

# Resumé

Les sciences de l'environnement et de la terre traitent souvent des phénomènes assez complexes. Les données environnementales sont généralement le résultat d'une interaction complexe entre plusieurs variables, qui dépendent à la fois des coordonnées spatiales et temporelles. Afin de mieux comprendre ces processus multivariés et de pouvoir construire des modèles de prédictions, il est essentiel de se limiter à un sous-ensemble des variables de taille optimale. Des différentes techniques de l'apprentissage automatique sont actuellement appliquées pour remédier au problème du fléau de la dimension. La vitesse du vent, déterminée par le gradient de pression et les forces de Coriolis, est l'un des phénomènes multivariés les plus complexes à l'échelle locale, où la topographie ainsi que d'autres variables peuvent modifier considérablement les modèles qui gouvernent ce phénomène.

Dans cette étude, un sous-ensemble de variables explicatives sera identifié permettant d'expliquer et de prédire au mieux les vitesses du vent en Suisse. Les performances des différentes méthodes sont évaluées afin de comprendre les difficultés liées à la variabilité spatio-temporelle propre au problème de prédiction de la vitesse du vent. Une analyse exhaustive portant sur 13 variables est réalisée afin de sélectionner un espace d'entrée de dimension minimale et optimale pour une meilleure prédiction. La sélection des variables est effectuée à l'aide de méthodes intrinsèques, intégrées et de filtrage. En particulier, baser sur le paramètre de lissage pour l'estimation du noyau de la densité de probabilité, la méthode de régression générale réseau de neurones est appliquée pour tester un nombre total de 730 vecteurs de sortie.

Les sous-ensembles de variables pertinents sont identifiés pour chaque vecteur de sortie à l'aide du réseaux de neurones à régression générale isotropique et anisotropique, le choix de ces variables est évalué par les erreurs quadratiques moyennes qui décrivent le compromis entre la variance et le biais. finalement, La durée de vie de chaque prédicteur est estimée comme étant la période (demi-vie) qui caractérise la persistance en terme d'importance des prédicteurs pertinents.