

Titre: L'apprentissage profond au service de la prédiction du productible
Title: solaire

Auteur: Saad Benslimane
Author:

Date: 2023

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Benslimane, S. (2023). L'apprentissage profond au service de la prédiction du productible solaire [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/55220/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/55220/>
PolyPublie URL:

Directeurs de recherche: Hanane Dagdougui, & Loubna Benabbou
Advisors:

Programme: Mathématiques appliquées
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

L'apprentissage profond au service de la prédiction du productible solaire

SAAD BENSLIMANE

Département de mathématiques et de génie industriel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Mathématiques appliquées

Août 2023

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

L'apprentissage profond au service de la prédiction du productible solaire

présenté par **Saad BENSLIMANE**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Luc ADJENGUE, président

Hanane DAGDOUGUI, membre et directrice de recherche

Loubna BENABBOU, membre et codirectrice de recherche

Arnaud ZINFLOU, membre

DÉDICACE

*À mes précieux parents,
piliers de force et de sagesse,
et à mes grands frères,
sources d'inspiration infinie. . .*

REMERCIEMENTS

Je tiens à exprimer ma plus profonde gratitude à Madame Hanane Dagdougui, ma directrice de recherche au sein du département de mathématiques et génie industriel, pour son encadrement, le soutien, qu'elle m'a apporté et pour m'avoir permis de réaliser cette maîtrise à Polytechnique Montréal. Sa connaissance du sujet et ses nombreux conseils m'ont aidé à découvrir et m'épanouir dans un domaine nouveau qu'est la recherche.

Je remercie également madame Loubna Benabbou, ma codirectrice de recherche au sein de l'unité départementale des sciences de la gestion à l'Université du Québec à Rimouski (UQAR), dont la précieuse expertise et les conseils avisés ont grandement contribué à l'accomplissement de ce travail.

Je souhaite également remercier chaleureusement les partenaires de ce projet à savoir l'Institut Québécois d'Intelligence Artificielle (Mila), l'Agence Marocaine pour l'Énergie Durable (Masen), l'Université du Québec à Rimouski (UQAR), et l'École Mohammadia d'Ingénieurs (EMI) pour leur soutien et leur engagement continu. Leurs contributions et leur engagement à promouvoir la recherche ont créé un environnement propice à l'apprentissage et à la découverte.

Je souhaite également remercier Monsieur Luc Adjengue président de jury et Monsieur Arnaud Zinflou membre de jury, d'avoir accepté de lire mon mémoire et d'y proposer les corrections et les améliorations nécessaires.

Finalement, je tiens à exprimer ma profonde gratitude au gouvernement du Québec pour avoir généreusement financé ce mémoire de maîtrise. Leur soutien financier a été essentiel pour bien mener cette recherche et contribue à l'avancement des connaissances dans ce domaine.

Cette thèse n'aurait pas vu le jour sans votre soutien, votre expertise et votre dévouement. Merci de m'avoir permis de repousser les limites de mes connaissances et de m'aider à réaliser cette étape importante de ma carrière académique.

RÉSUMÉ

Pour maintenir un équilibre entre la réduction des émissions de gaz à effet de serre et la croissance de la demande de l'électricité, il est très important de disposer d'une source d'énergie plus propre et plus durable comme l'énergie solaire Photovoltaïque (PV). Le sujet est d'importance capitale dans le contexte de la transition énergétique, qui place les ressources des énergies renouvelables au centre des efforts mondiaux de lutte contre le changement climatique. Cependant, à cause de leur nature intermittente, une prédiction précise du productible de l'énergie PV représente un défi majeur qui doit être surmonté pour faciliter l'intégration de l'énergie PV dans le réseau électrique.

Dans le premier chapitre « Introduction », l'importance de la prédiction de la production d'énergie solaire a été mise en évidence. En raison de l'intermittence inhérente à la production d'énergie solaire, causée par des facteurs tels que les conditions météorologiques changeantes et le cycle diurne, une prédiction précise est cruciale pour permettre une bonne intégration et gestion de l'énergie solaire PV dans le réseau électrique. Ce chapitre a également évoqué l'évolution rapide des technologies de l'énergie solaire et leur impact croissant sur la production d'énergie, soulignant encore plus la nécessité d'améliorer les méthodes de prévision.

Le deuxième chapitre « Revue de littérature », a examiné un large éventail de techniques existantes pour la prévision de la production d'énergie solaire. Cela comprenait les méthodes numériques, les méthodes statistiques traditionnelles, telles que les techniques de régression, ainsi que des approches basées sur l'intelligence artificielle et l'apprentissage automatique, telles que les réseaux de neurones artificiels, les forêts aléatoires et les machines à vecteurs de support. Cependant, la revue a également identifié une nouvelle génération de modèles d'apprentissage profond, en particulier ceux basés sur les transformateurs, qui ont démontré une capacité prometteuse à capturer des dépendances complexes et à long terme dans les données de séries temporelles, ce qui est crucial pour la prédiction de la production d'énergie solaire PV.

Le troisième chapitre « Article 1 », a présenté une étude empirique de l'application de divers modèles d'apprentissage profond pour la prédiction de la production d'énergie solaire PV. Cette étude a mis en évidence le modèle Pyraformer, qui a surpassé tous les autres modèles testés, y compris les réseaux neuronaux récurrents tels que LSTM et GRU, en termes de précision des prédictions. Cependant, l'étude a également révélé que tous les modèles testés avaient des difficultés à prévoir avec précision la production d'énergie solaire lorsque les

conditions météorologiques changeaient de manière significative d'un jour à l'autre. Cela a souligné le besoin d'intégrer des informations plus détaillées et à haute résolution sur les conditions météorologiques futures dans ces modèles pour améliorer leur performance.

En somme, ce mémoire a souligné l'importance des prévisions précises pour la production d'énergie solaire, a examiné l'état actuel des techniques de prédiction, et a identifié les transformateurs comme une approche prometteuse pour l'avenir. Elle a également mis en évidence des domaines pour de futures recherches, en particulier l'intégration de prévisions météorologiques à haute résolution et l'amélioration de la gestion des changements abrupts de conditions météorologiques.

ABSTRACT

To maintain a balance between the reduction of greenhouse gas emissions and the growth in demand for electricity, it is very important to have cleaner, more sustainable energy sources such as solar power. The subject is of paramount importance in the context of the energy transition, which places renewable energy resources at the center of global efforts to combat climate change. However, due to their intermittent nature, accurate prediction of the producibility of PV energy represents a major challenge that needs to be overcome to facilitate the integration of PV energy into the power grid.

In the first "Introduction" chapter, the importance of predicting solar energy production was highlighted. Due to the inherent intermittency of solar power generation, caused by factors such as changing weather conditions and the diurnal cycle, accurate prediction is crucial to enable proper integration and management of solar power PV into the power grid. This chapter also discussed the rapid evolution of solar energy technologies and their growing impact on power generation, further underlining the need for high-performance forecasting methods.

The second chapter, "Literature review", examined a wide range of existing techniques for forecasting solar energy production. This included numerical methods, traditional statistical methods such as regression techniques, as well as approaches based on artificial intelligence and machine learning, such as artificial neural networks, random forests and support vector machines. However, the review also identified a new generation of deep learning models, in particular those based on transformers, which have demonstrated a promising ability to capture complex and long-term dependencies in time series data, which is crucial for the prediction of solar power generation PV.

The third chapter « Paper 1 », presented an empirical study of the application of various deep learning models for the prediction of solar energy production PV. This study highlighted the Pyraformer model, which outperformed all other models tested, including recurrent neural networks such as LSTM and GRU, in terms of prediction accuracy. However, the study also revealed that all the models tested had difficulty in accurately predicting solar energy production when weather conditions changed significantly from day to day. This highlighted the need to incorporate more detailed, high-resolution information on future weather conditions into these models to improve their performance.

In sum, this thesis highlighted the importance of accurate forecasts for solar power generation, examined the current state of prediction techniques, and identified transformers as a

promising approach for the future. It also highlighted areas for future research, in particular the integration of high-resolution weather forecasts and improved management of abrupt changes in weather conditions.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	vii
TABLE DES MATIÈRES	ix
LISTE DES TABLEAUX	xii
LISTE DES FIGURES	xiii
LISTE DES SIGLES ET ABRÉVIATIONS	xv
CHAPITRE 1 INTRODUCTION	1
1.1 Mise en contexte	1
1.2 Problématique	2
1.3 Objectif Général	3
1.4 Structure du mémoire	3
CHAPITRE 2 REVUE DE LA LITTÉRATURE :	4
2.1 Modélisation de la cellule PV	4
2.1.1 Modèle idéal de cellule PV	5
2.1.2 Modèle de cellule PV à une diode	6
2.1.3 Modèle de cellule PV à deux diodes	7
2.1.4 Autres modèles de cellules PV	8
2.2 Modèles de Prévision Météorologique Numérique (NWP)	8
2.2.1 Rôle dans la prédiction du productible de l'énergie PV	8
2.2.2 Types de Modèles NWP	9
2.3 Approches Statistiques	10
2.3.1 Modèle Auto Régressif (AR)	10
2.3.2 Modèle Moyenne Mobile (MA)	10
2.3.3 Modèle Moyenne Mobile Auto-Régressif (ARMA)	11
2.3.4 Modèle Moyenne Mobile Intégrée Auto-Régressif (ARIMA)	11

2.4	Apprentissage automatique	12
2.4.1	Généralités	12
2.4.2	Modèles utilisés	12
2.5	Revue de littérature : prédiction du productible PV	19
CHAPITRE 3 MÉTHODOLOGIE DE RECHERCHE		21
3.1	Collecte et prétraitement des données	21
3.1.1	Description de données	21
3.1.2	Prétraitement des Données	22
3.2	Optimisation des Hyperparamètres	23
3.3	Modèles d'Apprentissage Profond Utilisés	24
3.3.1	CrossFormer	24
3.3.2	Non-Stationary Transformer (NonST)	25
3.3.3	Autoformer	26
3.3.4	Reformer	29
3.3.5	ETSformer	29
3.3.6	FEDformer	30
3.3.7	Informer	31
3.3.8	Pyraformer	32
3.3.9	Frequency improved Legendre Memory (FiLM)	33
3.3.10	LightTS	34
3.3.11	PatchTST	35
3.3.12	Multi-scale Isometric Convolution Network (MICN)	36
3.3.13	DLinear	38
CHAPITRE 4 ARTICLE 1: ONE DAY-AHEAD PV OUTPUT POWER FORECAS- TING: REVEALING THE POWER OF DEEP LEARNING MODELS BEYOND LSTM.		40
4.1	Introduction	40
4.2	Methodology	44
4.2.1	Direct approach	44
4.2.2	Evaluation indices	45
4.2.3	Benchmark Models	46
4.3	Data	49
4.3.1	Data normalization	50
4.3.2	Data splitting	51
4.3.3	Feature selection	51

4.4	Results and discussions	52
4.5	Limitation and challenges	55
4.6	Conclusion	55
4.7	Acknowledgment	56
CHAPITRE 5 CONCLUSION ET RECOMMANDATIONS		58
5.1	Synthèse des travaux	58
5.2	Limitations de la solution	59
5.3	Améliorations futures	59
RÉFÉRENCES		61

LISTE DES TABLEAUX

Table 4.1	Detailed information of the PV system from DKASC	50
Table 4.2	Results of the dry season for the region of Yulara in Australia	52
Table 4.3	Results of the wet season for the region of Yulara in Australia	53

LISTE DES FIGURES

Figure 2.1	Structure physique de la cellule PV [1]	5
Figure 2.2	Modèle idéal de la cellule PV [2]	6
Figure 2.3	Modèle de la cellule PV à une diode [2]	6
Figure 2.4	Modèle de la cellule PV à deux diodes [2]	7
Figure 2.5	Représentation d'un neurone [3]	13
Figure 2.6	Architecture d'un modèle ANN [3]	14
Figure 2.7	Représentation du module répétitif d'un modèle RNN [3]	14
Figure 2.8	Représentation du module répétitif d'un modèle LSTM [4]	15
Figure 2.9	Représentation du module répétitif d'un modèle GRU [4]	16
Figure 2.10	(à gauche) Représentation de l'Attention au produit de points mis à l'échelle. (à droite) Représentation de l'attention multi-têtes. [5] . . .	17
Figure 3.1	Structure interne de deux blocs CrossFormer consécutifs. SDA et LDA apparaissent alternativement dans différents blocs, LN représente la couche de normalisation. [6]	24
Figure 3.2	Non-Stationary Transformers. La stationnarisation des séries est adoptée comme une enveloppe sur le modèle de base pour normaliser chaque série entrante et dé-normaliser la sortie. L'attention déstationnaire remplace le mécanisme original d'attention pour approximer l'attention apprise à partir de séries non stationnaires, qui redimensionne les poids de dépendance temporelle actuels avec des facteurs dé-stationnaires appris τ , Δ . [7] . . .	26
Figure 3.3	Architecture de l'Autoformateur. L'encodeur élimine la partie tendance-cyclique à long terme par des blocs de décomposition en série (blocs bleus) et se concentre sur la modélisation des tendances saisonnières. Le décodeur accumule progressivement la partie tendance extraite des variables cachées. Les informations saisonnières passées du l'encodeur sont utilisées par l'auto-corrélation encodeur-décodeur (bloc vert central dans le décodeur). [8] . .	27
Figure 3.4	Auto-corrélation (à gauche) et agrégation des délais (à droite). La transformée de Fourier rapide est utilisée pour calculer l'autocorrélation $\mathcal{R}(\tau)$, qui reflète les similitudes de délai. Les sous-processus similaires sont ensuite regroupés dans le même index sur la base du délai τ sélectionné et agrégés par $\mathcal{R}(\tau)$. [8]	28
Figure 3.5	Architecture du modèle ETSformer. [9]	30

Figure 3.6	Structure du FEDformer. Le FEDformer se compose de N encodeurs et de M décodeurs. Le Frequency Enhanced Block (FEB, blocs verts) et le Frequency Enhanced Attention (FEA, blocs rouges) sont utilisés pour effectuer l'apprentissage de la représentation dans le domaine fréquentiel. Le FEB ou le FEA ont deux sous-versions (FEB-f et FEB-w ou FEA-f et FEA-w), où "-f" signifie l'utilisation de la base de Fourier et "-w" l'utilisation de la base d'ondelettes. Les blocs de décomposition par mélange d'experts (MOEDecomp, blocs jaunes) sont utilisés pour extraire les tendances saisonnières des données d'entrée. [10]	31
Figure 3.7	Architecture du modèle Informer. [11]	32
Figure 3.8	Architecture du modèle Pyraformer. [12]	32
Figure 3.9	Structure globale du FiLM (modèle de mémoire de Legendre amélioré en fréquence). LPU : Legendre Projection Unit. LPU_R : récupération inverse de la projection de Legendre. FEL : couche à fréquence améliorée. RevIn : bloc de normalisation des données. Les données d'entrée sont d'abord normalisées, puis projetées dans l'espace des polynômes de Legendre (mémoire C de la LPU). La mémoire LPU C est traitée avec FEL et génère la mémoire C' de sortie. Enfin, C' est reconstruit et normalisé à la série de sortie avec LPUR. Une structure multicale est employée pour traiter l'entrée de longueur $\{T, 2T, \dots, nT\}$. [13]	34
Figure 3.10	Vue d'ensemble de LightTS. Dans la partie I, le modèle capture les dépendances à court et long terme et extrait les caractéristiques de chaque série temporelle. Dans la partie II, le modèle apprend les interdépendances entre les différentes séries temporelles et fait des prédictions. [14]	35
Figure 3.11	Architecture globale du modèle PatchTST. [15]	36
Figure 3.12	Architecture globale du modèle MICN. [16]	37
Figure 4.1	Flow chart of the machine learning process for PV output power using the direct approach.	44
Figure 4.2	Heatmap representing Pearson correlation between the features for the region of Yulara in Australia.	51
Figure 4.3	Training time for the deep learning models used.	54
Figure 4.4	PV forecasting for the best five deep learning models used.	57

LISTE DES SIGLES ET ABRÉVIATIONS

PV	Photovoltaïque
IEA	Agence Internationale de l'Énergie
IRENA	Agence Internationale pour les Énergies Renouvelables
NREL	Laboratoire National de l'Énergie Renouvelable
NWP	Prévision Météorologique Numérique
GFS	Système de Prévision Globale
WRF	Recherche et Prévision Météorologiques
ECMWF	Centre Européen pour les Prévisions Météorologiques à Moyen Terme
HRRR	Réactualisation Rapide en Haute Résolution
RAP	Rafrâichissement Rapide
UM	Unifié du Met Office Britannique
AR	Auto-Régressif
MA	Moyenne Mobile
ARMA	Moyenne Mobile Auto-Régressif
ARIMA	Moyenne Mobile Intégrée Auto-Régressif
MLP	Perceptron Multicouche
RNN	Réseaux de Neurones Récurrents
LSTM	Réseaux à Mémoire à Long Terme
GRU	Unités Récurrentes à Portes
CNN	Réseaux Neuronaux Convolutionnel

CHAPITRE 1 INTRODUCTION

1.1 Mise en contexte

Les émissions de gaz à effet de serre ont un impact significatif sur le système climatique terrestre. Ces émissions retiennent la chaleur dans l’atmosphère, provoquant ainsi une hausse des températures mondiales et toute une série d’effets sur l’économie, l’environnement et la société [17–20]. Les sources traditionnelles d’énergie, telles que le charbon, le pétrole et le gaz naturel, contribuent fortement aux émissions de gaz à effet de serre. Selon l’Agence Internationale de l’Énergie (IEA), le charbon représentait 44 % des émissions mondiales de dioxyde de carbone liées à l’énergie en 2020 [21], alors que le pétrole et le gaz naturel contribuent également de manière significative aux émissions de gaz à effet de serre, représentant respectivement 34 % et 20 % des émissions [21].

D’autre part, la demande d’électricité a augmenté rapidement au cours des dernières décennies, sous l’effet de plusieurs facteurs, notamment l’augmentation de la population mondiale, la croissance économique, le changement de mode de vie (l’utilisation des appareils électroniques, l’électrification des transports, etc.) [22, 23]. Avec une demande d’électricité continuant d’augmenter, il est nécessaire de disposer des sources de production d’électricité plus propres et plus durables, telles que les sources d’énergie renouvelables comme l’énergie solaire Photovoltaïque (PV).

Plusieurs raisons justifient ce choix. Premièrement, la technologie PV est économiquement avantageuse. Le coût de la technologie PV a considérablement diminué au fil du temps grâce aux progrès technologiques, ce qui en fait une option compétitive pour la production d’électricité [24]. Ainsi, l’adoption de la technologie PV a créé de nombreuses opportunités d’emploi dans le secteur des énergies renouvelables. Selon un rapport de l’Agence Internationale pour les Énergies Renouvelables (IRENA), ce secteur employait 11,5 millions de personnes dans le monde en 2019, contre 10,3 millions en 2017 [25].

Deuxièmement, la technologie PV présente plusieurs avantages pour l’environnement. Selon le Laboratoire National de l’Énergie Renouvelable (NREL), un système PV résidentiel typique peut compenser 3 à 4 tonnes d’émissions de dioxyde de carbone par an [26]. Ainsi, l’énergie PV peut contribuer à réduire la pollution de l’air, qui peut avoir des effets néfastes sur la santé humaine et l’environnement [27]. D’autre part, la technologie PV ne nécessite pas de l’eau pour produire de l’électricité, contrairement aux méthodes traditionnelles de production d’électricité qui requièrent d’importantes quantités d’eau pour produire de l’électricité [26].

Finalement, la technologie PV permet d'accéder à une électricité fiable dans les zones où l'accès au réseau est limité ou inexistant. Cela peut avoir un impact significatif sur la qualité de vie des individus et des communautés [28]. La technologie PV peut contribuer à promouvoir l'équité énergétique en donnant accès à l'électricité aux communautés mal desservies. Ce qui contribue à réduire la pauvreté énergétique et à améliorer la situation sociale et économique de ces communautés [25, 29, 30].

La prédiction du productible d'un système PV peut offrir plusieurs avantages, notamment dans l'amélioration de la gestion de l'énergie. Cela permet aux opérateurs du système de planifier et d'optimiser la consommation d'énergie, ce qui peut contribuer à réduire les coûts énergétiques et à améliorer l'efficacité globale du système [31–33].

1.2 Problématique

La production d'énergie PV est influencée par plusieurs facteurs météorologiques tels que l'irradiance solaire, la température, la vitesse du vent, etc. Ces facteurs peuvent varier considérablement d'un jour à l'autre, voire d'une heure à l'autre, ce qui rend les prédictions plus difficiles en fonction de l'horizon de prédiction [34].

De plus, bien que les modèles numériques puissent être utilisés pour la prédiction de ces facteurs, leurs prédictions sont souvent imprécises dans des cas où il s'agit de prévoir la couverture nuageuse locale [35]. D'autre part, les modèles numériques nécessitent de grandes capacités de calcul, ce qui n'est pas toujours pratique pour les applications à temps réel [36].

D'autre part, bien que des modèles statistiques sont utilisés pour prédire le productible de l'énergie PV, ces modèles ont des difficultés à saisir des caractéristiques non linéaires et complexes et présument souvent que les tendances historiques se poursuivent dans le futur, ce qui n'est pas toujours le cas [37]. Cependant, les modèles d'apprentissage automatique ne nécessitent pas de connaissance a priori des données, et captent des relations plus complexes et non linéaires, afin d'améliorer la précision des prévisions [38].

La prédiction du productible de l'énergie solaire PV est un défi réel. Pour cette raison, l'utilisation des modèles d'apprentissage profond, qui sont capables de capturer des dépendances non linéaires complexes, s'est avérée prometteuse. Cependant, la majorité des études antérieures se concentrent sur des approches qui ont été dépassées par des nouvelles architectures d'apprentissage profond.

1.3 Objectif Général

Ce projet de maîtrise fait partie d'un projet de partenariat entre Polytechnique Montréal, l'Université du Québec à Rimouski (UQAR), l'Institut Québécois d'Intelligence Artificielle (Mila), l'Institut de Valorisation des Données (IVADO), l'École Mohammadia d'Ingénieurs (EMI) et l'Agence Marocaine pour l'Énergie Durable (Masen) dans le cadre du Programme de Coopération Climatique Internationale (PCCI) du ministère de l'Environnement et de la Lutte contre les Changements Climatiques (MELCC) au Québec. L'objectif global de ce projet est de développer un système d'aide à la décision pour prédire la ressource et le productible de l'énergie solaire, destiné aux centrales de production d'énergies renouvelables. L'objectif de ce projet de maîtrise est de proposer un modèle prédictif du productible de l'énergie solaire PV avec une précision suffisante pour des utilisations pratiques. Pour cela, nous allons développer une étude comparative de méthodes d'apprentissage profond récentes pour la prédiction du productible de l'énergie PV.

1.4 Structure du mémoire

Le présent mémoire est composé de quatre chapitres. Tout d'abord cette introduction qui traite l'importance de l'énergie solaire PV et son impact sur l'environnement, la société et l'économie. Ensuite, le deuxième chapitre propose une revue de la littérature abordant les différentes approches utilisées pour la prédiction du productible de l'énergie solaire PV. Le troisième chapitre présente une étude comparative des différents modèles d'apprentissage profond pour la prédiction du productible solaire. Finalement, le quatrième chapitre conclut ce mémoire en exposant les contributions, les limitations, et les améliorations potentielles.

CHAPITRE 2 REVUE DE LA LITTÉRATURE :

2.1 Modélisation de la cellule PV

La cellule PV constitue le noyau de la technologie de l'énergie solaire, qui convertit la lumière du soleil en électricité. Tout d'abord, la lumière du soleil est absorbée par un matériau absorbant dans la cellule PV, ce qui provoque l'excitation des électrons vers un état d'énergie plus élevé. Les électrons excités se déplacent à travers le matériau et génèrent un courant électrique.

Il existe plusieurs types de cellules PV en fonction du matériau utilisé, par exemple on trouve : 1) les cellules PV monocristallines qui sont fabriquées à partir d'une structure cristalline de silicium, elles offrent une haute efficacité et une longue durée de vie mais leur processus de production est coûteux [39], 2) Les cellules PV polycristallines qui sont composées de plusieurs fragments de silicium cristallin fondus, ces cellules offrent une efficacité et durée de vie moins que les cellules PV monocristallines mais restent moins chères [39], 3) Les cellules PV à couche minces, ce type de cellule est composée de plusieurs couches minces de matériau comme le silicium amorphe (a-Si), le sélénure de cuivre et d'indium gallium (CIGS) ou aussi le tellure de cadmium (CdTe) [39].

La simulation des cellules PV est une méthode souvent utilisée pour optimiser la conception des systèmes PV. Elle permet de modéliser le comportement électrique de la cellule et d'analyser les performances du système dans différentes conditions de fonctionnement. Dans ce chapitre, nous explorerons l'importance de la simulation des cellules PV comme des circuits, en mettant l'accent sur les avantages qu'elle procure. Parmi les avantages de la simulation des cellules PV :

- Aider à mieux comprendre la performance des cellules PV dans différentes conditions météorologiques (température, irradiance, etc.) [40].
- Optimiser la conception des systèmes PV, en testant les différentes configurations et technologies avant la mise en application réelle [41].
- Aider à prédire la production d'énergie des systèmes PV, afin d'améliorer la gestion de l'énergie et optimiser la consommation de l'énergie [40].
- Identifier les problèmes potentiels et procéder aux ajustements nécessaires pour augmenter l'efficacité du système PV, ainsi réduire les coûts énergétiques [42].

Les approches les plus utilisées pour la simulation des cellules PV sont les modèles numériques qui modélisent les cellules PV comme des circuits [43–47]. Dans la littérature, différents modèles de circuits pour les cellules PV ont été proposés. Un modèle de circuit approprié

est celui qui émule avec précision le comportement électrique de la cellule PV physique et qui n'est pas trop complexe. Par conséquent, un compromis approprié entre la précision et la simplicité doit être établi lors de la sélection du modèle de circuit de la cellule PV. Dans cette section, on va présenter les modèles de circuits de cellules PV les plus utilisés dans la littérature.

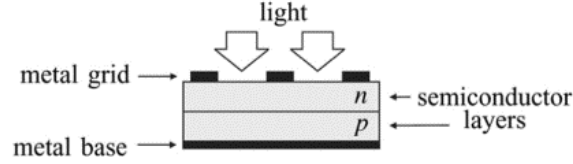


FIGURE 2.1 Structure physique de la cellule PV [1]

2.1.1 Modèle idéal de cellule PV

La cellule PV comporte deux couches de semi-conducteurs dopés différemment (Figure 2.1), dont la jonction P-N est exposée à la lumière. En l'absence de rayonnement solaire, la cellule PV fonctionne comme une simple diode à jonction P-N dont la courbe I-V est donnée par l'équation 2.1 de Shockley [48].

$$I_D = I_S \left[\exp \left(\frac{qV_D}{aKT} \right) - 1 \right] \quad (2.1)$$

Où :

- I_D est le courant de la diode en Ampère (A).
- V_D est la tension de la diode en Volte (V).
- I_S est le courant de saturation de la diode (A).
- a est le facteur d'idéalité de la diode.
- q est la valeur absolue de la charge électrique d'un électron en Coulomb (C).
- T est la température en Kelvin (K).
- K est la constante de Boltzman Joule/K.

La jonction P-N absorbe les photons de la lumière incidente et génère des paires électron-trou (ou porteurs) sous l'effet de l'irradiation [48]. Cela produit une différence de potentiel à travers la jonction, permettant aux porteurs de charge de circuler dans le circuit externe [48]. Ce phénomène est connu sous le nom d'effet PV, et le courant électrique qui en résulte est connu sous le nom de photo-courant I_{PV} [48]. L'ajout de I_{PV} à l'équation de Shockley fournit une description simple d'une cellule PV éclairée avec une source de courant et une diode à jonction P-N [48]. Le circuit dans la Figure 2.2 représente le modèle idéal de cellule PV. Le

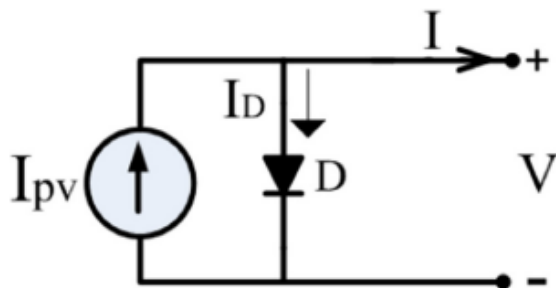


FIGURE 2.2 Modèle idéal de la cellule PV [2]

Le courant I est la superposition des courants I_{PV} et I_D . Le modèle idéal de cellule PV est exprimé par l'équation 2.2.

$$I = I_{PV} - I_S \left[\exp\left(\frac{qV_D}{aKT}\right) - 1 \right] \quad (2.2)$$

Il faut noter que le modèle de la cellule solaire PV ne prend pas en compte les effets de la résistance interne, ce qui ne permet pas d'établir une relation précise entre le courant et la tension de la cellule.

2.1.2 Modèle de cellule PV à une diode

Une résistance en série R_S est ajoutée au modèle d'une cellule idéale afin d'améliorer ses résultats. Bien que ce modèle soit simple, il révèle des failles lorsqu'il est exposé à des températures variables. Cette variante comprend désormais une résistance shunt R_P .

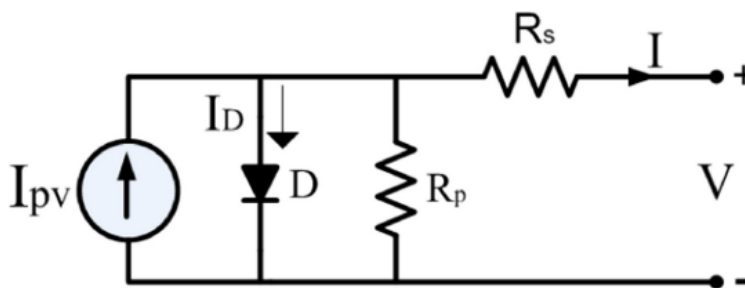


FIGURE 2.3 Modèle de la cellule PV à une diode [2]

Le modèle à une diode (Figure 2.3), se compose d'une source de courant et d'une diode avec des résistances connectées en série et en parallèle. La résistance en série représente la résistance (perte ohmique) offerte au flux de courant en raison du contact ohmique (contact

métal-semiconducteur) et des concentrations d'impuretés avec la profondeur de la jonction. Le courant de fuite à travers la jonction représente la résistance shunt R_P connectée en parallèle. Ce courant est exprimé par l'équation 2.3.

$$I = I_{PV} - I_D - \frac{V_D}{R_P} \quad (2.3)$$

Avec V_D représente la tension de la diode, exprimée par l'équation 2.4.

$$V_D = V + IR_S \quad (2.4)$$

Après modification de l'équation 2.2, le modèle de la cellule PV à une diode est exprimé par l'équation 2.5.

$$I = I_{PV} - I_S \left[\exp \left(\frac{q(V + IR_S)}{aKT} \right) - 1 \right] - \frac{V + IR_S}{R_P} \quad (2.5)$$

Où :

- I est le courant de sortie.
- V est la tension d'entrée.

Il faut noter la résistance en série affecte la tension de sortie, alors que la résistance en parallèle réduit le courant disponible [49].

2.1.3 Modèle de cellule PV à deux diodes

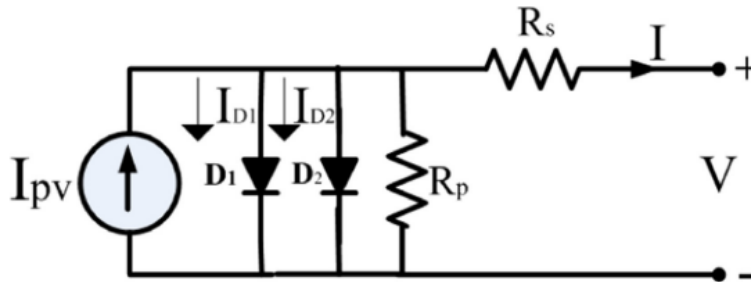


FIGURE 2.4 Modèle de la cellule PV à deux diodes [2]

Bien que le modèle à une diode présente certaines améliorations par rapport au modèle idéal, sa précision reste discutable. Afin de surmonter ce problème, un modèle de cellule PV à deux diodes est proposé. Comme on remarque sur la Figure 2.4, deux diodes sont connectées en parallèle à la source de courant. Comme dans le cas du modèle à une diode, la composante du courant I_{D1} traversant la première diode D_1 est identique à I_d . Le courant de recombinaison

I_{D2} dans la région de charge d'espace passe par la deuxième diode D_2 . Cela suggère que les courants de saturation d'une cellule solaire PV sont influencés par les deux termes de Shockley. La résistance série R_S et la résistance shunt R_P sont identiques au modèle à une diode. Contrairement au modèle à une diode, le modèle de la cellule PV à deux diodes est plus précis dans les niveaux d'ensoleillement faibles. L'équation 2.6 représente le modèle à deux diodes [50].

$$I = I_{PV} - I_{S1} \left[\exp \left(\frac{q(V + IR_S)}{a_1 KT} \right) - 1 \right] - I_{S2} \left[\exp \left(\frac{q(V + IR_S)}{a_2 KT} \right) - 1 \right] - \frac{V + IR_S}{R_P} \quad (2.6)$$

Avec : I_{S1} et I_{S2} sont les courants de saturation des deux diodes de la cellule PV.

2.1.4 Autres modèles de cellules PV

Différents modèles de cellules PV ont été introduits dans la littérature. Par exemple, le modèle à diode unique avec capacité [51], le modèle à trois diodes [52], les modèles à deux diodes modifiées [53], le modèle de dérive-diffusion [54] et le modèle à diodes multidimensionnelles [55]. Cependant, en raison de la complexité de ces modèles, leur application à la simulation des cellules PV est très limitée.

2.2 Modèles de Prévision Météorologique Numérique (NWP)

Alors que les conditions météorologiques ont un impact considérable sur le fonctionnement des systèmes PV, les modèles Prévision Météorologique Numérique (NWP) sont essentiels pour la prédiction du productible de l'énergie PV. Sur la base des conditions météorologiques courantes, les modèles NWP utilisent des simulations mathématiques pour prédire les conditions météorologiques. Les données utilisées pour initialiser ces modèles proviennent de diverses sources, telles que les satellites, les ballons météorologiques, les systèmes radars et les stations météorologiques [56]. Pour prédire l'état futur des conditions météorologiques, les modèles NWP sont complexes et nécessitent une grande capacité de calcul [57].

2.2.1 Rôle dans la prédiction du productible de l'énergie PV

L'énergie PV est produite à partir de la conversion directe de l'irradiance solaire en électricité à l'aide de cellules PV [58]. Le niveau d'irradiation solaire reçu par un panneau solaire influence directement la quantité d'électricité qu'il peut produire. Plus le niveau d'irradiation est élevé, le panneau solaire peut produire plus d'électricité. L'irradiation solaire peut varier en fonction de facteurs tels que l'heure de la journée, la saison, la situation géographique

et les conditions météorologiques courantes [59]. Par exemple, la couverture nuageuse peut réduire considérablement le niveau d'irradiation solaire [59]. Il est donc essentiel de prévoir avec précision l'irradiation solaire pour prédire la puissance de sortie du système PV [60].

D'autre part, le fonctionnement des cellules PV est également influencé par la température. L'efficacité d'une cellule PV, autrement dit sa capacité à convertir l'énergie solaire en énergie électrique, diminue à mesure que sa température augmente [61]. Par conséquent, dans les climats plus chauds ou pendant les heures les plus chaudes de la journée, un système PV peut produire moins d'énergie que ce qui est attendu du niveau d'irradiation solaire [62]. Les prévisions de température peuvent donc contribuer à améliorer la précision des prédictions du productible de l'énergie PV.

Les modèles NWP peuvent fournir des prédictions à la fois de l'irradiance solaire, de la température, ainsi que des paramètres météorologiques qui peuvent être utilisés comme données d'entrée des autres modèles (par exemple les modèles numériques ou d'autres modèles qu'on va voir par la suite) pour faire la prédiction du productible de l'énergie PV.

2.2.2 Types de Modèles NWP

Dans la littérature, on peut distinguer deux types de modèles NWP : modèle NWP global et modèle NWP régional. Les modèles globaux couvrent l'ensemble du globe, et permettent de prévoir les phénomènes météorologiques à grande échelle. Parmi ceux-ci, on a le modèle Système de Prévision Globale (GFS) qui fournit des prévisions pour l'ensemble du globe jusqu'à 16 jours à l'avance, avec une résolution de 13 km, et se met à jour quatre fois par jour (00 UTC, 06 UTC, 12 UTC et 18 UTC) [63]. Le modèle GFS est souvent utilisé pour des prévisions à long terme et offre une vision sur l'ensemble des conditions météorologiques [64,65]. D'autre part, on a le modèle du Centre Européen pour les Prévisions Météorologiques à Moyen Terme (ECMWF) qui fournit des prévisions des conditions météorologiques allant de 3 à 15 jours, pour l'ensemble du globe [66]. Le modèle ECMWF est connu par sa bonne précision et est considéré comme l'un des meilleurs modèles de prévision numérique [67,68].

Les modèles régionaux couvrent un pays ou une région limitée pour fournir une vision plus localisée sur les conditions météorologiques. Parmi ceux-ci, on a le modèle de Recherche et Prévision Météorologiques (WRF) qui est un modèle à l'échelle méso [69]. Ce modèle offre la possibilité d'être utilisé pour la recherche et la prévision opérationnelle, avec une résolution horizontale de 10 km [69]. Le modèle WRF est considéré comme l'un des modèles NWP les plus utilisés. On trouve aussi le modèle Réactualisation Rapide en Haute Résolution (HRRR), qui couvre une partie de l'Amérique du Nord avec une résolution de 3 km [70]. Le modèle HRRR est actualisé chaque heure, ce qui le rend adapté pour les prévisions à court terme [71].

Bien que les modèles NWP présentent de bons résultats en les combinant avec d'autres modèles prédictifs, ils restent limités dans certains points [71, 72]. D'un côté, ces modèles sont complexes et nécessitent des ressources de calcul de haute performance [73]. D'un autre côté, même les petits changements dans l'atmosphère peuvent avoir un impact significatif sur les prévisions météorologiques [73].

2.3 Approches Statistiques

Les approches statistiques sont des modèles mathématiques qui utilisent l'historique des séries temporelles pour faire des prédictions dans le futur. Contrairement aux modèles NWP, les approches statistiques donnent de bons résultats pour les petits horizons [74]. Dans cette section, nous allons faire le tour sur différents modèles statistiques utilisés pour la prédiction du productible de l'énergie PV.

2.3.1 Modèle Auto Régressif (AR)

Le modèle Auto-Régressif (AR) [75] est utilisé lorsqu'on remarque qu'il y a une corrélation entre la série temporelle et la version décalée d'elle-même. Autrement dit, le modèle AR explique la série temporelle en fonction de ses valeurs passées [75]. Cela est expliqué mathématiquement à l'aide de l'équation 2.7.

$$\forall t : X_t = \sum_{i=1}^p \alpha_i X_{t-i} + \epsilon_t \quad (2.7)$$

Où :

- X_t la série temporelle a expliqué.
- p est l'ordre du processus AR.
- ϵ_t est un bruit blanc.
- $(\alpha_1, \dots, \alpha_p)$ les paramètres du modèle.

2.3.2 Modèle Moyenne Mobile (MA)

Contrairement au modèle AR, le modèle Moyenne Mobile (MA) [75] n'utilise pas les valeurs passées pour exprimer la prédiction future. Le modèle MA utilise plutôt les erreurs des prédictions passées dans un modèle de régression [75]. Cela est expliqué mathématiquement à l'aide de l'équation 2.8.

$$\forall t : X_t = \sum_{i=1}^q \beta_i \epsilon_{t-i} + \epsilon_t \quad (2.8)$$

Où :

- X_t la série temporelle a expliqué.
- p est l'ordre du processus MA.
- ϵ_t est un bruit blanc.
- $(\beta_1, \dots, \beta_q)$ les paramètres du modèle.

Il faut noter que le modèle MA ne s'applique que si la série temporelle est stationnaire.

2.3.3 Modèle Moyenne Mobile Auto-Régressif (ARMA)

Le modèle Moyenne Mobile Auto-Régressif (ARMA) [75] est la combinaison des deux modèles AR et MA. Ce modèle utilise un mélange de termes autorégressifs (valeurs passées) et de termes de moyenne mobile (erreurs passées), ce qui permet de traiter des séries temporelles plus complexes [75]. Cela est expliqué mathématiquement à l'aide de l'équation 2.9.

$$\forall t : X_t = \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{i=1}^q \beta_i \epsilon_{t-i} + \epsilon_t \quad (2.9)$$

Où :

- p et q sont les ordres des processus AR et MA.
- ϵ_t est un bruit blanc.
- $(\alpha_1, \dots, \alpha_p)$ et $(\beta_1, \dots, \beta_q)$ les paramètres du modèle.

Pareil que le modèle MA, le modèle ARMA ne s'applique que si la série temporelle est stationnaire.

2.3.4 Modèle Moyenne Mobile Intégrée Auto-Régressif (ARIMA)

Le modèle Moyenne Mobile Intégrée Auto-Régressif (ARIMA) est considéré comme une généralisation du modèle ARMA qui ne peut s'appliquer que si la série temporelle est stationnaire [75]. Pour cela, le modèle ARIMA intègre le principe de la différenciation qui permet de rendre la série stationnaire [75]. Une fois la série est devenue stationnaire, le modèle ARMA est donc appliqué [75]. Le modèle ARIMA est noté $ARIMA(p, d, q)$ et possède trois paramètres qui sont :

- p est l'ordre du processus AR.
- q est l'ordre du processus MA.
- d est le nombre de différenciations nécessaires pour rendre la série stationnaire.

2.4 Apprentissage automatique

Dans cette section, nous allons explorer des généralités sur l'apprentissage automatique et les différents modèles utilisés dans la prédiction du productible de l'énergie PV. Le but final de cette section, est d'introduire et familiariser le lecteur aux concepts de l'apprentissage automatique que nous allons utiliser dans ce mémoire.

2.4.1 Généralités

L'apprentissage automatique est la branche connexionniste de l'intelligence artificielle qui permet à un système d'apprendre à faire des prédictions ou des classifications à partir des relations qui existent entre les données. Autrement dit, les modèles d'apprentissage automatique peuvent faire des tâches sans être explicitement programmés [76].

Il existe trois grandes familles de l'apprentissage automatique : l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement. L'apprentissage supervisé consiste à ce que le modèle s'entraîne sur des données labellisées, qu'on appelle données d'entraînement, pour apprendre à faire des prédictions ou des classifications sur des nouvelles données qu'on appelle données de test.

De l'autre côté, l'apprentissage non supervisé consiste à ce que le modèle s'entraîne sur des données non labellisées pour trouver des caractéristiques et des relations entre les données d'entrée. Le dernier type est l'apprentissage par renforcement qui consiste à ce que le modèle s'entraîne à se comporter dans un environnement en effectuant certaines actions et en observant les résultats.

2.4.2 Modèles utilisés

Dans le contexte de la prédiction du productible de l'énergie PV, il existe plusieurs modèles d'apprentissage automatique qui ont été utilisés dans les recherches antérieures [77, 78]. Dans cette partie, on va se concentrer sur les modèles d'apprentissage profond, qui est une sous-catégorie de l'apprentissage automatique.

Perceptron Multicouche (MLP)

Le modèle Perceptron Multicouche (MLP) [79] est considéré comme le modèle de base de l'apprentissage profond. Inspiré du système nerveux biologique, le modèle MLP se compose des unités élémentaires appelées neurones artificiels [79]. Comme le montre la Figure 2.5, un neurone est une fonction à n entrées x_1, x_2, \dots, x_n et une sortie y . Cela peut être exprimé à

l'aide de l'équation 2.10.

$$y = h \left(\Theta_0 + \sum_{i=1}^n \Theta_i x_i \right) \quad (2.10)$$

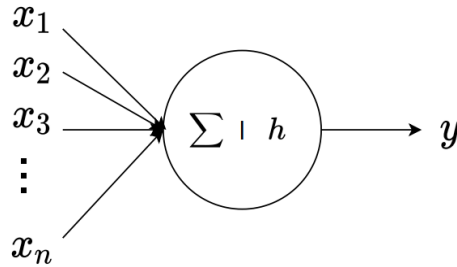


FIGURE 2.5 Représentation d'un neurone [3]

Avec h représente la fonction d'activation [80].

L'architecture du modèle MLP est composée d'une première couche appelée couche d'entrée, une ou plusieurs couches cachées et une dernière couche appelée couche de sortie. Comme le montre la Figure 2.6, chaque couche de la structure du modèle MLP se compose d'une ou plusieurs neurones artificiels. Chaque neurone est connecté avec les neurones de la couche précédente et la couche suivante du modèle. La première couche prend comme entrée les variables x_i des données utilisées, alors que la sortie y_i de la dernière couche représente les prédictions du modèle MLP.

D'après l'équation 2.10, les connexions entre les neurones d'un modèle MLP sont pondérées par $\Theta_i \in \mathbb{R}$ appelés des poids. Cela veut dire, que dans un modèle MLP, l'entrée d'un neurone est la combinaison linéaire pondérée des sorties de tous les neurones de la couche précédente. Certaines études [81, 82] ont démontré que les modèles MLP, avec une seule couche cachée et suffisamment de neurones, peuvent représenter toute fonction continue.

Bien que les modèles MLP ne sont pas adaptés aux problèmes des séries temporelles, il existe d'autres architectures basées sur le MLP comme LightTS [14] et TSMixer [83].

Réseaux de Neurones Récurrents (RNN)

Afin de gérer les données séquentielles comme les séries temporelles, les Réseaux de Neurones Récurrents (RNN) [84] ont été développés. Contrairement au modèle MLP qui considère chaque entrée individuellement, le modèle RNN maintient une forme de mémoire pour faire des prédictions à partir d'une séquence de données d'entrée. Cela signifie que les sorties de chaque étape du réseau sont utilisées comme entrées pour les étapes suivantes [85].

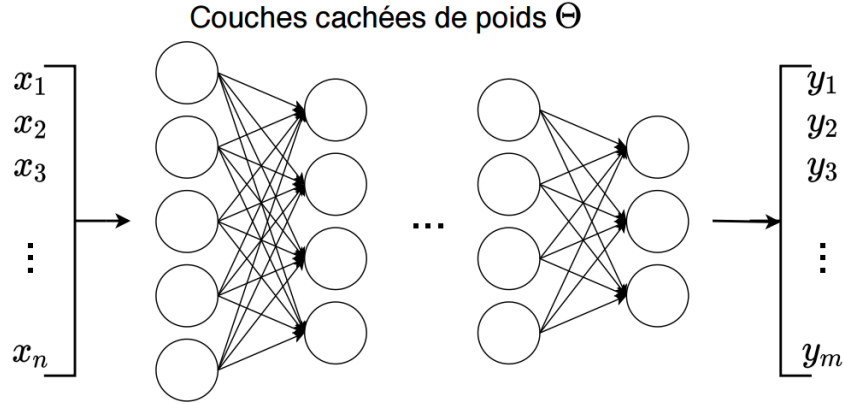


FIGURE 2.6 Architecture d'un modèle ANN [3]

Tout comme le modèle MLP, un modèle RNN est composé d'une couche d'entrée, une ou plusieurs couches cachées et une couche de sortie. Comme on peut voir sur la Figure 2.7, les neurones dans ce cas prennent comme entrée x_t qui représente l'état actuel, à l'aide de h_t qui maintient l'état interne du réseau, ils calculent la sortie y_t qui représente le nouvel état interne du réseau pour l'étape qui suit.

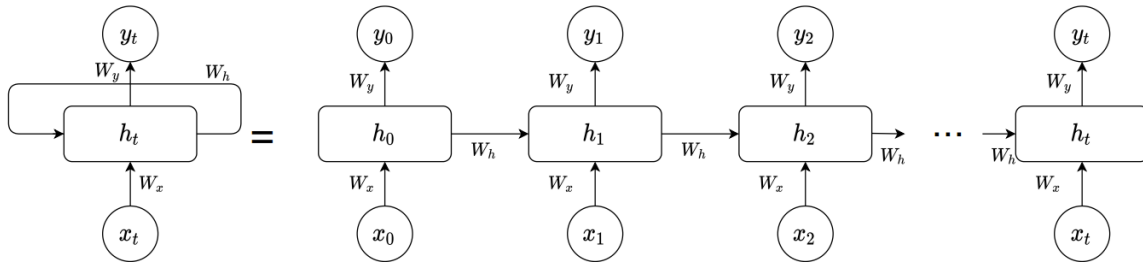


FIGURE 2.7 Représentation du module répétitif d'un modèle RNN [3]

Cela peut être expliqué mathématiquement à l'aide des équations 2.11a et 2.11b.

$$h_{t+1} = \sigma_h (W_x x_{t+1} + W_h h_t + b_h) \quad (2.11a)$$

$$y_{t+1} = \sigma_y (W_y h_{t+1} + b_y) \quad (2.11b)$$

Où :

- σ_x et σ_y sont deux fonctions d'activation.
- b_h et b_y sont les vecteurs de biais.
- W_x, W_h, W_y , sont des matrices de poids partagées sur des étapes.

Malheureusement ces modèles rencontrent aussi quelques problèmes. Premièrement, le problème de l'explosion gradient qui consiste à ce que la valeur du gradient peut augmenter d'une manière exponentielle, ce qui rend l'entraînement difficile [86]. Le deuxième problème est celui de la disparition du gradient, qui consiste à ce que la contribution de l'information diminue avec le temps. Cela empêche le modèle RNN à apprendre des séquences plus longues [86].

Les Réseaux à Mémoire à Long Terme (LSTM) [87], ont été conçus pour lutter contre ces problèmes. Il s'agit d'un type de RNN qui utilise des unités spéciales en plus des unités standard. Sur la Figure 2.8, on peut voir que les unités LSTM comprennent une cellule de mémoire qui peut conserver des informations en mémoire pendant de longues périodes. Un ensemble de portes contrôle le moment où l'information entre dans la mémoire, le moment où elle en sort et le moment où elle est oubliée. Cette architecture leur permet d'apprendre des dépendances à plus long terme, ce qui rend le modèle LSTM adapté à des problèmes de séries temporelles.

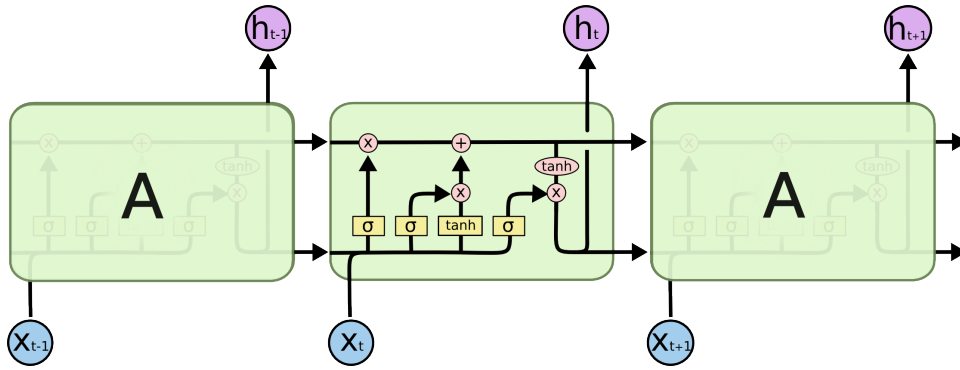


FIGURE 2.8 Représentation du module répétitif d'un modèle LSTM [4]

L'architecture du modèle LSTM peut être exprimée à l'aide des équations 2.12a, 2.12b, 2.12c, 2.12d, 2.12e et 2.12f.

$$f_t = \sigma(x_t U^f + h_{t-1} W^f) \quad (2.12a)$$

$$i_t = \sigma(x_t U^i + h_{t-1} W^i) \quad (2.12b)$$

$$o_t = \sigma(x_t U^o + h_{t-1} W^o) \quad (2.12c)$$

$$\hat{C}_t = \sigma(x_t U^g + h_{t-1} W^g) \quad (2.12d)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \hat{C}_t \quad (2.12e)$$

$$h_t = \tanh(C_t) \odot o_t \quad (2.12f)$$

Où :

- W et U sont les poids du réseau.
- σ c'est la fonction d'activation.
- i_t, f_t, o_t sont respectivement la porte d'entrée, la porte d'oublie et la porte de sortie.

Si les modèles LSTM améliorent considérablement la capacité du réseau à apprendre des dépendances à long terme par rapport aux RNN classiques, ils introduisent également le problème de la complexité de calcul. Chaque cellule LSTM possède trois portes différentes et un état de cellule, ce qui signifie qu'elle a plus de paramètres et nécessite donc plus de ressources de calcul. Cela peut poser un problème lorsque l'on travaille avec de très grands ensembles de données ou lorsque les ressources sont limitées.

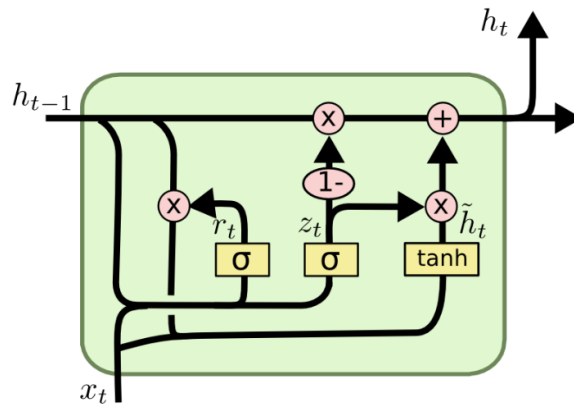


FIGURE 2.9 Représentation du module répétitif d'un modèle GRU [4]

Les Unités Récurrentes à Portes (GRU) [88] ont été introduites pour résoudre ce problème. Comme on peut voir sur la Figure 2.9, une GRU possède une porte de réinitialisation et une porte de mise à jour. La GRU est donc moins complexe sur le plan informatique que le LSTM, puisqu'elle a moins de paramètres. Cette réduction peut conduire à des performances similaires à celles de la LSTM, mais avec des temps d'apprentissage et de prédiction plus rapides.

L'architecture du modèle GRU peut être exprimée à l'aide des équations...

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (2.13a)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (2.13b)$$

$$\hat{h}_t = \tanh(W_h x_t + U_h (r_t \odot h_{t-1})) \quad (2.13c)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \hat{h}_t \quad (2.13d)$$

Où :

- W et U sont les poids du réseau.
- σ est la fonction d'activation.
- z_t, r_t sont respectivement la porte de mise à jour et la porte de réinitialisation.
- \odot représente le produit d'Hadamard.

Les Mécanismes d'attention

Le mécanisme d'attention a été proposé pour améliorer la précision des modèles traitant des données séquentielles, notamment la traduction linguistique et la reconnaissance vocale [5]. Le mécanisme d'attention est essentiellement une méthode qui permet aux modèles de se concentrer sur les parties les plus pertinentes des données d'entrée, ce qui est analogue à la façon dont les humains se concentrent sur des détails spécifiques lorsqu'ils traitent des informations.

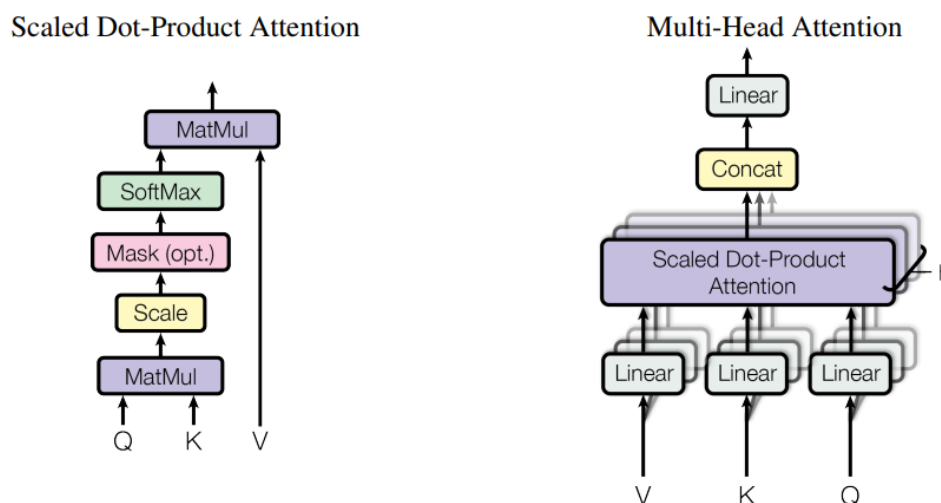


FIGURE 2.10 (à gauche) Représentation de l'Attention au produit de points mis à l'échelle. (à droite) Représentation de l'attention multi-têtes. [5]

Un mécanisme d'attention calcule une somme pondérée des valeurs d'entrée en fonction du contexte, où chaque poids indique l'importance ou la pertinence de la valeur correspondante. Le mécanisme d'attention reçoit trois entrées : la requête (Q), la clé (K) et la valeur (V). Il peut s'agir de vecteurs dérivés des données d'entrée. Dans le contexte d'un modèle de traduction automatique, par exemple, si on souhaite traduire une phrase de l'anglais au français, la requête peut être un vecteur représentant un mot ou une partie de la phrase

française que le modèle génère, tandis que les clés et les valeurs peuvent être des vecteurs représentant des parties de la phrase anglaise qu'on souhaite traduire.

D'après la Figure 2.10 à gauche, pour chaque requête, le mécanisme d'attention calcule un score pour chaque clé. Ce score détermine la pertinence de la valeur associée à la requête. La méthode de calcul du score est appelée produit de points mis à l'échelle. Les scores sont ensuite passés par une couche softmax, qui les transforme en probabilités. Cette opération garantit que la somme de toutes les probabilités est égale à 1. Après, les scores peuvent être interprétés comme la distribution de l'attention du modèle sur les valeurs (des scores plus élevés signifient une attention plus grande).

Enfin, chaque valeur est multipliée par son score softmax correspondant (désormais une probabilité) et les résultats sont additionnés. La somme obtenue est la sortie du mécanisme d'attention, qui est une somme pondérée où les poids sont déterminés par la distribution de l'attention calculée par le modèle. Cela peut être expliqué à l'aide de l'équation 2.14.

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (2.14)$$

Avec d_k est la dimension de la clé K . Le facteur $\frac{1}{\sqrt{d_k}}$ est utilisée pour mettre à l'échelle le produit QK^T .

L'attention multi-têtes représentée sur la Figure 2.10 à droite, est une extension du mécanisme d'attention. Ceci permet d'effectuer le mécanisme d'attention plusieurs fois en parallèle avec différentes transformations linéaires. Cela peut être expliqué à l'aide de l'équation 2.15.

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) W^O \quad (2.15)$$

Avec $\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$ et h représente le nombre de fois que le mécanisme d'attention a été effectué.

Le modèle Transformer [5], représente l'architecture d'apprentissage profond, utilisant le mécanisme d'attention, la plus utilisée pour les problèmes avec des données séquentielles. Contrairement aux LSTM et GRU, les Transformers peuvent capturer des dépendances même pour des données séquentielles plus longues. Ceci fait du transformer un modèle adapté à la prédiction des séries temporelles, notamment la prédiction du productible de l'énergie PV. Cependant, la précision de ce modèle diminue pour les longs horizons de prédiction. Pour surpasser cette limitation, plusieurs variations du modèle transformer ont été proposées. Chacune avec des améliorations pour gérer des données séquentielles plus longues et complexes et améliorer les prédictions du modèle.

2.5 Revue de littérature : prédiction du productible PV

Grâce au progrès de l'apprentissage automatique plusieurs études ont été publiées pour répondre au problème de la prédiction du productible de l'énergie PV. En effet, une étude comparative [89] a été proposée pour tester différents modèles d'apprentissage automatique et profond comme les machines à vecteurs de support (SVM), les forêts aléatoires (RF), RNN, LSTM et GRU. Les modèles ont pris comme entrée l'historique des paramètres météorologiques comme (l'irradiance, température, humidité, etc). Les résultats de cet article ont montré que le RNN performe mieux que les modèles testés avec une erreur RMSE de 3.52 %, alors que RF est le modèle le moins performant avec une erreur RMSE de 3.93 %. Une autre étude comparative a été proposée par [90] pour tester différents modèles d'apprentissage profond. Les auteurs ont comparé le LSTM, GRU, Réseaux Neuronaux Convolutionnel (CNN), et des modèles hybrides comme CNN-LSTM et CNN-GRU pour faire la prédiction de l'énergie PV avec des horizons variant de 1 jusqu'à 60 minutes. Cette fois-ci, les résultats ont montré que le LSTM est le modèle le plus performant parmi les modèles testés.

Afin de savoir si le type de la cellule PV peut influencer la performance du modèle prédictif, [91] ont proposé un modèle MLP pour faire la prédiction de l'énergie PV de trois différents types de cellules : Monocristalline, Polycristalline et cellule à couches minces. Le modèle MLP se composait de 7 couches cachées, et prenait comme données d'entrées l'historique des paramètres météorologiques comme l'irradiance, l'énergie PV, vitesse du vent et l'humidité. Les résultats ont montré que la performance du modèle MLP ne diminue pas beaucoup en changeant le type de la cellule PV (entre 96.5 % et 97.5 % de performance). L'étude [92] a proposé une machine d'apprentissage extrême (ELM), afin de voir l'impact de l'horizon de prédiction sur la performance du modèle. Les auteurs ont pris comme données d'entrée les paramètres météorologiques, afin de prédire le productible PV pour les prochaines 2h, 3h et 4h. les résultats ont montré que la performance du modèle ELM diminue en augmentant l'horizon de prédiction. Pour leur cas d'étude, les erreurs varient, selon la saison de l'année, entre 2.10 % et 6.22 % pour un horizon de 2 heures, entre 2.65 % et 7.43 % pour un horizon de 3 heures, entre 3.01 % et 7.74 % pour un horizon de 4 heures.

Dans [93], un modèle hybride combinant la transformée en ondelettes (WT) et le CNN a été proposé pour prédire le productible PV pour différents horizons. Comme le modèle proposé ne prend en considération que l'historique de l'énergie PV pour faire les prédictions, le fait d'ignorer les différents paramètres météorologiques qui influencent sur le productible PV représente un obstacle pour la méthode proposée. Une autre méthode proposée [94], combinant WT avec les réseaux de neurones à fonction de base radiale (RBFNN). Dans cette étude, les auteurs ont pris comme données d'entrée les valeurs de l'historique de l'irradiance,

productible PV et la température à l'instant t , $t-12$ heures et $t-20$ heures. Cette méthode a permis de diminuer les erreurs en RMSE du modèle RBFNN de 0.45-2.20 kW selon la saison, jusqu'à 0.32-1.57 kW selon la saison.

Plusieurs modèles de l'apprentissage profond pour les séries temporelles n'ont pas encore été testés pour la prédiction du productible de l'énergie PV. Dans le travail qui suit, on va comparer un ensemble de modèles d'apprentissage profond basés sur le MLP, RNN et surtout le Transformer, pour faire la prédiction du productible de l'énergie PV. Cette étude va nous permettre de tester de nouvelles approches et les comparer avec les modèles d'apprentissages profond les plus utilisés dans la littérature comme les LSTM et GRU.

CHAPITRE 3 MÉTHODOLOGIE DE RECHERCHE

Dans ce chapitre, nous allons traiter de manière détaillée la méthodologie de recherche utilisée dans notre étude comparative pour la prédiction du productible PV. Le chapitre est structuré de manière à fournir une explication détaillée de la conception de la recherche, y compris les méthodes utiles pour le prétraitement des données, les architectures d'apprentissage profond utilisées, l'optimisateur utilisées pour trouver la combinaison optimale des hyperparamètres des modèles, ainsi que les métriques d'évaluation adoptées pour évaluer les performances des modèles utilisés.

3.1 Collecte et prétraitement des données

3.1.1 Description de données

L'ensemble des données utilisées pour cette recherche provient du Desert Knowledge Australia Solar Centre (DKASC) [95] situé dans la région de Yulara, en Australie. Cette installation de pointe constitue une ressource inestimable pour les études sur l'énergie solaire, étant donné son emplacement stratégique dans un environnement aride et à forte irradiation, très propice à la production d'énergie PV.

L'ensemble des données du DKASC offre une gamme complète de mesures, y compris :

- l'irradiation horizontale globale (GHI) qui est une mesure de la puissance solaire reçue par unité de surface sur une surface plane et horizontale.
- La température qui est cruciale pour la performance des panneaux solaires, car des températures plus élevées peuvent réduire l'efficacité des cellules photovoltaïques.
- La vitesse du vent qui peut refroidir les panneaux solaires et donc potentiellement augmenter leur efficacité.
- La direction du vent qui peut influencer la quantité de poussière ou de débris qui s'accumule sur les panneaux, affectant ainsi leur performance.
- La pression de l'air qui peut avoir un impact sur la densité de l'air et donc sur la quantité d'énergie solaire dispersée ou absorbée par l'atmosphère.

L'ensemble des données couvre une période de plusieurs années, ce qui se traduit par des milliers de points de données qui offrent des informations sur la variabilité à court et à long terme. Cette étendue temporelle permet un entraînement plus robuste des modèles d'apprentissage profond, car elle capture les tendances saisonnières et les cycles diurnes qui peuvent influencer la production PV.

3.1.2 Prétraitement des Données

Normalisation des Données

La normalisation des données est une étape essentielle pour le prétraitement de données, en particulier les modèles d'apprentissage profond, qui sont très sensibles à la magnitude des entrées. La normalisation des données permet que chaque variable contribue de manière égale au processus d'apprentissage. Cela conduit à une convergence plus rapide pendant l'entraînement. Pour cette étude, la technique de `MinMaxScaler` a été choisie pour normaliser les données.

`MinMaxScaler` transforme chaque caractéristique en redimensionnant ses valeurs dans un intervalle prédéfini, généralement $[0, 1]$. Cette méthode préserve la forme des distributions et est moins sensible aux valeurs aberrantes que d'autres méthodes telles que la normalisation Z-Score (`StandardScaler`).

Il existe d'autres méthodes de normalisation comme le `RobustScaler` ou le `StandardScaler`, mais `MinMaxScaler` a été choisi pour les raisons suivantes :

- Simplicité et Interprétabilité : `MinMaxScaler` est simple à comprendre et à implémenter. Cela facilite l'interprétation des résultats du modèle.
- Conservation de la Forme de la Distribution : Contrairement à la normalisation Z-Score, `MinMaxScaler` ne modifie pas la distribution des données, ce qui est essentiel pour notre cas d'utilisation où nous voulons préserver la nature des variables.

Décomposition des Données

La décomposition des données en ensembles d'entraînement et d'évaluation est une étape cruciale pour l'évaluation des performances des modèles d'apprentissage profond. Dans cette étude, 80% des données ont été allouées à la phase d'entraînement et 20% à la phase d'évaluation.

Il existe d'autres méthode de décomposition comme la méthode K-Fold validation croisée [96], qui est une méthode populaire pour évaluer les performances des modèles. Cependant, elle n'a pas été utilisée pour les raisons suivantes :

- La méthode K-Fold nécessite d'entraîner le modèle K fois sur $K - 1$ sous-ensembles de données, ce qui peut être coûteux en termes de temps et de ressources pour des architectures d'apprentissage profond complexes.
- L'ensemble de données de DKASC est assez volumineux et diversifié (couvrant plusieurs années). Par conséquent, une simple partition en ensembles d'entraînement et d'évaluation est susceptible de capturer la variabilité inhérente des données.

3.2 Optimisation des Hyperparamètres

L'optimisation des hyperparamètres est une étape cruciale pour améliorer les performances des architectures d'apprentissage profond. Dans le contexte de cette étude comparative des modèles de prédiction de la production d'énergie PV, Optuna [97] a été utilisé comme outil d'optimisation des hyperparamètres.

Optuna est une bibliothèque open-source optimisée pour l'optimisation des hyperparamètres d'apprentissage automatique et profond en Python. Il utilise un certain nombre d'algorithmes heuristiques pour explorer efficacement l'espace des hyperparamètres et trouver les valeurs qui minimisent une certaine fonction objectif, généralement l'erreur de validation dans le contexte de l'apprentissage automatique et profond.

Il existe plusieurs raisons qui nous ont poussé à utiliser Optuna au lieu de la recherche par grille. Parmi les raisons on trouve :

- Efficacité computationnelle : La recherche par grille évalue toutes les combinaisons possibles d'hyperparamètres, ce qui peut devenir rapidement impraticable lorsque l'espace des hyperparamètres est grand ou lorsque le modèle est complexe et long à entraîner. Optuna, en revanche, utilise des algorithmes intelligents comme l'optimisation Bayésienne pour chercher de manière plus efficace dans l'espace des hyperparamètres.
- Optimisation Multi-Objectif : Optuna offre la possibilité d'optimiser plusieurs objectifs simultanément, ce qui peut être très utile lorsque l'on cherche à équilibrer les différentes métriques de performance.
- Flexibilité : Optuna permet de spécifier des espaces d'hyperparamètres continus, discrets ou conditionnels, offrant ainsi une grande flexibilité dans la définition du problème d'optimisation.
- Facilité d'intégration : Optuna peut être facilement intégré avec différentes bibliothèques d'apprentissage automatique et profond comme scikit-learn, TensorFlow et PyTorch, ce qui le rend extrêmement versatile.
- Pruning : Une caractéristique intéressante d'Optuna est le "pruning" automatique des essais. Si au cours d'un essai, il devient clair que les hyperparamètres en cours ne vont probablement pas conduire à une performance de modèle optimale, Optuna arrête automatiquement cet essai, économisant ainsi le temps et les ressources.
- Analyse et visualisation : Optuna fournit également des outils intégrés pour l'analyse et la visualisation des résultats d'optimisation, ce qui facilite l'interprétation et la communication des résultats.

En raison de ces avantages en termes d'efficacité computationnelle, de flexibilité et de fonctionnalités supplémentaires comme le pruning et l'optimisation multi-objectif, Optuna a été

choisi comme un outil d'optimisation des hyperparamètres plutôt que la recherche par grille traditionnelle.

3.3 Modèles d'Apprentissage Profond Utilisés

Dans le domaine de la prédiction de séries temporelles, l'utilisation des modèles d'apprentissage profond a récemment gagné en popularité grâce à leur capacité à modéliser des relations complexes et non-linéaires dans les données. Les séries temporelles, comme les données du productible de l'énergie PV, présentent souvent des structures saisonnières, des tendances et des dépendances temporelles. Dans cette étude, nous allons aborder une étude comparative entre plusieurs modèles d'apprentissage profond qui représentent l'état de l'art des modèles de prédiction des séries temporelles.

3.3.1 CrossFormer

Le coût de calcul et de mémoire du mécanisme d'attention est trop élevé pour les entrées visuelles en raison de la longueur de la séquence d'embedding. En outre, les transformateurs de vision existants ne parviennent pas à créer des interactions entre des caractéristiques d'échelles différentes, ce qui est essentiel pour la compréhension des images et des tâches telles que la segmentation d'instances. En effet, les embeddings sont générés à partir de parcelles de taille égale et sont soit inchangés, soit agrandis de manière uniforme, ce qui réduit la capacité d'attention à différentes échelles.

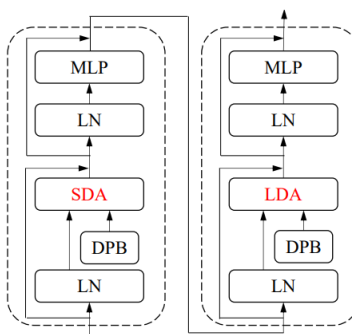


FIGURE 3.1 Structure interne de deux blocs CrossFormer consécutifs. SDA et LDA apparaissent alternativement dans différents blocs, LN représente la couche de normalisation. [6]

L'architecture de CrossFormer [6] est divisée en quatre étapes, chacune contient une couche d'embedding inter-échelle (Cross-Scale Embedding Layer ou CEL) et plusieurs blocs CrossFormer. La CEL génère des embeddings à différentes échelles. Les blocs CrossFormer comprennent une attention à longue et courte distance (Long Short-Distance Attention ou LSDA)

et un biais de position dynamique (Dynamic Position Bias ou DPB). La CEL utilise des couches convolutionnelles pour échantillonner et projeter des patches d'image à différentes échelles. Chaque bloc CrossFormer comprend LSDA pour réduire les coûts de mémoire et de calcul, et qui implique des modules d'attention à courte distance (SDA) et d'attention à longue distance (LDA), ainsi qu'un perceptron multicouche (MLP) avec des connexions résiduelles.

3.3.2 Non-Stationary Transformer (NonST)

Afin d'améliorer les performances des transformers dans la prévision des données de séries temporelles du monde réel. Bien que les transformers aient démontré leur efficacité dans la modélisation des séries temporelles, leurs performances peuvent être affectées lorsqu'ils sont confrontés à des données non stationnaires. Le modèle Non-Stationary Transformer [7] comporte deux parties complémentaires : La stationnarisation des séries pour atténuer la non-stationnarité des séries temporelles et l'attention déstationnaire pour réincorporer les informations non-stationnaires des séries brutes. Grâce à ces conceptions, les modèles Non-Stationary Transformers peuvent améliorer la prévisibilité des données et maintenir simultanément la capacité du modèle.

Comme le montre la figure 3.2, le modèle Non-Stationary Transformer contient deux opérations correspondantes : 1) Le module de normalisation, dans un premier temps, pour traiter les séries non stationnaires causées par la variation de la moyenne et de l'écart type. 2) Le module de dénormalisation, à la fin, pour retransformer les résultats du modèle avec les statistiques originales.

Module de Normalization

Pour atténuer la non-stationnarité de chaque série d'entrée, nous effectuons une normalisation sur la dimension temporelle par une fenêtre coulissante sur le temps. Pour chaque série d'entrée $x = [x_1, x_2, \dots, x_S]^T \in \mathbb{R}^{S \times C}$, nous la transformons par des opérations de translation et de mise à l'échelle et obtenons $x' = [x'_1, x'_2, \dots, x'_S]^T \in \mathbb{R}^{S \times C}$, où S et C représentent respectivement la longueur de la séquence et le nombre de variables. Le module de normalisation peut être formulé à l'aide de l'équation 3.1.

$$\mu_x = \frac{1}{S} \sum_{i=1}^S x_i, \quad \sigma_x^2 = \frac{1}{S} \sum_{i=1}^S (x_i - \mu_x)^2, \quad x'_i = \frac{1}{\sigma_x} \odot (x_i - \mu_x) \quad (3.1)$$

où $\mu_x, \sigma_x \in \mathbb{R}^{C \times 1}$, $\frac{1}{\sigma_x}$ signifie la division par élément et \odot est le produit par élément. Il

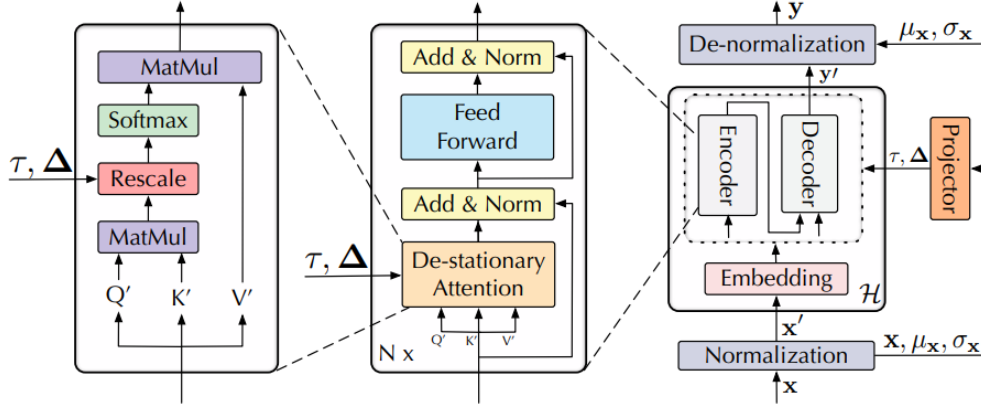


FIGURE 3.2 Non-Stationary Transformers. La stationnarisation des séries est adoptée comme une enveloppe sur le modèle de base pour normaliser chaque série entrante et dé-normaliser la sortie. L’attention déstationnaire remplace le mécanisme original d’attention pour approximer l’attention apprise à partir de séries non stationnaires, qui redimensionne les poids de dépendance temporelle actuels avec des facteurs déstationnaires appris τ, Δ . [7]

convient de noter que le module de normalisation réduit l’écart de distribution entre chaque série temporelle d’entrée, ce qui rend la distribution de l’entrée du modèle plus stable.

Module de Dénormalisation

Comme le montre la figure 3.2, après que le modèle de base \mathcal{H} a prédit la valeur future avec la longueur- O , nous adoptons la dénormalisation pour transformer la sortie du modèle $y' = [y'_1, y'_2, \dots, y'_O]^T \in \mathbb{R}^{O \times C}$ with μ_x et σ_x et obtenons $\hat{y} = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_O]^T$ comme résultats de prévision éventuels. Le module de dénormalisation peut être formulé à l’aide de l’équation 3.2.

$$y' = \mathcal{H}(x'), \quad \hat{y}_i = \sigma_x \odot (y'_i + \mu_x) \quad (3.2)$$

Grâce à la transformation en deux étapes, les modèles de base recevront des entrées stationnaires, qui suivent une distribution stable et sont plus faciles à généraliser. Cette conception rend également le modèle équivariant aux perturbations de translation et d’échelle des séries temporelles, ce qui favorise la prévision des séries dans le monde réel.

3.3.3 Autoformer

L’architecture Autoformer [8] transforme le modèle Transformer en une architecture de décomposition profonde (figure 3.3), comprenant le bloc de décomposition en série interne, le mécanisme d’autocorrélation et les encodeurs et décodeurs correspondants. le bloc de dé-

composition en série interne, le mécanisme d'auto-corrélation, et les encodeurs et décodeurs correspondants.

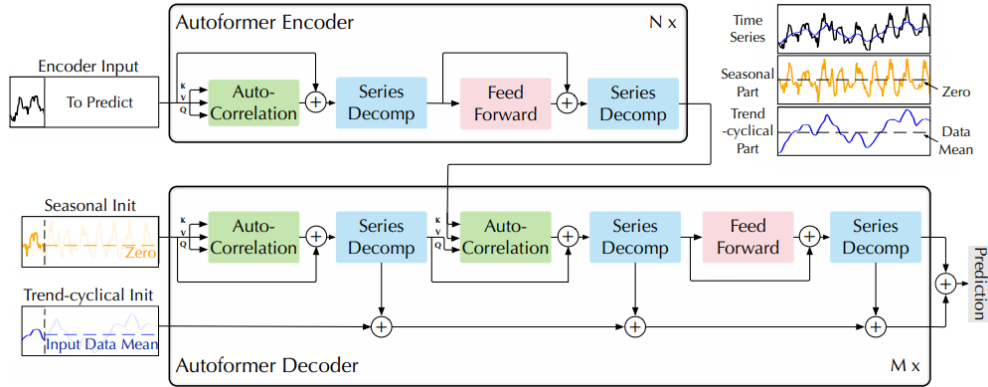


FIGURE 3.3 Architecture de l'Autoformateur. L'encodeur élimine la partie tendance-cyclique à long terme par des blocs de décomposition en série (blocs bleus) et se concentre sur la modélisation des tendances saisonnières. Le décodeur accumule progressivement la partie tendance extraite des variables cachées. Les informations saisonnières passées du l'encodeur sont utilisées par l'auto-corrélation encodeur-décodeur (bloc vert central dans le décodeur). [8]

Décomposition des séries

Le bloc de décomposition des séries dans le modèle Autoformer utilise le concept de décomposition pour séparer les séries en deux parties : la partie de tendance-cyclique et la partie saisonnière. La première reflète la progression à long terme de la série, tandis que la seconde en indique la saisonnalité.

La décomposition directe n'est pas réalisable pour les séries futures car l'avenir est inconnu. Pour résoudre ce problème, le bloc de décomposition des séries agit comme une opération interne de l'Autoformer. Il est capable d'extraire la tendance stationnaire à long terme à partir de variables cachées intermédiaires prédites. Concrètement, une moyenne mobile est adaptée pour lisser les fluctuations périodiques et mettre en évidence les tendances à long terme. Le processus est exprimé à l'aide de l'équation 3.3.

$$\begin{aligned}\mathcal{X}_t &= \text{AvgPool}(\text{Padding}(\mathcal{X})) \\ \mathcal{X}_s &= \mathcal{X} - \mathcal{X}_t\end{aligned}\tag{3.3}$$

where $\mathcal{X}_s, \mathcal{X}_t \in \mathbb{R}^{L \times d}$ désignent respectivement la partie saisonnière et la partie tendance-cyclique extraite.

Encodeur

Comme le montre la figure 3.3, l'encodeur se concentre sur la modélisation de la partie saisonnière. La sortie de l'encodeur contient les informations saisonnières passées et sera utilisée comme information croisée pour aider le décodeur à affiner les résultats de la prédiction. le décodeur à affiner les résultats de la prédiction.

Décodeur

Le décodeur comprend deux parties : la structure d'accumulation pour les composantes cycliques des tendances et le mécanisme d'autocorrélation empilé pour les composantes saisonnières (figure 3.3). Chaque couche du décodeur contient l'auto-corrélation interne et l'auto-corrélation encodeur-décodeur, qui peuvent respectivement affiner la prédiction et utiliser les informations saisonnières passées. Il convient de noter que le modèle extrait la tendance potentielle des variables cachées intermédiaires au cours du décodeur, ce qui permet à l'Autoformer d'affiner progressivement la prédiction de la tendance et d'éliminer les informations d'interférence pour la découverte des dépendances basées sur la période dans l'Auto-Corrélation.

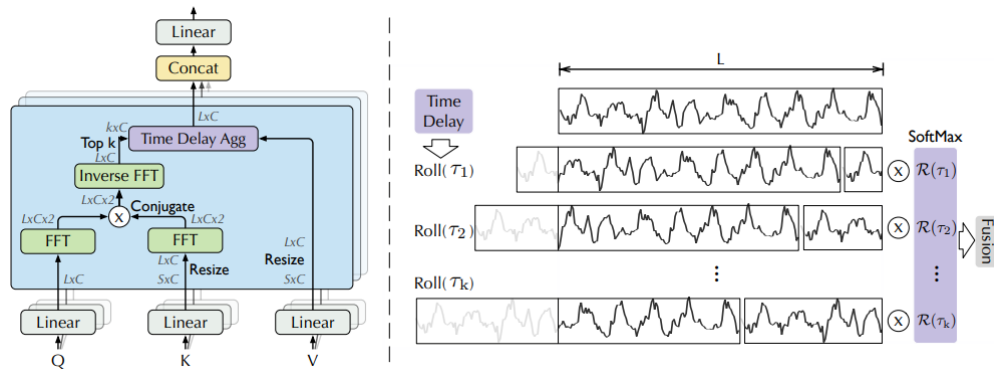


FIGURE 3.4 Auto-corrélation (à gauche) et agrégation des délais (à droite). La transformée de Fourier rapide est utilisée pour calculer l'autocorrélation $\mathcal{R}(\tau)$, qui reflète les similitudes de délai. Les sous-processus similaires sont ensuite regroupés dans le même index sur la base du délai τ sélectionné et agrégés par $\mathcal{R}(\tau)$. [8]

Mécanisme de l'Auto-Corrélation

Le mécanisme d'Auto-Corrélation, sur la figure 3.4, est proposé avec des connexions en série pour étendre l'utilisation des informations. L'auto-corrélation découvre les dépendances basées sur la période en calculant l'autocorrélation des séries et regroupe les sous-séries similaires par l'agrégation des délais.

3.3.4 Reformer

Le Reformer [98] est un modèle introduit pour améliorer l’efficacité des modèles Transformer. Il s’attaque au problème des coûts computationnel élevés et des besoins en mémoire, en particulier pour les longues séquences. Le Reformer remplace le mécanisme d’attention par produit de points par un mécanisme qui utilise le hachage sensible à la localité, ce qui réduit la complexité de l’attention de $O(L^2)$ à $O(L \times \log(L))$, où L est la longueur de la séquence. Il utilise également des couches résiduelles réversibles au lieu de couches résiduelles standard, ce qui réduit la nécessité de stocker les activations plusieurs fois.

Hachage Sensible à la Localité (LSH)

Le Hachage Sensible à la Localité est une technique utilisée pour trouver efficacement les voisins les plus proches dans des espaces de grande dimension. Il emploie un schéma de hachage spécialisé qui attribue un hachage $h(x)$ à chaque vecteur \mathbf{x} . L’unicité de ce schéma de hachage est qu’il est “sensible à la localité”, ce qui signifie que les vecteurs proches les uns des autres dans l’espace des données sont susceptibles d’avoir la même valeur de hachage avec une forte probabilité. À l’inverse, il est peu probable que des vecteurs éloignés aient la même valeur de hachage.

Le hachage est effectué par une méthode qui utilise des projections aléatoires. Pour générer b valeurs de hachage, une matrice aléatoire R de taille $[d_k, \frac{b}{2}]$ est d’abord créée. Le hachage $h(x)$ pour un vecteur donné x est ensuite défini comme $h(x) = \arg \max([xR; -xR])$, où $[u; v]$ signifie la concaténation des vecteurs u et v .

Cette méthode, est connue pour être efficace et est facile à mettre en œuvre, même pour des lots de vecteurs. Elle est particulièrement utile pour garantir que les vecteurs proches sont hachés dans le même seau avec une forte probabilité, et que ces seaux de hachage ont une taille similaire, également avec une forte probabilité.

3.3.5 ETSformer

Le modèle ETSformer [9], figure 3.5, est une architecture de encodeur-décodeur inspirée du lissage exponentiel et spécialisée dans la résolution du problème de la prédiction des séries temporelles. La méthodologie de conception de l’architecture repose sur trois principes clés :

- L’architecture tire parti de l’empilement de plusieurs couches pour extraire progressivement une série de représentations de niveau, de croissance et de saison à partir du résidu latent intermédiaire.
- Dans l’esprit du lissage exponentiel, le modèle extrait les modèles saisonniers saillants

tout en modélisant les composantes de niveau et de croissance en attribuant un poids plus élevé aux observations récentes.

- La prévision finale est une composition de composantes de niveau, de croissance et de saison, ce qui la rend interprétable par l'être humain.

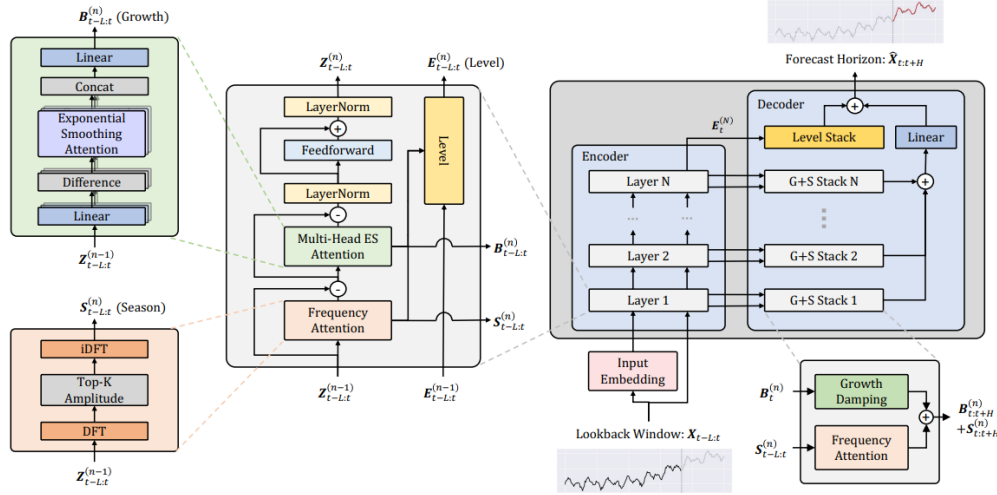


FIGURE 3.5 Architecture du modèle ETSformer. [9]

3.3.6 FEDformer

Inspirés par la décomposition des tendances saisonnières et l'analyse de la distribution, le modèle FEDformer [10] est une architecture de décomposition profonde, comme le montre la figure 3.6, comprenant un bloc à fréquence augmentée (FEB), une attention à fréquence augmentée reliant l'encodeur et le décodeur, et un bloc de décomposition par mélange d'experts (MOEDecomp).

L'encodeur adopte une structure multicouche comme suit dans l'équation 3.4.

$$\mathcal{X}_{en}^l = \mathbf{Encoder}(\mathcal{X}_{en}^{l-1}) \quad (3.4)$$

où $l \in \{1, \dots, N\}$ représente la sortie de la l -ième couche de l'encodeur et $\mathcal{X}_{en}^0 \in \mathbb{R}^{I \times D}$ est la série historique série historique intégrée.

Le décodeur adopte aussi une structure multicouche comme suit dans l'équation 3.5.

$$\mathcal{X}_{de}^l, \mathcal{T}_{de}^l = \mathbf{Decoder}(\mathcal{X}_{de}^{l-1}, \mathcal{T}_{de}^{l-1}) \quad (3.5)$$

où $l \in \{1, \dots, M\}$ représente la sortie de la l -ième couche du décodeur, et \mathcal{T}_{de}^l représente la composante de tendance dans la l -ième couche de décodeur.

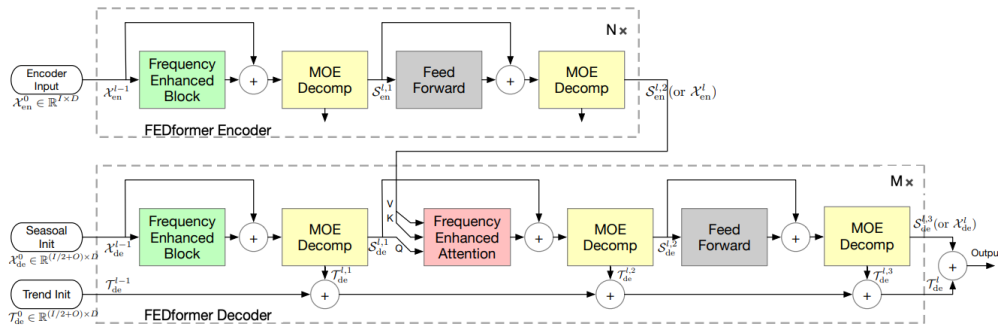


FIGURE 3.6 Structure du FEDformer. Le FEDformer se compose de N encodeurs et de M décodeurs. Le Frequency Enhanced Block (FEB, blocs verts) et le Frequency Enhanced Attention (FEA, blocs rouges) sont utilisés pour effectuer l'apprentissage de la représentation dans le domaine fréquentiel. Le FEB ou le FEA ont deux sous-versions (FEB-f et FEB-w ou FEA-f et FEA-w), où "-f" signifie l'utilisation de la base de Fourier et "-w" l'utilisation de la base d'ondelettes. Les blocs de décomposition par mélange d'experts (MOEDecomp, blocs jaunes) sont utilisés pour extraire les tendances saisonnières des données d'entrée. [10]

3.3.7 Informer

Deux limitations principales ont été identifiées dans les architectures actuelles de prévision des séries temporelles : l'erreur accumulée due au décodage dynamique et l'incapacité à saisir les dépendances à long terme dans les données. Le modèle Informer [11] vise à remédier à ces limitations en améliorant la capacité de prédiction pour les longues séquences. La figure 3.7 représente l'architecture du modèle Informer.

Encodeur

l'encodeur reçoit des entrées massives de longues séquences (série verte). L'auto-attention canonique est remplacé par l'auto-attention ProbSparse. Le trapèze bleu, sur la figure 3.7, est l'opération de distillation de l'auto-attention pour extraire l'attention dominante, ce qui réduit fortement la taille du réseau. Les répliques de l'empilement des couches augmentent la robustesse.

Décodeur

Le décodeur reçoit de longues séquences d'entrée, met les éléments cibles à zéro, mesure la composition pondérée de l'attention de la carte des caractéristiques et prédit instantanément les éléments de sortie (série orange) dans un style génératif.

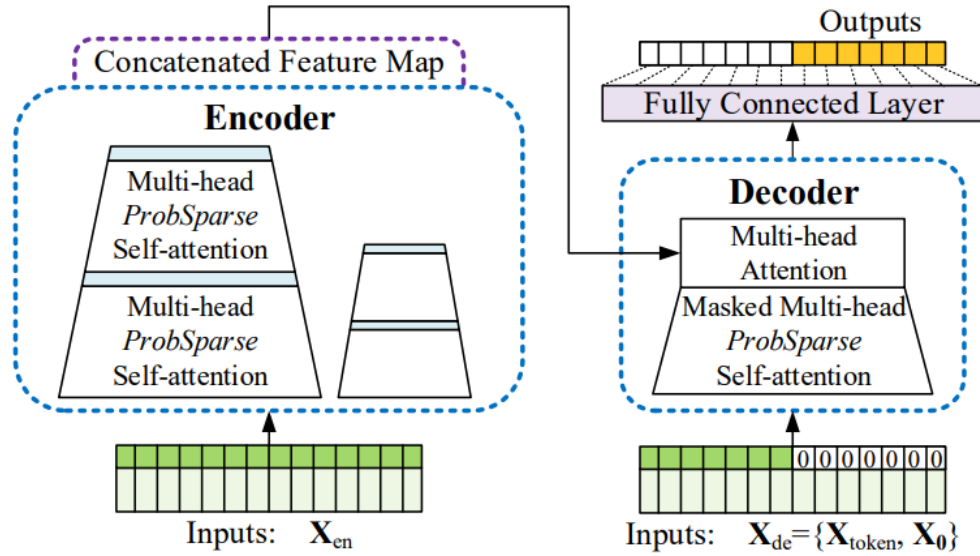


FIGURE 3.7 Architecture du modèle Informer. [11]

3.3.8 Pyraformer

L'architecture Pyraformer [12] vise à résoudre le problème de prévision des séries temporelles. L'objectif est de prédire les valeurs futures $z_{t+1:t+M}$ en se basant sur les valeurs passées observées $z_{1-L+1:t}$ et les covariables associées $x_{t-L+1:t+M}$. La figure 3.8 représente l'architecture du modèle Pyraformer. Le Module de Construction à Échelle Grossière (CSCM) résume la séquence intégrée à différentes échelles et construit une structure arborescente multi-résolution. Le Module d'Attention Pyramidal (PAM) est ensuite utilisée pour échanger des informations entre les nœuds de manière efficace.

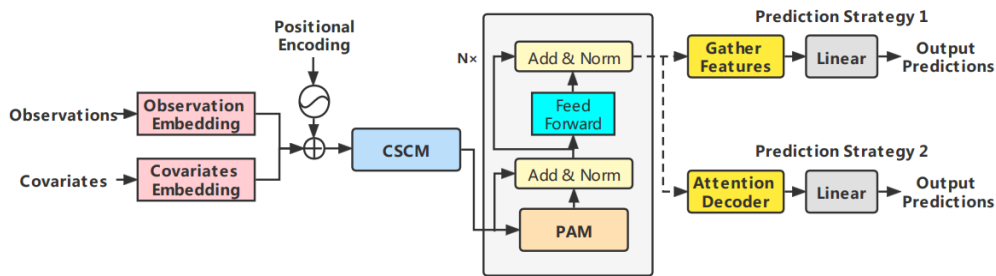


FIGURE 3.8 Architecture du modèle Pyraformer. [12]

Incorporation des Données

Au départ, l’architecture incorpore séparément les données observées, les covariables, et les positions. Ces incorporations sont ensuite additionnées ensemble, une approche similaire à celle utilisée dans le modèle Informer.

Arbre Multi-Résolution

Le Pyraformer utilise un Module de Construction à Échelle Grossière (CSCM) pour créer un arbre C-aire multi-résolution. Dans cet arbre, les nœuds à une échelle plus grossière résument les informations de C nœuds à une échelle plus fine. Cette structure arborescente permet au modèle de capturer et de résumer différentes échelles de données.

Module d’Attention Pyramidal (PAM)

Il s’agit d’un composant innovant conçu pour capturer les dépendances temporelles sur différentes plages de temps. Le PAM fonctionne sur l’arbre multi-résolution et passe des messages en utilisant des mécanismes d’attention au sein de cette structure graphique pyramidale.

Structure de Sortie

Enfin, l’architecture utilise différentes structures de réseau pour générer les prédictions finales, en fonction de la tâche en aval spécifique.

3.3.9 Frequency improved Legendre Memory (FiLM)

La structure globale du FiLM [13] est illustrée à la figure 3.9. Le FiLM cartographie une séquence $X \mapsto Y$, où $X, Y \in \mathbb{R}^{(T \times D)}$, en utilisant principalement deux sous-couches : la couche LPU (Legendre Projection Unit) et la couche FEL (Fourier Enhanced Layer). En outre, pour capturer les informations historiques à différentes échelles, un mélange d’experts à différentes échelles est mis en œuvre dans la couche LPU. Une couche supplémentaire facultative de normalisation des données, RevIN, est introduite pour améliorer encore la robustesse du modèle. Il convient de mentionner que FiLM est un modèle simple qui ne comporte qu’une couche de LPU et une couche de FEL.

Unité de projection de Legendre (LPU)

LPU est un modèle d’espace d’état : $C_t = AC_{t-1} + B_{x_t}$, où $x_t \in \mathbb{R}$ est le signal d’entrée, $C_t \in \mathbb{R}^N$ est l’unité de mémoire, et N est le nombre de polynômes de Legendre. LPU contient

deux matrices préfixées non entraînaibles A et B définies comme suit dans l'équation 3.6.

$$A_{nk} = (2n + 1) \begin{cases} (-1)^{n-k} & \text{if } k \leq n \\ 1 & \text{if } k \geq n \end{cases}, \quad B_n = (2n + 1)(-1)^n \quad (3.6)$$

La LPU comprend deux étapes, à savoir la projection et la reconstruction. La première étape projette le signal original vers l'unité de mémoire : $C = \mathbf{LPU}(X)$. L'étape suivante reconstruit le signal à partir de l'unité de mémoire : $X_{re} = \mathbf{LPU_R}(C)$.

Couche Améliorée en Fréquence (FEL)

Approximation de bas-rang : La FEL est dotée d'une matrice de poids unique pouvant être apprise ($W \in \mathbb{R}^{M' \times N \times N}$), ce qui est tout ce dont nous avons besoin pour apprendre à partir des données. Cependant, ce poids peut être important. Nous pouvons décomposer W en trois matrices $W_1 \in \mathbb{R}^{M' \times N' \times N'}$, $W_2 \in \mathbb{R}^{N \times N'}$ et $W_3 \in \mathbb{R}^{N' \times N}$ pour effectuer une approximation de bas rang ($N' \ll N$).

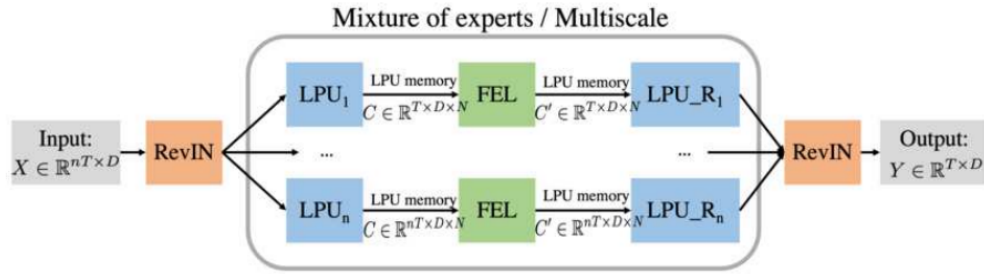


FIGURE 3.9 Structure globale du FiLM (modèle de mémoire de Legendre amélioré en fréquence). LPU : Legendre Projection Unit. LPU_R : récupération inverse de la projection de Legendre. FEL : couche à fréquence améliorée. RevIn : bloc de normalisation des données. Les données d'entrée sont d'abord normalisées, puis projetées dans l'espace des polynômes de Legendre (mémoire C de la LPU). La mémoire LPU C est traitée avec FEL et génère la mémoire C' de sortie. Enfin, C' est reconstruit et normalisé à la série de sortie avec LPUR. Une structure multiscalaire est employée pour traiter l'entrée de longueur $\{T, 2T, \dots, nT\}$. [13]

3.3.10 LightTS

La figure 3.10 présente l'architecture globale de LightTS [14]. Rappelons que la prévision de séries temporelles multivariées présente deux défis majeurs :

- Capturer les modèles locaux à court terme et les modèles globaux à long terme ;

— Capturer les interdépendances entre les différentes variables des séries temporelles.

LightTS se compose également de deux parties qui correspondent à ces deux défis. Dans la première partie, nous traitons différentes séries temporelles (c'est-à-dire les variables d'entrée) de manière indépendante sans tenir compte de leurs interdépendances. Cette partie vise à capturer les dépendances à court/long terme et à extraire les caractéristiques correspondantes de chaque série (le premier défi). Dans la seconde partie, nous concaténons toutes les séries temporelles et apprenons les corrélations entre les différentes variables d'entrée (le second défi). Les éléments clés de ces deux parties sont deux méthodes d'échantillonnage appelées échantillonnage continu et échantillonnage par intervalles, et trois blocs d'échange d'informations (IEBlock), dont nous décrivons les détails dans les sous-sections suivantes.

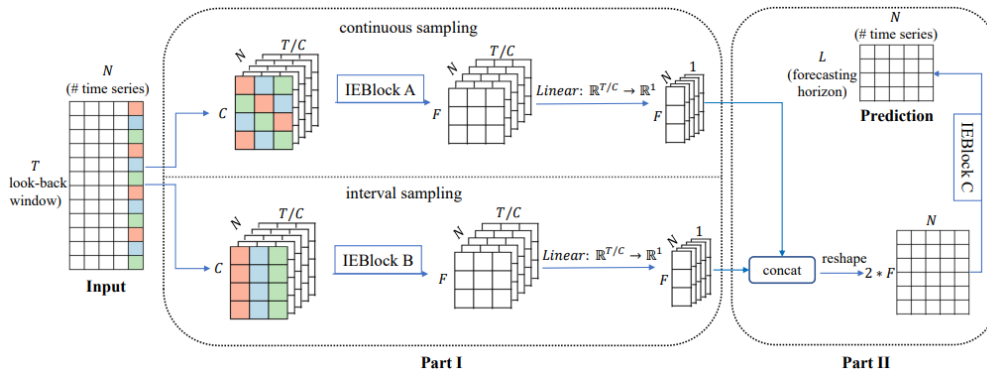


FIGURE 3.10 Vue d'ensemble de LightTS. Dans la partie I, le modèle capture les dépendances à court et long terme et extrait les caractéristiques de chaque série temporelle. Dans la partie II, le modèle apprend les interdépendances entre les différentes séries temporelles et fait des prédictions. [14]

3.3.11 PatchTST

Sur la figure 3.11, l'architecture globale du modèle PatchTST [15] est représentée. Les données de séries temporelles multivariées sont divisées en différents canaux. Ils partagent la même épine dorsale de transformateur, mais les processus d'acheminement sont indépendants.

Transformateur Backbone (supervisé)

Chaque série univariée d'un canal passe par l'opérateur de normalisation d'instance et est segmentée en parcelles, qui sont utilisées comme jetons d'entrée du transformateur. Ces parcelles sont utilisées comme jetons d'entrée du transformateur.

Transformateur Backbone (auto-supervisé)

Apprentissage de représentation auto-supervisé masqué avec PatchTST où les parcelles sont sélectionnées de manière aléatoire et mises à zéro. Le modèle reconstruit les parcelles masquées.

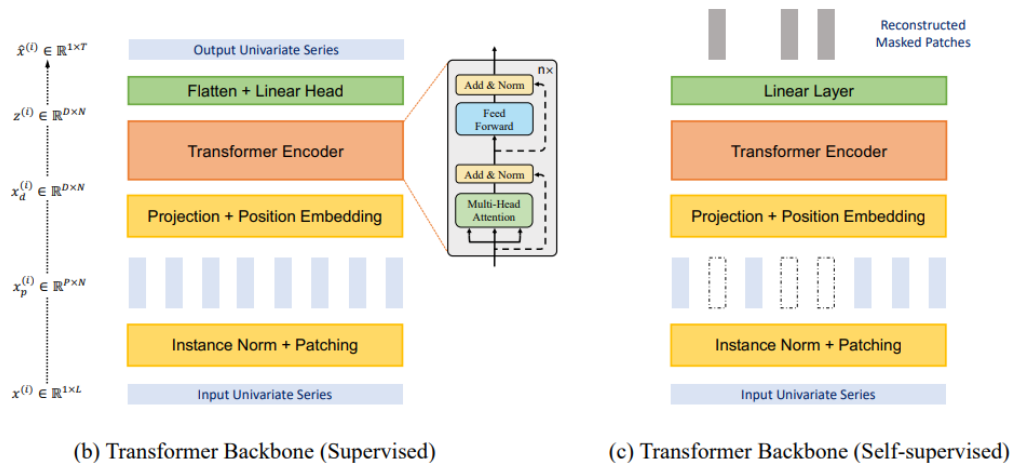
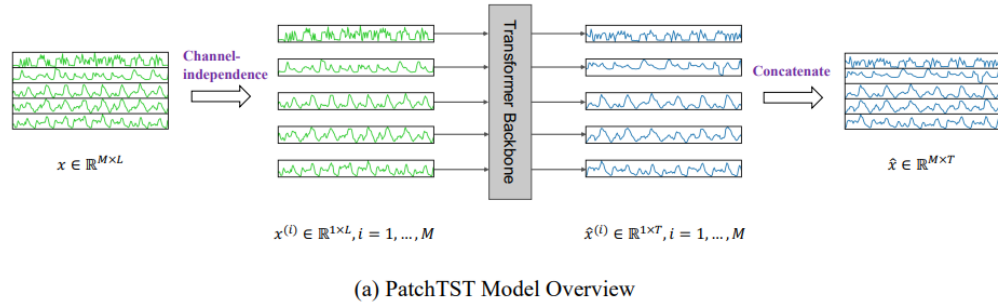


FIGURE 3.11 Architecture globale du modèle PatchTST. [15]

3.3.12 Multi-scale Isometric Convolution Network (MICN)

La structure globale du MICN est présentée à la figure 3.12. La tâche de prédiction des séries temporelles longues consiste à prédire une série future de longueur O sur la base d'une série passée de longueur I , ce qui peut être exprimé comme *entrée* – I – *prédire* – O , où O est beaucoup plus grand que I . Inspirés par les algorithmes traditionnels de décomposition des séries temporelles, le modèle MICN utilise un bloc de décomposition hybride multi-échelle (MHDecomp) pour séparer les schémas complexes des séries d'entrée. MICN utilise ensuite le bloc de prédiction saisonnière pour prédire les informations saisonnières et le bloc de prédiction tendancielle et cyclique pour prédire les informations tendancielle et cycliques. Finalement, les résultats des prédictions sont additionnés pour obtenir la prédiction finale Y_{pred} .

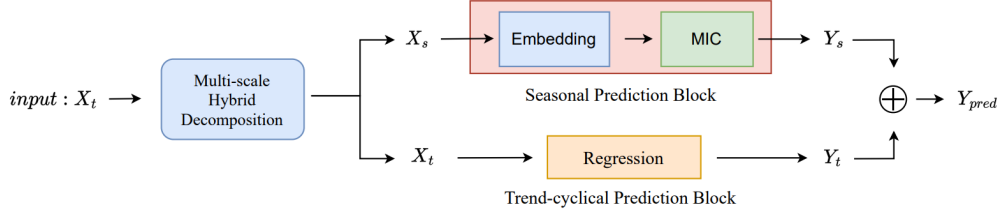


FIGURE 3.12 Architecture globale du modèle MICN. [16]

Décomposition Hybride Multi-échelle

Le bloc de décomposition hybride multi-échelle utilise plusieurs noyaux différents de l' $Avgpool(\cdot)$ et peut séparer plusieurs modèles différents de parties tendancielle, cycliques et saisonnières. Contrairement au bloc MOEDecomp de FEDformer, le modèle utilise de simples opérations de moyenne pour intégrer ces différents modèles parce qu'il est pas possible de déterminer le poids de chaque modèle avant d'apprendre ses caractéristiques [16]. En conséquence, cette opération de pondération est placée dans la partie Fusion du bloc prédiction saisonnière après la représentation des caractéristiques. Concrètement, pour la série d'entrée $X \in \mathbb{R}^{lI} \times d$ le processus est exprimé à l'aide de l'équation 3.7.

$$\begin{aligned} X_t &= \text{mean}(AvgPool(Padding(X))_{kernel_1}, \dots, AvgPool(Padding(X))_{kernel_n}) \\ X_s &= X - X_t \end{aligned} \quad (3.7)$$

Avec $X_t, X_s \in \mathbb{R}^{I \times d}$ représentent la partie tendancielle-cyclique et la partie saisonnière, respectivement.

Bloc de Prédiction Tendence-cyclique

Actuellement, Autoformer concatène la moyenne de la série originale et l'accumule ensuite avec la partie tendancielle et cyclique obtenue à partir du bloc de décomposition de la série intérieure. Mais il n'y a pas d'explication à ce sujet ni de preuve de son efficacité. le modèle MICN, on utilise une stratégie de régression linéaire simple pour faire une prédiction sur la tendance-cyclique, démontrant ainsi qu'une modélisation simple de la tendance-cyclique est également nécessaire pour les tâches de prévision des séries non stationnaires. Concrètement, pour la série trend-cyclique $X_t \in \mathbb{R}^{I \times d}$ obtenue avec le bloc MHDcomp, le processus est exprimé à l'aide de l'équation 3.8.

$$Y_t^{regre} = \text{regression}(X_t) \quad (3.8)$$

où $Y_t^{regre} \in \mathbb{R}^{O \times d}$ désigne la prédiction de la partie tendancielle à l'aide de la stratégie de régression linéaire. On utilise MICN-regre pour représenter le modèle MICN avec cette méthode de prédiction tendancielle et cyclique. À titre de comparaison, la moyenne de X_t est utilisée pour traiter les séries où la tendance-cyclique reste constante comme exprimé dans l'équation 3.9.

$$Y_t^{mean} = mean(X_t) \quad (3.9)$$

où $Y_t^{mean} \in \mathbb{R}^{O \times d}$ désigne la prédiction de la partie tendancielle. Nous utilisons MICN-mean pour représenter le modèle MICN avec cette méthode de prédiction tendance-cyclique.

Bloc de Préviation Saisonnière

Le bloc de prédiction saisonnière se concentre sur la modélisation plus complexe des parties saisonnières. Après avoir intégré la séquence d'entrée X_s , nous adoptons une convolution isométrique multi-échelle pour capturer les caractéristiques locales et les corrélations globales, et les branches de différentes échelles modélisent différents modèles sous-jacents de la série temporelle. Nous fusionnons ensuite les résultats des différentes branches pour compléter l'utilisation complète des informations de la séquence [16]. La méthode peut être résumée à l'aide de l'équation 3.10.

$$\begin{aligned} X_s^{emb} &= Embedding(Concat(X_s, X_{zero})) \\ Y_s^0 &= X_s^{emb} \\ Y_{s,l} &= MIC(Y_{s,l-1}), \quad l \in \{1, 2, \dots, N\} \\ Y_s &= Truncate(Projection(Y_{s,N})) \end{aligned} \quad (3.10)$$

où $X_{zero} \in \mathbb{R}^{O \times d}$ désigne les espaces réservés remplis de zéro et $X_s^{emb} \in \mathbb{R}^{(I+O) \times D}$ désigne la représentation intégrée de X_s . $Y_{s,l} \in \mathbb{R}^{(I+O) \times D}$ représente la sortie de la l-ième couche de convolution isométrique multi-échelle (MIC), et $Y_s \in \mathbb{R}^{O \times d}$ représente la prédiction finale de la partie saisonnière après une projection de la fonction linéaire avec $Y_{s,N} \in \mathbb{R}^{(I+O) \times D}$ et une opération de troncature.

3.3.13 DLinear

DLinear est une combinaison d'un schéma de décomposition utilisé dans Autoformer et FED-former avec des couches linéaires. Il décompose d'abord une entrée de données brutes en une composante de tendance par un noyau de moyenne mobile et une composante résiduelle (sai-

sonnière). Ensuite, deux couches linéaires à une couche sont appliquées à chaque composante, et nous additionnons les deux caractéristiques pour obtenir la prédiction finale. En traitant explicitement les tendances, DLinear améliore les performances d'un modèle linéaire lorsqu'il existe une tendance claire dans les données.

Dans le chapitre qui suit, nous allons entamer une étude comparative entre les différents modèles d'apprentissage profond qu'on vient d'expliquer dans ce chapitre. Plus de détails sur les données, la méthodologie et les résultats seront expliqués dans le chapitre qui suit.

CHAPITRE 4 ARTICLE 1 ONE DAY-AHEAD PV OUTPUT POWER FORECASTING REVEALING THE POWER OF DEEP LEARNING MODELS BEYOND LSTM.

Authors : Saad Benslimane, Hanane Dagdougui and Loubna Benabbou.

Manuscript submitted to: IEEE Transactions on Industrial Informatics.

Soumis le 24 juillet 2023

Abstract: Accurate forecasting of photovoltaic (PV) output power is crucial for efficient power systems, particularly in regions where PV power accounts for a significant portion of the total energy supply. Forecasting of PV power is achieved using physical models incorporating weather and solar data or statistical and machine learning models relying on past data and employing sophisticated algorithms to forecast PV output power. This paper presents a comparative study of various deep learning models to forecast one day ahead of PV output power. Recent literature has highlighted recurrent neural networks (RNNs) as the preferred choice for PV output power forecasting. However, our study aims to validate these claims and compare Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), two RNN models, against other deep learning models for time series forecasting. Our extensive empirical analysis is motivated by the increasing demand for accurate PV output power forecasting to enable better integration of renewable energy resources into the grid and promote energy sustainability. We evaluate the performance of each model across multiple metrics, including mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination, over a multi-year dataset for the region of Yulara in Australia.

Keywords: Solar energy, deep learning, photovoltaic (PV) output power forecasting.

4.1 Introduction

The rapid advancement of renewable energy sources, primarily driven by climate change concerns and accelerated by the swift population and economic growth, is a salient feature of the modern energy landscape [99]. Solar energy systems, specifically PV technology, are central to this development. They are rapidly transforming the structure of modern electricity grids due to their ability to harness the sun's abundant energy. Experts estimate that PV technology could meet up to a quarter of the global electricity demand by 2050 [100].

According to IEA's report, in 2021, renewable capacity additions expanded by 6%, reaching an impressive 295 Gigawatts (GW). This was primarily propelled by an increase in solar PV installations, which offset a 17% reduction in wind capacity additions that same year [101]. In

2022, solar PV technology emerged as a prominent player in the renewable energy industry, achieving a new milestone in its deployment. The growth in net additions, which amounted to around 220 GW, exhibited a remarkable increase of 35% compared to the preceding year. The prevalence of solar PV technology can be attributed to substantial expansions in China and the European Union. These regions collectively accounted for more than 85% of the yearly PV capacity increase [102].

In 2023, solar PV is anticipated to remain the leading global renewable capacity expansion source, contributing 65% to the growth. Distributed applications like residential and commercial systems will drive nearly half the global PV expansion. Looking ahead, projections for 2024 indicate that the yearly expansion of the solar photovoltaic (PV) market will persist, culminating in a nearly 310 GW capacity in 2024, signifying a growth rate of more than 7% compared to the previous year. The annual additions in significant markets such as China, the European Union, the United States, and India have been triggered by lower module prices, increased uptake of distributed PV systems, and policy initiatives aimed at promoting large-scale deployment [102].

The significant advantages of PV power plant deployment are related to their excellent economic viability in both the short and long term, the inexhaustibility of PV systems supply, and their long lifespan [103]. However, PV power plants are influenced by various factors like intermittency and time-variability, which can impact the stability and management of the power grid [104]. Thus, an efficient solution to this problem is to set powerful methods for forecasting PV output power.

The prediction horizon represents one of the factors that may affect the forecasting of PV output power [105]. We can define four categories of the prediction horizon: 1) Very short-term prediction is typically used in managing PV storage, electricity pricing, marketing, and supervising real-time electricity dispatch [106], 2) Short-term prediction is beneficial for regulating integrated renewable energy management systems [107], 3) Medium-term prediction is essential for maintenance planning of solar energy integrated systems [106], and finally 4) Long-term prediction is appropriate for considering seasonal changes, long-term power production, distribution, and solar energy rationing [107]. Through comprehension and implementation of diverse prediction horizons, it is possible to enhance the management and optimization of photovoltaic (PV) power, resulting in a more robust and effective power grid.

Generally, two main approaches can be used for PV output power forecasting: 1) Indirect approach, which operates in two steps. Initially, they concentrate on predicting meteorological characteristics such as solar irradiation, temperature, wind speed, etc. The second step includes using these forecasted meteorological parameters for forecasting PV output power.

As a result, this approach gives more nuanced knowledge by focusing on the connection between weather conditions and PV output power. 2) The direct approach, on the other hand, often uses historical meteorological parameters to forecast PV output power [108]. This approach directly correlates past weather conditions with PV output power, utilizing patterns and trends in past data to make projections about future PV power generation.

In academic literature, various algorithms have been described to provide accurate predictions of PV output power over varying forecasting time horizons, ranging from a few minutes to many days. Commonly, these algorithms are divided into four categories: statistical models, physical models, machine learning models, and hybrid models [109]. Statistical models, such as the Autoregressive Moving Average [110], Grey Theory [111], Kalman Filter [112], Bayesian Approach [113], and Markov Chain Model [114], are used to determine the mathematical relationships between historical meteorological parameters and PV output power [115]. Despite the utility of these models, it is essential to recognize that many of these statistical techniques have inherent limitations. Their capacity to manage such extensive forecasting horizons is limited [111].

On the other hand, physical models are based on numerical weather prediction (NWP) or cloud observations that simulate atmospheric dynamics [116]. These NWP-based models can be further divided into two distinct categories: global models and regional models. Both types can forecast solar energy and cloud cover by employing meteorological parameters like temperature, humidity, and wind speed, among others [117].

L. A. Fernandez-Jimenez and al. [118] employed two NWP-based models - the Global Forecasting System (GFS) and the Meso-scale NWP Model (MM5) - to forecast weather variables. Then, they used these forecasted weather variables from both GFS and MM5 as input for an artificial neural network model to forecast hourly energy generation. A similar methodology was utilized by Lima et al. [119], wherein they used the Weather Research and Forecasting model (WRF) to forecast irradiance over 24 hours with a 1-hour time resolution across North-eastern Brazil. The output of the WRF model was then used as input for an artificial neural network to minimize deviations and increase the confidence of the forecasted irradiance.

The other physical model subclass is the cloud observation-based model, which utilizes sky or satellite imaging to provide images of cloud movements [120, 121]. Researchers have used this model to convert digital sky images into solar irradiance and solar energy values [122]. In contrast to statistical methods, physical models provide more credible results for long-term predictions [105, 109]. However, a significant drawback of physical models is that they require a lot of computational resources due to the large amount of data required for frequent calibration [123]. This intensive computational demand can present processing speed and

resource allocation challenges.

On the other hand, machine learning models benefit from the fast development of computing power to provide higher performance than physical and statistical models. Machine learning models learn the mapping relationship between input and output variables through minimizing error [124]. Due to their potential abilities, machine learning models have been adopted across various domains, including power grid management [125], abnormal detection [126], energy consumption [127], and have notably emerged as a powerful tool for PV output power forecasting [128]. Using machine learning's inherent ability to identify patterns and dependencies within large datasets, these models can accurately forecast future outcomes based on historical data, thus improving the accuracy and reliability of forecasting PV output power.

Due to their strong predictive capabilities, several researchers have applied machine learning models to forecast PV output power. Key among these models are Random Forests (RF) [77], Support Vector Machines (SVM) [78], Artificial Neural Networks (ANN) [129], and Recurrent Neural Networks (RNNs) [104]. Each machine learning model contributes distinctive advantages to forecasting PV output power. Random Forests, for instance, are resistant to overfitting, SVM excels in high-dimensional feature spaces, ANN can model complex non-linear relationships, and RNNs are particularly suited to handling sequential data, making them a popular choice for time series problems such as PV output power forecasting. This variety of approaches illustrates the various machine learning techniques that can be applied to this problem, each contributing its assets to the endeavor.

In recent years, artificial intelligence (AI) has experienced a significant surge, with a particular emphasis on the rapid progression of deep learning technologies. A host of models have been developed to address various challenges, with a particular focus on solving time series problems. This article's primary objective is to explore and evaluate the application of novel deep learning models for PV output power forecasting. These models have evolved to handle complex relationships and dependencies, demonstrating superior capabilities compared to traditional methods. Specifically, we focus on one-day-ahead PV output power forecasting, a critical horizon for operational and trading decisions in the energy market. The main contributions of this work are explained as follows:

- We apply novel deep learning models for one-day ahead PV output power forecasting using meteorological parameters as inputs of the models to capture intricate relationships and dependencies that significantly affect PV output power output.
- We evaluate the performance of these deep learning models against the LSTM and GRU architectures that have been widely used in PV output power forecasting over the past years due to their ability to handle time series data and their proficiency in

capturing long-term dependencies.

4.2 Methodology

4.2.1 Direct approach

The indirect approach typically relies on forecasting future meteorological parameters, which are subsequently used for forecasting PV output power using machine learning models. Although this method has demonstrated effectiveness in certain circumstances, it is prone to error accumulation due to the two-stage forecasting process. Using NWP data as input for the deep learning model could help bypass the error accumulation issues in the indirect method. Given the context of this research, where we have access to only historical meteorological data, the direct approach could be more feasible. Consequently, this work will adopt the direct approach, forecasting PV output power for one day ahead using the past 48 hours of PV output power and meteorological parameters. This decision is based on practical data constraints and the desire to minimize the accumulation of forecasting errors.

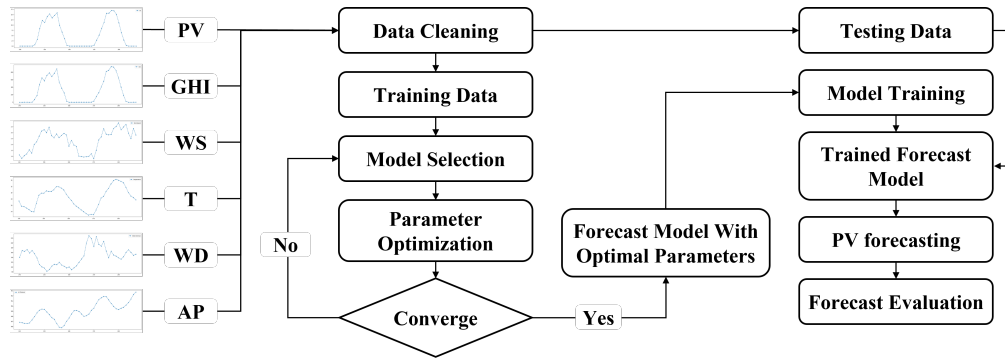


Figure 4.1 Flow chart of the machine learning process for PV output power using the direct approach.

The process followed for this work is shown in figure 4.1, which encompasses several key stages. The initial step of the process involves data preprocessing, which comprises feature selection, data normalization, and data splitting. This phase ensures the data is in an appropriate format for input into the machine learning models. It eliminates any issues, such as missing values or varying scales of measurement, that could impact the performance of the models. In the second step, the data is divided into three distinct datasets: training, validation, and testing. The training and validation datasets are utilized during model training and hyperparameter optimization, which dictate the model's learning process and ultimately

influence its performance. For the hyperparameter optimization step, we used Optuna [97]. On the other hand, the test dataset is reserved for the final evaluation stage, where it serves as unseen data to compare the performance of different deep learning architectures objectively. This methodology aims to rigorously test the effectiveness of various deep learning models in PV output power forecasting and provide valuable insights into their relative performance when applied to real-world data.

4.2.2 Evaluation indices

Performance estimation is crucial for evaluating the accuracy of each forecasting model. For a given data in the time series, the difference between the actual and the forecast values represents the fundamental core of each evaluation index. In this work, we will use the following evaluation indices:

Root Mean Square Error

The root mean square error (RMSE) measures the average value of the error using the square root of the mean square error between the actual values and the forecast values as shown in equation (4.1).

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (4.1)$$

Normalized Root Mean Square Error

The normalized root mean square error (nRMSE) is a variant of the RMSE. It is a relative measure of error that considers the target variable's scale. It enables the comparison of model correctness across multiple datasets or variables with differing scales. The nRMSE can be calculated using the equation (4.2).

$$nRMSE = \frac{100}{y_{max} - y_{min}} * RMSE \quad (4.2)$$

Mean Absolute Error

The mean absolute error (MAE) measures the absolute value of the average difference between the actual values and the forecast values as shown in equation (4.3).

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (4.3)$$

Normalized Mean Absolute Error

The normalized mean absolute error (nMAE), similar to the nRMSE, provides a relative measure of error that considers the target variable’s scale. It enables the comparison of model correctness across multiple datasets or variables with differing scales. The nMAE can be calculated using the equation (4.4).

$$nMAE = \frac{100}{y_{max} - y_{min}} * MAE \quad (4.4)$$

Coefficient of determination

The coefficient of determination (R^2) measures the strength of the relationship between the actual values and the forecast values using the equation 4.5.

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (4.5)$$

4.2.3 Benchmark Models

Transformer

This approach consists of an encoder-decoder structure, with the encoder and decoder composed of multiple identical layers. Each layer has two sub-layers: a multi-head self-attention mechanism and a position-wise fully connected feed-forward network. The self-attention mechanism allows the model to capture dependencies between different positions in the input sequence. The feed-forward network further processes the information, enabling the model to learn complex patterns [5].

Crossformer

The Crossformer addresses the limitations of the existing vision transformers in building interactions among features of different scales. The CrossFormer architecture consists of two main components: Cross-scale Embedding Layer (CEL) and Long Short Distance Attention (LSDA). CEL blends each embedding with multiple patches of different scales, providing the self-attention module with cross-scale features. LSDA splits the self-attention module into

a short-distance component and a long-distance component, preserving both small-scale and large-scale features in the embeddings. These two components enable cross-scale attention in the CrossFormer architecture [6].

Non-stationary Transformer (NonST)

It aims to enhance the performance of Transformers in forecasting real-world time series data. While Transformers have shown effectiveness in modeling time series, their performance can suffer when faced with non-stationary data. The Non-stationary Transformers framework consists of two modules: Series Stationarization and De-stationary Attention. Series stationarization normalizes the input time series, while De-stationary Attention approximates the attention learned from the raw series to recover the intrinsic non-stationary information. By incorporating both local variations and overall trends, the architecture improves the predictability of non-stationary time series while maintaining the capabilities of deep models [7].

Autoformer

This approach incorporates a series decomposition block and an Auto-Correlation mechanism. The series decomposition block allows the model to progressively decompose the long-term trend information from the predicted hidden variables. This enables the model to better capture the complex temporal patterns in the data. The Auto-Correlation mechanism is based on the series periodicity and conducts dependencies discovery and representation aggregation at the sub-series level [8].

Reformer

This method was introduced to improve the efficiency of the Transformer models. It addresses the problem of high training costs and memory requirements, especially for long sequences. The Reformer replaces the dot-product attention mechanism with one that uses locality-sensitive hashing, reducing the complexity of attention. It also uses reversible residual layers instead of standard residuals, reducing the need to store activations multiple times [98].

ETSformer

ETSformer introduces novel components called exponential smoothing attention (ESA) and frequency attention (FA) to replace the traditional self-attention mechanism in Transformers.

These components improve the accuracy and efficiency of the model by focusing on the temporal dependencies and seasonal patterns in the time-series data. ETSformer also redesigns the Transformer architecture with modular decomposition blocks, allowing it to decompose the time-series data into interpretable components such as level, growth, and seasonality [9].

FEDformer

It combines the Transformer with the seasonal-trend decomposition method. This allows the decomposition method to capture the global profile of the time series while the Transformer captures more detailed structures. In [10], the authors found that most time series have a sparse representation in well-known bases such as the Fourier transform. They developed a frequency-enhanced Transformer that takes advantage of this sparse representation [10].

Informer

It uses the ProbSparse self-attention mechanism, which reduces the computation and memory usage while capturing long-range dependencies. Also, the Informer model uses a self-attention distilling operation to prioritize dominant attention scores and reduce the total space complexity. This allows the model to handle long sequence inputs efficiently. Additionally, the Informer model uses a generative style decoder to predict long sequence outputs in a single forward step, avoiding error accumulation. [11].

Pyraformer

The approach is a pyramidal attention module to capture the data's short-term and long-term dependencies. It introduces a multi-resolution representation of the time series by summarizing features at different scales. The model achieves high prediction accuracy with low time and memory consumption, especially for long sequences. [12].

Frequency improved Legendre Memory (FiLM)

FiLM addresses the challenge of preserving historical information while avoiding overfitting noise. It incorporates Legendre Polynomials projections to approximate historical information, uses Fourier projection to remove noise, and includes a low-rank approximation to speed up computation. [13].

LightTS

The light deep learning architecture for multivariate time series forecasting is based on simple MLP-based structures and utilizes two down-sampling strategies, continuous sampling, and interval sampling, to capture short-term local patterns and long-term temporal dependencies. [14].

PatchTST

It is based on two key components: patching and channel independence. Patching involves dividing the time series into subseries-level patches, used as input tokens to the Transformer. This allows the model to retain local semantic information, reduce computation and memory usage, and attend to a long history. Channel independence means that each channel contains a single univariate time series, and the embedding and Transformer weights are shared across all the series within the channel. [15].

Multi-scale Isometric Convolution Network (MICN)

This approach combines local and global modeling perspectives to predict long-term series. It uses down-sampled convolution to extract the local features and isometric convolution to capture global correlations. This multi-scale branch structure allows the modeling of different patterns in the time series, and the final prediction is obtained by combining the results of different branches [16].

DLinear

It combines Autoformer and FEDformer's decomposition schemes with linear layers. Using a moving average kernel, the model decomposes data into a trend component and a remainder component. After this, two linear layers with a single layer are applied to each component, and the final prediction is obtained by summing the two features. When there is a distinct trend in the data, DLinear improves the efficacy of a standard linear model by explicitly managing the trend. [130].

4.3 Data

The time horizon forecasting and the input features of the forecasting model are some of the critical factors that can influence the forecasting accuracy of the model. The input features for PV output power forecasting models would be the meteorological parameters like solar

Table 4.1 Detailed information of the PV system from DKASC

Item	Value
Array Rating	22.56 kW
Panel Rating	327 W
Number Of Panels	69
Panel Type	SunPower SPR-327NE
Inverter Size / Type	SMA STP 25000TL-30 , 20000TL-30
Installed	19 March 2016
Array Tilt/Azimuth	Tilt = 10, Azi = 0 (Solar North)

irradiance, temperature, wind speed, etc. [105]. In this work, we will use a database for the region of Yulara in Australia to compare different deep learning architectures.

The region is Yulara, a city located in the northern territory of Australia. This region is known for its hot desert climate. During the summer months, temperatures are scorching and can reach around 35°-40°C, while during the winter months, temperatures can reach around 20°-25°C. Yulara receives abundant sunshine throughout the year, contributing to its arid conditions. Yulara is located in a desert region with low precipitation, with an annual rainfall average of around 200-300 millimeters, with most of it occurring in the winter months. Yulara experiences moderate winds. The database for this region is downloaded from the website of The Desert Knowledge Australia Solar Centre (DKASC) [95]. The dataset contains the values of PV output (kW) and meteorological features: Wind Speed (m/s), Temperature in Celsius (°C), Global Horizontal Radiation (GHI in W/m²), Wind Direction (Degrees), and Air Pressure (mmHg) from December 2016 to December 2021, at Yulara, Australia. The time resolution of the dataset is 1 hour. Detailed information about the PV system is shown in table 4.1.

4.3.1 Data normalization

Generally, the input features have different scales, which may cause a delayed or unstable learning process. Before filling the input data into the deep learning models, it is recommended to normalize them. MinMaxScaler technique (4.6) was used to rescale the input data. It converts the input data to a range between 0 and 1. Normalization will improve the accuracy and minimize the regression error of the models.

$$x_{normalized} = \frac{x_{actual} - x_{min}}{x_{max} - x_{min}} \quad (4.6)$$

where $x_{normalized}$ is the normalized input data, x_{actual} is the actual input, x_{min} is the minimum value of the input data, and x_{max} is the maximum value of the input data.

4.3.2 Data splitting

Training and evaluation are two essential steps in machine learning projects. For this work, we used 43 825 samples in the database representing five years (from 21 December 2016 until 21 December 2021). The first four years of the database are used for the training phase, and the last year is used for the evaluation phase.

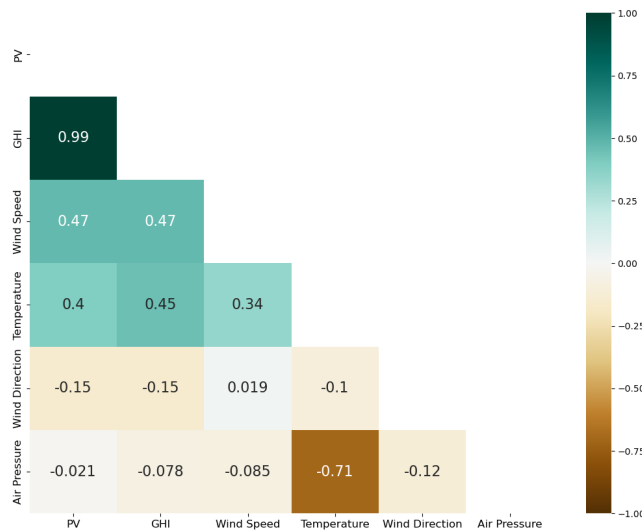


Figure 4.2 Heatmap representing Pearson correlation between the features for the region of Yulara in Australia.

4.3.3 Feature selection

The feature selection was performed using the Pearson correlation coefficient between the database features. The classification of the impact of these features on PV power forecasting designated the absolute value of correlation as very strong for values between 0.80 and 1.0, strong for values between 0.60 and 0.79, medium for values between 0.40 and 0.59, weak for values between 0.20 and 0.39, and very low for values between 0.00 and 0.19 [131].

As expected for the region of Yulara in Australia, we can see from figure 4.2 that the PV output power has a high correlation of 0.99 with the GHI. On the other hand, both temper-

ature and wind speed represent a medium correlation with the PV output power (0.4 and 0.48, respectively). At the same time, The wind direction and the air pressure have a very low correlation with the PV output power compared to the other features. Even if the PV output power is not well correlated with all meteorological parameters, we still have some reasons to keep them as inputs for the models. As we know, wind speed has a main role in heat dissipation and changing PV cell temperature. Thus the output cell voltage decrease if the PV cell temperature increase.

Table 4.2 Results of the dry season for the region of Yulara in Australia

Model	RMSE (kW)	MAE (kW)	nRMSE (%)	nMAE (%)	R^2
Pyraformer	1.2005	0.5230	5.9715	2.6016	0.9630
Informer	1.2857	0.5740	6.3954	2.8555	0.9576
MICN	1.3304	0.6026	6.6178	2.9974	0.9546
PatchTST	1.3433	0.6076	6.6822	3.0227	0.9537
Crossformer	1.4178	0.6579	7.0526	3.2728	0.9485
NonST	1.4239	0.6192	7.0829	3.0800	0.9480
FiLM	1.4380	0.6498	7.1531	3.2327	0.9470
DLinear	1.4545	0.7078	7.2350	3.5211	0.9458
FEDformer	1.5033	0.6915	7.4781	3.4400	0.9421
LightTS	1.5050	0.7314	7.4863	3.6382	0.9419
GRU	1.5793	0.7941	7.8557	3.9501	0.9361
Reformer	1.6678	0.6471	8.2963	3.2192	0.9287
LSTM	1.6776	0.8563	8.3449	4.2598	0.9279
Transformer	1.7148	0.6834	8.5298	3.3996	0.9246
ETSformer	2.2417	1.1917	11.1510	5.9277	0.8712
Autoformer	2.3157	1.1649	11.5190	5.7947	0.8626

4.4 Results and discussions

Table 4.2 and 4.3 comprehensively benchmark various deep learning models for PV output power forecasting in terms of their prediction accuracy for both dry and wet seasons. The analysis encompasses vital performance metrics such as RMSE, MAE, nRMSE, nMAE, R^2 .

Table 4.3 Results of the wet season for the region of Yulara in Australia

Model	RMSE (kW)	MAE (kW)	nRMSE (%)	nMAE (%)	R^2
Pyraformer	1.7597	0.8642	8.2957	4.0739	0.9346
Informer	1.9016	0.9074	8.9643	4.2779	0.9236
MICN	1.9915	0.9667	9.3882	4.5572	0.9162
PatchTST	2.0183	0.9872	9.5145	4.6540	0.9140
Crossformer	2.0321	1.0288	9.5797	4.8498	0.9128
NonST	2.0615	1.0117	9.7181	4.7693	0.9102
DLinear	2.0649	1.0625	9.7345	5.0089	0.9099
LightTS	2.1362	1.1025	10.0705	5.1973	0.9036
FiLM	2.1437	1.0690	10.1058	5.0395	0.9029
FEDformer	2.2309	1.1241	10.5170	5.2993	0.8949
GRU	2.2326	1.1280	10.5249	5.3175	0.8947
LSTM	2.2851	1.1977	10.7724	5.6463	0.8897
Reformer	2.3094	1.0799	10.8867	5.0910	0.8874
Transformer	2.3466	1.1033	11.0621	5.2010	0.8837
ETSformer	2.8153	1.5736	13.2717	7.4181	0.8326
Autoformer	2.9601	1.5672	13.9542	7.3881	0.8150

For the dry season (between April and September), the Pyraformer model stands out as the most accurate, scoring the lowest values for all error metrics RMSE, MAE, nRMSE, and nMAE (1.2005, 0.5230, 5.9715, 2.6016, respectively). In other words, the Pyraformer model achieved the closest forecasts to the actual values. In addition, it attains the highest R^2 value of 0.9630, suggesting that it can explain approximately 96.3% of the next 24 hours of PV output power during the dry season. On the other hand, recurrent neural networks such as GRU and LSTM perform poorly compared to the other deep learning models. The GRU and LSTM models have the highest error values across all error metrics, with LSTM landing the last place with error metrics RMSE, MAE, nRMSE, and nMAE equal to 1.6776, 0.8563, 8.3449, 4.2598, respectively. This indicates a less accurate forecast than the leading models, where GRU and LSTM can only explain approximately 93.6% and 92.7%, respectively, of the next 24 hours of Pv output power during the dry season.

We can say the same thing about the wet season (between October and March). the

Pyraformer model is the most accurate, scoring the lowest values for all error metrics RMSE, MAE, nRMSE, and nMAE (1.7597, 0.8642, 8.2957, 4.0739, respectively). Again, the Pyraformer model achieved the closest forecasts to the actual values. In addition, it attains the highest R^2 value of 0.9346, suggesting that it can explain approximately 93.4% of the next 24 hours of PV output power during the wet season. The same remark we can say about the GRU and LSTM models, they have the highest error values across all error metrics, with LSTM landing the last place with error metrics RMSE, MAE, nRMSE, and nMAE equal to 2.2851, 1.1977, 10.7724, 5.6463, respectively. This indicates a less accurate forecast than the leading models, where GRU and LSTM can only explain approximately 89.4% and 88.9%, respectively, of the next 24 hours of Pv output power during the wet season.

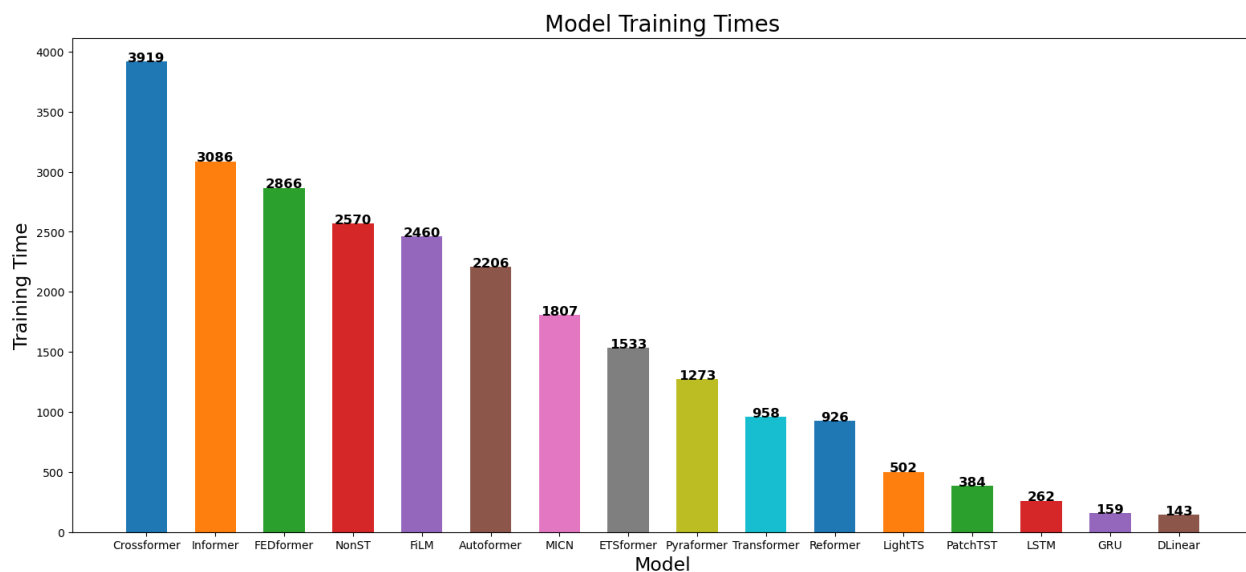


Figure 4.3 Training time for the deep learning models used.

Despite having lower error values than Pyraformer, the DLinear model is notable for its exceptional computational efficiency. From the figure 4.3, we can see that the Dlinear model has the minimum training time, measuring in at only 143 seconds. Despite this, its predictive accuracy outperforms the GRU and LSTM models, making it an attractive choice for situations with limited computational resources or time.

In conclusion, models such as Pyraformer, Informer, and DLinear demonstrate superior performance in terms of forecasting accuracy compared to recurrent neural networks. However, when considering computational efficiency, DLinear proves to be the fastest, even outpacing GRU and LSTM, suggesting a possible trade-off between prediction accuracy and computational efficiency across these models.

4.5 Limitation and challenges

Deep learning models, including those we evaluated in this work, primarily learn from patterns and trends in historical data to make future predictions. When forecasting PV output power, these models learn the relationship between historical PV output power, meteorological parameters (like temperature, humidity, wind speed, etc.), and future PV output power. Then they use these learned patterns to forecast future PV output power given future historical weather conditions. However, their performance can be limited in scenarios where weather conditions change drastically from one day to another, as these situations often deviate from the typical patterns the model has learned from the training data.

As we can see from figure 4.4, if the model has been primarily trained on data where sunny days are usually followed by similarly sunny days, it might struggle to accurately forecast the PV output power on a day that is sunny but is followed by a rainy or overcast day. This is because the change in weather conditions is an anomaly compared to the historical data, making it difficult for the model to forecast the outcome accurately. This is a common limitation of many machine learning and deep learning models. They excel at recognizing and exploiting patterns in the data they are trained on but can struggle with sudden, drastic changes that deviate from these learned patterns.

To improve the performance of the models in these scenarios, we need to incorporate more diverse data into the training set, including more instances of days with drastically changing weather conditions. Furthermore, investigating alternative model architectures or using additional categories of data (such as high-resolution weather forecasts) could help better the resiliency of the model to these significant day-to-day weather changes.

4.6 Conclusion

In conclusion, our study has thoroughly investigated the application of various state-of-the-art deep learning models for PV output power forecasting, focusing on one-day ahead forecasting using the past two days of meteorological parameters as inputs. Based on our error metrics analysis, we observed that the Pyraformer model has outperformed the other deep learning models tested and the traditional recurrent neural networks such as LSTM and GRU in terms of predictive accuracy. However, we also noted a common limitation across all the models. None of them could perform optimally when the weather drastically differed between two consecutive days. This underscores the need for incorporating high-resolution weather forecasts into these models. The inclusion of these forecasts serves as a potential avenue to improve their ability to handle abrupt changes in weather conditions that can

significantly affect PV output power. Additionally, there may be a need to include some clustering approaches. These could help in identifying the type of the day and modeling complex relationships in data, thus improving the accuracy of PV output power predictions.

4.7 Acknowledgment

We would like to express our gratitude to those who made this work possible. First, we would like to thank the Quebec government for financially supporting this work. We would also like to thank the Mila project team and Masen R&D team. Their contributions have greatly enriched the realization of this work.

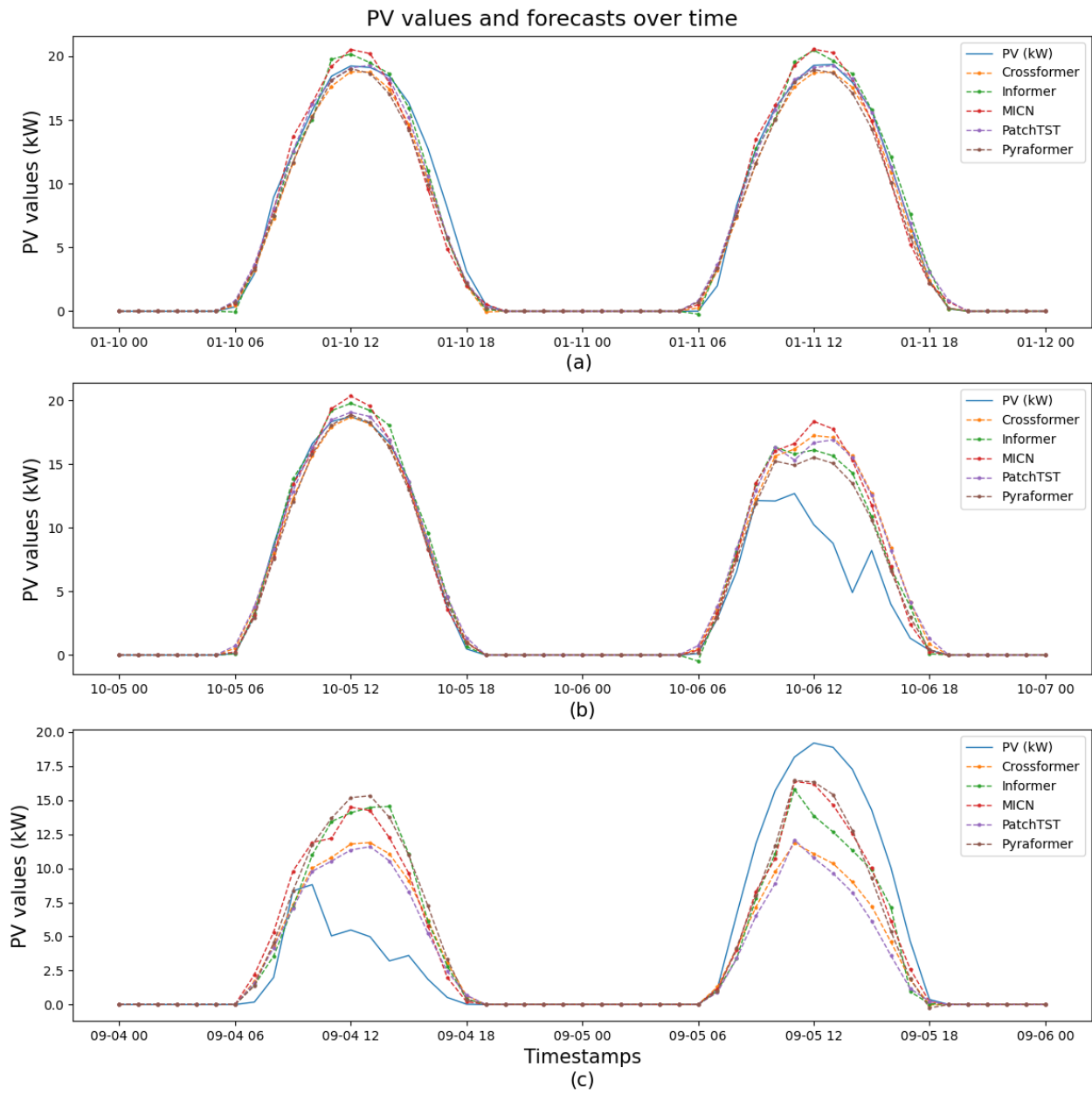


Figure 4.4 PV forecasting for the best five deep learning models used.

CHAPITRE 5 CONCLUSION ET RECOMMANDATIONS

Ce dernier chapitre clôture l'ensemble de ce mémoire de maîtrise dont l'objectif est de prédire le productible de l'énergie PV. Notre travail consiste à appliquer les modèles d'apprentissage profond sur des séries temporelles dans le cadre de la prédiction du productible de l'énergie PV. Cette problématique est d'actualité dans le contexte de la transition vers des ressources d'énergie renouvelables et durables.

5.1 Synthèse des travaux

La réalisation de cette étude a suivi plusieurs étapes qui ont contribué à l'accomplissement de notre objectif final qui est la prédiction du productible de l'énergie PV. La première phase de ce travail, se portait sur une revue de la littérature qui a été effectuée pour comprendre les différents modèles et approches utilisés dans la prédiction du productible de l'énergie PV. Cette revue nous a permis d'acquérir une compréhension solide des fondements théoriques et des méthodologies existantes.

Ensuite une méthodologie a été élaborée pour guider nos travaux. Nous avons commencé par identifier les différents modèles d'apprentissage profond comme les principales méthodes à explorer pour effectuer la prédiction du productible de l'énergie PV. Sur la base de cette méthodologie, nous avons collecté les données nécessaires pour la région de Yulara en Australie, telles que les données météorologiques (irradiance, température, vitesse du vent, etc.) et les mesures de production d'énergie PV, pour former et évaluer nos modèles de prédiction.

L'implémentation des modèles prédictifs a été une étape cruciale de notre travail. Nous avons utilisé des bibliothèques (Pytorch, Optuna, etc.) et des outils de programmation (Python) appropriés pour développer et entraîner les modèles, en tenant compte des spécificités de chaque modèle. Nous avons aussi préparé les données et nous avons configuré les paramètres des modèles.

Les résultats ont montré que des modèles d'apprentissage profond comme Pyraformer et Informer démontrent des performances supérieures en terme de précision de prédiction par rapport aux modèles LSTM et GRU, qui sont les plus utilisés dans la littérature. Les modèles Pyraformer et Informer arrivent à prédire en moyenne, respectivement, 94% et 93% du productible de l'énergie PV un jour à l'avance, contre 91% et 90% respectivement pour les modèles GRU et LSTM. En ce qui concerne l'efficacité de calcul, d'autres modèles comme DLinear s'avèrent être plus rapides, dépassant même les LSTM et GRU.

5.2 Limitations de la solution

Lors de la comparaison des modèles d'apprentissage profond, nous avons identifié certaines limitations inhérentes à ces modèles. Il est important de les prendre en compte pour une interprétation appropriée des résultats et pour orienter les futures recherches dans ce domaine.

Tout d'abord, l'une des principales limitations des modèles d'apprentissage profond est leur besoin en termes de données d'entraînement. Ces modèles nécessitent souvent de vastes ensembles de données pour atteindre leur plein potentiel de prédiction. Dans notre étude, bien que nous ayons utilisé des données disponibles, il est possible que la taille de l'ensemble de données n'ait pas été suffisamment grande pour capturer toutes les variations possibles dans le productible de l'énergie PV. Cela peut avoir un impact sur les performances et la généralisation des modèles.

Par ailleurs, un autre défi majeur des modèles d'apprentissage profond est la complexité de leurs architectures et de leurs paramètres. Ces modèles peuvent comporter de nombreux hyperparamètres qui doivent être réglés de manière appropriée pour obtenir les meilleures performances. Dans cette étude, nous avons effectué une optimisation des hyperparamètres, mais il est possible que d'autres combinaisons puissent conduire à de meilleures performances. De plus, la recherche de la meilleure architecture de modèle peut être un processus itératif, nécessitant des ressources de calcul considérables.

Enfin, il convient de mentionner que les modèles d'apprentissage profond peuvent être sensibles aux conditions spécifiques de l'ensemble de données et du contexte dans lequel ils sont appliqués. Autrement dit, les performances des modèles pourraient différer si des données provenant d'autres régions ou si d'autres périodes temporelles étaient utilisées. Il est donc important de considérer les particularités du contexte lors de l'application des modèles d'apprentissage profond à d'autres situations.

5.3 Améliorations futures

Parmi les améliorations à apporter dans les travaux futurs est d'augmenter la taille des données d'entraînement. Dans notre étude, nous avons utilisé des données disponibles, mais l'ajout de données supplémentaires provenant de différentes sources ou de périodes temporelles plus longues pourrait améliorer les performances des modèles.

Ensuite, l'optimisation des hyperparamètres peut être approfondie pour trouver les meilleures combinaisons pour chaque modèle. Dans notre étude, nous avons effectué une optimisation des hyperparamètres à l'aide de la bibliothèque Optuna, mais d'autres techniques, telles que la

recherche par grille ou l'optimisation bayésienne, pourraient être utilisées pour explorer plus efficacement l'espace des hyperparamètres et trouver les configurations optimales.

En résumé, des pistes d'amélioration pour les modèles d'apprentissage profond comparés dans ce travail incluent l'augmentation de la taille de l'ensemble de données et l'optimisation plus poussée des hyperparamètres. Ces améliorations potentielles peuvent contribuer à une meilleure performance prédictive, à une meilleure compréhension des modèles et à une application plus étendue dans le domaine de la prédiction du productible de l'énergie PV.

RÉFÉRENCES

- [1] M. G. Villalva, J. R. Gazoli et E. R. Filho, “Comprehensive approach to modeling and simulation of photovoltaic arrays,” *IEEE Transactions on Power Electronics*, vol. 24, n^o. 5, p. 1198–1208, 2009.
- [2] A. R. Jordehi, “Parameter estimation of solar photovoltaic (pv) cells : A review,” *Renewable and Sustainable Energy Reviews*, vol. 61, p. 354–371, 2016. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1364032116300016>
- [3] A. Bellahsen, “L’intelligence artificielle au service de l’optimisation de l’énergie électrique dans un réseau intelligent,” mémoire de maîtrise, Dép. de génie électrique, École Polytechnique de Montréal, Montréal, QC, 2020. [En ligne]. Disponible : <https://publications.polymtl.ca/5373/>
- [4] Colah’s blog, “Understanding lstm networks,” Aug 2015, consulté le 29 Juin 2023. [En ligne]. Disponible : <https://www.ncei.noaa.gov/products/weather-climatemodels/global%20forecasthttps://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- [5] A. Vaswani et et al., “Attention is all you need,” dans *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA, Dec 2017.
- [6] W. Wenxiao *et al.*, “Crossformer : A versatile vision transformer hinging on cross-scale attention,” arXiv preprint, 2021.
- [7] L. Y. *et al.*, “Non-stationary transformers : Exploring the stationarity in time series forecasting,” dans *36th Conference on Neural Information Processing Systems (NeurIPS 2022)*, 2022.
- [8] W. H. *et al.*, “Autoformer : Decomposition transformers with auto-correlation for long-term series forecasting,” dans *35th Conference on Neural Information Processing Systems (NeurIPS 2021)*, 2022.
- [9] W. Gerald *et al.*, “Etsformer : Exponential smoothing transformers for time-series forecasting,” Jun 2022.
- [10] Z. Tian *et al.*, “Fedformer : Frequency enhanced decomposed transformer for long-term series forecasting,” dans *Proceedings of the 39th International Conference on Machine Learning*, Jun 2022.
- [11] Z. Haoyi *et al.*, “Informer : Beyond efficient transformer for long sequence time-series forecasting,” Mar 2021.

- [12] L. Shizhan *et al.*, “Pyraformer : Low-complexity pyramidal attention for long-range time series modeling and forecasting,” dans *International Conference on Learning Representations 2022*, 2022.
- [13] Z. Tian *et al.*, “Film : Frequency improved legendre memory model for long-term time series forecasting,” dans *36th Conference on Neural Information Processing Systems (NeurIPS 2022)*, Sep 2022.
- [14] T. Zhang *et al.*, “Less is more : Fast multivariate time series forecasting with light sampling-oriented mlp structures,” *arXiv*, July 2022.
- [15] Y. Nie *et al.*, “A time series is worth 64 words : Long-term forecasting with transformers,” dans *International Conference on Learning Representations 2023*, 2023.
- [16] H. Wang *et al.*, “Micn : Multi-scale local and global context modeling for long-term series forecasting,” 2023. [En ligne]. Disponible : <https://openreview.net/pdf?id=zt53IDUR1U>
- [17] Intergovernmental Panel on Climate Change, “Global warming of 1.5°C,” IPCC, Rapport technique, 2018.
- [18] N. Watts *et al.*, “Health and climate change : policy responses to protect public health,” *The Lancet*, vol. 386, n°. 10006, p. 1861–1914, 2015.
- [19] Intergovernmental Panel on Climate Change, “Ipcc special report on the ocean and cryosphere in a changing climate,” Rapport technique, 2019. [En ligne]. Disponible : <https://www.ipcc.ch/srocc/>
- [20] U.S. Global Change Research Program, “Fourth national climate assessment,” Rapport technique, 2018. [En ligne]. Disponible : <https://nca2018.globalchange.gov/>
- [21] International Energy Agency, “Co2 emissions from fuel combustion,” Rapport technique, 2021. [En ligne]. Disponible : <https://www.iea.org/reports/co2-emissions-from-fuel-combustion-2021>
- [22] United Nations, “World population prospects 2019 : Highlights,” Rapport technique, 2019. [En ligne]. Disponible : https://population.un.org/wpp/Publications/Files/WPP2019_10KeyFindings.pdf
- [23] International Energy Agency, “World energy outlook 2020,” Rapport technique, 2020. [En ligne]. Disponible : <https://www.iea.org/reports/world-energy-outlook-2020>
- [24] U.S. Energy Information Administration, “Levelized cost and levelized avoided cost of new generation resources in the annual energy outlook 2020,” Rapport technique, 2020. [En ligne]. Disponible : https://www.eia.gov/outlooks/aeo/pdf/electricity_generation.pdf

- [25] International Renewable Energy Agency, “Renewable energy and jobs - annual review 2020,” Rapport technique, 2020. [En ligne]. Disponible : <https://www.irena.org/publications/2020/Jun/Renewable-Energy-and-Jobs-Annual-Review-2020>
- [26] National Renewable Energy Laboratory, “Benefits of renewable energy use,” 2021. [En ligne]. Disponible : <https://www.nrel.gov/research/re-benefits.html>
- [27] U.S. Department of Energy, “Solar energy technologies office,” 2021. [En ligne]. Disponible : <https://www.energy.gov/eere/solar/solar-energy-technologies-office>
- [28] International Energy Agency, “Energy access outlook 2017,” Rapport technique, 2017. [En ligne]. Disponible : <https://www.iea.org/reports/energy-access-outlook-2017>
- [29] Renewable Energy Policy Network for the 21st Century, “Renewables 2019 global status report,” Rapport technique, 2019. [En ligne]. Disponible : https://www.ren21.net/gsr-2019/chapters/chapter_02/chapter-02/
- [30] International Renewable Energy Agency, “Community energy : Accelerating the socio-economic benefits of renewable energy deployment,” Rapport technique, 2019. [En ligne]. Disponible : https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/Sep/IRENA_Community_Energy_2019.pdf
- [31] M. Houchati, M. H. Beitelmal et M. Khraisheh, “Predictive modeling for rooftop solar energy throughput : A machine learning-based optimization for building energy demand scheduling,” *Journal of Energy Resources Technology - Transactions of the ASME*, p. 1–15, 2021.
- [32] X. G. Agoua, R. Girard et G. Kariniotakis, “Short-term spatio-temporal forecasting of photovoltaic power production,” *IEEE Transactions on Sustainable Energy*, vol. 9, p. 538–546, 2018.
- [33] S. Wang *et al.*, “A new hybrid short-term interval forecasting of pv output power based on eemd-se-rvm,” *Energies*, 2019.
- [34] S. H. Nordahl, “Design of roof pv installation in oslo,” 2012. [En ligne]. Disponible : https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/257364/566421_FULLTEXT01.pdf?sequence=1
- [35] X. Chen *et al.*, “Response of u.s. west coast mountain snowpack to local sea surface temperature perturbations : Insights from numerical modeling and machine learning,” *Journal of Hydrometeorology*, vol. -1, p. 1045–1062, 2021.
- [36] K. Kurowski, M. Ciznicki et J. Węglarz, “Energy efficiency and performance modeling of stencil applications on manycore and gpu computing resources,” dans *2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID)*, Melbourne, VIC, Australia, 2020, p. 232–241.

- [37] S. Higgins, “Limitations to seasonal weather prediction and crop forecasting due to nonlinearity and model inadequacy,” Thèse de doctorat, London School of Economics and Political Science, 2015. [En ligne]. Disponible : <http://etheses.lse.ac.uk/id/eprint/3191>
- [38] G. Luís, J. Esteves et N. P. da Silva, “Energy forecasting using an ensemble of machine learning methods trained only with electricity data,” dans *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, 2020, p. 449–453.
- [39] L. M. Fraas, “Photovoltaics : Comparison of pv cell types,” 2015.
- [40] S. T. *et al.*, “Modelling and simulation of a photovoltaic cell for green instrumentation technology,” *Journal of Innovative Image Processing*, 2022.
- [41] B. Gyana Ranjan et M. Banaja, “Optimal frequency control of pv and fuel cell-based microgrid system,” *2021 1st International Conference on Power Electronics and Energy (ICPEE)*, p. 1–5, 2021.
- [42] V. Türkmenoglu et F. Güngör, “Matlab simulink & gui with pv cell simulation,” 2014.
- [43] S. Dash *et al.*, “Development of photovoltaic (pv) cell/module/array and non-uniform irradiance effect based on two-diode model by using pspice simulator,” dans *Proceedings of International Conference on Nascent Technologies in the Engineering Field (ICNTE 2015)*. IEEE, 2015, p. 1–6.
- [44] S. Vergura, “Scalable model of pv cell in variable environment condition based on the manufacturer datasheet for circuit simulation,” dans *Proceedings of IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC 2015)*. IEEE, 2015, p. 1481–1485.
- [45] R. Kadri *et al.*, “Modeling of the photovoltaic cell circuit parameters for optimum connection model and real-time emulator with partial shadow conditions,” *Energy*, vol. 42, p. 57–67, 2012.
- [46] N. Das, H. Wongsodihardjo et S. Islam, “Modeling of multi-junction photovoltaic cell using matlab/simulink to improve the conversion efficiency,” *Renewable Energy*, vol. 74, p. 917–924, 2015.
- [47] M. Awadallah et B. Venkatesh, “Estimation of pv module parameters from datasheet information using optimization techniques,” dans *Proceedings of IEEE International Conference on Industrial Technology (ICIT 2015)*. IEEE, 2015, p. 2777–2782.
- [48] V. J. Chin, Z. Salam et K. Ishaque, “Cell modelling and model parameters estimation techniques for photovoltaic simulator application : A review,” *Applied Energy*, vol. 154, p. 500–519, 2015. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0306261915006455>

- [49] H. Tsai, C. Tu et Y. Sui, “Development of generalized photovoltaic,” dans *WCECS 2008*, San Francisco, USA, October 2008.
- [50] K. Ishaque, “An improved modeling method to determine the model parameters of photovoltaic (pv) modules using differential evolution (de),” *Solar Energy*, vol. 85, n°. 9, p. 2349–2359, 2011. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0038092X11002374>
- [51] P. Suskis et I. Galkin, “Enhanced photovoltaic panel model for matlab-simulink environment considering solar cell junction capacitance,” dans *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, 2013, p. 1613–1618.
- [52] N. Kensuke *et al.*, “Analysis of multicrystalline silicon solar cells by modified 3-diode equivalent circuit model taking leakage current through periphery into consideration,” *Solar Energy Materials and Solar Cells*, vol. 91, n°. 13, p. 1222–1227, 2007. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0927024807001493>
- [53] K. Ken-ichi et M. Hiroyuki, “New two-diode model for detailed analysis of multicrystalline silicon solar cells,” *Japanese Journal of Applied Physics*, vol. 44, n°. 12R, p. 8314, dec 2005. [En ligne]. Disponible : <https://dx.doi.org/10.1143/JJAP.44.8314>
- [54] L. Matthew P. *et al.*, “Analytical drift-diffusion modeling of gaas solar cells incorporating a back mirror,” dans *2013 IEEE 39th Photovoltaic Specialists Conference (PVSC)*, 2013, p. 1063–1068.
- [55] S. Jing Jun, L. Kay-Soon et G. Shu Ting, “Multi-dimension diode photovoltaic (pv) model for different pv cell technologies,” dans *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*, 2014, p. 2496–2501.
- [56] P. A. Browne *et al.*, “Weakly coupled ocean–atmosphere data assimilation in the ecmwf nwp system,” *Remote Sensing*, vol. 11, n°. 234, 2019. [En ligne]. Disponible : <https://www.mdpi.com/2072-4292/11/3/234>
- [57] V. Bjercknes, “The problem of weather prediction, considered from the viewpoints of mechanics and physics,” *Meteorologische Zeitschrift*, p. 663–667, Dec 2009.
- [58] S. A. H. M. *et al.*, “Synthetic series of electricity generation through a photovoltaic system by using different panel temperature models,” *ResearchGate*, Sept 2021.
- [59] J. Cha *et al.*, “Investigation of applicability of impact factors to estimate solar irradiance : Comparative analysis using machine learning algorithms,” *Applied Sciences*, 2021.

- [60] X. Jiao *et al.*, “A graph neural network based deep learning predictor for spatio-temporal group solar irradiance forecasting,” *IEEE Transactions on Industrial Informatics*, vol. 18, n^o. 9, p. 6142–6149, 2022.
- [61] F. Mohammed, M. Bahgat et S. e. a. Elmasry, “Design of a maximum power point tracking-based pid controller for dc converter of stand-alone pv system,” *Journal of Electrical Systems and Inf Technol*, vol. 9, n^o. 9, 2022.
- [62] M. Q. AlOdat, “Experimental study of temperature influence on the performance of pv/t cell under jordan climate conditions,” *Journal of Ecological Engineering*, 2022.
- [63] National Centers for Environmental Information (NCEI), “Global Forecast System (GFS),” consulté le 23 Juin 2023. [En ligne]. Disponible : <https://www.ncei.noaa.gov/products/weather-climatemodels/global%20forecast>
- [64] S. Prakash *et al.*, “An early assessment of medium range monsoon precipitation forecasts from the latest high-resolution ncep-gfs (t1534) model over south asia,” *Pure and Applied Geophysics*, vol. 173, n^o. 7, p. 2215–2225, 2016. [En ligne]. Disponible : <https://doi.org/10.1007/s00024-016-1248-5>
- [65] Z. Ma *et al.*, “Spin-up characteristics with three types of initial fields and the restart effects on forecast accuracy in the grapes global forecast system,” *Geoscientific Model Development*, vol. 14, n^o. 1, p. 205–221, 2021. [En ligne]. Disponible : <https://gmd.copernicus.org/articles/14/205/2021/>
- [66] M. Quijal-Zamorano *et al.*, “Forecast of temperature-attributable mortality at lead times of up to 15 days for a very large ensemble of european regions,” dans *EGU General Assembly 2021, online*, Apr 2021, p. EGU21–4107. [En ligne]. Disponible : <https://doi.org/10.5194/egusphere-egu21-4107>
- [67] C. Ozen et A. Deniz, “A comprehensive country-based day-ahead wind power generation forecast model by coupling numerical weather prediction data and catboost with feature selection methods for turkey,” *Wind Engineering*, vol. 46, n^o. 5, p. 1359–1388, 2022.
- [68] P. Bauer *et al.*, “Aspects of ecmwf model performance in polar areas,” *Quarterly Journal of the Royal Meteorological Society*, vol. 142, n^o. 694, p. 583–596, 2014.
- [69] J. G. Powers *et al.*, “The weather research and forecasting model : Overview, system efforts, and future directions,” *Bulletin of the American Meteorological Society*, vol. 98, n^o. 8, p. 1717–1737, 2017. [En ligne]. Disponible : <https://journals.ametsoc.org/view/journals/bams/98/8/bams-d-15-00308.1.xml>
- [70] T. R. Lee *et al.*, “Using the u.s. climate reference network to identify biases in near-and subsurface meteorological fields in the high-resolution rapid refresh (hrrr) weather

- prediction model,” *Weather and Forecasting*, vol. 38, n°. 6, p. 879–900, 2023. [En ligne]. Disponible : <https://journals.ametsoc.org/view/journals/wefo/38/6/WAF-D-22-0213.1.xml>
- [71] T. Sarah A. *et al.*, “Differentiating freezing drizzle and freezing rain in hrrr model forecasts,” *Weather and Forecasting*, vol. 36, n°. 4, p. 1237–1251, 2021. [En ligne]. Disponible : <https://journals.ametsoc.org/view/journals/wefo/36/4/WAF-D-20-0138.1.xml>
- [72] L. Francis M. *et al.*, “Improved ecmwf forecasts of direct normal irradiance : A tool for better operational strategies in concentrating solar power plants,” *Renewable Energy*, vol. 163, p. 755–771, 2021. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0960148120313859>
- [73] F. Ján *et al.*, “Ecmwf short-term prediction accuracy improvement by deep learning,” *Scientific Reports*, vol. 12, n°. 1, p. 7898, 2022. [En ligne]. Disponible : <https://doi.org/10.1038/s41598-022-11936-9>
- [74] A. R. *et al.*, “A review and evaluation of the state-of-the-art in pv solar power forecasting : Techniques and optimization,” *Renewable and Sustainable Energy Reviews*, vol. 124, p. 109792, 2020. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1364032120300885>
- [75] A. Alsharif *et al.*, “Review of ml and automl solutions to forecast time-series data,” *Archives of Computational Methods in Engineering*, vol. 29, n°. 7, p. 5297–5311, 2022. [En ligne]. Disponible : <https://doi.org/10.1007/s11831-022-09765-0>
- [76] A. L. Samuel, “Some studies in machine learning using the game of checkers,” *IBM Journal of Research and Development*, vol. 3, n°. 3, p. 210–229, 1959.
- [77] C. Po-Han *et al.*, “Forecasting of solar photovoltaic system power generation using wavelet decomposition and bias-compensated random forest,” dans *2017 Ninth Annual IEEE Green Technologies Conference (GreenTech)*, 2017, p. 260–266.
- [78] A. Zendehboudi, M. Baseer et R. Saidur, “Application of support vector machine models for forecasting solar and wind energy resources : A review,” *Journal of Cleaner Production*, vol. 199, p. 272–285, 2018. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S095965261832153X>
- [79] J. Singh et R. Banerjee, “A study on single and multi-layer perceptron neural network,” dans *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, 2019, p. 35–40.
- [80] D. Shiv Ram, S. Satish Kumar et C. Bidyut Baran, “Activation functions in deep learning : A comprehensive survey and benchmark,” *Neurocomputing*, vol. 503, p.

- 92–108, 2022. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0925231222008426>
- [81] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Mathematics of Control, Signals, and Systems*, vol. 2, n^o. 4, p. 303–314, 1989.
- [82] L. Moshe *et al.*, “Multilayer feedforward networks with a nonpolynomial activation function can approximate any function,” *Neural Networks*, vol. 6, n^o. 6, p. 861–867, 1993. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0893608005801315>
- [83] S.-A. Chen *et al.*, “Tsmixer : An all-mlp architecture for time series forecasting,” *arXiv*, April 2023.
- [84] M. Kaur et A. Mohta, “A review of deep learning with recurrent neural network,” dans *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2019, p. 460–465.
- [85] I. Goodfellow, Y. Bengio et A. Courville, *Deep Learning*. MIT Press, 2016.
- [86] R. Pascanu, T. Mikolov et Y. Bengio, “On the difficulty of training recurrent neural networks,” dans *International Conference on Machine Learning*, 2013, p. 1310–1318.
- [87] S. Hochreiter et J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, n^o. 8, p. 1735–1780, 1997.
- [88] J. Chung *et al.*, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” *arXiv*, Dec 2014, accessed : May 27, 2023. [En ligne]. Disponible : <http://arxiv.org/abs/1412.3555>
- [89] R. A. Ramadhan *et al.*, “Comparison of physical and machine learning models for estimating solar irradiance and photovoltaic power,” *Renewable Energy*, vol. 178, p. 1006–1019, 2021. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0960148121009459>
- [90] A. Mellit, A. M. Pavan et V. Lughi, “Deep learning neural networks for short-term photovoltaic power forecasting,” *Renewable Energy*, vol. 172, p. 276–288, 2021. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0960148121003475>
- [91] A. Erduman, “A smart short-term solar power output prediction by artificial neural network,” *Electrical Engineering*, vol. 102, p. 1441–1449, 2020. [En ligne]. Disponible : <https://doi.org/10.1007/s00202-020-00971-2>
- [92] Y. Ceyhun et A. Hakan, “A kernel extreme learning machine-based neural network to forecast very short-term power output of an on-grid photovoltaic power plant,” *Energy*

- Sources, Part A : Recovery, Utilization, and Environmental Effects*, vol. 43, n^o. 4, p. 395–412, 2021.
- [93] W. Huaizhi *et al.*, “Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network,” *Energy Conversion and Management*, vol. 153, p. 409–422, 2017.
- [94] M. Paras *et al.*, “Forecasting power output of solar photovoltaic system using wavelet transform and artificial intelligence techniques,” *Procedia Computer Science*, vol. 12, p. 332–337, 2012, complex Adaptive Systems 2012. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1877050912006710>
- [95] “Desert Knowledge Australia Centre. Download Data : Yulara Solar System.” consulté le 03 Septembre 2022. [En ligne]. Disponible : <https://dkasolarcentre.com.au/download?location=yulara>
- [96] R. Zhu, W. Guo et X. Gong, “Short-term photovoltaic power output prediction based on k-fold cross-validation and an ensemble model,” *Energies*, vol. 12, n^o. 7, 2019. [En ligne]. Disponible : <https://www.mdpi.com/1996-1073/12/7/1220>
- [97] T. Akiba *et al.*, “Optuna : A next-generation hyperparameter optimization framework,” dans *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019, p. 2623–2631.
- [98] K. N., K. L. et L. A., “Reformer : The efficient transformer,” dans *International Conference on Learning Representations 2020*, 2020.
- [99] I.-Y. Joo et D.-H. Choi, “Distributed optimization framework for energy management of multiple smart homes with distributed energy resources,” *IEEE Access*, vol. 5, p. 15 551–15 560, Aug 2017.
- [100] G. M. Wilson et et al., “The 2020 photovoltaic technologies roadmap,” *Journal of Physics D : Applied Physics*, vol. 53, p. 493001, 2020.
- [101] IEA, “Renewable energy market update outlook for 2022 and 2023,” France, May 2022.
- [102] —, “Renewable energy market update outlook for 2023 and 2024,” France, June 2023.
- [103] P. Esa, H. Hannele et K. Tiina, “Inter-sectoral effects of high renewable energy share in global energy system,” *Renewable Energy*, vol. 136, p. 1119–1129, 2019. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S096014811831156X>
- [104] L. Gangqiang *et al.*, “Recurrent neural networks based photovoltaic power forecasting approach,” *Energies*, vol. 12, n^o. 13, p. 2538, 2019. [En ligne]. Disponible : <https://doi.org/10.3390/en12132538>

- [105] R. Ahmed *et al.*, “A review and evaluation of the state-of-the-art in pv solar power forecasting : Techniques and optimization,” *Renewable and Sustainable Energy Reviews*, vol. 124, p. 109792, 2020. [En ligne]. Disponible : <https://doi.org/10.1016/j.rser.2020.109792>
- [106] D. Utpal Kumar *et al.*, “Forecasting of photovoltaic power generation and model optimization : A review,” *Renewable and Sustainable Energy Reviews*, vol. 81, p. 912–928, 2018. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1364032117311620>
- [107] B. Manoja Kumar, M. Irani et N. Niranjana, “Solar photovoltaic power forecasting using optimized modified extreme learning machine technique,” *Engineering Science and Technology, an International Journal*, vol. 21, n^o. 3, 2018.
- [108] W. Xiaoyang *et al.*, “Comparative study of machine learning approaches for predicting short-term photovoltaic power output based on weather type classification,” *Energy*, vol. 240, p. 122733, 2022. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0360544221029820>
- [109] W. Huaizhi *et al.*, “A review of deep learning for renewable energy forecasting,” *Energy Conversion and Management*, vol. 198, 2019.
- [110] N. H. *et al.*, “Univariate time series prediction of solar power using a hybrid wavelet-arma-narx prediction method,” dans *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, 2016, p. 1–5.
- [111] W. Lifeng *et al.*, “Using a novel multi-variable grey model to forecast the electricity consumption of shandong province in china,” *Energy*, vol. 157, p. 327–335, 2018. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0360544218309812>
- [112] M. Raul V.A. *et al.*, “Estimating photovoltaic power generation : Performance analysis of artificial neural networks, support vector machine and kalman filter,” *Electric Power Systems Research*, vol. 143, p. 643–656, 2017. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0378779616304576>
- [113] B. Antonio *et al.*, “A bayesian method for short-term probabilistic forecasting of photovoltaic generation in smart grid operation and control,” *Energies*, vol. 6, n^o. 2, p. 733–747, 2013.
- [114] Y.-z. Li, R. Luan et J.-c. Niu, “Forecast of power generation for grid-connected photovoltaic system based on grey model and markov chain,” dans *2008 3rd IEEE Conference on Industrial Electronics and Applications*, 2008, p. 1729–1733.

- [115] H. Jianming *et al.*, “Research and application of a hybrid model based on meta learning strategy for wind power deterministic and probabilistic forecasting,” *Energy Conversion and Management*, vol. 173, p. 197–209, 2018. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S019689041830788X>
- [116] S. Sobri, S. Koochi-Kamali et N. Abd. Rahim, “Solar photovoltaic generation forecasting methods : A review,” *Energy Conversion and Management*, vol. 156, p. 459–497, 2018. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0196890417310622>
- [117] C. Chi Wai, B. Serge et K. Jan, “Cloud motion and stability estimation for intra-hour solar forecasting,” *Solar Energy*, vol. 115, p. 645–655, 2015. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0038092X15001565>
- [118] F.-J. L. Alfredo *et al.*, “Short-term power forecasting system for photovoltaic plants,” *Renewable Energy*, vol. 44, p. 311–317, 2012. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0960148112001516>
- [119] L. Francisco J.L. *et al.*, “Forecast for surface solar irradiance at the brazilian northeastern region using nwp model and artificial neural networks,” *Renewable Energy*, vol. 87, p. 807–818, 2016. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0960148115304249>
- [120] S. Yu *et al.*, “Resilient wide-area damping control using grhdp to tolerate communication failures,” *IEEE Transactions on Smart Grid*, vol. 10, n^o. 3, p. 2547–2557, 2019.
- [121] A. Tuohy *et al.*, “Solar forecasting : Methods, challenges, and performance,” *IEEE Power and Energy Magazine*, vol. 13, n^o. 6, p. 50–59, 2015.
- [122] A.-M. J., B. F.J. et P. C., “Solar irradiance forecasting at one-minute intervals for different sky conditions using sky camera images,” *Energy Conversion and Management*, vol. 105, p. 1166–1177, 2015. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0196890415008365>
- [123] A.-M. J. et B. F.J., “Solar radiation forecasting in the short- and medium-term under all sky conditions,” *Energy*, vol. 83, p. 387–393, 2015. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0360544215001917>
- [124] W. Huaizhi *et al.*, “Taxonomy research of artificial intelligence for deterministic solar power forecasting,” *Energy Conversion and Management*, vol. 214, p. 112909, 2020. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0196890420304477>
- [125] G. Li *et al.*, “Neural networks-aided insider attack detection for the average consensus algorithm,” *IEEE Access*, vol. 8, p. 51 871–51 883, 2020.

- [126] Y. Xiaofeng *et al.*, “Hierarchical quality-relevant feature representation for soft sensor modeling : A novel deep learning strategy,” *IEEE Transactions on Industrial Informatics*, vol. 16, p. 3721–3730, 2020.
- [127] W. Qiao *et al.*, “A novel hybrid prediction model for hourly gas consumption in supply side based on improved whale optimization algorithm and relevance vector machine,” *IEEE Access*, vol. 7, p. 88 218–88 230, 2019.
- [128] T. Yang, B. Li et Q. Xun, “Lstm-attention-embedding model-based day-ahead prediction of photovoltaic power output using bayesian optimization,” *IEEE Access*, vol. 7, p. 171 471–171 484, 2019.
- [129] M. M.O. *et al.*, “Design of experiments using artificial neural network ensemble for photovoltaic generation forecasting,” *Renewable and Sustainable Energy Reviews*, vol. 135, p. 110450, 2021. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1364032120307371>
- [130] A. Zeng *et al.*, “Are transformers effective for time series forecasting ?” Aug 2022.
- [131] A. Razin *et al.*, “Computationally expedient photovoltaic power forecasting : A lstm ensemble method augmented with adaptive weighting and data segmentation technique,” *Energy Conversion and Management*, vol. 258, p. 115563, 2022. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0196890422003594>