

gapIT : Un outil visuel pour l'imputation de valeurs manquantes en hydrologie

Olivier Parisot, Laura Giustarini, Olivier Faber, Renaud Hostache, Ivonne Trebs,
and Mohammad Ghoniem

Centre de Recherche Public – Gabriel Lippmann, Belvaux, Luxembourg
parisot@lippmann.lu

Résumé. Les données manquantes sont problématiques en hydrologie, car elles gênent le calcul de statistiques interannuelles et sur de longues périodes, ainsi que l'analyse et l'interprétation de la variabilité des données. Dans cet article, nous présentons gapIT, une plateforme d'analyse de données permettant d'inspecter visuellement les données manquantes et ensuite de choisir la méthode de correction adéquate. Nous avons utilisé l'outil pour estimer les données manquantes dans des séries temporelles correspondant aux débits mesurés par des stations hydrométriques du Luxembourg.

1 Introduction

Traditionnellement, les données hydrométriques se présentent sous la forme de séries temporelles, représentant des mesures effectuées régulièrement par des stations : ces mesures peuvent concerner différents aspects comme les hauteurs et les débits de l'eau dans les cours d'eau, les quantités de précipitations, etc. Comme ces mesures sont souvent prélevées par un réseau distribué de capteurs, le problème des données manquantes est inévitable. Allant d'une simple valeur manquante à une longue plage de valeurs manquantes, les lacunes peuvent avoir des causes multiples : dysfonctionnement des capteurs, maintenance des stations de mesure, erreurs humaines, etc. (Harvey et al. (2010)).

Le réseau hydrométrique au Luxembourg fournit un bon cas d'utilisation. Il est constitué de différentes stations hydrométriques permettant de mesurer notamment les débits des cours d'eau. Les mesures sont ensuite fréquemment utilisées dans les modèles numériques de prévision hydrologique ou pour calculer des statistiques sur les écoulements (e.g. temps de retour des crues ou des sécheresses).

En conséquence, lorsque certaines séries de mesures présentent beaucoup de lacunes (par exemple : les données de la station de Hallerbach au Luxembourg, Figure 1), cela pose de nombreux problèmes et il est nécessaire d'apporter un soin très particulier à combler ces lacunes avec une bonne précision.

Afin de combler ces lacunes, les méthodes classiques d'analyse de données ont été appliquées dans le domaine hydrologique (Salas (1980)), et des travaux récents tentent de fournir des solutions toujours plus efficaces (Harvey et al. (2010); Mwale et al. (2012)). Or, il est souvent difficile de choisir la bonne méthode de calcul parmi toutes celles qui existent car