INRS - Centre Eau Terre Environnement

ENSEMBLE OF ARTIFICIAL NEURAL NETWORKS FOR SEASONAL FORECASTING OF WIND SPEED IN EASTERN CANADA

Par

Pia Milena Leminski

Mémoire présenté pour l'obtention du grade de
Maître es Sciences, M.Sc.
en sciences de la Terre
(bidipl. accueil INRS-Univ. de Reykjavik, énergie renouvelable)

Jury d'évaluation

Président du jury et examinateur interne	André St-Hilaire INRS ETE
Examinateur externe	Maryam R. Al Shehhi Khalifa University
Directeur de recherche	Taha B.M.J. Ouarda INRS ETE
Codirecteur de recherche	Ármann Gylfason Reykjavik University

© Droits réservés de Pia Milena Leminski, April 2024

RÉSUMÉ

L'exploitation efficace des ressources en énergie éolienne, y compris les progrès dans les méthodes de prévision saisonnière et d'analyse des tendances, est impérative pour le progrès durable de la production d'énergie éolienne. Alors que les données relatives aux températures et aux précipitations font l'objet d'une attention considérable dans l'analyse des tendances, il existe une lacune notable dans l'exploration des corrélations entre les indices climatiques et la vitesse du vent. Cet article propose l'utilisation d'un ensemble de réseaux neuronaux artificiels pour prévoir les vitesses du vent sur la base des indices d'oscillation du climat et évalue ses performances. Un premier examen indique un signal de corrélation entre les indices climatiques et les vitesses du vent d'ERA5 pour l'étude de cas sélectionnée dans l'est du Canada. Les prévisions sont faites pour la saison avril-mai-juin et sont basées sur les indices climatiques les plus corrélés des saisons précédentes. Une prédiction ponctuelle est effectuée avec un ensemble de 20 membres, qui est vérifié par validation croisée de type "leave-on-out". Les résultats obtenus sont analysés en termes d'erreur quadratique moyenne, de biais et de score de compétence et montrent une performance compétitive avec les prévisions numériques de vent de pointe de SEAS5, les surpassant dans plusieurs régions. Un modèle relativement basique avec une seule unité dans la couche cachée et un taux de régularisation de 10^{-2} fournit des résultats prometteurs, en particulieren particulier dans les régions où le nombre d'indices pris en compte est plus élevé. Cette étude s'inscrit dans le cadre des efforts déployés au niveau mondial pour améliorer la précision des prévisions en introduisant une nouvelle approche.

Mots-clés Vent, EANN, prévisions saisonnières, ERA5

ABSTRACT

Efficient harnessing of wind energy resources, including advancements in weather and seasonal forecasting and climate projections, is imperative for the sustainable progress of wind power generation. While temperature and precipitation data receive considerable attention in interannual variability and seasonal forecasting studies, there is a notable gap in exploring correlations between climate indices and wind speeds. This paper proposes the use of an ensemble of artificial neural networks to forecast wind speeds based on climate oscillation indices and assesses its performance. An initial examination indicates a correlation signal between climate indices and wind speeds of ERA5 for the selected case study in eastern Canada. Forecasts are made for the season April-May-June and based on most correlated climate indices of preceding seasons. A pointwise forecast is conducted with a 20 member ensemble, which is verified by leave-on-out cross-validation. The results obtained are analyzed in terms of root mean squared error, bias, and skill score and show a competitive performance with state-of-the-art numerical wind predictions from SEAS5, outperforming them in several regions. A relatively basic model with a single unit in the hidden layer and regularization rate of 10^{-2} provides promising results, especially in areas with a higher number of indices considered. This study adds to global efforts to enable more accurate forecasting by introducing a novel approach.

Keywords Wind, EANN, Seasonal forecasting, ERA5

SYNOPSIS

Introduction

Il est impératif d'utiliser efficacement les ressources renouvelables pour garantir des progrès durables dans la production d'énergie et atténuer les effets du changement climatique. Parmi les sources d'énergie renouvelables, l'énergie éolienne présente un potentiel important et devient une pierre angulaire dans les efforts mondiaux de transition vers des alternatives énergétiques plus propres. Dans ce contexte, la gestion et l'optimisation de l'infrastructure éolienne jouent un rôle essentiel pour répondre à la demande croissante d'énergie dans le monde.

Au Canada, l'énergie éolienne apparaît comme l'option la plus économiquement réalisable pour accroître la production d'électricité, comme le souligne l'Association canadienne des énergies renouvelables (CanREA, 2023). Les provinces de l'est du Canada, notamment l'Ontario, le Québec et Terre-Neuve, présentent un potentiel prometteur pour l'exploitation de l'énergie éolienne afin de répondre aux besoins énergétiques croissants de la région. Par conséquent, ces régions ont été sélectionnées pour des études de cas détaillées afin de comprendre la dynamique de l'intégration de l'énergie éolienne. Cependant, les perspectives prometteuses de l'énergie éolienne, la prévision reste un défi important pour son intégration transparente dans les réseaux électriques existants. Les prévisions permettent un équilibrage efficace de l'énergie, une planification optimale de la production et du stockage, et éclairent les décisions cruciales concernant la répartition et la maintenance de l'énergie. Il est essentiel de relever ce défi pour maximiser la fiabilité et l'efficacité des systèmes d'énergie éolienne.

Les prévisions climatiques saisonnières constituent une solution viable pour améliorer la précision des prévisions en matière d'énergie éolienne. En fournissant des informations avancées sur les conditions climatiques à venir, les décideurs peuvent atténuer les risques associés aux conditions météorologiques défavorables et tirer parti des conditions climatiques favorables. Ces prévisions reposent sur la compréhension des interactions entre l'atmosphère et les composantes plus lentes du système climatique, telles que l'océan, qui donnent lieu à une variabilité climatique à basse fréquence.

Les oscillations climatiques sont de puissants indicateurs des changements climatiques saisonniers et jouent un rôle essentiel dans l'élaboration des prévisions climatiques saisonnières. Plusieurs centres météorologiques opérationnels dans le monde utilisent des modèles de circulation globale pour générer des prévisions d'anomalies climatiques saisonnières. Ces prévisions offrent des projections des conditions météorologiques moyennes saisonnières, couvrant généralement une période allant jusqu'à trois mois avant la saison cible. Si les corrélations entre les oscillations climatiques et les régimes de température et de précipitations ont fait l'objet de nombreuses recherches, l'analyse des tendances en matière de vitesse du vent présente des lacunes. Malgré le rôle critique de la vitesse du vent dans la production d'énergie éolienne, elle a reçu comparativement moins d'attention dans les études d'analyse des tendances. Il est essentiel de combler cette lacune dans la recherche pour acquérir une compréhension globale des facteurs influençant la production d'énergie éolienne et améliorer la précision des prévisions dans le secteur des énergies renouvelables.

Des prévisions fiables sont essentielles pour une mise en œuvre réussie de l'énergie éolienne dans le système électrique et pour une programmation efficace. Les approches modernes comprennent une variété de modèles autonomes basés sur des données ainsi que des modèles hybrides composés de modèles physiques et de modèles basés sur des données pour différents horizons temporels. Les prévisions saisonnières sont souvent utilisées pour avoir une meilleure idée générale des régimes de vent et des capacités. Les modèles en place pour fournir des prévisions saisonnières sont basés sur les conditions atmosphériques et de surface ainsi que sur des modèles numériques de prévision météorologique. Des prévisions basées sur des téléconnexions ont été faites pour des facteurs climatiques tels que les précipitations et la température ; cependant, il y a un manque général d'évaluation des vents basée sur les oscillations climatiques.

Méthodes

Cette étude explore une approche innovante qui exploite les corrélations entre les indices climatiques et la vitesse du vent. Les indices climatiques pris en compte sont énumérés dans le tableau 2.1.

Cette approche utilise les données ERA5 du Centre européen pour les prévisions météorologiques à moyen terme pour la période 1982-2022, avec une résolution spatiale de 0,25° $\times$ 0,25°. La zone d'étude couvre l'est du Canada, de 55° de longitude ouest à 92° de longitude ouest et de 44° de latitude nord à 65° de latitude nord. Les données sont ramenées à une résolution de 1° $\times$ 1° et un masque terre-mer est appliqué. La variable dépendante est l'anomalie normalisée de la vitesse moyenne du vent en avril-mai-juin (AMJ) dans la zone d'étude, en se référant aux statistiques à long terme de 1950-2021.

Les prédicteurs sont des indices climatiques corrélés, et le coefficient de corrélation de Pearson est utilisé pour évaluer leur relation linéaire avec la variable dépendante. Les valeurs moyennes saisonnières des indices climatiques sont utilisées comme covariables, la corrélation étant calculée pour les variables prédictives décalées, de février-mars-avril (FMA) de l'année précédente à décembre-janvier-février (DJF) de l'année en cours. Les indices climatiques présentant des corrélations statistiquement significativ sur la base loi de Student à 95 % (Wilks, 2011) sont inclus dans un ensemble de prédicteurs possibles. Enfin, la saison décalée qui présente la corrélation la plus élevée est sélectionnée pour chaque indice climatique présentant des corrélations statistiquement significatives. Pour chaque indice climatique présentant des corrélations statistiquement significatives. Toutes les variables ont été linéairement détriangées avant l'analyse et une correction du biais basée sur Alessandrini *et al.* (2013) a été effectuée. Le Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT) n'a cessé d'améliorer ses systèmes de prévisions saisonnières depuis 1997, avec notamment le lancement du système de cinquième génération, SEAS5, en novembre 2017, qui a remplacé SEAS4. SEAS5 intègre des améliorations dans les modèles et les conditions initiales, en particulier dans le modèle atmosphérique du système de prévision intégré (IFS). Initialement axé sur les prévisions d'ensemble à grande échelle pour 10 à 46 jours, SEAS5 a été affiné pour améliorer la précision grâce à diverses modifications.

Les prévisions de la vitesse du vent sont affectées par des imprécisions provenant des limites de la simulation précise de tous les processus pertinents contribuant à la variabilité du climat. Pour y remédier, les prévisions dynamiques sont soumises à une phase d'ajustement des biais afin de s'aligner sur les bases d'observations, d'atténuer les imprécisions des prévisions et de garantir la fiabilité des prévisions. L'industrie de l'énergie éolienne reconnaît la nécessité d'ajuster les biais liés à la vitesse du vent pour assurer la fiabilité des décisions. Dans la présente étude, la correction des biais est effectuée conformément à l'équation suivante.

EANN

Les réseaux neuronaux artificiels (RNA), utilisés dans l'apprentissage automatique, offrent une alternative puissante aux modèles statistiques et de régression conventionnels (Dave & Dutta, 2014). Cette étude utilise une structure RNA feedforward comprenant des couches d'entrée, cachées et de sortie. La couche d'entrée contient des prédicteurs dont le nombre varie pour chaque point de grille. La couche cachée fonctionne avec une ou trois unités, tandis que la couche de sortie se compose d'une seule unité. Pour améliorer la stabilité et la généralisation, le modèle utilise des ensembles redondants, combinant plusieurs ANN (Shu & Burn, 2004; Sharkey, 2012; Kai & Salamon, 1990). L'implémentation se fait en Python 3.10, en utilisant les bibliothèques Tensorflow et Keras, avec un ensemble de 20 membres générés via l'algorithme de bagging (Breiman, 1996). Cet algorithme crée divers sous-ensembles en rééchantillonnant l'ensemble de données original avec remplacement, chaque modèle étant entraîné sur un sous-ensemble différent et leurs prédictions étant combinées par le biais d'une moyenne. Pour atténuer le surajustement, la régularisation L2 est appliquée, augmentant la fonction de perte avec le carré de la magnitude des coefficients, en utilisant des constantes de régularisation de 10^{-2} et 10^{-5} pour quatre ensembles d'hyperparamètres. Chaque ANN est formé via l'algorithme de rétropropagation, en ajustant la matrice de poids sur la base du gradient de la fonction de perte, avec un paramètre de taux d'apprentissage fixe de 10^{-3}. La fonction d'activation tangente hyperbolique est employée pour la couche cachée et l'activation linéaire pour la couche de sortie. L'apprentissage s'arrête soit lorsque le gradient de la fonction de perte tombe en dessous de 10^{-2}, soit après 10 000 époques. Le processus de formation est illustré avec la figure 2.1.

Vérification et évaluation des prévisions

Dans cette étude, l'évaluation des prévisions saisonnières implique l'évaluation de leur performance par rapport à une norme de référence, un processus connu sous le nom d'évaluation de la qualité des prévisions. La validation croisée "Leave-one-out" est utilisée pour l'évaluation de l'ensemble, dans laquelle chaque année de 1982 à 2021 est exclue une par une pour correspondre aux périodes de prévisions rétrospectives et de prévisions SEAS5. Il en résulte un ensemble d'entraînement comprenant 70 échantillons, dérivés des années entre 1951 et 2021, ainsi que des statistiques à long terme de la période 1951-1981. Les anomalies normalisées sont calculées pour la période d'entraînement de 70 ans, ainsi que pour l'année exclue en utilisant les mêmes statistiques à long terme. Les modèles sont entraînés avec des sous-ensembles de l'ensemble d'entraînement générés par le Bagging et utilisés pour prédire l'observation omise, ce qui permet d'obtenir 20 prédictions agrégées par leur moyenne. Les mesures d'évaluation comprennent l'erreur quadratique moyenne (RMSE), qui évalue la précision des prévisions par rapport à l'ampleur des erreurs, et le biais, qui mesure l'erreur moyenne des prévisions. Le score de compétence RMSE (RMSESS) est utilisé pour évaluer la précision par rapport aux prévisions climatologiques, calculé en pourcentage sur la base de la RMSE de la climatologie.

Étude et résultats

Les résultats, présentés dans la figure 2.2, montrent les corrélations entre chaque indice climatique considéré pour différentes saisons précédant la saison de prévision AMJ. La colonne de gauche montre la corrélation avec la saison AMJ exactement un an avant la saison AMJ en question. Viennent ensuite les saisons JJA, ASO, OND et DJF. Les corrélations existantes confirment l'objectif d'utiliser des indices climatiques corrélés comme prédicteurs. En outre, la figure 2.3 met en évidence les saisons les plus influentes pour chaque indice, soulignant la complexité spatio-temporelle de ces relations. Parallèlement, une analyse approfondie des hyperparamètres du modèle révèle que des configurations simples peuvent rivaliser avec des modèles plus complexes. En particulier, le premier des quatre ensembles d'hyperparamètres, caractérisé par une seule unité dans la couche cachée et un taux de régularisation de 10^{-2}, donne des résultats remarquables, comme l'illustre le tableau 2.2. Il en résulte une RMSE moyenne globale de 0,868 m/s et un biais de -0,084. L'impact du nombre de membres de l'ensemble est minime, comme le montre la figure 2.5. Enfin, la comparaison avec le modèle SEAS5, présentée dans la figure 2.6, met en évidence les avantages potentiels du modèle EANN dans certaines régions. Dans un premier temps, il apparaît clairement que les performances des deux modèles varient en fonction de la région examinée. La précision relative de SEAS5 varie de -10 % à 12 %, tandis que les performances d'EANN varient de -10 % à 14 %. SEAS5 n'atteint une précision relative de plus de 10 % qu'en deux points situés à l'extrême nord-ouest de la zone sélectionnée. Sur la base de cette étude, une approche de prévision de la vitesse du vent avec un EANN basé sur des indices climatiques s'avère rentable à des fins opérationnelles, car un tel modèle est plus facile à mettre en œuvre et moins coûteux en termes de calcul que les modèles dynamiques. Cette étude contribue aux efforts mondiaux visant à permettre des prévisions plus précises grâce à l'introduction d'une nouvelle approche. Dans le prolongement, une hybridation basée sur cette approche serait intéressante. Le modèle étant plus performant que le modèle SEAS5 dans certains domaines, il présente un intérêt supplémentaire en tant que stratégie d'amélioration. En outre, il serait utile d'effectuer des combinaisons supplémentaires d'hyperparamètres pour saisir les capacités de performance étendues du modèle.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AMJ	Season April-May-June
AMO	Atlantic Multi-decadal Oscillation
ANN	Artificial Neural Networks
ARIMA	Autoregressive Integrated Moving Average
ASO	Season August-September-October
DJF	Season December-January-February
EANN	Ensemble of Artificial Neural Networks
ECMWF	European Centre for Medium-Range Weather Forecasts
ENSO	El Niño-Southern Oscillation
ERA5	ECMWF Reanalysis v5
INRS	Institut national de la recherche scientifique
JJA	Season June-July-August
NAO	North Atlanctic Oscillation
NWP	Numerical weather prediction
OND	Season October-November-December
PDO	Pacific Decadal Oscillation
WPF	Wind Power Forecasting

1 GENERAL INTRODUCTION

The sustainable progress of renewable energy generation is highly dependent on the efficient use of renewable resources, which is vital for combating climate change. As the global shift towards cleaner energy sources gains momentum, effective management and enhancement of wind energy are key in meeting the increasing energy needs worldwide.

The efficient use of sustainable resources is vital for the sustainable progression of renewable energy production. For instance, in wind power, the control and enhancement of wind power are crucial in fulfilling the rising worldwide energy needs. Accurate predictions facilitate effective energy management, generation and storage planning, and impact decisions related to dispatch and maintenance. In Canada, wind energy is the most economical way to generate new electricity CanREA (2023). According to Dehghani-Sanij *et al.* (2022), the eastern Canadian provinces (Ontario, Quebec, New Brunswick, Nova Scotia, Prince Edward Island and Newfoundland and Labrador) have great potential to meet the growing demand for wind energy, emphasizing the selected area for a case study.

Forecasting remains the biggest challenge in integration of wind power (Hanifi *et al.*, 2020).

Utilizing seasonal climate prediction is a valuable method to anticipate wind speeds in the context of wind energy. Seasonal climate forecasting allows decision makers to take advanced actions to minimize risks associated with adverse climate conditions or maximize the benefits from favorable ones. It relies on the interactions of the atmosphere with slower components of the climate system, such as the ocean, which produce modes of low-frequency variability (WMO, 2020). Climate oscillations are strong indicators for seasonal climate shifts and have been used for practical forecasting in various areas. In fact, the field of forecasting seasonal weather fluctuations has evolved into a regular operational practice in meteorological forecasting services. Numerous operational meteorological centers worldwide regularly generate seasonal climate anomaly predictions based on global circulation models (Weisheimer *et al.*, 2020). In contrast to traditional weather predictions, seasonal forecasts offer projections of seasonal-average weather conditions, usually for a period of up to three months preceding the target season.

The number of available studies investigating the correlation between climate oscillations and temperature and precipitation is increasing. However, the field lacks an analysis of correlations between climate indices and wind speeds, confirming the fact that wind speed records have been studied much less than other climate and weather measures such as temperature and precipitation data when it comes to long-term variability analysis (Pryor *et al.*, 2020).

This research links the correlations between teleconnection and wind speeds in eastern Canada.

1.1 Wind energy

Fundamentally, wind energy operates by transforming the kinetic energy of air into electrical energy. Generally, wind is plentiful, free of charge, widely available, and its production is environmentally friendly (Irena, 2019). In comparison to other uses of renewable energy technologies, wind power generation stands out due to its advanced technology, well-established infrastructure, and cost-effectiveness. Wind and solar energy are the primary sources of energy production in the scenario of net zero emissions by 2050 and are expected to have a growing significance in the future energy landscape. (Fung *et al.*, 1981). To stay on track with this path, the amount of wind capacity added each year must increase significantly until 2030 (IEA, 2022).
Wind turbines can be set up alone or in clusters to form wind farms, depending on the area and the climate conditions. To identify the most suitable location for wind farms and the ideal turbine heights, numerous evaluations must be performed. Wind speed has been identified as the most crucial factor in decision-making.(Rediske *et al.*, 2021; Giani *et al.*, 2020). In the field of wind energy, onshore wind farms are the most established and widely utilized globally. On the contrary, offshore wind farms are still in their early stages of development, with the disparities between the two technologies attributed to technical challenges and costs (Ahmed *et al.*, 2020).
In Canada, wind energy is the most economical way to generate new electricity (CanREA, 2023). The majority of installed wind power in Canada is located in the provinces of Ontario, Quebec, and Alberta. According to Dehghani-Sanij *et al.* (2022), the eastern Canadian territories have great potential to meet the growing demand for wind energy, which qualifies and emphasizes the area's selection for a case study.

1.2 Teleconnection

Teleconnection refers to a significant and persistent atmospheric link between weather or climate conditions in different parts of the world. These connections often occur over large distances and involve the exchange of energy and information between remote regions. Teleconnections play a crucial role in shaping global and regional climate patterns. The most well-known teleconnections often involve large-scale climate phenomena (Liu & Alexander, 2007; An *et al.*, 2020).

The best known example is the El Niño-Southern Oscillation (ENSO), which is a major teleconnection pattern associated with periodic warming (El Niño) and cooling (La Niña) of sea surface temperatures in the central and eastern equatorial Pacific Ocean. The impacts of the El Niño Southern Oscillation (ENSO) are experienced in different parts of the globe, affecting weather conditions, rainfall, and temperatures (McPhaden *et al.*, 2020).
Another well known example is the North Atlantic Oscillation (NAO), which is a teleconnection pattern in the north Atlantic region, characterized by variations in atmospheric pressure between the Icelandic Low and the Azores High. It affects the strength and position of the jet stream, influencing

the weather conditions in Europe, North America, and parts of Africa.
In the northeastern region of Canada, there is a significant change from moist polar conditions to dry polar conditions under the influence of a positive North Atlantic Oscillation (+NAO). In contrast, in the eastern United States, both types of polar weather are much less frequent during positive NAO (Sheridan, 2003). A third example of a climate oscillation is the Pacific Decadal Oscillation (PDO). The PDO is a longer-term teleconnection pattern in the north Pacific Ocean, characterized by anomalies in sea surface temperature. It can influence climate conditions over decades and is associated with changes in ocean atmospheric patterns (Johnson *et al.*, 2020). PDO is also an example of an index, with a predictability influenced by climate change, making is less predictable (Li *et al.*, 2020)
Teleconnections play a crucial role in comprehending and forecasting climate variability worldwide. Alterations in one geographical area can trigger a chain reaction on atmospheric conditions in remote areas, impacting ecosystems, agriculture, and human undertakings. Scientists use various indices and models to study and monitor teleconnections, helping to improve climate predictions and assessments. In this study, the focus is on determining how indices are correlated with wind developments and how they can be used for forecasting.

When looking at existing studies dealing with the relation of climatogoy and wind speed, Naizghi & Ouarda (2017) take wind data from six ground stations in the UAE and analyze the presence of periodicity, trends, and change points in wind speed time series. They also investigate the relationship between wind speed and climate oscillations using the cross-wavelet transform and wavelet coherence analysis. Multiple linear regression models are developed to represent wind speed based on various climate indices. The results show patterns of spatial distribution of wind speed in the UAE, with north and northwest winds blowing from the sea throughout the year. The study also identified significant trends in wind speed at different stations, some showing increasing trends and others showing decreasing trends. Teleconnection analysis reveals correlations between wind speed and climate indices. In conclusion, the paper states that wavelet techniques can be effective in studying wind speed variability and teleconnections, providing valuable information for the evaluation of wind energy potential.

A more specific reference point for selection of climate indices to consider is delivered by Ouarda & Charron (2021), who evaluated the potential for wind energy and wind farm design by modeling the probability distribution of wind speed using short-term data. The study used wind speed data from 20 meteorological stations in Quebec, Canada, with record periods of more than 30 years. The data were analyzed using non-stationary statistical models, considering covariates such as atmospheric circulation indices as NAO and PNA and temporal trends. The shape of the wind speed distribution varied depending on the covariates, indicating the importance of considering interannual variability and long-term trends in wind speed modeling. The proposed approach has the potential to predict the potential of wind energy during the projected lifetime of wind farms. The study suggested that future research should explore other methods for non-stationary modeling of wind speed data. It also recommended the use of reanalysis data sets that combine models with observational data for more accurate and fine-grained spatial information. In addition, Vincent *et al.* (2015) indicate the influence of PDO and AMO on the climate of Canada. Finally, Woldesellasse *et al.* (2020) confirm the use of teleconnection indices, wind speed, and its direction to be used in a nonlinear correlation analysis to predict wind speed, indicating an adequate estimation of the variability.

1.3 Seasonal forecasts based on climate variability

Seasonal forecasting is the practice of predicting climate conditions over a period of several months, typically more than one month but less than one year. It is relevant because it allows society to anticipate and prepare for climate anomalies, such as droughts or favorable conditions, to minimize risks and maximize benefits in sectors such as agriculture, food security, health, energy, water management and disaster risk reduction (WMO, 2020; Troccoli *et al.*, 2008).

Looking at existing forecasting tools based on climate variabilities, various approaches have been explored, with an overall lack of seasonal forecast models of wind speed. Numerous studies have been conducted to analyze long-term forecasts of precipitation and temperature, yet only a few have focused on long-term forecasting of wind speed (Pryor *et al.*, 2020), despite its increasing importance due to the rapid growth of wind energy as a power source. This makes the study approach for this project interesting and allows for similar approaches as the literature dealing with e.g. precipitation forecasts.

Several studies have included analysis of climate indexes and correlations between indices and weather phenomena, such as precipitation and temperature. A study by Ouarda & Charron (2021) demonstrates a climate indices analysis through a nonstationary Weibull probability density function that was used to model the long-term distribution of hourly wind speed, with climate indices and time as covariates. Another study by Chandran *et al.* (2016) examines the relationship between climate indices and seasonal precipitation and temperature in the United Arab Emirates (UAE). It aims to gain a better understanding of the climate oscillations that affect precipitation and temperature in the region. Work by Schepen *et al.* (2012) for example provides a thorough analysis of the use of lagged climate indices as predictors of seasonal rainfall in Australia.

As one of the biggest challenges in wind power implementation is accurate forecasting, improving and developing new forecasting methods is the key to reliable expansion.

1.4 Neural Networks in forecasting

Neural networks are a powerful tool to forecast long-term wind speed and recognize general patterns (Azad *et al.*, 2014). As mentioned, use of ANN based forecasting is currently further explored in analysis of precipitation and temperature.

For example the work of Mekanik *et al.* (2013) focuses on the development and comparison of Artificial Neural Network (ANN) models and Multiple Regression (MR) models for the forecast of rainfall. It analyzes the ability of El Nino Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) values to forecast future spring rainfall in Victoria, Australia. The results show that the ANN models have a better generalization ability and lower errors compared to the MR models, suggesting their potential for rainfall forecasting using large-scale climate modes. Ouarda *et al.* (2020) examine data from three rainfall stations in the United Arab Emirates, providing a non-stationary framework. They compare the uncertainties associated with stationary and nonstationary models, quantifying them through the mean confidence intervals computed by parametric bootstrapping.

Shu & Burn (2004) provide an interesting framework for the use of ensembles of artificial neural networks for a pooled flood frequency analysis. Valuable conclusions include an ensemble size of at least 10 ANNs and the higher accuracy of an ensemble compared to a single ANN. Another successful instance, Pinheiro & Ouarda (2023), confirms the use of EANNs based on the climate indices ONI, TNA and TSA for forecasting precipitation in Brazil. The EANN model demonstrates itself as a well-calibrated model with moderate confidence.

As for the use of neural networks in wind speed forecasting, the study conducted by Li (2022) provides a spatiotemporal analysis of wind data using deep residual networks, but focuses on very short-term prediction based on relationships among adjacent turbines. In their review, Wang *et al.* (2021) summarizes a variety of machine learning methods for wind speed and wind power forecasting. The benefits of the planned approach of an EANN are confirmed by advantages such as low computational load and potential for hybridization. The authors point out that the primary obstacles in predicting wind speed using deep neural networks are uncertainties in the data, incomplete features, and complex nonlinear relationships.

With the existing literature taken into account, the relevance of this suggested research becomes evident. Furthermore, it becomes clear that a teleconnection delivers a strong basis, and investigation of correlations can be further analyzed.

1.5 Ensemble of artificial neural networks

In this research, a model based on neural networks is selected to be built. Artificial neural networks simplify complex dynamics and functions into mathematical models inspired by the structure and function of the brain. Their significance lies in their ability to handle massive data sets and perform tasks that were previously difficult or impossible to automate effectively. They continue to advance with deep learning architectures that enable even deeper insights from the data.

1.5.1 Structure and Function of Artificial Neural Networks

Artificial neural networks (ANNs) are computational models inspired by the structure and function of biological neurons. They process information through interconnected layers of artificial neurons, mimicking the way neurons in the brain transmit and process signals. ANNs perform computations by combining inputs with respective weights to produce a weighted sum, known as preactivation, and then applying a chosen function (linear or nonlinear) to determine the neuron's output, in a process called activation (Yegnanarayana, 2009).

These networks are usually made up of input, hidden, and output layers. The input layer is where external data are received, the hidden layers are responsible for processing and changing these data, and the output layer produces the final result, such as classifications or predictions. The hidden layers are vital in improving the network's expressiveness by facilitating nonlinear transformations of inputs. They enable ANNs to efficiently approximate complex functions. (Piccinini, 2008; Miller *et al.*, 1990; Rosenblatt, 1958). The depth of the network, which refers to the number of hidden layers it has, affects its capability to learn complex functions and generalize to new instances. Deep networks, which consist of multiple hidden layers, can capture intricate features and relationships present in the data. The network's depth brings significant advantages as deep networks can leverage compositional structures in functions by reusing features calculated in lower layers. This hierarchical strategy improves accuracy and generalization, particularly in tasks with large datasets such as image and speech recognition, natural language processing, and reinforcement learning (Abraham, 2005; Zou *et al.*, 2009). Latest improvements include techniques such as pretraining, initialization, regularization, and normalization, along with the introduction of rectified linear units, and generally have significantly improved performance (Kriegeskorte & Golan, 2019).

Training deep neural networks involves an incremental adjustment of the weights using gradient descent. This process efficiently computes derivatives of the error with respect to each weight, allowing the network to iteratively improve its performance. The stochastic gradient descent strikes a balance between stability and efficiency by using one sample per epoch of training examples.

By modifying the weights and biases, artificial neural networks (ANNs) have the capability to learn and adjust to different tasks, rendering them flexible instruments in the fields of machine learning and artificial intelligence, and the preferred method for this particular case study.

1.6 Dataset

The analysis is based on ECMWF Reanalysis 5 (ERA5) data which is a high-resolution global atmospheric data set created by the European Centre for Medium-Range Weather Forecasts (ECMWF). It integrates diverse data from weather stations and satellites with a numerical weather model to offer a thorough account of previous weather conditions. ERA5 provides hourly data, a wide range of meteorological variables, worldwide coverage from 1940 to the current time, and is well-known for its consistent and high-quality data. It is widely used for climate research, weather forecasting, and environmental monitoring. In terms of grid spacing, ERA5 uses a regular grid with grid points spaced at horizontally 31 km intervals with global coverage. The vertical resolution varies with altitude levels, providing data at multiple atmospheric levels from the surface to the upper atmosphere.

The wind data is derived from Hersbach *et al.* (2023) and comes divided into its components u and v. The u and v components of the wind refer to the east-west and north-south velocity components of the wind vector, respectively. The u component represents the wind movement from west to east (east) when positive and from east to west (west) when negative. Conversely, the v component indicates the wind movement northwards when positive and southwards when negative. Together, these components provide a description of the horizontal movement of the wind. Wind speed is calculated as

$$\mathrm{windspeed} = \sqrt{\mathrm{u}^2 + \mathrm{v}^2}. \tag{1.1}$$

Before the data is implemented, it is prepared by adjustment for bias, including calibration, to improve statistical properties Torralba *et al.* (2017) and ensure reliability and accuracy.

1.7 Structure and Case study

This project focuses on the analysis of wind characteristics, investigating the impacts of climate oscillations on seasonal wind speed interannual variability with the purpose of building a seasonal wind speed EANN forecasting model. It involves teleconnection analysis to identify climate oscillation indices that affect wind speed and the development of pointwise nonlinear models of wind speed that take into account teleconnection patterns. ERA5 gridded wind speed and climate indices data are used to conduct the teleconnection analysis and EANN fitting.

The case study defined for the initial project is in eastern Canada. A map of the selected area is shown in figure 1.1, showing coverage from 55° W to 92° W in longitude and from 44° N to 65° N in latitude. The selection covers the province of Quebec, New Brunswick, Prince Edward Island and vast majorities of Ontario, Nova Scotia and Newfoundland and Labrador as well as a small section of the southeastern part of Nunavut. The examined case study covers an area of more than $5 \times 10^6\ \mathrm{km}^2$

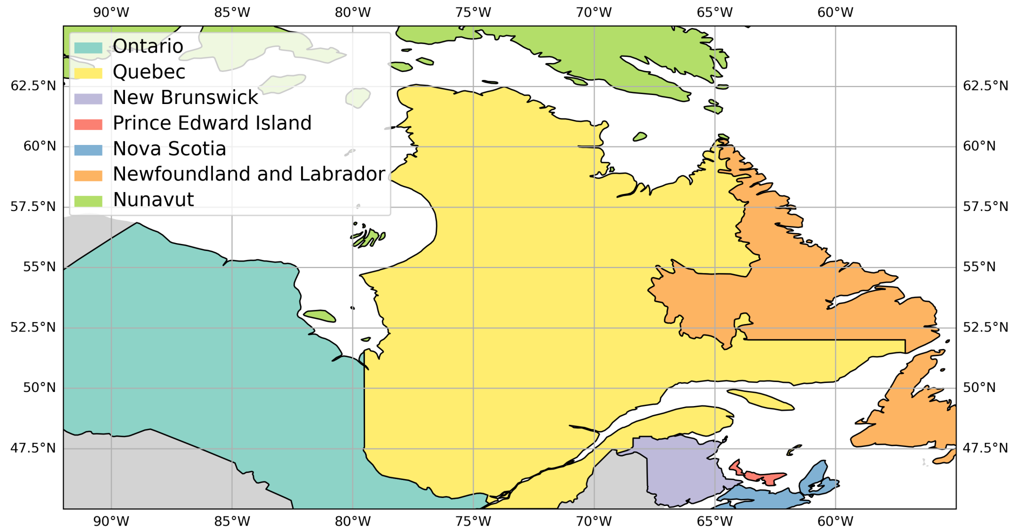


Figure 1.1 : Selected area for case study covering eastern Canada

2 PAPER

Ensemble of artifical neural networks for seasonal forecasting of wind speed in eastern Canada

Ensemble de réseaux de neurones artificiels pour la prévision saisonnière de la vitesse du vent dans l'Est du Canada

Authors

Pia Leminski[1], Enzo Pinheiro[1], Taha Ouarda[1]

[1] INRS - ETE

Abstract

Efficient harnessing of wind energy resources, including advancements in weather and seasonal forecasting and climate projections, is imperative for the sustainable progress of wind power generation. While temperature and precipitation data receive considerable attention in interannual variability and seasonal forecasting studies, there is a notable gap in exploring correlations between climate indices and wind speeds. This paper proposes the use of an ensemble of artificial neural networks to forecast wind speeds based on climate oscillation indices and assesses its performance. An initial examination indicates a correlation signal between climate indices and wind speeds of ERA5 for the selected case study in eastern Canada. Forecasts are made for the season April-May-June and based on most correlated climate indices of preceding seasons. A pointwise forecast is conducted with a 20 member ensemble, which is verified by leave-on-out cross-validation. The results obtained are analyzed in terms of root mean squared error, bias, and skill score and show a competitive performance with state-of-the-art numerical wind predictions from SEAS5, outperforming them in several regions. A relatively basic model with a single unit in the hidden layer and regularization rate of 10^{-2} provides promising results, especially in areas with a higher number of indices considered. This study adds to global efforts to enable more accurate forecasting by introducing a novel approach.

Keywords: Wind, EANN, Seasonal forecasting, ERA5

2.1 Introduction

The effective utilization of renewable resources is essential for the sustainable advancement of renewable power generation and a crucial component in mitigating climate change. As the world increasingly transitions toward cleaner and more environmentally friendly energy sources, the management and optimization of wind energy plays a pivotal role in meeting the growing global energy demands. Reliable forecasts enable efficient energy balancing, generation, and storage planning, and affect dispatch and maintenance choices. In Canada, wind energy is the most economical way to generate new electricity (CanREA, 2023). According to (Dehghani-Sanij *et al.*, 2022), the eastern Canadian territories (Ontario, Quebec, New Brunswick, Nova Scotia, Prince Edward Island and Newfoundland and Labrador) have great potential to meet the growing demand for wind energy, emphasizing the selected area for a case study.

Forecasting remains the biggest challenge in integration of wind power (Hanifi *et al.*, 2020). Seasonal climate forecasting allows decision makers to take advanced actions to minimize risks associated with adverse climate conditions or maximize the benefits of favorable ones. It relies on the interactions of the atmosphere with slower components of the climate system, such as the ocean, which produce modes of low-frequency variability (WMO, 2020). Climate oscillations are strong indicators for seasonal climate changes and have been used for practical forecasting in various areas. In fact, the field of forecasting seasonal weather fluctuations has evolved into a regular operational practice in meteorological forecasting services. Numerous operational meteorological centers around the world regularly generate seasonal climate anomaly predictions based on global circulation models (Weisheimer *et al.*, 2020). In contrast to traditional weather predictions, seasonal forecasts offer projections of seasonal-average weather conditions, usually for a period of up to three months preceding the target season. There are numerous studies available investigating the correlation between climate oscillations and temperature and precipitation. However, the field lacks an analysis of correlations between climate indices and wind speeds, confirming the fact that wind speed records have been studied much less than other climate and weather measures such as temperature and precipitation data when it comes tolong-term variability (Pryor *et al.*, 2020).

2.2 Background

2.2.1 Wind Forecasting

The literature extensively discusses various wind forecasting methods, frequently classified as physical-based, data-driven, and hybrid methodologies (Ouyang *et al.*, 2019). For long-term forecasting, physical approaches, such as numeric weather prediction (NWP), are standard and produce highly accurate predictions under stable weather conditions (Potter & Negnevitsky, 2006;

Candy *et al.*, 2009), but are expensive and require a significant amount of time (Tascikaraoglu & Uzunoglu, 2014). In comparison, data-driven models extract functional relationships from the data themselves to construct models that illustrate the connections between wind power and other input variables (Yan & Ouyang, 2019). Statistical approaches encompass models based on time series and (artificial) neural networks (ANN) (Naizghi & Ouarda, 2017; González-Sopeña *et al.*, 2021; Kim & Hur, 2020).

Hybrid approaches, which integrate different methods or combine short- and medium-term models, have demonstrated improved forecasting capabilities. In particular, the incorporation of neural networks significantly improves the utility of NWPs (Soman *et al.*, 2010). A recent emerging approach considers the spatial relationship between wind speeds, using the correlation between speeds at various locations. Models predicting wind speed, often employing neural networks or adaptive neuro-fuzzy networks, leverage data from specific and nearby locations, considering temporal lags (Khazaei *et al.*, 2022; Woldesellasse *et al.*, 2020). Another method that is gaining increasing attention is looking into transformer architectures for wind speed forecasting, which find promising results both for standalone models and for transformer-based enhancements (Bentsen *et al.*, 2023; Sarkar *et al.*, 2023; Huang *et al.*, 2023). Early studies, such as Watson *et al.* (1994), demonstrate the potential for fossil fuel savings of 15% to 25% by incorporating different levels of wind power through a combination of NWP models, a physical flow model, and a statistical model for forecasting.

The current wind power forecasting (WPF) method generates deterministic spot predictions based on conditional expectations. Efforts to enhance prediction accuracy have been a focal point of research (Costa *et al.*, 2008; Giebel *et al.*, 2011). Due to the unpredictable nature of atmospheric conditions, inherent uncertainty is associated with every forecasting method. Recent years have seen a growing emphasis on uncertainty prediction and probabilistic forecasting in prediction theory, particularly in weather forecasting (Palmer, 2002), providing quantitative insights into the variability associated with wind power generation and serving as optimal inputs for decision-making under uncertain conditions (Zhang *et al.*, 2014; Ouarda & Charron, 2021).
While NWP remains dominant in weather forecasting, machine learning-driven weather forecasting is emerging as a viable alternative. Recent cases showcase the potential of machine learning-driven weather forecasting, particularly in areas where traditional NWP may be relatively weak (Lam *et al.*, 2023). ANNs have been applied in wind forecasting and spatial correlations for several decades (Alexiadis *et al.*, 1999; Corotis *et al.*, 1977; Giebel *et al.*, 2003). The latest studies, such as Ti *et al.* (2021) demonstrate that ANN models made a reasonable prediction of wind power and outperformed analytical models due to their flexibility and capacity to capture nonlinearity in the process. Ensemble forecasting has led to significant advances in NWP models in recent years. It offers multiple equally probable scenarios of meteorological variables, as opposed to the single scenario typically provided by traditional NWP Zhang *et al.* (2014). Studies such as Wu *et al.* (2022) examine the correlation between wind surface data and climate features, such as, for example, temperature, dew point, and sea level pressure, confirming the correlation between

wind speeds and climate features. Modern hybrid modeling approaches, which incorporate machine learning techniques, exhibit promising potential (Lipu *et al.*, 2021). For instance, Singh *et al.* (2019) demonstrate succesful hybridization of autoregressive integrated moving average (ARIMA) and ANN compared to separate models for wind power forecasting in Denmark.

2.2.2 Case

An extensive assessment of teleconnection-based ensemble learning approaches for wind prediction remains a notable gap in the current research domain. Hence, this study seeks to address this gap by thoroughly assessing the seasonal wind prediction capabilities of an ensemble of artificial neural networks (EANN) model based on large-scale climate oscillations. The performance of the EANN is in contrast to that of the current state-of-the-art SEAS5 model.
Eastern Canada was selected as a study due to its potential to expand wind power (Dehghani-Sanij *et al.*, 2022) and its unique energy profile that produces the potential to synthesize hydropower with wind power (Kapica *et al.*, 2021). In addition, influence of El Niño-Southern Oscillation (ENSO) on precipitation, as well as Atlantic Multi-decadal Oscillation (AMO) on temperature North America (Zhao & Brissette, 2022) emphasize the investigation of correlation of climate indices and wind.

The analysis is performed for the April-May-June (AMJ) season as this season currently provides a significant share of wind power production in Canada (StatisticsCanada, 2024). Furthermore, the AMJ period is critical for potential integration with hydropower and thus essential to evaluate. The goal of this study is to evaluate the performance of an ensemble of artificial neural networks for forecasting wind speeds in eastern Canada based on teleconnection.

The rest of the document is structured as follows: In the "Data" section, the datasets are introduced. The methodology employed to build the EANN and the validation process are detailed in the Methods section 2.3. A comparison of statistical and dynamical models is made in the Results and Discussion section 2.4, which also examines the correlation between model performance, index correlation, and seasonal dependence. Lastly, the Summary section provides a recap, an analysis of the key findings, and suggestions for future research.

2.3 Methods

2.3.1 Data

The predictant data used in this approach are the ERA5 data of the European Centre for Medium-Range Weather Forecasts (ECMWF). The reanalysis data are available in grid form with a spatial resolution of 0.25° x 0.25° and were extracted for the period 1950-2022. The selected study area extends from 55°W to 92°W in longitude and from 44°N to 65°N in eastern Canada. For later analysis, the data were further coarsened to a resolution of 1° x 1°. A land-sea mask was applied to disregard grid points in large bodies of water. The dependent variable is the 10m average wind speed standardized anomalies of the AMJ computed at each point in the grid of the selected area. The reference period for computing long-term statistics spans 1950-2021.

The predictors are made up of correlated climate indices. The Pearson correlation coefficient has been used as an indicator of the linear relationship between predictors and predictants. The covariables are seasonal mean values of any of the listed climate indices. At each grid point of the selected area, correlation is computed between predictors (lagging from the previous year FMA to the current year DJF) and the AMJ predicitant. Climate indices with statistically significant correlations based on a two-tailed t-student test at 95% (Wilks, 2011) are included in a pool of possible predictors. Finally, the lagged season that results in the highest correlation is selected for each climate index with statistically significant correlations. All variables were linearly detrended prior to analysis. A list of the monthly climate indices considered can be found in table 2.1.

Table 2.1 : List of climate indices selected for the case study and their symbols retrieved from NOAA (nd)

Climate index	Symbol
Atlantic Meridional Mode	AMM
Atlantic Multidecadal Oscillation	AMO
Arctic Oscillation	AO
Dipole Mode Index	DMI
North Atlantic Oscillation	NAO
Niño 4	Nino4
Niño 1+2	Nino12
Pacific Decadal Oscillation	PDO
Pacific North American Index	PNA

2.3.2 SEAS5

Since 1997, ECMWF has continuously improved its seasonal forecast systems. The fifth-generation system, SEAS5, launched in November 2017, replacing SEAS4. SEAS5 incorporates advancements in models and initial conditions, especially in the Integrated Forecast System's (IFS) atmosphere model. The project began with the extended range ensemble forecast for 10 to 46 days, leading to the development of SEAS5 with enhanced accuracy through various modifications (Johnson *et al.*, 2019).

Wind speed predictions are influenced by inaccuracies caused by limitations in accurately simulating all relevant processes that contribute to climate variability (Doblas-Reyes *et al.*, 2013). Dynamical forecasts require a bias adjustment phase to statistically mirror the observational baseline, reduce prediction inaccuracies, and establish reliable predictions. The wind energy industry has recognized the need to de-bias wind speed to ensure reliability for decision-making purposes (Alessandrini *et al.*, 2013). In this study, a bias correction has been performed according to (Leung *et al.*, 1999; Torralba *et al.*, 2017)

$$\mathrm{y_{ij}} = (\mathrm{x_{ij}} - \bar{\mathrm{x}})\frac{\sigma_{\mathrm{ref}}}{\sigma_{\mathrm{e}}} + \bar{\mathrm{o}}. \tag{2.1}$$

where y_{ij} denotes the new seasonal mean for each forecast year i and member j after bias correction. x_{ij} represents the seasonal average of each forecast for each year i and member j. $\bar{x}$ signifies the ensemble mean of the seasonal averages. σ_{ref} denotes the reference standard deviation. σ_e represents the interannual standard deviation of the ensemble members. $\bar{o}$ stands for the climatological mean of the reference dataset, which is added to the bias-corrected values to obtain the final result. Applied independently to each grid cell, this procedure refines the ensemble wind speed forecast, aligning it with the mean and standard deviation of the reference forecast (Torralba *et al.*, 2017).

2.3.3 Artificial Neural Network Ensembles

ANNs represent models used in machine learning (ML) and have emerged as a viable alternative to traditional regression and other statistical models in terms of their utility (Dave & Dutta, 2014).

This implemented model uses the method of Pinheiro & Ouarda (2023), adapted for a flexible selection of covariables. The model has a feed forward structure, comprising of an input, hidden, and output layer. The input layer of this multilayer perceptron ANN has the same number of units as predictors, potentially differing for each gridpoint. The hidden layer is run with one as well as three units and the output layer is made up of one unit. To create the ensemble, several ANNs are combined in redundant ensembles, which makes the model more stable and improves generalizability (Shu & Burn, 2004; Sharkey, 2012; Kai & Salamon, 1990). The model is implemented in Python 3.10, with the libraries Tensorflow and Keras. The number of members of the ensemble is set to 20, as a minimum of 10 members is typically suggested. Shu & Burn (2004). The ensemble members are created using Bagging Algorithim (Breiman, 1996). This algorithm works by resampling with replacement the original sample to generate diversified subsets of the original dataset. Each model is trained using a different subset, which are then combined through regular mean to generate the final prediction. This methodology was shown to improve overall predictive performance and reduce overfitting compared to single models (Shu & Burn, 2004; Shu & Ouarda, 2007). Another measure to reduce overfitting is L2 regularization, which is implemented to increase the loss function by adding the squared magnitude of coefficients, multiplied by a regularization constant of 10^{-2} and 10^{-5} as different sets of hyperparameters. Together with the different numbers of units in the hidden layer, this results in a total of four different sets of hyperparameters run. Each ANN is subject to training via the conventional backpropagation algorithm, which adjusts the weight matrix by computing the gradient of the loss function with a learning rate parameter fixed at 10^{-3}. The network operates a hyperbolic tangent activation function for the hidden layer and linear activation for the output layer. The training process is completed when either the loss function gradient reaches

below 10^{-2} or the number of epochs reaches 10^4. The modeling process
The training process is illustrated in Figure 2.1.

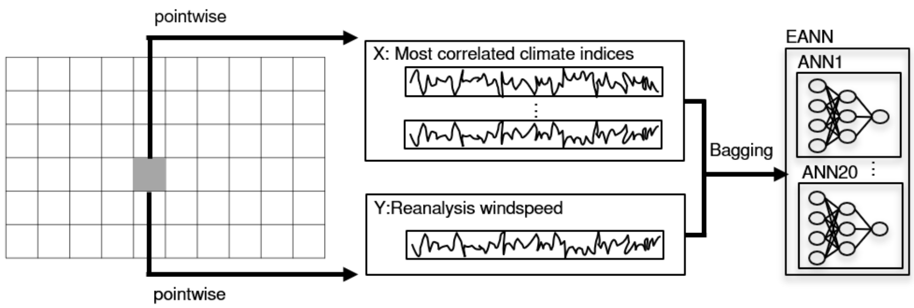


Figure 2.1 : Schematic illustration of training process

2.3.4 Forecast verification and evaluation

Seasonal predictions must be evaluated against a standard or reference to gauge their comprehensive performance referred to as forecast-quality assessment (Doblas-Reyes *et al.*, 2013; Torralba *et al.*, 2017). In this study, the ensemble assessment is carried out using leave-one-out cross-validation. In cross-validation, the year to be predicted is removed from the sample that generates the forecast. This is done for each year between 1982 and 2021 to match SEAS5 hindcast and forecast periods, leaving out the years one by one while calculating the long-term mean and standard deviations of the observation based on the remaining 39 years in addition to the period 1951-1981, resulting in a training set with 70 samples.
Subsequently, standardized anomalies were calculated for the 70-year period, covering the years between 1951 and 2021. In addition, anomalies for the excluded year were calculated using the same long-term statistics as the training set.The models were then trainined with subsets of the training set generated through Bagging and applied to predict the omitted observation. This results in 20 predictions that are aggregated by calculating their mean.
The statistical evaluation measures used include the root mean squared error (RMSE), as seen in equation 2.2 which evaluates the accuracy of the forecasts, considering magnitude of errors,

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\bar{\mathrm{y}}_{\mathrm{i}} - \mathrm{o}_{\mathrm{i}})^2} \tag{2.2}$$

where y_i and o_1 are the i-th of the n pairs of ensemble average and observation. Bias, which measures the mean error of the forecasts, is calculated as the difference between the predicted values and the actual outcomes, see equation 2.3.

$$\mathrm{Bias} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\bar{\mathrm{y}}_{\mathrm{i}} - \mathrm{o}_{\mathrm{i}}) \tag{2.3}$$

The RMSE skill score (RMSESS) is a thorough metric used to evaluate the accuracy relative to climatological forecasts (Wilks, 2011). In order to compute it, the RMSE of the climatology $\mathrm{RMSE_{clim}}$ is defined, with iteration over k, representing each time period for which forecasts are made and verified.

$$\mathrm{RMSE_{clim}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\bar{o} - o_k)^2} \tag{2.4}$$

this can then be used to calculate the comparative RMSESS in [%].

$$\mathrm{RMSESS} = (1 - \frac{\mathrm{RMSE}}{\mathrm{RMSE_{clim}}}) * 100 \tag{2.5}$$

2.4 Results and Discussion

2.4.1 Climate indices as predictants

As the model concept is based on the dependency of wind speed standardized anomalies on climate indices, figure 2.2 shows the correlations between each climate index considered for different seasons leading up to the forecast season AMJ. The left column shows the correlation with the AMJ season exactly one year before the season AMJ of interest. From there on follow the seasons JJA, ASO, OND and DJF.

Going through the indices, different intensities of signals can be observed. Overall, the indices NINO12, PDO, NAO and AO provide the highest values correlations, indicating high predictive power. The areas with the overall highest correlations throughout the indices are north eastern coastal area, central area, eastern area, and south western area. It becomes clear that a different correlation pattern can be identified for each index. The first illustrated index, NINO4, yields a high correlation for some grid points, especially in the AMJ season for the coastal region and relatively stable prediction zones throughout the seasons. The second index, NINO12, shows an overall higher correlation than NINO4, providing stronger level correlations ranging lower than -0.2 and higher than 0.3 in the southwestern area for AMJ and JJA and northeastern costal area in ASO, OND, and DJF. The index NPI shows highest correlations in the DJF and ASO seasons and a generally large area of consistent positive correlation for DJF, as well as negative for the ASO season. The PDO index shows a strong correlation with the northeastern coastal area in JJA and OND. It also indicates a correlation in the central southern region, most significant in the JJA and ASO seasons. PNA shows weaker overall correlations, with some stronger signals in the central and eastern regions in JJA and the coastal regions of Hudson Bay in DJF. The NAO indicates a strong correlation for the ASO season in the northwestern area of the examined area. The same area shows a strong negative correlation in the ASO season, expanding across the northern edge of the area, also featuring a very central area. The AMM correlation displays a comparatively

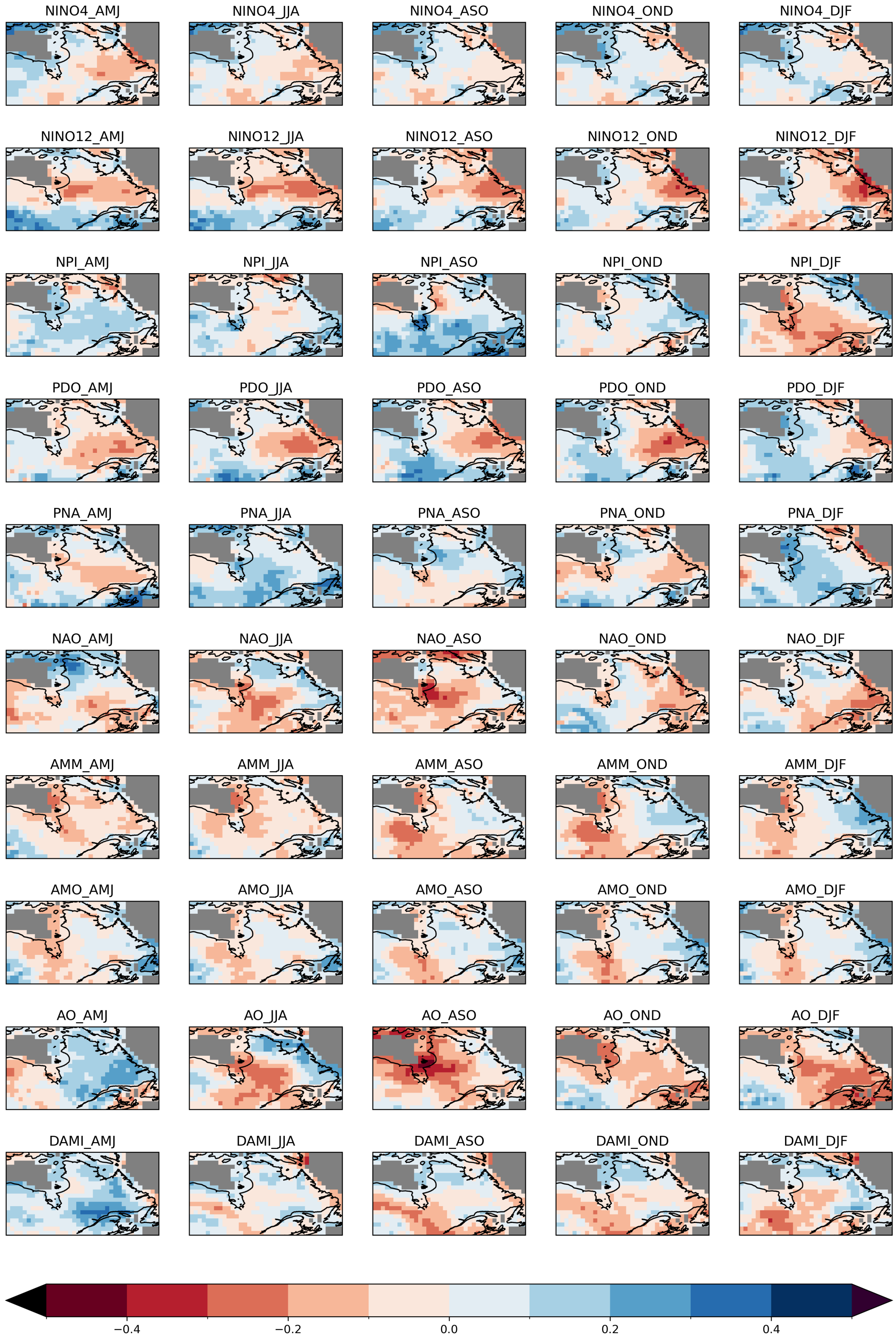


Figure 2.2 : Seasonal correlation of different selected climate indices with ERA5 data for season AMJ covering the years 1982-2021

high signal for a distinct area in northeastern Ontario. It also has a correlation with the eastern coast of Newfoundland and Labrador. The AO index provides strong correlation patterns for the southeastern section of the map in DJF and for the northwestern section in ASO.

The existing correlations confirm the use of correlated climate indices as predictors.

2.4.2 Season selection

In order to further assess the most relevant season for the prediction process, as well as the result of selected indices, Figure 2.3 shows the highest correlated seasons for each of the individual indices. In this case, all seasons with overlap are considered, resulting in a total of 11 seasons leading up to AMJ, beginning with FMA of the previous year and going until DJF before the AMJ season. When looking at the seasons selected for the forecast, results differ vastly between indices and areas. Initially, it can be stated that NINO4, as well as AMO were the two indices with the fewest overall gridpoints selected, only covering small groups of used seasons for the northern and coastal parts in the case of NINO4 and eastern coastal parts, as well as southern central part for AMO respectively. Both of these indices indicate a prominent selection of the ASO, as well as the AMJ season. The indices NINO12, PDO, AMM and DMI show a greater contribution to forecasts made with a diverse selection of seasons. In NINO12 and DMI, the AMJ and DJF seasons are featured most prominently, whereas PDO shows a diverse mix of selected seasons. AMM shows the most frequent use of the SON and DJF season. Next, the indices NPI, NAO, PNA and AO show the overall fullest selection coverage. The areas covered bear resemblance, aligned in the northeast coastal area, eastern to central southern areas, and dispersed central coverage. NPI, and AO show significant use of the ASO season, with AOs large areas of different seasons used stadnind out. PNA and NAO show dispersed use of various seasons.
Overall, the figure emphasizes the high variability of seasonal correlation in terms of spatial distribution and season selection. The complexity of correlation emphasizes the need for a straightforward approach to further assess the application of index based forecasting.

2.4.3 Different sets of hyperparameters

Four different sets of hyperparameters were run in the configured model. The sets differ in number of units in the hidden layer and regularization rate. A summary of different configurations is provided in Table 2.2 below along with the overall mean spatial RMSE and Bias values.

In addition, Figure 2.4 shows boxplots of the RMSE as well as Bias spatial distribution of the different sets of hyperparameters.

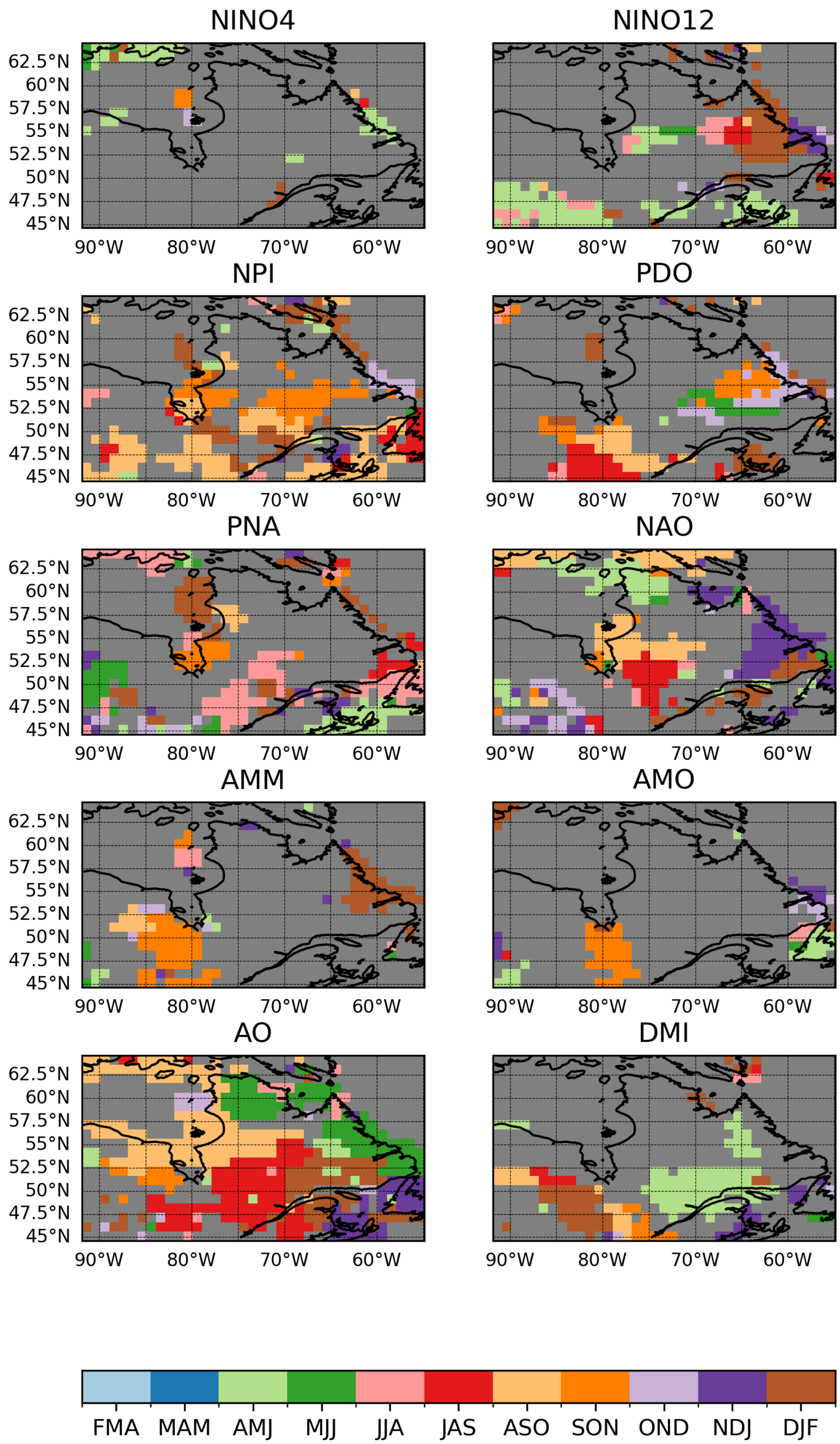
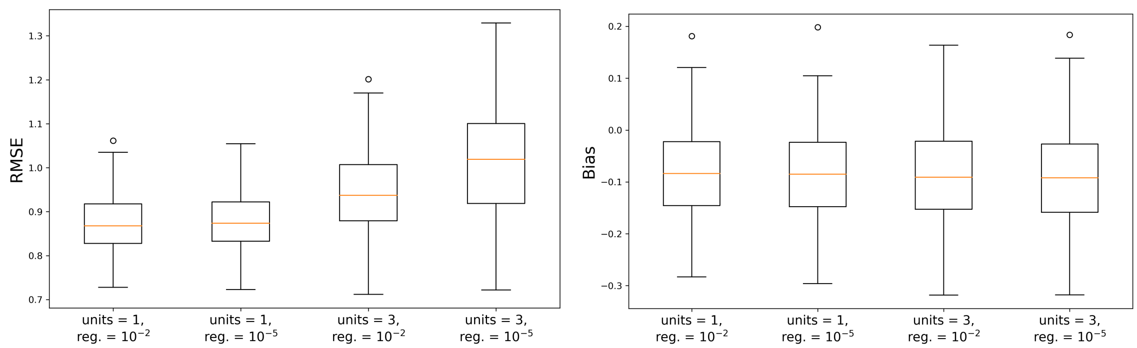


Figure 2.3 : Seasonal correlation used for prediction for each of the utilized indices

Table 2.2 : **Overview of different sets of hyperparameters trained, showing spatial mean values of RMSE and Bias**

Regularization	Units	RMSE	Bias
10^{-2}	1	0.868	-0.084
10^{-5}	1	0.871	-0.084
10^{-2}	3	0.893	-0.087
10^{-5}	3	0.924	-0.088



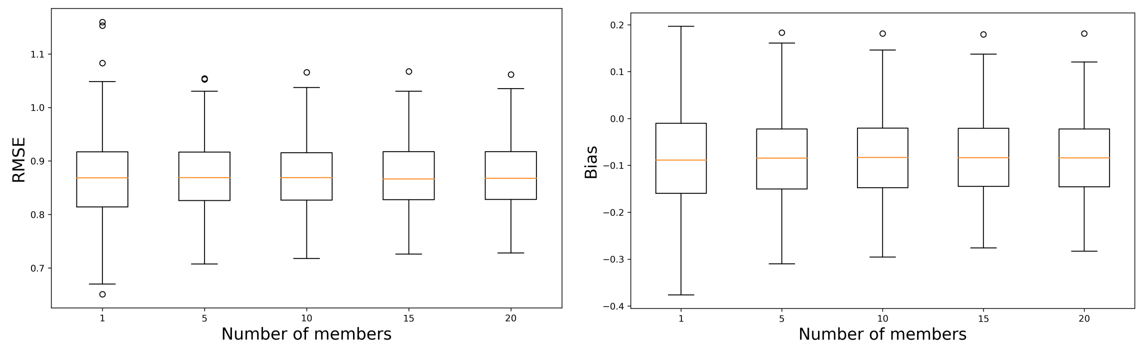
(a) RMSE **(b) Bias**

Figure 2.4 : Comparison of different sets of hyperparameters in terms of 2.4a RMSE and 2.4b Bias

As can be seen in the boxplots, the first set of hyperparameters with a regularization rate of 10^{-2} and 1 unit in the hidden layer performs best with an average RMSE of 0.868. The next hyperparameter configuration, reducing the regularization to 10^{-5} shows no significant change in RMSE. When looking at set 3 with 3 units and a regularization of 10^{-2}, an increase in the RMSE distribution is visible. For the last set, 3 units and a lower regularization of 10^{-5}, the RMSE increases even more, with a significantly wider spread of 25th and 75th percentiles. In terms of bias, it becomes clear that all four sets of hyperparameters feature a slight negative bias. The first two configurations perform very similarly and configurations 3 and 4 show a slight negative shift of the overall distribution. However, when comparing the median bias of the configurations, it becomes clear that the value is very much consistent throughout varying configurations. In summary it can be stated that a basic model consisting of just one unit in the hidden layer and a higher regularization of 10^{-2} in this case outperforms the other three. This indicates that a simple model design is a better choice than a more complex model. Thus, the first combination of hyperparameters is selected for discussion in the following sections. Running additional sets of hyperparameters could prove valuable, for instance, lower convergence tolerance values could also be tested.

2.4.4 Ensemble size

As part of the analysis, the effect of ensemble size on error and bias spatial distribution was examined. The results for the first configuration of hyperparameters are displayed as boxplots in Figure 2.5.



(a) RMSE **(b) Bias**

Figure 2.5 : Comparison of different number of members in terms of 2.5a RMSE and 2.5b Bias

The figure shows the changes in RMSE and Bias distributions depending on the use of 1, 5, 10, 15 or 20 members in the ensemble. As the boxplots indicate, the size of the ensemble had a small impact on the median values of both RMSE and bias using the selected set of hyperparameters. However, there is a clear narrowing of both metrics distributions from a single model to a 5-member ensemble, indicating a decrease in the error and bias spatial variability, whereas there is a narrowing of 25th and 75th percentiles as the number of members is increased. For future applications of this approach, proceeding with an ensemble of 10 members would be appropriate, as there is no significant improvement with a number of members exceeding 10.

2.4.5 Comparison SEAS5

In order to classify the performance of the built EANN model, a comparison was carried out with the state-of-the-art seasonal forecasting model SEAS5. Therefore, the RMSESS was calculated in relation to climatology (see Equation 2.5) for both models and is shown in Figure 2.6.

Initially, it becomes clear that the two models vary in performance depending on the region examined. The relative accuracy of SEAS5 ranges from -10 % to 12 %, whereas the EANN performance ranges from -10 % to 14 %. SEAS5 only reaches a relative accuracy greater than 10% in two points in the very northwest sub-region of the selected area. Overall, both models perform well in the central-west, central-south, and central-east regions. SEAS5 outperforms EANN north and south of Hudson Bay and in the very central area of the selected region. The EANN has better performance in the southernmost region, especially in southern Quebec (between 80W and 65W) and southeastern Newfoundland and Labrador (eastward of 60W) , beating climatology by up to 13% in some regions. Climatological forecasts outperform both models in the southwest, east of Hudson Bay and along the Hudson Strait. In these regions the wind speed is harder to predict due to reduced selection of indices, see Figure 2.3. An additional reason for the lower performance of the EANN model in the northern Quebec area could be high variability of annual wind speeds with an existing and projected overall increase, especially in the area of Nunavik (Jeong & Sushama, 2019).

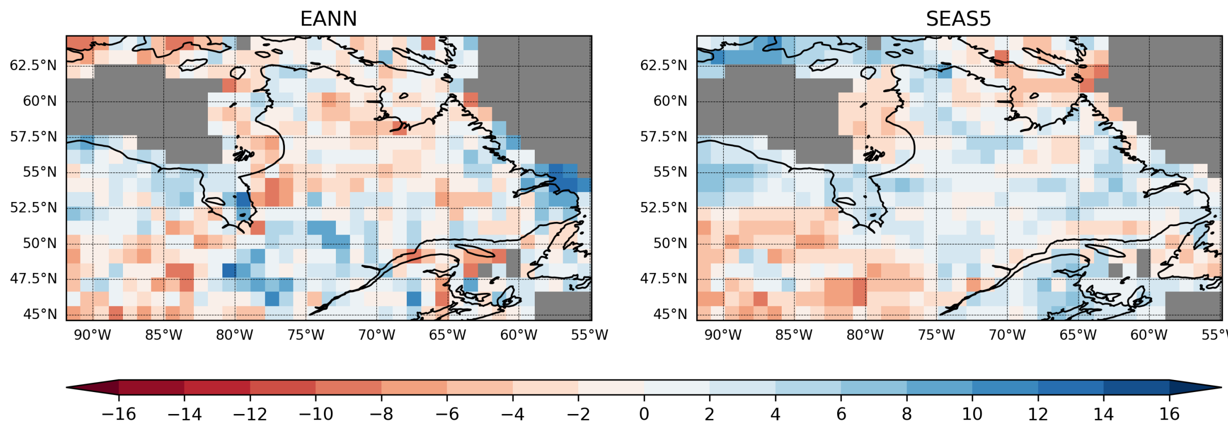


Figure 2.6 : RMSESS in [%] of the built EANN model (left) and SEAS5 (right)

The study demonstrates a correlation between climate indices and wind speeds. In order to assess the influence of the different indices, it is of interest to determine which correlation of which index or combination of indices produces strong results. A well-performing area, such as that east of the twin islands in the southern Hudson Bay at 53°N 79°W, is using the indices NPI, PNA, NAO and AO. The strongest corresponding correlations are above 0.3 for NPI and below -0.2 for NAO. Another strong performing area (2.6) east of the Akami-Uapishk-KakKasuak-Mealy Mountains National Park Reserve, ranging from 54°N 55°W to 56°N 58°W, is forecast based on NINO12, NPI, PDO, PNA, NAO, AMM, AMO and AO (Figure 2.3). When considering the level of correlation, it becomes clear that these indices also show comparably high correlation values, ranging above and below 0.2 and -0.2 respectively (2.2) A contributing factor to the model's good performance is likely the use of NAO winter season indices, as Ouarda & Charron (2021) found NAO to have high impact on wind speed distributions in eastern Canada during winter months. To assess whether the performance of the model is rather dependent on the level of correlation versus the number of predictants, it is interesting to consider the weakly performing area in the very northeast of the case study. Only two indices, NINO4 and PDO, indicate a correlation above or below 0.2 (Figure 2.2), emphasizing influence of number of predictants as well as level of correlation. Moreover, this also provides information of the predictive power of these two indices, as for instance NINO4 shows overall low correlation values for the entire area. The results suggest that the indices NAO, AO and NPI have strong predictive power. The season with the most predictive power for the area selected is SON, followed by the subsequent season ASO, indicating a high predictive power throughout the months August to November.

Another aspect to consider is a varying quality of ERA5 data for the area of the case study. As ERA5's wind data is most accurate for flat and sea-level sites with mainly homogeneous land cover (Gualtieri, 2021), heterogeneous land cover could be a source of potential challenge for the model.

When considering the topography of the terrain, the performance of the EANN model aligns with the findings of Brune *et al.* (2021) of ERA5 performing the strongest in flat terain. Specifically the flat areas in northern Ontario, as well as central and south-western Quebec show a good model performance.

2.5 Summary

Reliable forecasting is essential for the successful implementation of wind energy in the power system and effective scheduling. Modern approaches feature a variety of stand-alone data-based models as well as hybrid models consisting of physical and data-driven models for different time horizons. Seasonal forecasting is often used to get a better overall idea of wind patterns and capacities. The models in place to provide seasonal forecasts are based on atmospheric and land surface conditions as well as numerical weather prediction models. Teleconnection-based forecasts have been conducted for climate factors, such as precipitation and temperature; however, there is an overall lack of evaluation of the wind based on climate oscillations. This study confirms the potential to forecast wind speeds based on correlated climate indices, producing results competitive to state-of-the-art numerical weather predictions. The results reveal that a model with a basic configuration provides competitive results. For example, running an ensemble consisting of a one unit hidden layer and regularization rate of 10^{-2} proved to be the best performance among the different sets of hyperparameters tested, resulting in an overall mean RMSE of 0.868 and bias of -0.084. Based on this study, an approach to forecast wind speeds with an EANN based on climate indices proves profitable for operational purposes, since such a model is easier to implement and computationally cheaper than dynamical models. This study adds to global efforts to enable more accurate forecasting by introducing a novel approach. In continuation, hybridization based on this approach would be of interest. As the model performs better than SEAS5 in certain areas, it is worth exploring as an enhancement strategy. In addition, it would be beneficial to run additional combinations of hyperparameters to grasp extended performance capabilities of the model.

3 GENERAL DISCUSSION AND CONCLUSION

3.1 Additional discussion

The paper written for the thesis presents the research carried out on a climate oscillation-based wind speed forecasting approach using an EANN. In addition to the discussion section of the preceding paper, this chapter shows two additional plots and discusses the results in a broader context.

3.1.1 Spatial results of different hyperparameters

To add to the numerical and box-plotted results (Table 2.2 and Figure 2.4) presented in the article, the following Figure 3.1 shows the areal distribution of the RMSE values for different sets of hyperparameters.

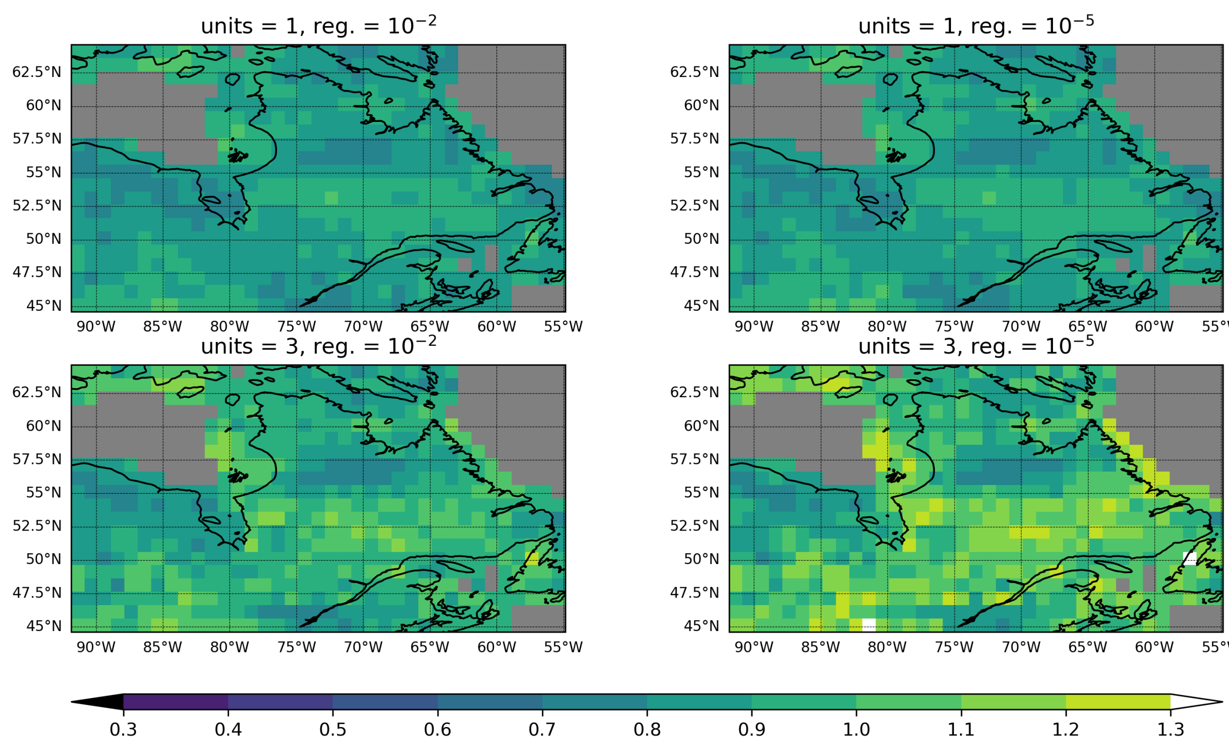


Figure 3.1 : Areal distribution of RMSE for four different sets of hyperparameters

Confirming previous observations, the strong performance of the set of hyperparameters with 1 unit and a regularization of 10^{-2} is the strongest. The RMSE of this first configuration ranges between 0.7 and 1, with a few single-grid points reaching 1.1. The maps show a clear decrease in performance between the different sets of hyperparameters, aligned with the box plot results in Figure 2.5a. The most noticeable deterioration occurs in the coastal areas of western Quebec, along the Hudson Bay, and on the coast of Newfoundland and Labrador. Especially the two configurations with 3 units, displayed in the bottom of the plot, indicate a significantly higher RMSE, varying around 1 and in the case of the fourth configuration even reaching above 1.2 in the mentioned inferior performing areas.

Overall, the distribution of the RMSE aligns with performance in terms of skill score, as depicted compared to SEAS in Figure 2.6. There are a few notable exceptions, such as the area spreading from 68°W to 75°W at around 56°N. Whilst this area has a constant comparably low RMSE of 0.6-0.7 throughout the four sets of hyperparameters, it performs rather mediocre in terms of comparison with climatology.

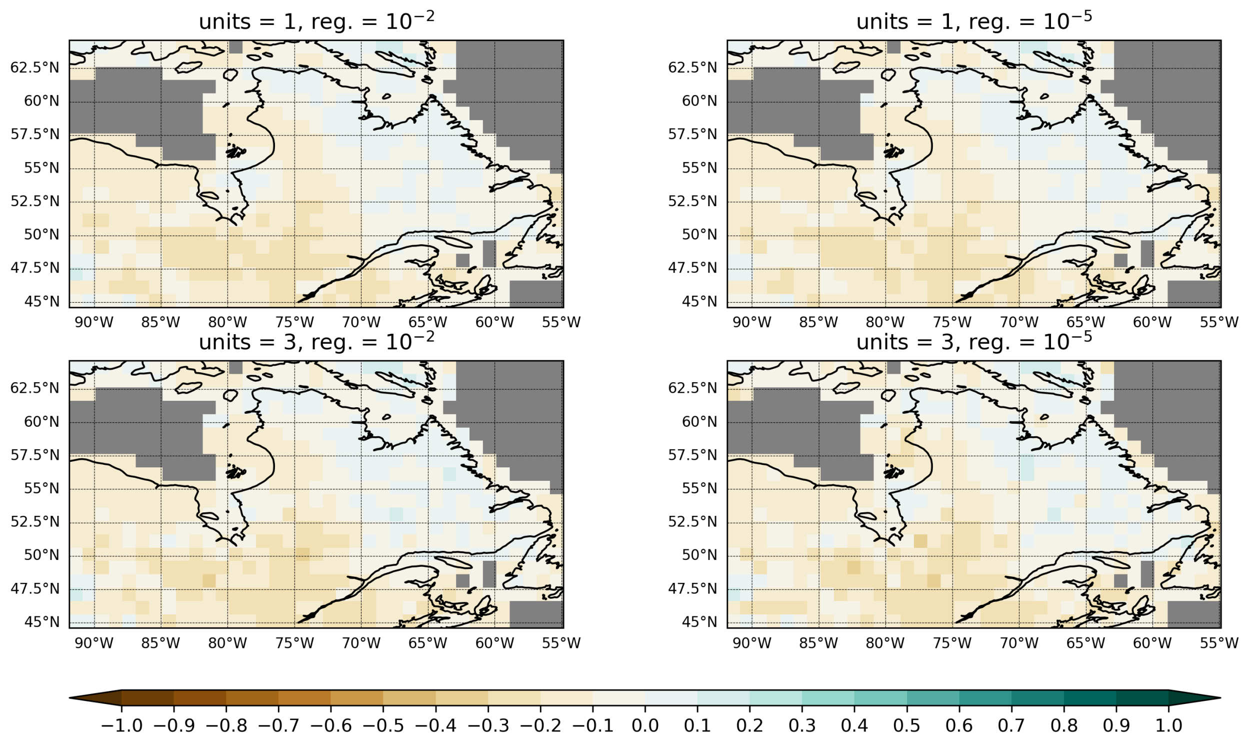


Figure 3.2 : Areal distribution of bias for four different sets of hyperparameters

Figure 3.2 shows a steady bias spatial distribution throughout the different sets of hyperparameters. There is a significantly lower bias in the northeast of Quebec and Newfoundland and Labrador, as well as in the covered section of Nunavut, ranging around 0. In comparison, the rest of Quebec and Ontario have an overall slight negative bias, ranging between -0.1 and -0.3.

3.1.2 General discussion

In general, this modeling approach provides valuable information on the use of seasonal climate indices for forecasting wind speed. In terms of parameters selected for the model, the suggested use of at least 10 members in an ensemble by Shu & Burn (2004) was confirmed. There is no significant improvement noticeable for a number of members exceeding 10. The model is subject to a slight bias which is constant throughout the different sets of hyperparameters.

As for the signal strength and season selection of the considered climate indices, it became clear that the number and intensity of correlation of the indices considered tend to lead to better performance of an area.

Overall, the performance of the built model shows promising results, beating climatalogy in several areas, as well as SEAS5 in the southern part of Quebec and the coastal area of Newfoundland and Labrador (Figure 2.6). Additional comparisons with more wind forecasting models could be considered to further assess the performance of the model.

One of the biggest challenges of the model is the potentially variable quality of ERA5 data for different areas, depending on factors such as data availability, but also influencing factors such as terrain flatness and land cover.

3.2 Conclusion

This research aimed to develop and evaluate the performance of an EANN for wind speed forecasting in eastern Canada based on climate oscillation indices and ERA5 reanalysis data.
Analysis of the used climate indices confirms high correlations with the indices AO, NAO and NPI. The most selected seasons, for which the prediction for the AMJ season was based on, are ASO and SON. The training of a 20-member ensemble resulted, in the case of the best performing set of hyperparameters, in an overall RMSE of 0.868 and a bias of -0.084. Thus, the proposed model results in low-bias estimates and comparable accuary to state-of-the-art dynamical model wind speed seasonal forecasting. The model performs best in terms of skill score in the western part of the examined area, northern Ontario, as well as central to southern Quebec region and coastal area of Newfoundland and Labrador. Moreover, comparison with SEAS5 and the built EANN model shows that such models present distinguished forecasting characteristics, indicating a potentially beneficial combination of both approaches.
The analysis performed illustrates the benefit of a higher regularization rate of 10^{-2} over a lower 10^{-5} for the proposed model. A single unit in the hidden layer also resulted in a stronger performance than the alternatively tested 3. Further research is of interest in determining the fit of the model in terms of hyperparameter selection. For instance, running the model with a smaller convergence tolerance would provide further insight if a better fitting could be achieved.

BIBLIOGRAPHY

Abraham A (2005). Artificial neural networks. *Handbook of measuring system design*. doi:10.1002/0471497398.mm421.

Ahmed SD, Al-Ismail FS, Shafiullah M, Al-Sulaiman FA & El-Amin IM (2020). Grid integration challenges of wind energy: A review. *Ieee Access*, 8:10857–10878. doi:10.1109/ACCESS.2020.2964896.

Alessandrini S, Sperati S & Pinson P (2013). A comparison between the ecmwf and cosmo ensemble prediction systems applied to short-term wind power forecasting on real data. *Applied energy*, 107:271–280. doi:10.1016/j.apenergy.2013.02.041.

Alexiadis M, Dokopoulos P & Sahsamanoglou H (1999). Wind speed and power forecasting based on spatial correlation models. *IEEE Transactions on Energy Conversion*, 14(3):836–842. doi:10.1109/60.790962.

An S, Wang C & Mechoso C (2020). *Teleconnections in the atmosphere*. Cambridge University Press, pages 54–88. doi:10.1017/9781108610995.003.

Azad HB, Mekhilef S & Ganapathy VG (2014). Long-term wind speed forecasting and general pattern recognition using neural networks. *IEEE Transactions on Sustainable Energy*, 5(2):546–553. doi:10.1109/TSTE.2014.2300150.

Bentsen LØ, Warakagoda ND, Stenbro R & Engelstad P (2023). Spatio-temporal wind speed forecasting using graph networks and novel transformer architectures. *Applied Energy*, 333: 120565. doi:10.1016/j.apenergy.2022.120565.

Breiman L (1996). Bagging predictors. *Machine learning*, 24:123–140. doi:10.1007/BF00058655.

Brune S, Keller JD & Wahl S (2021). Evaluation of wind speed estimates in reanalyses for wind energy applications. *Advances in Science and Research*, 18:115–126. doi:10.5194/asr-18-115-202.

Candy B, English SJ & Keogh SJ (2009). A comparison of the impact of quikscat and windsat wind vector products on met office analyses and forecasts. *IEEE transactions on geoscience and remote sensing*, 47(6):1632–1640. doi:10.1109/TGRS.2008.2009993.

CanREA (2023). *Canadian Renewable Energy Association: Wind Energy*. https://renewablesassociation.ca/wind-energy/.

Chandran A, Basha G & Ouarda T (2016). Influence of climate oscillations on temperature and precipitation over the united arab emirates. *International Journal of Climatology*, 36(1):225–235.

Corotis RB, Sigl AB & Cohen MP (1977). Variance analysis of wind characteristics for energy conversion. *Journal of Applied Meteorology and Climatology*, 16(11):1149–1157. doi:10.1175/1520-0450(1977)016<1149:VAOWCF>2.0.CO;2.

Costa A, Crespo A, Navarro J, Lizcano G, Madsen H & Feitosa E (2008). A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews*, 12(6):1725–1744. doi:10.1016/j.rser.2007.01.015.

Dave VS & Dutta K (2014). Neural network based models for software effort estimation: a review. *Artificial Intelligence Review*, 42(2):295–307. doi:10.1007/s10462-012-9339-x.

Dehghani-Sanij AR, Al-Haq A, Bastian J, Luehr G, Nathwani J, Dusseault MB & Leonenko Y (2022). Assessment of current developments and future prospects of wind energy in canada. *Sustainable Energy Technologies and Assessments*, 50:101819. doi:10.1016/J.SETA.2021.101819.

Doblas-Reyes FJ, García-Serrano J, Lienert F, Biescas AP & Rodrigues LR (2013). Seasonal climate predictability and forecasting: status and prospects. *Wiley Interdisciplinary Reviews: Climate Change*, 4(4):245–268. doi:10.1002/wcc.217.

Fung K, Scheffler R & Stolpe J (1981). Wind energy-a utility perspective. *IEEE Transactions on Power Apparatus and Systems*, (3):1176–1182. doi:10.1109/TPAS.1981.316586.

Giani P, Tagle F, Genton MG, Castruccio S & Crippa P (2020). Closing the gap between wind energy targets and implementation for emerging countries. *Applied Energy*, 269:115085. doi:10.1016/J.APENERGY.2020.115085.

Giebel G, Brownsword R, Kariniotakis G, Denhard M & Draxl C (2011). *The state-of-the-art in short-term prediction of wind power: A literature overview*. doi:10.11581/DTU:00000017.

Giebel G, Landberg L, Badger J & Sattler K (2003). *Using ensemble forecasting for wind power*.

González-Sopeña J, Pakrashi V & Ghosh B (2021). An overview of performance evaluation metrics for short-term statistical wind power forecasting. *Renewable and Sustainable Energy Reviews*, 138:110515. doi:10.1016/j.rser.2020.110515.

Gualtieri G (2021). Reliability of era5 reanalysis data for wind resource assessment: a comparison against tall towers. *Energies*, 14(14):4169. doi:10.3390/en14144169.

Hanifi S, Liu X, Lin Z & Lotfian S (2020). A critical review of wind power forecasting methods—past, present and future. *Energies*, 13(15):3764. doi:10.3390/en13153764.

Hersbach H, Bell B, Berrisford P, Biavati G, Horányi A, Muñoz Sabater J, Nicolas J, Peubey C, Radu R, Rozum I, Schepers D, Simmons, A.and Soci C, Dee D & Thépaut JN (2023). Era5 hourly data on single levels from 1940 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. doi:10.24381/cds.adbb2d47.

Huang S, Yan C & Qu Y (2023). Deep learning model-transformer based wind power forecasting approach. *Frontiers in Energy Research*, 10:1055683. doi:10.3389/fenrg.2022.1055683.

IEA (2022). World energy outlook 2022. *Paris, France: International Energy Agency (IEA)*.

Irena I (2019). Future of wind: Deployment, investment, technology, grid integration and socio-economic aspects. *Abu Dhabii*.

Jeong DI & Sushama L (2019). Projected changes to mean and extreme surface wind speeds for north america based on regional climate model simulations. *Atmosphere*, 10(9):497. doi:10.3390/atmos10090497.

Johnson SJ, Stockdale TN, Ferranti L, Balmaseda MA, Molteni F, Magnusson L, Tietsche S, Decremer D, Weisheimer A, Balsamo G *et al.* (2019). Seas5: the new ecmwf seasonal forecast system. *Geoscientific Model Development*, 12(3):1087–1117. doi:10.5194/gmd-12-1087-2019.

Johnson ZF, Chikamoto Y, Wang SYS, McPhaden MJ & Mochizuki T (2020). Pacific decadal oscillation remotely forced by the equatorial pacific and the atlantic oceans. *Climate Dynamics*, 55:789–811. doi:10.1007/s00382-020-05295-2.

Kai L & Salamon P (1990). Neural network ensembles. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 12. doi:10.1109/34.58871.

Kapica J, Canales FA & Jurasz J (2021). Global atlas of solar and wind resources temporal complementarity. *Energy Conversion and Management*, 246:114692. doi:10.1016/j.enconman.2021.114692.

Khazaei S, Ehsan M, Soleymani S & Mohammadnezhad-Shourkaei H (2022). A high-accuracy hybrid method for short-term wind power forecasting. *Energy*, 238:122020. doi:10.1016/j.energy.2021.122020.

Kim Y & Hur J (2020). An ensemble forecasting model of wind power outputs based on improved statistical approaches. *Energies*, 13(5):1071. doi:10.3390/en13051071.

Kriegeskorte N & Golan T (2019). Neural network models and deep learning. *Current Biology*, 29(7):R231–R236. doi:10.1016/j.cub.2019.02.034.

Lam R, Sanchez-Gonzalez A, Willson M, Wirnsberger P, Fortunato M, Alet F, Ravuri S, Ewalds T, Eaton-Rosen Z, Hu W, Merose A, Hoyer S, Holland G, Vinyals O, Stott J, Pritzel A, Mohamed S & Battaglia P (2023). Learning skillful medium-range global weather forecasting. *Science*, 382:1416–1421. doi:10.1126/science.adi2336.

Leung LR, Hamlet AF, Lettenmaier DP & Kumar A (1999). Simulations of the enso hydroclimate signals in the pacific northwest columbia river basin. *Bulletin of the American Meteorological Society*, 80(11):2313–2330. doi:10.1175/1520-0477(1999)080<2313:SOTEHS>2.0.CO;2.

Li H (2022). Short-term wind power prediction via spatial temporal analysis and deep residual networks. *Frontiers in Energy Research*, 10:920407. doi:10.3389/fenrg.2022.920407.

Li S, Wu L, Yang Y, Geng T, Cai W, Gan B, Chen Z, Jing Z, Wang G & Ma X (2020). The pacific decadal oscillation less predictable under greenhouse warming. *Nature Climate Change*, 10(1): 30–34. doi:10.1038/s41558-019-0663-x.

Lipu MH, Miah MS, Hannan M, Hussain A, Sarker MR, Ayob A, Saad MHM & Mahmud MS (2021). Artificial intelligence based hybrid forecasting approaches for wind power generation: Progress, challenges and prospects. *IEEE Access*, 9:102460–102489. doi:10.1109/ACCESS.2021.3097102.

Liu Z & Alexander M (2007). Atmospheric bridge, oceanic tunnel, and global climatic teleconnections. *Reviews of Geophysics*, 45(2). doi:10.1029/2005RG000172.

McPhaden MJ, Santoso A & Cai W (2020). Introduction to el niño southern oscillation in a changing climate. *El Niño Southern Oscillation in a changing climate*, pages 1–19. doi:10.1002/9781119548164.ch1.

Mekanik F, Imteaz M, Gato-Trinidad S & Elmahdi A (2013). Multiple regression and artificial neural network for long-term rainfall forecasting using large scale climate modes. *Journal of Hydrology*, 503:11–21. doi:10.1016/j.jhydrol.2013.08.035.

Miller WT, Glanz FH & Kraft LG (1990). Cmac: An associative neural network alternative to back-propagation. *Proceedings of the IEEE*, 78(10):1561–1567. doi:10.1109/5.58338.

Naizghi MS & Ouarda TB (2017). Teleconnections and analysis of long-term wind speed variability in the uae. *International Journal of Climatology*, 37:230–248. doi:10.1002/joc.4700.

NOAA (n.d.). *Climate Indices: Monthly Atmospheric and Ocean Time Series*. https://psl.noaa.gov/data/climateindices/list/.

Ouarda TB & Charron C (2021). Non-stationary statistical modelling of wind speed: A case study in eastern canada. *Energy Conversion and Management*, 236. doi:10.1016/j.enconman.2021.114028.

Ouarda TB, Charron C & St-Hilaire A (2020). Uncertainty of stationary and nonstationary models for rainfall frequency analysis. *International Journal of Climatology*, 40:2373–2392. doi:10.1002/joc.6339.

Ouyang T, Zha X, Qin L, He Y & Tang Z (2019). Prediction of wind power ramp events based on residual correction. *Renewable energy*, 136:781–792. doi:10.1016/j.renene.2019.01.049.

Palmer TN (2002). The economic value of ensemble forecasts as a tool for risk assessment: From days to decades. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 128(581):747–774. doi:10.1256/003590002321042289.

Piccinini G (2008). Some neural networks compute, others don't. *Neural networks*, 21(2-3):311–321. doi:10.1016/j.neunet.2007.12.010.

Pinheiro E & Ouarda TB (2023). Short-lead seasonal precipitation forecast in northeastern brazil using an ensemble of artificial neural networks. *Scientific Reports*, 13(1):20429. doi:10.1038/s41598-023-47841-y.

Potter CW & Negnevitsky M (2006). Very short-term wind forecasting for tasmanian power generation. *IEEE Transactions on power systems*, 21(2):965–972. doi:10.1109/TPWRS.2006.873421.

Pryor SC, Barthelmie RJ, Bukovsky MS, Leung LR & Sakaguchi K (2020). Climate change impacts on wind power generation. *Nature Reviews Earth & Environment*, 1(12):627–643. doi:10.1038/s43017-020-0101-7.

Rediske G, Burin H, Rigo P, Rosa C, Michels L & Siluk J (2021). Wind power plant site selection: A systematic review. *Renewable and Sustainable Energy Reviews*, 148:111293. doi:10.1016/j.rser.2021.111293.

Rosenblatt F (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386. doi:10.1037/h0042519.

Sarkar MR, Anavatti SG, Dam T, Pratama M & Al Kindhi B (2023). Enhancing wind power forecast precision via multi-head attention transformer: An investigation on single-step and multi-step forecasting. *2023 International Joint Conference on Neural Networks (IJCNN)*, IEEE, pages 1–8. doi:10.1109/IJCNN54540.2023.10191444.

Schepen A, Wang Q & Robertson D (2012). Evidence for using lagged climate indices to forecast australian seasonal rainfall. *Journal of Climate*, 25(4):1230–1246. doi:10.1175/JCLI-D-11-00156.1.

Sharkey AJ (2012). *Combining artificial neural nets: ensemble and modular multi-net systems*. Springer Science & Business Media. doi:10.1007/978-1-4471-0793-4.

Sheridan SC (2003). North american weather-type frequency and teleconnection indices. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 23(1):27–45. doi:10.1002/joc.863.

Shu C & Burn DH (2004). Artificial neural network ensembles and their application in pooled flood frequency analysis. *Water Resources Research*, 40(9). doi:10.1029/2003WR002816.

Shu C & Ouarda TB (2007). Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space. *Water Resources Research*, 43(7). doi:10.1029/2006WR005142.

Singh PK, Singh N & Negi R (2019). Wind power forecasting using hybrid arima-ann technique. *Ambient Communications and Computer Systems: RACCCS-2018*, Springer, pages 209–220. doi:10.1007/978-981-13-5934-7_19.

Soman SS, Zareipour H, Malik O & Mandal P (2010). A review of wind power and wind speed forecasting methods with different time horizons. *North American power symposium 2010*, IEEE, pages 1–8. doi:10.1109/NAPS.2010.5619586.

StatisticsCanada (2024). *Electric power generation, monthly generation by type of electricity*. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=2510001501.

Tascikaraoglu A & Uzunoglu M (2014). A review of combined approaches for prediction of short-term wind speed and power. *Renewable and Sustainable Energy Reviews*, 34:243–254. doi:10.1016/j.rser.2014.03.033.

Ti Z, Deng XW & Zhang M (2021). Artificial neural networks based wake model for power prediction of wind farm. *Renewable Energy*, 172:618–631. doi:10.1016/j.renene.2021.03.030.

Torralba V, Doblas-Reyes FJ, MacLeod D, Christel I & Davis M (2017). Seasonal climate prediction: a new source of information for the management of wind energy resources. *Journal of Applied Meteorology and Climatology*, 56(5):1231–1247. doi:10.1175/JAMC-D-16-0204.1.

Troccoli A, Harrison M, Anderson DL & Mason SJ (2008). *Seasonal climate: forecasting and managing risk*. volume 82. Springer Science & Business Media. doi:10.1007/978-1-4020-6992-5.

Vincent LA, Zhang X, Brown R, Feng Y, Mekis E, Milewska E, Wan H & Wang X (2015). Observed trends in canada's climate and influence of low-frequency variability modes. *Journal of Climate*, 28(11):4545–4560. doi:10.1175/JCLI-D-14-00697.1.

Wang Y, Zou R, Liu F, Zhang L & Liu Q (2021). A review of wind speed and wind power forecasting with deep neural networks. *Applied Energy*, 304:117766. doi:10.1016/j.apenergy.2021.117766.

Watson S, Landberg L & Halliday J (1994). Application of wind speed forecasting to the integration of wind energy into a large scale power system. *IEE Proceedings-Generation, Transmission and Distribution*, 141(4):357–362. doi:10.1049/ip-gtd:19941215.

Weisheimer A, Befort DJ, MacLeod D, Palmer T, O'Reilly C & Strømmen K (2020). Seasonal forecasts of the twentieth century. *Bulletin of the American Meteorological Society*, 101(8):E1413–E1426. doi:10.1049/ip-gtd:19941215.

Wilks DS (2011). *Statistical methods in the atmospheric sciences*. volume 100. Academic press.

WMO (2020). *Guidance on operational practices for objective seasonal forecasting*.

Woldesellasse H, Marpu PR & Ouarda TB (2020). Long-term forecasting of wind speed in the uae using nonlinear canonical correlation analysis (nlcca). *Arabian Journal of Geosciences*, 13:1–9.

Wu H, Meng K, Fan D, Zhang Z & Liu Q (2022). Multistep short-term wind speed forecasting using transformer. *Energy*, 261. doi:10.1016/j.energy.2022.125231.

Yan J & Ouyang T (2019). Advanced wind power prediction based on data-driven error correction. *Energy conversion and management*, 180:302–311. doi:10.1016/j.enconman.2018.10.108.

Yegnanarayana B (2009). *Artificial neural networks*. PHI Learning Pvt. Ltd.

Zhang Y, Wang J & Wang X (2014). Review on probabilistic forecasting of wind power generation. *Renewable and Sustainable Energy Reviews*, 32:255–270. doi:10.1016/j.rser.2014.01.033.

Zhao C & Brissette F (2022). Impacts of large-scale oscillations on climate variability over north america. *Climatic Change*, 173(1-2):4. doi:10.1007/s10584-022-03383-2.

Zou J, Han Y & So SS (2009). Overview of artificial neural networks. *Artificial neural networks: methods and applications*, pages 14–22. doi:10.1007/978-1-60327-101-1_2.