Proc. of the 9th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Zhongqiang Liu, Jian Dai and Kate Robinson
Published by Research Publishing, Singapore. ©2025 The Author(s), CC-BY 4.0 licence.
doi: 10.3850/GRF-25280825_isgsr-101-P123-cd

Prediction of Cold Wave-Induced UndergroundPipe Failures under Climate Change Using Machine Learning

Qiang Zhang

Urban Mobility Institute, Tongji University, Shanghai, China. E-mail: zhangqiang917@tongji.edu.cn

Qunfang Hu

Shanghai Institute of Disaster Prevention and Relief, Tongji University, Shanghai, China. E-mail:huqunf@tongji.edu.cn

Delu Che

College of Civil Engineering, Tongji University, Shanghai, China. E-mail:1811402@tongji.edu.cn

Fei Wang

Shanghai Institute of Disaster Prevention and Relief, Tongji University, Shanghai, China. E-mail:wangf @tongji.edu.cn

ZhanSu

Shanghai Institute of Disaster Prevention and Relief, Tongji University, Shanghai, China. E-mail:2305219@tongji.edu.cn

Abstracts: The low temperatures caused by cold waves under current global climate change seriously threaten the operational safety of underground pipes, especially water supply pipes. This study presents a data-driven method to identify potential underground water supply pipe failures during cold waves using Light Gradient Boosting Machine (LightGBM). Furthermore, the impact of various factors on pipe failures during cold waves is explored through the interpretability analysisof the machine learning model. In the study area, it is found that the average daily pipe failures rise to 0.87 during cold waves, a 31.8% increase from the 0.66 average during non-cold wave periods. The model prediction results demonstrate that the LightGBM model achieves superior performance in predicting pipe failures during cold waves with accuracy (0.9398) and AUC (0.9524). The fragility of the underground pipe network during cold waves is fundamentally related to the operational condition of the pipes, and the sharp drop in temperature over a short period is the key triggering factor.

Keywords: Climate change; Cold wave; Underground pipe; Machine learning; Failure prediction.

1. Introduction

Under the backdrop of current global climate change, the frequency and intensity of extreme weather events, such as cold waves, are gradually increasing(Horton et al., 2015).This trend is exacerbating the operational risks and safety challenges for urban underground infrastructure.In particular, large-scale underground pipe failures during cold waves have emerged as a critical and recurrent urban disaster(Hu et al., 2024b).For example, in January 2021, Shanghai experienced two consecutive cold waves, leading to 70 bursts in large-diameter pipes over 500 mm anda 14.8% increase in daily water supply(Hu et al., 2023).Mitigating underground pipe failures during cold waves is crucial for maintaining urban safety.

Previous studies reveal the adverse effects of low temperatures on the service performance of underground pipelines.The number of underground pipe failures each year shows a clear positive correlation with the frequency and intensity of low temperatures(Bruaset and Sægrov, 2018). The incidence of pipe failure increases significantly with seasonal decreases in temperature, with the majority of failures occurring during the winter season(Pietrucha-Urbanik, 2015).Furthermore, cold waves increase mechanical responses such as stress and displacementin underground pipes, withthe joints of shallow-buried pipes being particularly vulnerable to the effects of cold waves(Barton et al., 2019; Hu et al., 2024a). However, there is currently limited research on methods for predicting underground pipe failures during cold waves, which can guide water utilities to effectively take proactive measures to reduce the number of failures and associated losses.

The main objective of this study is topropose a machine learning-based method for predicting underground pipe failures induced by cold waves and to investigate the impact of various factors on these failures.This study offers valuable insights and tools to address the challenges posed by extreme cold events on the operational safety of urban underground infrastructure under the backdrop of global climate change.

2. Materials

2.1. *Cold wave-induced pipe failure records*

This study takes the underground water supply pipe with a diameter above 500mm in the central urban area of a city in eastern China as the research object. Data from various sources are acquired and organized for this study, including the following datasets: (i) 3957 historical records of pipe failures from November 2003 to January 2021, (ii) historical weather data from the National Oceanic and Atmospheric Administration, and (iii) geographic information system data for the underground pipe network.

Between November 2003 and January 2021, the study area experienced a total of 52 cold wavesbased on the definition and classification of cold waves in current Chinese standards (China Meteorological Administration, 2017). The pipe failure records occurring during these 52 cold wave periods are selected from a total of 3957 records. Since the recovery of soil temperature lags behind air temperature, pipe failure records occurring within three days after the end of the cold wave are also included in the analysis.This results in a total of 448 pipe failure records associated with cold waves between November 2003 and January 2021(see Fig. 1).During this period, the average daily number of pipe failures during cold waves is 0.87, and the number during non-cold wavesis 0.66. The number of pipe failures during cold waves increases by 31.8% compared to the non-cold wave period.Moreover,the average daily number of failures rises as the minimum temperature during cold waves decreases, with a sharp increase in failures occurring when the minimum temperature drops below -5°C (see Fig. 2).

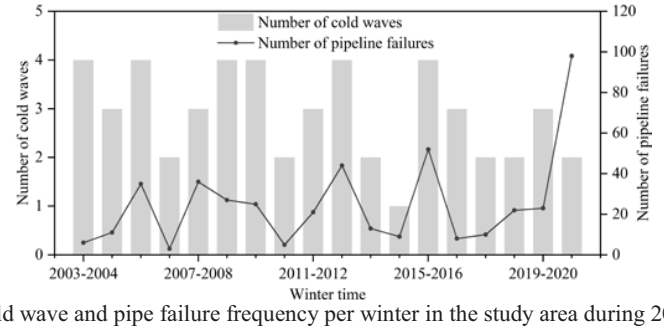


Fig. 1. Cold wave and pipe failure frequency per winter in the study area during 2003-2021.

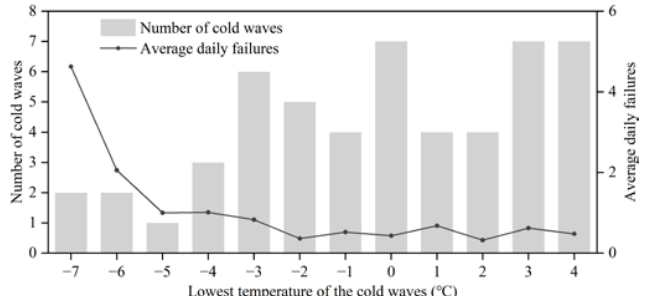


Fig. 2. The number of cold waves at various minimum temperatures and the average daily failures.

2.2.*Databaseestablishment*

To minimize the impact of data uncertainty on the model establishment, this study focuses on the three most common pipe material categories in the database: Ductile iron pipe (DI), Cast iron pipe (CI), and Steel pipe (ST). From the 448 records, a total of 351 complete pipe failure records for these three types of pipes areselected as positive samples. These 351 positive samples are then divided into training and testing positive samples in chronological order. Specifically, 271 pipe failure records from 2003 to 2019 are used as training positive samples, while 80 pipe failure records from 2020 to 2021 are used as testing positive samples. In addition, negative samples are randomly selected from pipes that remained operational during cold waves, ensuring the sample size for negative samples is ten times larger than that of the positive samples.This results in 2710 training negative samples and 800 testing negative samples. Finally, the positive and negative samples are merged to create a comprehensive dataset consisting of 2981 training samples and 880 testing samples, with a training-to-testing set ratio of approximately 3.4:1 (seeTable 1).

Table 1 Detailed distribution of dataset samples.

Dataset	Positive samples	Negative samples	Total
Training set	271	2710	2981
Testing set	80	800	880
Total	351	3510	3861

The development of machine learning models relies on the selection of appropriate features as inputs. In this study, 14 key features are extracted from three categories: pipe features, temperature features, and cold wave features (see Table 2). The pipe features represent the intrinsic conditions of the pipes and have been widely used and proven effective for pipe failure prediction (Liu et al., 2023).The temperature andcold wave features are extracted to comprehensively represent the patterns of temperature changes during cold wave periods. During data pre-processing, continuous features are normalized using the z-score method, while categorical features are encoded using label encoding. To address the issue of data imbalance, the Adaptive Synthetic Sampling (ADASYN) technique is applied to the training set.

Table 2 Key features for water pipe failures during cold wave periods.

Item	Features	Name	Type
Cold wave feature	The highest temperature at the beginning (℃)	HTB	Continuous
	The lowest temperature during the cold wave (℃)	LT	Continuous
	Cooling rate (℃/d)	CR	Continuous
	Heating rate (℃/d)	HR	Continuous
	Duration of Cold Wave	DCW	Continuous
	Cold wave level	CWL	Discrete
Temperature feature	The lowest temperature on the day of failure (℃)	LTF	Continuous
	The maximum temperature difference in the three days before failure (℃)	MTT	Continuous
	The number of days with the lowest daily temperature in days before failure	DLT	Continuous
Pipe feature	Pipe material	PM	Discrete
	Pipe diameter (mm)	PD	Continuous
	Pipe age (year)	PA	Continuous
	Pipe buried depth (m)	PBD	Continuous
	Pipe length (m)	PL	Continuous

3.Methodology

3.1.*Machine learning algorithm*

Predicting underground pipe failures induced by cold waves is a binary classification task, where the outcomes are either failure(positive) or non-failure (negative). In this study, an ensemble machine learning models with robust classification capabilities, namely Light Gradient Boosting Machine (LightGBM), is developed to predict cold wave-inducedunderground pipe failures by data-driven extraction of incident features.LightGBM is an efficient implementation of the Gradient Boosting Decision Tree (GBDT) algorithm, specifically optimized for training on large datasets with better speed and lower memory usage (Ke et al., 2017). It uses multi-threaded parallel histograms to accelerate the training process and performs data pre-processing based on Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB). During the model training, the classification threshold is set to 0.5,and hyperparameter optimization is performed using 5-fold cross-validation and grid search techniques. The model is subsequently trained using these optimal hyperparameters and evaluated on the testing set.

3.2.*Model evaluation*

Five key metrics, namely accuracy, recall, specificity,maintenance rate (MR), and the Area Under the Curve (AUC), are chosen to evaluate the performance of prediction models. Specifically, accuracy quantifies the proportion of samples correctly predicted in the entire dataset, while recall and specificity respectively characterize the proportion of actual positive and actual negative samples correctly predicted. MR is defined as the proportion of samples predicted as positive among all samples, where the samples predicted as positive are considered to be pipes requiring maintenance. The lower the MR, the fewer pipes require maintenance, indicating higher cost-effectiveness. AUC, the area under the Receiver Operating Characteristic Curve (ROC), provides a comprehensive evaluation of the model's predictive performance, with higher values indicating the stronger classification ability of the model.

4. Results and Discussion

The LightGBM model demonstrates excellent performance in predictingcold wave-induced pipe failure, achieving anaccuracy of 0.9398 and an AUC of 0.9524, both exceeding 0.9 (see Fig. 3). This highlights the relevance of the extracted features in reflecting the critical factors leading to pipe failures during cold waves, as well as the robust capability of machine learning model in capturing the complex non-linear relationships among these key factors.Specifically, the recall of 0.75 indicates that the model is capable of identifying a majority ofcoldwave-induced pipe failures, while the high specificity of 0.9587 illustratesits effectiveness inaccurately recognizing pipes that remain operational during cold waves.The low MR of 0.1057 demonstrates the model's superior cost-effectiveness. This helps avoid a significant number of missed predictions and false positives, realizing the mitigation of a large number of potential pipe failures at a low maintenance cost.Additionally, water utility companies can fine-tune the model's classification thresholdto optimize the trade-off between failure mitigation and maintenance cost according to their operational needs.

Theimportance scoreforeachfeaturein LightGBM is calculated based on the split gain (see Fig. 4). The four most important features are PL, PA, MTT, and PBD, with three of them related to the operational condition of the pipes, followed by temperature-related factors such as LTF and CR.Thisfindingdemonstrates that the operational condition of the pipes is the fundamental factor influencing pipe failures during cold waves, while the sharp drop in temperature over a short period serves as the crucial trigger leading to these failures.During

cold waves, low temperatures cause frost heave in the soil and leadto severe mechanical strain on the pipes,thereby threatening the structural safetyofthe pipe.Furthermore, underground pipes across each urban district experience nearly identical air temperature fluctuations,consequently, the weak points of the underground pipe network during cold waves are fundamentally associated with the operational condition of the pipes.

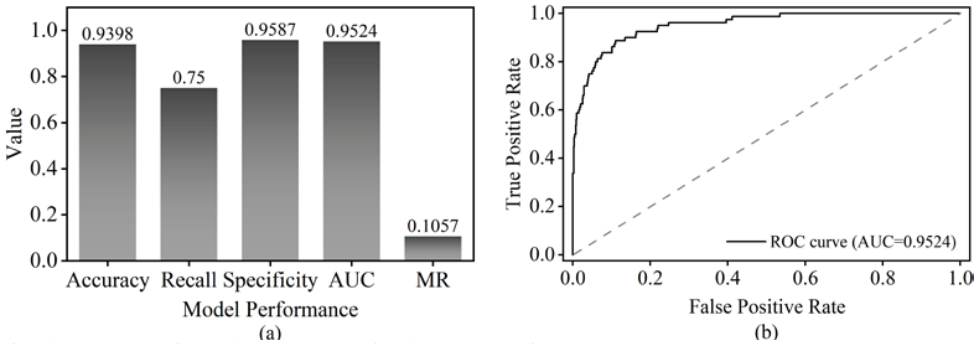


Fig. 3. Modelperformance on the testing set: (a) evaluation metrics, (b)ROC curve.

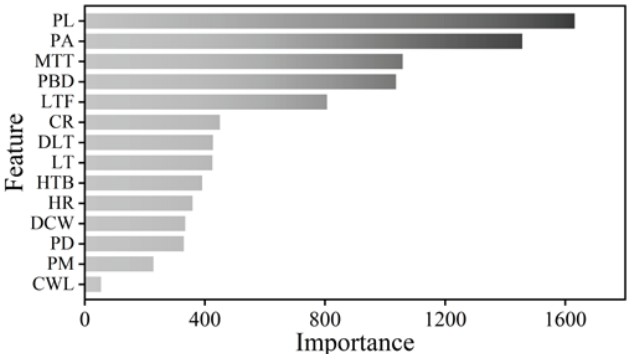


Fig. 4. The gain importance of features.

5. Conclusion

Thisstudyproposed a machine learning-based approach to predict cold wave-induced underground pipe failures and toinvestigate the impact of cold waves on pipe failures. In the study area, the average daily pipe failures increase to 0.87 during cold waves, representing a 31.8% rise compared to the average of 0.66 during non-cold wave periods. The LightGBM model demonstrates outstanding performance in predicting potential pipe failures during cold waves,achieving accuracy (0.9398) and AUC (0.9524). The fragility of the underground pipe network during cold waves isfundamentally related to the operational condition of the pipes, and the sharp drop in temperature over a short period serves asthe key triggering factor.This study provides valuable insights into the safety and risk management of underground and geotechnical engineering amidst global climate change.

Acknowledgement

The support from the National Key Research and Development Program of China (2022YFC3801000), Shanghai Science and Technology Innovation Action Plan project (22dz1201200, 22dz1200402) is greatly appreciated.

References

Barton, N.A., Farewell, T.S., Hallett, S.H., Acland, T.F., 2019. Improving pipe failure predictions: Factors affecting pipe failure in drinking water networks. Water Research 164, 114926.

Bruaset, S., Sægrov, S., 2018. An Analysis of the Potential Impact of Climate Change on the Structural Reliability of Drinking Water Pipes in Cold Climate Regions. Water 10, 411.

China Meteorological Administration, 2017. Grade of cold wave.

Horton, D.E., Johnson, N.C., Singh, D., Swain, D.L., Rajaratnam, B., Diffenbaugh, N.S., 2015. Contribution of changes in atmospheric circulation patterns to extreme temperature trends. Nature 522, 465–469.

Hu, Q., Ayinde, O., Liu, W., 2024a. Mechanical response of buried water pipes to traffic loading before and after extreme cold waves. Transportation Geotechnics 49, 101418.

Hu, Q., Che, D., Song, Z., Wang, F., Fang, H., He, L., Zhu, H., 2023. Impact characteristics and fortification standards of cold wave disasters in urban water supply systems. Journal of Tongji University (Natural Science) 51, 188–196.

Hu, Q., Che, D., Wang, F., He, L., 2024b. Analyzing the effects of extreme cold waves on urban water supply network safety: A case study from 2020 to 2021. Urban Climate 58, 102146.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. Presented at the Neural Information Processing Systems.

Pietrucha-Urbanik, K., 2015. Failure analysis and assessment on the exemplary water supply network. Engineering Failure Analysis 57, 137–142.

Liu, W., Xie, Z., Song, Z., 2023. Predicting Water Pipe Failures Using Deep Learning Algorithms. Journal of Infrastructure Systems 29(3).