

# ORIENTER LA RECHERCHE DE FUITES PAR L'ESTIMATION DE LA PROBABILITE D'ETAT FUYARD DES CONDUITES

Marie Collet<sup>1</sup>, Vincent Couallier<sup>2</sup>, Nicolas Rodriguez<sup>1</sup>, Alain Husson<sup>1</sup>, Yves Le Gat<sup>1</sup>, Eddy Renaud<sup>1</sup>

## MOTS-CLES

Gestion patrimoniale des réseaux d'eau, recherche de fuite, modèle statistique zéro-augmenté, données manquantes.

## CHAPÔ

Pour contribuer à l'optimisation de la lutte contre les fuites du réseau d'eau potable de la SWDE\*, INRAE† a développé un nouveau modèle statistique permettant d'estimer la probabilité d'état fuyard des conduites en s'appuyant sur l'historiques des campagnes de recherche de fuites.

## INTRODUCTION

La maîtrise des fuites des réseaux d'eau potable est un enjeu fort des collectivités locales. Pour optimiser les actions mises en œuvre, une prédiction de l'état fuyard des tronçons composant le réseau est recherchée. Les données disponibles pour obtenir de telles prédictions sont celles issues de campagnes de recherche de fuite (RF) ciblées, stoppées par la SWDE lorsqu'une fuite est trouvée. De ce fait, l'état fuyard détecté d'un tronçon (considéré comme l'individu statistique) lors d'une campagne de RF peut être modélisé comme une variable binaire Z, valant 1 quand le tronçon a été trouvé fuyard et 0 sinon. Pour autant, la valeur 0 couvre deux phénomènes bien distincts : le tronçon a été inspecté et il n'était pas fuyard ou le tronçon n'a pas été inspecté alors qu'il faisait partie du secteur visé par la campagne (Figure 1).

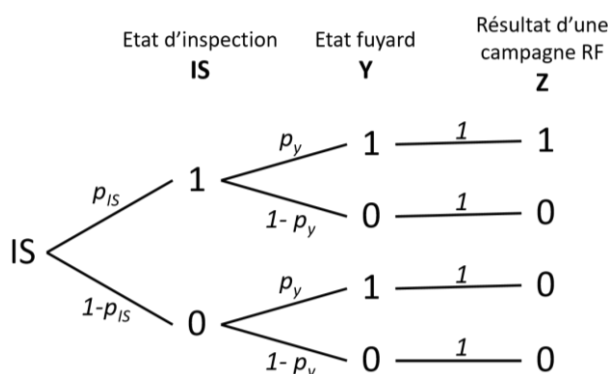


Figure 1. Schéma du modèle de données. Z est la seule variable observée. Inspiré de Moriña et al., 2021.

## MATERIEL ET METHODES

Afin de distinguer les deux cas d'absence de fuite détectée ( $Z=0$ ), nous avons développé un modèle basé sur le principe des modèles zéro-augmenté (Blasco-Moreno et al., 2019; Mullahy, 1986) en

<sup>1</sup> INRAE, ETTIS, F-33612 Cestas, France.

<sup>2</sup> IMB, Université de Bordeaux, F-33405 Talence, France.

\* Société Wallonne Des Eaux

† Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

considérant notre variable aléatoire d'observation  $Z$  comme étant le produit de deux variables aléatoires latentes non systématiquement observées,  $Y$  étant l'état réel de l'individu (état fuyard du tronçon) et  $IS$  étant l'indicateur d'observation de l'individu lors de la campagne de RF (Figure 1). Pour simplifier, on supposera que les variables binaires de fuite  $Y$  et d'inspection  $IS$  satisfont chacune un modèle de régression logistique. D'où, sous réserve d'indépendance de  $Y$  et  $IS$ , nous avons  $P(Z = 1|W_i = w_i, X_i = x_i) = \frac{1}{1+e^{-\theta'w_i}} \times \frac{1}{1+e^{-\beta'x_i}}$ , avec  $W_i$  et  $X_i$  les vecteurs de covariables des tronçons respectivement pour l'état d'inspection et l'état fuyard, et  $\theta'w_i$  la combinaison linéaire  $\theta_0 + \theta_1 w_1 + \dots + \theta_p w_p$ , avec  $w = (1, w_1, \dots, w_p)$  (respectivement  $\beta'x_i$ ).

Nous avons travaillé sur 185000 conduites situées dans les 1884 secteurs hydrauliques de la SWDE concernés par de la RF depuis 2018. Les données comprennent les caractéristiques des tronçons (date de pose, matériau, longueur, diamètre...), ainsi que l'historique de la RF et des réparations sur 4 ans (2018-2021). 4766 fuites ont été trouvées lors de 17834 campagnes de RF.

## RESULTATS ET DISCUSSION

La probabilité moyenne d'inspection des tronçons lors d'une campagne est estimée à 3.8% et celle d'état fuyard à 8.4% (Figure 2). 650 conduites qui représentent 1.2% du linéaire total ont un risque d'état fuyard supérieur à 90%. Le modèle estime qu'en moyenne les 100 conduites les plus à risque de fuir sont anciennes (60 ans), longues (400m), avec beaucoup de branchements (12) et ont déjà cassées de nombreuses fois (9). Ces résultats, cohérents avec les remontées du terrain, soulignent l'intérêt d'utiliser ce nouveau modèle qui corrige le biais d'une régression logistique classique qui considérerait une probabilité d'inspection des conduites de 100%. Il devrait donc permettre d'améliorer la prise de décision pour cibler la RF et remplacer les conduites.

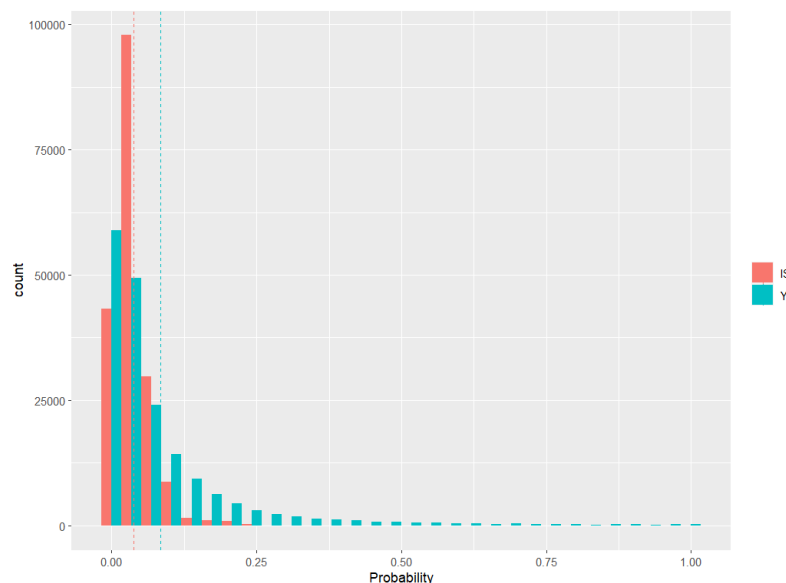


Figure 2. Distribution des probabilités d'état fuyard ( $Y$ , en vert) et d'inspection ( $IS$ , en orange) des tronçons du réseau d'eau potable de la SWDE. Les moyennes sont représentées par les traits verticaux pointillés.

## BIBLIOGRAPHIE

- Blasco-Moreno, A., Pérez-Casany, M., Puig, P., Morante, M., Castells, E., 2019. What does a zero mean? Understanding false, random and structural zeros in ecology. *Methods Ecol. Evol.* 10, 949–959. <https://doi.org/10.1111/2041-210X.13185>
- Moriña, D., Puig, P., Navarro, A., 2021. Analysis of zero inflated dichotomous variables from a Bayesian perspective: application to occupational health. *BMC Med. Res. Methodol.* 21, 277. <https://doi.org/10.1186/s12874-021-01427-2>
- Mullahy, J., 1986. Specification and testing of some modified count data models. *J. Econom.* 33, 341–365. [https://doi.org/10.1016/0304-4076\(86\)90002-3](https://doi.org/10.1016/0304-4076(86)90002-3)