

# IDENTIFICATION AUTOMATIQUE DES SOURCES DE VIBRATIONS PAR INTELLIGENCE ARTIFICIELLE POUR LE SUIVI DE CHANTIER

## AUTOMATIC IDENTIFICATION OF VIBRATION SOURCES BY ARTIFICIAL INTELLIGENCE FOR SITE MONITORING

Etienne REY<sup>1</sup>, Nickolas STELZENMULLER<sup>2</sup>

<sup>1</sup> GEOLITHE, Crolles, France

<sup>2</sup> GEOLITHE INNOV, Crolles, France

**RÉSUMÉ** – Dans le cadre de l'instrumentation pour suivi et contrôle des vibrations sur les enjeux/ouvrages avoisinant les chantiers de génie civil, nous proposons ici l'utilisation d'un algorithme d'intelligence artificielle permettant une identification automatisée des vibrations enregistrées, afin de faciliter le suivi régulier à distance par les différents acteurs du chantier, et de discriminer les sources vibratoires liées au chantiers du bruit environnant.

**ABSTRACT** – As part of the instrumentation for monitoring and controlling vibrations on the stakes/structures neighboring civil engineering sites, we propose here the use of an artificial intelligence algorithm allowing automated identification of the recorded vibrations, in order to facilitate regular remote monitoring by the various actors on the site, and to discriminate between the vibration sources linked to the sites and the surrounding noise.

### 1. Introduction

Des dispositifs de suivi et de contrôle des vibrations sont de plus en plus souvent installés sur les enjeux avoisinant les chantiers de génie civil générateurs de vibrations (vibrofonçage ou battage de profilés métalliques dans le sol, terrassement par vibrocompactage, brise-roche hydraulique ou minage, chantiers d'amélioration de sols par compactage dynamique ou vibro-flottation, creusement ou déroctage à l'explosif...), afin de se prémunir d'éventuels dommages sur les ouvrages avoisinants par effet direct des vibrations.

En plus des systèmes d'alerte généralement installés sur site (alarmes, alertes par SMS ou e-mail), un suivi régulier à distance est de plus en plus fréquemment demandé par les différents acteurs du chantier (MOE, AMO, entreprise...). Généralement le bureau de contrôle en charge du suivi procède à une validation des détections comme étant liées aux travaux, avec pour objectif d'écarter d'éventuelles fausses alertes. Cependant, lorsque les travaux ont lieu en période nocturne ou le week-end, une automatisation de l'identification des signaux prend tout son sens, pour gagner en réactivité, éviter des astreintes coûteuses et, d'une manière plus générale, de permettre une plus grande autonomie à l'entreprise de travaux dans la gestion de ses émissions vibratoires.

Dans ce cadre, et afin de discriminer les sources de vibration générées par les travaux du chantier faisant l'objet du suivi d'autres sources éventuelles de vibrations environnantes (train, circulation non attribuable aux travaux, artefact de mesure...), nous avons testé l'utilisation d'un algorithme spécifique d'intelligence artificielle pour identifier la signature de différentes sources vibratoires émises par un chantier de la région dont GEOLITHE avait en charge le suivi vibratoire : le remplacement du tablier métallique du viaduc ferroviaire franchissant la rivière Ay au nord du département de l'Ardèche, et sur la ligne de fret régionale entre Lyon et Marseille.

## 2. Chantier test

### 2.1. Contexte

Le chantier de remplacement du tablier métallique du viaduc ferroviaire de l'Ay comportait une première phase qui consistait à battre des tubes métalliques de fondation provisoires sur les berges et dans le lit du cours, destinés à préparer la dépose par translation latérale (sur des traverses métalliques, dans le sens de la rivière et perpendiculairement aux voies) de l'ancien tablier, puis le lancement et la mise en place par le même procédé du nouveau tablier. Ces tubes provisoires en acier avaient des diamètres atteignant 1220 mm pour 12 de longueur.

La seconde phase consistait à mettre en fiche par battage des tubes d'acier de diamètre 1400 mm et de longueur moyenne 14 mètres, en position définitive, à proximité immédiate (moins de 1 mètre) des culées maçonnées de l'ouvrage existant. C'est durant cette phase délicate, effectuée sous coupure ferroviaire, qu'ont été réalisés les tests d'identification. Les tubes, remplis de béton après mise en fiche à leur profondeur définitive ont pour vocation de servir de fondation aux culées du nouveau tablier, plus lourd que l'ancien modèle.

Compte tenu de la manutention de longs tubes à très faible distance des ouvrages ferroviaires (notamment des voies) ces travaux ont été réalisés sous coupure de circulation ferroviaires durant 4 week-ends consécutifs du printemps 2021.

### 2.2. Type de travaux

La mise en fiche des tubes a été réalisée essentiellement par battage, après un positionnement (mise à la verticale) au vibreur. L'outil utilisé pour le battage était un marteau IHC S-70 de masse frappante 5,5 Tonne et d'énergie de frappe 70 kJ par coup (Figure 1).



Figure 1. Mise en fiche des tubes définitifs par battage

### 2.3. Instrumentation sur site et suivi à distance

Le dispositif de suivi vibratoire était constitué de 4 capteurs nommés V1 à V4, qui étaient déplacés selon l'avancement du chantier, au plus proche du quadrant encours de réalisation (Figure 2).

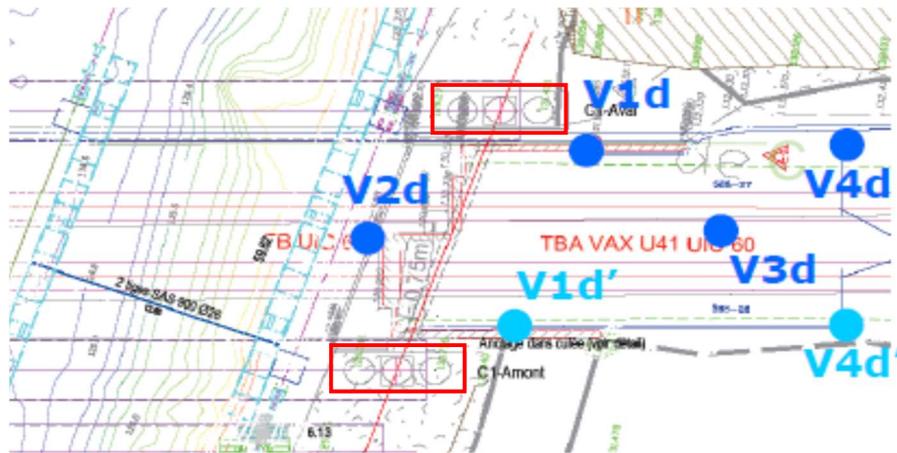


Figure 2. A gauche : implantation des capteurs de vibration sur la culée C1 (rive droite) côté aval (V1d à V4d). Les 2 capteurs excentrés par rapport aux voies (V1d et V4d) sont déplacés en positions V1d' et V4d' lorsque le chantier passe dans le quadrant C1-amont. Les pieux définitifs faisant l'objet des travaux sont encadrés en rouge). A droite : implantation du capteur V4 sur le plot de fondation d'un poteau caténaire.

La chaîne de mesurage était composée de 4 capteurs triaxiaux (géophones tri-directionnels) de fréquence propre 2 Hz ou 4.5 Hz, reliés par 2 à 2 centrales de mesure de type SAVIB 6v2 (marque Myotis). En plus de stockage permanent des valeurs crêtes pour chacune des voies sous forme d'histogramme des maxima, chaque station fonctionne en mode zéro-crête, c'est-à-dire qu'en cas de dépassement du seuil en vitesse particulière prédéfini par l'utilisateur (généralement 75% du seuil admissible pour chacun des ouvrages instrumentés), elle enregistre les signaux temporels immédiatement consécutifs au dépassement du seuil, échantillonnés à 1000 Hz sur les 6 voies (2 x 3 voies / capteur) simultanément, pour une durée de 10 secondes.

### 2.4. Suivi à distance

Ces signaux sont immédiatement transmis à distance sur un serveur FTP, distant pour traitement du signal, affichage, identification (« manuelle » ou automatique) de la nature des signaux comme issus (ou non) des travaux du chantier (battage en l'occurrence) et enfin validation.

## 3. Méthodologie suivie

### 3.1 Problématique

Les données qui proviennent d'évènements vibratoires sont des séries temporelles, mesurées en vitesse particulière (en mm/s). Chaque fois que le seuil de vitesse est dépassé un fichier est créé et transmis, dans lequel il y a six séries temporelles de 10 secondes (1 par voie de chacun des capteurs). Une série temporelle typique d'une opération de battage est affichée sur la figure 3.

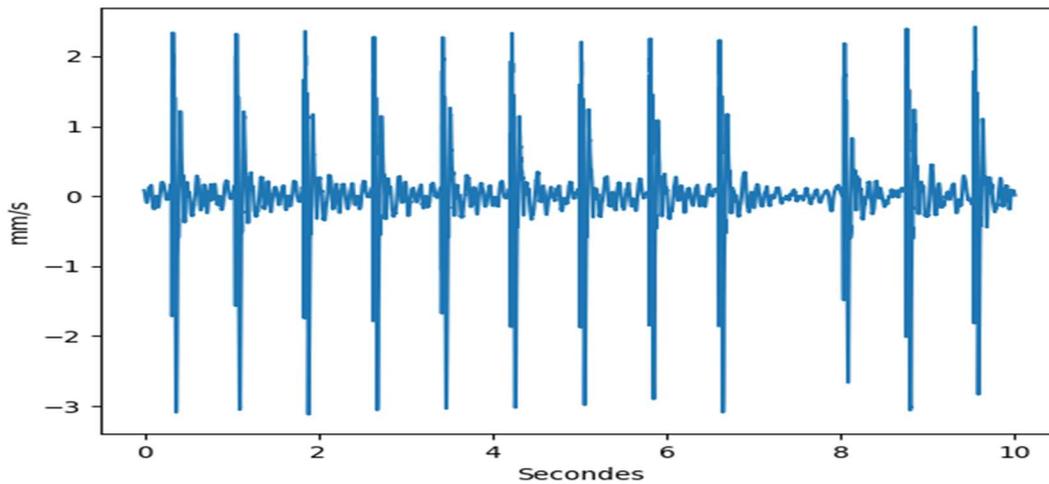


Figure 3. Série temporelle typique ayant pour origine le battage des tubes métalliques.

Dans notre cas d'étude, le problème se résume à une classification binaire de séries temporelles, l'objectif étant de discriminer les événements vibratoires causés par une opération de battage des autres sources de vibrations (principalement des trains en l'occurrence).

Les séries temporelles de vibration ayant d'autres origines que le battage sont beaucoup plus diverses, allant de signaux au spectre large (voir figure. 4, début du signal) à des pics isolés (figure 2, seconde partie du signal).

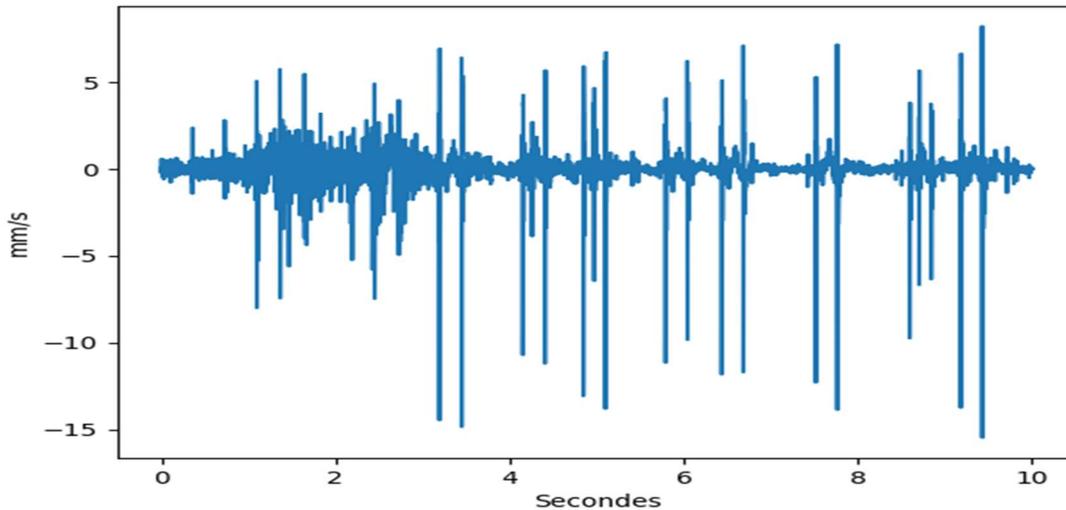


Figure 4. Série temporelle d'origine autre qu'une opération de battage.

À première vue, les signaux représentés sur les figures 1 et 2 peuvent paraître faciles à classer à partir de leur contenu fréquentiel. Cependant, une classification par analyse spectrale, basée sur le pic principal dans le spectrogramme par exemple, n'a pas une précision très élevée. Ce manque de précision est expliqué par la diversité de caractéristiques dans chaque catégorie («battage» et «non-battage»). Afin d'augmenter la précision et robustesse de la classification nous avons adopté une approche d'apprentissage automatique.

### 3.2. Algorithme utilisé

Un algorithme d'apprentissage automatique a été appliqué à cette problématique de classification, à savoir une variante de l'approche «Random forest», conçue pour l'exploitation de caractéristiques spectrales pour la classification de séries temporelles (Fawaz et al., 2019 ; Ruiz et al., 2021). Cet algorithme RISE («Random Interval Spectral Ensemble») classe des séries temporelles utilisant des caractéristiques comme des coefficients d'auto-régression et des coefficients d'autocorrélations sur plusieurs échelles de temps. En principe, il n'est pas nécessaire de spécifier en amont des caractéristiques typiques de chaque catégorie, par exemple la plage de fréquence typique de battage. L'algorithme utilise plutôt les données labélisées pour entraîner un modèle. Ce modèle comprend les caractéristiques spectrales les plus pertinentes pour la classification de ces données.

### 3.3 Apprentissage

L'apprentissage du modèle d'une manière supervisée nécessite une base de données labélisées, c'est-à-dire validées par un autre moyen (analyse experte et/ou observations réalisées sur site). La construction de cette base de données est l'étape la plus chronophage, car souvent les données doivent être labélisées «à la main» (analyse experte par un spécialiste).

Dans notre cas d'étude, nous avons utilisé comme support le portail de surveillance Evorisk, développé spécifiquement par Géolithe, qui permet notamment d'organiser (classer) et afficher tout événement, pour faciliter sa labélisation. Par ailleurs, chaque événement / détection enregistrée sur site donne lieu à la création d'un fichier dans lequel il y a six séries temporelles, donc la labélisation d'un événement nous fournit six séries temporelles labélisées. Ces éléments nous ont permis de construire rapidement une base de données labélisée suffisante pour réaliser la phase d'apprentissage automatique par le modèle RISE.

Pour éviter un biais dans le modèle, nous avons sélectionné des séries temporelles pour en avoir un nombre comparable dans chaque catégorie (« battage » et « non-battage »). Ce jeu de données a été ensuite divisé comme suit : 75% pour l'entraînement du modèle et 25% pour l'évaluation de modèle. Ces chiffres sont résumés dans le tableau 1.

Tableau 1 : Nombre d'événements et de séries temporelles par catégorie

Type	Événements labélisés	Séries temporelles	Séries temporelles pour l'entraînement	Séries temporelles pour l'évaluation
Battage	249	1494	1120	374
Non-battage	255	1530	1148	382

Afin que le modèle soit capable de classer correctement les événements, le jeu de données d'entraînement doit inclure des exemples d'événements atypiques. A titre d'exemple, le battage d'un tube métallique avec un marteau donne des vibrations généralement régulières en fréquence et en amplitude sur l'ensemble de la série temporelle de dix secondes. Cependant, cette opération peut être parfois atypique, par exemple lorsque par exemple elle s'arrête avant la fin de l'enregistrement de 10 secondes (Figure 5). Même s'il apparaît assez évident que cet événement est une opération de battage, si le jeu de données d'entraînement ne comprend pas suffisamment d'exemples d'événements atypiques, le modèle risque d'être surentraîné sur des données plus (trop) typiques, et de

ne pas être capable de reconnaître des cas comme celui de la figure 5. La taille minimum de jeux de données nécessaire pour l'entraînement d'un modèle est une fonction de la fréquence relative de ces événements atypiques.

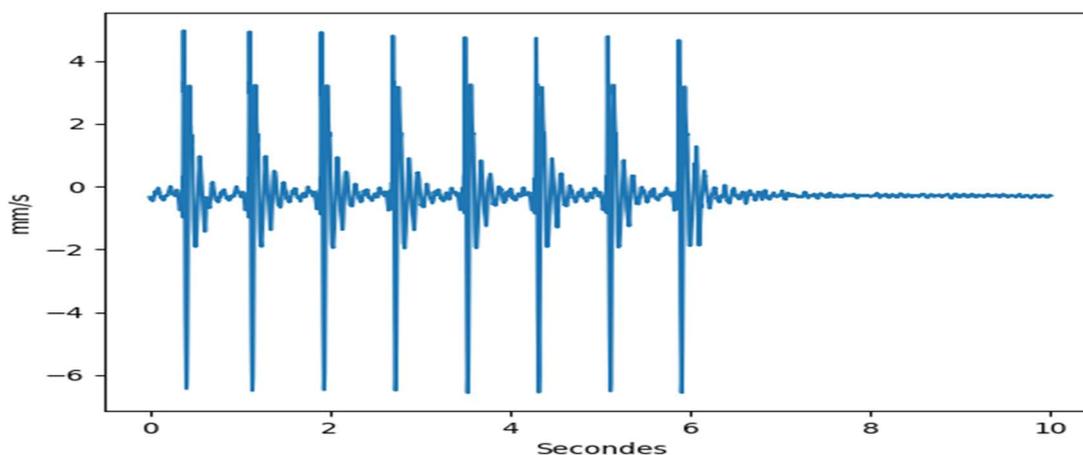


Figure 5 : Une opération de battage atypique

#### 4. Résultats

Une fois entraîné, le modèle peut classifier de nouveaux événements vibratoires (inconnus). Le résultat d'une telle évaluation est une probabilité que l'événement appartienne à la catégorie «battage». Un histogramme représentant ces probabilités, suite à l'évaluation du jeu de données réservé à cet effet, est affiché en figure 6. Rappelons qu'un modèle idéal donnerait une distribution bimodale concentrée sur 0 et sur 1. Tout chevauchement représente une ambiguïté dans la classification des événements.

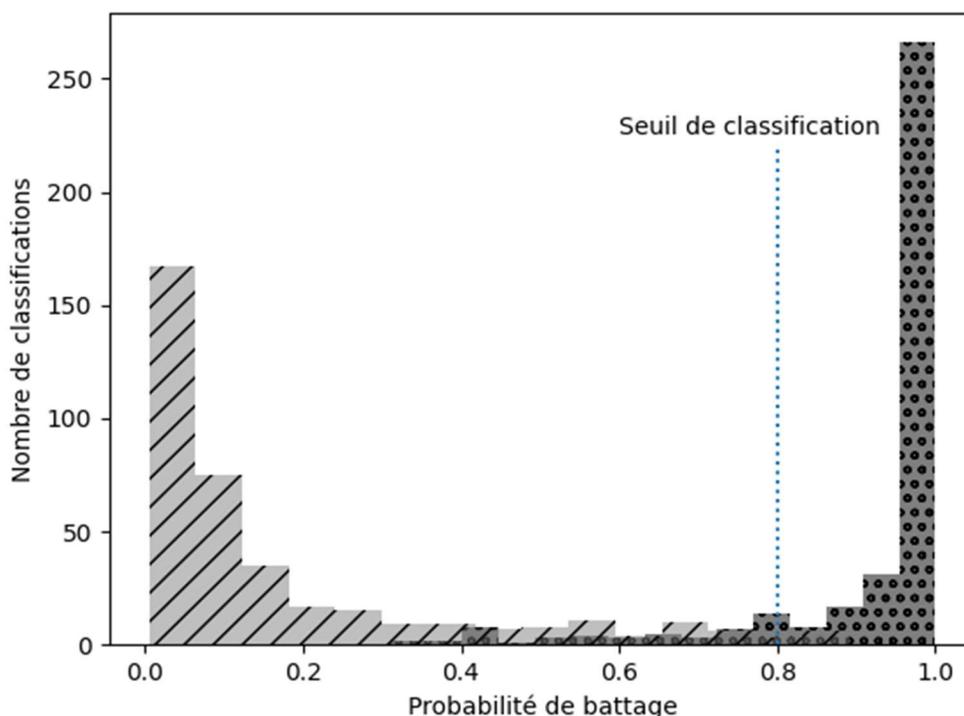


Figure 6. Histogramme de la probabilité de battage, avec un seuil de classification arbitrairement choisi à 0.8.

Malgré le léger chevauchement de ces 2 catégories, on peut définir un seuil de probabilité tel que tout résultat au-dessus de ce seuil appartient à la catégorie «battage» et tout résultat au-dessous appartient à la catégorie «non-battage». Par exemple, le choix d'un seuil défini à 0.8 favorisera la classification d'événements en «non-battage» (Figure 6), autrement dit, avec ce choix de paramètres, il sera plus probable d'avoir un événement battage classifié en non-battage qu'un événement non-battage classifié en battage. Les résultats du modèle avec ce seuil sont présentés dans le tableau 2.

Tableau 2 : Résultats de la classification binaire de jeu de données d'évaluation, avec un seuil de probabilité de 0.8

Observation	Prédiction	
	Battage	Non-battage
Battage	326	48
Non-battage	7	375
Taux de vrais battages	97,9 %	
Taux de vrais non-battages		88,7 %

Le taux de « vrais battages » (vrais positifs) est défini par les prédictions correctes de battage divisés par le nombre total d'événements de battage (labellisés par identification experte). Parallèlement, le taux de « vrais non-battages » (vrais négatifs) est défini par les prédictions correctes de non-battage divisés par le nombre total d'événements labellisés comme non-battage. Le taux de « vrais battages », plus élevé que le taux de vrais non-battage (97,9% vs 88,7%), dépend du choix de ce seuil. Dans l'exemple présenté plus haut nous souhaitons minimiser la probabilité d'une classification d'un événement non-battage en battage (minimisation des fausses alertes). Quoi qu'il en soit, les pourcentages de réussite obtenus sont très encourageants, et nous envisageons d'appliquer la méthode à d'autres signatures vibratoires.

## 5. Conclusions et perspectives

La démarche proposée précédemment apparaît prometteuse, avec des résultats très positifs, qui permettent d'identifier de manière automatisée les vibrations émises par la technique de battage avec un taux de réussite important (97.8 %). On retiendra la possibilité d'influer légèrement sur le taux de réussite (vrai positifs ou négatifs) à travers le choix du seuil de probabilité. Cette technique constitue une aide précieuse aux entreprises ou maîtres d'œuvre pour aller vers plus d'autonomie en termes de suivi des vibrations.

En termes de perspectives, cette méthode est adaptable à une classification de plusieurs catégories, par exemple : battage, vibrofonçage, autres... Ce travail est en cours, ainsi que l'extension de la base de données pour inclure des données d'autres sites et l'optimisation des hyperparamètres du modèle. La constitution d'une base de données représentative d'une technique de travaux particulière, transposable à d'autres sites/chantiers car peu, voire pas dépendant d'un site particulier, constitue une perspective intéressante car elle vise à écourter, voire idéalement à supprimer, la phase d'apprentissage qui peut nécessiter en phase initiale un nombre d'évènements importants.

**Références bibliographiques**

Fawaz H.I., et al. (2019). Deep learning for time series classification: a review. Data mining and knowledge discovery, Vol. 33.4 (2019): 917-963.

Ruiz A.P. et al. (2021). The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances. Data mining and knowledge discovery, Vol. 35.2 : 401-449. doi:10.1007/s10618-020-00727-3.