Abstract

This study explores the application of the K-Nearest Neighbor (K-NN) algorithm in predicting the operational conditions of natural gas pipeline transmission networks, utilizing data collected from 2020 to 2021. The experiments conducted involve variations in the prediction time range (168-730 hours) and prediction steps (6-24 hours). In this research, the use of the K-Nearest Neighbor (KNN) algorithm was chosen due to its ability to handle large time-series datasets without requiring certain data distribution assumptions. KNN has the advantage of handling non-linear relationships between variables and can adapt to complex patterns. In the "168 hours Slicing with 6 hours Step Ahead" experiment, K-NN excelled. MAE for pressure was 0.314 with R-Squared at 0.985. Volume rate predictions achieved a MAE of 0.665 and R-Squared of 0.956. These metrics showcase K-NN's robust accuracy in pressure and volume rate predictions. The results' evaluation reveals performance variations dependent on the prediction time span experiments. Although K-NN yields reasonably accurate estimates for pressure and volume rate, discernible differences in performance emerge, especially at shorter prediction timescales. This underscores the necessity for a meticulous approach in parameter selection to enhance the model's consistency and accuracy. The findings not only emphasize the algorithm's potential but also underscore the importance of a nuanced parameter selection process to achieve optimal results. The study sheds light on the challenges faced when forecasting operational conditions of natural gas pipelines and accentuates the importance of refining the K-NN's performance. This research contributes valuable insights into the application of K-NN in forecasting natural gas pipeline conditions, providing a foundation for further refinement and improvement of the algorithm's performance in the context of dynamic operational environments.

Index Terms: Forecasting, K-Nearest Neighbor (K-NN), Natural Gas Pipeline, Operational Conditions, Time Series Analysis.