

Étude de comblement de lacunes : Cas des séries pluviométriques observées du réseau de l'ONM

KERTALI Fouzia 1 ^{*a}

^a 1 Office national de la météorologie, Dar El Beida, Alger

Correspondant: fouzia.khadidja@yahoo.com

Résumé

Les données manquantes sont fréquentes dans les séries météorologiques et peuvent entraîner d'importantes erreurs dans l'estimation statistique des paramètres. Ce travail expose les différentes méthodes théoriques de traitement des séries lacunaires. La méthode de double masse est choisie dans cette étude pour rattraper des données manquantes d'une série de données pluviométriques. L'efficacité de cette méthode est mise en exergue après un second test appliqué sur un autre jeu de données.

Keywords comblement de lacune, méthode de double masse, donnée pluviométrique manquante, homogénéité.

1. Introduction

La climatologie est la science du climat. Mais son domaine d'application n'est pas restreint à ce dernier. Il s'agit d'une discipline beaucoup plus vaste. Elle emprunte à d'autres sciences des notions ou des résultats dont elle a besoin en faisant appel à des moyens techniques de plus en plus sophistiqués. On peut citer quelques unes : toutes les sciences concernant l'atmosphère comme la physique, la biologie, l'agronomie, l'hydrologie, l'économie, l'informatique et surtout les statistiques pour le traitement et l'utilisation rationnelle des données. [1] Pourquoi corriger la série climatique? Parmi les questions qui agitent les scientifiques à propos du changement climatique revient continuellement celle de la comparaison de la variabilité climatique actuelle avec celle du passé. A l'échelle du dernier millénaire, les mesures directes ne sont pas disponibles, il faut donc les reconstruire. Cependant, depuis le XIX^{ème} siècle, nous disposons de nombreuses longues séries d'observation instrumentales. Mais leur qualité doit être étudiée au préalable.

Par définition, une série d'observation météorologique[2] est dite homogène lorsque les conditions de mesure n'ont pas varié au cours du temps. L'homogénéité ou plutôt l'absence d'homogénéité des longues séries instrumentales en climatologie est un problème connu depuis longtemps, le déplacement des postes climatologique au cours du temps, la modification des sites de mesure, de l'instrumentation, des méthodes de calcul des paramètres météorologiques[3] et les changements d'observateurs, parmi tant d'autres, vont se traduire par autant de sauts dans les séries de données. Or, ces ruptures artificielles peuvent être du même ordre

de grandeur que les phénomènes qu'on cherche à mettre en évidence dans les séries climatiques, comme les tendances, les cycles, etc. Leur correction est donc indispensable avant toute étude climatique sérieuse.

2. Méthode de comblement des données manquantes

En statistique, on parle de valeur manquante lorsqu'on ne dispose pas d'observation pour une variable donnée ou pour un individu donné. Le problème de la gestion des données manquantes est un vaste sujet. Les données manquantes ne peuvent pas être ignorées lors d'une analyse statistique. Selon leur proportion et leur type, des solutions différentes vont être choisies. On pourra soit tirer les variables ou les individus présentant des données manquantes ou imputer des valeurs aux données manquantes ou encore développer des méthodes (ou algorithmes) qui permettent de mener les analyses en présence de données manquantes.

En statistique, on définit trois types de données manquantes :

- Données manquantes complètement aléatoires (MCAR)
- Données manquantes aléatoires (MAR)
- Données manquantes non aléatoires (NMAR)

Les modèles mathématiques les plus connues pour combler les données manquantes sont :