods

Wei Dai[a], Selen Cremaschi[a, *], Hariprasad J. Subramani[b], Haijing Gao[b]
[a] Chemical Engineering Department, Auburn University, Auburn, AL 36849, USA
[b] Chevron Energy Technology Company, Houston, TX, USA

[*]selen-cremaschi@auburn.edu



*Abstract*

The transport of solids in multiphase flows is common practice in energy industries due to the unavoidable extraction of solids from oil and gas bearing reservoirs. The persistent collision of solids to the pipeline can lead to erosion, i.e., the removal of internal surface of the pipeline. Reliable estimates of erosion rates are essential for designing and safely operating pipelines that transport solids. Prediction of erosion rates in multiphase flow is a complex problem due to the lack of accurate models for predicting particle movements in the flow and their impact velocities to the wall. The erosion-rate calculations also depend on the accuracy of the flow regime predictions in the pipeline. The comparisons of existing model predictions to experimental data revealed that the predictions might differ by several orders of magnitude for some operating conditions. The goal of this paper is to introduce a computational framework that estimates the model-prediction uncertainty of erosion-rate models. The inputs are a model predicting erosion rates and a database containing erosion-rate measurements at various operating conditions. The framework utilizes a non-parametric regression analysis, Gaussian Process Modeling (GPM), for estimating the model-prediction uncertainty. We compare two approaches for clustering the data prior to training GPMs: (1) a flow regime based clustering, and (2) a new clustering approach introduced in this paper. The results reveal that the new data clustering approach significantly shrinks the confidence intervals of the uncertainty estimates.

*Keywords*