



UNIVERSITE DE CORSE-PASCAL PAOLI
ECOLE DOCTORALE ENVIRONNEMENT ET SOCIETE
UMR CNRS 6134
Sciences Pour l'Environnement



Thèse présentée pour l'obtention du grade de
DOCTEUR EN MECANIQUE DES FLUIDES, ENERGETIQUE,
THERMIQUE, COMBUSTION, ACOUSTIQUE

Mention : Energétique, Génie des procédés

Soutenue publiquement par
Alexis FOUILLOY

Le 24 Septembre 2019

Comparaison de méthodes d'apprentissage automatique
de prévision de la ressource solaire pour une application à
une gestion optimisée des réseaux intelligents

Directeurs

M. Gilles NOTTON, Dr-HDR, Université de Corse
M. Cyril VOYANT, Dr-HDR, Université de Corse

Rapporteurs

M. Philippe LAURET, PR, Université de La Réunion
M. Fawaz MASSOUH, Dr-HDR, Arts et Métiers ParisTech

Jury :

M. Cédric JOIN, PR, Université de Lorraine
M. Jean-François MUZY, DR-CNRS, Université de Corse
Mme Marie-Laure NIVET, Dr, Université de Corse
M. Gilles NOTTON, Dr-HDR, Université de Corse
M. Giuseppe Marco TINA, PR, Université de Catane
M. Cyril VOYANT, Dr-HDR, Université de Corse
M. Philippe LAURET, PR, Université de La Réunion
M. Fawaz MASSOUH, Dr-HDR, Arts et Métiers ParisTech

Résumé

Les enjeux relatifs à la production énergétique future, notamment en termes d'utilisation de ressources locales et « propres », conduisent les producteurs d'électricité à se tourner de plus en plus vers les sources renouvelables d'énergie et particulièrement les sources intermittentes que sont le vent et le soleil. Le problème est que leur caractère intermittent et aléatoire oblige les gestionnaires de réseau à limiter leur intégration au mix énergétique. Il est alors nécessaire de coupler différents systèmes de production pour garantir la stabilité du réseau et la sécurité des moyens de production. Afin de faciliter ces opérations de gestion et d'optimiser l'intégration des énergies renouvelables intermittentes, le solaire dans notre cas, il est nécessaire de s'intéresser à la prévision de la ressource. Dans le but de connaître à l'avance l'énergie disponible et de permettre une gestion optimale du couplage entre systèmes de production conventionnels et intermittents. Au cours de cette étude, nous avons développé et étudié un large panel de modèles de prévision du rayonnement solaire (persistance, persistance intelligente, filtre de Kalman, ARMA, réseau de neurones, processus Gaussien, machine à vecteurs de support, arbres de régressions simples, élagués, renforcés, ensachés et forêts aléatoires), pour des horizons de prévision utiles aux gestionnaires de réseaux, et appliqués à des données en provenance de différents sites. Ces travaux ont été réalisés dans le cadre d'un projet de recherche Horizon 2020, le projet TILOS pour « Technology Innovation for the Local Scale Optimum Integration of Battery Energy Storage » qui consiste en l'installation d'une centrale hybride solaire, éolienne et stockage par batteries NaNiCl_2 sur une petite île de l'archipel du Dodécanèse. Les horizons de prévision testés sont de 1 à 6 heures par pas de temps horaire (6h/1h) pour les 4 sites de mesures pour la prévision du rayonnement global horizontal. Les sites sont répartis en Europe dans des zones géographiques qui possèdent des climats différents : Ajaccio (Corse, France), Tilos (Dodécanèse, Grèce), Nancy (Grand Est, France) et Odeillo (Languedoc Roussillon, France). Nous avons caractérisé chaque site par variabilité des données, on entend par là leur tendance à varier fortement ou non avec le temps. Les principaux résultats de ces travaux sont que les prévisions sur des données sur des sites à faible variabilité peuvent être réalisés par des modèles simples. Plus la variabilité est élevée, plus la prévision est difficile à réaliser et des modèles plus complexes doivent être utilisés (basés sur l'apprentissage automatique et l'apprentissage d'ensemble) pour obtenir de meilleurs résultats. Nous avons par ailleurs utilisé une prévision probabiliste pour donner une plage de confiance de la prévision au gestionnaire de réseau. Etant donné que nous disposons de mesures des composantes directe normale et diffuse horizontale pour un des quatre sites, nous avons confronté nos modèles à ces prévisions. Il apparaît que le rayonnement direct normal est difficile à prévoir, notamment à cause de sa forte variabilité, et que les forêts aléatoires sont les plus probants. Enfin des modèles ont été développés pour être insérés dans le système automatique de gestion de l'énergie, appliqués au rayonnement global incliné, avec un horizon de 10 minutes par pas de temps de 1 minute et un horizon de 2 heures avec des pas de temps de 10 et 15 minutes. Il apparaît que les modèles d'apprentissage automatique donnent tous sensiblement de bons résultats et que le choix de l'un ou l'autre sera plutôt réalisé en fonction des contraintes techniques et pratiques des outils.

Mots clés : prévision, rayonnement global, intelligence artificielle, séries temporelles, apprentissage supervisé, stationnarité.

Abstract

The stakes relating to future energy production, particularly in terms of the use of local and "clean" resources, are leading electricity producers to turn more and more towards renewable sources of energy and particularly intermittent sources: the wind and the sun. The problem is that their intermittent and random nature forces network operators to limit their integration into the energy mix. It is then necessary to couple different production systems to guarantee the stability of the network and the security of the means of production. In order to facilitate these management operations and to optimize the integration of intermittent renewable energies, solar energy in our case, it is necessary to focus on the forecast of the resource. In order to know in advance, the available energy and to allow an optimal management of the coupling between conventional and intermittent production systems. In this study, we have developed and studied a wide range of solar radiation prediction models (persistence, smart persistence, Kalman filter, ARMA, neural network, Gaussian process, support vector machine, simple regression trees, pruned, boosted, bagged and random forests), for forecast horizons useful to network managers, and applied to data from different sites. This work was carried out as part of a Horizon 2020 research project, the TILOS project for "Technology Innovation for the Local Scale Optimum Integration of Battery Energy Storage" which consists of the installation of a solar, wind and solar hybrid power station. NaNiCl₂ battery storage on a small island in the Dodecanese archipelago. The forecast horizons tested are from 1 to 6 hours per hour time step (6h / 1h) for the 4 measurement sites for horizontal global radiation prediction. The sites are spread across Europe in geographical areas with different climates: Ajaccio (Corsica, France), Tilos (Dodecanese, Greece), Nancy (Grand Est, France) and Odeillo (Languedoc Roussillon, France). We have characterized each site by data variability, which means their tendency to vary strongly or not with time. The main results of this work are that the forecasts on data on sites with low variability can be realized by simple models. The higher the variability, the more predictive and difficult to achieve, and more complex models must be used (based on machine learning and overall learning) for better results. We also used a probabilistic forecast to give a confidence range of the forecast to the network manager. Since we have measurements of the normal and diffuse horizontal direct components for one of the four sites, we compared our models to these predictions. It appears that normal direct radiation is difficult to predict, in particular because of its high variability, and random forests are the most convincing. Finally, models have been developed to be inserted in the automatic energy management system, applied to inclined global radiation, with a horizon of 10 minutes per time step of 1 minute and a horizon of 2 hours with no time steps 10 and 15 minutes. It appears that the machine learning models all give significantly good results and that the choice of one or the other will rather be made according to the technical and practical constraints of the tools.

Key words: forecasting, global radiation, artificial intelligence, time series, supervised learning, stationarity.