Predicting Blue–Green Infrastructure Underperformance Using Integrated Machine-Learning Models

Prédiction des performances insuffisantes des infrastructures bleues-vertes et à l'aide de modèles d'apprentissage automatique intégrés

Virginia Smith[1], Musfiqur Rahman[1], Ruixuan Wang[2], Achira Amur[1], Shiva Arabi[1], Peleg Kremer[3], Xun Jiao[2], Bridget Wadzuk[1]
[1]Department of Civil and Environmental Engineering, Villanova University, virginia.smith@villanova.edu ; mrahman@villanova.edu ; aamur@villanova.edu ; shiva.arabi@villanova.edu ; Bridget.Wadzuk@villanova.edu ;
[2]Department of Electrical and Computer Engineering, Villanova University, rwang@villanova.edu
[3]Department of Geography and the Environment, Villanova University, peleg.kremer@villanova.edu

RÉSUMÉ

L'évaluation régulière des performances des infrastructures vertes et bleues est essentielle à une gestion durable des eaux pluviales en milieu urbain, mais elle reste difficile en raison de la complexité des environnements urbains, de la variabilité des systèmes d'infrastructures vertes et bleues et des ressources financières et humaines limitées. Les approches traditionnelles d'estimation des performances insuffisantes reposent sur des modèles hydrologiques basés sur la physique, qui soit simplifient à l'excès les conditions réelles, soit nécessitent des données volumineuses et des calculs complexes. Cette étude présente un cadre intégré, basé sur les données, pour évaluer le risque de performance des infrastructures vertes en reliant deux composantes prédictives : (1) des modèles d'apprentissage automatique qui prévoient les scores d'inspection des infrastructures vertes, et (2) des modèles qui estiment la durée de la stagnation des eaux après les épisodes pluvieux. À l'aide d'un ensemble de données complet provenant de la ville de Philadelphie, nous développons des modèles d'apprentissage automatique qui prédisent les scores d'inspection en fonction des inspections antérieures et des variables environnementales et sociodémographiques locales. Parallèlement, nous exploitons un ensemble unique de données hydrologiques observées sur plusieurs sites d'infrastructures vertes – combinées à des données météorologiques – pour prédire la durée de stagnation de l'eau, un indicateur direct de la performance du système. En combinant ces deux modèles, nous générons des prévisions spécifiques aux événements concernant le risque de dysfonctionnement des infrastructures vertes à l'échelle de la ville. La capacité de prédire la durée de stagnation de l'eau à l'aide de données facilement accessibles représente une avancée importante vers une gestion proactive, évolutive et durable des systèmes d'infrastructures vertes urbaines.

ABSTRACT

Regular performance assessment of blue–green infrastructure (BGI) is critical for sustainable urban stormwater management, yet it remains difficult due to the complexity of urban settings, variability across BGI systems, and limited financial and staffing resources. Traditional approaches for estimating underperformance rely on physics-based hydrological models, which either oversimplify real-world conditions or require extensive data and computational effort. This study introduces an integrated, data-driven framework for evaluating BGI performance risk by linking two predictive components: (1) machine-learning models that forecast BGI inspection scores, and (2) models that estimate ponding duration following storm events. Using a comprehensive dataset from the City of Philadelphia, we develop ML models that predict inspection scores based on historical inspections and local environmental and sociodemographic variables. In parallel, we leverage a unique set of observed hydrologic data from multiple BGI sites—combined with weather inputs—to predict ponding time as a direct indicator of system performance. By coupling these two models, we generate event-specific forecasts of underperformance risk for BGI across the city. The ability to predict ponding duration using readily available data represents an important advancement toward proactive, scalable, and sustainable management of urban BGI systems.

KEYWORDS

blue-green infrastructure, data-driven, machine learning, stormwater, sustainable management

1 INTRODUCTION

The combined effects of urbanization and climate change amplify the risks posed by surface water runoff in urban environments. Blue-green infrastructures (BGIs) are nature-based, sustainable solutions that mitigate these impacts while providing significant environmental benefits. BGI is also an increasingly prevalent engineering strategy around the world to provide resilience and offset the impacts of urban stormwater. As BGI receive inflow from the dynamic, and often messy, urban landscape, factors such as litter, land cover changes, vegetation cover, and weather patterns can impact the function and performance of infiltration based BGIs over time. It is important to analyze their efficiency and effectiveness to make sure they are functioning well over time (Amur et al., 2022; Catalano de Sousa et al., 2016; Meerow et al., 2021; William et al., 2019). As these systems become increasingly prolific in the urban landscape assessing effectiveness is increasing onerous, requiring ample costs and manpower. BGI performance (or underperformance) assessment is often done through ponding duration or overflow, aligning with compliance regulations. Ponding duration provides an understanding of the overall behaviour within the system and the system's ability to manage the ponded runoff. Typically, this is estimated through physics derived hydrological modelling, requiring either simplifications of the highly dynamic nature hydrology of these systems (limiting accuracy) or complex data inputs (limiting the spatial and temporal resolution) (Al Mehedi et al., 2023).

Data-driven Machine Learning (ML) helps to discover complex, non-linear relationships between different variables where physics-based models are constrained by certain parameters. Recent advancements linking observational data and ML have created an opportunity to improve prediction of BGI performance and identify systems at risk. Historical monitoring data from BGI systems can help to identify patterns, trends, and information to make decisions for ensuring system performance (Al Mehedi et al., 2023; Han et al., 2021; Islam et al., 2022). This type of model is highly dynamic and evolving in nature, which is a helpful tool for continually updating data for BGI evaluation (Wadzuk et al., 2021) and accounting for the dynamics of the urban landscape. This study integrates two ML models, 1) predicting BGI ponding duration and 2) predicting inspection score, to predict BGI risk of underperformance.

2 METHODS

2.1 Inspection Prediction

The Philadelphia Water Department (PWD) conducts regular BGI inspections to ensure functionality and inform maintenance. Using inspection records from the PWD (i.e., a dataset from 2016-2020 was used), this study assessed the performance of basin, bioinfiltration, and permeable pavement BGI systems to predict the overall inspection rating of BGI systems. To address the labor-intensive nature of inspection-driven assessments, we develop an ML-based framework that applies Random Forest, LightGBM, CatBoost, and Gradient Boosting models with K-fold cross-validation to evaluate the capacity of environmental and social context features to predict overall BGI performance. Data preprocessing was done followed by an analysis of the correlation between each inspection data type and the final BGI performance rating. We used 20 environmental variables summarized within a defined buffer around each BGI sites such as land cover, land use, maintenance need index (MNI). Social Vulnerability Index (SVI) features consisting of four thematic areas are also included to examine the influence of socioeconomic condition on BGI performance and further support a fairness-oriented analysis. To identify the most influential environmental and social features with performance ratings, we used Pearson correlation analysis. Finally, a range ML models with cross validation were employed to assess the predictability of the overall BGI rating.

2.2 Ponding Length Prediction

The length of water ponding can be used to identify performance of BGI and as an indicator of a maintenance needs (Al Mehedi et al., 2023; Amur et al., 2022). This study incorporated both observed and simulated data to predict ponding depth. All the observed data were gathered from Villanova Center for Resilient Water System (VCRWS) database (Smith et al., 2023). Several types of BGI were equipped with pressure transducers to collect these pressure data and to calculate water depth data. The precipitation is monitored using a tipping bucket rain gauge. A hydrological model, USEPA Storm Water Management Model (SWMM 5.2), was developed using observed data from Villanova databasae and calibrated to complement observed data to enhance the abilities of the ML model.

Data preprocessing and storm identification were conducted using Python in Google Colab. The initial data cleaning process included detection and handling of missing values and outliers. Storm events were characterized by hydrographs displaying clear rising limbs, peaks, and recession limbs. For each identified storm at each site, we calculated characteristic features describing the dry conditions and weather. We supplemented these features with pre-existing site parameters: contributing drainage area, footprint area, hydraulic control, and type of system. We used linear regression (LR), random forest (RF), extreme gradient boosting (XGB), Catagorical Boosting (CatBoost), Light Gradient-Boosting Machine (LightGBM) and support vector machine (SVM) models. Models were developed using features (precipitation, storm duration, average intensity, maximum intensity, antecedent dry days, Julian days) as the predictor and one feature (ponding duration) as the target. The coefficient of determination R-squared along with MAE and MAPE were used in this study to evaluate the model performance.

2.3 Performance Prediction

The framework for predicting BGI performance builds upon the inspection and ponding predictive models, and integrates predicted ponding duration with site-specific predicted inspection data through a sensitivity-based adjustment process. A sensitivity analysis is used to identify a range of ponding durations based on site characteristics for a given storm event and BGI unit and altering the hydraulic conductivity of the system (KSAT). Scenarios, varying the KSAT, are generated by running SWMM simulations. The KSAT represents an effective infiltration parameter, which for this study is correlated with inspection score (a high infiltration rate represents better performance, while a lower infiltration rate represents worse performance). Running a series of scenarios with a range of KSAT values produced a dataset of ponding times for a BGI system per event. A new ML model for ponding duration, which includes a predicted site inspection score, is then used to predict ponding duration. This new ponding duration value is based on an adjusted KSAT related to the inspection score. If the site is predicted to be in poor condition, the poor inspection score will result in a lower condition KSAT, causing a longer ponding duration. For high predicted inspection score of well-maintained sites an ideal condition KSAT value is selected (matching the design), resulting in a lower ponding time. The predicting ponding time can then be used to indicate if the BGI is at risk of being out of compliance with regulatory requirements.

3 RESULTS AND DISCUSSION

The ML models are evaluated to predict BGI performance ratings using environmental features, social vulnerability features, and a combined feature set. Across all models, environmental features outperform social features when used independently. Combing these two types of features improves prediction accuracy. LightGBM achieved highest average accuracy (61.7%) when all features were included. While RF and CB also shows strong stable performance across feature groups. It can be said that these two feature groups combinedly provide strong prediction of BGI functionality and integrating both features improves model reliability.

In ponding duration prediction model, across all ML algorithms. RF, SVM, CatBoost, and LightGBM consistently achieved the highest predictive accuracy, with average R-squared values exceeding 0.80. Models are run with both simulated and observed data to increase the range of scenarios. The models show deviations for high and low ponding duration as model tries to fit these extreme values thus reducing their ability to generalize unseen data. Model performance is improved by binning the database by storm size. It was found that storm duration, BGI type, size of BGI have highest importance in predicting ponding duration compared to the other input features.

A predictive model for risk of underperformance combines the predictive model for inspection score with the ponding duration model. This allows for a new ponding duration based on inspection duration. The models are generally successful at capturing the trends and fluctuations of the datasets. Most of the model shows that in predicting performance of BGI storm duration are very important. However, there is limited observed data for underperforming systems, which limits the ability to test this framework using real world BGI. The team is now developing a monitoring network of sites to capture this data.

4 CONCLUSIONS AND FUTURE WORK

The findings of this study suggest that more targeted, efficient, and proactive municipal maintenance practices are feasible using a data-driven approach. This type of approach as the potential to enhance BGI longevity and effectiveness. More broadly, ML models prove especially useful in extracting patterns and insights from complex, irregular datasets such as inspection records, offering a scalable tool for infrastructure management. The developed ML models show strong predictive capabilities for both inspection score and ponding duration model. Combing these two models has the potential to develop an integrated framework to forecast risk of underperformance.

The developed framework is needed validation, so collection of inspection data along with ponding data for storms is on-going. This project has the potential to aid in the optimization of BGI maintenance and management, reduce labor-intensive inspections, and support evidence-based decision-making for urban stormwater infrastructure. It can assist city agencies in prioritizing maintenance efforts, allocating resources more efficiently, and responding proactively to potential system underperformance.

LIST OF REFERENCES *(only for scientific papers)*

Al Mehedi, M. A., Amur, A., Metcalf, J., McGauley, M., Smith, V., & Wadzuk, B. (2023). Predicting the performance of green stormwater infrastructure using multivariate long short-term memory (LSTM) neural network. Journal of Hydrology, 625, 130076.

Amur, A., Wadzuk, B., & Traver, R. (2022). A 15 year analysis of precipitation and rain garden response. Hydrological Processes, 36(11), e14736.

Catalano de Sousa, M. R., Montalto, F. A., & Gurian, P. (2016). Evaluating green infrastructure stormwater capture performance under extreme precipitation. Journal of Extreme Events, 3(02), 1650006.

Han, H., Choi, C., Jung, J., & Kim, H. S. (2021). Deep learning with long short term memory based sequence-to-sequence model for rainfall-runoff simulation. Water, 13(4), 437.

Islam, S. M. N., Thakur, N., Garg, K., & Gupta, A. (2022). A recent survey on lstm techniques for time-series data forecasting: Present state and future directions. Applications of Artificial Intelligence, Big Data and Internet of Things in Sustainable Development, 123–132.

Meerow, S., Natarajan, M., & Krantz, D. (2021). Green infrastructure performance in arid and semi-arid urban environments. Urban Water Journal, 18(4), 275–285.

Smith, V. B., McGauley, M. W., Newman, M., Garzio-Hadzick, A., Kurzweil, A., Wadzuk, B. M., & Traver, R. (2023). A relational data model for advancing stormwater infrastructure management. Journal of Sustainable Water in the Built Environment, 9(1), 4022023.

Wadzuk, B., Gile, B., Smith, V., Ebrahimian, A., Strauss, M., & Traver, R. (2021). Moving toward dynamic and data-driven BGI maintenance. Journal of Sustainable Water in the Built Environment, 7(4), 2521003.

William, R., Gardoni, P., & Stillwell, A. S. (2019). Reliability-based approach to investigating long-term clogging in green stormwater infrastructure. Journal of Sustainable Water in the Built Environment, 5(1), 4018015.