

APPRENTISSAGE STATISTIQUE SUR MODÈLES MÉTÉOROLOGIQUES POUR L'ÉOLIEN

Aurélie Fischer

*LPSM, Université Paris Diderot
8 Place Aurélie Nemours
75013 PARIS
aurelie.fischer@univ-paris-diderot.fr*

Résumé. Cet exposé est consacré au problème dit de “réduction d’échelle” d’une quantité météorologique à partir des sorties de modèles numériques de prévisions météorologiques. Il est essentiel dans le domaine de l’énergie éolienne de disposer de prévisions précises de la vitesse du vent aux emplacements des parcs éoliens. Ces prévisions sont aussi cruciales pour la prévention des dommages liés à des vents violents. Nous étudions les performances de méthodes d’apprentissage statistique pour la reconstruction et la prévision de la vitesse du vent à partir des sorties du modèle du Centre européen de prévisions météorologiques à moyen terme (ECMWF), en utilisant tout d’abord les données de vitesse du vent locales fournies par la plateforme d’observation du Site Instrumental de Recherche par Télédétection Atmosphérique (SIRTA), puis des données observées à l’échelle de la France.

Mots-clés. Modélisation de la vitesse du vent, réduction d’échelle, prévisions numériques météorologiques, méthode d’apprentissage statistique.

Abstract. This talk addresses the problem of downscaling a meteorological quantity from outputs of Numerical Weather Prediction models. For wind energy application, it is essential to have at hand accurate forecasts of wind speed at wind farms locations. These forecasts are also crucial to prevent severe wind damages. We study the performances of statistical learning methods for reconstructing and forecasting wind speed from the European Center of Medium-range Weather Forecasts (ECMWF) model outputs, using first local wind speed data from SIRTA observation platform, and then data observed all over France.

Keywords. Wind speed modeling, downscaling, numerical weather prediction, statistical learning methods.

1 Introduction

Dans cet exposé, nous présenterons un travail appliqué sur le thème des énergies renouvelables, issu d’une collaboration ayant débuté pendant le projet ANR FOREWER

Forecasting and Risk Evaluation of Wind Energy Production, entre des mathématiciens membres du Laboratoire de Probabilités, Statistique et Modélisation (LPSM) et des physiciens membres du Laboratoire de Météorologie dynamique (LMD).

Les énergies renouvelables sont intermittentes, ce qui rend difficile la prévision de production d'énergie. Or, comme l'électricité peut difficilement être stockée, il est important d'être en mesure d'équilibrer production et consommation d'électricité. Dans le contexte de l'énergie éolienne, dans le but de faire de la prévision de production d'énergie électrique, il est nécessaire de disposer de prévisions locales précises de la vitesse du vent. Plus généralement, les prévisions de vent jouent aussi un rôle important pour la prévention des dommages et accidents liés aux intempéries.

Dans cette perspective, notre objectif ici consiste à modéliser le lien entre des mesures réelles de vitesse du vent et des données de modèles météorologiques numériques.

2 Etude sur le site du SIRTA

Notre premier travail dans cette direction (Alonzo et al., 2018) est basé sur les données de vitesse du vent observées sur le site instrumental de recherche par télédétection atmosphérique (SIRTA), situé sur le campus de l'Ecole Polytechnique. Il s'agit de mesures de vent à 10m obtenues grâce à un anémomètre, et de données de vent à 100m obtenues à l'aide d'une télédétection par laser (ou LIDAR = Light Detection And Ranging), avec un pas de temps de 10 minutes, pour la période 2011-2015.

A partir de ces observations, nous cherchons à modéliser la vitesse du vent en fonction de données météorologiques provenant de modèles numériques : nous avons utilisé les données du Centre Européen pour les Prévisions Météorologiques à Moyen terme (ECMWF). Environ 200 variables sont disponibles : altitude géopotentielle, vorticit , plusieurs données de température, cisaillement du vent...

Nous nous attachons à reconstruire la vitesse du vent à 10m et à 100m. Notons que les prévisions de vitesse du vent du modèle ECMWF sont meilleures à 100m qu'à 10m, où l'on observe invariablement un biais. Nous comparons les résultats de modélisation obtenus en utilisant plusieurs stratégies basées sur une régression linéaire ou une forêt aléatoire. Plus précisément, on considère 7 procédures de régression linéaire :

- avec le vent à 10m comme unique variable explicative,
- avec toutes les variables ECMWF disponibles,
- avec sélection de variables de type pas à pas,
- avec une pénalisation Lasso,
- avec uniquement le vent à 10m, mais en construisant 8 modèles selon la direction du vent
- avec toutes les variables, 8 modèles selon la direction,
- avec sélection de variables, 8 modèles selon la direction.

Concernant les forêts aléatoires, on utilise toutes les variables disponibles, et l'on applique

la méthode classique ainsi que celle consistant à construire 8 modèles selon la direction du vent.

3 Etude à l'échelle de la France

Afin de disposer d'une étude systématique à l'échelle de la France, nous avons ensuite étendu l'étude effectuée sur le site du SIRTA à un ensemble de sites répartis sur le territoire français, en utilisant des observations de la base de données Integrated Surface Database (ISD), qui rassemble des données météorologiques de 400 stations françaises environ (Goutham et al., 2019). Ici, les données observées concernent 171 stations, pour lesquelles on dispose de mesures horaires du vent à 10m pour la période 2010-2017.

Nous examinons les performances de plusieurs méthodes d'apprentissage : régression linéaire, k plus proches voisins, forêts aléatoires, boosting.

D'un point de vue physique, on peut s'attendre à ce que la qualité du modèle ECMWF dépende de différents facteurs géographiques. Il est particulièrement intéressant de quantifier l'apport des méthodes statistiques dans les zones où le modèle numérique est le moins précis. La Figure 1 présente sur une carte de la France un aperçu des résultats dans le cas d'une forêt aléatoire. Nous constatons notamment que l'apport des méthodes statistiques par rapport au modèle ECMWF est particulièrement intéressant pour les stations situées près du littoral.

4 Prévisions à court terme

A la suite du travail de modélisation, nous avons également obtenu des résultats préliminaires très encourageants concernant la prévision à court terme de la vitesse du vent. Il ressort de l'étude qu'adjointre aux prévisions des modèles météorologiques numériques l'information provenant de vraies données observées dans le passé peut s'avérer très utile.

Bibliographie

Alonzo, B. and Plougonven, R. and Mougeot, M. and Fischer, A. and Dupré, A. and Drobinski, P. (2018), From Numerical Weather Prediction outputs to accurate local surface wind speed : statistical modeling and forecasts. In *Renewable Energy : Forecasting and Risk Management*, Springer Proceedings in Mathematics & Statistics.

Goutham, N. and Alonzo, B. and Dupré, A. and Plougonven, R. and Doctors, R. and Liao, L. and Mougeot, M. and Fischer, A. and Drobinski, P. (2019), Using machine learning methods to improve surface wind from the outputs of a Numerical Weather Prediction model.

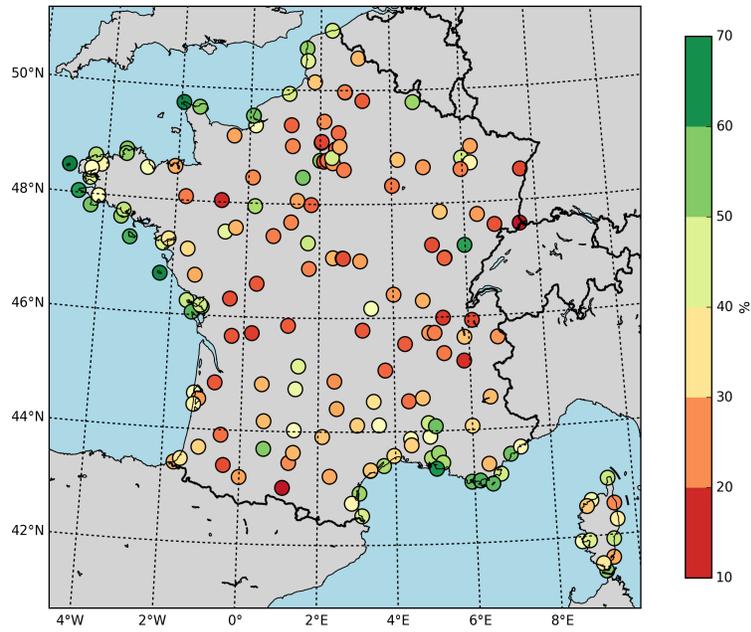


FIGURE 1 – Variation relative de l'erreur d'un modèle d'apprentissage statistique basé sur une forêt aléatoire (racine carrée de l'erreur quadratique moyenne) par rapport à l'erreur du modèle numérique ECMWF.