

MAINTENANCE PREDICTIVE DES OUVRAGES D'ART AVEC DES FONDATIONS EN SITES AQUATIQUES

PREDICTIVE MAINTENANCE FOR BRIDGES CROSSING WATERWAYS

Tianyu WANG^{1,2}, Philippe REIFFSTECK², Christophe CHEVALIER², Chi-Wei CHEN¹, Franziska SCHMIDT³

¹ EMF-EXP, Département des Ouvrages d'Art, DTR, DGII, SNCF Réseau, La Plaine Saint Denis Cédex, 93212, France

² Univ Gustave Eiffel, GERS-SRO, F-77454 Marne-la-Vallée, France

³ Univ Gustave Eiffel, MAST-EMGCU, F-77454 Marne-la-Vallée, France

RÉSUMÉ – L'affouillement est l'une des principales causes de défaillance des ponts. Cette étude a établi un modèle par machine learning (ML) pour la prédiction du risque d'affouillement des ponts en utilisant les données de la SNCF. Il s'avère que l'algorithme "extreme gradient boosting" (XGBoost) surpasse les autres algorithmes et obtient des résultats très encourageants. La méthodologie proposée dans cet article permet aux ingénieurs d'avoir une estimation du risque d'affouillement des piles de pont de manière précise et rapide.

ABSTRACT – Scour is one of the major causes for bridge failures. This study established a machine learning (ML) based framework for bridge scour risk prediction by using SNCF data. It is found that *extreme gradient boosting* (XGBoost) algorithm outperforms other algorithms and achieves very encouraging results. The proposed methodology in this paper allows engineers to evaluate scour risk of bridge piers in an accurate yet rapid way.

1. Introduction

Le nombre de catastrophes naturelles a beaucoup augmenté au cours des 50 dernières années en raison de l'effet du changement climatique. Les inondations sont l'une des catastrophes naturelles les plus courantes (*Increase in Natural Disasters on a Global Scale by Ten Times*). Des études ont montré que les dommages induits par les événements hydrauliques sont les principales causes de défaillance des ponts. Entre 1989 et 2000 aux Etats-Unis, plus de 50% des ruptures de ponts ont été causées par des affouillements et autres facteurs hydrauliques (Wardhana & Hadipriono, 2003). En 1978, en France, le pont Wilson s'est effondré, et près de 100 000 personnes ont été impactées par les conséquences dans la ville de Tours. Il s'agissait d'un pont en maçonnerie de 15 arches construit en 1765-1778.

La politique actuelle de la maintenance des ponts avec fondations en site aquatique se base principalement par l'inspection sur le terrain. Les méthodes d'inspection couramment utilisées sont le pataugeage, la plongée, ou parfois l'essai d'impact pour voir ou détecter l'affouillement de la pile du pont (approche couramment utilisée par le Railway Technical Research Institute). Sur la base des résultats de l'inspection, un niveau de risque est donné par l'ingénieur expert et, le cas échéant, des travaux de confortement sont programmés, tels que des enrochements, des gabions, etc. Il convient de noter que le processus de décision est influencé par la subjectivité des ingénieurs, les questions budgétaires, l'impact social, etc.

L'affouillement des ponts induit par l'action hydraulique est un processus complexe et évalué de manière très souvent empirique. Dans la pratique, les guides actuels (Cerema,

2019; Takayanagi T. et al., 2018) s'inscrivent dans le cadre de la maintenance préventive ou réactive, ce qui signifie que les travaux de régénération sont effectués régulièrement ou lorsque les ouvrages ont été gravement endommagés. Compte tenu de la nature stochastique des cours d'eau, une telle politique de maintenance peut ne pas être en mesure d'assurer la sécurité des ouvrages de manière optimale.

Grâce au développement du big data et de l'intelligence artificielle (IA), il existe aujourd'hui différents types d'applications ML dans le domaine des assurances, de la médecine, etc. Du point de vue des ingénieurs ferroviaires, l'évaluation du risque d'affouillement a toujours été un défi majeur puisque le danger et la vulnérabilité des ponts face à l'affouillement sont complexes à évaluer. Cependant, les modèles ML, qui sont bons pour trouver des relations implicites entre variables, devraient aider à révéler les mécanismes sous-jacents du système sous-sol - cours d'eau - fondation. Dans ce but, cette étude vise à construire une maintenance prédictive basée sur le ML pour l'affouillement des ponts, de sorte à évaluer le risque et programmer à temps les travaux de maintenance correspondants.

Les données utilisées dans cette étude ont été fournies par la SNCF. Après cette introduction, l'article est organisé comme suit : la section 2 présente la méthodologie concernant le choix des caractéristiques d'entrée et des algorithmes ML. Les résultats des phases d'entraînement, puis de tests utilisant les algorithmes de la section 2 sont ensuite présentés dans la section 3. La section 4 présente les défis à relever pour appliquer une politique de maintenance prédictive basée sur le ML.

2. Méthodologie

La maintenance prédictive est une méthode dans laquelle la durée de vie des éléments structuraux importants est prédite sur la base d'une inspection ou d'un diagnostic. Par rapport à une maintenance périodique préventive, la maintenance prédictive est basée sur les conditions réelles (Mokhtab et al., 2018). Les approches ML sont utilisées pour la définition de l'état réel du système ainsi que pour la prévision de ses états futurs (Susto et al., 2015). Pour développer un modèle d'affouillement de pont basé sur le ML, les quatre étapes essentielles sont respectivement : la collecte des données, le stockage et le prétraitement des données, le développement du modèle et la maintenance du modèle. Le présent document se concentre principalement sur l'étude des étapes de collecte de données et de développement du modèle.

2.1 Collecte des données

Les données utilisées dans cette étude proviennent des rapports d'inspection de la SNCF. Les infrastructures ferroviaires sont inspectées régulièrement en plongeant ou en pataugeant. Après chaque inspection, il est demandé aux ingénieurs de donner un niveau de risque. L'affouillement des ponts est un phénomène naturel complexe : selon les directives pratiques actuelles, les caractéristiques du pont, le milieu environnant ainsi que les observations sur site peuvent tous influencer le niveau de risque au final. Dans cette étude, 18 variables/mesures ont été sélectionnées comme données d'entrée. Elles sont numériques et catégoriques. Finalement, l'ensemble de données contient 208 jeux de mesures (208 piles de pont) sur 75 ponts. La Table 1 montre un exemple de jeu de données.

Table 1. Un exemple de jeu de données (pour une pile de pont)

Type d'écoulement	Pente du lit (%)	Débit de crue (m ³ /s)	Rapport d'encaissement	Topographie
Fluvial	0.19	97.82	15.36	Plaine
Sinuosité du cours d'eau	Matériaux du lit	Forme de la pile	Type de la fondation	Protection de la fondation
Quasi rectiligne	Sable	Circulaire ou cylindrique	Superficielle	Oui
Historique de l'affouillement	Historique des inondations	Susceptible de l'affouillement	Classement des chaînes	Classement des berges
Non	Non	Oui	Mauvais	Moyen
Dislocation ou déformation	Affouillement local	Classement des autres avaries		
Oui	Non	Très bien		

2.2 Développement du modèle

Une fois les données collectées et prétraitées, elles peuvent être entraînées dans des modèles ML. Cette étude utilise l'apprentissage supervisé en supposant que les données de sortie sont le niveau de risque donné par les ingénieurs (2 classes : élevé ou faible). Pour construire un modèle ML, 70% des données sont utilisées pour l'entraînement du modèle, et le reste pour son test. Une matrice de confusion (Table 2) a été utilisée pour évaluer les performances du modèle, à partir de laquelle quatre mesures peuvent être calculées : l'exactitude, la précision, le rappel et le taux de faux positifs (indiqués dans les Eq. (1) - Eq. (4)).

Table 2. Matrice de confusion

		Classe estimée	
		Positive	Négative
Classe réelle	Positive	Vrais positifs (VP)	Faux négatifs (FN)
	Négative	Faux positifs (FP)	Vrais négatifs (VN)

$$\text{Exactitude} = \frac{VP+VN}{VP+FN+FP+VN} \quad (1)$$

$$\text{Précision} = \frac{VP}{VP+FP} \quad (2)$$

$$\text{Rappel} = \frac{VP}{VP+FN} \quad (3)$$

$$\text{Taux de faux positif} = \frac{FP}{FP+VN} \quad (4)$$

Trois algorithmes couramment utilisés pour résoudre les problèmes géotechniques ont été choisis dans cette étude. Il s'agit respectivement de la machine à vecteurs de support (SVM – "Support Vector Machine") (Cortes & Vapnik, 1995), de la forêt d'arbres de décision (RF – "Random Forest") (Breiman, 2001) et de l'"eXtreme Gradient Boosting" (XGBoost) (Chen & Guestrin, 2016).

La machine à vecteurs de support (SVM) est basée sur une fonction linéaire plus un terme de biais. Pour séparer les points de données, un hyperplan optimal est défini dans le SVM. L'objectif est de s'assurer que l'hyperplan présente la distance maximale entre les points de données des deux classes.

La forêt d'arbres de décision (RF) est une méthode d'apprentissage d'ensemble qui regroupe de nombreux arbres de décision. Un arbre unique est un classificateur faible en raison de sa variance élevée. Pour surmonter ce problème, la RF forme une forêt en générant un grand nombre d'arbres de décision. L'objectif de RF est de faire croître chaque arbre de décision individuellement en adoptant la méthode de ré-échantillonnage *bagging*.

XGBoost est proposé par (Chen & Guestrin, 2016) sur la base de modèles d'arbres. Il a été reconnu dans un grand nombre de compétitions Machine Learning et de data mining. Lors du défi Kaggle 2015, près de 60% des solutions gagnantes publiées utilisaient XGBoost.

Il est développé en utilisant l'algorithme boosting de gradient (gradient boosting) dans le cadre de l'apprentissage d'ensemble (ensemble learning). Par rapport à un seul arbre de décision (également connu sous le nom de CART, Breiman et al., 1984), l'apprentissage d'ensemble génère plusieurs arbres de décision afin d'obtenir une grande précision dans la prédiction. Dans l'algorithme d'amplification du gradient, il combine plusieurs apprenants faibles en un apprenant fort en ajoutant séquentiellement des prédicteurs pour corriger son prédécesseur. L'erreur résiduelle est minimisée par l'ajout d'arbres CART. Par rapport à l'algorithme traditionnel de boosting de gradient (gradient boosting), l'algorithme XGBoost ajoute un terme régularisé pour pénaliser la complexité du modèle afin d'éviter "overfitting". Une présentation mathématique plus détaillée concernant l'algorithme XGBoost peut se référer au travail de Chen et Guestrin (2016).

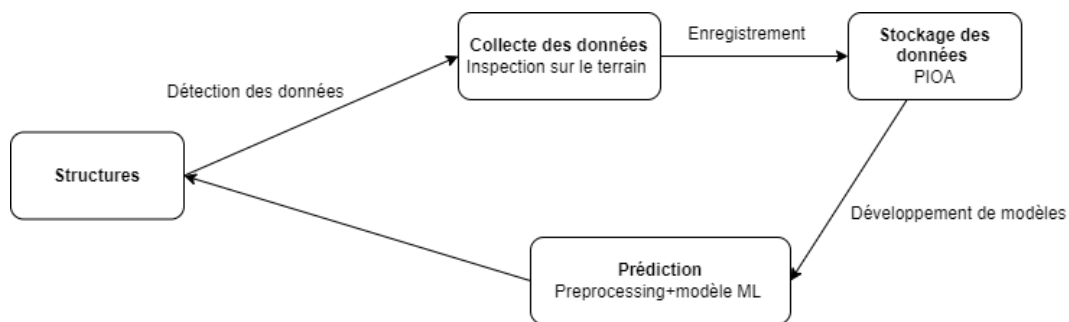


Figure 1. Présentation schématique d'un flux de travail basé sur le ML pour l'affouillement des ponts

La Figure 1 représente schématiquement le flux de travail d'un système de maintenance basé sur le ML pour l'affouillement des ponts. Les données recueillies par les ingénieurs sur le terrain sont d'abord enregistrées dans PIOA (Patrimoine Informatisé Ouvrages d'Arts) qui est une plateforme interne et numérisée de la SNCF destinée aux infrastructures ferroviaires. Ensuite, les données sont prétraitées et utilisées par un modèle ML pré-entraîné. Le niveau de risque donné par le modèle ML est considéré comme un complément à la décision de l'ingénieur. Les travaux correspondants de contre-mesure à l'affouillement seront programmés par la suite.

3. Résultats actuels

Aléatoirement, 70 % des données ont été utilisées pour entraîner le modèle et les 30 % restants pour le tester. Afin de trouver l'algorithme optimal, les mêmes données ont été entraînées et testées en utilisant trois algorithmes différents : SVM, RF et XGBoost. Les résultats de l'entraînement et du test sont présentés respectivement dans le Table 3 et le Table 4.

Table 3. Résultats sur les données d'entraînement

	Exactitude	Précision	Rappel	Taux de faux positif
SVM	0.89	0.93	0.91	0.17
RF	0.92	0.91	0.99	0.24
XGBoost	0.97	0.96	0.99	0.01

Table 4 Résultats sur les données de test

	Exactitude	Précision	Rappel	Taux de faux positif
SVM	0.87	0.94	0.91	0.25
RF	0.91	0.90	0.98	0.42
XGBoost	0.92	0.98	0.92	0.08

On peut observer que parmi les trois algorithmes d'apprentissage, le classificateur XGBoost a la plus grande exactitude, la plus grande précision et le plus faible taux de faux positif pour l'ensemble des algorithmes. RF a le rappel le plus élevé, mais il a toujours le taux de faux positif le plus élevé. En conclusion, le modèle basé sur XGBoost a une performance très satisfaisante par rapport aux autres modèles ML, compte tenu de toutes les mesures de performance du modèle proposé.

4. Défis et travaux futurs

Bien que le classificateur XGBoost ait des performances satisfaisantes, cela ne signifie pas que la méthode proposée est prête à être appliquée sur le terrain. L'un des plus grands défis est l'interprétabilité du modèle.

Le modèle ML est souvent considéré comme une boîte noire. Bien que des résultats de prédiction encourageants aient été montrés dans divers types de problèmes de génie civil en utilisant différentes métriques, La société et la communauté scientifique ne s'intéressent pas seulement aux performances des modèles. L'obtention de bons résultats avec des mesures sélectionnées n'indique effectivement pas que le mécanisme derrière le problème est révélé, mais plutôt que l'algorithme spécifique peut prédire correctement dans l'ensemble des situations établies et dont les données sont considérées comme suffisamment représentatives du phénomène. En outre, l'évaluation du risque d'affouillement d'un pont est liée à sa défaillance. Dans un tel projet à haut risque, une mauvaise classification peut éventuellement menacer l'ouvrage, son environnement et avoir un coût humain et financier important. Par conséquent, il est nécessaire de savoir comment chaque prédiction est faite à partir du modèle de boîte noire et dans quelle mesure il faut faire confiance au modèle.

Outre l'interprétabilité du modèle, pour s'assurer de l'application d'un modèle ML, d'autres aspects doivent également être étudiés pour augmenter la performance et la fiabilité du modèle (Shahin, 2013). Il s'agit respectivement de : la robustesse du modèle (la capacité prédictive à généraliser sur une gamme de données similaire à celle utilisée pour

l'entraînement du modèle), l'extrapolation du modèle (la capacité du modèle à prédire de manière appropriée en dehors de la gamme des données utilisées pour la calibration du modèle) et l'incertitude du modèle (la quantification de l'incertitude de la prédiction du modèle).

En ce qui concerne la méthodologie proposée, il n'existe pas de relation explicite entre le nombre de variables et le nombre de données. Le classificateur entraîné sera testé sur un plus grand nombre de ponts à la SNCF pour voir sa précision et sa robustesse dans la prédiction. En parallèle, le modèle pourrait être optimisé avec plus de cas de test dans le futur.

5. Conclusions

La vulnérabilité des ponts à l'affouillement s'est considérablement accrue ces dernières années sous l'effet du changement climatique. Or, les gestionnaires d'infrastructures doivent être préparés à des catastrophes naturelles imprévues et aux éventuelles conséquences sur leurs ouvrages. Dans ce but, une politique de maintenance prédictive basée sur le ML a été présentée dans cette étude. Avec les données actuelles, le classificateur XGBoost surpasse les autres algorithmes et obtient de très bons résultats pour l'exactitude (0,92), la précision (0,98) et le rappel (0,92) ainsi qu'un faible taux de faux positifs (0,08). Cependant, en plus de l'obtention de performances élevées, l'interprétabilité du modèle doit aussi être prise en compte, en particulier lorsqu'il s'agit d'un projet à haut risque où une mauvaise classification peut avoir un impact énorme. Les travaux futurs se concentreront donc sur le test d'un plus grand nombre de cas afin de prouver la robustesse du modèle et de rendre le modèle transparent et facile à comprendre pour les ingénieurs utilisateurs.

Un processus plus détaillé concernant la construction du modèle ML a été soumis au journal *European Journal of Environmental and Civil Engineering*, qui est actuellement en cours de révision.

6. Références bibliographiques

- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Cerema. (2019). *Analyse de risque des ponts en site affouillable*.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13-17-Aug, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/bf00994018>
- Increase in Natural Disasters on a Global Scale by Ten Times*. (n.d.). Retrieved November 26, 2021, from <https://www.visionofhumanity.org/global-number-of-natural-disasters-increases-ten-times/>
- Mokhatab, S., Poe, W. A., & Mak, J. Y. (2018). Handbook of natural gas transmission and processing: Principles and practices. *Handbook of Natural Gas Transmission and Processing: Principles and Practices*, 1–826. <https://doi.org/10.1016/C2017-0-03889-2>
- Shahin, M. (2013). Artificial Intelligence in Geotechnical Engineering: Applications, Modeling Aspects, and Future Directions. *Metaheuristics in Water, Geotechnical and Transport Engineering*, 169–204. <https://doi.org/10.1016/B978-0-12-398296-4.00008-8>
- Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2015). Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Transactions on Industrial Informatics*, 11(3), 812–820. <https://doi.org/10.1109/TII.2014.2349359>

- Takayanagi T., Naito N., Manome R., N. O. (2018). Evaluation method using score table for identifying bridge piers vulnerable to scouring in Japan. In *Scour and Erosion IX* (1st Editio). CRC Press.
- Wardhana, K., & Hadipriono, F. C. (2003). Study of Recent Building Failures in the United States. *Journal of Performance of Constructed Facilities*, 17(3), 151–158. [https://doi.org/10.1061/\(asce\)0887-3828\(2003\)17:3\(151\)](https://doi.org/10.1061/(asce)0887-3828(2003)17:3(151))