

Vers une méthode de caractérisation et de quantification des incertitudes dans le cadre d'une fusion de données hétérogènes multicapteurs dans le domaine de la pollution atmosphérique

Aymeric Ambert*, Mickael Germain*, Yacine Bouroubi*

* Département de géomatique appliquée, Faculté des lettres et sciences humaines
2500, boulevard de l'Université, Sherbrooke (Québec) J1K 2R1
aymeric.ambert@usherbrooke.ca

Résumé. La lutte contre la pollution atmosphérique est un enjeu majeur du 21^e siècle. La gestion des données massives liées à la diversification des supports de mesure est un défi et engendre des problématiques inédites en termes de volume de qualification et de traitement de l'information. Les méthodes de fusion de données sont autant de solutions au problème posé par l'utilisation de données massives. La prise en compte de chaque capteur en tant qu'élément d'information dans le cadre de la fusion entraîne néanmoins un risque quant à l'incertitude globale des données à prendre en considération. Le présent article vise à établir une approche pour réduire l'incertitude relative pour chaque source de données en utilisant la fusion évidentielle. Se basant sur un modèle attributaire de données existantes, l'article propose de définir des indicateurs de performances permettant de valider ou non un tel modèle.

1 Introduction et état de l'art

Le 22 septembre 2021, l'OMS (Organisation Mondiale de la Santé) a rendu public son nouveau guide concernant la pollution atmosphérique considérant en préambule que "la pollution atmosphérique est l'une des principales menaces environnementales pour la santé" (OMS, 2021). Lutter contre la pollution atmosphérique est donc devenu un enjeu majeur de santé publique. L'utilisation d'un capteur de pollution correspond à un besoin spécifique et doit prendre en compte la précision, l'incertitude et la capacité d'adaptation à de potentiels nouveaux besoins de celui-ci (ex : Piedrahita et al. (2014)). Pour y parer, la recherche et l'industrie, via l'essor de nouvelles technologies comme l'intelligence artificielle, s'orientent de plus en plus vers l'utilisation big data, c'est-à-dire la massification des données, notamment pour pallier l'incomplétude et l'incertitude de données mesurées (Hariri et al., 2019). Dans le contexte de la pollution atmosphérique et du présent article, il s'agit justement de fusionner des données issues de capteurs de différents types. Il existe déjà de nombreux travaux de recherche liés à la fusion de données (Khaleghi et al., 2013). Néanmoins la notion d'incertitude revêt un caractère tout à fait singulier, car sa prise en compte pour chaque capteur impacte l'information globale