UQAC
Université du Québec
à Chicoutimi

MODÈLES PRÉDICTIFS DES PERTES ÉLECTRIQUES SUR LES LIGNES DE TRANSPORT D'ÉNERGIE

PAR MOISE TSOMANYA KUMAVI

MÉMOIRE PRÉSENTÉ À L'UNIVERSITÉ DU QUÉBEC À CHICOUTIMI EN VUE DE L'OBTENTION DU GRADE DE MAÎTRISE ÈS SCIENCES (M. SC.) EN INFORMATIQUE

QUÉBEC, CANADA

© MOISE TSOMANYA KUMAVI, 2026

REMERCIEMENTS

Je tiens tout d'abord à exprimer ma profonde gratitude à mon directeur de recherche, M. Kevin Bouchard, pour son encadrement, sa disponibilité, ses conseils et la confiance qu'il m'a accordée tout au long de ce travail. Son accompagnement a été essentiel dans l'orientation scientifique de ce mémoire et dans la réalisation de ce projet de recherche.

Je remercie également mon co-directeur, M. Bruno Bouchard, pour ses recommandations, son soutien et ses remarques constructives, qui ont contribué à améliorer la qualité de ce travail. Mes remerciements s'adressent aussi à l'ensemble du corps professoral du Département d'informatique et de mathématique (DIM) de l'Université du Québec à Chicoutimi, pour la qualité de l'enseignement reçu et pour l'environnement académique favorable à mon développement scientifique et professionnel.

Je souhaite également remercier mes gestionnaires chez Hydro-Québec, M. Alexandre Pitre et M. Zakaria Sadeq, pour leur encadrement, leur confiance et l'opportunité de travailler sur un projet concret lié aux enjeux réels du réseau de transport électrique. J'adresse aussi mes sincères remerciements à M. Julien Witty pour ses conseils, son expertise en intelligence artificielle et son appui dans la réflexion et l'implémentation de l'architecture des modèles développés dans ce mémoire.

Je remercie particulièrement mon grand frère, Komlan Mawuli Adje, pour son soutien, ses conseils et ses encouragements constants tout au long de mon parcours. Sa présence et son accompagnement ont été d'une grande importance dans les différentes étapes de ma formation.

Enfin, je remercie ma famille, mes amis, mes collègues ainsi que toutes les personnes qui, de près ou de loin, ont contribué à la réussite de mon parcours. Leur soutien, leurs encouragements et leur confiance ont été une source précieuse de motivation tout au long de cette maîtrise.

RÉSUMÉ

Les pertes d'énergie sur les réseaux de transport constituent un enjeu majeur pour l'efficacité et la fiabilité des systèmes électriques, particulièrement dans les régions disposant de réseaux étendus et soumis à des conditions environnementales extrêmes. Au Québec, la société d'État Hydro-Québec (HQ) exploite le plus grand réseau hydroélectrique d'Amérique du Nord. Ce réseau se caractérise par des lignes à très haute tension, une forte variabilité saisonnière et des conditions opérationnelles complexes qui rendent la prédiction précise des pertes particulièrement difficile. Les modèles déterministes classiques, basés sur les lois de la physique, peinent à représenter la complexité du réseau en conditions réelles, notamment face à la dynamique des charges et à la diversité des sources de données. Il devient donc nécessaire de recourir à des approches de modélisation plus avancées, capables d'intégrer la complexité des réseaux électriques modernes. Dans ce mémoire, nous proposons une approche de prédiction des pertes techniques à granularité fine, fondée sur une modélisation ligne par ligne et sur la comparaison de plusieurs familles de modèles. Cette étude inclut des méthodes statistiques (ARIMA et Prophet), des algorithmes d'apprentissage par ensemble (Random Forest et XGBoost), ainsi qu'un modèle LSTM doté d'un mécanisme d'attention. Le jeu de données construit combine des variables électriques, des données météorologiques associées aux postes connectés aux lignes, ainsi que des caractéristiques calendaires. Il couvre vingt ans d'historique de pertes de transport, fournies grâce à la collaboration entre l'équipe CYNERGIA et HQ, permettant ainsi d'évaluer ces modèles dans un contexte industriel réel. Les résultats indiquent, dans le cadre expérimental étudié, que le modèle LSTM avec attention obtient les meilleures performances parmi les modèles comparés, avec un RMSE de 0.064 et un R^2 de 0.97. Ce travail met en évidence le potentiel des architectures récurrentes attentionnelles pour affiner la prédiction des pertes électriques et soutenir la planification opérationnelle des réseaux à grande échelle.

L’IA générative a été utilisée pour des suggestions ou des corrections mineures (orthographe, grammaire), ainsi que pour l’amélioration d’éléments visuels, tels que l’ajout d’un titre ou d’une légende. Niveau d’utilisation de l’intelligence artificielle.



TABLE DES MATIÈRES

Table des matières . . . . . **iv**

Liste des tableaux . . . . . **vi**

Liste des figures . . . . . **vii**

I Introduction **1**

II Revue de la littérature **11**

2.1 Fondements physiques des pertes techniques dans les réseaux de transport . . 12

2.2 Approches statistiques et modèles de séries temporelles 16

2.3 Méthodes d'apprentissage automatique pour la prédiction des pertes 19

2.4 Apprentissage profond et modèles séquentiels 23

2.5 Granularité des approches de prédiction des pertes 29

2.6 Disponibilité des données réelles et limites des approches simulées 33

2.7 Synthèse des lacunes et positionnement du mémoire 36

III Méthodologie pour l'analyse comparative de plusieurs modèles pour la prédiction des pertes d'énergie **39**

3.1 Architecture générale du pipeline 40

3.2 Formulation du problème de prédiction 42

3.3 Sources de données et construction de la variable cible 43
3.4 Fusion, nettoyage et préparation des données 47
3.5 Analyse exploratoire des données 49
3.6 Découpage temporel des données et stratégie de validation 50
3.7 Modèles développés et comparés 51
3.8 Fonction de coût, optimisation et arrêt anticipé 55
3.9 Évaluation des performances et diagnostics 56
3.10 Organisation expérimentale et suivi des modèles 57
3.11 Synthèse de la méthodologie 59

IV Résultats et discussion 60
4.1 Jeu de données et protocole expérimental 61
4.2 Comparaison globale des modèles 62
4.3 Résultats des modèles statistiques 63
4.4 Résultats des modèles d'apprentissage automatique 64
4.5 Résultats du modèle LSTM avec attention 65
4.6 Discussion comparative 67
4.7 Limites observées 68
4.8 Synthèse des résultats 69

V CONCLUSION ET PERSPECTIVES 70

Bibliographie 73

LISTE DES TABLEAUX

3.1 Configuration de base du modèle LSTM avec attention. 55

3.2 Principaux artefacts sauvegardés pour assurer la reproductibilité des expérimentations. 58

4.1 Comparaison des performances des modèles sur l'ensemble de test. 62

LISTE DES FIGURES

1.1 Lignes de transport et centrales électriques d'Hydro-Québec (Source : `https://www.hydroquebec.com/transenergie/fr/`) 4

2.1 Influence du courant et des conditions locales sur les pertes techniques d'une ligne de transport. 14

2.2 Principe de décomposition additive utilisé par Prophet. 19

2.3 Cadre général d'un modèle supervisé pour la prédiction des pertes techniques. 21

2.4 Architecture récurrente LSTM couplée à un mécanisme d'attention temporelle pour la prédiction des pertes. 24

2.5 Architecture générale du modèle Transformer et décomposition du mécanisme d'auto-attention (adapté de Vaswani et al. [76]). 28

2.6 Comparaison entre une approche globale et une approche ligne par ligne pour la prédiction des pertes techniques. 32

2.7 Principe d'une séparation chronologique des données pour la validation des modèles de séries temporelles. 36

3.1 Architecture générale du pipeline de prédiction des pertes techniques. 41

3.2 Évolution conjointe normalisée de la température et des pertes pour une ligne donnée. 46

3.3 Distribution des pertes en fonction de l'heure de la journée. 49

3.4 Variation mensuelle moyenne des pertes pour différentes années. 50

4.1 Évolution des pertes d'entraînement et de validation du modèle LSTM avec attention. 66

4.2 Analyse des résidus, comparaison entre valeurs réelles et prédites, et distribution des erreurs absolues pour le modèle LSTM avec attention. 67

LISTE DES ABRÉVIATIONS

ARIMA AutoRegressive Integrated Moving Average
CCR Centre de Conduite du Réseau
CYNERGIA Cybersécurité Énergie Intelligence Artificielle
ECMWF European Centre for Medium-Range Weather Forecasts
EMS Energy Management System
ERA5 ECMWF Reanalysis v5
HQ Hydro-Québec
HRDPS High Resolution Deterministic Prediction System
IA Intelligence Artificielle
KNN K-Nearest Neighbors
kV Kilovolt
LIME Local Interpretable Model-agnostic Explanations
LSTM Long Short-Term Memory
MAE Mean absolute error
MAPE Mean absolute percentage error
ML Apprentissage Automatique
MSE Mean squared error
RMSE Root mean squared error
RNN Recurrent Neural Network
SARIMA Seasonal AutoRegressive Integrated Moving Average
SCADA Supervision Contrôle et Acquisition de Données
SHAP SHapley Additive exPlanations
sMAPE Symmetric Mean absolute percentage error
TFT Temporal Fusion Transformer
VAR Vector Autoregression
XGBoost eXtreme Gradient Boosting

CHAPITRE I

INTRODUCTION

Dans un contexte mondial de transition énergétique, où les réseaux électriques doivent intégrer une proportion croissante d'énergies renouvelables, acheminer l'énergie sur de longues distances, répondre à une demande en constante évolution et maintenir un haut niveau de fiabilité ; la maîtrise des flux d'énergie constitue un enjeu stratégique majeur pour assurer l'équilibre, la stabilité et l'efficacité des systèmes électriques modernes [38, 36].

Le Centre de Conduite du Réseau (CCR), chargé d'assurer en temps réel l'équilibre entre la production et la consommation d'électricité sur l'ensemble du réseau, joue un rôle essentiel dans cette dynamique [15]. Cependant, la complexité croissante des systèmes électriques, combinée à la variabilité des sources renouvelables, aux fluctuations de la demande et à l'évolution des conditions d'exploitation, rend la gestion optimale de ces flux de plus en plus difficile.

Le transport d'énergie électrique sur de longues distances constitue un élément central des systèmes énergétiques modernes. Il assure non seulement l'acheminement de l'électricité vers les centres de consommation urbains et industriels, mais contribue également à la stabilité et à la résilience des réseaux grâce aux interconnexions régionales et transfrontalières [4]. Cependant, ce transport massif d'énergie s'accompagne d'un défi technique majeur : les pertes

électriques en ligne. Celles-ci se manifestent principalement sous forme de chaleur par effet Joule, en raison de la résistance des conducteurs, mais peuvent également inclure des pertes électromagnétiques, des courants de fuite ainsi que des effets de couronne sur les lignes à haute tension. L'existence de ces pertes impose une contrainte importante sur la planification de la production, car leur anticipation précise est nécessaire pour ajuster la puissance injectée sur le réseau et répondre adéquatement à la demande.

Dans la littérature, on distingue généralement deux grandes catégories de pertes d'énergie dans les réseaux électriques. La première correspond aux pertes techniques, qui sont directement liées aux caractéristiques physiques du réseau, notamment la résistance des conducteurs, les courants de fuite ainsi que les effets inductifs et capacitifs inhérents aux lignes à haute tension. Ces pertes dépendent, entre autres, des conditions de charge, de la longueur des lignes, du niveau de tension, de la température des conducteurs et de l'efficacité des infrastructures de transport [10]. La seconde catégorie regroupe les pertes non techniques, qui proviennent plutôt de facteurs externes au comportement physique du réseau, tels que les imprécisions de mesure, les erreurs humaines, les vols d'électricité ou encore les défaillances des systèmes de comptage [14, 64, 62, 59].

Dans le cadre de cette étude, l'attention est portée exclusivement sur les pertes techniques, lesquelles représentent la composante la plus directement liée aux caractéristiques physiques du réseau de transport et à ses conditions d'exploitation. Ces pertes constituent une part significative de l'électricité injectée sur le réseau [46, 61], réduisant ainsi l'efficacité énergétique globale du système électrique [9, 68]. Leur prédiction constitue donc un enjeu important pour améliorer la planification énergétique, optimiser l'exploitation du réseau et limiter les écarts entre l'énergie produite et l'énergie effectivement disponible après transport.

Au Canada, la gestion des pertes énergétiques revêt une importance stratégique en raison de l'immensité du territoire et des très longues distances de transport de l'électricité. Plus spécifiquement, au Québec, la société d'État Hydro-Québec exploite l'un des réseaux électriques les plus étendus d'Amérique du Nord, comprenant 62 centrales hydroélectriques, plus de 34 000 kilomètres de lignes à haute tension ainsi que 536 postes de transformation (voir figure 1.1). Le réseau est également relié aux réseaux voisins du Canada et des États-Unis par 17 interconnexions, offrant une capacité maximale d'importation de 11 150 mégawatts [1]. HQ figure ainsi parmi les plus grands producteurs mondiaux d'hydroélectricité [2].

Le réseau de transport d'HQ se distingue notamment par ses lignes à très haute tension, dont plusieurs à 735 kV, par une forte variabilité saisonnière de la demande énergétique ainsi que par des conditions environnementales parfois extrêmes, incluant le froid intense, les épisodes de givre et les vents importants. Ces facteurs influencent directement le comportement électrique des lignes de transport, notamment par la variation de la résistance des conducteurs, les fluctuations de charge en période hivernale et les contraintes opérationnelles imposées aux infrastructures. Cette dynamique complexifie considérablement la modélisation et la prédiction des pertes d'énergie, particulièrement dans un contexte où les conditions météorologiques et les profils de charge peuvent varier fortement d'une région à l'autre du réseau.

Dans ce contexte, la capacité à anticiper avec précision les pertes d'énergie sur le réseau de transport constitue un enjeu fondamental pour la planification énergétique et l'équilibrage de la charge. Une prédiction fiable permet de mieux aligner les décisions de production avec l'énergie effectivement disponible après transport, réduisant ainsi les risques de déséquilibre

1. Statistique provenant du site officiel de la Régie de l'énergie du Québec (regie-energie.qc.ca)
2. Source : Gouvernement du Québec (quebec.ca)

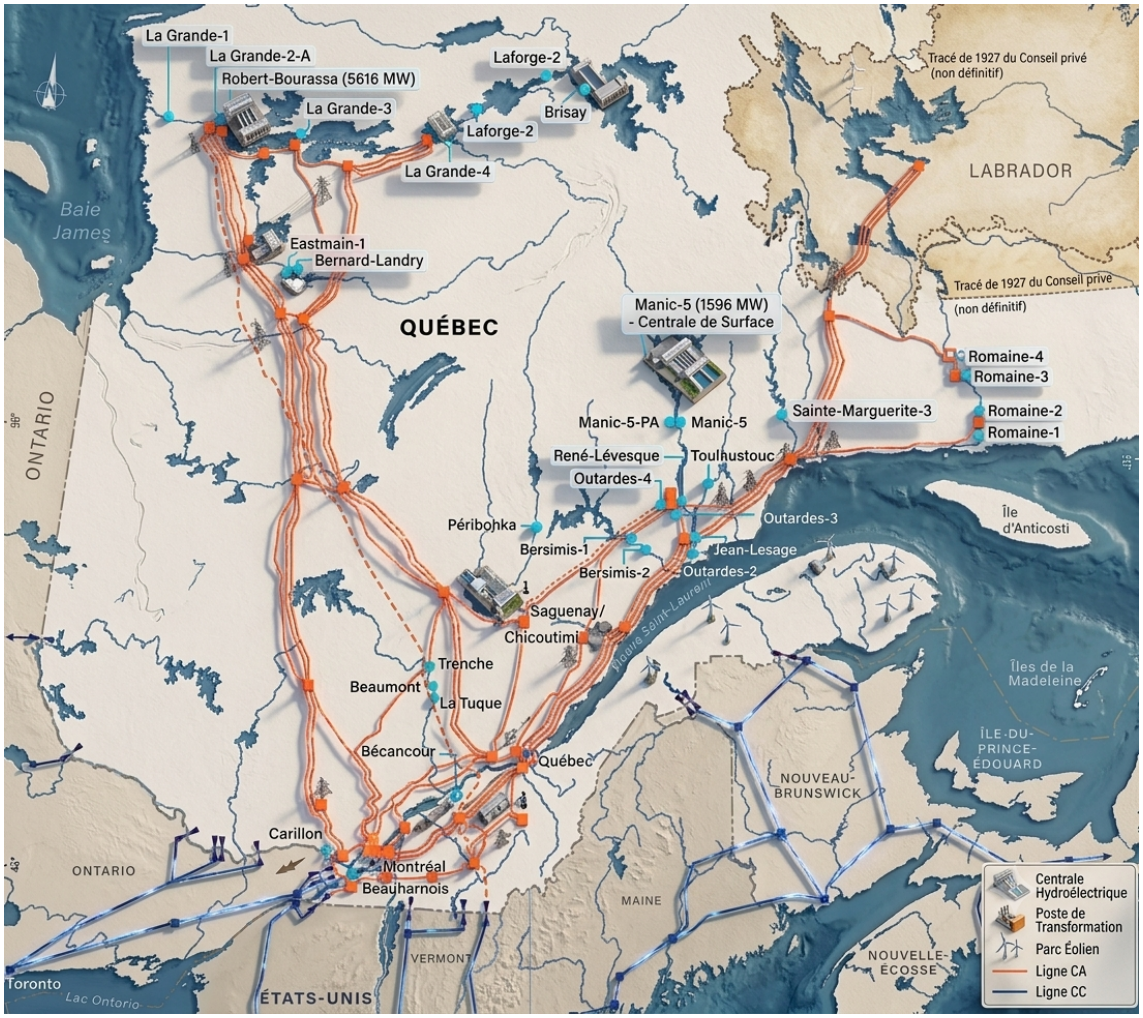

Figure 1.1 : Lignes de transport et centrales électriques d'Hydro-Québec (Source : `https://www.hydroquebec.com/transenergie/fr/`)

entre l'offre et la demande. Cette nécessité justifie le recours à des approches de modélisation et de prédiction plus avancées, capables de tenir compte de la complexité des réseaux électriques modernes.

Traditionnellement, la prédiction des pertes d'énergie dans les réseaux électriques s'appuie sur des modèles déterministes fondés sur les lois physiques de l'électricité. Bien que ces approches soient interprétables et reposent sur une compréhension explicite du comportement du réseau, elles présentent certaines limites en contexte réel, notamment lorsque les réseaux possèdent des topologies complexes, des charges fortement variables et des données issues de sources hétérogènes [52, 54]. Par ailleurs, le domaine de l'énergie demeure fortement contraint par des enjeux de confidentialité, de sécurité industrielle et de protection des infrastructures critiques. Les données détaillées relatives aux flux électriques, aux tensions, aux charges ou encore aux pertes sont rarement accessibles au public, ce qui limite la disponibilité de jeux de données ouverts et représentatifs.

En raison de ces contraintes, de nombreux travaux de recherche s'appuient sur des données simulées ou sur des environnements de test simplifiés. Bien que ces approches facilitent le développement et la validation initiale des modèles, elles reproduisent difficilement l'ensemble des conditions réelles d'exploitation d'un réseau électrique à grande échelle. Les variations saisonnières, les événements climatiques extrêmes, les fluctuations imprévisibles de la demande ainsi que les interactions complexes entre les différentes composantes du réseau y sont souvent partiellement représentées, voire absentes. Cette simplification réduit la représentativité des résultats obtenus et limite la capacité des modèles développés à généraliser efficacement lorsqu'ils sont appliqués à des données réelles et à des contextes opérationnels variés.

Une autre limite importante observée dans plusieurs travaux de la littérature concerne le niveau de granularité utilisé pour la prédiction des pertes. En effet, plusieurs approches existantes s'intéressent principalement à une prédiction globale des pertes à l'échelle du réseau électrique complet. Bien que cette formulation permette d'obtenir une vision générale du comportement énergétique du réseau, elle peut masquer les particularités locales propres à chaque ligne de transport. Or, les pertes techniques dépendent fortement de conditions locales telles que les caractéristiques physiques des lignes, les profils de charge, la température ambiante, le vent, le givre et d'autres variables météorologiques propres à l'environnement géographique traversé par chaque infrastructure. L'utilisation de variables globales, comme une météo moyenne sur l'ensemble du réseau ou des indicateurs électriques agrégés, peut ainsi limiter la précision des modèles, particulièrement dans le cas d'un réseau étendu couvrant des zones géographiques aux conditions climatiques très différentes.

Cependant, malgré l'importance opérationnelle des pertes techniques dans les réseaux de transport, leur prédiction demeure complexe en raison de la nature non linéaire du phénomène, de la variabilité temporelle des charges, de l'influence des conditions météorologiques et de la rareté des données réelles disponibles. De plus, les approches globales tendent à agréger les pertes à l'échelle du réseau complet, ce qui peut lisser les variations locales et masquer les comportements propres à chaque ligne de transport. Le verrou scientifique de cette recherche réside donc dans la capacité à développer des modèles prédictifs suffisamment précis et généralisables pour représenter les pertes techniques d'un réseau de transport réel, tout en tenant compte des spécificités locales propres à chaque ligne.

Face à ces limites, les approches basées sur l'Intelligence Artificielle (IA) et l'Apprentissage Automatique (ML) suscitent un intérêt croissant, puisqu'elles permettent de développer des

modèles prédictifs capables d'extraire des relations complexes à partir de données historiques et opérationnelles. Ces approches offrent la possibilité de dépasser certaines limites des modèles classiques en capturant des relations non linéaires, des dépendances temporelles et des interactions multivariées propres aux réseaux électriques modernes. Elles constituent ainsi une avenue prometteuse pour améliorer la prédiction des pertes d'énergie dans des conditions réelles d'exploitation.

Cette étude a été rendue possible grâce à une collaboration entre l'équipe CYNERGIA et HQ, laquelle a permis un accès privilégié à un vaste ensemble de données réelles couvrant vingt années d'historique des pertes et des paramètres électriques du réseau de transport. Cet ensemble de données constitue un atout majeur, puisqu'il permet non seulement de pallier la rareté des données publiques dans ce domaine, mais également de représenter plus fidèlement la complexité opérationnelle d'un réseau électrique à grande échelle. Il fournit ainsi un cadre particulièrement pertinent pour le développement, l'entraînement et l'évaluation de modèles dédiés à la prédiction des pertes d'énergie dans des conditions réelles d'exploitation.

Différentes familles de modèles sont étudiées afin d'évaluer leur capacité à prédire les pertes techniques propres à chaque ligne de transport, en exploitant des données historiques réelles composées de variables électriques et environnementales issues du réseau d'HQ. Les approches statistiques classiques basées sur les séries temporelles, telles que les modèles AutoRegressive Integrated Moving Average (ARIMA), Seasonal AutoRegressive Integrated Moving Average (SARIMA), Holt-Winters, Vector Autoregression (VAR) et Prophet, sont considérées en raison de leur simplicité, de leur interprétabilité et de leur capacité à modéliser les tendances temporelles. Cependant, ces méthodes peuvent présenter des limites lorsqu'il s'agit de capturer

des relations non linéaires, des dépendances temporelles complexes et l'influence simultanée de multiples variables opérationnelles et environnementales.

Les méthodes d'apprentissage automatique, telles que les forêts aléatoires (Random Forest), la méthode des K-Nearest Neighbors (KNN), les machines à vecteurs de support et les méthodes d'apprentissage par ensemble comme XGBoost, offrent une meilleure capacité à capturer les interactions complexes présentes dans les données énergétiques. Plus récemment, les modèles d'apprentissage profond, notamment les architectures LSTM, les mécanismes d'attention et les Transformers temporels, ont montré un intérêt particulier pour l'analyse des séries temporelles énergétiques, notamment en raison de leur capacité à modéliser des dépendances séquentielles et des interactions temporelles complexes.

Ainsi, la présente étude s'inscrit dans une démarche de modélisation prédictive des pertes d'énergie sur le réseau de transport d'HQ, reposant sur l'analyse de données réelles et l'évaluation comparative de différentes approches de modélisation. Contrairement aux approches globales généralement rencontrées dans la littérature, qui visent à prédire les pertes à l'échelle du réseau complet, ce travail adopte une approche désagrégée à granularité plus fine, fondée sur la prédiction des pertes au niveau des lignes de transport. Plus précisément, un modèle prédictif distinct est entraîné pour chaque ligne étudiée afin de mieux tenir compte des spécificités locales influençant les pertes, notamment les conditions météorologiques propres à l'environnement de la ligne, les caractéristiques électriques associées et les dynamiques temporelles propres à chaque infrastructure. Cette approche vise à éviter l'effet de lissage associé aux modèles globaux et à mieux représenter les conditions locales propres à chaque ligne de transport.

L'objectif principal de ce mémoire est donc d'évaluer la performance de différentes familles de modèles pour la prédiction des pertes techniques dans le réseau de transport d'HQ, en

considérant une approche ligne par ligne. Plus précisément, cette étude examine l'efficacité de modèles statistiques classiques, d'algorithmes d'apprentissage automatique et de modèles d'apprentissage profond, afin d'identifier les approches les plus adaptées à la dynamique complexe du réseau de transport. Cette démarche vise non seulement à comparer les performances prédictives de différentes méthodes, mais également à positionner l'approche ligne par ligne par rapport aux approches globales fondées sur des données agrégées.

Les principales contributions de ce mémoire peuvent ainsi être résumées comme suit. Premièrement, il exploite un ensemble de données réelles, rare et étendu, couvrant vingt ans d'exploitation du réseau de transport d'HQ. Deuxièmement, il propose une comparaison systématique de plusieurs familles de modèles, incluant des approches statistiques, des méthodes d'apprentissage automatique et des modèles d'apprentissage profond. Troisièmement, il adopte une approche de prédiction désagrégée au niveau des lignes de transport, permettant de mieux représenter les spécificités locales du réseau et de réduire la perte d'information associée à l'utilisation de variables globales ou agrégées. Cette contribution méthodologique est particulièrement pertinente pour les réseaux étendus, où les conditions météorologiques, les charges et les caractéristiques physiques peuvent varier considérablement d'une ligne à l'autre.

Ces contributions ont donné lieu à une première valorisation scientifique dans le cadre d'un article de conférence de huit pages, publié en libre accès dans la série Procedia Computer Science d'Elsevier et indexé dans les bases de données Scopus et EI Compendex. L'article a été présenté lors de la 9e Conférence internationale sur les données émergentes et l'industrie (EDI40), tenue conjointement avec la 17e conférence ANT à Istanbul, en Turquie, du 14 au 16 avril 2026. Cette publication met en évidence l'intérêt de l'approche proposée pour la

prédiction des pertes d'énergie à partir de données réelles issues d'un réseau de transport à grande échelle.

Le reste du mémoire est structuré comme suit. Le chapitre 2 présente la revue de la littérature ainsi qu'une discussion sur les travaux connexes. Le chapitre 3 décrit la méthodologie proposée, incluant la préparation des données, les modèles utilisés et la stratégie d'évaluation. Le chapitre 4 présente et analyse les résultats expérimentaux. Enfin, le chapitre 5 conclut le mémoire en discutant des limites de l'étude ainsi que des perspectives de recherche futures.

CHAPITRE II

REVUE DE LA LITTÉRATURE

La prédiction des pertes techniques dans les réseaux de transport d'électricité s'inscrit dans un champ de recherche interdisciplinaire, à l'intersection de l'ingénierie électrique, de l'analyse des séries temporelles et de l'IA. Elle mobilise à la fois la compréhension des phénomènes physiques associés au transport de l'énergie, l'exploitation de données historiques et temporelles, ainsi que des modèles d'apprentissage automatique capables de capturer des relations non linéaires entre les variables électriques, météorologiques et opérationnelles [71, 32, 75, 31].

Comme indiqué dans l'introduction, ces pertes dépendent de plusieurs facteurs interdépendants, notamment les caractéristiques physiques des lignes, les conditions de charge, les niveaux de tension, la température des conducteurs et les conditions environnementales locales [69, 67, 66, 7]. Leur comportement est donc difficile à représenter à l'aide d'un modèle unique, en particulier dans le cas d'un réseau étendu, soumis à de fortes variations saisonnières et météorologiques. Cette complexité a conduit à une évolution progressive des méthodes de prédiction, depuis les approches déterministes fondées sur les lois physiques jusqu'aux méthodes statistiques, aux algorithmes d'apprentissage automatique et aux modèles d'apprentissage profond [39, 72, 41].

Dans le cadre de ce mémoire, la revue de la littérature se concentre sur les travaux liés à la prédiction des pertes techniques dans les réseaux de transport. Elle présente d'abord les fondements physiques des pertes en ligne, puis les modèles statistiques basés sur les séries temporelles. Elle analyse ensuite les méthodes d'apprentissage automatique et d'apprentissage profond appliquées aux données énergétiques. Enfin, elle met en évidence les principales limites observées dans la littérature, notamment la rareté des données réelles, l'utilisation fréquente de données simulées ou agrégées, ainsi que le manque d'approches fondées sur une granularité fine au niveau des lignes de transport.

2.1 FONDEMENTS PHYSIQUES DES PERTES TECHNIQUES DANS LES RÉSEAUX DE TRANSPORT

Avant d'aborder les modèles prédictifs fondés sur les données, il est nécessaire de rappeler les mécanismes physiques à l'origine des pertes techniques. Cette compréhension permet de situer les approches statistiques et d'apprentissage automatique par rapport aux modèles traditionnels utilisés en ingénierie électrique. Elle permet également d'identifier les variables explicatives pertinentes, telles que le courant, la tension, la résistance des conducteurs, la température ambiante et les conditions météorologiques locales.

2.1.1 ORIGINE ET FORMULATION PHYSIQUE DES PERTES

Historiquement, la prédiction des pertes techniques dans les réseaux de transport reposait principalement sur des modèles déterministes fondés sur les équations de la physique du courant alternatif [52, 54]. Les pertes ohmiques, qui représentent une composante importante des pertes techniques, sont généralement décrites par la relation suivante :

$$P_{\text{loss}} = I^2 R, \tag{2.1}$$

où P_{loss} représente les pertes par effet Joule, I le courant circulant dans la ligne et R la résistance électrique du conducteur [48]. Dans le cas d'une ligne triphasée équilibrée, cette relation est souvent exprimée sous la forme :

$$P_{\text{loss}} = 3I^2 R. \tag{2.2}$$

La résistance du conducteur dépend notamment du matériau, de la longueur de la ligne, de sa section et de la température du conducteur [69, 23]. Cette dépendance peut être représentée par :

$$R(T) = R_0 \left[1 + \alpha(T - T_0)\right], \tag{2.3}$$

où R_0 désigne la résistance à la température de référence T_0, T la température du conducteur et α le coefficient de température du matériau. Cette relation montre que les conditions météorologiques, notamment la température, le vent et les phénomènes de refroidissement ou d'échauffement des conducteurs, peuvent influencer indirectement les pertes techniques.

La figure 2.1 illustre le fait que les pertes techniques ne dépendent pas uniquement de la puissance transportée, mais également des caractéristiques physiques et environnementales propres à chaque ligne de transport. En particulier, la résistance du conducteur R(T) varie avec la température ambiante, elle-même influencée par les conditions météorologiques locales telles que la température de l'air et le vent [7, 67]. Cette dépendance introduit une variabilité spatiale et temporelle des pertes qui ne peut être capturée par un modèle agrégé. Cette observation

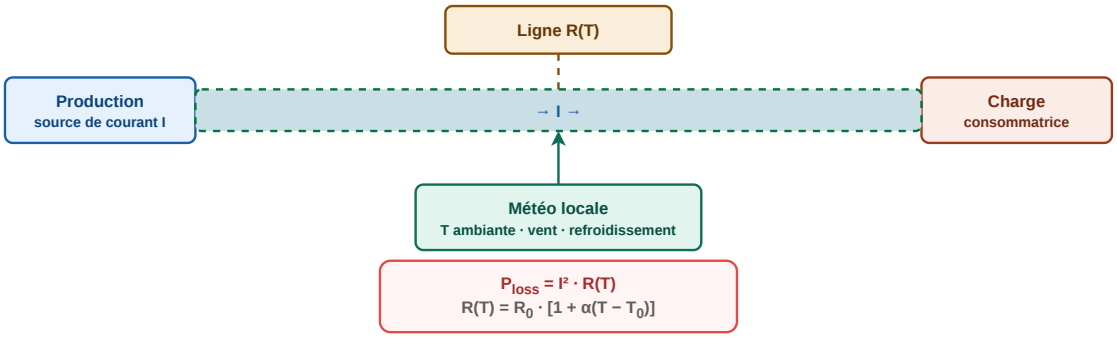

Figure 2.1 : Influence du courant et des conditions locales sur les pertes techniques d'une ligne de transport.

justifie l'intérêt d'une approche à granularité fine, dans laquelle les variables locales associées à chaque infrastructure, telles que le courant de transit, la résistance thermique du conducteur et les conditions météorologiques, sont explicitement prises en compte dans le processus de modélisation.

2.1.2 FORCES ET LIMITES DES APPROCHES DÉTERMINISTES

Les approches déterministes présentent l'avantage de reposer sur des principes physiques clairement interprétables, notamment les équations de flux de puissance, les lois de Kirchhoff et les modèles électriques des lignes, des transformateurs et des équipements du réseau. Elles sont largement utilisées dans les outils de calcul de flux de puissance, les logiciels de simulation de réseaux ainsi que dans les applications d'analyse des systèmes de gestion de l'énergie (EMS), souvent alimentées par des mesures issues des systèmes de Supervision Contrôle et Acquisition de Données (SCADA). Lorsque la topologie du réseau, les paramètres électriques des lignes et les conditions d'exploitation sont connus avec suffisamment de précision, ces approches

permettent de représenter explicitement le comportement électrique du réseau et d'estimer les tensions, les courants, les flux de puissance ainsi que les pertes associées [74, 17, 13].

Toutefois, l'application de ces approches devient plus exigeante lorsque le réseau est vaste, fortement interconnecté et soumis à des variations rapides de charge, de topologie et de conditions environnementales. La précision des modèles déterministes dépend en effet de la disponibilité de paramètres détaillés et régulièrement mis à jour, tels que les caractéristiques électriques des lignes, les conditions de charge, les niveaux de tension, les flux de puissance, l'état de la topologie du réseau ainsi que, dans certains cas, les conditions météorologiques locales. Dans un contexte opérationnel réel, ces informations ne sont pas toujours disponibles avec une résolution temporelle, spatiale ou qualitative suffisante. De plus, les pertes techniques ne résultent pas uniquement des équations électriques de base, mais peuvent aussi être influencées par des phénomènes dynamiques liés à la variabilité temporelle de la demande, aux conditions climatiques, à l'évolution de la température des conducteurs et aux interactions entre différentes zones du réseau [71, 56, 31, 32].

Ces limites expliquent l'intérêt croissant pour les approches fondées sur les données, capables d'apprendre à partir d'observations historiques les relations complexes entre les variables électriques, opérationnelles, environnementales et les pertes mesurées. Ces approches ne visent pas nécessairement à remplacer la connaissance physique du réseau, mais plutôt à la compléter lorsque certains phénomènes sont difficiles à formaliser explicitement ou lorsque les paramètres nécessaires aux modèles déterministes ne sont pas disponibles avec une précision suffisante [75, 71, 53].

2.2 APPROCHES STATISTIQUES ET MODÈLES DE SÉRIES TEMPORELLES

Les modèles statistiques constituent l'une des premières familles d'approches fondées sur les données pour la modélisation et la prédiction des séries temporelles. Ils sont particulièrement adaptés lorsque la variable à prédire présente une structure temporelle identifiable, telle qu'une tendance, une saisonnalité, une cyclicité ou une dépendance aux valeurs passées. Dans le domaine énergétique, ces modèles ont été largement mobilisés pour la prédiction de la charge électrique, de la demande, de la consommation et de la production d'énergie renouvelable. Ils sont également utilisés comme modèles de référence pour comparer des approches plus récentes d'apprentissage automatique et profond appliquées à des variables énergétiques liées au fonctionnement du réseau et, indirectement, aux pertes [49, 37, 6].

2.2.1 MODÈLES AUTORÉGRESSIFS, ARIMA, SARIMA ET VAR

Les approches statistiques basées sur les séries temporelles ont été largement utilisées dans le domaine énergétique en raison de leur simplicité, de leur interprétabilité et de leur capacité à modéliser les tendances, les saisonnalités et les dépendances temporelles [51]. Les modèles autorégressifs, les modèles ARIMA, leurs extensions saisonnières SARIMA, les modèles VAR ainsi que les méthodes de lissage exponentiel, telles que Holt-Winters, figurent parmi les approches classiques les plus courantes pour la prédiction de variables énergétiques, notamment la charge électrique, la demande, la consommation et certains indicateurs opérationnels liés au réseau [47, 37, 49, 30].

Un modèle ARIMA peut être représenté de manière générale par :

$$\phi(B)(1-B)^d y_t = \theta(B)\varepsilon_t, \tag{2.4}$$

où y_t représente la série temporelle à l'instant t, B l'opérateur de retard, d l'ordre de différenciation, $\phi(B)$ la composante autorégressive, $\theta(B)$ la composante moyenne mobile et ε_t le terme d'erreur.

Les modèles SARIMA ajoutent une composante saisonnière, ce qui les rend particulièrement pertinents pour les données énergétiques influencées par des cycles journaliers, hebdomadaires ou annuels. Les modèles VAR permettent quant à eux de représenter simultanément plusieurs séries temporelles interdépendantes, ce qui peut être utile lorsque la variable à prédire dépend de plusieurs grandeurs électriques corrélées. Ces méthodes nécessitent généralement moins de données que les modèles d'apprentissage automatique complexes ; elles sont relativement faciles à interpréter et reposent sur un cadre mathématique bien établi pour l'analyse des séries temporelles.

Cependant, leur efficacité demeure limitée lorsque les relations entre les variables sont fortement non linéaires ou lorsque les pertes dépendent simultanément de nombreux facteurs électriques, temporels et environnementaux [49]. Les hypothèses de linéarité, de stationnarité et de stabilité de la structure temporelle peuvent alors devenir restrictives dans le cas des réseaux électriques modernes , caractérisés par une forte variabilité de la charge, des conditions climatiques et des modes d'exploitation.

Cependant, ces modèles deviennent moins adaptés lorsque les séries présentent des relations fortement non linéaires, une variabilité temporelle importante ou une dépendance simultanée à plusieurs facteurs électriques, temporels et environnementaux [49]. Cette limite est particulièrement importante dans les réseaux électriques modernes [50, 33], où la charge, les conditions météorologiques, l'intégration des énergies renouvelables et les modes d'ex-

ploitation influencent directement les grandeurs à prédire, y compris les pertes techniques [56, 32, 75].

2.2.2 MODÈLE PROPHET ET DÉCOMPOSITION ADDITIVE

Le modèle Prophet [3], développé par Facebook[5, 25], introduit par Taylor et Letham, constitue une approche additive de prévision des séries temporelles, particulièrement adaptée aux données présentant des tendances, des changements de régime et des saisonnalités marquées . Il repose sur une décomposition explicite de la série temporelle sous la forme :

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t, \tag{2.5}$$

où $g(t)$ représente la tendance non périodique, $s(t)$ les composantes saisonnières, $h(t)$ les effets associés à des jours particuliers ou à des événements connus, et ε_t le terme d'erreur résiduelle.

Cette structure rend Prophet intéressant pour certaines séries énergétiques caractérisées par des cycles temporels réguliers, comme la demande électrique [16], la consommation ou la production, qui peuvent être influencées par des variations journalières, hebdomadaires, annuelles, ainsi que par les conditions climatiques et les jours particuliers. La figure 2.2 illustre le fonctionnement général de cette décomposition.

Toutefois, malgré son interprétabilité et sa capacité à modéliser explicitement les tendances et les saisonnalités, Prophet peut être limité lorsque la variable à prédire dépend d'interactions

3. Lien de la documentation officielle du modèle Prophet de Facebook : `https://facebook.github.io/prophet/`

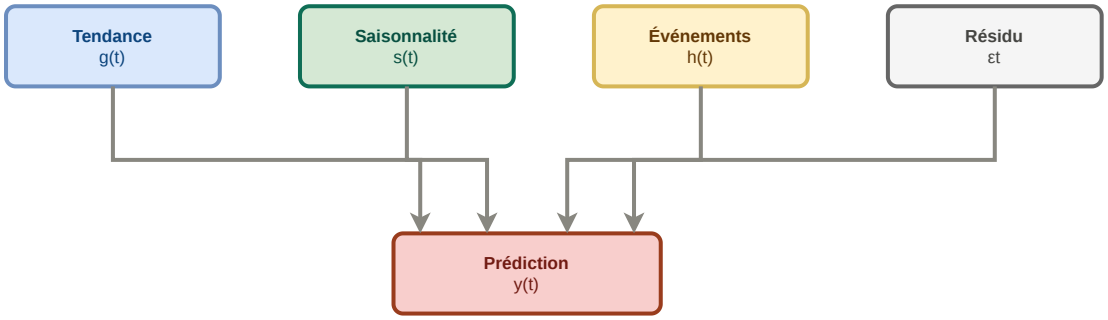

Figure 2.2 : Principe de décomposition additive utilisé par Prophet.

complexes entre plusieurs variables explicatives ou de comportements fortement non linéaires [8]. Dans le cas des réseaux électriques, où les pertes peuvent dépendre simultanément de facteurs électriques, temporels, opérationnels et environnementaux [61], cette limite justifie l'utilisation de modèles plus flexibles, notamment les méthodes d'apprentissage automatique et profond.

2.3 MÉTHODES D'APPRENTISSAGE AUTOMATIQUE POUR LA PRÉDICTION DES PERTES

L'apprentissage automatique représente une évolution importante par rapport aux modèles statistiques classiques, en particulier lorsque la variable à prédire dépend de relations non linéaires et multivariées [26, 41]. Contrairement aux approches statistiques linéaires ou aux modèles déterministes fondés sur une formalisation explicite du réseau, ces méthodes permettent d'apprendre directement à partir des données les relations complexes entre les variables d'entrée et les pertes observées. Elles sont donc particulièrement adaptées lorsque les pertes dépendent simultanément de variables électriques, temporelles, opérationnelles et environnementales, dont les interactions sont difficiles à modéliser explicitement [32].

2.3.1 CADRE GÉNÉRAL DE L'APPRENTISSAGE SUPERVISÉ

Dans un cadre supervisé, la prédiction des pertes consiste à apprendre une fonction f_θ reliant un ensemble de variables d'entrée $\mathbf{x}_t$ à une valeur cible y_t représentant les pertes observées :

$$\widehat{y} * t = f * \theta(\mathbf{x}_t), \tag{2.6}$$

où $\widehat{y} * t$ désigne la perte prédite et θ les paramètres du modèle. L'entraînement consiste généralement à ajuster ces paramètres afin de minimiser une fonction de coût mesurant l'écart entre les valeurs observées et les valeurs prédites. Dans le cas d'un problème de régression, une fonction couramment utilisée est l'erreur quadratique moyenne :

$$\mathcal{L}(\theta) = \frac{1}{n} \sum *t = 1^n \left(y_t - \widehat{y}_t\right)^2 . \tag{2.7}$$

La figure 2.3 illustre la logique générale d'un pipeline d'apprentissage supervisé appliqué à la prédiction des pertes. Les variables d'entrée peuvent inclure des grandeurs électriques telles que la charge, la puissance active et réactive, la tension ou le courant, ainsi que des variables temporelles et environnementales comme l'heure, le jour, la saison, la température, l'humidité ou la vitesse du vent. Dans le cadre d'une approche ligne par ligne, ce vecteur d'entrée est défini spécifiquement pour chaque ligne de transport, afin de mieux tenir compte de ses caractéristiques électriques et de ses conditions locales.

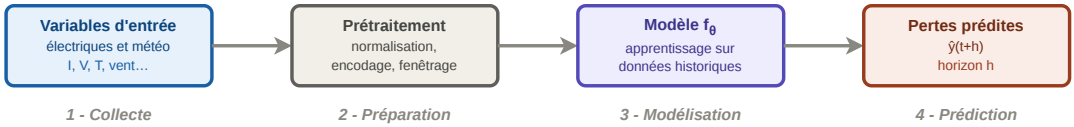

Figure 2.3 : Cadre général d'un modèle supervisé pour la prédiction des pertes techniques.

2.3.2 FORÊTS ALÉATOIRES, SVR, KNN ET MÉTHODES DE GRADIENT BOOSTING

Parmi les méthodes supervisées classiques utilisées pour les problèmes de régression figurent les machines à vecteurs de support pour la régression, les forêts aléatoires, les KNN, ainsi que les méthodes d'apprentissage par ensemble telles que XGBoost et LightGBM [42, 29, 51, 58]. Dans le contexte des pertes électriques, les travaux récents portant sur les réseaux de transport mettent particulièrement en évidence l'usage de modèles comme le SVR, XGBoost, les forêts aléatoires et les réseaux neuronaux pour apprendre les relations entre les pertes observées et les variables d'exploitation du réseau [71].

Ces algorithmes peuvent intégrer simultanément plusieurs variables d'entrée, telles que la charge, la tension, le courant, la puissance active ou réactive, la température, l'humidité, la vitesse du vent ou d'autres variables météorologiques. Ils peuvent ainsi capturer empiriquement des relations non linéaires et des interactions entre les conditions électriques, temporelles et environnementales, sans nécessiter une formalisation explicite de tous les phénomènes physiques en jeu [31].

Les forêts aléatoires reposent sur l'agrégation de plusieurs arbres de décision entraînés sur des sous-échantillons des données et des sous-ensembles de variables. Pour un problème de régression, leur prédiction peut être représentée par :

$$\widehat{y} = \frac{1}{M} \sum_{m=1}^{M} T_m(\mathbf{x}), \tag{2.8}$$

où T_m désigne le m-ième arbre de décision et M le nombre total d'arbres. Cette agrégation permet généralement de réduire la variance du modèle et d'améliorer sa robustesse face au bruit [12].

Les méthodes de gradient boosting, comme XGBoost et LightGBM, construisent progressivement un modèle additif sous la forme :

$$\widehat{y} = F_0(\mathbf{x}) + \sum_{m=1}^{M} \eta f_m(\mathbf{x}), \tag{2.9}$$

où $F_0(\mathbf{x})$ représente le modèle initial, f_m un modèle faible, généralement un arbre de décision, et η le taux d'apprentissage. À chaque itération, un nouveau modèle est ajouté afin de corriger les erreurs résiduelles ou, plus généralement, d'optimiser la fonction de coût par descente de gradient. Cette stratégie explique les bonnes performances des méthodes de gradient boosting dans de nombreux problèmes de régression [11, 19, 34].

Malgré leurs avantages, les méthodes d'apprentissage automatique classiques présentent certaines limites. Leur performance dépend fortement de la qualité et de la représentativité des données d'entrée, du choix des variables explicatives, du réglage des hyperparamètres ainsi que de la stratégie de validation utilisée. De plus, des modèles tels que les forêts aléatoires, les SVR ou les méthodes de gradient boosting sont généralement appliqués à des données tabulaires. Ils ne modélisent donc pas explicitement la structure séquentielle des séries temporelles. La prise en compte de la dépendance temporelle repose alors souvent sur une ingénierie manuelle des caractéristiques, par exemple par l'ajout de variables retardées, de moyennes mobiles, de variables calendaires ou de fenêtres glissantes [22, 21].

Cette dépendance à l'ingénierie des caractéristiques peut devenir limitante lorsque les pertes évoluent selon des dynamiques complexes, influencées simultanément par l'historique de charge, les variations de tension et de puissance, les conditions météorologiques locales ainsi que les changements d'exploitation du réseau [21, 75, 22].

2.4 APPRENTISSAGE PROFOND ET MODÈLES SÉQUENTIELS

Les modèles d'apprentissage profond ont récemment suscité un intérêt important pour la prédiction des séries temporelles énergétiques. Leur capacité à apprendre automatiquement des représentations complexes à partir des données les rend adaptés aux problèmes où les relations entre les variables sont non linéaires, temporellement dépendantes et multivariées. Dans le contexte des pertes techniques, ces modèles peuvent représenter les dépendances temporelles entre les grandeurs électriques, les conditions d'exploitation, les variables météorologiques et les pertes observées [20, 75, 70].

2.4.1 RÉSEAUX RÉCURRENTS, MODÉLISATION SÉQUENTIELLE PAR LSTM ET MÉCANISME D'ATTENTION TEMPORELLE

Les réseaux de neurones récurrents, ou RNN, sont conçus pour le traitement des données séquentielles. À chaque pas de temps t, le modèle met à jour un état caché à partir de l'observation courante et de l'état précédent, selon l'équation :

$$\mathbf{h}_t = \sigma\left(W_x \mathbf{x} * t + W_h \mathbf{h} * t - 1 + \mathbf{b}\right), \tag{2.10}$$

où $\mathbf{x}_t$ représente le vecteur d'entrée, $\mathbf{h}_t$ l'état caché, W_x et W_h les matrices de poids, $\mathbf{b}$ le vecteur de biais et σ une fonction d'activation non linéaire. Les RNN classiques restent toutefois limités par les problèmes d'évanouissement et d'explosion du gradient, ce qui peut restreindre leur capacité à capturer des dépendances temporelles à long terme [35, 75].

Les architectures LSTM répondent en partie à cette limitation en introduisant une cellule mémoire et un système de portes qui contrôlent l'information conservée, oubliée ou transmise au fil du temps [35].

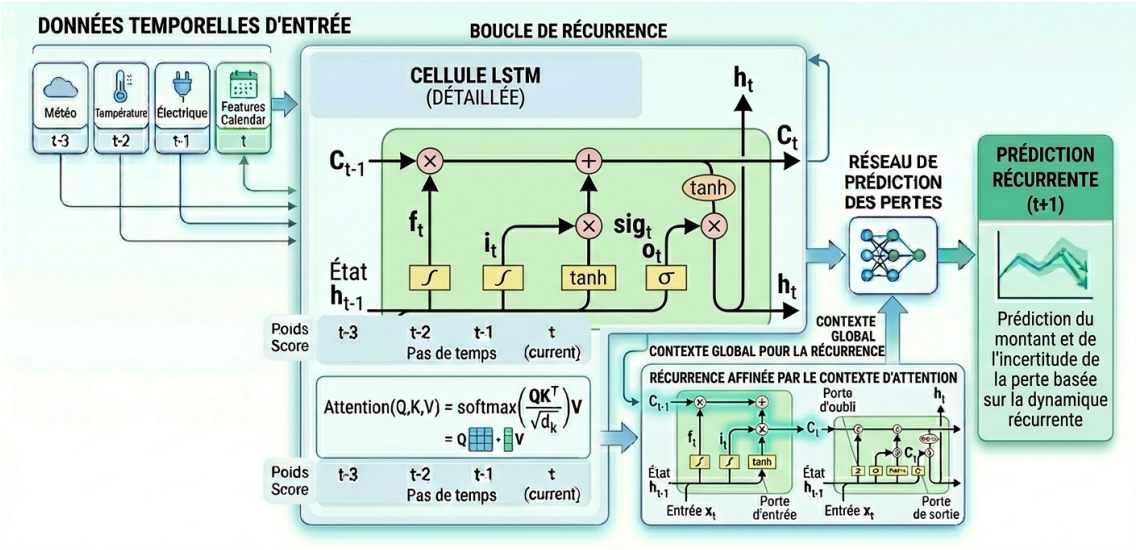

Figure 2.4 : Architecture récurrente LSTM couplée à un mécanisme d'attention temporelle pour la prédiction des pertes.

Dans une approche de prédiction des pertes fondée sur des séquences temporelles, le modèle peut recevoir en entrée une fenêtre glissante de données multivariées comprenant des variables électriques, météorologiques, opérationnelles et calendaires [75]. Pour une ligne donnée, cette séquence peut être notée :

$$\mathbf{X} * t - L + 1 : t = (\mathbf{x} * t - L + 1, \mathbf{x} * t - L + 2, \ldots, \mathbf{x} * t)\,, \tag{2.11}$$

où L désigne la longueur de la fenêtre temporelle utilisée comme entrée du modèle.

À chaque pas de temps, la cellule LSTM met à jour son état de mémoire $\mathbf{C}_t$ et son état caché $\mathbf{h}_t$ à l'aide de trois portes principales : la porte d'oubli $\mathbf{f}_t$, la porte d'entrée $\mathbf{i}_t$ et la porte de sortie $\mathbf{o}_t$. Ces mécanismes sont régis par les équations suivantes :

$$\mathbf{f}_t = \sigma\left(W_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f\right), \tag{2.12}$$

$$\mathbf{i}_t = \sigma\left(W_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i\right), \tag{2.13}$$

$$\widetilde{\mathbf{C}}_t = \tanh\left(W_C[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_C\right), \tag{2.14}$$

$$\mathbf{C}_t = \mathbf{f}_t \odot \mathbf{C}_{t-1} + \mathbf{i}_t \odot \widetilde{\mathbf{C}}_t, \tag{2.15}$$

$$\mathbf{o}_t = \sigma\left(W_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o\right), \tag{2.16}$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{C}_t), \tag{2.17}$$

où $\odot$ désigne le produit de Hadamard, σ la fonction sigmoïde, tandis que les matrices W et les vecteurs $\mathbf{b}$ représentent respectivement les poids et les biais appris par le modèle [35].

Afin de renforcer la capacité du modèle à identifier les instants les plus pertinents dans une séquence historique, un mécanisme d'attention temporelle peut être ajouté aux représentations produites par le LSTM [60, 1, 65]. Dans le cas d'une attention de type *scaled dot-product*, les poids d'attention sont calculés à partir de matrices de requêtes Q, de clés K et de valeurs V :

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \tag{2.18}$$

où d_k désigne la dimension des clés. L'application de la fonction *softmax* permet d'obtenir des poids normalisés qui attribuent une importance variable aux observations passées [55, 76].

Comme l'illustre la figure 2.4, le mécanisme d'attention peut produire un vecteur de contexte à partir des états cachés du LSTM. Ce vecteur résume les informations temporelles jugées les plus pertinentes pour la prédiction. Il est ensuite transmis à une couche de sortie, généralement une couche entièrement connectée, afin de prédire la perte à l'instant futur, par exemple $\widehat{y}_{t+1}$. Cette combinaison entre mémoire récurrente et attention temporelle permet de mieux exploiter les dépendances à court et à long terme présentes dans les séries énergétiques multivariées.

2.4.2 MÉCANISMES D'ATTENTION GLOBALE ET ARCHITECTURES BASÉES SUR LES TRANSFORMERS

Parallèlement aux structures récurrentes, les mécanismes d'attention neuronale et les architectures de type Transformers offrent une perspective particulièrement intéressante pour le traitement des séries temporelles [73, 2, 76]. Contrairement aux approches récurrentes classiques, qui traitent l'information de manière séquentielle, les Transformers s'appuient sur l'auto-attention (*self-attention*) afin de mettre en relation différents pas de temps d'une séquence, indépendamment de leur distance temporelle. Pour préserver l'ordre chronologique indispensable à l'analyse des réseaux électriques, ces architectures intègrent un codage positionnel (*positional encoding*) [2]. Cette propriété globale permet de mieux modéliser les dépendances à long terme et d'exploiter plus efficacement les interactions entre les observations historiques .

Le mécanisme d'attention produit scalaire mis à l'échelle, introduit précédemment à l'équation (2.18), constitue le cœur des architectures Transformers [55, 76]. Dans ce cadre, il est

utilisé sous forme d'auto-attention, c'est-à-dire que les matrices de requêtes, de clés et de valeurs sont construites à partir de la même séquence d'entrée. Cette formulation permet au modèle d'évaluer les corrélations croisées entre les différents pas de temps d'une séquence et d'attribuer dynamiquement des poids aux observations historiques selon leur pertinence pour la prédiction.

Comme l'illustre la figure 2.5, issue du modèle original introduit par Vaswani et al.[76], cette architecture s'affranchit des contraintes de récurrence. Elle permet une parallélisation massive lors de la phase d'entraînement, puisque les dépendances temporelles ne sont plus traitées de manière successive, mais simultanée sur l'ensemble de la fenêtre temporelle.

Dans le cadre de la prévision de variables énergétiques, les mécanismes d'attention sont particulièrement pertinents pour analyser des séries temporelles multivariées influencées par des facteurs électriques, temporels, opérationnels et environnementaux [55, 76]. Pour la prédiction des pertes techniques sur le réseau de transport, cette capacité permet d'isoler et de surpondérer des périodes historiques critiques, telles que des épisodes de forte charge, des variations rapides de transit de puissance ou des extrêmes météorologiques, lorsque ces événements influencent directement les pertes futures.

Néanmoins, l'application de Transformers purs aux séries temporelles énergétiques présente plusieurs défis. L'auto-attention standard devient coûteuse lorsque la longueur des séquences augmente, en raison de sa complexité quadratique par rapport au nombre de pas de temps [45, 79]. De plus, ces modèles nécessitent un réglage minutieux des hyperparamètres, des stratégies de régularisation adaptées et un volume de données conséquent afin de limiter le risque de surapprentissage. Ces difficultés ont motivé le développement de variantes optimisées

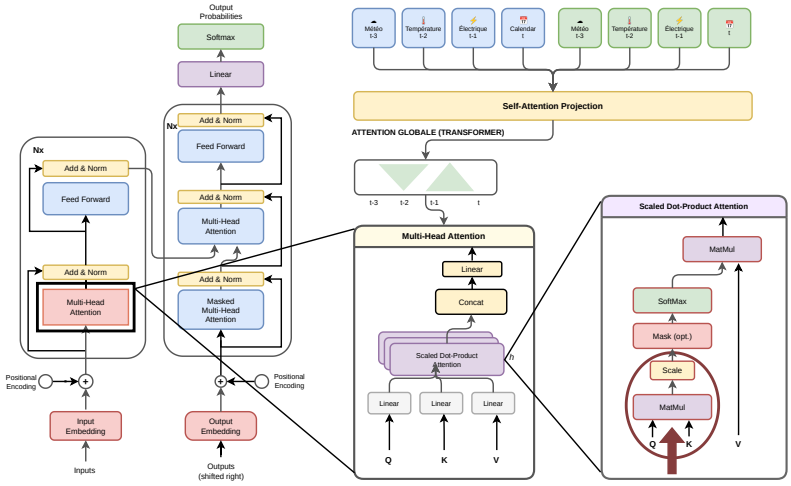

Figure 2.5 : Architecture générale du modèle Transformer et décomposition du mécanisme d'auto-attention (adapté de Vaswani et al. [76]).

pour les séquences temporelles longues, telles qu'Informer [79], Autoformer ou le Temporal Fusion Transformer (TFT) [43, 78].

Ainsi, dans le cadre d'une comparaison entre modèles statistiques, d'apprentissage automatique et d'apprentissage profond, les Transformers constituent une famille d'architectures pertinente à évaluer. Toutefois, pour un contexte opérationnel où l'on cherche à concilier performance prédictive, stabilité d'entraînement et exploitation de séquences temporelles multivariées, une approche hybride combinant LSTM et mécanisme d'attention peut représenter une alternative plus progressive [57, 18]. Cette combinaison permet de conserver l'aptitude native du LSTM à

modéliser la dynamique séquentielle locale, tout en utilisant l'attention pour capter les relations à plus long terme nécessaires à la prédiction [40].

2.5 GRANULARITÉ DES APPROCHES DE PRÉDICTION DES PERTES

Au-delà du choix de la famille de modèles, la granularité de la prédiction constitue un enjeu central. Une même architecture peut produire des résultats différents selon que les pertes sont prédites à l'échelle du réseau complet, d'une zone électrique ou d'une ligne de transport. Cette question est particulièrement importante dans le cadre de ce mémoire, car l'approche proposée repose sur une prédiction ligne par ligne.

2.5.1 APPROCHES GLOBALES FONDÉES SUR DES VARIABLES AGRÉGÉES

Dans plusieurs travaux, les pertes sont étudiées à l'échelle globale du réseau ou d'une zone électrique, à partir de variables agrégées. Cette approche présente l'avantage de réduire la complexité du problème et de fournir une vision générale du comportement énergétique du réseau. Par exemple, Tulensalo et al. [75] proposent un modèle LSTM pour la prédiction des pertes d'un réseau électrique à partir de séries temporelles horaires combinant des données de marché, des variables calendaires et des conditions météorologiques locales. De même, Pandžić et al. [56] étudient la prévision des pertes actives dans un réseau de transport à l'échelle opérationnelle d'un gestionnaire de réseau, en utilisant des données disponibles au niveau du système de transport.

Une approche globale peut être représentée par :

$$\widehat{Y}_{t+h}^{\text{global}} = f\left(\mathbf{X} * t^{\text{global}}\right), \tag{2.19}$$

où $\widehat{Y} * t + h^{\text{global}}$ désigne la prédiction globale des pertes à l'horizon h, et $\mathbf{X}_t^{\text{global}}$ un ensemble de variables agrégées à l'échelle du réseau ou de la zone étudiée. Une telle approche peut être pertinente lorsque l'objectif est d'obtenir une estimation globale des pertes pour la planification ou l'analyse macroénergétique.

Toutefois, cette agrégation peut masquer les spécificités locales propres aux différentes lignes de transport. Les pertes techniques d'une ligne dépendent notamment de ses caractéristiques physiques, de sa longueur, de son niveau de tension, du courant qui la traverse, de la puissance transitée, de son environnement géographique et des conditions météorologiques locales auxquelles elle est exposée. Cette dépendance locale est particulièrement importante pour certaines composantes des pertes techniques, comme les pertes corona, dont la sensibilité aux conditions météorologiques autour des lignes de transport a été mise en évidence par Gupta et al. [31]. Dans un réseau étendu, les lignes peuvent traverser des régions soumises à des conditions climatiques différentes. L'utilisation d'une météo moyenne ou de variables électriques globales peut alors introduire un effet de lissage et réduire la capacité du modèle à représenter précisément les pertes associées à certaines lignes.

2.5.2 APPROCHE DÉSAGRÉGÉE AU NIVEAU DES LIGNES DE TRANSPORT

Une alternative aux approches globales consiste à modéliser les pertes à une échelle plus fine, notamment au niveau des lignes de transport. Cette approche désagrégée vise à préserver les spécificités locales de chaque ligne, qui peuvent être masquées lorsque les variables électriques ou météorologiques sont agrégées à l'échelle du réseau complet. Cette logique est cohérente avec les travaux de Ding et al. [21], qui traitent explicitement la prédiction des pertes de lignes

de transport à très haute tension en exploitant des variables électriques et météorologiques associées aux lignes étudiées.

Dans une telle formulation, un modèle spécifique peut être associé à chaque ligne de transport :

$$\widehat{y} * \ell, t + h = f * \ell \left(\mathbf{x} * \ell, t, \mathbf{x} * \ell, t - 1, \ldots, \mathbf{x} * \ell, t - p \right), \tag{2.20}$$

où $\widehat{y} * \ell, t + h$ représente la perte prédite en MW pour la ligne ℓ à l'horizon h, f_ℓ le modèle associé à cette ligne, $\mathbf{x}_{\ell,t}$ le vecteur de variables propres à la ligne au temps t, et p la profondeur temporelle considérée.

Cette approche permet d'intégrer des variables électriques, temporelles et environnementales propres à chaque infrastructure. Elle est particulièrement pertinente dans le cas des réseaux étendus, où les lignes peuvent être exposées à des conditions climatiques différentes selon leur localisation géographique. Contrairement à une approche fondée sur des variables moyennes à l'échelle du réseau, la modélisation ligne par ligne permet donc de mieux représenter les dynamiques locales susceptibles d'influencer les pertes techniques.

Lorsque les pertes sont exprimées en unités absolues, par exemple en MW, les prédictions obtenues à l'échelle des lignes peuvent ensuite être agrégées afin d'obtenir une estimation globale des pertes horaires sur le périmètre étudié :

$$\widehat{Y} * t + h^{\text{lignes}} = \sum * \ell = 1^{L} \widehat{y}_{\ell,t+h}, \tag{2.21}$$

où L représente le nombre total de lignes considérées. Cette formulation permet de conserver l'information locale propre à chaque ligne tout en reconstruisant une vision agrégée des pertes sur l'ensemble des lignes étudiées.

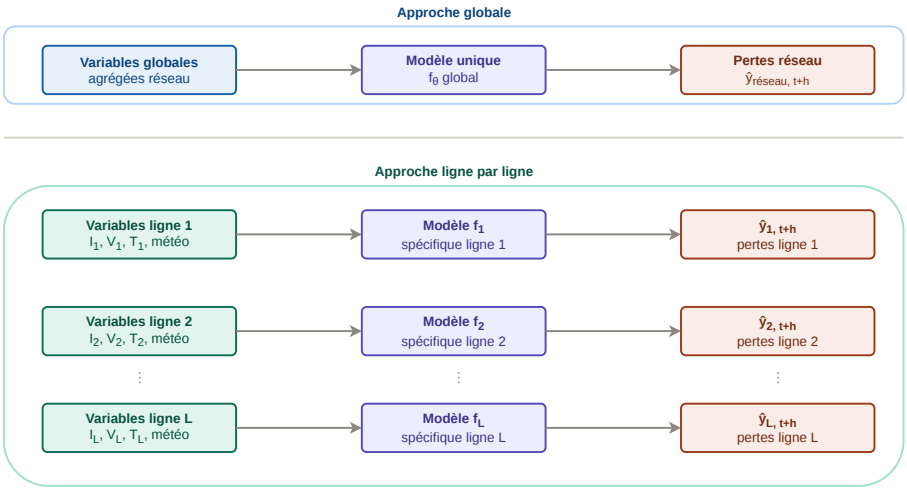

Figure 2.6 : Comparaison entre une approche globale et une approche ligne par ligne pour la prédiction des pertes techniques.

La figure 2.6 résume la différence entre une approche globale et une approche désagrégée. Cette stratégie est cohérente avec l'objectif du présent mémoire, qui vise à comparer différentes familles de modèles dans un cadre de prédiction ligne par ligne, plutôt que de produire uniquement une prédiction agrégée à l'échelle du réseau complet.

2.6 DISPONIBILITÉ DES DONNÉES RÉELLES ET LIMITES DES APPROCHES SIMULÉES

Un enjeu important dans la littérature sur la prédiction des pertes électriques concerne la disponibilité limitée des données réelles et publiques issues des réseaux électriques [3, 44]. Les enregistrements détaillés des réseaux électriques sont rarement diffusés publiquement en raison de contraintes liées à la sécurité, à la confidentialité industrielle et à la protection des infrastructures critiques [24]. Cette situation limite la reproductibilité de certains travaux et rend difficile la comparaison directe entre les méthodes proposées dans la littérature.

En l'absence de jeux de données publics, plusieurs études s'appuient sur des réseaux de test standardisés, des données simulées ou des jeux de données synthétiques [28, 44, 3]. Ces environnements contrôlés sont utiles pour tester des algorithmes, comparer des architectures de modèles ou analyser certains scénarios. Toutefois, ils ne reproduisent pas toujours toute la complexité d'un réseau industriel exploité sur une longue période. Les données simulées peuvent notamment ne pas refléter certains phénomènes observés en exploitation réelle, tels que les changements de configuration du réseau, les indisponibilités d'équipements, le bruit de mesure, les valeurs manquantes, les variations locales de la demande ou certains événements climatiques.

Cette limite est particulièrement importante pour la prédiction des pertes techniques, car les performances d'un modèle peuvent dépendre fortement des conditions réelles d'exploitation et de la qualité des mesures disponibles. Un modèle performant sur des données simulées peut voir sa précision diminuer lorsqu'il est appliqué à des données opérationnelles réelles.

Dans ce contexte, l'accès à un historique de données réelles issues du réseau de transport d'Hydro-Québec constitue un atout méthodologique important. Ces données permettent d'évaluer les modèles en conditions réelles d'exploitation et de mesurer leur capacité à prédire les pertes en présence de variations électriques, temporelles et météorologiques effectivement observées sur le réseau. Cette disponibilité de données opérationnelles réelles renforce ainsi la pertinence de la comparaison entre modèles statistiques, modèles d'apprentissage automatique et modèles d'apprentissage profond proposée dans ce travail.

2.6.1 MÉTRIQUES D'ÉVALUATION ET VALIDATION TEMPORELLE

La comparaison des modèles nécessite des métriques cohérentes avec l'objectif de prédiction. Dans les problèmes de régression énergétique, les métriques les plus utilisées incluent notamment l'erreur absolue moyenne (MAE), l'erreur quadratique moyenne (MSE), la racine de l'erreur quadratique moyenne (RMSE), l'erreur absolue moyenne en pourcentage (MAPE) ainsi que le coefficient de détermination (R^2) [77, 63, 27]. Elles peuvent être exprimées comme suit :

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n} \left| y_t - \widehat{y} * t \right|, \tag{2.22}$$

$$\text{MSE} = \frac{1}{n}\sum * t = 1^{n} \left(y_t - \widehat{y} * t \right)^2, \tag{2.23}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum * t = 1^{n} \left(y_t - \widehat{y} * t \right)^2}, \tag{2.24}$$

$$\text{MAPE} = \frac{100}{n}\sum * t = 1^{n} \left| \frac{y_t - \widehat{y} * t}{y_t} \right|, \tag{2.25}$$

$$R^2 = 1 - \frac{\sum * t = 1^{n} \left(y_t - \widehat{y} * t \right)^2}{\sum * t = 1^{n} \left(y_t - \overline{y} \right)^2}, \tag{2.26}$$

où y_t représente la valeur observée, $\widehat{y}_t$ la valeur prédite, $\overline{y}$ la moyenne des valeurs observées et n le nombre d'observations évaluées.

La MAE mesure l'erreur moyenne en valeur absolue et s'exprime dans la même unité que la variable cible, ici les MW. La MSE pénalise davantage les erreurs importantes en raison du terme quadratique, mais elle s'exprime en MW^2, ce qui la rend moins directement interprétable. La RMSE, qui correspond à la racine carrée de la MSE, conserve cette sensibilité aux grandes erreurs tout en revenant à l'unité d'origine, soit les MW. La MAPE exprime l'erreur en pourcentage, mais son interprétation doit être faite avec prudence lorsque les valeurs observées y_t sont faibles ou proches de zéro. Enfin, le coefficient R^2 mesure la proportion de variance expliquée par le modèle par rapport à une prédiction fondée sur la moyenne des observations. Une valeur proche de 1 indique une bonne capacité explicative, tandis qu'une valeur négative peut indiquer que le modèle est moins performant qu'une prédiction naïve basée sur la moyenne.

Dans le cas des séries temporelles, il est important d'éviter les stratégies de validation qui mélangent aléatoirement les observations passées et futures, car elles peuvent introduire une fuite d'information. Une séparation chronologique des données est donc plus appropriée ; les modèles sont entraînés sur une période historique, validés sur une période ultérieure, puis testés sur une période encore plus récente. Cette stratégie permet de mieux représenter les conditions réelles de prédiction, où les observations futures ne sont pas disponibles au moment de l'entraînement.

La figure 2.7 présente le principe d'une validation temporelle simple. Dans le cadre d'une approche ligne par ligne, cette séparation doit être appliquée de manière cohérente à chaque ligne de transport, afin que les modèles soient comparés sur des périodes d'entraînement, de

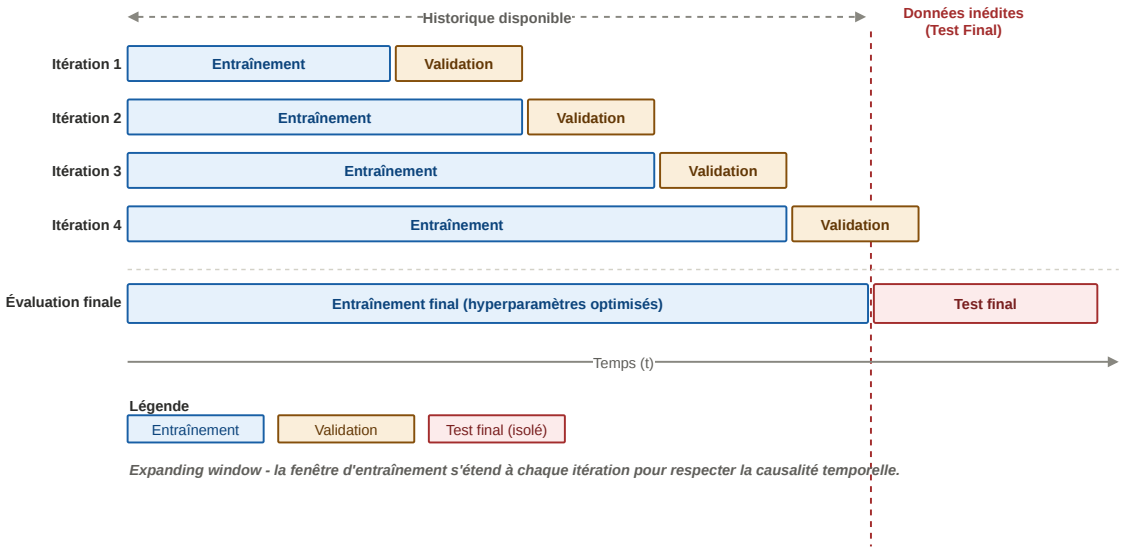

Figure 2.7 : Principe d'une séparation chronologique des données pour la validation des modèles de séries temporelles.

validation et de test équivalentes. Cette cohérence est importante pour éviter qu'un modèle soit évalué sur une période plus favorable qu'un autre et pour garantir une comparaison équitable entre les différentes familles de modèles.

2.7 SYNTHÈSE DES LACUNES ET POSITIONNEMENT DU MÉMOIRE

Cette dernière section synthétise les principaux constats issus de la littérature et précise le positionnement du présent mémoire.

2.7.1 LACUNES IDENTIFIÉES DANS LA LITTÉRATURE

La littérature montre que les approches de prédiction des pertes techniques ont fortement évolué au cours des dernières années. Les modèles physiques offrent une base rigoureuse et interprétable, mais nécessitent des informations détaillées sur le réseau, sa topologie, ses

paramètres électriques et ses conditions d'exploitation. Leur application peut donc devenir exigeante lorsque le réseau est vaste, dynamique et soumis à des variations importantes de charge ou de conditions environnementales.

Les modèles statistiques constituent une première famille d'approches fondées sur les données. Ils sont simples, interprétables et utiles pour représenter certaines structures temporelles, notamment les tendances, les saisonnalités et les dépendances aux valeurs passées. Cependant, leur efficacité peut être limitée lorsque les pertes dépendent de relations fortement non linéaires et multivariées. Les méthodes d'apprentissage automatique et d'apprentissage profond offrent alors des alternatives plus flexibles, capables d'exploiter simultanément plusieurs variables électriques, temporelles et environnementales. Leur performance dépend cependant de la qualité des données, du choix des variables, du réglage des hyperparamètres et de la stratégie de validation adoptée.

Plusieurs lacunes demeurent importantes. La disponibilité limitée de données réelles et publiques issues des réseaux de transport réduit la reproductibilité des études et rend difficile la comparaison directe entre les méthodes proposées dans la littérature. De plus, plusieurs travaux s'appuient sur des réseaux de test, des données simulées, des jeux de données synthétiques ou des variables fortement agrégées, ce qui peut réduire la représentativité des résultats par rapport aux conditions réelles d'exploitation. Enfin, les approches globales de prédiction des pertes à l'échelle du réseau complet peuvent masquer les comportements locaux propres aux lignes de transport.

2.7.2 POSITIONNEMENT DU PRÉSENT MÉMOIRE

Le présent mémoire se positionne par rapport à ces lacunes en proposant une comparaison de plusieurs familles de modèles à partir de données réelles issues du réseau de transport d'Hydro-Québec. L'étude ne suppose pas a priori la supériorité d'une approche particulière, mais vise plutôt à évaluer empiriquement la capacité de modèles statistiques, de modèles d'apprentissage automatique et de modèles d'apprentissage profond à prédire les pertes techniques dans un contexte opérationnel réel.

L'originalité du mémoire réside également dans l'adoption d'une modélisation à granularité fine, fondée sur la prédiction des pertes au niveau des lignes de transport. Cette approche vise à mieux représenter les conditions locales qui influencent les pertes, notamment les grandeurs électriques propres à chaque ligne, les variations temporelles et les conditions météorologiques associées à son environnement. Comme les pertes sont exprimées en MW, les prédictions obtenues ligne par ligne peuvent ensuite être agrégées afin d'estimer les pertes globales sur le périmètre étudié.

Ainsi, la revue de la littérature met en évidence les fondements qui structurent la suite du mémoire. Elle rappelle d'abord l'importance des connaissances physiques dans la compréhension des pertes, puis souligne l'intérêt de comparer plusieurs familles de modèles prédictifs à partir de données historiques réelles. Elle met enfin en avant la pertinence d'une modélisation ligne par ligne pour mieux représenter les spécificités locales du réseau de transport.

CHAPITRE III

MÉTHODOLOGIE POUR L'ANALYSE COMPARATIVE DE PLUSIEURS MODÈLES POUR LA PRÉDICTION DES PERTES D'ÉNERGIE

Ce chapitre présente la méthodologie retenue pour développer, entraîner et évaluer des modèles de prédiction des pertes techniques sur les lignes de transport d'HQ. La démarche s'inscrit dans la continuité du positionnement établi dans l'introduction et dans la revue de littérature. Elle repose sur une approche à granularité fine, dans laquelle la prédiction n'est pas réalisée uniquement à l'échelle globale du réseau, mais au niveau de chaque ligne de transport étudiée. Ce choix méthodologique vise à mieux représenter les spécificités locales des infrastructures, notamment la longueur des lignes, le niveau de tension, les postes connectés, les conditions électriques observées et les conditions météorologiques propres à l'environnement géographique de chaque ligne.

La méthodologie comporte cinq étapes principales. La première consiste à construire la variable cible à partir des mesures horaires de puissance active en entrée et en sortie des lignes. La deuxième étape vise à enrichir ces données électriques avec des variables météorologiques issues d'une correspondance spatiale entre les postes et les points de grille d'ECMWF Reanalysis v5 (ERA5), ainsi qu'avec des variables calendaires et structurelles. La troisième étape porte sur la fusion, le nettoyage, la synchronisation temporelle et la préparation des

données. La quatrième étape consiste à entraîner et comparer plusieurs familles de modèles, soit des modèles statistiques, des modèles d'apprentissage automatique tabulaire et un modèle séquentiel de type LSTM avec mécanisme d'attention. La cinquième étape concerne l'évaluation des performances, la définition des diagnostics graphiques utilisés dans le chapitre des résultats et la sauvegarde des artefacts nécessaires à la reproductibilité des expérimentations.

3.1 ARCHITECTURE GÉNÉRALE DU PIPELINE

La figure 3.1 présente l'architecture générale du pipeline expérimental. Les données électriques issues de l'environnement opérationnel sont d'abord ingérées dans une plateforme analytique infonuagique, puis structurées dans des tables de données. Les variables électriques, météorologiques et calendaires sont ensuite assemblées dans un jeu de données d'entraînement. Les modèles sont développés et suivis dans cet environnement, tandis que les expériences sont consignées à l'aide d'outils de suivi et de gestion du cycle de vie des modèles d'apprentissage automatique. Cette organisation permet de séparer clairement les étapes d'acquisition des données, d'ingénierie des variables, d'entraînement, de validation et d'inférence.

Cette architecture est cohérente avec l'objectif du mémoire, car elle permet de produire des modèles distincts par ligne de transport tout en conservant une chaîne de traitement reproductible. Elle facilite également le suivi des résultats ligne par ligne, ce qui est essentiel pour analyser les différences de comportement entre les infrastructures.

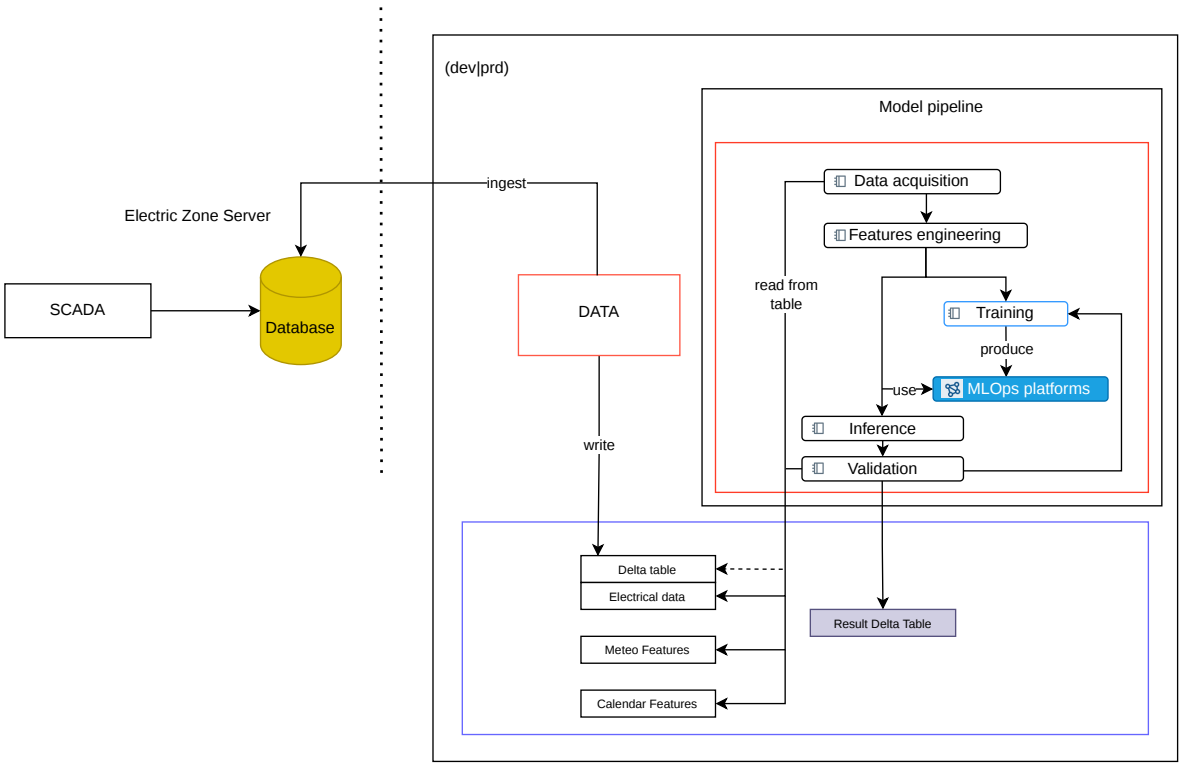

Figure 3.1 : Architecture générale du pipeline de prédiction des pertes techniques.

3.2 FORMULATION DU PROBLÈME DE PRÉDICTION

L'approche proposée est formulée comme un problème de prédiction supervisée. Pour chaque ligne de transport ℓ et chaque instant t, un ensemble de variables explicatives est utilisé pour prédire la perte technique associée à cette même ligne à un horizon futur h. Contrairement à une approche globale qui apprendrait un modèle unique à partir de variables agrégées sur l'ensemble du réseau, la méthodologie retenue repose sur l'entraînement d'un modèle distinct pour chaque ligne étudiée. Cette formulation permet de préserver l'information locale et de limiter l'effet de lissage associé à l'utilisation de variables globales.

La prédiction ligne par ligne peut être représentée par :

$$\widehat{y}_{\ell,t+h} = f_\ell \left(\mathbf{X}_{\ell,t-L+1:t},\, \mathbf{z}_{\ell,t+h} \right), \tag{3.1}$$

où $\widehat{y}_{\ell,t+h}$ désigne la perte prédite pour la ligne ℓ à l'horizon h, f_ℓ le modèle propre à cette ligne, $\mathbf{X}_{\ell,t-L+1:t}$ une fenêtre temporelle de longueur L contenant les observations passées, et $\mathbf{z}_{\ell,t+h}$ les variables connues ou disponibles pour l'horizon de prédiction, comme certaines variables calendaires ou météorologiques.

Dans le cas des modèles tabulaires, tels que Random Forest et XGBoost, les dépendances temporelles sont représentées au moyen de variables dérivées, notamment des retards temporels, des statistiques mobiles et des variables calendaires. Dans le cas du modèle LSTM avec attention, la structure séquentielle est directement fournie au modèle sous forme de fenêtres temporelles multivariées.

3.3 SOURCES DE DONNÉES ET CONSTRUCTION DE LA VARIABLE CIBLE

3.3.1 DONNÉES ÉLECTRIQUES ISSUES DU RÉSEAU DE TRANSPORT

Les données électriques utilisées dans cette étude proviennent de l'historique d'exploitation du réseau de transport. Elles contiennent les mesures horaires de puissance active associées aux postes connectés par les lignes de transport, ainsi que des informations structurelles telles que la longueur des lignes, le niveau de tension nominal en Kilovolt (kV) et les postes reliés à chaque ligne. La logique de construction de la variable cible repose sur le suivi de la puissance active qui entre dans un poste, transite par une ligne, puis ressort par un autre poste.

Pour une ligne de transport ℓ connectant un poste amont et un poste aval, la perte active horaire est calculée par la différence entre la puissance active mesurée à l'entrée et la puissance active mesurée à la sortie :

$$y_{\ell,t} = P^{\text{entree}}_{\ell,t} - P^{\text{sortie}}_{\ell,t}, \tag{3.2}$$

où $y_{\ell,t}$ représente la perte technique de la ligne ℓ à l'instant t, $P^{\text{entree}}_{\ell,t}$ la puissance active injectée à l'entrée de la ligne et $P^{\text{sortie}}_{\ell,t}$ la puissance active mesurée en sortie. Cette formulation est appliquée en respectant la convention de signe retenue pour les flux de puissance. Les observations incohérentes, par exemple celles associées à des conventions de mesure différentes ou à des anomalies de capteurs, sont traitées lors de l'étape de contrôle de qualité.

Lorsque l'analyse nécessite une comparaison relative entre les lignes, le taux de perte peut également être calculé sous la forme :

$$L_{\ell,t} = \frac{P^{\text{entree}}_{\ell,t} - P^{\text{sortie}}_{\ell,t}}{P^{\text{entree}}_{\ell,t}} \times 100. \tag{3.3}$$

Dans la suite de la méthodologie, la variable cible principale demeure la perte active horaire calculée au niveau de chaque ligne, tandis que le taux de perte peut servir à certaines analyses comparatives.

3.3.2 LOCALISATION DES POSTES ET ENRICHISSEMENT MÉTÉOROLOGIQUE

Les pertes techniques peuvent être influencées par les conditions environnementales locales, notamment la température, le vent, l'humidité, les précipitations et les phénomènes associés aux conditions hivernales. Pour tenir compte de cette dépendance, les postes connectés à chaque ligne sont associés à leurs coordonnées géographiques. Ces coordonnées sont ensuite mises en correspondance avec les points de grille du modèle météorologique ERA5 afin d'obtenir des variables météorologiques localisées autour des postes.

Dans le cadre de l'entraînement, les variables météorologiques sont construites à partir des données ERA5. Pour la phase d'inférence à un horizon de 48 h, les variables météorologiques futures sont remplacées par les prévisions issues du modèle météo High Resolution Deterministic Prediction System (HRDPS).

Pour chaque ligne ℓ, on note $\mathcal{P}_\ell$ l'ensemble des postes connectés à cette ligne. Pour une variable météorologique m, les valeurs associées aux postes sont agrégées afin de produire des variables représentatives de l'environnement local de la ligne. Les agrégations utilisées

incluent notamment la moyenne, le minimum et le maximum :

$$m_{\ell,t}^{\text{moy}} = \frac{1}{|\mathcal{P}_\ell|} \sum_{p \in \mathcal{P}_\ell} m_{p,t}, \tag{3.4}$$

$$m_{\ell,t}^{\min} = \min_{p \in \mathcal{P}_\ell} m_{p,t}, \tag{3.5}$$

$$m_{\ell,t}^{\max} = \max_{p \in \mathcal{P}_\ell} m_{p,t}. \tag{3.6}$$

Cette stratégie évite l'utilisation d'une météo moyenne à l'échelle du réseau complet. Elle est cohérente avec l'approche désagrégée adoptée dans ce mémoire, car elle conserve une information météorologique localisée autour de chaque ligne de transport. Les variables ainsi construites permettent aux modèles d'exploiter les contrastes climatiques entre les différentes zones traversées par les lignes.

La figure 3.2 illustre un exemple d'évolution conjointe normalisée de la température et des pertes pour une ligne donnée. Cette visualisation ne vise pas à établir une relation causale directe, mais elle montre que les pertes et les variables météorologiques peuvent présenter des dynamiques temporelles distinctes et parfois associées. Elle justifie donc l'intégration des conditions environnementales dans le jeu de données final.

3.3.3 VARIABLES CALENDAIRES ET VARIABLES STRUCTURELLES

Les variables calendaires sont ajoutées afin de représenter les régularités temporelles susceptibles d'influencer les pertes. Elles incluent notamment l'heure de la journée, le jour de la semaine, le mois, la saison, les jours fériés et certains indicateurs liés aux périodes particulières. Les variables cycliques, comme l'heure ou le mois, sont encodées sous forme sinusoïdale afin

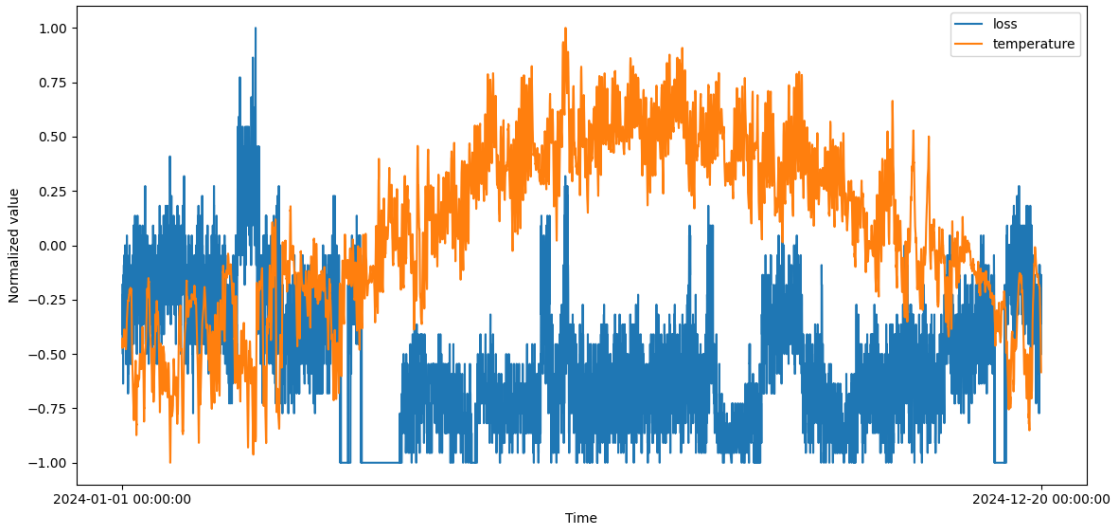

Figure 3.2 : Évolution conjointe normalisée de la température et des pertes pour une ligne donnée.

d'éviter les discontinuités artificielles entre deux valeurs proches, par exemple entre 23 h et 0 h :

$$\text{heure}_{\sin}(t) = \sin\left(2\pi \frac{\text{heure}(t)}{24}\right), \tag{3.7}$$

$$\text{heure}_{\cos}(t) = \cos\left(2\pi \frac{\text{heure}(t)}{24}\right). \tag{3.8}$$

Les variables structurelles décrivent les propriétés propres aux lignes, telles que la longueur, le niveau de tension nominale et le nombre de postes associés. Ces variables fournissent aux modèles des informations stables sur les caractéristiques physiques de chaque infrastructure.

3.4 FUSION, NETTOYAGE ET PRÉPARATION DES DONNÉES

3.4.1 FUSION TEMPORELLE ET SPATIALE DES DONNÉES

La construction du jeu de données final repose sur la fusion de plusieurs sources hétérogènes. Les données électriques sont d'abord structurées par ligne et par pas de temps horaire. Les variables météorologiques sont ensuite associées aux postes connectés à chaque ligne, puis agrégées à l'échelle de la ligne. Enfin, les variables calendaires sont générées directement à partir des horodatages.

Le vecteur final de variables explicatives associé à une ligne ℓ et à un instant t peut être noté :

$$\mathbf{x}_{\ell,t} = \left[\mathbf{x}_{\ell,t}^{\text{elec}},\ \mathbf{x}_{\ell,t}^{\text{meteo}},\ \mathbf{x}_{t}^{\text{cal}},\ \mathbf{x}_{\ell}^{\text{struct}}\right], \tag{3.9}$$

où $\mathbf{x}_{\ell,t}^{\text{elec}}$ regroupe les variables électriques, $\mathbf{x}_{\ell,t}^{\text{meteo}}$ les variables météorologiques agrégées, $\mathbf{x}_{t}^{\text{cal}}$ les variables calendaires et $\mathbf{x}_{\ell}^{\text{struct}}$ les caractéristiques propres à la ligne.

3.4.2 TRAITEMENT DES VALEURS MANQUANTES ET ABERRANTES

Les données issues d'un réseau réel peuvent contenir des valeurs manquantes, des doublons, des changements de convention de mesure ou des valeurs aberrantes liées à des anomalies de capteurs, à des interventions d'exploitation ou à des événements particuliers. Une étape de contrôle de qualité est donc appliquée avant l'entraînement des modèles.

Les valeurs manquantes sont traitées en fonction de leur durée et de leur fréquence. Les absences courtes peuvent être interpolées lorsque la continuité temporelle de la variable le

permet. Les périodes plus longues ou les observations incohérentes sont écartées afin de limiter l'introduction de bruit artificiel dans l'apprentissage. Les valeurs aberrantes sont détectées à partir de règles statistiques, telles que l'écart interquartile ou l'analyse des scores normalisés, puis sont corrigées ou retirées selon leur nature.

Une attention particulière est portée à la fuite de données. Toutes les opérations nécessitant l'apprentissage de paramètres, notamment la normalisation, sont ajustées sur les données d'entraînement seulement, puis appliquées aux jeux de validation et de test. Cette précaution est essentielle pour évaluer les performances dans des conditions proches d'une utilisation réelle.

3.4.3 NORMALISATION DES VARIABLES

Les variables numériques sont mises à l'échelle afin de faciliter l'apprentissage des modèles, en particulier pour les réseaux de neurones. La normalisation standard consiste à transformer une variable x selon :

$$x' = \frac{x - \mu_{\text{train}}}{\sigma_{\text{train}}}, \tag{3.10}$$

où μ_{train} et σ_{train} représentent respectivement la moyenne et l'écart type calculés uniquement sur le jeu d'entraînement. Dans l'approche ligne par ligne, cette normalisation peut être réalisée séparément pour chaque ligne afin de tenir compte des différences d'échelle entre les infrastructures étudiées.

3.5 ANALYSE EXPLORATOIRE DES DONNÉES

Avant l'entraînement des modèles, une analyse exploratoire a été réalisée afin d'identifier les principales structures temporelles présentes dans les pertes. Les visualisations produites mettent en évidence des variations horaires, saisonnières et interannuelles, ce qui confirme l'intérêt d'intégrer des variables calendaires et météorologiques au processus de prédiction.

La figure 3.3 illustre l'effet de l'heure de la journée sur la distribution des pertes. Les différences observées entre les heures indiquent que les pertes ne sont pas uniformes au cours d'une journée. Cette variabilité peut être liée aux cycles de charge, aux conditions d'exploitation et aux habitudes de consommation. Elle justifie l'ajout de variables calendaires telles que l'heure, le mois, la saison, les jours ouvrables et les jours fériés.

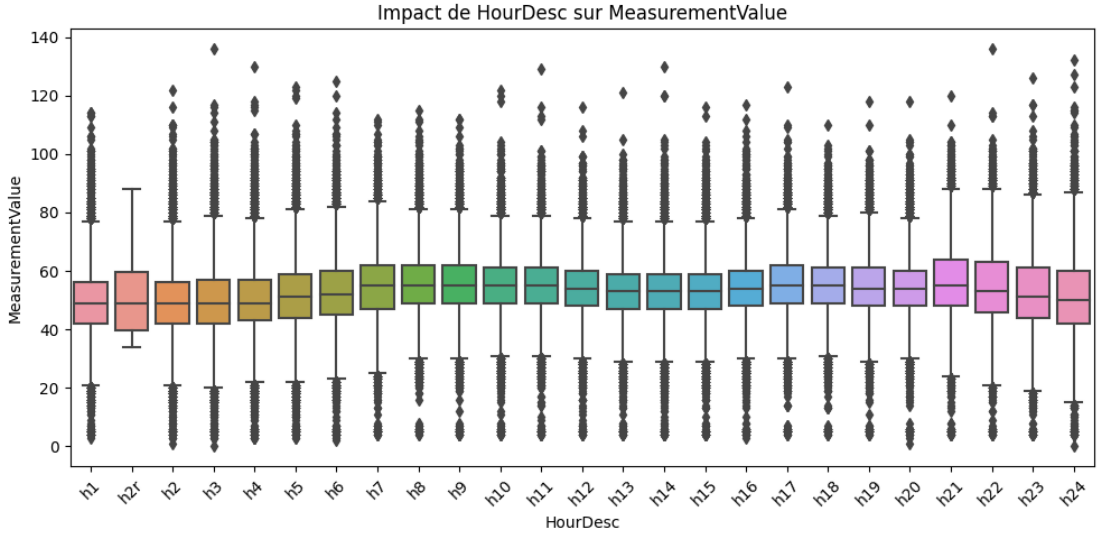

Figure 3.3 : Distribution des pertes en fonction de l'heure de la journée.

La figure 3.4 présente l'évolution mensuelle moyenne des pertes pour plusieurs années. Elle montre l'existence d'une variabilité saisonnière importante, cohérente avec les fluctuations de

charge et les conditions climatiques propres au Québec. Cette observation confirme la pertinence des variables temporelles et météorologiques pour la prédiction des pertes techniques.

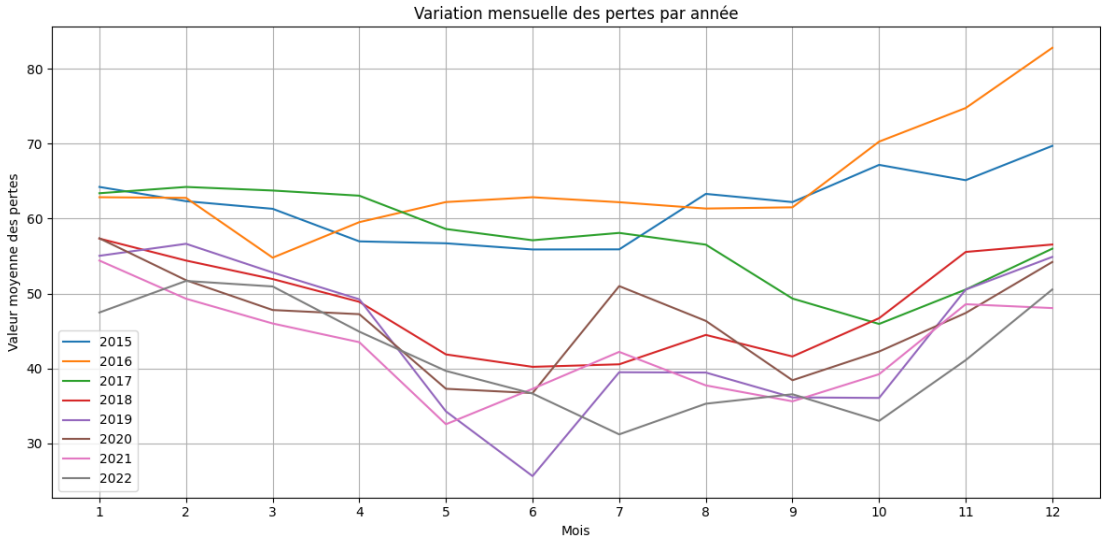

Figure 3.4 : Variation mensuelle moyenne des pertes pour différentes années.

Ces analyses exploratoires ne remplacent pas l'évaluation expérimentale des modèles. Elles permettent toutefois d'orienter la construction du jeu de données et de confirmer l'intérêt des variables temporelles, météorologiques et structurelles dans la prédiction des pertes techniques.

3.6 DÉCOUPAGE TEMPOREL DES DONNÉES ET STRATÉGIE DE VALIDATION

Comme les données étudiées sont des séries temporelles, le découpage des jeux d'entraînement, de validation et de test est réalisé de manière chronologique. Les observations les plus anciennes sont utilisées pour l'entraînement, une période intermédiaire sert à la validation et les observations les plus récentes sont réservées au test final. Cette stratégie permet d'éviter de mélanger des observations futures avec des observations passées et d'obtenir une évaluation plus réaliste de la capacité prédictive des modèles.

Dans cette étude, le découpage suit une logique temporelle de type entraînement, validation et test. Une configuration possible consiste à utiliser environ 70 % des observations pour l'entraînement, 15 % pour la validation et 15 % pour le test. Les proportions exactes sont ajustées selon la disponibilité des données pour chaque ligne, mais l'ordre chronologique est toujours conservé.

Pour les modèles séquentiels, les données sont transformées en fenêtres glissantes. Une fenêtre de longueur L contient les observations passées utilisées comme entrée du modèle :

$$\mathbf{X}_{\ell,t-L+1:t} = \left(\mathbf{x}_{\ell,t-L+1}, \mathbf{x}_{\ell,t-L+2}, \ldots, \mathbf{x}_{\ell,t}\right). \tag{3.11}$$

Le modèle produit ensuite une prédiction à plusieurs horizons futurs. Dans le cas du modèle LSTM avec attention, l'horizon principal considéré est de 48 h.

3.7 MODÈLES DÉVELOPPÉS ET COMPARÉS

Trois grandes familles de modèles sont comparées afin de couvrir différents niveaux de complexité et différentes hypothèses de modélisation. Les modèles statistiques classiques servent de références temporelles interprétables. Les modèles d'apprentissage automatique tabulaire permettent de capturer des relations non linéaires entre les variables électriques, météorologiques, calendaires et structurelles. Enfin, le modèle LSTM avec mécanisme d'attention vise à exploiter explicitement la structure séquentielle des données et les dépendances temporelles à court et moyen terme.

3.7.1 MODÈLES STATISTIQUES CLASSIQUES

Les modèles ARIMA et Prophet sont utilisés comme approches statistiques de référence. ARIMA permet de modéliser les dépendances temporelles linéaires à partir des valeurs passées de la série. Pour une ligne donnée, la série des pertes peut être notée $\{y_{\ell,t}\}_{t=1}^{T}$. Un modèle ARIMA cherche à représenter la dynamique de cette série par une combinaison de termes autorégressifs, de différenciation et de moyennes mobiles.

Prophet repose sur une décomposition additive de la série temporelle :

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t, \tag{3.12}$$

où $g(t)$ représente la tendance, $s(t)$ la saisonnalité, $h(t)$ les effets calendaires particuliers et ϵ_t le bruit résiduel. Ce modèle est pertinent pour explorer la contribution des composantes saisonnières et calendaires dans la prédiction des pertes.

Ces modèles sont entraînés séparément pour les lignes étudiées lorsque la quantité de données disponible est suffisante. Ils constituent des points de comparaison utiles pour évaluer l'apport des méthodes d'apprentissage automatique et d'apprentissage profond.

3.7.2 MODÈLES D'APPRENTISSAGE AUTOMATIQUE TABULAIRE

Random Forest et XGBoost sont utilisés pour modéliser les relations non linéaires entre les pertes et les variables électriques, météorologiques, calendaires et structurelles. Contrairement aux modèles statistiques univariés, ces algorithmes peuvent exploiter simultanément un grand nombre de variables explicatives. Cependant, comme ils ne reçoivent pas directement une

séquence temporelle, des caractéristiques dérivées sont ajoutées pour représenter la dynamique des pertes.

Les variables dérivées incluent notamment les retards temporels de la variable cible et de certaines variables explicatives, les moyennes mobiles, les écarts types sur fenêtres temporelles, les variations horaires, ainsi que les variables calendaires. Le problème d'apprentissage peut alors être écrit sous la forme :

$$\widehat{y}_{\ell,t+h} = f_{\ell}^{\mathrm{ML}} \left(\phi(\mathbf{x}_{\ell,t}, \mathbf{x}_{\ell,t-1}, \ldots, \mathbf{x}_{\ell,t-p}) \right), \tag{3.13}$$

où $\phi(\cdot)$ désigne la fonction de construction des caractéristiques tabulaires à partir de l'historique disponible.

Random Forest repose sur l'agrégation de plusieurs arbres de décision :

$$\widehat{y}_{\ell,t} = \frac{1}{M} \sum_{m=1}^{M} T_{m,\ell}(\mathbf{x}_{\ell,t}), \tag{3.14}$$

où M est le nombre d'arbres et $T_{m,\ell}$ le m-ième arbre entraîné pour la ligne ℓ.

XGBoost construit progressivement un modèle additif composé d'arbres successifs :

$$\widehat{y}_{\ell,t} = \sum_{m=1}^{M} \eta f_{m,\ell}(\mathbf{x}_{\ell,t}), \tag{3.15}$$

où $f_{m,\ell}$ représente un arbre ajouté à l'itération m et η le taux d'apprentissage.

3.7.3 MODÈLE LSTM AVEC MÉCANISME D'ATTENTION

Le modèle LSTM avec mécanisme d'attention constitue l'approche séquentielle principale de cette étude. Il est conçu pour produire des prédictions à un horizon de 48 h pour chaque ligne de transport. Le modèle reçoit en entrée une fenêtre temporelle multivariée contenant les variables électriques, météorologiques, calendaires et structurelles associées à la ligne.

L'architecture générale comprend quatre composantes. La première correspond à la fenêtre d'entrée multivariée. La deuxième est composée d'une ou plusieurs couches LSTM qui transforment la séquence d'entrée en une suite d'états cachés. La troisième applique un mécanisme d'attention temporelle afin de pondérer les instants passés selon leur contribution à la prédiction. La dernière composante est une couche dense qui produit les pertes prédites pour l'horizon souhaité.

Pour une séquence d'entrée, le LSTM produit une suite d'états cachés $(\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_L)$. Le mécanisme d'attention calcule un score e_t pour chaque état caché, puis une pondération α_t :

$$e_t = \mathbf{v}^\top \tanh(W_h \mathbf{h}_t + \mathbf{b}), \tag{3.16}$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{L} \exp(e_j)}. \tag{3.17}$$

Le vecteur de contexte transmis à la couche de sortie est ensuite obtenu par une somme pondérée des états cachés :

$$\mathbf{c} = \sum_{t=1}^{L} \alpha_t \mathbf{h}_t. \tag{3.18}$$

Cette formulation permet au modèle d'accorder plus d'importance aux instants de la fenêtre temporelle qui contribuent davantage à la prédiction des pertes.

Une configuration de base utilisée pour les expérimentations est présentée dans le tableau 3.1. Ces paramètres constituent une base expérimentale et peuvent être ajustés selon la ligne étudiée, la taille du jeu de données et les performances observées en validation.

Tableau 3.1 : Configuration de base du modèle LSTM avec attention.

Paramètre	**Valeur de base**
Longueur de la fenêtre temporelle	24 h
Horizon de prédiction principal	48 h
Dimension cachée	128
Nombre de couches LSTM	2
Dropout	0,2
Taille de lot	1024
Taux d'apprentissage	3×10^{-3}
Nombre maximal d'époques	30
Patience pour l'arrêt anticipé	5

3.8 FONCTION DE COÛT, OPTIMISATION ET ARRÊT ANTICIPÉ

L'entraînement des modèles de régression vise à minimiser l'écart entre les pertes observées et les pertes prédites. Pour les modèles neuronaux, la fonction de coût principale est l'erreur quadratique moyenne :

$$\mathrm{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2 , \tag{3.19}$$

où y_i représente la valeur observée et $\widehat{y}_i$ la valeur prédite. L'erreur quadratique moyenne pénalise davantage les grandes erreurs, ce qui est utile lorsque les écarts importants doivent être fortement pris en compte.

Afin de limiter le surapprentissage, l'entraînement du modèle LSTM avec attention utilise un mécanisme d'arrêt anticipé. À chaque époque, la perte calculée sur le jeu de validation est comparée à la meilleure perte observée jusque-là. Si aucune amélioration n'est observée pendant un nombre défini d'époques, l'entraînement est arrêté et le meilleur modèle est conservé.

Les courbes d'apprentissage sont utilisées comme outils de diagnostic afin de suivre l'évolution de la perte d'entraînement et de validation au fil des époques. Elles permettent de vérifier la convergence du modèle, de détecter un éventuel surapprentissage et de comparer le comportement d'apprentissage d'une ligne à l'autre. Comme ces courbes reflètent des résultats expérimentaux propres aux modèles entraînés, elles sont présentées et discutées dans le chapitre consacré aux résultats.

3.9 ÉVALUATION DES PERFORMANCES ET DIAGNOSTICS

L'évaluation des modèles repose sur plusieurs métriques complémentaires afin de caractériser différents aspects de la performance prédictive. La racine de l'erreur quadratique moyenne mesure l'ampleur moyenne des erreurs en conservant l'unité de la variable cible. L'erreur absolue moyenne est moins sensible aux grandes erreurs que la RMSE. Le coefficient de détermination R^2 mesure la part de variance expliquée par le modèle. Enfin, le Symmetric Mean absolute percentage error (sMAPE) permet de comparer l'erreur relative tout en réduisant certaines instabilités du MAPE lorsque les valeurs observées sont faibles ou proches de zéro.

Les principales métriques utilisées sont définies comme suit :

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y}_i|, \tag{3.20}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}, \tag{3.21}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}, \tag{3.22}$$

$$\text{sMAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{2|y_i - \widehat{y}_i|}{|y_i| + |\widehat{y}_i| + \epsilon}, \tag{3.23}$$

où ϵ est une petite constante ajoutée afin d'éviter une division par zéro.

Au-delà des métriques numériques, des diagnostics graphiques sont produits pour analyser le comportement des modèles. Le graphique des valeurs réelles en fonction des valeurs prédites permet de vérifier si les points sont proches de la diagonale idéale. La distribution des résidus permet d'observer si les erreurs sont centrées autour de zéro et d'identifier la présence d'erreurs extrêmes. Le graphique des résidus en fonction des prédictions permet de détecter des structures non modélisées ou une variance d'erreur qui augmente avec le niveau de perte. Les prédictions sont également visualisées ligne par ligne afin de repérer les infrastructures pour lesquelles le modèle produit des prédictions trop lissées, des pics mal représentés ou des comportements atypiques. Ces figures, étant directement liées aux performances obtenues, sont réservées au chapitre des résultats.

3.10 ORGANISATION EXPÉRIMENTALE ET SUIVI DES MODÈLES

Les expérimentations sont organisées de manière à assurer la traçabilité des données, des modèles et des résultats. Les données préparées sont stockées dans un environnement de

catalogue de données, puis utilisées par les scripts d'entraînement. Pour chaque ligne de transport, les modèles sont entraînés et évalués séparément. Cette organisation permet d'obtenir des métriques propres à chaque ligne, mais aussi des métriques agrégées sur l'ensemble des lignes étudiées.

Un outil de suivi des expérimentations consigne, pour chaque exécution, les métriques d'entraînement et de validation, les hyperparamètres, les artefacts du modèle ainsi que les figures de diagnostic. Pour le modèle LSTM avec attention, les artefacts sauvegardés incluent notamment les poids du modèle, le normaliseur ajusté sur les données d'entraînement, la configuration du modèle et la liste des lignes entraînées.

Tableau 3.2 : Principaux artefacts sauvegardés pour assurer la reproductibilité des expérimentations.

Artefact	**Description et rôle**
Poids du modèle	Rechargement des paramètres du modèle entraîné pour la validation, le test ou l'inférence.
Normaliseur	Application de la même mise à l'échelle lors des phases de validation, de test et d'inférence.
Configuration du modèle	Conservation des hyperparamètres, des dimensions d'entrée, de la fenêtre historique et de l'horizon de prédiction.
Journal des métriques	Suivi et comparaison des performances d'apprentissage entre les différents modèles et configurations.
Figures de diagnostic	Documents produits pour l'analyse des résultats, incluant la convergence, les valeurs prédites, les résidus et les erreurs absolues.
Fichier de suivi des lignes	Registre des lignes incluses dans l'entraînement.

L'organisation expérimentale distingue deux types de traitements. Le premier correspond au workflow d'entraînement, qui prépare les données, entraîne un modèle par ligne, évalue les performances et sauvegarde les artefacts. Le second correspond au workflow d'inférence, qui

assemble les variables nécessaires pour les horizons futurs et applique les modèles enregistrés afin de produire les prédictions de pertes.

3.11 SYNTHÈSE DE LA MÉTHODOLOGIE

La méthodologie proposée met en place un cadre complet pour la prédiction des pertes techniques sur les lignes de transport. Elle combine la construction d'une variable cible à partir des mesures de puissance active, l'enrichissement des données par des variables météorologiques locales issues d'une correspondance spatiale avec les postes, l'ajout de variables calendaires et structurelles, ainsi qu'une stratégie d'apprentissage supervisé adaptée aux séries temporelles.

Le choix d'un modèle distinct par ligne constitue l'élément central de cette méthodologie. Il permet de préserver les spécificités locales qui peuvent être perdues dans une approche globale fondée sur des variables agrégées. La comparaison entre ARIMA, Prophet, Random Forest, XGBoost et LSTM avec attention permet ensuite d'évaluer différentes hypothèses de modélisation, depuis les approches statistiques interprétables jusqu'aux architectures séquentielles capables de représenter des dépendances temporelles complexes.

Les étapes décrites dans ce chapitre préparent l'analyse expérimentale présentée au chapitre suivant, où les performances des modèles seront comparées à l'aide de métriques quantitatives et de diagnostics graphiques par ligne.

CHAPITRE IV

RÉSULTATS ET DISCUSSION

Ce chapitre présente les résultats obtenus pour la prédiction des pertes techniques sur les lignes de transport d'HQ. Conformément à la méthodologie proposée, l'évaluation est réalisée à une granularité fine, c'est-à-dire au niveau des lignes de transport, plutôt qu'à l'échelle globale du réseau. Cette stratégie vise à mieux représenter les comportements locaux propres à chaque ligne, en tenant compte des variables électriques, environnementales et calendaires associées à son fonctionnement.

Les modèles évalués appartiennent à trois familles principales. La première regroupe les modèles statistiques de séries temporelles, notamment ARIMA et Prophet. La deuxième comprend les modèles d'apprentissage automatique classiques, soit Random Forest et XGBoost. La troisième correspond au modèle d'apprentissage profond proposé, fondé sur une architecture LSTM enrichie par un mécanisme d'attention temporelle. Cette dernière approche est évaluée afin de mesurer l'apport d'une modélisation séquentielle dans la prédiction des pertes techniques horaires.

4.1 JEU DE DONNÉES ET PROTOCOLE EXPÉRIMENTAL

Le jeu de données utilisé provient de l'historique du réseau de transport d'HQ et couvre une période de 252 mois, soit de janvier 2005 à décembre 2025. Les données sont organisées à l'échelle horaire et regroupent des variables électriques, météorologiques et calendaires. Après nettoyage, normalisation et fusion des différentes sources de données, environ 160 000 observations et 47 variables explicatives ont été retenues pour l'entraînement, la validation et le test des modèles.

La variable cible correspond aux pertes techniques calculées au niveau de chaque ligne de transport. Pour une ligne donnée, ces pertes sont obtenues à partir de la différence entre la puissance active entrante et la puissance active sortante. Cette formulation permet d'évaluer les modèles à l'échelle locale des lignes, tout en conservant la possibilité d'agréger les prédictions afin d'obtenir une vision plus globale des pertes sur l'ensemble du périmètre étudié.

Afin de respecter la structure temporelle des données et d'éviter toute fuite d'information entre les périodes d'apprentissage et de test, une séparation chronologique a été adoptée. Les 70 % premières observations ont été utilisées pour l'entraînement, les 15 % suivantes pour la validation et les 15 % restantes pour le test final. Tous les modèles ont été évalués sur le même ensemble de test afin de garantir une comparaison équitable.

Les performances sont analysées à l'aide de plusieurs métriques complémentaires, notamment la racine de l'erreur quadratique moyenne (RMSE), l'erreur absolue moyenne (MAE), le coefficient de détermination (R^2) et l'erreur relative. Lorsque les pertes présentent des valeurs faibles ou proches de zéro, le MAPE peut devenir instable. Dans ce contexte, le sMAPE peut être utilisé comme un indicateur relatif plus robuste.

4.2 COMPARAISON GLOBALE DES MODÈLES

Le tableau 4.1 présente les performances obtenues par les différents modèles sur l'ensemble de test. Les résultats permettent de comparer les approches statistiques, les modèles d'apprentissage automatique et le modèle LSTM avec attention.

Tableau 4.1 : Comparaison des performances des modèles sur l'ensemble de test.

Modèle	**RMSE**	**sMAPE**	R^2	**MAE**
ARIMA	0.145	8.5 %	0.72	0.205
Prophet	0.130	7.1 %	0.80	0.095
Random Forest	0.117	6.2 %	0.84	0.090
XGBoost	0.093	5.1 %	0.90	0.075
LSTM avec attention	0.064	3.7 %	0.97	0.045

Les résultats montrent une progression claire des performances lorsque l'on passe des modèles statistiques aux modèles d'apprentissage automatique, puis vers le modèle LSTM avec attention. Les modèles statistiques constituent des bases de comparaison utiles, mais leurs performances demeurent limitées face à la nature non linéaire et multivariée des pertes techniques. Les modèles Random Forest et XGBoost améliorent les résultats grâce à leur capacité à capturer des interactions complexes entre les variables électriques, météorologiques et calendaires.

Le modèle XGBoost obtient les meilleures performances parmi les modèles classiques, avec un RMSE de 0.093 et un $\boldsymbol{R^2}$ de 0.90. Toutefois, il ne modélise pas explicitement les dépendances temporelles à long terme entre les observations successives. À l'inverse, le modèle LSTM avec un mécanisme d'attention exploite directement les séquences horaires et permet d'accorder un poids plus important aux instants passés les plus pertinents pour la prédiction. Les résultats obtenus montrent que cette architecture réduit l'erreur de prédiction par rapport à XGBoost, ce

qui confirme l'intérêt de la modélisation séquentielle pour la prédiction des pertes techniques, ligne par ligne.

4.3 RÉSULTATS DES MODÈLES STATISTIQUES

Les modèles statistiques de séries temporelles constituent la première base de comparaison. Le modèle ARIMA obtient une performance inférieure à celle des autres approches, avec un RMSE de 0.145 et un coefficient de détermination de 0.72. Cette performance s'explique par la structure essentiellement linéaire du modèle, qui permet de représenter certaines dépendances temporelles locales, mais qui demeure limitée lorsque les pertes dépendent simultanément de plusieurs variables exogènes.

Le modèle Prophet[4] améliore les performances par rapport à ARIMA réduisant le RMSE de 0.015. Cette amélioration peut être attribuée à sa capacité à représenter les tendances et les composantes saisonnières. Cette propriété est pertinente dans le contexte des pertes électriques, qui peuvent être influencées par des cycles journaliers, hebdomadaires et annuels. Toutefois, Prophet demeure limité lorsqu'il s'agit de modéliser les interactions complexes entre les variables électriques, météorologiques et opérationnelles.

Ces résultats confirment que les modèles statistiques sont utiles comme références, notamment en raison de leur simplicité et de leur interprétabilité. Cependant, leur capacité de prédiction reste limitée lorsque la dynamique des pertes devient fortement non linéaire, multivariée et dépendante des conditions locales de chaque ligne.

4. Lien de la documentation officielle du modèle Prophet de Facebook : `https://facebook.github.io/prophet/`

4.4 RÉSULTATS DES MODÈLES D'APPRENTISSAGE AUTOMATIQUE

Les modèles Random Forest et XGBoost améliorent les performances par rapport aux approches statistiques. Cette amélioration montre l'intérêt des méthodes d'apprentissage automatique pour modéliser les relations non linéaires entre les variables explicatives et les pertes techniques.

Le modèle Random Forest obtient des performances proches de celles de Prophet en termes de RMSE et d'erreur relative, tout en réduisant légèrement le MAE de 0.005. Cette performance indique que l'agrégation de plusieurs arbres de décision permet de mieux représenter certaines interactions entre variables, tout en limitant le risque de surapprentissage.

Le modèle XGBoost obtient les meilleurs résultats parmi les modèles classiques évalués. Sa structure fondée sur le gradient boosting lui permet de corriger progressivement les erreurs des arbres précédents et de mieux capturer les interactions non linéaires. Dans le contexte de ce mémoire, cette propriété est particulièrement importante, car les pertes techniques dépendent à la fois des conditions électriques de la ligne, des variables météorologiques locales et des effets calendaires.

Malgré ces performances, XGBoost reste un modèle principalement tabulaire. Il dépend fortement des variables construites lors de l'ingénierie des caractéristiques, notamment les variables retardées, les agrégats temporels ou les encodages calendaires. Il ne traite donc pas directement la séquence complète des observations passées comme le fait un modèle récurrent. Cette limite justifie l'évaluation du modèle LSTM avec attention.

4.5 RÉSULTATS DU MODÈLE LSTM AVEC ATTENTION

Le modèle LSTM avec attention a été développé afin de mieux exploiter la structure séquentielle des données horaires. Contrairement aux modèles statistiques et aux modèles d'apprentissage automatique classiques, cette architecture traite une fenêtre temporelle d'observations passées et apprend les dépendances entre les états successifs. Le mécanisme d'attention complète cette structure en permettant au modèle d'identifier les instants de la séquence qui contribuent le plus à la prédiction.

Dans le cadre de ce mémoire, cette propriété est particulièrement pertinente. Les pertes techniques ne dépendent pas uniquement de l'état instantané d'une ligne, mais également de l'évolution récente de la charge, de la température, du vent et des autres variables explicatives. Le modèle LSTM avec attention est donc mieux adapté pour représenter les effets retardés, les transitions rapides et les dépendances temporelles qui peuvent être partiellement manquées par les modèles classiques.

La figure 4.1 présente l'évolution des pertes d'entraînement et de validation du modèle LSTM avec attention. La diminution rapide des deux courbes au cours des premières époques indique que le modèle apprend efficacement les structures présentes dans les données. La stabilisation progressive des pertes suggère que le modèle atteint ensuite un régime de convergence. L'absence d'écart croissant important entre les courbes d'entraînement et de validation indique que le modèle ne présente pas de signe marqué de surapprentissage sur l'exemple présenté.

La figure 4.2 présente trois diagnostics complémentaires. Le graphique des résidus permet d'analyser la dispersion des erreurs autour de zéro. Le nuage de points entre valeurs réelles

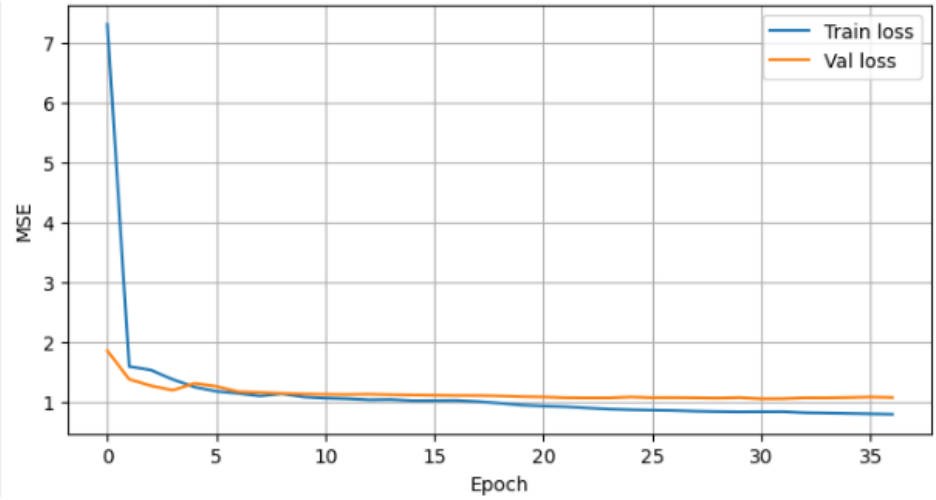

Figure 4.1 : Évolution des pertes d'entraînement et de validation du modèle LSTM avec attention.

et valeurs prédites permet d'évaluer la capacité du modèle à reproduire les pertes observées. Enfin, la distribution des erreurs absolues montre la concentration des erreurs autour de faibles valeurs, tout en mettant en évidence quelques erreurs plus importantes associées à des situations plus difficiles à prédire.

Les diagnostics visuels confirment l'intérêt du modèle LSTM avec attention. La proximité des points avec la diagonale dans le graphique réel-prédit indique une bonne capacité à reproduire la tendance générale des pertes. La distribution des erreurs absolues montre qu'une grande partie des erreurs est concentrée autour de faibles valeurs, ce qui traduit une bonne précision globale. Les erreurs les plus importantes semblent associées à des variations brusques ou à des régimes de fonctionnement moins fréquents, ce qui constitue une limite normale dans un contexte de données opérationnelles réelles.

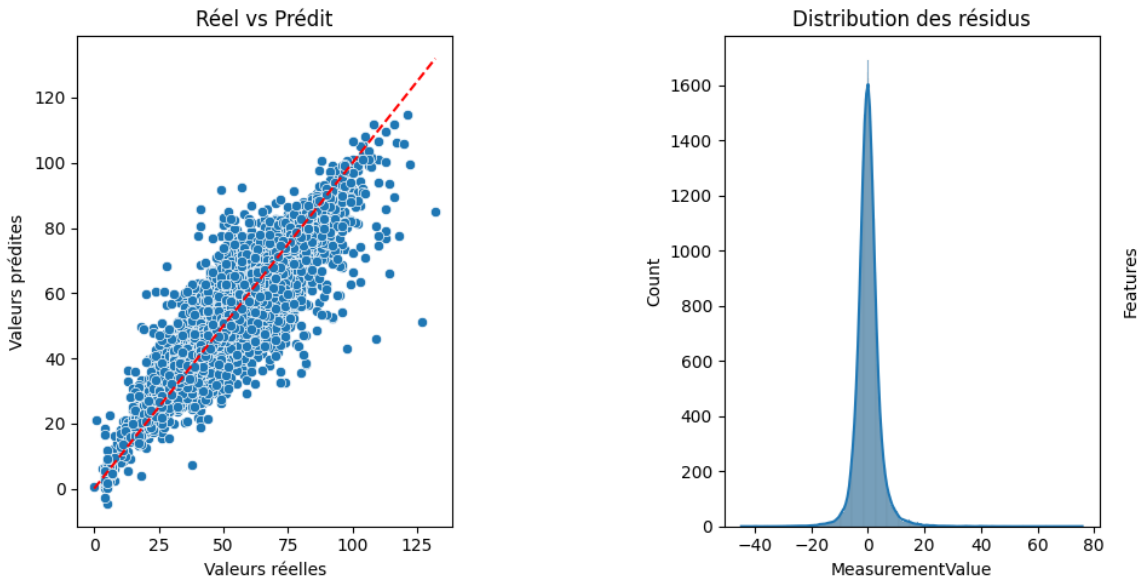

Figure 4.2 : Analyse des résidus, comparaison entre valeurs réelles et prédites, et distribution des erreurs absolues pour le modèle LSTM avec attention.

4.6 DISCUSSION COMPARATIVE

La comparaison des modèles met en évidence trois constats principaux. Premièrement, les modèles statistiques, bien qu'interprétables, ne suffisent pas à représenter toute la complexité des pertes techniques. Leur structure repose principalement sur les dépendances temporelles de la variable cible et demeure limitée lorsque plusieurs variables exogènes interagissent de manière non linéaire.

Deuxièmement, les modèles d'apprentissage automatique classiques, en particulier XGBoost, améliorent nettement les performances. Cette amélioration confirme l'importance des variables électriques, météorologiques et calendaires construites dans la méthodologie. XGBoost constitue ainsi une référence solide pour la prédiction tabulaire des pertes techniques.

Troisièmement, le modèle LSTM avec attention apporte une amélioration supplémentaire avec un RMSE de 0.064, un R^2 de 0.97 et une diminution du MAE de 0.03 par rapport au modèle XGBoost. Cette amélioration s'explique par sa capacité à exploiter directement les séquences temporelles, au lieu de dépendre uniquement de variables retardées construites manuellement. Le mécanisme d'attention renforce cette capacité en permettant au modèle de sélectionner les instants les plus informatifs dans la fenêtre temporelle considérée.

Ces résultats sont cohérents avec le positionnement du mémoire. La prédiction ligne par ligne permet de préserver les spécificités locales de chaque infrastructure, tandis que le modèle LSTM avec attention permet de mieux exploiter la dynamique temporelle propre à chaque ligne. L'association de ces deux éléments constitue donc une contribution importante du travail réalisé.

4.7 LIMITES OBSERVÉES

Malgré les performances obtenues, certaines limites doivent être soulignées. Premièrement, la qualité des prédictions dépend fortement de la qualité des données électriques et météorologiques disponibles. Les valeurs manquantes, les erreurs de mesure, les changements de configuration du réseau ou les périodes d'exploitation atypiques peuvent affecter la performance des modèles.

Deuxièmement, le modèle LSTM avec attention nécessite davantage de ressources de calcul que les modèles statistiques et les modèles d'apprentissage automatique classiques. Cette exigence est particulièrement importante dans une approche ligne par ligne, où un modèle distinct est entraîné pour chaque ligne de transport.

Troisièmement, même si le mécanisme d'attention améliore l'interprétabilité temporelle du modèle, l'architecture demeure moins directement interprétable qu'un modèle statistique ou qu'un modèle d'arbres accompagné d'analyses d'importance des variables. Une analyse plus approfondie des poids d'attention et des contributions des variables pourrait donc être envisagée afin de renforcer l'explicabilité du modèle.

4.8 SYNTHÈSE DES RÉSULTATS

Les résultats obtenus montrent que la prédiction des pertes techniques bénéficie fortement de l'utilisation de modèles capables de représenter les interactions entre variables électriques, météorologiques et temporelles. Les modèles statistiques offrent une base interprétable, mais limitée. Les modèles d'apprentissage automatique, notamment XGBoost, améliorent significativement les performances en capturant des relations non linéaires. Le modèle LSTM avec attention constitue l'approche la plus performante, car il combine la modélisation séquentielle des données horaires avec une attention portée aux instants les plus pertinents pour la prédiction.

Dans l'ensemble, ces résultats confirment la pertinence de l'approche proposée dans ce mémoire. La granularité ligne par ligne permet de mieux représenter les particularités locales du réseau de transport, tandis que le modèle LSTM avec attention permet d'exploiter la dynamique temporelle des pertes. Cette combinaison apparaît particulièrement adaptée à la prédiction des pertes techniques dans un réseau réel, étendu et soumis à des conditions d'exploitation variables.

CHAPITRE V

CONCLUSION ET PERSPECTIVES

Les pertes techniques sur les lignes de transport représentent un enjeu important pour les opérateurs de réseaux électriques, puisqu'elles influencent l'efficacité énergétique, la planification de la production et l'équilibre entre l'énergie injectée et l'énergie disponible après transport. Dans le contexte du réseau d'HQ, leur prédiction demeure complexe en raison de l'étendue du réseau, de la variabilité des charges, des conditions météorologiques et des caractéristiques propres à chaque ligne.

Ce mémoire a porté sur le développement de modèles prédictifs des pertes techniques à partir de données réelles du réseau de transport d'HQ. Contrairement aux approches globales souvent utilisées dans la littérature, l'approche proposée repose sur une prédiction ligne par ligne. Cette granularité permet de mieux tenir compte des spécificités locales des infrastructures, notamment les variables électriques, les caractéristiques des lignes, les postes connectés, les conditions météorologiques locales et les effets calendaires.

La méthodologie développée a permis de construire un jeu de données combinant des mesures électriques, des données météorologiques et des variables calendaires. Plusieurs familles de modèles ont ensuite été comparées, incluant des modèles statistiques, des modèles d'apprentissage automatique et un modèle LSTM avec un mécanisme d'attention.

Les résultats montrent que les modèles statistiques, tels que ARIMA et Prophet, constituent des références utiles, mais présentent des limites lorsque les pertes dépendent de relations non linéaires et multivariées. Parmi les approches classiques d'ML étudiées, XGBoost offre les meilleures performances. Toutefois, le modèle LSTM avec mécanisme d'attention se révèle le plus performant parmi l'ensemble des approches évaluées, grâce à sa capacité à exploiter la dynamique temporelle des données horaires et à mieux capturer les variations associées aux pertes techniques.

Ces résultats mettent en évidence l'intérêt de la prédiction ligne par ligne. En préservant les caractéristiques propres à chaque ligne de transport, cette approche limite l'effet de lissage associé aux modèles globaux fondés sur des données agrégées. Elle apparaît donc particulièrement pertinente pour les réseaux étendus, où les conditions climatiques, les profils de charge et les caractéristiques physiques peuvent varier fortement d'une infrastructure à l'autre.

Certaines limites demeurent toutefois présentes. La qualité des prédictions dépend de la qualité des données électriques et météorologiques utilisées, ainsi que de la cohérence temporelle entre les différentes sources. De plus, les modèles d'apprentissage profond, bien qu'efficaces, restent moins interprétables que les modèles statistiques ou les modèles fondés sur des arbres de décision, ce qui constitue un enjeu important pour une utilisation en contexte industriel.

Plusieurs perspectives peuvent être envisagées. Les travaux futurs pourraient intégrer davantage de connaissances physiques dans les modèles prédictifs, notamment par des approches hybrides combinant lois physiques et IA. L'amélioration de l'explicabilité des modèles, à l'aide de méthodes telles que SHapley Additive exPlanations (SHAP), Local Interpretable Model-agnostic Explanations (LIME) ou l'analyse des poids d'attention, constituerait également

une piste importante. Une autre perspective consisterait à explorer des architectures avancées dédiées aux séries temporelles longues, telles qu'Informer, Autoformer ou le TFT. Enfin, l'extension de l'approche à un plus grand nombre de lignes, à différents horizons de prédiction et à des données complémentaires, comme l'état de la topologie du réseau ou les informations de maintenance, permettrait de renforcer la robustesse et l'utilité opérationnelle des modèles.

En conclusion, ce mémoire montre que la prédiction des pertes techniques sur les lignes de transport peut être améliorée par l'utilisation de données réelles enrichies et par une approche ligne par ligne. Cette contribution ouvre la voie au développement d'outils intelligents d'aide à la décision pour l'exploitation et la planification des réseaux électriques de grande envergure.

BIBLIOGRAPHIE

[1] H. Abbasimehr et R. Paki, "Improving time series forecasting using lstm and attention models," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, n°. 1, p. 673–691, 2022.

[2] S. Ahmed, I. E. Nielsen, A. Tripathi, S. Siddiqui, R. P. Ramachandran, et G. Rasool, "Transformers in time-series analysis : A tutorial," *Circuits, systems, and signal processing*, vol. 42, n°. 12, p. 7433–7466, 2023.

[3] I. Aravena, C.-C. Sun, R. Shi, S. Majumder, W. Yan, J.-Y. Joo, L. Xie, et J. Wang, "Open power system datasets and open simulation engines : A survey toward machine learning applications," *IEEE Open Access Journal of Power and Energy*, vol. 12, p. 353–365, 2025.

[4] M. Aredes, R. F. S. Dias, A. F. C. Aquino, C. Portela, et E. H. Watanabe, "Power electronics applications in bulk power transmission over long distances," dans *2010 IEEE International Symposium on Industrial Electronics*, 2010, p. 3757–3765.

[5] J. Asha, S. Rishidas, S. SanthoshKumar, et P. Reena, "Analysis of temperature prediction using random forest and facebook prophet algorithms," *Lecture Notes on Data Engineering and Communications Technologies*, vol. 46, p. 432 – 439, 2020, cited by : 11. [En ligne]. Disponible : https://www.scopus.com/inward/record.uri?eid=2-s2.0-85083445122&doi=10.1007%2f978-3-030-38040-3_49&partnerID=40&md5=ddc2949519a2b84f91b872aead52b248

[6] M. E. Athanasopoulou, J. Deveikyte, A. Mosca, I. Peri, et A. Provetti, "A hybrid model for forecasting short-term electricity demand," dans *Proceedings of the Second ACM International Conference on AI in Finance*, ser. ICAIF '21. New York, NY, USA : Association for Computing Machinery, 2022. [En ligne]. Disponible : https://doi.org/10.1145/3490354.3494371

[7] I. Baran, M. Costea, et T. Leonida, "Power losses on overhead lines under various loading regimes and weather conditions," dans *2017 5th International Symposium on Electrical and Electronics Engineering (ISEEE)*, 2017, p. 1–6.

[8] T. Bashir, C. Haoyong, M. F. Tahir, et Z. Liqiang, “Short term electricity load forecasting using hybrid prophet-lstm model optimized by bpnn,” *Energy Reports*, vol. 8, p. 1678–1686, 2022. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S2352484721015067

[9] P. Boopathy, M. Liyanage, N. Deepa, M. Velavali, S. Reddy, P. K. R. Maddikunta, N. Khare, T. R. Gadekallu, W.-J. Hwang, et Q.-V. Pham, “Deep learning for intelligent demand response and smart grids : A comprehensive survey,” *Computer Science Review*, vol. 51, p. 100617, 2024. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1574013724000017

[10] J. Breido, V. Kaverin, et S. Voytkevich, “Distribution of power losses on high-voltage supports,” dans *Annals of DAAAM and Proceedings of the International DAAAM Symposium*, vol. 29, n°. 1. Danube Adria Association for Automation and Manufacturing, DAAAM, 2018, Conference paper, p. 0329 – 0337, cited by : 5 ; All Open Access, Gold Open Access. [En ligne]. Disponible : https://www.scopus.com/inward/record.uri?eid=2-s2.0-85060111018&doi=10.2507%2f29th.daaam.proceedings.047&partnerID=40&md5=a7161009263f43f39c244dfbf65e0089

[11] L. Breiman, “Bagging predictors,” *Machine Learning*, vol. 24, n°. 2, p. 123–140, Aug 1996. [En ligne]. Disponible : https://doi.org/10.1007/BF00058655

[12] ——, “Random forests,” *Machine learning*, vol. 45, n°. 1, p. 5–32, 2001.

[13] M. Brown Do Coutto Filho, *Introduction*. Cham : Springer Nature Switzerland, 2024, p. 1–16. [En ligne]. Disponible : https://doi.org/10.1007/978-3-031-63288-4_1

[14] V. Bugade, O. Shinde, P. Potdar, et A. Patil, “Iot based theft detection in three phase distribution line,” dans *2022 IEEE Pune Section International Conference (PuneCon)*, 2022, p. 1–4.

[15] E. Caro, *Application of monte carlo method to electric power systems : Computation of measurement correlations*. Nova Science Publishers, Inc., 2020. [En ligne]. Disponible : https://www.scopus.com/inward/record.uri?eid=2-s2.0-85145869847&partnerID=40&md5=a7483ca9e5e7429b3be3bcb6ceca0479

[16] S. Chaturvedi, E. Rajasekar, S. Natarajan, et N. McCullen, “A comparative assessment of sarima, lstm rnn and fb prophet models to forecast total and peak monthly energy demand for india,” *Energy Policy*, vol. 168, p. 113097, 2022. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0301421522003226

[17] G. Cheng, Y. Lin, A. Abur, A. Gómez-Expósito, et W. Wu, “A survey of power system state estimation using multiple data sources : Pmus, scada, ami, and beyond,” *IEEE Transactions on Smart Grid*, vol. 15, n°. 1, p. 1129–1151, 2024.

[18] S. Debnath, M. Uzzal Mia, M. Abubakkar, M. Robiul Islam, M. Shafiqul Islam Mridul, et A. Kishor Biswas, "Hybrid multi-scale deep learning enhanced electricity load forecasting using attention-based convolutional neural network and lstm model," *IEEE Access*, vol. 14, p. 13 423–13 444, 2026.

[19] S. Deng, X. Dong, L. Tao, J. Wang, Y. He, et D. Yue, "Multi-type load forecasting model based on random forest and density clustering with the influence of noise and load patterns," *Energy*, vol. 307, p. 132635, 2024. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0360544224024095

[20] S. Dhulipala, B. Maram, R. Rao, M. Naved, K. S. S. J, et M. Nidhya, "A comparative analysis of deep learning and ensemble methods for time series forecasting," dans *2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)*, 2025, p. 1713–1716.

[21] C. Ding, Y. Zhou, Q. Ding, et Z. Wang, "Loss prediction of ultrahigh voltage transmission lines based on eemd–lstm–svr algorithm," *Frontiers in Energy Research*, vol. Volume 10 - 2022, 2022. [En ligne]. Disponible : https://www.frontiersin.org/journals/energy-research/articles/10.3389/fenrg.2022.811745

[22] Q. Dong, R. Huang, C. Cui, D. Towey, L. Zhou, J. Tian, et J. Wang, "Short-term electricity-load forecasting by deep learning : A comprehensive survey," *Engineering Applications of Artificial Intelligence*, vol. 154, p. 110980, 2025. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0952197625009807

[23] M. Draz, F. Pagel, et S. Albayrak, "Probabilistic risk assessment in power systems with high wind energy penetration," *IEEE Access*, vol. 12, p. 140 097–140 111, 2024.

[24] V. Dvorkin et A. Botterud, "Differentially private algorithms for synthetic power system datasets," *IEEE Control Systems Letters*, vol. 7, p. 2053–2058, 2023.

[25] M. Elseidi, "A hybrid facebook prophet-arima framework for forecasting high-frequency temperature data," *Modeling Earth Systems and Environment*, vol. 10, n°. 2, p. 1855 – 1867, 2024, cited by : 15. [En ligne]. Disponible : https://www.scopus.com/inward/record.uri?eid=2-s2.0-85173691485&doi=10.1007%2fs40808-023-01874-4&partnerID=40&md5=1cf0b9e32387579ec35d6e0bbbf21138

[26] R. Fateh, A. Darif, M. Pouliquen, M. Frikel, et S. Safi, "Machine learning for nonlinear system identification : Kernels and ann," dans *2024 3rd International Conference on Embedded Systems and Artificial Intelligence (ESAI)*, 2024, p. 1–6.

[27] A. Ghasempour et M. Martínez-Ramón, "Electric load forecasting using multiple output gaussian processes and multiple kernel learning," dans *2023 IEEE Symposium on Industrial Electronics & Applications (ISIEA)*, 2023, p. 1–6.

[28] M. Gillioz, G. Dubuis, et P. Jacquod, “A large synthetic dataset for machine learning applications in power transmission grids,” *Scientific Data*, vol. 12, n°. 1, Jan. 2025. [En ligne]. Disponible : http://dx.doi.org/10.1038/s41597-025-04479-x

[29] R. Guo, Z. Zhao, T. Wang, G. Liu, J. Zhao, et D. Gao, “Degradation state recognition of piston pump based on iceemdan and xgboost,” *Applied Sciences*, vol. 10, n°. 18, p. 6593, 2020.

[30] S. Guo, J. Guo, X. Li, et N. Zhang, “Rainfall prediction in henan province based on the prophet-sarima model,” dans *Proceedings of the 2024 International Conference on Big Data Mining and Information Processing*, ser. BDMIP ’24. New York, NY, USA : Association for Computing Machinery, 2025, p. 286–290. [En ligne]. Disponible : https://doi.org/10.1145/3735014.3735914

[31] P. K. Gupta, K. Tuttelberg, et J. Kilter, “Weather dependency of corona losses on 330 kv overhead transmission lines,” *International Journal of Electrical Power & Energy Systems*, vol. 155, p. 109537, 2024. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S014206152300594X

[32] ——, “Forecasting corona losses on high voltage transmission lines using machine learning,” *IEEE Transactions on Power Delivery*, vol. 40, n°. 5, p. 2696–2705, 2025.

[33] M. Hasan, Z. Mifta, S. J. Papiya, P. Roy, P. Dey, N. A. Salsabil, N.-U.-R. Chowdhury, et O. Farrok, “A state-of-the-art comparative review of load forecasting methods : Characteristics, perspectives, and applications,” *Energy Conversion and Management : X*, vol. 26, p. 100922, 2025. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S2590174525000546

[34] T. K. Ho, “The random subspace method for constructing decision forests,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, n°. 8, p. 832–844, 1998.

[35] S. Hochreiter et J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, n°. 8, p. 1735–1780, 11 1997. [En ligne]. Disponible : https://doi.org/10.1162/neco.1997.9.8.1735

[36] H. Holttinen, A. Tuohy, M. Milligan, E. Lannoye, V. Silva, S. Müller, et L. Söder, “The flexibility workout : Managing variable resources and assessing the need for power system modification,” *IEEE Power and Energy Magazine*, vol. 11, n°. 6, p. 53–62, 2013.

[37] D. Hulak et G. Taylor, “Investigating an ensemble of arima models for accurate short-term electricity demand forecasting,” dans *2023 58th International Universities Power Engineering Conference (UPEC)*, 2023, p. 1–6.

[38] R. Islam, S. Kabir, A. Shufian, M. S. Rabbi, et M. Akteruzzaman, “Optimizing renewable energy management and demand response with ant colony optimization : A pathway to enhanced grid stability and efficiency,” dans *2025 IEEE Texas Power and Energy Conference (TPEC)*, 2025, p. 1–6.

[39] J. Janjanin, A. Kurćubić, et Ž. Jovanović, “Forecast of transmission system losses by ems,” dans *2019 2nd International Colloquium on Smart Grid Metrology (SMAGRIMET)*, 2019, p. 1–5.

[40] X. Ji, H. Huang, D. Chen, K. Yin, Y. Zuo, Z. Chen, et R. Bai, “A hybrid residential short-term load forecasting method using attention mechanism and deep learning,” *Buildings*, vol. 13, n°. 1, p. 72, 2023.

[41] K. P. Juhász, I. Vokony, B. Hartmann, F. Béres, et V. Könyves, “Machine learning-based transmission loss prediction and mitigation strategies for the hungarian network,” dans *2025 IEEE Kiel PowerTech*, 2025, p. 1–6.

[42] H. T. Lestari, C. O. Sereati, M. Siregar, et K. O. Bachri, “Predicting transmission losses using eemd – svr algorithm,” *International Journal of Power Electronics and Drive Systems*, vol. 16, n°. 3, p. 2122 – 2129, 2025, cited by : 0 ; All Open Access, Gold Open Access. [En ligne]. Disponible : https://www.scopus.com/inward/record.uri?eid=2-s2.0-105015379355&doi=10.11591%2fijpeds.v16.i3.pp2122-2129&partnerID=40&md5=b4356ef1635a7c6ffb247eb25abf4b26

[43] B. Lim, S. Ö. Arık, N. Loeff, et T. Pfister, “Temporal fusion transformers for interpretable multi-horizon time series forecasting,” *International Journal of Forecasting*, vol. 37, n°. 4, p. 1748–1764, 2021.

[44] N. Lin, P. Palensky, et P. P. Vergara, “Energydiff : Universal time-series energy data generation using diffusion models,” *IEEE Transactions on Smart Grid*, vol. 16, n°. 5, p. 4252–4265, 2025.

[45] A. L’Heureux, K. Grolinger, et M. A. Capretz, “Transformer-based model for electrical load forecasting,” *Energies*, vol. 15, n°. 14, p. 4993, 2022.

[46] v. Manimegalai, V. Mohanapriya, P. R. Raaghav, D. Ranjan, S. Ranjan, et S. Selvamurugan, “Leveraging smart meter technology to combat energy theft and empower consumption monitoring,” dans *2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, 2024, p. 264–269.

[47] T. Matijašević, T. Antić, et T. Capuder, “A systematic review of machine learning applications in the operation of smart distribution systems,” *Energy Reports*, vol. 8, p. 12 379–12 407, 2022. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S2352484722017929

[48] C. Moldoveanu, I. Ionita, V. Brezoianu, S. Zaharescu, S. Grigorescu, I. Hategan, et M. Marcolt, “Romanian innovative system for power losses real time monitoring and analysis in electrical transmission and distribution systems,” dans *2021 9th International Conference on Modern Power Systems (MPS)*, 2021, p. 1–6.

[49] A. T. Mustafa et O. S. A.-D. Al-Yozbaky, “Forecasting energy demand and generation using time series models : A comparative analysis of classical, grey, fuzzy, and intelligent approaches,” *Franklin Open*, vol. 12, p. 100350, 2025. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S2773186325001380

[50] S. A. Nabavi, S. Mohammadi, N. H. Motlagh, S. Tarkoma, et P. Geyer, “Deep learning modeling in electricity load forecasting : Improved accuracy by combining dwt and lstm,” *Energy Reports*, vol. 12, p. 2873–2900, 2024. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S2352484724005614

[51] S. Nageshwaran et V. Nageshwaran, “Hybrid deep learning and xgboost models for enhanced energy forecasting : A comparative analysis,” dans *2025 IEEE Technology and Engineering Management Society Conference - Global (TEMSCON Global)*, 2025, p. 1–5.

[52] B.-C. Neagu, G. Grigoras, M. Gavrilas, R. Porumb, et I. Tristiu, “Technical losses estimation in low voltage distribution network using deterministic methods,” dans *The 17th International Conference Interdisciplinarity in Engineering*, L. Moldovan et A. Gligor, édit. Cham : Springer Nature Switzerland, 2024, p. 546–560.

[53] Q.-H. Ngo, B. L. Nguyen, T. V. Vu, J. Zhang, et T. Ngo, “Physics-informed graphical neural network for power system state estimation,” *Applied Energy*, vol. 358, p. 122602, 2024. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0306261923019669

[54] D. F. Niste, R. Tîrnovan, S. Pavel, H. Beleiu, C. Andrei, et M. Misaroș, “Electricity losses in focus : Detection and reduction strategies—state of the art,” *Applied Sciences (Switzerland)*, vol. 15, n°. 7, 2025, cited by : 2; All Open Access, Gold Open Access. [En ligne]. Disponible : https://www.scopus.com/inward/record.uri?eid=2-s2.0-105002275328&doi=10.3390%2fapp15073517&partnerID=40&md5=1ad14ccc9ff8a66c96ab27100d44a505

[55] Z. Niu, G. Zhong, et H. Yu, “A review on the attention mechanism of deep learning,” *Neurocomputing*, vol. 452, p. 48–62, 2021.

[56] F. Pandžić, I. Sudić, T. Capuder, et I. Pavičić, “On the practical aspects of machine learning based active power loss forecasting in transmission networks,” *IET Generation, Transmission & Distribution*, vol. 18, n°. 14, p. 2452–2463, 2024. [En ligne]. Disponible : https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/gtd2.13205

[57] V. Pentsos, S. Tragoudas, J. Wibbenmeyer, et N. Khdeer, “A hybrid lstm-transformer model for power load forecasting,” *IEEE Transactions on Smart Grid*, vol. 16, n°. 3, p. 2624–2634, 2025.

[58] V. Perifanis, C. L. Athanasiadis, N. Pavlidis, P. S. Efraimidis, et T. A. Papadopoulos, “Towards detecting non-technical losses in electrical power systems with machine learning,” dans *2024 3rd International Conference on Energy Transition in the Mediterranean Area, SyNERGY MED 2024*, 2024, Conference paper, cited by : 1. [En ligne]. Disponible : https://www.scopus.com/inward/record.uri?eid=2-s2.0-85215580510&doi=10.1109%2fSyNERGYMED62435.2024.10799450&partnerID=40&md5=5be340b3485af4f2e0271863ebf47436

[59] M. Pečjak, M. Symponi, et D. Stratogiannis, “Detecting suspicious behaviour in power consumption with fuzzy logic,” dans *2023 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*, 2023, p. 1–6.

[60] A. Pranolo, Y. Mao, A. P. Wibawa, A. B. P. Utama, et F. A. Dwiyanto, “Robust lstm with tuned-pso and bifold-attention mechanism for analyzing multivariate time-series,” *IEEE Access*, vol. 10, p. 78 423–78 434, 2022.

[61] D. Pál, Ľ. Beňa, J. Urbanský, et M. Oliinyk, “The impact of reconfiguration on power losses in smart networks,” dans *2020 21st International Scientific Conference on Electric Power Engineering (EPE)*, 2020, p. 1–5.

[62] L. M. R. Raggi, F. C. L. Trindade, V. C. Cunha, et W. Freitas, “Non-technical loss identification by using data analytics and customer smart meters,” *IEEE Transactions on Power Delivery*, vol. 35, n°. 6, p. 2700–2710, 2020.

[63] N. S. M. Salleh, A. Suliman, et B. N. Jørgensen, “Comparison of electricity load prediction errors between long short-term memory architecture and artificial neural network on smart meter consumer,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13051 LNCS, p. 600 – 609, 2021, cited by : 0. [En ligne]. Disponible : https://www.scopus.com/inward/record.uri?eid=2-s2.0-85120530402&doi=10.1007%2f978-3-030-90235-3_52&partnerID=40&md5=2a74707944e6202837cf4a16a2d5dc60

[64] A. Shah, W. Mesbah, et A. T. Al-Awami, “An algorithm for detaching technical losses from non-technical losses in distribution systems,” dans *2021 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2021, p. 1–5.

[65] S. A. H. Shah, U. Ahmed, M. Bilal, A. R. Khan, S. Razzaq, I. Aziz, et A. Mahmood, “Improved electric load forecasting using quantile long short-term memory network with dual attention mechanism,” *Energy Reports*, vol. 13, p. 2343–2353, 2025. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S2352484725000617

[66] J. V. Shitikova et S. N. Shelyug, “Calculation of power losses in feeders 0,4 kv with incorrect information about the load current dynamic management model of innovations generations,” dans *2017 XX IEEE International Conference on Soft Computing and Measurements (SCM)*, 2017, p. 199–202.

[67] G. Shvedov et A. Shchepotin, “Analysis of error range in calculation of load electric power losses in wires of overhead transmission lines when neglecting influence of meteorological factors,” dans *2019 International Ural Conference on Electrical Power Engineering (UralCon)*, 2019, p. 368–372.

[68] J. Smith et L. Wang, “Ai-based estimation of energy losses in power grids,” dans *Procedia Energy Conference*. Elsevier, 2023, p. 89–96.

[69] G. Stana et V. Brazis, “Theoretical and practical study of the effect of dc grid wire temperature variation on transmission losses,” dans *2023 International Symposium on Fundamentals of Electrical Engineering (ISFEE)*, 2023, p. 1–6.

[70] F. Succetti, A. Rosato, R. Araneo, et M. Panella, “Multidimensional feeding of lstm networks for multivariate prediction of energy time series,” dans *2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, 2020, p. 1–5.

[71] I. Sudić, M. Mesar, B. Franc, T. Capuder, T. Ivanković, K. Pavić, et I. Pavić, “Short-term transmission system losses forecast based on supervised machine learning,” dans *2020 International Conference on Smart Systems and Technologies (SST)*, 2020, p. 199–204.

[72] I. Sudić, T. Capuder, B. Franc, et M. Mesar, “Estimating the value of power losses forecast in the transmission system by comparing different electricity procurement strategies,” dans *2022 18th International Conference on the European Energy Market (EEM)*, 2022, p. 1–5.

[73] S. Thundiyil, S. Shalamzari, J. Picone, et S. McKenzie, “Transformers for modeling long-term dependencies in time series data : A review,” dans *2023 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 2023, p. 1–5.

[74] W. F. Tinney et C. E. Hart, “Power flow solution by newton’s method,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-86, n°. 11, p. 1449–1460, 1967.

[75] J. Tulensalo, J. Seppänen, et A. Ilin, “An lstm model for power grid loss prediction,” *Electric Power Systems Research*, vol. 189, p. 106823, 2020. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0378779620306246

[76] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, et I. Polosukhin, “Attention is all you need,” dans *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus,

S. Vishwanathan, et R. Garnett, édit., vol. 30. Curran Associates, Inc., 2017. [En ligne]. Disponible : https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[77] T. Vu, G. S. Thirunavukkarasu, M. Seyedmahmoudian, S. Mekhilef, et A. Stojcevski, "Comparative analysis of regression models for household appliance energy consumption prediction using extreme gradient boosting," dans *2023 33rd Australasian Universities Power Engineering Conference (AUPEC)*, 2023, p. 1–6.

[78] J. Zhao, F. Chu, L. Xie, Y. Che, Y. Wu, et A. F. Burke, "A survey of transformer networks for time series forecasting," *Computer Science Review*, vol. 60, p. 100883, 2026.

[79] H. Zhou, S. Zhang, J. Peng, et et al., "Informer : Beyond efficient transformer for long sequence time-series forecasting," dans *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, n°. 12, 2021, p. 11 106–11 115.